DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES
DER FAKULTÄT FÜR CHEMIE UND PHARMAZIE
DER LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Proteomics investigations of immune activation

ANNIKA FRAUENSTEIN

aus

REGENSBURG

2022

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Matthias Mann betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe von mir erbracht.

München, den 06.04.2022

Annika Frauenstein

Dissertation eingereicht am 12.04.2022

1. Gutachter: Prof. Dr. Matthias Mann

2. Gutachter: Prof. Dr. Felix Meissner

Mündliche Prüfung am 04.07.2022

1 Abstract

Cellular information flow is facilitated by sophisticated molecular networks comprised primarily of proteins. Biological signals are communicated by dynamic network rearrangements induced by covalent and frequently reversible protein modifications, as well as by interactions between various protein sets. While precise orchestration of these signaling networks is required for cell-type and context-specific tasks, their dysregulation is frequently associated with disease, as signal transmission can become inefficient, overshooting, or misdirected. However, the interplay of protein modification and interaction in immune signaling remains poorly understood. Experimental methodologies that reflect the fundamental principles of dynamic signaling network assembly are thus critical for establishing causal relationships between cellular phenotypes and developing tactics for targeted interference. Proteomics based on mass spectrometry has evolved into a versatile method for addressing a wide variety of problems. Apart from investigating the expression of proteins, mass-spectrometry can decode protein-protein interaction networks and detect chemical modifications of proteins.

Throughout my Ph.D. I focused on dynamic protein-protein interactions and post-translational changes in immune cells. I developed and employed affinity enrichment followed by mass spectrometry (AE-LC-MS/MS) to investigate the dynamic interaction between PPIs and PTMs of 19 bait proteins in response to TLR activation or drug perturbation. I was able to functionally evaluate novel PPIs and PTMs regulating NFkB activation. I also studied proteolysis of signal peptides in human and murine tissues by establishing a proteomic meta-analysis workflow. Thereby, I provide mass spectral evidences for signal-peptide cleavages and was able to double the currently experimentally confirmed signal peptide cleavage sites. Additionally, I devised a detailed step-by-step protocol for analyzing secretomes by proteomics and collaborated in two projects examining the secretomes of pyroptosis and TNF-induced necroptosis, respectively. In an additional collaboration, I investigated the influence of arginine on the metabolism and development of multinuclear giant cells with a multi-omics approach involving transcriptomics, proteomics, and metabolomics.

Contents

1 Abstract 3

2 Introduction 6

2.1 Mass Spectrometry (MS)- based proteomics 6

2.1.1 Bottom-up and top-down proteomics 6

2.1.2 Sample preparation for shotgun proteomics 8

2.1.3 Liquid chromatography 9

2.1.4 MS Analysis: Instrumentation 10

2.1.5 Fragmentation techniques 11

2.1.6 Acquisition methods 12

2.1.7 Protein identification via database searches and de novo sequencing 13

2.1.8 Protein quantification 14

2.2 Interaction proteomics 16

2.2.1 AP-MS with antibodies raised against endogenous proteins 17

2.2.2 AP-MS with epitope tags 17

2.2.3 Interactor calling in AP-MS 19

2.2.4 New approaches for identification of protein-protein interactions 20

2.3 Post-translational protein modifications (PTMs) 21

2.3.1 Phosphorylation 21

2.3.2 Isgylation 23

2.3.3 Proteolysis 25

2.4 Secretomics 27

2.5 Biological rational: TLR signaling 28

3 Aims of the thesis 31

4 Publications 32

4.1 Project 1: Identification of Covalent Modifications Regulating Immune Signaling Complex Composition and Phenotype 32

4.2 Project 2: Revisiting signal peptide cleavage by neo N-terminal peptide meta-analysis 55

4.2.1 Abstract 57

4.2.2 Introduction 57

4.2.3 Results 59

4.2.4 Discussion 64

4.2.5 Figures and figure legends 66

4.2.6 Supplementary figures and figure legends 71
4.2.7 Methods 76
4.3 Project 3: Quantitative Proteomics of Secreted Proteins 79
4.4 Project 4: Quantitative and Dynamic Catalogs of Proteins Released during Apoptotic and Necroptotic Cell Death 93
4.5 Project 5: Proteomics reveals distinct mechanisms regulating the release of cytokines and alarmins during pyroptosis 112
4.6 Project 6: Environmental arginine controls multinuclear giant cell metabolism and formation 134
5 Discussion 150
5.1 Interaction proteomics 150
5.2 Post-translational modifications 152
6 References 154
7 Addendum 170

2 Introduction

2.1 Mass Spectrometry (MS)- based proteomics

Proteins are the ultimate executors of genetic information, orchestrating cellular processes. They are translated from mRNA, which is transcribed from DNA. Whole-genome sequencing efforts of the human genome have uncovered more than 20 000 protein-coding genes [1]. Co-transcriptional and co-translational processes lead to the emergence of more than 6 million estimated proteoforms [2-5],

Mass spectrometry-based proteomics investigates the proteome – the entirety of expressed proteins at a specific time and location. Recent breakthroughs in MS-based proteomics workflows, instrumentation and analyses enable the detection, identification, and relative or absolute quantification of thousands of proteins in a single experiment. Detailed proteomic abundance atlases of human tissues have delivered spectral evidence for over 13 000 proteins [6]. Approximately 90% of the human proteome has been discovered to date, leaving 10% of proteins unidentified [7].

2.1.1 Bottom-up and top-down proteomics

Two major MS-based proteomic strategies are utilized to characterize proteins comprehensively. These approaches are referred to as top-down or bottom-up proteomics. Bottom-up protein analysis examines peptides derived from proteins via experimental proteolysis - also termed shotgun proteomics [8-10]. Bottom-up techniques begin with the extraction of proteins from biological material (e.g., tissues or cells) and their proteolytic digestion into small pieces for effective fragmentation by proteases (e.g., Trypsin, Lys-C) [11]. The resulting complex peptide mixtures are separated by reversed phased high-pressure liquid chromatography (HPLC) based on hydrophobicity and coupled online to a mass spectrometer. The technique is widely used in many proteomic studies, and there are several bioinformatic tools for data interpretation, including identification and quantification [10]. Sample preparation, liquid chromatography, and mass spectrometry technology have evolved dramatically in recent years, enabling proteomic analysis of minute sample quantities and paving the way for single-cell proteomics [12].

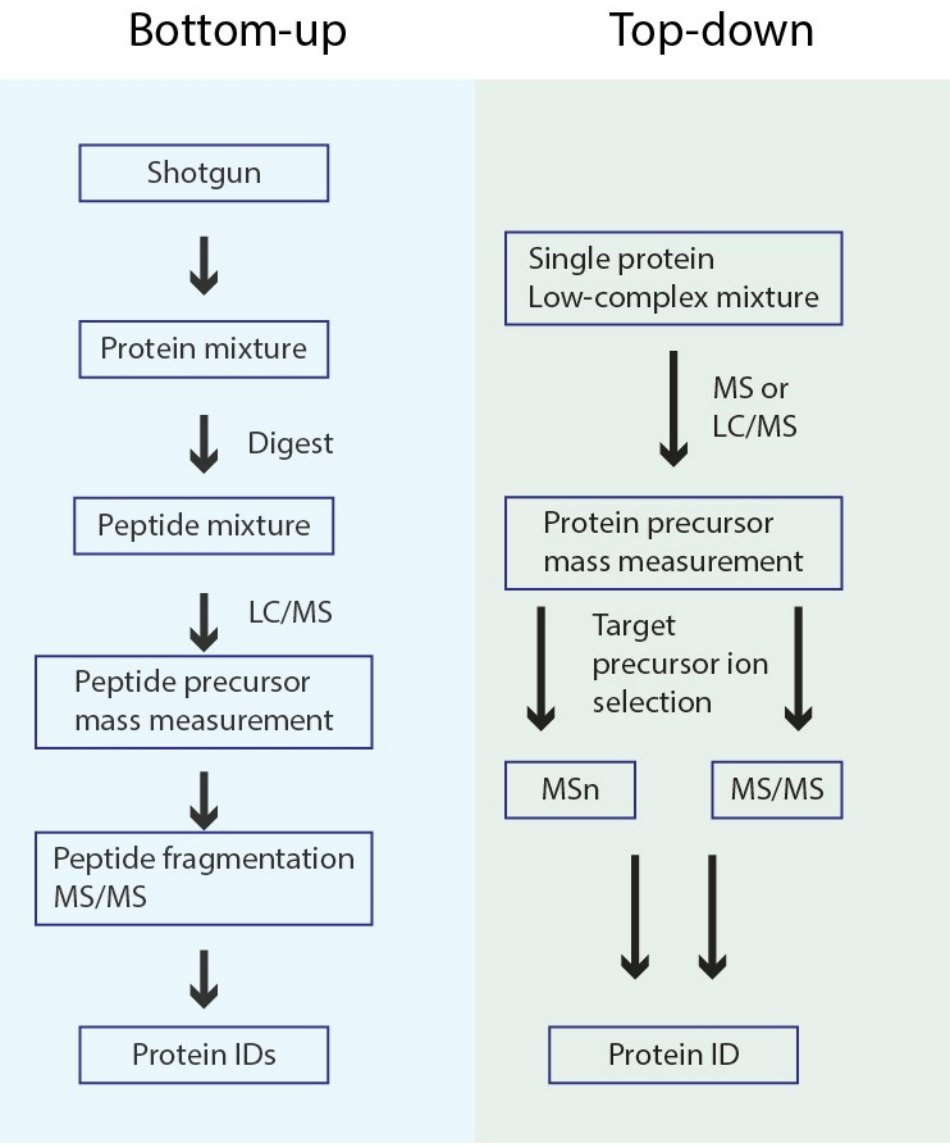


Figure 1: A schematic approach for the bottom-up and top-down characterization and identification of proteins using mass spectrometry. Adapted from [13].

Top-down proteomics studies proteins in their native state and is the ideal technique for delineating unambiguous, and proteoform resolved molecular information. Analyses can be carried out under both denaturing [14] and native conditions [15]. Often, pre-analysis enrichment of proteins of interest is conducted [16-18], however global proteoform characterizations in conjunction with high-resolution front-end protein separation are also being developed [19]. Top-down proteomics is currently applied for large biomolecule therapeutics and diagnostics like antibodies, but due to advancements in sample prep, protein separation and instrumentation may be a future avenue for the discovery of novel biomarkers and disease mechanisms [20].

2.1.2 Sample preparation for shotgun proteomics

Methods for sample preparation that are robust and reproducible are critical for any shotgun proteomics experiment. Before LC-MS, sample preparation for label-free proteomics workflows includes cell/tissue lysis, protein denaturation, cysteine reduction and alkylation, proteolytic digestion and desalting, and removal of interferences (e.g., detergents, salts, chemicals, lipids, metabolites). Many protocols exist and frequently vary according to the research question.

Fully denaturing techniques with ionic or chaotropic detergents (SDC, SDS, Urea, or Guanidinium Chloride) are frequently used to homogenize and extract proteins from biological samples efficiently. Native cell lysis (freeze-thawing, bead-milling, cell-cracking, low concentrations of non-ionic, weak detergents like NP-40) is used for interaction proteomics, co-fractionation studies, or spatial proteomics experiments when intact protein complexes are of interest. In addition, native lysis is used in conjunction with pan protease inhibitors to avoid proteolysis during the lysis and subsequent affinity purification processes. While cytosolic proteins are easily examined using native interaction proteomics processes, membrane proteins require specific detergents such as the non-ionic glycoside detergents DDM or DDM/CHS [21]. Notably, detergent-assisted lysis is frequently combined with sonication to aid in the mechanical solubilization of proteins, cell lysis, and shearing of genomic DNA.

Cysteine reduction and alkylation are used in practically every proteomic experiment to disrupt disulfide links, facilitate protease access, and prevent disulfide bond reassembly – excluded are, e.g., studies on cysteine modifications [22]. Dithiothreitol (DTT) [23] and tris(2-carboxyethyl)phosphine (TCEP) [24] are two commonly used chemical agents for cystin reduction. Alkylating agents are iodoacetamide (IAA) or chloroacetamide (CAA), leading to the covalent modification of cysteine thiol groups by carbamidomethylation.

Most frequently proteins are digested using endoproteinase Lys-C and trypsin resulting in basic amino acids (Lysine, Arginine) at the C-terminus of each peptide. Thereby, double-charged peptide species at both the N- and C-termini under acidic HPLC conditions are generated. On average, Trypsin-generated peptides are ten amino acids in length [11]. Because some of the peptides formed by Trypsin/Lys-C are either too short or too long for efficient MS analysis and data-analysis, additional proteases such as AspN, LysN, ArgC, GluC, or chymotrypsin can be used to increase sequence coverage [25].

Interfering chemicals such as detergents, salts, lipids, metabolites, DNA, and polysaccharides are removed from samples before proteomics analysis. A milestone in the field was the development of a microcolumn tip-based micro purification system termed Stop and Go Extraction (StageTips), which is based on solid-

phase extraction with reversed-phase material (octadecyl carbon chain - C18) [26]. Further developments and optimization of reproducibility and adaptability in sample preparation resulted in the invention of the 'in-Stage-Tip' (iST) workflow [27], in which the entire sample preparation can be completed in a single vessel using the mild detergent SDC in combination with the solid-phase extraction material Styroldivinylbenzol – Reversed-Phase Sulfonate (SDB-RPS). Another single-vessel sample preparation method (SP3) that is suitable with liquid handling robots is based on carboxyl-coated paramagnetic nanoparticles that enable the use of powerful detergents (e.g., SDS) by allowing for washing steps with organic solvents[28]. MStern is another high-throughput sample preparation approach that uses polyvinylidene fluoride (PVDF) membranes and is suited for processing extremely dilute samples [29].

Materials can be pre-fractionated online or offline before MS analysis, to obtain a more comprehensive proteome coverage. Pre-fractionation techniques include strong anion exchange (SAX) and strong cation exchange (SCX), both of which have a high degree of orthogonality with reversed-phase chromatography [9, 30, 31]. In addition, separation on polarity (HILIC) and hydrophobicity (ERLIC) are often employed for pre-fractionating phosphopeptides [32]. Offline high pH reversed-phase separation combined with fraction concatenation leads to high orthogonality of online acidic reversed-phase chromatography. While fractionation is time-consuming and requires a high protein starting material, a fully automated fractionator dubbed the "loss-less nano spider" fractionator based on high pH fractionation has been invented [33]. This fractionator automatically concatenates dissimilar fractions to obtain complete chromatograms [34].

2.1.3 Liquid chromatography

Chromatographic separation of complete proteome samples, containing thousands of unique peptides, is required for maximum peptide identification by proteomics. The fundamental premise of the LC separation of peptides is their affinity for a stationary material (column) and a mobile phase (solvent gradient elution). Due to its high resolving power (peak capacity), reproducibility, and resilience, reversed-phase HPLC is utilized to separate peptide mixtures in the majority of bottom-up proteomics research [35]. The mass spectrometer is often operated in positive mode, necessitating positively charged peptide species as analytes. Thus, in reverse phase chromatography, peptides are initially loaded onto the stationary phase (C-18 silica) using their affinity for the stationary material due to coulomb or hydrophobic contact at acidic pH. The solubility of peptides is gradually enhanced by increasing the amount of organic solvent (acetonitrile), and they are eluted from the column into the mass spectrometer. The number of detected peptides increases linearly with the liquid chromatography’s peak capacity and resolving power

[36]. Column length, column diameter, column filling, and gradient length affect the resolution of chromatographic separation. Significant improvements in resolution, speed and sensitivity were achieved by using long columns filled with small silica particles (1.7 um in diameter) and operated at extremely high pressures (up to 1000 bar) [37, 38]. By lowering the backpressure created by long columns and small particle sizes, column heating devices contribute to future advancements in chromatographic performance [39].

2.1.4 MS Analysis: Instrumentation

New high-performance instrumentation that provides both high resolution and high mass measurement accuracies for MS1 and MS2 levels has resulted from major developments over the last decade. A mass spectrometer is typically composed of an ion source, a mass analyzer, and a detector.

Electrospray ionization (ESI) is by far the most frequently used technology for ionizing analytes (peptides or proteins) [40, 41]. First, the dissolved analytes eluting from the capillary column are subjected to a high voltage. Then, a strong electrical field is used to transfer the analytes to gas phase in the electrospray ion source. Another ionization technique is matrix aided laser desorption ionization (MALDI), which uses a laser beam to vaporize dry analytes embedded in a matrix [42]. All measurements in this thesis were carried out on Thermo Fisher Scientific Orbitrap mass spectrometers, which include Q Exactive HF, Q Exactive HF-X and Exploris.

2.1.4.1 Mass analyzers

There are different kinds of mass analyzers frequently used in proteomics investigations: Quadrupole mass analyzers, Ion trap analyzers, time-of-flight, FT-ICR and orbitrap mass analyzers. They ensure sensitivity, mass resolution, mass accuracy, and high quality MS/MS spectra [43]. In my projects, orbitrap mass analyzers were employed.

Quadrupole mass analyzers can differentiate between and filter ions of a certain m/z. They contain four cylindrical or hyperbolic rods (quadrupole) inside a vacuum chamber. A radio frequency (RF) and directed current (DC) allow only ions of a certain m/z to pass through the quadrupole in a stable trajectory [44].

Ion trap analyzers accumulate or “trap” ions of a selected mass range for some time before MS or MS/MS analysis commences. 2D linear ion traps, and 3D ion traps are also termed Paul Trap [45]. In 3D traps, ions are trapped between hyperbolic ring electrodes and hyperbolic electrode plates by an oscillating RF field and a superimposed DC electric field. Selective ejection of specific ions is achieved by varying the RF potential. Linear traps are similar to quadrupoles, but a potential field is applied to the end of the rods to

trap the ions. This ion trapping can lead to increased sensitivity; however, ion trap mass analyzers suffer from low resolving power (single unit mass resolution).

In **TOF instruments**, ions are accelerated and separated based on the time to travel through a field-free region, correlating to their m/z [46]. Mass is strictly a function of the time between initial acceleration and detection (time-of-flight), as the kinetic energy and the length of the flight tube remain constant. TOF analyzers can analyze small and big ions (ranging from few Da- 100 kDa) due to high ion transmission efficiencies.

The **FT-ICR** measures the cyclotron frequency of ions in a fixed magnetic field to reveal m/z [47]. Ions are caught in a Penning trap, similar to the 3D trap but with a magnetic instead of an electrical field. Thereby, the ions circle in a plane perpendicular to the magnetic field. Upon coherent excitation with a broadband RF field ions are raised to a higher cyclotron orbit [48]. Receiver plates detect the ions, and their intensity and time signal is converted into a frequency spectrum via Fourier transformation. Thus, the ion m/z is directly proportional to the cyclotron frequency.

In the **orbitrap mass analyzers** – a type of an ion-trap mass analyzer - ions are induced by an electrical field to oscillate in a trap. Fourier transformation converts the signal of the oscillating ions from time to frequency [49, 50]. The orbitrap mass analyzer is built from an inner spindle electrode covered by two hollow outer concave electrodes. Between the inner and outer electrodes, a linear electric field is generated by a voltage potential. Ions enter the orbitrap through a hole in one of the outer electrodes and start axial harmonical oscillation, influenced by the conical shape of the electrodes, the electrical field, and the tangential velocity of the ions, which creates opposing centrifugal forces. The outer electrodes detect the oscillating ions, and the signal is transformed to the frequency domain (axial harmonic frequency) by Fourier transformation, which is proportional to m/z. Due to their high mass resolution and versatility, orbitrap mass analyzers are broadly employed in proteomics.

2.1.5 Fragmentation techniques

The mass of intact peptides as determined by the MS1 scan is insufficient to unambiguously identify peptides, much less to detect and locate PTMs. Therefore, the complete peptide (precursor ion) is fragmented into product ions (MS/MS or MS2) to determine the amino acid sequence and position. The target precursor ions are selected within a certain m/z range, fragmented, and the mass of the fragment ions is determined [51]. The following MS/MS fragmentation modes are frequently employed in proteomics experiments: Collision-induced dissociation (CID) and higher-energy collisional dissociation (HCD) fragment peptides at the peptide bonds, generating b, a, and c-ions from the peptide's N-terminus,

and y, x, and z-ions from the peptide's C-terminus [52]. Fragmentation occurs when precursor ions collide with an inert gas (He, N_2 or Ar).

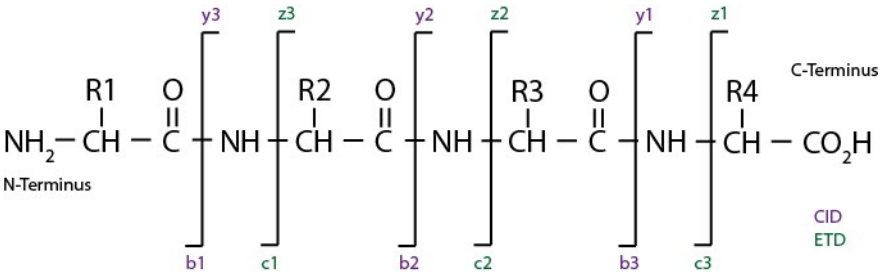


Figure 2: Fragmentation of peptides forms different ion species. b and y type ions are generated by CID/ HCD, while c- and z-type ions are generated by ETD.

2.1.6 Acquisition methods

Shotgun proteomics is often performed using discovery-driven data-dependent acquisition (DDA) techniques, in which peptides are fragmented based on their signal intensity. In my thesis, I used DDA as an acquisition method, predominantly. In DDA, the mass (m/z) of eluting peptides is determined using a survey scan (MS1, full scan), and the N most abundant precursor ions are selected for subsequent fragmentation (MS2), yielding fragment-ion spectra. Chromatographic elution of peptides in long gradients takes longer than a survey scan plus MS/MS scan, and accordingly numerous MS2 scans can follow a single MS survey scan to achieve sequencing of a maximum number of precursors. A dynamic exclusion window is set to the median elution time for the scanned peptides to prevent resequencing of the same peptide. Due to the semi-stochastic nature of intensity-driven precursor selection in DDA, identifications can be difficult to reproduce across large datasets containing highly complex mixtures, such as full-proteomes – typically for low abundance peptides. Match-between-runs — a MaxQuant technique for enhancing peptide identification and quantification – permits the transfer of MS/MS data between samples by aligning retention times and MS1 features. Due to increase in scan speed and resolution of modern mass spectrometers, data-independent acquisition (DIA) has gained appeal in recent years for discovery-driven proteomics [53, 54]. While DDA techniques pick a predetermined number of precursors, DIA simultaneously fragments all precursor ions within a given time window (broadband DIA) or mass window [55]. Co-eluting peptides within a particular mass window are co-fragmented, resulting in highly convoluted MS2 spectra [56]. Compared to DDA, higher levels of data completeness (fewer missing values), accuracy, and identification rates can be achieved [57, 58]. Spectral interpretations of DIA relies on sophisticated software solutions and pre-recorded or *in silico* high-quality spectrum libraries [57].

2.1.7 Protein identification via database searches and de novo sequencing

Protein identification relies on the interpretation of high resolution MS1 and MS2 spectra of the corresponding peptides. Peptide database search engines such as the commercial Mascot [59] and open-source Andromeda [60] are used for automated spectra identification. An *in silico* digest of all proteins of the studied organism is used as the reference to match observed and theoretical MS2 spectra with the MS1 mass as a constraint. The matches are evaluated with a probability-based scoring model, assessing the chance of theoretical and measured matches occurring randomly [61]. In MaxQuant [62], the quality of peptide spectrum (PSMs) matches is further scored by the posterior error probability (PEP), which is a measure of peptide length, charge, number of PTMs, and missed cleavages, and the Andromeda score. The estimation and control of false-positive identifications are conducted on the level of peptide-spectrum matches and protein group level in a target-decoy search strategy [63], which often employs a reverse database and enables calculating the false-discovery rate (FDR), which is usually set to 1%. It has been discussed that increasing the search space -e.g., by searching multiple genomes or PTMs - counter-intuitively leads to lower identification rates because of the higher probability of retrieving false -positives [64]. As MS/MS spectra can be convoluted and contain co-eluting peptides, MaxQuant enables the search of signals resulting from co-fragmenting additional precursors – termed “second peptide” search [60]. Finally, peptides are assembled to proteins with an FDR control based on protein-level PEPs – a product of the individual peptide PEP [60]. Peptides can be unique for proteins or shared between different proteins – for example, if several protein isoforms are present in the sample. The shared non-unique peptides – termed “razor peptides” - are attributed to the protein with the highest number of shared peptides following the principle of Occam's razor.

High-resolution spectral data from state-of-the-art MS instruments also paves the way to de novo sequencing, a database-independent strategy of spectrum identification [65]. The peptide amino acid sequences are directly inferred from MS/MS spectra. In the case of proteome analyses of less-well characterized model systems [66], microbial communities [67], splice-variants [68], mutations, and novel PTMs or multiple PTM searches [69], the prospect of de novo sequencing is alluring. The de novo technique is compelling for the assembly of full-length monoclonal antibodies [70, 71]. The computational tools for conducting de novo searches have improved significantly over the last few years. Instead of performing brute-force exhaustive searches [72] with all possible peptide sequences corresponding to the precursor mass, sub-sequencing methods [73, 74] and graph-based models [75, 76] with vertices corresponding to fragment ion peaks and edges to mass-differences between peaks [77] have gained in popularity. Even though many software solutions for de novo sequencing exist (i.e., PepNovo [78],

pNovo+[79], Novor [80], and the commercial PEAKS [81]), the technique has been used less than database searches so far. I have applied PEAKS (taggraph) to perform global PTM searches on both bait and interactor proteins in the MIP-MS project. In comparison to database-centric searches the de novo algorithms turned out to be less sensitive in identifying and quantifying specific PTMs.

2.1.8 Protein quantification

Unravelling function and dynamics of biological systems is catalyzed by the advent of quantitative proteomics experiments [82]. The main challenge is that peptide intensities measured by MS-instruments do not directly correlate with the peptide abundance in the biological specimen, as the molecular composition impacts ionization efficiency. However, both label-free and isotope-label-based quantitation methods have been developed to study differential protein expression within large dynamic ranges (1-10^5 or 10^6 and up to 10^{10} in plasma [83]). Experiments are often designed to determine protein abundance differences between conditions, e.g., genotypes, and drug treatments. To quantify proteins, both relative and absolute quantification can be performed.

2.1.8.1 Chemical and metabolic labeling strategies

Metabolic labeling requires cell lines to be cultured in medium with enriched stable isotopes. Stable Isotope labeling by amino acids in cell culture (SILAC) leads to the incorporation of “heavy” amino acids (often Arginine or Lysine with 13C, 15N, 2H) into the proteome [84, 85]. Also, whole organisms - e.g., mice [86], flies [87], plants [88] - have been reported to be SILAC labeled. SILAC labeling leads to a specific mass shift of the heavy compared to the light population. The multiplexing capability of SILAC is limited to the comparison of 3 conditions. Due to the mass-difference introduced by the isotopes, the quantification occurs at the survey scan level (MS1 level), where the intensity difference between each SILAC peptide pair is used as a measure for protein abundance difference [89]. Differential PTM analyses have also been performed with SILAC strategies [90, 91]. However, increased spectral complexity at the MS1 level due to individual precursors for isotope and non-isotope labeled peptides leads to lower numbers of identified peptides. As SILAC benefits from complete metabolic incorporation of isotope-labelled amino acids, analysis of primary, non-dividing tissues remains challenging. To address this issue and remove the time-consuming labeling step, entire labeled proteomes (super SILAC, [92]) or specific SILAC-labelled protein epitope signature tags (PrESTs, [93]) to use as spike-in standards have been invented. Targeted quantification strategies for absolute quantification of specific proteins involve the synthesis of peptides of interest with stable isotopes (AQUA [94, 95]) or the digest of artificial proteins (QconCAT [96]) to be added as an internal standard. Absolute quantification can be achieved by calculating the ratio between

the endogenous peptide and the internal standard with a known amount, but it is still limited to few proteins per sample.

Chemical labeling incorporates stable isotopes at the peptide level post digestion. Thereby, it can be readily used for primary cells or body fluids and tissues. Tandem Mass Tags (TMT) [97] or isobaric tags for relative absolute quantitation (iTRAQ) [98] are designed on the principle that isobaric labels are fragmented to generate so-called reporter ions in the low mass range, that can be accurately quantified. The labels often contain an amine-reactive NHS-group readily reacting with the N-Terminus and Lysines of peptides, a fragmentable chemical moiety, and a mass normalization. HCD [99-102], ETD [103] or CID [104, 105] is required for inducing linker fragmentation. TMT reagents with different isotope incorporations (13C, 15N) allow multiplexing of six [106] and even ten [107] samples per single experiment. However, high precision by isobaric tagging is accompanied by imperfect accuracy due to a systematic underestimation of ratios by co-fragmenting ions within the isolation window of the targeted precursor [108, 109]. Running an additional isolation and fragmentation cycle (MS3) has been suggested to eliminate the ratio compression problem [110]. Other isobaric tag strategies with more labile linkers have shown that quantitation of the labeled peptide (complementary reporter ions) instead of the reporter circumvents the ratio distortion, as well [111].

2.1.8.2 Label free quantitative proteomics

LFQ approaches are broadly applied and very popular and have been mainly used in my thesis. They are characterized by a relatively easy experimental setup and low cost compared to labeling strategies. LFQ has been performed by spectral counting or intensity-based methods. Instead of being confined to multiplexing capabilities of chemical or metabolic strategies, LFQ approaches theoretically allow the comparison of unlimited number of conditions. Indeed, higher number of protein identifications are often observed in LFQ experiments [112]. Importantly, additional sample preparation steps that are often required for chemical labelling strategies can introduce variation into the samples [113] and lead to the loss of proteins [114]. However, as single samples are compared across different LC-MS/MS runs, data quality can suffer from non-robust sample preparation, LC-performance issues and the stochastic sequencing by data-dependent acquisition methods.

Intensity-based label-free quantitation is grounded on the precursor signal intensity extracted from the extracted ion chromatogram (XIC). Similar to spectral counts, the intensity is also correlating with the initial protein abundance [115]. For intensity calculation either the peak height or area under the curve at specific retention times are employed.

2.2 Interaction proteomics

Protein function is often determined by their interaction with other proteins and also post translational modification status. To understand how signaling networks are orchestrated, knowledge on protein complexes composition is critical, as “no protein is an island entire of itself” [116]. To study protein-protein-interactions many techniques have been developed, including yeast two-hybrid (Y2H) [117, 118], protein-fragment complementation assay (PCA) [119], LUMIER [120] and FRET [121]. Mass spectrometry-based methods for the characterization of protein-protein interactions have advanced considerably over the last two decades. Qualitative approaches analyze the basal protein-protein interactions to elucidate protein-protein interaction networks, while quantitative approaches analyze interaction dynamics across different biological conditions [122].

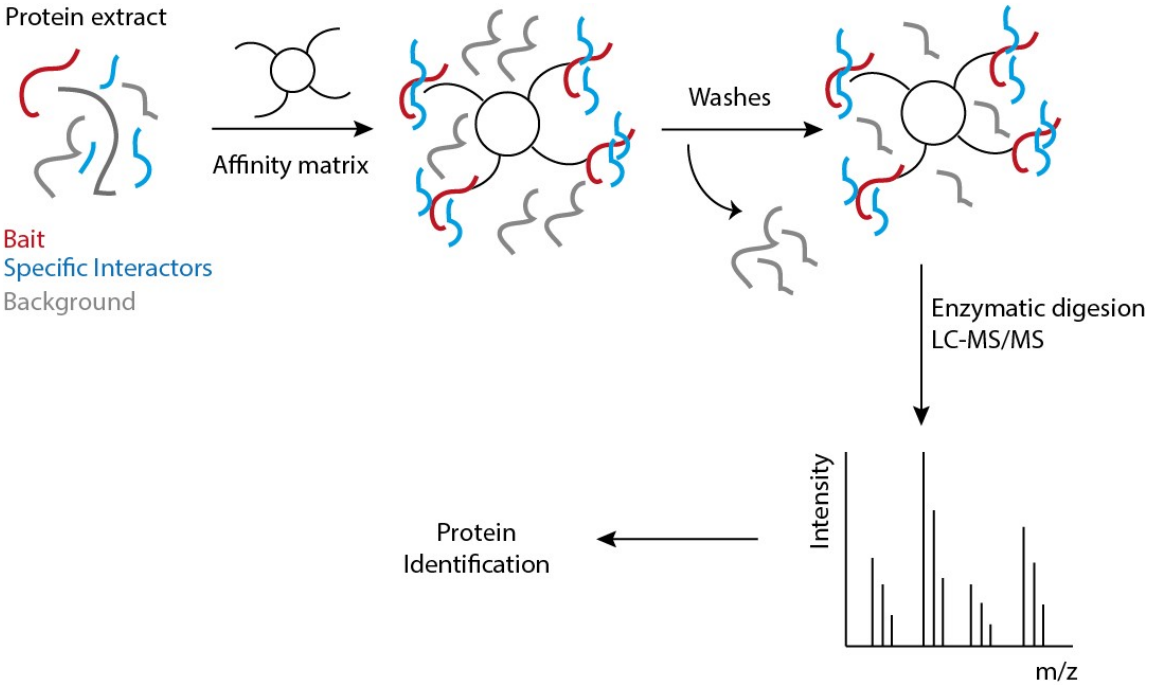


Figure 3: The pull-down-MS assay procedure is depicted schematically. Adapted from [123].

A long-standing and broadly used method to study PPIs is Affinity purification mass spectrometry (AP-MS). Here, proteins of interest in the soluble phase are captured by affinity matrix composed of an immobilized ligand on a solid support (agarose or magnetic beads). Ligands can be either antibodies raised against endogenous proteins of interest or against epitope tags, which requires prior expression – often overexpression – of tagged proteins. Purified proteins and complexes are directly subjected to LC-MS/MS analysis to unbiasedly identify and quantify their interactome – all interacting proteins in a protein complex. Due to the type of ligand, solid phase material, stringency of washes, background binding

proteins are co-enriched with protein complexes of interest. Historically, methods have tried to limit the background binding proteins to an absolute minimum by employing tandem purification strategies. However, newer concepts implemented the idea of affinity-enrichment MS, where quantitative MS is used to discriminate between background and protein complex members [124]. In my thesis, I have employed AE-MS to study the interactome of proteins involved in immune signaling in human immune cells.

2.2.1 AP-MS with antibodies raised against endogenous proteins

Focus on the endogenous protein is the main benefit of using antibodies directly raised against proteins of interest. Multiple isoforms can be studied, simultaneously, while there would be the need of individual experiments with epitope-tagging approaches. Both monoclonal and polyclonal antibodies have been used for AP-MS experiments, and also synthetic antibody-like molecules have been employed [125]. There are several challenges when working with protein specific antibodies: (1) High quality antibodies are often not available and generating specific antibodies of high quality is often associated with significant costs. (2) Antibodies may interrupt protein-protein interactions, if antibody epitopes are situated in protein-regions important for interactions. (3) As antibodies have varying affinity and specificities, affinity-purification conditions need to be adjusted with every experiment. (4) False positive interactors may be detected, as controls including isotype-matched antibodies against unrelated proteins or pre-immune serum are often lacking, especially due to the cross-reactivity of many antibodies. Reasonable controls for antibody specificity are knockout animals or knock-out/ knock-down cell-lines for the bait protein of interest [126]. Alternatively, multiple antibodies targeting different epitopes on the protein of interest can also reduce the number of false positives due to antibody cross-reactivity [127]. Especially for cell-lines or tissues, but also when the same protein of interest is compared against multiple cell-types or conditions, antibody-based AP-MS is a powerful technique [128].

2.2.2 AP-MS with epitope tags

Proteins can be fused to an epitope tag, that can be used for affinity purification. Here, the open reading frame (ORF) of a protein of interest is either C- or N-terminally appended by the sequence of the epitope tag, resulting in fusion proteins with affinity handles for subsequent affinity purification. In rare examples where N- or C-termini are presumably sterically hidden in bigger complexes, epitope tags have also been inserted in the middle of proteins [129]. One of the major advantages of epitope-based AP-MS is using the same epitope tag (and thereby also affinity matrix) to purify multiple proteins. This transferability of the enrichment strategy between different pulldowns allows for streamlined controls due to similar background binding proteins and opens the possibility to high-throughput interactome studies. Many

different tags including peptide tags (small) and protein tags (big) have been described for AP-MS experiments (see Table 1: Affinity tags for AP-MS experiments). Also, tandem affinity tags have been described: they allow the purification of proteins of interest with several rounds of purification using different affinity moieties. For our study, we have chosen the His-tag in conjunction with an IMAC resin, due to the following reasons: (1) His-Tag is a small peptide tag; (2) The affinity matrix for His-tag enrichments is not proteinogenic, thereby on-bead digests after affinity purification is possible. Proteinogenic matrices like antibodies or streptavidin might lead to masking of interactors due to their high abundance as matrix proteins (3); The very stable His-IMAC complex binding also allows for denaturing and very stringent affinity washing procedures. (4) The IMAC affinity matrix is economically the cheapest choice. Further, we have compared His-IMAC to gold-standard AP-MS approaches like Strep-tag and Flag-tag and have found similar results in terms of interacting proteins identified, bait protein sequence coverage and total bait protein intensity.

Table 1: Affinity tags for AP-MS experiments

Tag	**Sequence/ MW**	**Affinity resin**	**Original reference**
Peptide tags			
c-myc	EQKLISEEDL	Anti-c-myc (9E10)	[130]
FLAG	DYKDDDDK	Anti-FLAG (M1, M2, M5)	[131]
HA	YPYDVPDYA	Anti-HA (12CA5)	[132]
His-Tag	HHHHHH	Ni^{2+}/Co^{2+}-NTA/CMA	[133]
Strep-tag II	WSHPQFEK	Streptavidin	[134]
Protein tags			
GST	26 kDa	Glutathione	[135]
GFP	26.9 kDa	Anti-GFP	[136]
Protein A	45 kDa	IgG	[137]

Adapted from: [128]

An important factor to consider in every epitope-tagging approach, is the required over-expression of exogenous fusion proteins which can be achieved either by transfection or viral transduction. Transfection strategies are leading to transient cell-lines and are often limited to standard model cell-lines like HEK293T or HeLA cell-lines. Viral transduction results in stable cell-lines and is also applicable to non-dividing

primary cells and non-standard cell-lines. However, in such experiments the exogenous protein expression is neither under the control of the endogenous promoter, nor are introns or untranscribed regions (3' or 5' UTRs) involved in the expression regulation. Accordingly, major disadvantages of exogenous protein overexpression are potential effects on protein folding, protein localization and protein regulation. Additionally, due to overexpression, proteome expression differences compared to wildtype cells have been observed [138]. Whenever proteins are exogenously expressed, a total proteome analysis is valuable to control for possible expression differences in transgenic cell-lines. Several concepts have been devised in the past to limit false-positive or false-negative interactor calling in experiments involving AP-MS with tagged bait proteins [139-141]: (1) Weak overexpression of proteins of interest under weak promoters [142-145]. (2) Protein expression induction by inducible promoters (i.e., Tetracyclin inducible promoters) [146-148] in combination with single-genome integration of the epitope-tagged gene of interest via the Flp-In system [149]. (3) Bacterial artificial chromosomes (BACs) contain the native gene architecture with most regulatory elements to enable close to endogenous epitope-tagged protein expression [150, 151]. This versatile system has been used in whole proteome covering interactome studies [152]. (4) With the advent of gene-editing technologies like CRISPR and the recently described prime-editing methods, true endogenous gene epitope-tagging has been achieved with relatively little effort and high efficiency [153]: it consists of a prime editing guide RNA (pegRNA) and a prime editor (PE) to edit – either by inserting of deleting – the gene of interest employing both a Cas9 nickase and a reverse transcriptase, in concert.

2.2.3 Interactor calling in AP-MS

Even though affinity purification in theory only enriches protein complexes of interest, many additional background proteins are also co-enriched. This is due to non-specific interaction of proteins with the antibody, streptavidin (e.g., in vivo biotinylated proteins) or IMAC (proteins containing multiple His in adjacency) or also unspecific binding to the bead material (e.g., agarose, sepharose, magnetic beads). Additionally, background proteins are highly organism, tissue and even cell-type specific. In the past, different strategies to discriminate between true interactors and background proteins have been employed: Pre-MS-analysis reduction of background proteins (biochemical) or post-MS-analysis reduction of contaminants (bioinformatic).

Biochemical reduction of contaminants: Stringent washing steps with high concentrations of salt and detergent can reduce the background-bound proteome, which can result in not identifying weaker and more transient interactions [154]. Considering shorter purification times and low amounts of affinity

matrix also reduces non-specific interactors [155-157]. Additionally, purifying proteins with multiple purification steps via tandem affinity purification (TAP, [158]) has been successfully applied in many organisms, including yeast [159, 160] and mammalian cells [161-163] . In TAP, the non-specifically binding proteins from the first AP round can be removed by a second round of AP employing different affinities and affinity matrices. Nevertheless, even the most stringent protocol might still lead to enrichment of background proteins, highlighting the use of proper controls.

Bioinformatic reduction of contaminants post MS analysis: MS based proteomics delivers quantitative information on proteins in a sample. Accordingly, a true interactor will be enriched over all the bait-specific enrichment fractions, while the background remains unchanged. Many computational approaches have been established both for label-free and metabolic/chemical labelling datasets, to distinguish between bait/ interactors and background binding proteins. The simplest, but also most erroro-prone solution is fixing an enrichment threshold for putative interactors. Standard statistical tests in combination with threshold enrichment like QUBIC [124, 151, 152], probabilistic approaches like CompPASS [142, 145, 164] or SAINT [165] have been used successfully for identifying significant interactors. In addition, the Contaminant Repository for Affinity Purification (CRAPome) built on the SAINT algorithm describes common false-positive interactors in a cell-type and enrichment strategy dependent manner [166].

2.2.4 New approaches for identification of protein-protein interactions

While static interactions can be described very reliably with AP-MS approaches, the more dynamic or transient interactions are harder to capture. Many factors such as PTMs, conformational changes influence the stability of PPIs [167]. With AP-MS as *ex vivo* approaches, especially time of purification experiment, the dilution effect during lysis procedures, stringency including salt and detergent conditions and temperature can influence the achieved results [123]. Time-controlled SILAC experiments have been employed to comprehensively describe transient versus static interactions via AP-MS: SILAC labelled and non-labelled samples are mixed before purification, allowing transient interactors to switch between light and heavy forms [168]. In my project, I have studied transient PPIs in response to TLR activation in human immune cells with an AP-MS method. Other strategies to study transient interactions including stabilizing PPIs with crosslinking approaches and proximity dependent labelling are described below.

Cross-linking mass spectrometry (XL-MS): To stabilize transient PPIs, cross-linking approaches [169] either based on formaldehyde [170-172], photo-inducible amino acids [173] or NHS-chemistry (i.e., DSP)

[174] have been described. These crosslinkers consist of two reactive groups separated by spacer arms of defined length [175], which can either lead to crosslinks between two spatially close proteins, between reactive amino acids within the same protein. Crosslinking has been used together with AP-MS in targeted approaches, where interactions of a particular protein are of interest [176]. However, also in global interactome studies, XL-MS has been employed. The main challenge of cross-linking mass spectrometry (XL-MS) involves the deconvolution of the heterogenous samples, as cross-linked peptides result in

Proximity dependent labelling (PDL): Recently, proximity dependent labeling (PDL) methods have emerged, identifying potential interactors dependent on spatial proximity. The main methods include: BioID [177], Ascorbate Peroxidase (APEX) [178] and especially for membrane proteins selective proteomic proximity labeling assay (SPPLAT- HRP based) [179] or enzyme mediated activation if radical sources (EMRAS – HRP based) [180, 181]. In PDL, a bait protein of interest is fused i.e., to a promiscuous biotin ligase (BirA/ BirA*) or an ascorbate peroxidase (APEX), which biotinylate neighboring proteins with reactive biotin species. The biotinylated proteins are captured via Streptavidin affinity purification and further subjected to LC-MS analysis to unbiasedly identify interacting proteins. Thereby, proximity labelling approaches allow a “snapshot” of PPI rearrangements in dynamic processes [182].

2.3 Post-translational protein modifications (PTMs)

Protein modifications after protein translation can regulate protein activity by affecting protein interaction, protein localization, protein folding and protein stability. PTMs occur in a time and localization specific manner. Reversible and also irreversible, enzymatically catalyzed and also no-catalyzed chemical protein modifications have been observed. Proteins can be modified by relatively small chemical groups (e.g., phosphorylation, methylation, Acetylation), more complex molecules (e.g., glycosylation, isoprenylation) or even small proteins (e.g., ubiquitinylation, sumoylation, isgylation). A special form of protein modification is proteolysis that can be observed during signal peptide cleavage and I studied in this PhD Thesis. While antibody-based approaches in combination with immunoblotting allow the study of specific protein modifications, MS based proteomics is so far the only appropriate method to comprehensively identify and quantify PTMs [183]. In the following section, I describe selected PTMs, that were studied in depth by targeted (phosphorylation, Isgylation) and global approaches (proteolysis by signal peptidases) in the course of this PhD thesis.

2.3.1 Phosphorylation

Many cellular processes are regulated by phosphorylation [184]. Protein phosphorylation is induced by kinases and reversed by phosphatases, a tightly controlled equilibrium to modulate protein activity across

signal transduction pathways [185, 186]. Up to 23% of intracellular adenosine-triphosphate (ATP) is used by protein kinases for substrate phosphorylation [187, 188]. Dysregulation of protein phosphorylation can be the origin of various diseases [189, 190] and deeper understanding of phosphorylation induces signaling pathways might herald the discovery of novel therapeutic reagents [191]. Over 50 kinase inhibitors - including small molecules and targeted antibodies - have been approved by clinical trials [192, 193] and attribute to the importance of generating knowledge about protein phosphorylation.

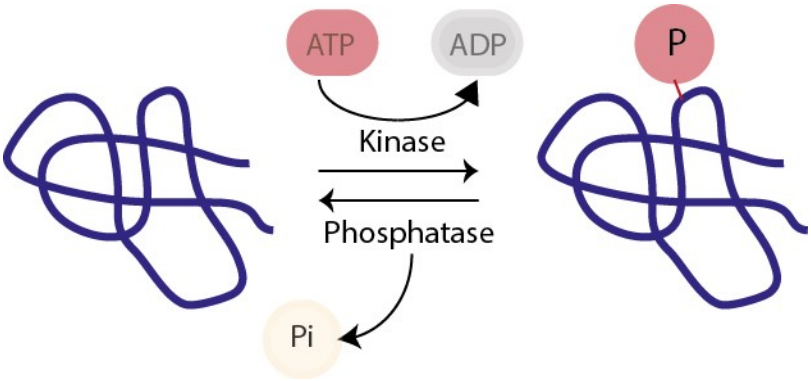


Figure 4: Reversible phosphorylation and dephosphorylation of phosphoproteins by kinases and phosphatases.

Protein phosphorylation is catalyzed by kinases and includes the transfer of phosphate from ATP to the hydroxyl group of serine, threonine or tyrosine [194], leading to the covalent addition of phosphate group PO_3. Recently, also non-canonical phosphorylation of the amino acids histidine, aspartic acid, glutamic acid, arginine, cysteine and lysine has been described in detail [195-197]. However, despite the importance of protein kinases, the majority of described phosphorylation sites have no known protein kinase or biological function [198]. To study the biological impact of phosphosites identified in global screening approaches, the site-specific mutation of phosphorylated amino acids due to chemical similarity has been employed in the past: to mimic constitutive phosphorylation Ser/ Thr are mutated to Asp/Glu, while phosphorylation deficient mutants are generated by mutating to Ala, respectively. I have applied this technique to further characterize phosphorylation sites identified on TRAF2.

In the past decade, global phosphoproteomics has advanced considerably due to new highly sensitive and accurate MS instruments, measurement strategies, novel algorithms and software as well as progress in sample preparation, which altogether allows the routine identification of several thousand phosphosites in single DDA shotgun experiments [199]. Even higher coverage in combination with shorter gradients (> 20 000 phosphosites in 15 min) can be achieved by DIA based measuring methods [200]. As phosphorylation is a substoichiometric PTM, phosphoproteomics benefits especially from robust

enrichment methods, which can be based either on immobilized metal affinity chromatography (IMAC: Fe^{3+}) [201] or metal oxide affinity chromatography (MOAC) with enhancers [202-204] – e.g., titanium oxide (TiO_2, [205]). Higher specificity for MOAC based enrichment of phosphopeptides in comparison to non-phosphorylated peptides was achieved by addition of organic acids [206, 207]. As both IMAC and MOAC favor serine and threonine phosphorylation, antibody dependent enrichments are employed for global studies of tyrosine phosphorylation [208].

2.3.2 Isgylation

Interferon stimulated gene 15 (ISG15) belongs to the protein family of ubiquitin-like modifiers [209] and is induced by interferon to execute antiviral activity [210]. Additionally, ISG15 can be induced by multiple other stimuli like retinoic acid, LPS and genotoxic stress [211-213]. Both intracellular and viral proteins can be covalently modified by ISG15, by conjugating the C-terminal glycine residue of the LRLRGG motif, that can also be found in ubiquitin, to a lysine of the target protein [214] – a process termed isgylation. Isgylation requires the sequential action of three enzymes: an E1 (UBA7), an E2 (UBE2L6) and so far three described E3 enzymes (HERC5/6, TRIM25, HHARI). The covalent modification with ISG15 can be reversed by the ISG15 specific protease termed USP18 or UBP43 [215]. Isgylation of intracellular proteins has been shown to affect trafficking and protein stability, by competing with ubiquitin [216, 217]. Additionally, ISG15 has been reported to have antiviral roles as a free intracellular molecule through non-covalent interactions [218, 219] and also was suggested to have extracellular functions as a cytokine [220-222].

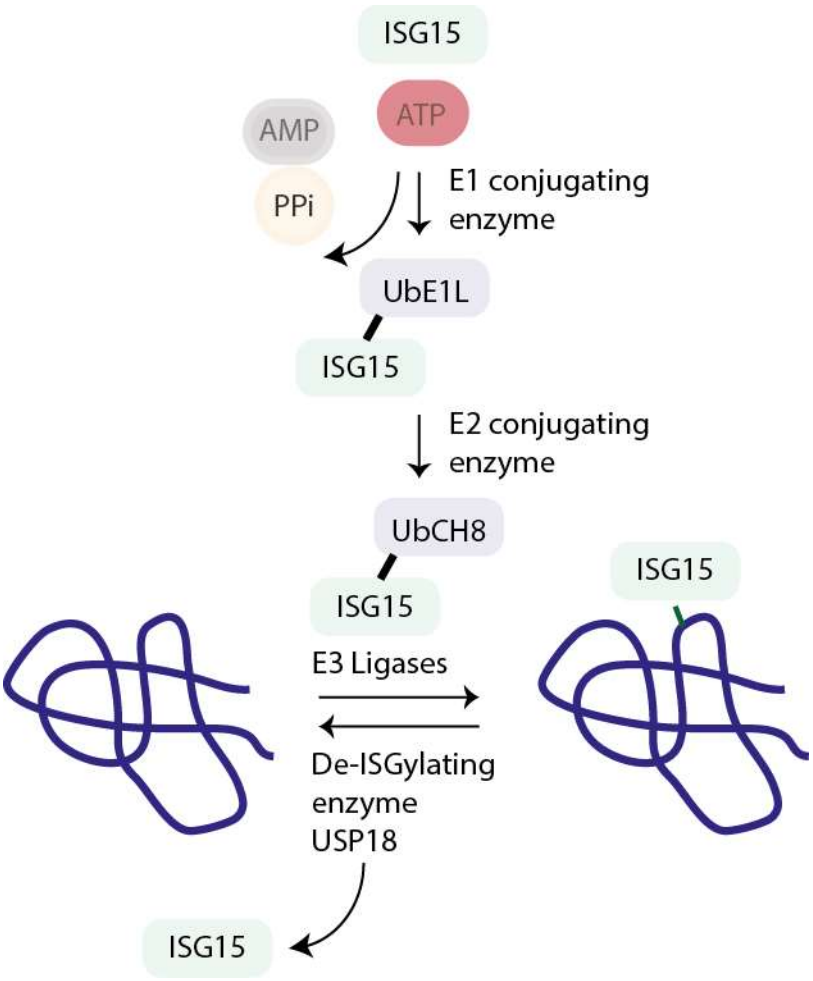


Figure 5: Schematic representation of protein isgylation and de-isgylation. Adapted from [223].

In typical MS-based proteomics workflows, enzymatic digestion with trypsin leads to typical peptide remnants of two glycine residues (di-Gly) on lysine by ubiquitin/ ubiquitin-like proteins and the release of conjugated ubiquitin/ ubiquitin-like proteins. The mass difference of 114 Da introduced by the di-Gly adduct can be readily identified and quantified. As ubiquitinylation and also isgylation are substoichiometric PTMs, enrichment of modified peptides is required prior to MS analysis: this is usually achieved by employing antibodies raised against di-Gly [224-226]. However, ubiquitin-like modifications like isgylation, neddylation, sumoylation are attributed only a small amount (< 6%) of all the observed di-Gly remnant peptides [227]. Therefore, additional strategies are required to detect protein isgylation. Initially, isgylation target identification was achieved by expressing ectopic tagged ISG15 expression and affinity purification [228-230]. As we have identified ISG15 as an interaction partner of TRAF2, we have indirectly described TRAF2 isgylation by combining TRAF2 isgylation-deficient mutation, pulldown and GlyGly enrichment. Other global isgylome studies were conducted with the combination of GlyGly

enrichment *Isg15*-deficient mice in comparison to wildtype animals and additionally knock-in of inactive USP18 mutants, to abrogate deisgylation [231]. However, missing highly specific antibodies, global quantitative isgylome studies are dependent on prior genomic modification.

2.3.3 Proteolysis

The hydrolysis of peptide bonds by proteases is an irreversible post-translational modification process, that gives rise to multiple isoforms for native proteins. Proteolytic cleavage results in neo-N or neo-C protein termini and frequently causes change in structure and function [232]. In humans, about 3% of the human protein coding genes are estimated to be proteases, corresponding roughly to 560 potential proteases [233]. Proteases a categorized by their catalytic mechanism into metallo, serine, cysteine, aspartic and threonine classes [234]. Proteolysis orchestrates processes such as fertilization, tumor metastasis, angiogenesis, cell death, aging, maturation and growth by controlling the activity of cytokines, hormones, growth factors and other enzymes [reviewed in:235]. Hereditary diseases, developmental complications and disease manifestations are characterized by aberrant proteolysis [236-238]. To identify protease substrates, three major methods have been employed, including library screening (synthetic peptide or phage display libraries), bioinformatics and proteomics approaches. As only short peptide sequences are presented in library approaches, the influence of exosites and protein three-dimensional structure is not accounted for in library approaches [239].

More reliable substrate identification can be achieved by proteomics approaches screening for native substrates in intact cellular context either by triggering the activity of proteases or artificial addition of the protease of interest to a cellular lysate. One of the challenges of protease substrate identification in intact cellular context is the possibility of multiple substrate cleavage by diverse proteases. Thereby, proper controls – e.g., different genotypes, specific protease inhibitors - are needed for the *bona fide* discovery of specific protease substrates. Due to the low abundance of protease cleaved peptides, identification of protease substrates is often coupled to enrichment of neo N-/C-termini peptides.

Cleavage products can be isolated either by positive [240] or negative [241] enrichment of newly formed N-termini. Positive enrichment directly enriches N-terminal peptides and is frequently combined with biotin-avidin affinity purification, e.g., by (1) enzymatic biotinylation of N-termini via subtiligase and engineered mutant enzymes [242, 243] or (2) chemoselective biotinylation of α-amines of proteins [244]. Also, TMT-tagging of α-amines in combination with prior blocking of lysine residues has been described [245]. Negative enrichment involves the removal of non-N or C-terminal peptides, thereby enriching the

terminal peptides, also allowing the analysis of modified N-termini e.g., by acetylation or N-terminal glutamine cyclization [246].

2.3.3.1 Signal peptide cleavage by SPP

Signal peptide peptidases (SPPs) and signal peptide peptidase-like protease (SPPLs) – presenilin homologues - are intramembrane-cleaving proteases sharing a conserved Y/FD and GxGD motif in the catalytic center [247, 248]. As polytopic membrane proteins, the active sites of SPPs are localized in two adjacent transmembrane domains and induce the peptide bond hydrolysis of substrates. Experiments including transition state analogue inhibitors and site-specific-mutation of the catalytic Asp corroborate SPP/SPPLs are aspartyl proteases [249-251]. The physiological importance of SPP/ SPPLs is emphasized by their conserved occurrence in eukaryotes, comprising of fungi, protozoa, plants and animals [247-249]. Five members (SPP, SPPL2a, SPPL2b, SPPL2c, and SPPL3) of the SPP/SPPL family are classified in mammals [252], but number and also specificity of SPP/SPPLs is highly organism specific. The subcellular localization of the individual members is differing: SPP localizes to the endoplasmic reticulum [253], SPPL3 to the Golgi apparatus [254], SPPL2a to the endo/lysosomal compartments [255], SPPL2b to the plasma membrane [254] and SPPL2C to the ER and ER-Golgi intermediate compartment (ERGIC) [254, 256].

Typical substrates of SPP contain signal sequence-derived signal peptides [257], however also other substrates of SPP/ SPPLs been reported, recently: this includes type II transmembrane proteins without predicted signal sequences e.g., TNF-a [250] and CD74 [258, 259].

The mechanism of co-translational signal peptide cleavage is a multistep process and has not been fully elucidated at the molecular level, yet [reviewed in:260]: protein translation is initiated in the cytoplasm and temporarily halted or slowed by the binding of the signal recognition particle (SRP) to the emerging signal peptide. SRP mediates the transfer of the ribosome-bound nascent protein chain to the endoplasmic reticulum in a SRP receptor dependent manner, which leads to the transfer of the nascent protein chain to the translocon, which possesses a so called lateral gate. As translation resumes, the N-terminus of the signal peptide which first faces the ER-lumen, flips-turns to face the cytosol in a type II transmembrane protein reminiscent way: this step is crucial for the signal peptide to fully move into the lateral gate, therefore clearing the space for the elongating protein chain. At the luminal surface, the signal peptide is cleaved of by SPP and chain elongation resumes until the STOP codon is reached, initiating the ribosome dislocation from the ER. While the structure of the translocon is conserved through all domains of life [261], signal peptides are comprised of variable primary structures specific for every protein and organism.

Signal peptides are localized at the N-terminus of nascent proteins (e.g., membrane proteins, secretory proteins). The tripartite structure of signal sequences is defined by a cationic N-terminal region, a hydrophobic h-region with at least six non-hydrophilic amino acids and a C-terminal region, including the cleavage site of signal peptidases [262]. Hereditary diseases (e.g., Classic Ehlers-Danlos Syndrome [263], Crigler-Najjar Disease [264], Hereditary Angioedema [265], etc.) with reduced protein expression are correlated with mutations in the signal sequences, in particular an insertion of a polar amino acid into the h-region [266]. Interestingly, signal peptide dependent control of the efficiency of protein compartmentalization [267] and attenuation of protein translocation in response to ER stress [268] can regulate protein expression. For example, an insertion of two amino acids in the h-region of the human lutenizing hormone receptor (LHR) – a commonly observed genetic polymorphism – enhances the translocation of LHR and thereby the protein abundances, which can have a negative impact on cancer survival rates in women [269].

The prediction of signal peptides and the resulting subcellular localization has been studied deeply by bioinformatic approaches and involves both (1) the discrimination between signal peptide containing proteins and non-secretory proteins and (2) prediction of the actual signal peptide cleavage position to determine the exact N-terminus of the protein [270]. Prediction algorithms have advanced with methodological developments in the field of bioinformatics, starting out from simple statistics and weight matrices and maturing into employing artificial neural network algorithms [271, 272], hidden Markov models [273], support vector machines [274] and deep and recurrent ANNs [275]. Initially, distinguishing between signal peptides and N-terminal transmembrane helices was challenging and lead to false positive predictions from N-terminal transmembrane regions. An advanced version of SignalP [276] based on neural network-based methods solved this issue by including two negative control datasets of (1) cytosolic/ nuclear proteins and (2) proteins without signal peptides and transmembrane regions in the beginning of the protein [277]. To date, the occurrence of most signal peptides is mostly predicted and not experimentally verified (Project 2). In our project 2, we have applied meta-analysis of proteomics data to comprehensively describe signal peptide cleavage in human and murine tissues.

2.4 Secretomics

The secretome is composed of all the components secreted by cells, tissues or organs, including extracellular matrix proteins, enzymes, growth factors, inflammatory cytokines, exosomes, and microvesicles [278]. Secretory proteins are fundamental for multidirectional intercellular communication governing processes such as proliferation, growth, migration and metabolic regulation [279]. Secretory

proteins have a wide range of functions, from roles in the immune system to neurotransmitters in the nervous system.

Secretomics is a subfield of proteomics studying the secretome that offers a powerful method for identifying and quantifying proteins secreted by a sender cell population under particular biological conditions [280, 281]. Secretome studies have been used, i.e., in the analysis of drug resistance [282], inflammation [280, 283], extracellular matrix remodeling [284], and tumor metastasis [285-288]. Secretome research is currently dominated by unbiased label-free qualitative and quantitative proteomics analyses based on LC-MS/MS methods, in which secretome samples are trypsin-digested either in-gel or in-solution. Secretome processing can be done using a variety of methods other than MS, including DNA microarray, RNA sequencing, Serial Analysis of Gene Expression (SAGE), and antibody or bead array [289]. Next to MS, another important method for secretome analysis is analysis with antibody array. Antibody array-based secretome analysis has been used in the study of cardiovascular disease and liver cancer [290, 291]. A possible drawback of antibody arrays is the quality and specificity of the antibodies (e.g., recognizing highly glycosylated secreted proteins) and the limitation of comprehensiveness in comparison to unbiased proteomics approaches. I have applied MS-based secretomics workflows in various collaborative projects (Project 4, Project 5) and have further written a step-by-step protocol for performing secretomics in immune cells (Project 3).

The majority of secretome studies in mammalian cells are conducted *in vitro*, with cells of interest first being cultured in serum-supplemented medium to obtain a sufficient number of cells for analysis. Prior to the start of the experiment, serum-containing medium is depleted and exchanged for serum-free medium, as highly abundant proteins in serum would otherwise lead to MS interference. While secretome analysis in cultured serum-containing conditioned media is possible, it necessitates comprehensive protein or peptide fractionation [292]. Alternatively, azidohomoalanine (AHA) labeling was used for secretome analysis from cells grown in serum supplemented with unnatural amino acids [293]. Duration of the experiment in serum-free conditions – usually just normal culturing medium without addition of FCS - has to be carefully assessed due to decreasing cell viability in the absence of serum [294]. Alternatively, for more challenging cells like primary cells, approaches with special serum-free medium – i.e., medium also used for the industrial expression of antibodies – can be applied.

2.5 Biological rational: TLR signaling

Toll-like receptors (TLRs) are expressed by a variety of cell types, including macrophages, monocytes, dendritic cells (DCs), neutrophils, natural killer cells, and fibroblasts [295, 296]. TLRs are type I

transmembrane proteins with leucine-rich repeats (LRRs) in their ectodomains, transmembrane domains and intracellular Toll–interleukin-1 (IL-1) receptor (TIR) domains. Microbe-associated molecular patterns (MAMPs) produced by microbes and damage-associated molecular patterns (DAMPs) produced by dying or wounded cells are recognized by LRRs. Adaptor molecules bind to the TIR domains to activate downstream signaling pathways. Numerous species, most notably mammals, express the TLR family, and thirteen distinct TLR forms have been identified. TLR1 to TLR9 are conserved between mice and humans. However, a retroviral insertion made the TLR10 molecule inactive in mice. Humans lack TLR11, -12, and -13. Active TLRs have a distinct localization pattern. TLR1, -2, -4, -5, -6, and -10 are expressed on the surface of the cell, while TLR3, -7, -8, -9, -11, -12, and -13 are expressed in the endosome [297, 298]. Each TLR has a distinct role in terms of MAMP detection and immune responses, as shown by studies on mice defective in each TLR.

TLRs on the cell surface identify a wide range of ligands including elements of microbial membranes, such as lipids, lipoproteins, and proteins [299]. TLR4 senses lipopolysaccharides (LPS) derived from Gram-negative bacteria and syncytial virus envelope proteins, trypanosome glycoinositol phospholipids, and heat shock proteins 60 and 70, S100A8 originating from dying cells [300, 301] and free fatty acids [302, 303]. Many transmembrane accessory molecules – such as CD14 and MD-2 - play a role in the activation of TLR signaling. CD14 and MD-2 molecules recognize LPS and internalize TLR4 into endosomes, activating the TLR4 signaling cascade via SYK and phospholipase C-2 [304]. Another member of the TLR family - TLR2 - works in concert with TLR1 or TLR6 to identify a diverse array of MAMPs derived from Gram-positive bacteria, including lipoproteins, peptidoglycans, lipoteichoic acids, zymosan, and mannan. TLR2-TLR1 heterodimers recognize triacylated lipoproteins, while the TLR2-TLR6 heterodimers recognize diacylated lipopeptides. We have used TLR2 activation of monocytes, to study dynamic PPIs and PTMs of several important players of the TLR signaling cascade via affinity purification in publication 1.

When MAMPs and DAMPs bind to their respective TLRs, adaptor molecules - such as MyD88, TRIF, TIRAP, or TRAM - are recruited to the receptor. There are two major pathways downstream of TLR4: MyD88-dependent and MyD88-independent (TRIF-dependent).

In the MYD88-dependent pathway, TLR4 recognition of a ligand induces the formation of a complex between the TIR-domain-containing adaptor molecule MyD88 and an IRAK family molecule through the adaptors' cytoplasmic region [305]. Following complex formation, IRAK4 phosphorylates IRAK1 to release IRAK1 from MyD88 [306, 307]. IRAK1 forms an association with TRAF6, a RING-domain E3 ubiquitin ligase. Then, TRAF6, in conjunction with the ubiquitin E2-conjugating enzyme complex (UBC13 and UEV1A),

promotes the polyubiquitination (K63-linked) of TRAF6 and the TAK1 protein kinase complex [308, 309]. TAK1 interacts through ubiquitin chains with the IKK complex and activates two distinct signaling pathways, the IKK complex– NF-κB and the IKK complex–MAPK. The IKK complex - composed of IKKα, IKKβ, and NEMO (NF-κB critical modulator) - phosphorylates and then degrades the NF-κB inhibitory protein IB, allowing NF-κB to translocate to the nucleus. Also, TAK1 activates the MAPK signaling pathway, triggering the inflammatory response [295, 298].

In the MYD88-independent pathway, other TIR-domain-containing adaptor proteins, such as TRIF and TRAM, function as TLR4 adaptor proteins. When LPS binds to TLR4, TRIF is recruited to the cytosolic region of TLR4 and activates IRF3 and NF-κB, resulting in the expression of type I IFNs and pro-inflammatory cytokines. TRIF activates downstream signaling via TRAF6 and TRAF3. Interaction with TRAF6 leads to RIP-1 binding, and this complex then cooperatively activates TAK1, activating NF-κB and inducing the expression of pro-inflammatory cytokines. Unlike TRAF6, TRAF3 activates the IKK-related kinases TBK1 and IKKi, which phosphorylate IRF3 to translocate to the nucleus, resulting in the expression of type I IFNs [295, 298].

3 Aims of the thesis

In my thesis, I developed and applied different cellular, biochemical and MS-based quantitative proteomics methodologies, to investigate the protein-protein interactions and post-translational modifications in immune cells and established a computational workflow for the detection of the signal peptide peptidase induced proteolysis of signal proteins.

First, I developed a streamlined and scalable method to simultaneously study dynamically regulated PPIs and PTMs, which are usually analyzed in experimentally distinct workflows. The method is based on the enrichment of epitope-tagged proteins with the His-IMAC affinity matrix and provides a cost-effective and non-proteinogenic matrix for AP-MS experiments, which we termed MIP-APMS (Modifications, Interactions and Phenotypes by Affinity Purification Mass Spectrometry). We evaluated and optimized all steps of MIP-APMS, including (1) epitope-tagging of proteins of interest and mammalian cell transduction, (2) affinity purification conditions for optimal interaction network and PPI enrichment, (3) followed by MS-based quantification and identification of PTMs and PPIs, and (4) ultimate biochemical and phenotypic validation of interactors and PTMs in primary human immune cells. We demonstrate the pipeline's discovery potential by probing dynamically assembled protein communities in human monocyte immune signaling using Toll-like receptor (TLR) 2 activation and MAPK14 inhibition. Our analysis of 19 protein complexes discovered over 50 previously unknown PTMs, including phosphorylation, acetylation, methylation, isgylation, and other less well documented chemical changes, as well as an interaction network spanning over 300 PPIs.

Second, I studied the proteolysis of signal peptide containing proteins by signal peptide peptidase to elucidate the exact N-Terminus of these proteins. To this end, I developed and applied a tailored bioinformatics strategy to identify neo-N-terminal peptides from publicly deposited MS raw files. This meta-analysis identified not only the cleavage sites predicted by prediction algorithms, but also alternative cleavage sites in the vicinity of the predicted cleavage site. Annotated enrichment analyses revealed a particular high amount of alternative cleavage sites in transmembrane proteins.

Furthermore, I devised a step-by-step proteomics protocol for secreted proteins, wrote a Commentary on macrophages and contributed to collaborative research projects assessing the paracrine functions of (1) immune cells undergoing apoptosis and necroptosis, (2) and pyroptosis as well as (3) the influence of arginine on the metabolism and development of multinuclear giant cells.

4 Publications

4.1 Project 1: Identification of Covalent Modifications Regulating Immune Signaling Complex Composition and Phenotype

Protein-protein interactions and post-translational modifications are typically studied independently. In this project, I present a cost-effective experimental and analytical approach for simultaneously identifying all detectable post-translational modifications (PTMs) and protein interactions (PPIs) in relevant signal transduction pathways using mass spectrometry (MS)-based proteomics. This technology permits the interrogation of protein signaling complex composition and function at single-amino-acid resolution in response to receptor activation or pharmacological inhibition in nearly any cell type, including primary cells. The method's application to monocytes elucidates the functional relationships between numerous previously unknown PTMs and PPIs involved in immune signaling, including the first description of TRAF2 isgylation.

This method – Modifications, Interactions, and Phenotypes by Affinity Purification Mass Spectrometry (MIP-APMS) – entails a rapid and streamlined pathway for cloning and transduction of tagged proteins of interest into functionalized reporter cells, followed by affinity chromatography and MS-based quantification of all enriched proteins with their covalent modifications. MIP-APMS elucidates the time-resolved interactions of over 600 proteins and over 100 modifications, including phosphorylation, acetylation, and methylation, as well as chemically less well-characterized covalent modifications at certain signal transduction phases. Due to the lack of systems-wide approaches to study protein isgylation, the unique discovery potential of MIP-APMS is highlighted by the identification of an activating function for isgylation in monocyte innate immune signaling. Our pipeline enables rapid validation of novel hypotheses, such as the co-regulation of various protein interactions and alterations with a distinct functional cellular phenotype, after CRISPR-Cas9-mediated knockout or site-specific mutation in the same experimental system. As a result, we established the functional significance of TRAF2 isgylation and phosphorylation, the interaction of ARHGEF18 with MAP3K7, and the identification of the target versus off-target effects of small molecule p38 inhibitors in human monocytes.

MIP-APMS may thus enable rapid and systematic exploration of dynamic protein communities in a wide variety of fields of biology, and serves as a template for how innovative technology reveals a myriad of previously unknown functional molecular checkpoints in cellular signaling cascades.

The corresponding paper was published in Molecular Systems Biology [310].

I have performed all experiments, developed and implemented the bioinformatics methods. Together with Felix Meissner, I conceived the data analysis, interpreted the data and wrote the manuscript.

Article

molecular systems biology

Identification of covalent modifications regulating immune signaling complex composition and phenotype

Annika Frauenstein[1], Stefan Ebner[1], Fynn M Hansen[2], Ankit Sinha[2], Kshiti Phulphagar[1], Kirby Swatek[3], Daniel Hornburg[4], Matthias Mann[2] & Felix Meissner[1,5,*]

Abstract

Cells signal through rearrangements of protein communities governed by covalent modifications and reversible interactions of distinct sets of proteins. A method that identifies those post-transcriptional modifications regulating signaling complex composition and functional phenotypes in one experimental setup would facilitate an efficient identification of novel molecular signaling checkpoints. Here, we devised modifications, interactions and phenotypes by affinity purification mass spectrometry (MIP-APMS), comprising the streamlined cloning and transduction of tagged proteins into functionalized reporter cells as well as affinity chromatography, followed by MS-based quantification. We report the time-resolved interplay of more than 50 previously undescribed modification and hundreds of protein–protein interactions of 19 immune protein complexes in monocytes. Validation of interdependencies between covalent, reversible, and functional protein complex regulations by knockout or site-specific mutation revealed ISGylation and phosphorylation of TRAF2 as well as ARHGEF18 interaction in Toll-like receptor 2 signaling. Moreover, we identify distinct mechanisms of action for small molecule inhibitors of p38 (MAPK14). Our method provides a fast and cost-effective pipeline for the molecular interrogation of protein communities in diverse biological systems and primary cells.

Keywords mass spectrometry; posttranslational modifications; protein interactions; proteomics; signaling networks

Subject Categories Immunology; Proteomics

DOI 10.15252/msb.202010125 | Received 16 November 2020 | Revised 8 July 2021 | Accepted 8 July 2021

Mol Syst Biol. (2021) 17: e10125

Introduction

Cellular functions rely on complex molecular networks that are mainly composed of proteins (Seet *et al*, 2006; Pan *et al*, 2012). Cell type- and context-specific functions require a tight orchestration of signaling, and their dysregulation is often associated with pathology (Arkin *et al*, 2014). Experimental approaches that quantitatively capture the mechanisms of dynamic signaling networks are therefore highly valuable for establishing causal links to cellular phenotypes and the development of strategies for targeted interference.

Traditionally, the analysis of signal transduction mechanisms has focused on proteins with annotated functions in a given biological pathway. Pathway activation is probed with antibodies that determine the abundance of posttranslational modifications (PTMs) or interaction of selected proteins (protein–protein interactions, PPIs). Although valuable for testing pre-defined molecular states of selected proteins, the utility of this approach is limited by antibody availability, and prior knowledge of molecular and functional relationships. While employing antibodies would be applicable irrespective of the cell type, the discrimination of direct and indirect, as well as antibody-bound and bait-bound protein interactors, is often challenging because of limited antibody specificity (Marcon *et al*, 2015). Conversely, while epitope tagging of selected proteins provides an alternative that guarantees specific enrichment with stable background binders—a defined set of proteins adhering to the affinity matrix—not all cell types are amenable to efficient genetic manipulations. An optimal strategy would therefore combine efficient and antibody-independent enrichment with universal applicability for eukaryotic cell types.

Mass spectrometry (MS)-based proteomics allows the detection of PTM and PPIs without prior knowledge. In recent years, MS-based proteomics has advanced tremendously and transitioned from identifying only a few proteins to comprehensively quantifying cellular proteomes and identifying modified proteins and protein interactions on a large scale (Larance & Lamond, 2015; Aebersold &

1 Experimental Systems Immunology, Max Planck Institute of Biochemistry, Martinsried, Germany
2 Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Martinsried, Germany
3 Department of Molecular Machines and Signaling, Max Planck Institute of Biochemistry, Martinsried, Germany
4 Department of Genetics, School of Medicine, Stanford University, Stanford, CA, USA
5 Institute of Innate Immunity, Department of Systems Immunology and Proteomics, Medical Faculty, University of Bonn, Bonn, Germany
*Corresponding author (lead contact). Tel: +49 22828751210; E-mail: felix.meissner@uni-bonn.de

Mann, 2016). As such, it provides systems-wide views of cellular states with immense discovery potential, as indicated by large-scale efforts to map the entire interactomes in yeast (Gavin *et al*, 2002; Ho *et al*, 2002; Krogan *et al*, 2006), drosophila (Guruharsha *et al*, 2011), and human (Hein *et al*, 2013; Hein *et al*, 2015; Huttlin *et al*, 2015), kinase and phosphatase interactomes (Gingras *et al*, 2007; Couzens *et al*, 2013; Yao *et al*, 2017; Buljan *et al*, 2020) as well as global views of specific PTMs (Choudhary *et al*, 2009; Humphrey *et al*, 2015; Lescarbeau *et al*, 2016; Liu *et al*, 2018).

Although it is well appreciated that the interplay of PTMs and PPIs determines how biological responses are regulated, MS-based technologies are almost always used to investigate PTMs and PPIs separately, and rely on distinct biochemical and analytical strategies. Hence, the analysis of PPIs is bait-centric, and selected proteins are affinity-enriched together with their interacting partners (Paul *et al*, 2011). By contrast, PTM analysis generally focuses on a single modification type (e.g., phosphorylation), wherein modified peptides of all cellular proteins are affinity-enriched. Alternatively, in order to classify PTMs on specific proteins, affinity purification mass spectrometry (APMS) approaches with stringent washing and lysis conditions have been performed at the expense of PPI elucidation (Stutz *et al*, 2017; Pankow *et al*, 2019; Karayel *et al*, 2020). Consequently, these two molecular modes of protein regulations are experimentally disconnected, hampering the discovery of the relationships between PTMs and PPIs in cellular signaling pathways. Furthermore, easy methods to simultaneously monitor different PTM types in a single sample are missing. Conventional enrichment strategies for distinct PTMs vary widely, and hence, mapping of multiple PTMs usually requires several sequential or parallel biochemical steps. This requires large amounts of starting material and results in low-sample throughput, while comprehensiveness is still limited as the enrichment strategies are tailored toward known biochemical properties of selected PTMs. A method that would capture in an unbiased manner all detectable PTMs in protein complexes of interest is therefore needed so as to comprehensively pinpoint molecular signaling checkpoints in complex biological systems.

The functional evaluation of emerging PTMs and PPIs is a common bottleneck in systems-wide discovery approaches. While initial screens are often performed in an experimental system that closely resembles cellular physiology, experimental validation of hits among all discovered candidates frequently relies on loss- or gain-of-function experiments in cell lines to achieve the necessary throughput. However, desirable would be an experimental setup that facilitates both discovery and validation in primary cells.

To develop a method for the systematic dissection of cellular signaling checkpoints by simultaneous PTM and PPI mapping in one experiment, we devised a streamlined pipeline—Modifications, Interactions and Phenotypes by APMS (MIP-APMS). We evaluated and technically optimized all steps of MIP-APMS, comprising (i) the epitope tagging of proteins of interest and mammalian cell transduction, (ii) affinity purification conditions for optimal interaction network and PPI enrichment, (iii) followed by MS-based PTM and PPI quantification and identification, and (iv) ultimate biochemical and phenotypic validation of interactors and PTMs in primary human immune cells. Integration of multiple MIP-APMS experiments generates dynamic signal transduction networks and pinpoints time-resolved co-regulations of PTMs and PPIs in sequential signal transduction steps. We show the discovery potential of our pipeline by interrogating dynamically assembling protein communities in human monocyte immune signaling using Toll-like receptor (TLR) 2 activation and MAP kinase MAPK14 inhibition as paradigms. Our screen encompassing 19 protein complexes identified more than 50 previously undescribed PTMs, including phosphorylation, acetylation, methylation, ISGylation as well as other less well-described chemical modifications and elucidated an interaction network spanning more than 300 PPIs. We used the modular concept of MIP-APMS to test emerging data-driven hypotheses to validate PTMs and PPIs regulating immune signaling in reporter and primary cells. In this way, MIP-APMS enables the streamlined validation of crosstalk between different layers of protein regulation with broad applicability.

Results

Experimental and proteomics strategy for interrogating dynamic signal transduction networks

We devised MIP-APMS for the identification and perturbation of the functional checkpoints of cellular signaling pathways. MIP-APMS involves the following four stages with the indicated time frames (see Graphical Abstract, Figs 1 and EV1A):

1 Cloning of genes encoding epitope-tagged proteins and transduction of specialized cell types.
2 Streamlined quantification of various types of PTMs together with PPIs.
3 Implementation of an analytical strategy to pinpoint genetic or pharmacological signaling network perturbations.
4 Direct biochemical and functional evaluation of novel biological regulations in the same experimental system.

Universal cloning and transduction strategy

To enable interrogation of signaling cascades, we employed a cost-effective method for epitope tagging of proteins with a restriction enzyme-free approach, called restriction enzyme-free seamless ligation cloning extract (SLiCE) cloning (Zhang *et al*, 2012). A modified weak phosphoglycerate kinase (PGK) promoter controls the expression of the GOIs, which are flanked by attL sites. Thereby, our vector system is compatible with commercial DNA assembly cloning strategies such as the NEBuilder platform or Gateway, which had been used before (Lambert *et al*, 2014). As shown previously, employing lentiviral transduction for amphotrophic gene transfer extends the scope from readily transfectable cell lines, e.g., human embryonic kidney (HEK) cells, to non-dividing and terminally differentiated cells of primary origin (Huttlin *et al*, 2015; Samavarchi-Tehrani *et al*, 2018). In particular for application with primary immune cells, transduction is advantageous as other methods can activate innate immune signaling pathways and induce cell death (Fernandes-Alnemri *et al*, 2009; Hornung *et al*, 2009; Gaidt *et al*, 2017). As a relevant and challenging experimental model system, we chose human monocytes, because these cells are not easily transfectable and execute a broad spectrum of cellular programs by the dynamic intracellular propagation of molecular signals downstream of cell surface receptors. For method development and

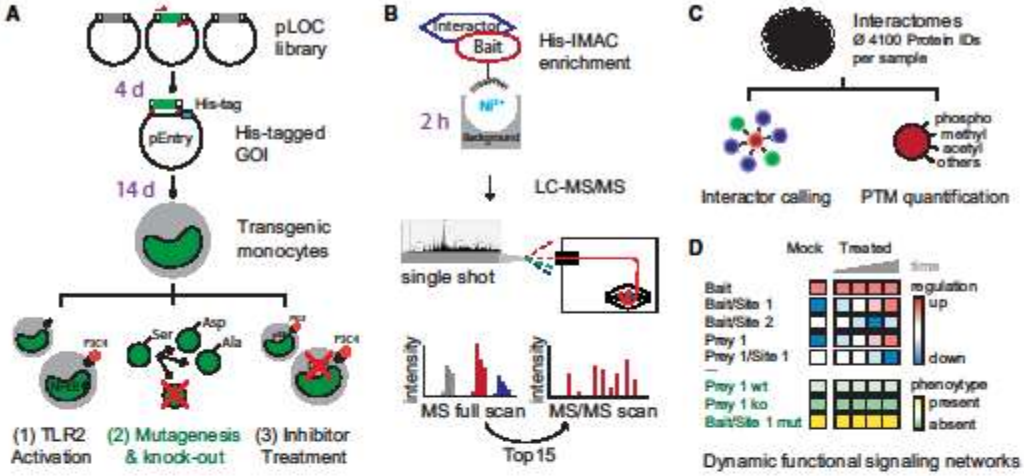


Figure 1. Experimental strategy for interrogating dynamic signal transduction networks with MIP-APMS.

A Generation of polyclonal transgenic cell lines by lentiviral transduction of genes encoding epitope-tagged wild-type or variant bait proteins. Analysis of PTMs and PPIs upon cellular activation (exemplified for the TLR1/2 activation by the agonist Pam3CSK4, P3C4), or pharmacological signal perturbation (exemplified by MAPK14 inhibitors). Time frames for the individual steps are indicated in violet.

B Single-step His-IMAC affinity enrichment and single-run liquid chromatography–tandem mass spectrometry (LC-MS/MS).

C Investigation of dynamic signal network topologies by simultaneous analysis of PPIs and multiple different PTMs. The numbers indicate analysis steps.

D Schematic representation of PTM and PPI dynamics as a function of cellular phenotypes. Wt, wild type; ko, knockout; mut, mutation.

Data Information: See also Figs EV1–EV4.

phenotypic screening, we employed the monocytic cell line U937 and validated our results with primary cells. We achieved 92 (± 5) % cellular transduction efficiency after antibiotic marker selection (Fig EV1B). We further demonstrated the universality of our approach with primary human macrophages differentiated from peripheral monocytes (Fig EV1C and D, Table EV1) and primary human T cells (Fig EV1E, Table EV1).

We carefully characterized the functional properties of generated cell lines: The average copy numbers of the endogenous protein counterparts to the tagged proteins were 3.1 million per cell, increasing only slightly to an average of 4.3 million copies upon transduction (Fig EV2A). Importantly, global protein expression levels remained stable within cells upon expression of epitope-tagged bait proteins (Fig EV2B and C). We specifically confirmed that transduced cells exhibit no background immune activation by assessing expression levels of proteins involved in immune- and infection-associated pathways (Fig EV2D and E) and retain their full activation potential by assessing NFkB activity with Luciferase reporter assays (Fig EV2F).

Simultaneous enrichment of PTMs and PPIs

Next, to study interdependency of PTMs and PPIs in signaling cascades, we evaluated biochemical enrichment strategies for epitope-tagged proteins with MS-based proteomics using high-performance liquid chromatography (HPLC) coupled to a linear quadrupole Orbitrap mass spectrometer (Q Exactive HF, Thermo) operated in a data-dependent acquisition mode (Fig 1B) (Scheltema *et al*, 2014). We systematically compared typical short epitope tags: Flag-tag (Hopp *et al*, 1988), Strep-tag (Schmidt & Skerra, 2007), and polyhistidine tag (Hochuli *et al*, 1988). To quantitatively compare epitope tag-based enrichments, identification and label-free quantification (LFQ) were performed in the MaxQuant environment (Cox & Mann, 2008). While > 1,000 proteins were shared between all three enrichments (Fig EV3B), His-IMAC enrichment identified more background binding proteins. Exemplified for MAPK14, our results show high overlap of known interactors for Strep-tag and His-tag IPs with on-bead digestion, whereas Flag-tag and Strep-tag with elution yielded lower numbers of significant interactors (Fig EV3A, Table EV1). Notably, the highest median bait protein sequence coverage (Fig EV3C), highest intensity of MAPK14 (Fig EV3D), and highest number of significantly interacting proteins were achieved with His-IMAC.

Accordingly, we incorporated His-IMAC in the MIP-APMS protocol and further optimized the protocol for high bait enrichment and high-sequence coverage by titrating imidazole concentration in lysis and wash buffers, respectively (Fig EV3E and F). Following method optimization, the respective bait proteins were among the highest enriched proteins after MIP-APMS (Fig EV3G, Table EV1). We achieved a median sequence coverage of 70% for bait proteins (Fig EV3H), opening up the possibility of directly identifying and quantifying PTMs, such as phosphorylation, acetylation, or methylation as well as other less well-studied covalent protein modifications on

any bait protein (Fig 1C). Differently modified peptides were not analyzed separately, as in typical proteomics workflows, but instead the selected enriched proteins represented all present and detectable proteoforms. This made it possible to simultaneously quantify the differently modified and unmodified versions of peptides. MIP-APMS enables the efficient and cost-effective and robust analysis of PTMs and PPIs in a single experiment.

Dynamic signaling network analysis

To study how signaling networks rearrange upon cellular activation, we integrated quantitative PTM and PPI information from multiple MIP-APMS experiments. This enabled quantitative analysis of sequential steps in signal transduction, since it allowed for dynamic PTM and PPI crosstalk to be resolved providing a basis to identify molecular switches in signal transduction networks. We observed enrichments and de-enrichment of prey proteins in protein complex of interest and also dynamically regulated PTMs on both bait and prey proteins (Fig 1D, regulation up/down).

Biochemical and functional evaluation of novel biological regulations in the same experimental system

To validate our findings in follow-up studies, we employed the same experimental system used for discovery. We investigated the alterations in dynamic signaling networks of proteins mutated on single amino acid sites discovered in our study. Furthermore, by transforming our model system into NFkB reporter cells, we were able to reveal functional effects on NFKB activation of novel PTMs and PPIs by CRISPR-Cas9-mediated gene knockout and site-specific gene mutations, respectively (Fig 1D, phenotype, Fig 1A). As described below in more detail, we were able to derive functional molecular checkpoints in monocyte signal transduction networks.

Signaling networks of kinases, signaling adapters, and caspases in monocytes

We tested our MIP-APMS approach by interrogating the molecular composition of protein communities in mammalian cells *in situ*. Specifically, we investigated innate immune signaling complexes, assembled various protein classes, such as kinases, caspases, and tumor necrosis factor (TNF) receptor-associated factors (TRAFs) in human monocytes.

We generated 19 transgenic monocytic U937 cell lines and analyzed them with MIP-APMS, as described above. This identified and quantified an average of 4,106 proteins per measurement, including non-specifically binding proteins as expected for non-stringent APMS conditions (Trinkle-Mulcahy *et al*, 2008; Rees *et al*, 2011). We observed high median intra-bait and inter-bait Pearson correlations (> 0.9) between biological replicates (Fig EV3I) and between different cell lines (Fig EV3J). This highlights the overall reproducibility of the devised workflow. To discriminate specifically interacting proteins from background binders common to all baits, we compared enrichments from single vs. all other cell lines with a standard statistical test (two-sided *t*-test) at a stringent false discovery rate (FDR) of 1% to correct for multiple hypothesis testing (Hein *et al*, 2015; Keilhauer *et al*, 2015; Hubel *et al*, 2019). This resulted in a small fraction of significantly interacting proteins (378 proteins in total, with a median of 16 interactors per bait) compared to a large proportion of background binders (Table EV1, Fig EV4A). Notably, distinct protein intensity differences and *P*-values clearly distinguish specific bait and prey from unspecific background proteins (Fig EV4B and C). MIP-APMS prioritizes bait-specific preys, as proteins enriched in multiple experiments—including interconnected interactors—show lower enrichment differences and *P*-values (Fig EV4D) by unbiased statistical interactor calling (see Materials and Methods). We compared our LFQ intensity and *t*-test-based strategy to the results of the SAINT algorithm (spectral count based) exemplary for MAPK14 and identified largely similar interactors (Fig EV4E).

The identified interactors included previously described as well as novel proteins (Fig 2A, Table EV2). Unsupervised hierarchical clustering of label-free quantification (LFQ) intensity profiles of the significant interactors grouped specific interactors of corresponding bait proteins together (Fig EV4F). To determine the topology of the detected protein interaction network, we assembled proteins according to shared interactors. This enabled the identification of signaling hubs through common connections of bait and prey proteins that clustered together in the network (Fig 2B). The analysis recapitulated many known interactions, including the TRAF2–BIRC2 CORUM complex (Ruepp *et al*, 2010) involving the binary interaction of TRAF2 and BIRC2, supplemented by such players as TRAF1, TBK1, TANK, and IKBKE (Wu *et al*, 2005). Some TRAF2 interactors, such as RIPK1, CASP8, and TNF (Hsu *et al*, 1996), were not detected in this experiment perhaps because they require distinct context-dependent cellular activation, e.g., through TNFR. These

Figure 2. Dissection of protein signaling networks in human monocytes using MIP-APMS.

A Percentage of previously described interactors (green) and novel interactors (blue), and the count of significant interactors (FDR < 0.01, enrichment > 2) per bait protein (median interactor count: 16).
B Protein–protein interaction network of clustered interaction data. Edges indicate interactions, with shared interactions connecting the individual MIP-APMS experiments. Red nodes correspond to bait proteins, green nodes to interactors reported in the literature, and blue nodes to novel interactors.
C Numbers of acetylations, methylations, and phosphorylations identified on bait proteins and interactors.
D Percentage of PTMs identified on bait proteins and interactors.
E Numbers of PTMs on bait proteins/interactors of individual pull-downs.
F Numbers of novel and described (Uniprot-annotated) acetylations, methylations, and phosphorylations.
G Unsupervised clustering (Pearson correlation) of the *z*-scored intensity profiles of all PPIs (357) and PTMs (37) upon TLR2 activation, partitioned in seven clusters.
H Dynamic profiles of co-regulated PTMs and PPIs with close network proximity, from the indicated clusters; median *z*-scored intensity of each time point (blue: median, gray: confidence interval = 0.95, method: loess); *n*, number of proteins in clusters 1–7. Selected proteins from each cluster are indicated, with the bait proteins in parentheses.

Data Information: See also Fig EV3, Table EV2 for PPIs, and Table EV3 for PTMs.

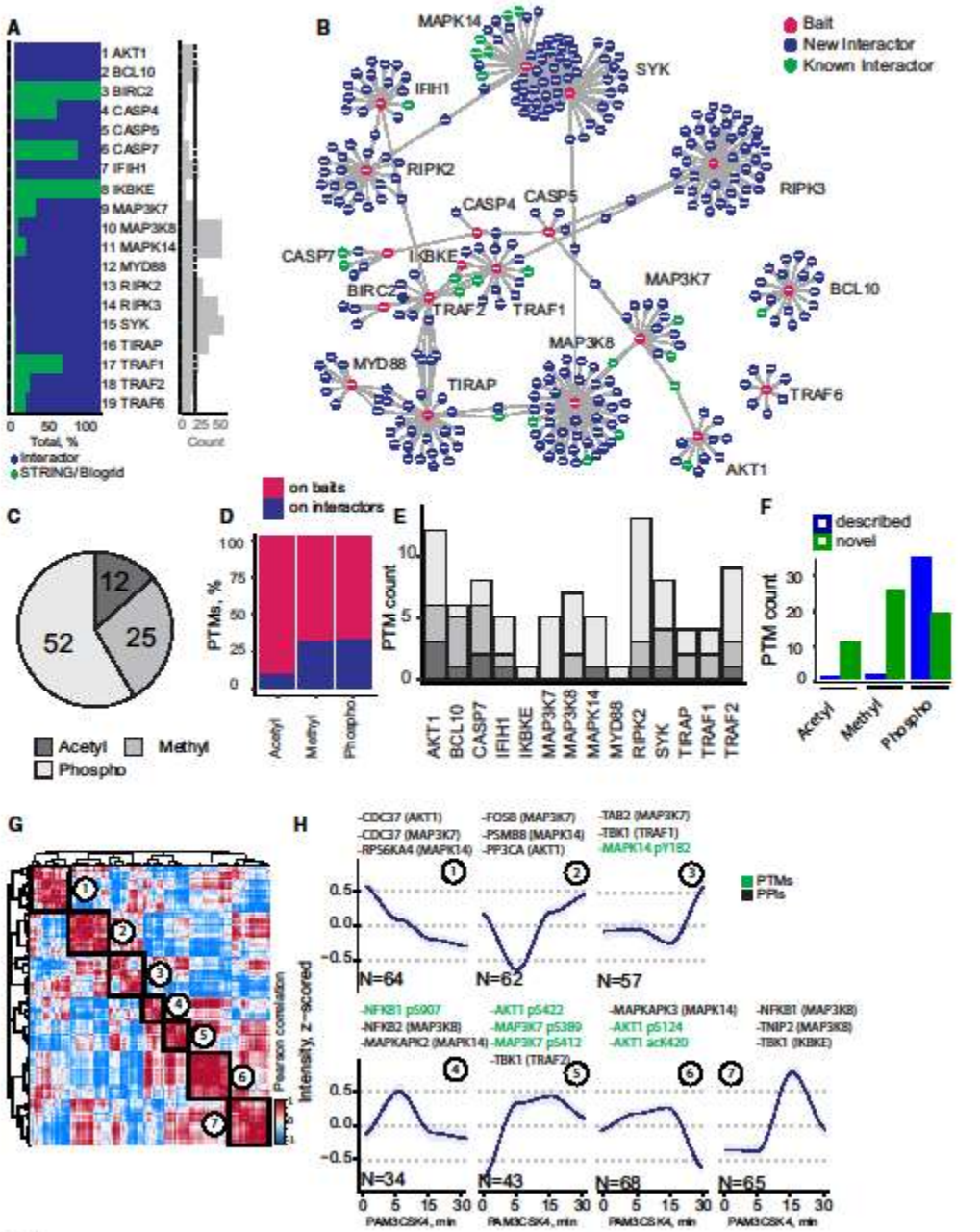

A
1 AKT1
2 BCL10
3 BIRC2
4 CASP4
5 CASP5
6 CASP7
7 IFIH1
8 IKBKE
9 MAP3K7
10 MAP3K8
11 MAPK14
12 MYD88
13 RIPK2
14 RIPK3
15 SYK
16 TIRAP
17 TRAF1
18 TRAF2
19 TRAF6
0 50 100
Total, %
0 25 50
Count
Interactor
STRING/Biogrid
B
Bait
New Interactor
Known Interactor
MAPK14
SYK
IFIH1
RIPK2
RIPK3
CASP4
CASP5
CASP7
IKBKE
BIRC2
TRAF2
TRAF1
MAP3K7
BCL10
MAP3K8
MYD88
TIRAP
TRAF6
AKT1
C
12
25
52
Acetyl
Methyl
Phospho
D
on baits
on interactors
PTMs, %
100
75
50
25
0
Acetyl
Methyl
Phospho
E
PTM count
10
5
0
AKT1
BCL10
CASP7
IFIH1
IKBKE
MAP3K7
MAP3K8
MAPK14
MYD88
RIPK2
SYK
TIRAP
TRAF1
TRAF2
F
described
novel
PTM count
30
20
10
0
Acetyl
Methyl
Phospho
G
Pearson correlation
H
-CDC37 (AKT1)
-CDC37 (MAP3K7)
-RPS6KA4 (MAPK14)
-FOSB (MAP3K7)
-PSMB8 (MAPK14)
-PP3CA (AKT1)
-TAB2 (MAP3K7)
-TBK1 (TRAF1)
-MAPK14 pY182
PTMs
PPIs
Intensity, z-scored
0.5
0.0
-0.5
N=64
N=62
N=57
-NFKB1 pS907
-NFKB2 (MAP3K8)
-MAPKAPK2 (MAPK14)
-AKT1 pS422
-MAP3K7 pS389
-MAP3K7 pS412
-TBK1 (TRAF2)
-MAPKAPK3 (MAPK14)
-AKT1 pS124
-AKT1 acK420
-NFKB1 (MAP3K8)
-TNIP2 (MAP3K8)
-TBK1 (IKBKE)
N=34
N=43
N=68
N=65
0 5 15 30
PAM3CSK4, min

Figure 2.

observations validated the utility of MIP-APMS for the interrogation of intracellular signaling networks.

To identify and quantify PTMs in the same experimental setup, we re-analyzed our data with phosphorylation, acetylation, and methylation as variable posttranslational modifications. Even though we did not enrich for PTMs, we identified and quantified PTMs spanning phosphorylations, acetylations, and methylations on baits as well as prey proteins (88 PTMs on 19 bait proteins). Phosphorylation was the most abundant PTM in the dataset (52 sites), followed by methylation (25 sites) and acetylation (11 sites) (Fig 2C). While the majority of PTMs were detected on bait proteins, some (31 PTMs on 10 proteins) were also detected on prey proteins (26% of all known PTMs; Fig 2D). A remarkable 74% of the studied bait proteins or their respective interactors were posttranslationally modified, with some proteins, e.g., AKT1 and RIPK2, harboring more than 10 PTMs (Fig 2E). Notably, MIP-APMS identified 52 previously undescribed PTMs, in particular methylation and acetylation sites (Fig 2F). Furthermore, an unbiased analysis of covalent peptide modifications using the dependent peptide algorithm in MaxQuant, the string-based search algorithm Taggraph—based on a de novo search in PEAKS—and MS Fragger (Devabhaktuni *et al*, 2019)—revealed a series of less well-described covalent modifications on MAPK14 (Fig EV4G). Twenty-six modifications were shared between search engines (2.3% of all modifications for dependent peptides, 1.5% PEAKS/Taggraph, and 0.9% MS Fragger, Fig EV4H). Out of these 26 modifications, six were reproducibly identified and quantified in all replicates (Fig EV4I). To distinguish biologically regulated from other—for example—sample preparation-introduced modifications, we quantified the identified modifications upon cell activation with specific searches in MaxQuant. Notably, only MAPK14 phosphorylation was differentially regulated between conditions. Moreover, acetylation, methylation, and phosphorylation detected on TRAF2, MAPK14, and MAP3K7 with specific searches were missed by open searches (dependent peptides of MaxQuant and PEAKS/Taggraph; Fig EV4J). This demonstrates that MIP-APMS can discover novel PTMs in signaling complexes; however, comparisons across search engines and confirmation with specific search strategies are advisable to increase confidence.

To capture the dynamics of cellular signal transduction, we next analyzed how the intracellular networks rearrange upon cellular activation via cell surface receptors. We stimulated cells via TLR2, as this pattern recognition receptor is prominently expressed in monocytes and induces a robust pro-inflammatory program that involves activation of the transcription nuclear factor (NF) κB pathway (Oliveira-Nascimento *et al*, 2012; Rieckmann *et al*, 2017). We analyzed the dynamic signaling networks downstream of TLR2 using time course experiments in biological quadruplicates. Upon stimulation with the lipopeptide Pam3CysK (PAM3CSK4), cellular signaling was activated (Fig EV4K), and stable vs. dynamic PTMs and PPIs could be distinguished. Because of the short time frame of kinetic investigations (within 30-min post-cellular activation), we did not normalize protein levels to expression-induced protein abundance changes. On average, we detected two statistically significant dynamic PPIs and one dynamic PTMs per bait (Fig EV5A; Table EV2 and EV3). Our data suggest that phosphorylation is the most

Figure 3. N-Terminal phosphorylation of TRAF2 and ISG15 is dynamic functional regulators downstream of TLR2.

A Volcano plot representing the interactome of TRAF2 (measured 15x in biological replicate) compared against all other pull-downs in the control group. The results of the *t*-tests are represented in volcano plots, which show the protein enrichment versus the significance of the enrichment. Numbers indicate enrichment ranks with the heatmap labels of (C) serving as the legend. Significant interactors of TRAF2 (two-tailed *t*-test, FDR < 0.01, enrichment > 4) are colored in blue (novel interactors) and green (known interactors).

B Interactors of TRAF2 (blue: novel interactors, green: known interactors) with interconnecting proteins between different baits colored in gray.

C Hierarchical clustering of significant interactors of TRAF2 upon activation with significant hits in at least one time point denoted with an asterisk. Cell activation was performed for 5, 15, and 30 min with the TLR2 ligand PAM3CSK4 (P3C4).

D Intensity profile of the TRAF2 interactor TANK upon activation, normalized to TRAF2 bait. LFQ intensity.

E Hierarchical clustering of the TRAF2 PTMs (acetylation, methylation, and phosphorylation) upon activation, with significant hits (*t*-test) in at least one time point denoted with an asterisk.

F Intensity profile of the phosphorylation of TRAF2 on Thr7 upon activation, normalized to TRAF2 bait intensity. Central band of the boxplot shows the median, boxes represent the IQR, 3 biological replicates were performed for UT, and 4 biological replicates were performed for additional time points. *P*-values were calculated by *t*-test. Asterisks indicate significant differences. *P-value < 0.05.

G Intensity profile of TRAF2 interactors ISG15 and TRAF1 in different TRAF2 phospho-variants, normalized to TRAF2 wild-type intensities. Central band of the boxplot shows the median, boxes represent the IQR, and 4 biological replicates were performed for every condition. *P*-values were calculated by *t*-test. Asterisks indicate significant differences. **P-value < 0.01, ***P-value < 0.001.

H Induction of NFκB determined based on luciferase luminescence in TLR2-activated U937 NFκB reporter cells transfected with genes encoding different TRAF2 phospho-variants. Bar represents the median, error bars represent the standard deviation, and 4 biological replicates were performed for additional time points. *P*-values were calculated by *t*-test. Asterisks indicate significant differences. ***P-value < 0.001.

I MS/MS Spectrum containing GlyGly modification K320 on TRAF2 after *GlyGly* enrichment on TRAF2 MIP-APMS.

J Differences and *P*-values of ISG15 intensity in TRAF2 K->R mutants compared against TRAF WT.

K Intensity profile of TRAF2 interactors ISG15 and TANK in TRAF2 K->R mutants, normalized to TRAF2 wild-type intensities. Central band of the boxplot shows the median, boxes represent the IQR, and four biological replicates were performed for every condition. *P*-values were calculated by *t*-test. Asterisks indicate significant differences. **P-value < 0.001.

L Induction of NFκB determined based on luciferase luminescence in TLR2-activated U937 NFκB reporter cells transfected with genes encoding different TRAF2 K→R mutants (each bar represents a mean from three independent measurements; error bars represent the standard deviation; ***P-value < 0.001).

M Intensity profile of TRAF2 interactors ISG15 and TANK in TRAF2-K388R and S11D mutants in human primary macrophages. Central band of the boxplot shows the median, boxes represent the IQR, and three biological replicates were performed for every condition. *P*-values were calculated by *t*-test. Asterisks indicate significant differences. **P-value < 0.01, *P-value < 0.05, ***P-value < 0.001.

Data Information: Experiments in (A–L) were performed in U937 cell lines. Gray boxes indicate missing values. IQR stands for interquartile range and represents the 25th to 75th percentile. See also Table EV1–EV4.

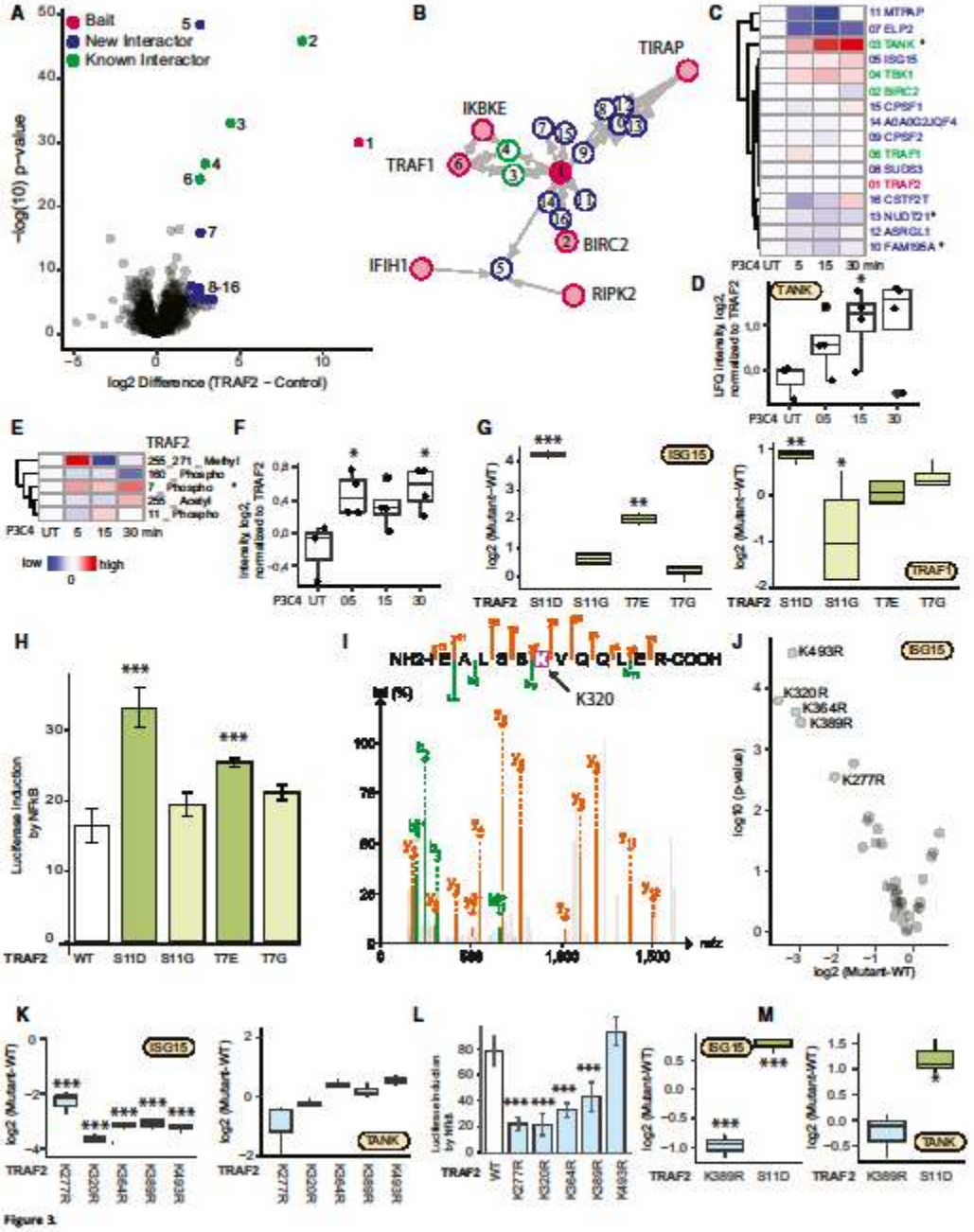

A
Bait
New Interactor
Known Interactor
−log(10) p−value
log2 Difference (TRAF2 − Control)
B
TIRAP
IKBKE
TRAF1
IFIH1
BIRC2
RIPK2
C
11 MTRAP
07 ELP2
03 TANK *
05 ISG15
04 TBK1
02 BIRC2
15 CPSF1
14 A0A0G2JQF4
09 CPSF2
06 TRAF1
08 SUDS3
01 TRAF2
16 CSTF2T
13 NUDT21*
12 ASRGL1
10 FAM98A *
P3C4 UT 5 15 30 min
D
TANK
LFQ intensity log2, normalized to TRAF2
P3C4 UT 05 15 30
E
TRAF2
255_271_Methyl
160_Phospho
7_Phospho *
255_Acetyl
11_Phospho
P3C4 UT 5 15 30 min
low high
F
Intensity log2, normalized to TRAF2
P3C4 UT 05 15 30
G
ISG15
log2 (Mutant−WT)
TRAF2 S11D S11G T7E T7G
TRAF1
H
Luciferase induction by NFκB
TRAF2 WT S11D S11G T7E T7G
I
NH2-F E A L S S K V Q Q L E R-COOH
K320
Int (%)
m/z
J
ISG15
K493R
K320R
K364R
K389R
K277R
log10 (p-value)
log2 (Mutant-WT)
K
ISG15
TANK
log2 (Mutant-WT)
TRAF2 K277R K320R K364R K389R K493R
L
Luciferase induction by NFκB
TRAF2 WT K277R K320R K364R K389R K493R
M
ISG15
TANK
log2 (Mutant-WT)
TRAF2 K389R S11D

Figure 3.

dynamic PTM in the tested setting, followed by methylation (Fig EV5B).

Next, to study PTM and PPI interdependency, we correlated all PTM and PPI intensities and clustered them unbiasedly over the time course of TLR2 activation (Fig 2G, Table EV3). We detect the dynamic co-regulation on both molecular layers (PPIs and PTMs), identifying correlating and anti-correlating PTMs and PPIs during signaling pathway activation. We identified seven clusters with distinct kinetics, some peaking early (Fig 2H, Cluster: 4,5) and others late (Fig 2H, Cluster: 7) upon pathway activation, as well as up- (Fig 2H, Cluster: 4,5,6,1) vs. down-regulated (Fig 2H, Cluster: 1,2,3) PTMs and PPIs. Interestingly, interactors identified in more than one MIP-APMS experiment (e.g., CDC37: Cluster 1) were in close network proximity. Our approach facilitated an unbiased discovery of time-resolved molecular connections between dynamic PTMs and PPIs, exemplified by the correlated interaction of MAP3K8 interactors (NFKB1, NFKB2) and NFκB1 phosphorylation (Cluster 4,7), or the anti-correlated phosphorylation of the C-terminal kinase domain of AKT1 and the interaction with CDC37 (Clusters 1, 5). This demonstrates that the sensitivity and robustness of MIP-APMS enable the simultaneously determination of cellular signaling network rearrangements by PPIs, PTMs, and their interplay. We conclude that MIP-APMS is sufficiently sensitive and robust to capture dynamic signaling networks in mammalian cells *in situ*. It generates highly reproducible data that may be used for the discovery of novel dynamic PTMs in signal transduction cascades, and simultaneous evaluation of multiple PTMs and PPIs in signaling networks.

Dynamic phosphorylations and ISGylations regulate TRAF2 downstream of TLR2

We next evaluated MIP-APMS for the discovery of novel molecular checkpoints in intracellular immune signaling. We focused on significantly regulated PPIs (FDR < 0.01) and PTMs (P-value < 0.05) identified for TRAF2 and MAP3K7, and examined their biochemical and phenotypic relevance through network perturbations mediated by gain- and loss-of-function mutations.

TRAF2 is a central adaptor protein in TNF signaling and regulates pro-inflammatory cytokine production through NFκB and JNK signaling pathways (Borghi *et al*, 2016). As described above, the MIP-APMS analysis confirmed previously reported TRAF2 interactors, such as TNF receptor-associated factor TRAF1, baculoviral IAP repeat-containing protein BIRC2 (cIAP2), TRAF family member-associated NFκB activator TANK, and serine/threonine-protein kinase TBK1. In addition, we identified ELP2 and ISG15 as novel components of the TRAF2 complex (Fig 3A and B) and TANK, a negative regulator of TRAF2 (Cheng & Baltimore, 1996), as dynamically recruited to the TRAF2 complex. By contrast, the majority of other TRAF2 interactors remained unchanged upon activation (Fig 3 C and D). While most other PTMs remained unchanged upon signal pathway activation, the analysis revealed dynamic N-terminal phosphorylations on Thr7 and Ser11 of TRAF2 (Fig 3E and F). Thus, the interactome and PTMs of TRAF2 are dynamically regulated upon NFκB activation via TLR2.

To test whether these dynamic N-terminal phosphorylations affected the composition and function of the TRAF2 protein complex, we used the MIP-APMS streamlined workflow to generate protein phospho-variants, in which specific Ser or Thr residues were changed to Gly, or to Asp/Glu to mimic phosphorylation. We probed the resulting signaling network rearrangements using MIP-APMS and found a specific enrichment for ubiquitin-like protein ISG15 and TRAF1 by the phospho-mimetic TRAF2 variants compared to wild-type TRAF2 (Fig 3G; Table EV4). These data suggest that N-terminal phosphorylation of TRAF2 at both Thr7 and Ser11 stabilizes a protein complex with ISG15 and TRAF1.

To further assess the functional relevance of the N-terminal TRAF2 phosphorylation on cellular regulation, we introduced the phospho-mimetic and phospho-dead TRAF2 variants into NFκB reporter monocyte cell lines. TLR2-induced NFκB activation was elevated with TRAF2 N-terminal phospho-mimetics, whereas the phospho-dead variants showed activation comparable to that of wild-type TRAF2, indicating that N-terminal phosphorylation boosted downstream signal transduction (Fig 3H).

ISG15 is a ubiquitin-like protein that covalently modifies target proteins on lysine residues in a process called ISGylation (Loeb & Haas, 1992; Zhang & Zhang, 2011). After tryptic digest, isgylated peptides harbor GlyGly modifications on lysines that can be readily detected by LC-MS/MS. As we did not directly detect GlyGly-modified peptides, we combined MIP-APMS with GlyGly enrichment and indeed identified two GlyGly modification sites on TRAF2 (Positions K27, K320; Fig 3I). To deduce the impact of ISGylation on the TRAF2 interaction network, we performed site-directed mutagenesis of TRAF2 lysines and subjected the K→R mutant cell lines to MIP-APMS. Out of the total 32 K→R mutants, 5 showed strong (more than 4×) and significant depletion of ISG15 in the TRAF2 complex (Fig 3J, Table EV4). Interestingly, the most regulated site—K320—was also identified by our initial GlyGly enrichment, suggesting an ISGylation of TRAF2. Reduced ISG15 levels in the interactomes of certain K→R mutants further support this observation. In contrast, TANK levels—a TRAF2 complex member—remained unaltered in the different TRAF2 mutants, pointing toward a specific partial perturbation of the TRAF2 protein community by K→R site-directed mutagenesis (Fig 3K). We excluded potential clonal or TRAF2 mutant expression effects on ISG15 levels by comparing ISG15 levels of transgenic monocyte interactomes to full proteomes (Fig EV5D). Unchanged ISG15 intensities upon stringent MIP-APMS conditions (6 M GdmCl) in a TRAF2 MIP-APMS experiment as well as no evident interaction of recombinant ISG15 and TRAF2 in a size exclusion-based binding assay further support the covalent ISGylation of TRAF2 (Fig EV5E). Functional analysis of the K→R mutants revealed reduction in NFkB activation for K277R, K320R, K364R, and K389R mutants, suggesting that ISGylation of TRAF2 may act as a positive regulator downstream of TLR2 (Fig 3L). To expand our findings to primary human macrophages, we selected the novel phospho-mimetic TRAF2 mutant S11D and lysine-mutant K389R. These experiments confirm ISG15 enrichment in the TRAF2-S11D complex and depletion in the TRAF2-K389R complex (Fig 3M).

ARHGEF18 and FOSB are functional regulators downstream of TLR2

To further explore the utility of MIP-APMS for discovery of new interactors, we evaluated functional interactions of MAP3K7. MAP3K7 (TAK1) is a central kinase of the MAPK signaling pathway, with crucial roles in the activation of TRAF6 downstream of TLRs

and other receptors (B-cell receptor, TNF receptor) (Landström, 2010) and known as a major regulator of NFκB signaling (Sato *et al*, 2005). The MIP-APMS analysis recapitulated the TNFα/NFκB signaling complex 7 (CORUM) consisting of TAB1, TAB2, TAB3, and CDC37 (Fig 4A and B). Upon TLR2 activation, TAB1 and SNX17 were depleted from the MAP3K7 complex (Fig 4C), while phosphorylation of MAP3K7 on Ser389 increased, significantly (Fig 4D). This revealed dynamic regulation of both PTMs and PPIs during pathway execution.

From the nine previously unknown interactors, we selected the guanine nucleotide exchange factor ARHGEF18 and the transcriptional regulators FOSB and FOXK2 for functional hypothesis testing. Because MAP3K7 is implicated in NFκB activation, we used CRISPR to knock out the respective genes in monocytic NFκB reporter cells and determined the pathway activity by luciferase induction that directly correlates with the activation of NFkB (Fig 4E). Upon deletion of genes encoding TLR2 and MYD88 (the receptor and proximal adaptor of PAM3CSK4, respectively (Li *et al*, 2010)), we observed an almost complete inhibition of NFκB activation. CRISPR knockout of *MAP3K7*, and the interactors *ARHGEF18* and *FOSB*, led to a partial reduction of NFκB activation, thereby linking this PPI to a functional downstream phenotype in the signaling cascade (all

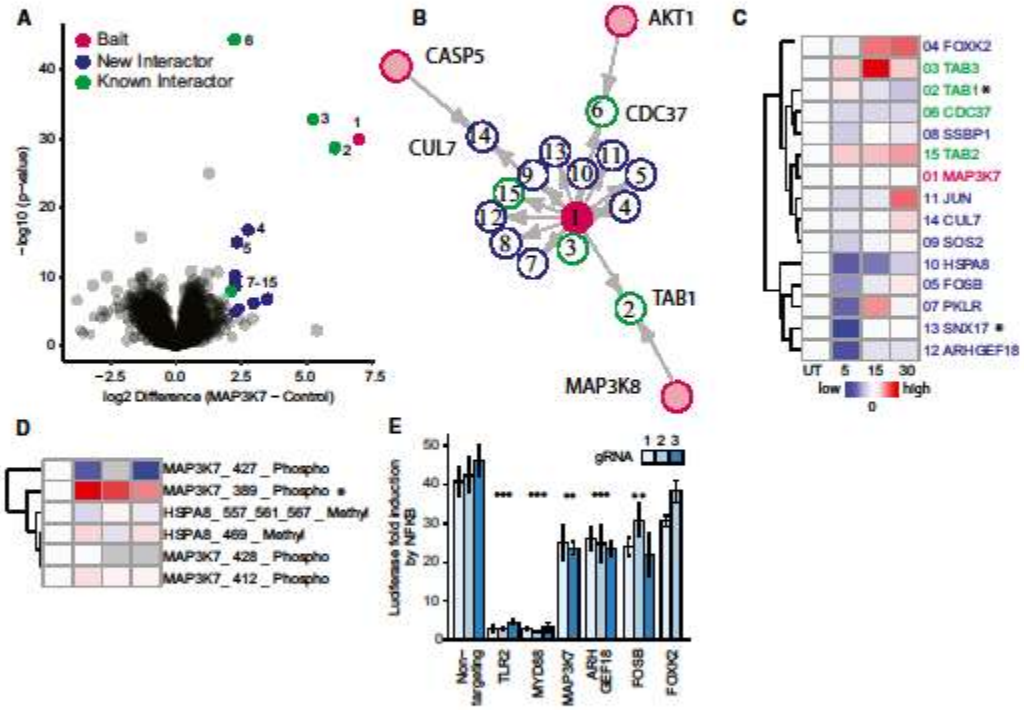


Figure 4. ARHGEF18 and FOSB are functional regulators downstream of TLR2.

A Volcano plot representing the interactome of MAP3K7 (measured 16× in biological replicates) compared against all other pull-downs in the control group. The results of the *t*-tests are represented in volcano plots, which show the protein enrichment versus the significance of the enrichment. Numbers indicate enrichment ranks with the heatmap labels of (C) serving as the legend. Significant interactors of MAP3K7 (two-tailed *t*-test, FDR < 0.01, enrichment > 4) are colored in blue (novel interactors) and green (known interactors).

B Interactors of MAP3K7 (blue: novel interactors, green: known interactors) with interconnecting proteins between different baits colored in gray.

C Heatmap of significant interactors of MAP3K7 upon activation, with significant hits in at least one time point (*t*-test, *P*-value < 0.05) denoted with an asterisk. Cell activation was performed for 5, 15, and 30 min with the TLR2 ligand PAM3CSK4 (P3C4).

D Heatmap of MAP3K7 PTMs (phosphorylation) upon activation, with significant hits (*t*-test, *P*-value < 0.05) in at least one time point denoted with an asterisk.

E Induction of NFκB determined based on luciferase luminescence in U937 NFκB reporter cells with CRISPR-Cas9 knockouts of the potential novel interactors of MAP3K7 upon TLR2 activation (each bar represents a mean of four independent measurements; error bars represent the standard deviation; *P*-values were calculated by *t*-test. Asterisks indicate significant differences. ***P-value < 0.001, **P-value < 0.01).

Data Information: Gray boxes indicate missing values. See also Appendix Figs S1–S34, Tables EV1 and EV2.

PTMs and PPIs of the characterized bait proteins are available in Appendix Figs S1–S14). We verified CRISPR-KO of ARHGEF18 and FOSB by Western blot analysis (Fig EV5F).

Hence, our MIP-APMS strategy can interrogate the functional relevance of individual molecular switches in a streamlined manner on the levels of PTMs as well as PPIs in signal transduction networks.

Dissecting drug mode of action for MAPK14 inhibitors with MIP-APMS

Small molecules are often used to interfere with specific cellular functions and are the mainstay of the drug industry. Definition of the target engagement of small molecules is a major challenge in drug discovery and novel proteomics approaches have been devised for this purpose (Schirle *et al*, 2012). We reasoned that MIP-APMS could enable the identification of signaling network rearrangements induced by small molecules, providing a unique proteomic perspective on the mode of drug action. We selected previously described pharmacological inhibitors of p38 kinase (MAPK14) (JX-401 (Friedmann *et al*, 2006), sorafenib (Edwards & Emens, 2010), and skepinone-L (Koeberle *et al*, 2011)) and analyzed their mode of perturbation of the cellular signaling network assemblies involving MAPK14 (Fig 5A–C).

The obtained data indicated that skepinone-L and sorafenib interfered with the physiological intracellular signaling network of MAPK14 to a greater extent than JX-401 (Fig 5D; Table EV5). Further, interestingly, sorafenib and skepinone-L perturbed the interactions within the core complexes differently. While MAPKAPK5, a downstream substrate of MAPK14 (New *et al*, 1998), was depleted in the MAPK14 protein complex upon treatment with both sorafenib and skepinone-L (Fig 5D), only sorafenib reduced the binding of RPS6KA4 (MSK1) and PTPN7 to MAPK14, and even more so upon cellular activation with TLR2 ligands (Fig 5E). Further, both sorafenib and skepinone-L induced hyper-phosphorylation of the MAPK14 phospho-loop on Tyr182, whereas an N-terminal phosphorylation site (Ser2) remained unaltered (Fig 5B and F). This indicated that PTMs and PPIs of MAPK14 are altered upon inhibitor treatment.

Further, MIP-APMS also allowed testing of drug off-target effects (Fig EV5G–I). MAP3K7 phosphorylation on Ser367, Ser412, and Ser445 was significantly altered, and both JUN and TAB2 were depleted from the MAPK14 complex upon treatment with sorafenib. This suggests that the MAP3K7 protein complex, reported to be an upstream activator of MAPK14 (Martín-Blanco, 2000), is in part targeted by MAPK14-specific inhibitors. Enrichment of ELP2 (JX-401, Skepinone-L) and TBK1 (JX-401) was observed in the TRAF2 signaling complex. PTMs on TRAF2 were not affected by the inhibitor treatment. Hence, MIP-APMS can be used to dynamically resolve the interactome and PTM changes upon small molecule treatment and provides information on molecular relationships in signal transduction networks that facilitate understanding of drug mode of action.

Discussion

Cellular processes are orchestrated by signal transduction pathways that depend on PTMs and PPIs. However, how PTMs and PPIs collaborate in structuring the dynamic signaling network topologies remains incompletely understood, in part because of the laborious experimental approaches involved in dissecting these interactions. Here, we describe MIP-APMS, a combined streamlined cell line generation and proteomics approach to interrogate functional signal transduction networks in intracellular signaling pathways. We quantified more than 370 PPIs and 80 PTMs across innate protein signaling cascades in human monocytes upon receptor activation or drug treatment. Among these are 50 previously undescribed PTMs, including those for which specific enrichment methods are less streamlined, such as ISGylation. Our approach revealed biochemical connections between PTMs and PPIs, as well as protein subnetworks that regulate cellular programs dependent on site-specific PTMs.

We employed MIP-APMS for streamlined and selected interference with protein subnetworks. Demonstrating this principle for the site-specific manipulation of protein phosphorylation as well as ISGylation on TRAF2 yielded differential interactomes of mutated proteins as well as altered cellular physiology. In this way, structural insights into interaction interfaces between protein complexes and crucial PTMs for stabilizing interacting proteins can be revealed. To our knowledge, this is the first description of protein ISGylation augmenting NFkB activity. We disturbed protein interaction networks of the kinase MAPK14 with small molecules to shed additional light on the drug mode of action of kinase inhibitors. Both skepinone-L and sorafenib changed the protein interaction network by a different mode of action, whereas both inhibitors lead to phosphorylation of the MAPK14 phospho-loop.

Modifications, interactions and phenotypes-APMS experiments with temporal resolution further allow the elucidation of co-regulations at different biochemical layers—adding to our understanding of molecular connections along the sequential steps of signal transduction. By further increasing temporal resolution, it may become possible to resolve the causalities between regulation on the PTM and PPIS levels in even greater detail.

For epitope tagging, we employed constructs from the pLOC library (GE Healthcare); however, other cDNA libraries or gene synthesis can readily be employed with polymerase chain reaction (PCR) to obtain DNA fragments with respective homologous overhangs. We employed the cost-effective, non-commercial SLICE cloning strategy; however, commercial solutions using NEBuilder or Gateway are possible with our vector system. An advantage of the small peptide tag chosen for the enrichment strategy in the current study is that it results in little steric interference with physiological protein–protein interactions.

According to our evaluation, Strep-tag and His-tag-based enrichments resulted in high bait sequence coverage; however, His-tag captured known interaction partners most comprehensive. By optimizing a non-stringent lysis procedure with low detergent and salt concentrations and also low temperature in the MIP-APMS protocol, we aimed to capture PTMs together with stable as well as transient interactions. According to our analysis, on average 12.3% of the PTMs and 5.5% of the PPIs are dynamic; however, as biochemical procedures impact recovery of interactors and different thresholds for significance calling are employed, comparability of PTM and PPI dynamics across studies remains challenging. Incorporating chemical cross-linking approaches (Holding, 2015; Liu & Heck, 2015)

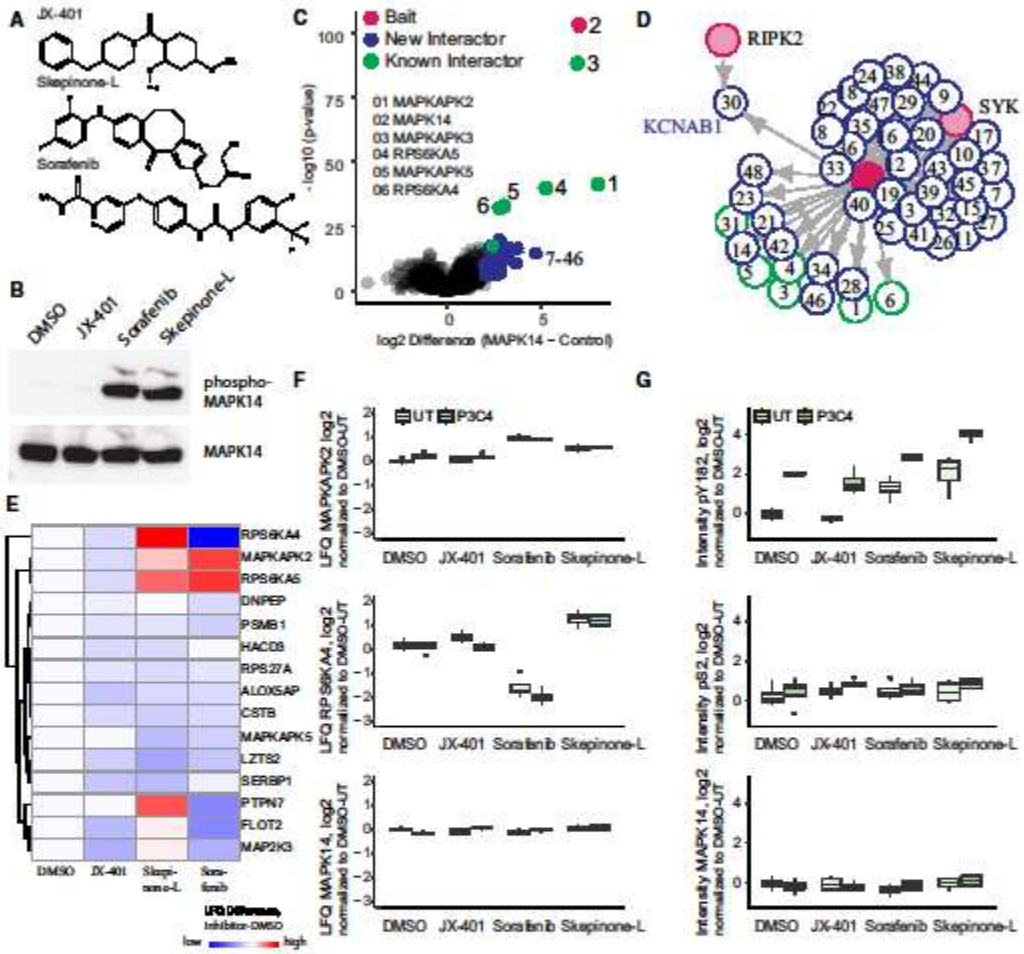


Figure 5. Dissecting drug mode of action for MAPK14 inhibitors with MIP-APMS.

A Chemical structures of MAPK14 inhibitors (JX-401, skepinone-L, and sorafenib).

B Phosphorylation of MAPK14 in U937 WT after treatment with the inhibitors, analyzed by Western blotting using an alpha-phospho-MAPK14 antibody. Total MAPK14, detected by alpha-MAPK14 antibody, was used a loading control.

C The interactome of MAPK14 compared against all other pull-downs in the control group. The results of the t-tests are represented in volcano plots, which show the protein enrichment versus the significance of the enrichment. Numbers indicate enrichment ranks with the heatmap labels of (C) serving as the legend. Only the top interactors of MAPK14 are numbered. The complete list can be found in Table EV1.

D Interactors of MAPK14 (blue: novel interactors, green: known interactors) with interconnecting proteins between different baits colored in gray.

E Heatmap of MAPK14 interactors significantly altered upon treatment with the different MAPK14 inhibitors, with significant hits in at least one treatment (t-test, *P*-value < 0.05) denoted with an asterisk. Treatments were normalized to DMSO control. The complete list can be found in Table EV5.

F LFQ intensity profiles of the MAPK14 interactors RPS6KA4 and MAPKAPK2 and MAPK14 after treatment with different MAPK14 inhibitors, normalized to MAPK14 bait intensity. Drug mode of action was analyzed in the presence (P3C4, 0.5 μg/ml, 30 min) or absence of P3C4 after inhibitor treatment. Central band of the boxplot shows the median, boxes represent the IQR, and 4 biological replicates were performed for every condition.

G Intensity profiles of MAPK14 phosphorylation on positions Ser2 and Tyr182 and MAPK14 protein intensity after treatment with different MAPK14 inhibitors, normalized to MAPK14 bait intensity. Drug mode of action was analyzed in the presence (P3C4, 0.5 μg/ml, 30 min) or absence of P3C4 after inhibitor treatment. Central band of the boxplot shows the median, boxes represent the IQR, and 4 biological replicates were performed for every condition.

Data information: Gray boxes indicate missing values. Bars represent median, error bars s.d. See also Fig EV5, Table EV5.

could further stabilize transient interactors. Combining MIP-APMS with structural information of the bait protein as well as its interactome can potentially reveal distinct interaction interfaces of protein complexes that are perturbed by site-specific covalent modification or drug action. It would be particularly interesting to integrate protein cross-linking with the PTM status, e.g., of wild-type vs. mutant variant proteins of interest, so that altered structural interaction interfaces can be resolved in addition to differential PTM-dependent PPIs.

We have already explored the strategy of disturbing protein interaction networks using small molecules to determine the effect of drugs on protein complexes. Whereas other proteomics approaches are aimed at identifying drug targets (Molina *et al*, 2013), MIP-APMS elucidates changes in protein communities involving the selected targets. It may thus serve as an additional drug discovery tool to resolve target protein network properties or off-target effects.

Despite its advantages, MIP-APMS currently has some limitations, which can be addressed by developing the method further in the future. These include the possibility that bait protein levels are different from those of endogenous proteins and incomplete protein sequence coverage. MIP-APMS is based on epitope-tagged bait proteins, which are introduced into target cells by lentiviral cellular transduction. Although this strategy enables rapid experiments and functional interrogation with high bait throughput, protein production levels may deviate from endogenous levels with ectopic expression instead of genome editing (Lackner *et al*, 2015). To address this and avoid excessive overproduction of the bait proteins (see Fig EV1B), we employed an engineered weak PGK promoter, as opposed to the commonly used strong cytomegalovirus (CMV) promoter (Qin *et al*, 2010). In general, we recommend total proteome measurements as described in the current study to evaluate whether normalization of changes caused by bait introduction is required.

Using MIP-APMS, we achieved 70 % sequence coverage for bait proteins using a single-enzyme protein digestion strategy combined with data-dependent acquisition. To further increase sequence coverage and map PTMs on bait proteins even more comprehensively, additional proteases, e.g., chymotrypsin or GluC, could be used. As MIP-APMS does not include a second enrichment step, the method preferentially quantifies abundant PTMs on bait and prey proteins. Ubiquitinylation, neddylation, and ISGylation are known as sub-stoichiometric PTMs and special biochemical enrichment or MS methods are commonly used for their detection (Kim *et al*, 2011; Wagner *et al*, 2011; Bustos *et al*, 2012; Hansen *et al*, 2021). We show that MIP-APMS combined with GlyGly enrichment facilitates the bait-centric identification of ubiquitin-like modification sites, exemplified for TRAF2. In the future, the total measuring time per sample at a comparable proteomics depth may be further reduced by using data-independent acquisition strategies and short LC gradients (Bruderer *et al*, 2017; Bache *et al*, 2018). Reproducibility, precision and accuracy of modified peptide quantification may be increased further by using isobaric labeling strategies as opposed to LFQ (Hogrebe *et al*, 2018; Virreira Winter *et al*, 2018).

In conclusion, MIP-APMS provides a versatile platform for parallel and time-resolved determination of PPIs and all PTMs of protein complexes in all transducible cells. It quantitatively resolves dynamic signaling network topologies and has broad applicability for the monitoring of virtually all coordinated intracellular programs. Owing to its conceptual design, emerging hypotheses on PTM and PPI involvement in selected signaling cascades are readily testable by protein mutation or loss of function impact on cellular phenotypes.

Materials and Methods

Reagents and Tools table

Reagent or Resource	Source	Identifier
Experimental Models: Cell Lines		
Human: HEK293T	ATCC	CRL-3216
Human: U937	ATCC	CRL-1593.2
Recombinant DNA		
CRISPR vector	Transomics	TEIA1002
Gene synthesis	This paper	Appendix Table S6
pLOC vectors	GE Healthcare	Appendix Table S5
pMD2.G	Addgene	#12259
psPAX	Addgene	#12260
Antibodies		
Anti-rabbit IgG, HRP-linked antibody	Cell Signaling	7074
GAPDH (14C10) Rabbit mAb	Cell Signaling	2118
Phospho-p38 MAPK (Thr180/Tyr182) antibody	Cell Signaling	9211

Reagents and Tools table (continued)

Reagent or Resource	Source	Identifier
ARHGEF18	Sigma	HPA042689
MAP3K7	R&D	MAB5307
FOSB	R&D	AF2214
Bacterial and Virus Strains		
Cignal Lenti NFκB Reporter (luc)	Qiagen	CLS-013L
XL1-Blue Competent Cells	Agilent Technologies	200249
Oligonucleotides		
PCR and cloning primers	This study	Appendix Table S1–S4
Chemicals, enzymes, and other reagents		
Blasticidin	Invivogen	ant-bl-1
cOmplete™, Mini Protease Inhibitor Cocktail	Sigma	4693132001
DMEM	Life Technologies	31966047
JX-401	Santa Cruz Biotechnology	CAS 349087-34-9
LysC	Wako-Chemicals	129-01541
Ni-IDA Agarose	Jena Bioscience	AC-310-25
PAM3CSK4	Invivogen	tlrl-pms
Passive Lysis 5X Buffer	Promega	E1941
Phosstop—20 TABLETS	Sigma	4906837001
Phusion® High-Fidelity DNA Polymerase	New England Biolabs	M0530S
Polybrene	Sigma	107689
Polyethylenimine, Linear, MW 25000, Transfection Grade (PEI 25K)	Polysciences	23966-1
Puromycin	Invivogen	ant-pr-5
RPMI-1640	Life Technologies	72400054
Skepinone-L	Merck	506374-5MG
Sorafenib	Santa Cruz Biotechnology	CAS 284461-73-0
SwaI	New England Biolabs	R0604L
T4 DNA ligase reaction buffer	New England Biolabs	B0202S
Trypsin	Sigma	T6567-1mg
Critical commercial assays		
Dual-Luciferase® Reporter Assay System	Promega	E1910
QUIKChange II XL Site-Directed Mutagenesis Kit	Agilent	#200521
Oligonucleotides		
Oligos for pLOC cloning, site-directed mutagenesis, and gRNA cloning	This paper	Appendix Table S6
Software and Algorithms		
MaxQuant	(Cox & Mann, 2008)	http://www.biochem.mpg.de/5111795/maxquant
Perseus	(Tyanova et al, 2016)	http://www.biochem.mpg.de/5111810/perseus
R	NA	https://www.r-project.org/
Ggplot2	(Wickham, 2016)	https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf
Igraph	NA	http://igraph.org/r/
CHOPCHOP	(Labun et al, 2016)	http://chopchop.cbu.uib.no/index.php

Methods and Protocols

Experimental design

All experiments were performed in replicate. No aspect of the study was blinded. Sample size was not predetermined, and no outliers were excluded from analyses.

Molecular biology

Entry vector design

We made use of restriction enzyme-free seamless ligation cloning extract (SLiCE) cloning using universal primer pairs to insert coding sequences of genes from the Precision LentiORF Collection (pLOC) library (GE Healthcare) into target vectors under the control of a modified weak phosphoglycerate kinase (PGK) promoter, introducing C-terminal epitope tags into the encoded proteins (Zhang *et al*, 2012). Our vector system is compatible with commercial DNA assembly cloning strategies such as the NEB Builder platform or Gateway due to AttI sites flanking the GOIs.

Entry vectors for SLiCE cloning were derived from the pLOC library (GE Healthcare). The vectors include a blasticidin resistance cassette for antibiotic-assisted cell-line selection and an IRES-GFP for FACS sorting. An efficient entry site for SLiCE cloning of the GOI was integrated by SLiCE cloning: The original pLOC vector was PCR-amplified using primers 1 and 2 containing overhangs with a SwaI restriction enzyme (New England Biolabs) cutting site for plasmid linearization, attL1/2 sites for homologous recombination-based SLiCE cloning, and a His-GSG-Flag-tag for GOI epitope tagging. SLiCE cloning was performed as previously described (Zhang *et al*, 2012). Briefly, 300 ng of the amplified pLOC vector, 1:10 (v/v) SLiCE extract (in-house), and 1:10 (v/v) T4 ligase buffer (New England Biolabs) was incubated for 1 h at 37°C. After incubation, the SLiCE mixture was used to transform XL1-blue bacteria (in-house) by heat shock. The transformants were selected on LB plates supplemented with 100 μg/ml ampicillin (LB-Amp plates) after overnight incubation at 37°C. Positive clones were identified by sequencing using primer 3.

The CMV promoter in the modified pLOC vector was exchanged for a weak PGK promoter by SLiCE cloning: the modified pLOC vector was PCR-amplified using primers 4 and 5 (see Appendix Table S1), the weak PGK promoter with homologous ends to the modified pLOC vector was *de novo* synthesized (see Appendix Table S6), and the two fragments were combined by SLiCE, as described before (Zhang *et al*, 2012). Briefly, 300 ng of the amplified pLOC vector, 100 ng of the synthesized weak PGK promoter fragment, 1:10 (v/v) SLiCE extract, and 1:10 (v/v) T4 ligase buffer (New England Biolabs) were incubated for 1 h at 37°C. After incubation, the SLiCE mixture was used to transform XL1-blue bacteria by heat shock. The transformants were selected on LB-Amp plates after overnight incubation at 37°C. Positive clones were identified by sequencing using primer 6 (see Appendix Table S1), as above. The obtained vector was used in subsequent cloning steps as an entry vector, called pLOC entry vector (pLOC-PGKweak-C-HisGSGFlag-BLASTICIDIN).

Cloning for epitope tagging

Open-reading frame (ORF) clones were obtained from the Precision LentiORF Collection. GOI (see Appendix Table S5) were PCR-amplified from the pLOC library (GE Healthcare) using the universal primers 7, 8, and 9. The CDS of MAP3K7 with attL1/attL2 overhangs was obtained by gene synthesis (Thermo Fisher Scientific). The pLOC entry vector was digested with the restriction enzyme SwaI (New England Biolabs) according to the manufacturer's instructions. Then, 300 ng of linearized pLOC entry vector, 100 ng of amplified GOI, 1:10 (v/v) SLiCE extract, and 1:10 (v/v) T4 ligase buffer (New England Biolabs) were incubated for 1 h at 37°C. After incubation, the SLiCE mixture was used to transform XL1-blue bacteria by heat shock. The transformants were selected on LB-Amp plates supplemented with 10% (v/v) glucose after overnight incubation at 37°C. Positive clones were identified by sequencing using primers 4 and 10 (see Appendix Table S1).

Site-directed mutagenesis of selected phosphosites

For the site-directed mutagenesis of the N-terminal TRAF2 phosphosites and TRAF2 K→R mutants, the QUIKChange II XL site-directed mutagenesis kit (Agilent) was employed. The site-directed mutagenesis was performed by PCR amplification of pLOC-TRAF2 using specific primers (see Appendix Table S2), according to the manufacturer's instructions.

Molecular biology and protein purification for TRAF2-ISG15 binding assays

ISG15, TRAF2, and influenza B virus NS1B were cloned into pCoofy vector as a N-terminal His-GST fusion. Plasmids were transformed into Rosetta (DE3) pLacI cells, grown in TB medium, and expression induced with 200 μM IPTG at OD_{600} 0.4–0.8. After induction, cultures were grown for 16 h at 18°C. Cells were re-suspended in lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 2 mM β-mercaptoethanol, protease inhibitor cocktail [Roche]) and lysed by sonication. Proteins were purified in tandem with His- and glutathione resin. Purified proteins were cleaved overnight with His-3C PreScission protease at 4°C. Following cleavage, the His-GST tag and His-3C protease were removed by a His pull-down. Proteins were either further purified by SEC (Superdex 75 10/300 GL, GE Life Sciences) or immediately buffer exchanged into storage buffer (50 mM Tris, 150 mM NaCl, 2 mM DTT). Proteins were concentrated and flash-frozen in liquid nitrogen.

Cell biology

Cell culture

U937 cells (CRL-1593.2) were purchased from the ATCC. The cells were cultured according to the manufacturer's instructions, in RPMI-1640 medium (Life Technologies) supplemented with 100 U/ml penicillin (GIBCO), 100 μg/ml streptomycin (GIBCO), and 10% (v/v) heat-inactivated fetal bovine serum (GIBCO; complete RPMI medium). The cells were incubated at 37°C under 5% CO_2.

HEK293T cells (CRL-3216) were purchased from ATCC. The cells were cultured according to the manufacturer's instructions, in DMEM (Life Technologies) supplemented with 100 U/ml penicillin (GIBCO), 100 μg/ml streptomycin (GIBCO), 1× Glutamax (GIBCO), and 10% heat-inactivated fetal bovine serum (complete DMEM medium). The cells were incubated at 37°C under 5% CO_2.

Primary human monocytes were obtained by culturing primary human monocytes enriched from buffy coats as described previously (Rieckmann *et al*, 2017). Primary human macrophages were differentiated in RPMI-1640 medium (Life Technologies) supplemented with 100 U/ml penicillin (GIBCO), 100 μg/ml streptomycin (GIBCO),

10 % (v/v) heat-inactivated fetal bovine serum (GIBCO), and 50 ng/ml M-Csf. The cells were incubated at 37°C under 5% CO_2.

NFκB reporter cell lines

U937 cell lines were transduced with Cignal Lenti NFκB-reporter constructs (Qiagen) according to the manufacturer's instructions. The transductants were selected in the presence of puromycin (5 μg/ml) for 14 days to establish stable cell lines.

Cell lines for epitope-tagged bait proteins

For lentivirus production, HEK293T cells (2×10^6, one six-well) were transfected with sequence-validated pLOC-GOI vectors using polyethyleneimine (Polysciences) as a transfection reagent. Helper plasmids pMD2.G, psPAX, and the pLOC vector harboring the GOI were combined in a ratio of 1:1.5:2. After 4-h incubation in complete RPMI medium at 37°C under 5% CO_2, the transfection mix was removed and fresh complete RPMI medium was added. Lentiviral supernatant was collected after 48-h incubation at 37°C under 5% CO_2, centrifuged (500 g, 5 min), filtered (0.45 μm), and supplemented with 8 μg/ml polybrene (Sigma). Then, the virus (complete supernatant of one six-well) was added to 0.2 Mio U937 cells or U937-NFkB Reporter Cell lines, incubated for 4 h at 37°C under 5% CO_2, following which fresh medium was added. Selection pressure with blasticidin (10 μg/ml; Invivogen) was introduced after 48 h. The cells were cultured for 2 weeks under the selective pressure and then directly used in MIP-APMS experiments.

Transduction of primary human macrophages was performed as previously described (Berger *et al*, 2011). In short, 10 Mio macrophages were transduced with a mix of VPX-Vlps and pLOC lentivirus (v/v, 50 %) in the presence of polybrene (8 μg/ml), incubated for 4 h at 37°C under 5% CO_2, following which fresh medium was added. Cells were harvested after 72h and then directly used in MIP-APMS experiments.

HEK293T (10 Mio) cells were transfected with pLOC-MAPK14-HisGSGFlag or pLOC-MAPK14-Strep using polyethyleneimine (Polysciences) as a transfection reagent. After 4-h incubation in complete DMEM medium at 37°C under 5% CO_2, the transfection mix was removed and fresh complete DMEM medium was added. Cells were harvested after 72h and then directly used in MIP-APMS, Flag-MS, and Strep-MS experiments.

TLR2 activation of U937 cells

Cells (5 Mio suspension) were seeded in deep-well 24-well plates, with one plate was used per cell line. TLR2 activation with PAM3CSK4 (0.5 μg/ml; Invivogen) was performed in a reverse time course and in quadruplicate, for 30, 15, 5, and 0 min at 37°C under 5% CO_2. The 0 min time point was not treated with PAM3CSK4. The cells were harvested by centrifugation and flash-frozen and stored at −80°C until MIP-APMS.

Drug mode of action on MAPK14

Cells (5 Mio) were seeded in deep-well 24-well plates. The cells were treated with MAPK14 inhibitors (sorafenib: 10 μM; skepinone-L: 80 nM; and JX-401: 10 μM) for 2 h at 37°C under 5% CO_2 in quadruplicate. Inhibitor-treated cells and controls either harvested directly or were activated with PAM3CSK3 (P3C4, 0.5 μg/ml; Invivogen) for 30 min at 37°C under 5% CO_2. Cells were harvested by centrifugation and frozen until MIP-APMS.

CRISPR/Cas9 Knockout

CRISPR knockout experiments were performed to identify potential novel interactors of MAP3K7 (see Appendix Table S3). For effective delivery of gRNA and Cas9, the transEDIT gRNA Plus Cas9 Expression vector with blasticidin was purchased from Transomics. For the experiment, gRNAs were designed using the web tool CHOP-CHOP (Labun *et al*, 2016) and cloned into the transEDIT vector according to the manufacturer's instructions (the primer list is provided in Appendix Table S3). Virus for each gRNA was produced as explained above (Cell Lines for Epitope-Tagged Bait Proteins). The U937-NFκB reporter cells were transduced and co-selected using puromycin (5 μg/ml) and blasticidin (10 μg/ml) at 37°C under 5% CO_2.

Luciferase reporter assay

U937-NFκB reporter cells (5×10^4) were seeded in quadruplicate on the day before the experiment. The cells were activated with PAM3CSK4 (0.5 μg/ml; Invivogen) for 6 h and harvested in passive lysis buffer (Promega). Luminescence of *Renilla* luciferase was determined in a dual-luciferase reporter assay (Promega), according to the manufacturer's instructions, using a microplate reader (Tecan).

Biochemistry

Western blots

One million U937 cells were stimulated, washed in PBS, and lysed in buffer (4% SDS, 40 mM HEPES [pH 7.4, 10 mM DTT] supplemented with protease inhibitors [Sigma-Aldrich, 4693159001]). Samples were centrifuged (16,000 g, 10 min), Li-LDS sample buffer was added to a final concentration of 1×, and the supernatant was incubated (5 min, 95°C). Proteins were separated on 12% Novex Tris-glycine gels (Thermo Fisher Scientific, XP00120BOX) and transferred onto PVDF membranes (Merck Millipore, IPVH00010) or Nitrocellulose membranes (Amersham, 10600002). Membranes were blocked in 5% BSA in PBST, and antibodies were diluted in 2% BSA in PBST. Antibodies used for immunoblotting were as follows (diluted 1:1,000): phospho-p38 MAPK (Thr180/Tyr182) antibody (Cell Signaling, 9211), GAPDH (14C10) rabbit mAb (Cell Signalling, 2118), p38 MAPK (R&D, AF8691), ARHGEF18 (Sigma, HPA042689), MAP3K7 (R&D, MAB5307), POSB (R&D, AF2214) and anti-rabbit IgG, HRP-linked antibody (Cell Signaling, 7074).

His-IMAC enrichment

Frozen pellets of 19 cell lines with bait proteins containing 9x His-tags in deep-well 24-well plates were defrosted (5 min, 37°C). The cells were re-suspended in 800 μl of lysis buffer (10 mM HEPES [pH 7.5; Gibco], 50 mM NaCl [Sigma], 20 mM imidazole [Sigma], 0.05% NP-40 [Thermo Fisher], 1 mM $MgCl_2$ [Sigma], 50 U/ml benzonase [in-house], protease inhibitors [Roche, 1 tablet per 50 ml], and phosphatase inhibitors [Roche, 1 tablet per 50 ml]), incubated for 15 min on ice, and cleared by centrifugation (500 g, 5 min, 4°C). Supernatants were transferred to deep-well 96-well plates already containing equilibrated Ni-IDA beads (JenaBioScience GmbH, 50 μl slurry per well). The plates were incubated at 4°C for 1 h, shaking. The beads were washed three times (10 mM HEPES [pH 7.5], 50 mM NaCl, and 20 mM imidazole), and the supernatant was removed completely before proceeding.

Flag-enrichment

Frozen pellets of HEK293T-MAPK14-HisGSGFlag (1xFlag) and control cell lines were defrosted (5 min, 37°C). The cells were re-suspended in 800 µl of lysis buffer (10 mM HEPES [pH 7.5; Gibco], 50 mM NaCl [Sigma], 0.05 % NP-40 [Thermo Fisher], 1 mM $MgCl_2$ [Sigma], 50 U/ml benzonase [in-house], protease inhibitors [Roche, 1 tablet per 50 ml], and phosphatase inhibitors [Roche, 1 tablet per 50 ml]), incubated for 15 min on ice, and cleared by centrifugation (500 *g*, 5 min, 4°C). Supernatants were transferred to deep-well 96-well plates already containing equilibrated anti-Flag M2 agarose gel (Sigma, 50 µl slurry per well). The plates were incubated at 4°C for 1 h, with shaking at over 1,500 rpm. The beads were washed three times (10 mM HEPES [pH 7.5], 50 mM NaCl) and the supernatant was removed completely before proceeding.

Strep-enrichment

Frozen pellets of HEK293T-MAPK14-Strep (1× Strep-tag II) and control cell lines were defrosted (5 min, 37°C). The cells were re-suspended in 800 µl of lysis buffer (10 mM HEPES [pH 7.5; Gibco], 50 mM NaCl [Sigma], 0.05 % NP-40 [Thermo Fisher], 1 mM $MgCl_2$ [Sigma], 50 U/ml benzonase [in-house], protease inhibitors [Roche, 1 tablet per 50 ml], and phosphatase inhibitors [Roche, 1 tablet per 50 ml]), incubated for 15 min on ice, and cleared by centrifugation (500 *g*, 5 min, 4°C). Supernatants were transferred to deep-well 96-well plates already containing equilibrated MagStrep "type3" beads (iba, 50 µl slurry per well). The plates were incubated at 4°C for 1 h, with shaking at over 1,500 rpm. The beads were washed three times (10 mM HEPES [pH 7.5], 50 mM NaCl), and the supernatant was removed completely before proceeding with sample preparation for on-bead digestion. For elution, beads were incubated with 50 µl 1× buffer BXT (IBA Lifesciences) and purified proteins were eluted at room temperature for 30 min with constant shaking at 1,100 rpm on a ThermoMixer C incubator as described previously (Gordon *et al*, 2020). Proportional amounts of bead and elution were analyzed.

Combination of MIP-APMS with GlyGly enrichment

We used 500 Mio TRAF2-U937 cells and performed His-IMAC enrichment as described above adjusted for higher input. The sample was digested as explained below under sample preparation. Peptide desalting was performed on SepPack C18 columns as per the manufacturer's instruction. After elution, peptides were lyophilized overnight. The lyophilized sample was reconstituted in 900 µl cold immunoaffinity purification buffer (IAP; 50 mM MOPS, pH 7.2, 10 mM Na_2HPO_4, 50 mM NaCl). For the enrichment of diGly remnant containing peptides, antibodies of the PTMScan® Ubiquitin Remnant Motif (K-ε-GG) Kit (Cell Signaling Technology [CST] were first cross-linked to beads). For this, one vial of antibody-coupled beads was washed three times with 1 ml cold cross-linking buffer (100 mM sodium tetraborate decahydrate, pH 9.0), followed by 30-min incubation in 1 ml cross-linking buffer (20 mM dimethylpimipimidate in cross-linking wash buffer) for 30 min at room temperature and gentle agitation. After two consecutive washes with 1 ml cold quenching buffer (200 mM ethanolamine, pH 8.0) and 2-h incubation in 1 ml cold quenching, crosslinked beads were washed three times with 1 ml cold IAP buffer and 1/24 was immediately used for immunoaffinity purification. For this, peptides were added to crosslinked antibody beads and incubated for 2 h at 4°C under gentle agitation. After incubation, beads were sequentially washed two times with cold IAP buffer and five times with cold ddH_2O in GF-StageTips. Thereafter, peptides were eluted twice with 50 µl 0.15% TFA into SDB-RPS StageTips. Eluted peptides were loaded onto stationary material and washed once with 200 µl 0.2% TFA and once with 200 µl 0.15% TFA/ 2% ACN. Peptides were eluted from SDB-RPS StageTips with 60 µl 1.25 % ammonium hydroxide (NH_4OH)/80 % ACN and dried using a SpeedVac centrifuge (Eppendorf, Concentrator plus). For mass spectrometry, dried peptides were re-suspended in 9 µl A* (2% ACN, 0.1% TFA).

Analytical size-exclusion chromatography binding assays

Binding assays were performed with ISG15 (1–157 aa) and TRAF2 variants (1–185 aa) on a Vanquish HPLC system (Thermo Fisher Scientific) using an AdvanceBio size-exclusion chromatography column (Agilent Technologies). As a positive control for ISG15 binding, the influenza B virus NS1B protein (1–103 aa) was used. Prior to analytical sizing, the column was pre-equilibrated with SEC buffer (50 mM Tris pH 7.5, 150 mM NaCl, 2 mM DTT). ISG15 (30 µM) was mixed with TRAF2 variants or NS1B (25 µM) prior to injection on the column. Fractions were mixed with SDS sample buffer and resolved on a 4–20% gradient SDS/PAGE. Gels were visualized by Coomassie staining.

Quantitative proteomics analysis

MIP-APMS sample preparation

After His-IMAC, the beads were re-suspended in 50 µl of 8 M urea and 40 mM HEPES (pH 8.0). LysC digestion (Wako, 0.5 µg/µl, 1 µl) was performed for 3 h at room temperature - 25°C (with shaking, 1,500 rpm). Afterward, the samples were diluted (1:6) with water and digested with trypsin (Sigma; 0.5 µg/µl, 1 µl) for 16 h (room temperature, with shaking, 1,500 rpm). The digests were centrifuged (5 min, 500 *g*), and the supernatants were transferred to new 96-well plates. Cysteines were reduced by the addition of dithiothreitol (1 mM, room temperature, 1,500 rpm, 30 min), before proceeding to cysteine alkylation with iodoacetamide (55 mM, room temperature, 30 min, dark). Excess iodoacetamide was quenched by adding thiourea (100 mM, room temperature, 10 min) prior to acidification for peptide desalting with trifluoroacetic acid (TFA; final concentration: 1% v/v). Peptides were loaded onto C18 StageTips (EmporeTM, IVA-Analysentechnik). They were then eluted with 80% acetonitrile, dried using a SpeedVac, and re-suspended in a solution of 2% acetonitrile, 0.1% TFA, and 0.5% acetic acid.

Whole-proteome MS sample preparation

Cells were lysed in SDC-lysis buffer and digested with LysC and trypsin, as described previously (Kulak *et al*, 2017). Peptides were desalted on stacked poly(styrene-divinylbenzene) reversed-phase sulfonate plugs and eluted with a mixture of 80 % acetonitrile, 19% ddH_2O, and 1% ammonia. MS measurements were performed in replicate ($n = 3$) using Q Exactive HF (Thermo Fisher Scientific).

LC-MS/MS

Peptides were separated using an EASY-nLC 1200 HPLC system (Thermo Fisher Scientific) coupled online to the Q Exactive HF and Q Exactive HF-X mass spectrometer via a nanoelectrospray source (Thermo Fisher Scientific), as described before (Scheltema *et al*,

2014; Kelstrup *et al*, 2018). Peptides were loaded in buffer A (0.5% formic acid) on in-house packed columns (75 μm inner diameter and 20 cm long; packed with 1.9-μm C18 particles from Dr. Maisch GmbH, Germany). Peptides were eluted using a nonlinear 95-min gradient of 5–60% buffer B (80% acetonitrile and 0.5% formic acid) at a flow rate of 300 nl/min and a column temperature of 55°C. The operational parameters were monitored in real-time by using the SprayQC software (in-house) (Scheltema & Mann, 2012). The Q Exactive HF and Q Exactive HF-X were operated in a data-dependent acquisition positive mode with a survey scan range of 300–1,650 *m/z* and a resolution of 60,000–120,000 at *m/z* 200. Up to 15 most abundant isotope patterns with a charge of > 1 were isolated using a 1.8 Thomson (Th) isolation window and subjected to high-energy collisional dissociation fragmentation at a normalized collision energy of 27. Fragmentation spectra were acquired with a resolution of 15,000 at *m/z* 200. Dynamic exclusion of sequenced peptides was set to 20 s to reduce repeated peptide sequencing. Thresholds for ion injection time and ion target values were set to 20 ms and 3E6 for the survey scans, and 55 ms and 1E5 for the MS/MS scans. Data were acquired using the Xcalibur software (Thermo Scientific).

Quantification and statistical analysis

Peptide identification and LC-MS/MS data analysis

MaxQuant software (version 1.5.3.16) was used to analyze MS raw files. MS/MS spectra were searched against the human Uniprot FASTA database (version July 2015, 91,645 entries) and a common contaminants database (247 entries) by the Andromeda search engine (Cox & Mann, 2008). Cysteine carbamidomethylation was set as a fixed modification, and N-terminal acetylation and methionine oxidation were set as variable modifications. To identify and quantify phosphorylation, acetylation, and methylation, variable modification search was consecutively performed. Enzyme specificity was set to trypsin, with a maximum of two missed cleavages and a minimum peptide length of seven amino acids. FDR of 1% was applied at the peptide and protein level. Peptide identification was performed with an allowed initial precursor mass deviation of up to 7 ppm and an allowed fragment mass deviation of 20 ppm. Nonlinear retention time alignment of all analyzed samples was performed using MaxQuant. Peptide identifications were matched across all samples within a time window of 1 min of the aligned retention times. Protein identification required at least one "razor peptide" in MaxQuant. A minimum ratio count of 1 was required for valid quantification events using the MaxQuant's LFQ algorithm (MaxLFQ). Data were filtered for the presence of common contaminants and peptides only identified by site modification, and hits to the reverse database (Cox & Mann, 2008) were excluded from further analysis.

Dependent peptide in MaxQuant analysis was performed to analyze unbiased PTMs on MAPK14 with standard parameters (FDR < 0.01, Mass bin size 0.0065 Da). For TagGraph analysis, sequence interpretations were first analyzed with the de novo search engine Peaks. Peaks analysis was performed with 10 ppm precursor mass tolerance and 0.01 Da fragment mass tolerance (Ma *et al*, 2003). TagGraph analysis was performed using human Uniprot FASTA database (version July 2015, 91,645 entries), with FDR cutoff of 0.1, and all other settings remained to unchanged as present in the software distribution.

Interactor calling

We integrate differences in intensity and abundance as described before (Keilhauer *et al*, 2015). We employ an AE-MS workflow with quantitative MS, which means that we use not only the information for protein identification but also for protein quantification for post-experiment interactor calling. To determine which proteins are substantially enriched (i.e., bait and prey proteins), AE-MS employs standard statistical testing (*t*-test) with a multiple hypothesis correction (FDR 0.01 for multiple hypothesis testing). In detail, each quantified protein had to be identified with more than one peptide and in more than 60% of replicates of at least one cell line to be considered valid. Protein LFQ intensities were log-transformed to the base of 2 and missing values imputed from a random normal distribution centered on the detection limit (width = 0.3, Down Shift 1.8). Samples were clustered by using Pearson correlation into different control groups in the Perseus environment leading to three separate groups (see Fig EV3K). To identify the interactors, a two-tailed Student's *t*-test (permutation-based FDR < 0.01 with 250 randomizations, enrichment > 2) with a minimum of 10 valid values in the first group was performed in the Perseus environment, using all other cell lines in the respective control group (Tyanova *et al*, 2016). Here, the baits were loaded as first group and second group mode was selected as "complement". Significant interactors were compared to the STRING and Biogrid databases (Szklarczyk *et al*, 2015; Chatr-Aryamontri *et al*, 2017) and overlaps were denoted in the Figs.

SAINT analysis via crapome

MAPK14 His IPs and controls (U937 transduced with His-Tag) were performed in triplicates and uploaded to the SAINT-based Crapome server (https://reprint-apms.org) (Mellacheruvu *et al*, 2013). As Experiment Type, we selected single-step epitope tag APMS and spectral counts as quantitation Type. As external controls, we selected PBMC (cell/tissue type), agarose (affinity support), and Q Exactive (Instrument type). The primary empirical fold change score (FC-A) was calculated by user controls using average for combining replicates (number of virtual controls = 10). The secondary fold change score was calculated by all controls (user + external controls) using geometric mean for combining replicates (number of virtual controls = 3). The probabilistic SAINT Score was calculated by user controls (combining replicates: average) and 10 virtual controls. Saint options were 2,000 n-burn, 4,000 n-iter, 0 LowMode, 1 MinFold, and 1 Normalize.

Analysis of dynamic PTMs and PPIs

Prior to the analysis of dynamic PPIs, LFQ intensities of significant interactors of each replicate were normalized to the LFQ intensities of the respective bait proteins to avoid loading artifacts.

$$\text{LFQ}-\text{intensity}(\text{prey}-\text{protein})_{\text{normalized}} = \frac{\text{LFQ intensity}(\text{prey}-\text{protein})}{\text{LFQ intensity}(\text{bait}-\text{protein})}.$$

A *two-tailed Student's t*-test (*P*-value < 0.05) was performed on the previously identified significant interactors comparing unactivated conditions versus activated conditions at different time points. Significant dynamic preys were reported with an asterisk in the heatmaps.

Conversely, intensities of modified peptides of each replicate were normalized to the intensity of the respective protein intensity to decrease the total coefficient of variation. PTMs that had valid values in at least 3 replicates of at least one time point were considered for the analysis. No imputation was performed.

$$\text{Intensity (modified peptide of protein } X)_{\text{normalized}} = \frac{\text{Intensity(modified peptide of protein } X)}{\text{Intensity (protein } X)}.$$

A *two-tailed Student's t*-test (P-value < 0.05) was performed on the previously identified significant interactors and modified peptides, respectively, comparing un-activated conditions versus activated conditions at different time points. Significant dynamic preys/PTMs were reported with an asterisk in the heatmaps.

Unsupervised clustering

Intensities of dynamically regulated PPIs (357) and PTMs (178) upon TLR 2 activation were filtered for at least 70% valid values and normalized per time point (PPIs to bait protein intensity and PTMs to protein intensity of the modified protein as explained above). The median of each time point was calculated and then Z-scored. Pearson correlation was calculated between each of the PPIs and PTMs, and results were visualized by hierarchical clustering. The data were clustered and median z-scored intensities (confidence interval: 0.95) were plotted against the time course of TLR2 activation (method = loess, $y \sim x$). N shows the number of PPIs/PTMs corresponding to each cluster.

Analysis of whole-proteome data

Full proteomes were measured in triplicates as described under peptide identification and LC-MS/MS data analysis. Data were filtered for the presence of common contaminants and peptides only identified by site modification, and hits to the reverse database (Cox & Mann, 2008) were excluded from further analysis. As a requirement, each quantified protein had to be identified with more than one peptide and in more than 60% of replicates of at least one cell line to be considered valid. Protein LFQ intensities were log-transformed to the base of 2 and missing values imputed from a random normal distribution centered on the detection limit (width = 0.3, Down Shift:1.8). To identify differentially expressed proteins between wildtype and transduced cell lines, a two-tailed Student's *t*-test (permutation-based FDR < 0.05 with 250 randomizations, enrichment > 2) with a minimum of two valid values in the first group was performed in the Perseus environment, using all other cell lines in the respective control group (Tyanova *et al*, 2016). Copy numbers were calculated with the Perseus Plugin Proteomic Ruler, which normalizes protein intensity to the molecular mass of each protein (Wiśniewski *et al*, 2014).

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD010996. The datasets produced in this study are available in the following database: https://www.ebi.ac.uk/pride/

Project accession: PXD010996 (http://www.ebi.ac.uk/pride/archive/projects/PXD010996).

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Felix Meissner (felix.meissner@uni-bonn.de).

Expanded View for this article is available online.

Acknowledgements

We thank I. Paron, C. Deimi, J. Mueller, and K. Mayr for MS assistance; G. Sowa, J. Baum, V. Boeck S. Dewitz, and A. Ullrich for technical assistance; M. Oroshi for computer and database support; and M. Tanzer, J. Swietlik, J. Rieckmann, M. Steger, J. Bader, and C. Luber for helpful discussions. This work was funded by the Max Planck Society for the Advancement of Science and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID 165054386 (SFB 914), Project-ID 360372040 (SFB 1335), and Project-ID 408885537 (TRR274). Open Access funding enabled and organized by Projekt DEAL.

Author contributions

AF, SE, FMH, KP, and KS performed experiments. AF developed and implemented bioinformatics methods. AF and FM conceived the data analysis and interpreted the data. AS and DH assisted in data analysis. FM and MM conceived the study. FM supervised the experiments. AF and FM wrote the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Aebersold R, Mann M (2016) Mass-spectrometric exploration of proteome structure and function. *Nature* 537: 347

Arkin MR, Tang Y, Wells JA (2014) Small-molecule inhibitors of protein-protein interactions: progressing towards the reality. *Chem Biol* 21: 1102 – 1114

Bache N, Geyer PE, Bekker-Jensen DB, Hoerning O, Falkenby L, Treit PV, Doll S, Paron I, Müller JB, Meier F *et al* (2018) A novel LC system embeds analytes in pre-formed gradients for rapid, ultra-robust proteomics. *Mol Cell Proteomics* 17: 2284 – 2296

Berger G, Durand S, Goujon C, Nguyen X-N, Cordeil S, Darlix J-L, Cimarelli A (2011) A simple, versatile and efficient method to genetically modify human monocyte-derived dendritic cells with HIV-1–derived lentiviral vectors. *Nat Protoc* 6: 806 – 816

Borghi A, Verstrepen L, Beyaert R (2016) TRAF2 multitasking in TNF receptor-induced signaling to NF-κB, MAP kinases and cell death. *Biochem Pharmacol* 116: 1 – 10

Bruderer R, Bernhardt OM, Gandhi T, Xuan Y, Sondermann J, Schmidt M, Gomez-Varela D, Reiter L (2017) Optimization of experimental parameters in data-independent mass spectrometry significantly increases depth and reproducibility of results. *Mol Cell Proteomics* 16: 2296 – 2309

Buljan M, Ciuffa R, van Drogen A, Vichalkovski A, Mehnert M, Rosenberger G, Lee S, Varjosalo M, Pernas LE, Spegg V *et al* (2020) Kinase interaction network expands functional and disease roles of human kinases. *Mol Cell* 79: 504 – 520.e9

Bustos D, Bakalarski CE, Yang Y, Peng J, Kirkpatrick DS (2012) Characterizing ubiquitination sites by peptide-based immunoaffinity enrichment. *Mol Cell Proteomics* 11: 1529 – 1540

Chatr-Aryamontri A, Oughtred R, Boucher L, Rust J, Chang C, Kolas NK, O'Donnell L, Oster S, Theesfeld C, Sellam A *et al* (2017) The BioGRID interaction database: 2017 update. *Nucleic Acids Res* 45(D1): D369–D379

Cheng G, Baltimore D (1996) TANK, a co-inducer with TRAF2 of TNF- and CD 40L-mediated NF-kappaB activation. *Genes Dev* 10: 963–973

Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, Olsen JV, Mann M (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325: 834

Couzens A, Knight J, Kean M, Teo G, Weiss A, Dunham W, Lin Z, Bagshaw R, Sicheri F, Pawson T *et al* (2013) Protein interaction network of the mammalian Hippo pathway reveals mechanisms of kinase-phosphatase interactions. *Sci Signal* 6: rs15

Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotech* 26: 1367–1372

Devabhaktuni A, Lin S, Zhang L, Swaminathan K, Gonzalez CG, Olsson N, Pearlman SM, Rawson K, Elias JE (2019) TagGraph reveals vast protein modification landscapes from large tandem mass spectrometry datasets. *Nat Biotechnol* 37: 469–479

Edwards JP, Emens LA (2010) The multikinase inhibitor sorafenib reverses the suppression of IL-12 and enhancement of IL-10 by PGE(2) in murine macrophages. *Int Immunopharmacol* 10: 1220–1228

Fernandes-Alnemri T, Yu J-W, Datta P, Wu J, Alnemri ES (2009) AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. *Nature* 458: 509–513

Friedmann Y, Shriki A, Bennett ER, Golos S, Diskin R, Marbach I, Bengal E, Engelberg D (2006) JX401, A p38α inhibitor containing a 4-benzylpiperidine motif, identified via a novel screening system in yeast. *Mol Pharmacol* 70: 1395

Gaidt MM, Ebert TS, Chauhan D, Ramshorn K, Pinci F, Zuber S, O'Duill F, Schmid-Burgk JL, Hoss F, Buhmann R *et al* (2017) The DNA inflammasome in human myeloid cells is initiated by a STING-cell death program upstream of NLRP3. *Cell* 171: 1110–1124.e1118

Gavin A-C, Bösche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon A-M, Cruciat C-M *et al* (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415: 141

Gingras A-C, Gstaiger M, Raught B, Aebersold R (2007) Analysis of protein complexes using mass spectrometry. *Nat Rev Mol Cell Biol* 8: 645–654

Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, O'Meara MJ, Rezelj VV, Guo JZ, Swaney DL *et al* (2020) A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* 583: 459–468

Guruharsha KG, Rual J-F, Zhai B, Mintseris J, Vaidya P, Vaidya N, Beekman C, Wong C, Rhee DY, Cenaj O *et al* (2011) A protein complex network of *Drosophila melanogaster*. *Cell* 147: 690–703

Hansen FM, Tanzer MC, Brüning F, Bludau I, Stafford C, Schulman BA, Robles MS, Karayel O, Mann M (2021) Data-independent acquisition method for ubiquitinome analysis reveals regulation of circadian biology. *Nat Commun* 12: 254

Hein MY, Sharma K, Cox J, Mann M (2013) Proteomic analysis of cellular systems A2. In *Handbook of Systems Biology*, Marian Walhout AJ, Vidal M, Dekker J (eds), San Diego, CA: Academic Press: 3–25.

Hein MY, Hubner NC, Poser I, Cox J, Nagaraj N, Toyoda Y, Gak IA, Weisswange I, Mansfeld J, Buchholz F *et al* (2015) A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* 163: 712–723

Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams S-L, Millar A, Taylor P, Bennett K, Boutilier K *et al* (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 415: 180

Hochuli E, Bannwarth W, Döbeli H, Gentz R, Stüber D (1988) Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Bio/Technology* 6: 1321–1325

Hogrebe A, von Stechow L, Bekker-Jensen DB, Weinert BT, Kelstrup CD, Olsen JV (2018) Benchmarking common quantification strategies for large-scale phosphoproteomics. *Nat Commun* 9: 1045

Holding AN (2015) XL-MS: Protein cross-linking coupled with mass spectrometry. *Methods* 89: 54–63

Hopp TP, Prickett KS, Price VL, Libby RT, March CJ, Pat Cerretti D, Urdal DL, Conlon PJ (1988) A short polypeptide marker sequence useful for recombinant protein identification and purification. *Bio/Technology* 6: 1204–1210

Hornung V, Ablasser A, Charrel-Dennis M, Bauernfeind F, Horvath G, Caffrey DR, Latz E, Fitzgerald KA (2009) AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* 458: 514–518

Hsu H, Huang J, Shu H-B, Baichwal V, Goeddel DV (1996) TNF-dependent recruitment of the protein kinase RIP to the TNF receptor-1 signaling complex. *Immunity* 4: 387–396

Hubel P, Urban C, Bergant V, Schneider WM, Knauer B, Stukalov A, Scaturro P, Mann A, Brunotte L, Hoffmann HH *et al* (2019) A protein-interaction network of interferon-stimulated genes extends the innate immune system landscape. *Nat Immunol* 20: 493–502

Humphrey SJ, Azimifar SB, Mann M (2015) High-throughput phosphoproteomics reveals *in vivo* insulin signaling dynamics. *Nat Biotechnol* 33: 990

Huttlin EL, Ting L, Bruckner RJ, Gebreab F, Gygi MP, Szpyt J, Tam S, Zarraga G, Colby G, Baltier K *et al* (2015) The BioPlex network: a systematic exploration of the human interactome. *Cell* 162: 425–440

Karayel O, Tonelli F, Virreira Winter S, Geyer PE, Fan Y, Sammler EM, Alessi D, Steger M, Mann M (2020) Accurate MS-based Rab10 phosphorylation stoichiometry determination as readout for LRRK2 activity in Parkinson's disease. *Mol Cell Proteomics* 19: 1546–1560

Keilhauer EC, Hein MY, Mann M (2015) Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). *Mol Cell Proteomics* 14: 120–135

Kelstrup CD, Bekker-Jensen DB, Arrey TN, Hogrebe A, Harder A, Olsen JV (2018) Performance evaluation of the Q Exactive HF-X for shotgun proteomics. *J Proteome Res* 17: 727–738

Kim W, Bennett EJ, Huttlin EL, Guo A, Li J, Possemato A, Sowa ME, Rad R, Rush J, Comb MJ *et al* (2011) Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol Cell* 44: 325–340

Koeberle SC, Romir J, Fischer S, Koeberle A, Schattel V, Albrecht W, Grütter C, Werz O, Rauh D, Stehle T *et al* (2011) Skepinone-L is a selective p38 mitogen-activated protein kinase inhibitor. *Nat Chem Biol* 8: 141

Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP *et al* (2006) Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* 440: 637

Kulak NA, Geyer PE, Mann M (2017) Loss-less nano-fractionator for high sensitivity, high coverage proteomics. *Mol Cell Proteomics* 16: 694–705

Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E (2016) CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res* 44(W1): W272–W276

Lackner DH, Carré A, Guzzardo PM, Banning C, Mangena R, Henley T, Oberndorfer S, Gapp BV, Nijman SMB, Brummelkamp TR *et al* (2015) A generic strategy for CRISPR-Cas9-mediated gene tagging. *Nat Commun* 6: 10237

Lambert J-P, Tucholska M, Pawson T, Gingras A-C (2014) Incorporating DNA shearing in standard affinity purification allows simultaneous identification of both soluble and chromatin-bound interaction partners. *J Proteomics* 100: 55–59

Landström M (2010) The TAK1–TRAF6 signalling pathway. *Int J Biochem Cell Biol* 42: 585–589

Larance M, Lamond AI (2015) Multidimensional proteomics for cell biology. *Nat Rev Mol Cell Biol* 16: 269

Lescarbeau RS, Lei L, Bakken KK, Sims PA, Sarkaria JN, Canoll P, White FM (2016) Quantitative phosphoproteomics reveals Wee1 kinase as a therapeutic target in a model of proneural glioblastoma. *Mol Cancer Ther* 15: 1332–1343

Li X, Jiang S, Tapping RI (2010) Toll-like receptor signaling in cell proliferation and survival. *Cytokine* 49: 1–9

Liu F, Heck AJR (2015) Interrogating the architecture of protein assemblies and protein interaction networks by cross-linking mass spectrometry. *Curr Opin Struct Biol* 35: 100–108

Liu JJ, Sharma K, Zangrandi L, Chen C, Humphrey SJ, Chiu Y-T, Spetea M, Liu-Chen L-Y, Schwartzer C, Mann M (2018) *In vivo* brain GPCR signaling elucidated by phosphoproteomics. *Science* 360: eaao4927

Loeb KR, Haas AL (1992) The interferon-inducible 15-kDa ubiquitin homolog conjugates to intracellular proteins. *J Biol Chem* 267: 7806–7813

Ma B, Zhang K, Hendrie C, Liang C, Li M, Doherty-Kirby A, Lajoie G (2003) PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun Mass Spectrom* 17: 2337–2342

Marcon E, Jain H, Bhattacharya A, Guo H, Phanse S, Pu S, Byram G, Collins BC, Dowdell E, Fenner M *et al* (2015) Assessment of a method to characterize antibody selectivity and specificity for use in immunoprecipitation. *Nat Methods* 12: 725

Martín-Blanco E (2000) p38 MAPK signalling cascades: ancient roles and new functions. *BioEssays* 22: 637–645

Mellacheruvu D, Wright Z, Couzens AL, Lambert J-P, St-Denis NA, Li T, Miteva YV, Hauri S, Sardiu ME, Low TY *et al* (2013) The CRAPome: a contaminant repository for affinity purification–mass spectrometry data. *Nat Methods* 10: 730

Molina DM, Jafari R, Ignatushchenko M, Seki T, Larsson EA, Dan C, Sreekumar L, Cao Y, Nordlund P (2013) Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* 341: 84

New L, Jiang Y, Zhao M, Liu K, Zhu W, Flood LJ, Kato Y, Parry GC, Han J (1998) PRAK, a novel protein kinase regulated by the p38 MAP kinase. *EMBO J* 17: 3372–3384

Oliveira-Nascimento L, Massari P, Wetzler LM (2012) The role of TLR2 in infection and immunity. *Front Immunol* 3: 79–79

Pan CQ, Sudol M, Sheetz M, Low BC (2012) Modularity and functional plasticity of scaffold proteins as p(l)acemakers in cell signaling. *Cell Signal* 24: 2143–2165

Pankow S, Bamberger C, Yates JR (2019) A posttranslational modification code for CFTR maturation is altered in cystic fibrosis. *Sci Signal* 12: eaan7984

Paul FE, Hosp F, Selbach M (2011) Analyzing protein–protein interactions by quantitative mass spectrometry. *Methods* 54: 387–395

Qin JY, Zhang L, Clift KL, Hulur I, Xiang AP, Ren B-Z, Lahn BT (2010) Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS One* 5: e10611

Rees JS, Lowe N, Armean IM, Roote J, Johnson G, Drummond E, Spriggs H, Ryder E, Russell S, St Johnston D *et al* (2011) *In vivo* analysis of proteomes and interactomes using Parallel Affinity Capture (iPAC) coupled to mass spectrometry. *Mol Cell Proteomics* 10: M110.002386

Rieckmann JC, Geiger R, Hornburg D, Wolf T, Kveler K, Jarrossay D, Sallusto F, Shen-Orr SS, Lanzavecchia A, Mann M *et al* (2017) Social network architecture of human immune cells unveiled by quantitative proteomics. *Nat Immunol* 18: 583

Ruepp A, Waegele B, Lechner M, Brauner B, Dunger-Kaltenbach I, Fobo G, Frishman G, Montrone C, Mewes HW (2010) CORUM: the comprehensive resource of mammalian protein complexes—2009. *Nucleic Acids Res* 38 (suppl_1): D497–D501

Samavarchi-Tehrani P, Abdouni H, Samson R, Gingras A-C (2018) A versatile lentiviral delivery toolkit for proximity-dependent biotinylation in diverse cell types *. *Mol Cell Proteomics* 17: 2256–2269

Sato S, Sanjo H, Takeda K, Ninomiya-Tsuji J, Yamamoto M, Kawai T, Matsumoto K, Takeuchi O, Akira S (2005) Essential function for the kinase TAK1 in innate and adaptive immune responses. *Nat Immunol* 6: 1087–1095

Scheltema RA, Mann M (2012) SprayQc: a real-time LC–MS/MS quality monitoring system to maximize uptime using off the shelf components. *J Proteome Res* 11: 3458–3466

Scheltema RA, Hauschild J-P, Lange O, Hornburg D, Denisov E, Damoc E, Kuehn A, Makarov A, Mann M (2014) The Q exactive HF, a benchtop mass spectrometer with a pre-filter, high performance quadrupole and an ultra-high field orbitrap analyzer. *Mol Cell Proteomics* 13: 3698–3708

Schirle M, Bantscheff M, Kuster B (2012) Mass spectrometry-based proteomics in preclinical drug discovery. *Chem Biol* 19: 72–84

Schmidt TGM, Skerra A (2007) The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc* 2: 1528–1535

Seet BT, Dikic I, Zhou M-M, Pawson T (2006) Reading protein modifications with interaction domains. *Nat Rev Mol Cell Biol* 7: 473

Stutz A, Kolbe C-C, Stahl R, Horvath GL, Franklin BS, van Ray O, Brinkschulte R, Geyer M, Meissner F, Latz E (2017) NLRP3 inflammasome assembly is regulated by phosphorylation of the pyrin domain. *J Exp Med* 214: 1725–1736

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP *et al* (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 43(Database issue): D447–D452

Trinkle-Mulcahy L, Boulon S, Lam YW, Urcia R, Boisvert FM, Vandermoere F, Morrice NA, Swift S, Rothbauer U, Leonhardt H *et al* (2008) Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J Cell Biol* 183: 223–239

Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Meth* 13: 731–740

Virreira Winter S, Meier F, Wichmann C, Cox J, Mann M, Meissner F (2018) EASI-tag enables accurate multiplexed and interference-free MS2-based proteome quantification. *Nat Methods* 15: 527–530

Wagner SA, Beli P, Weinert BT, Nielsen ML, Cox J, Mann M, Choudhary C (2011) A proteome-wide, quantitative survey of *in vivo* ubiquitylation sites reveals widespread regulatory roles. *Mol Cell Proteomics* 10: M111.013284

Wickham H (2016) *ggplot2: elegant graphics for data analysis*. New York, NY: Springer-Verlag. ISBN 978-3-319-24277-4. https://ggplot2.tidyverse.org

Wiśniewski JR, Hein MY, Cox J, Mann M (2014) A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards. *Mol Cell Proteomics* 13: 3497

Wu CJ, Conze DB, Li X, Ying SX, Hanover JA, Ashwell JD (2005) TNF-α induced c-IAP1/TRAF2 complex translocation to a Ubc6-containing compartment and TRAF2 ubiquitination. *EMBO J* 24: 1886

Yao Z, Darowski K, St-Denis N, Wong V, Offensperger F, Villedieu A, Amin S, Malty R, Aoki H, Guo H *et al* (2017) A global analysis of the receptor tyrosine kinase-protein phosphatase interactome. *Mol Cell* 65: 347–360

Zhang D, Zhang D-E (2011) Interferon-stimulated gene 15 and the protein ISGylation system. *J Interferon Cytokine Res* 31: 119–130

Zhang Y, Werling U, Edelmann W (2012) SLICE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res* 40: e55

4.2 Project 2: Revisiting signal peptide cleavage by neo N-terminal peptide meta-analysis

Signal peptides play a key role in protein targeting and protein translocation. They are short N-terminal peptide sequences targeting proteins to secretion or specific organelles. Signal peptide cleavage is mediated by enzymes calls signal peptide peptidases, which leads to the post-translational modification of the N-terminus of proteins harboring signal peptides.

Typically, signal peptide cleavage is computationally predicted using algorithms such as SignalP, as experimental approaches are frequently time consuming and focused on a single protein. Enzymatic digest, i.e. by signal peptide peptidases, is expected to result in non-tryptic peptides that can be identified by semi-specific digest search algorithms. I present a proteomics meta-analysis workflow for investigating signal peptide cleavage in previously reported datasets by analyzing the global results of N-terminal semi-specific search. I combined the spectral information from over 30 distinct MS-based proteomics datasets covering over 13000 rawfiles in order to identify signal peptide cleavage sites. In total, I found 1315 SignalP-predicted signal peptide cleavage sites in human and murine tissues, which more than doubles the amount of experimental evidence now available for these two species. Notably, I identify 412 additional cleavage sites near the expected cleavage site (-3/+7 aa). A significant challenge is differentiating genuine SPP cleavage sites from random sample preparation or non-SPP enzymatically produced neo-N-terminal peptides. Here, I show non-linear occurrence of neo-N-terminal peptides around the predicted cleavage site of proteins containing signal peptides. My investigation of functional annotations for proteins classified into several cleavage categories depending on mode and number of cleavages show that signal peptides from cytokines, growth factors, developmental proteins, and differentiation proteins are cleaved as predicted. Additionally, transmembrane proteins have either a single predicted cleavage site or an alternate cleavage site in the vicinity of the predicted cleavage site. I observed sequence motifs aligned to the predicted cleavage site varied between proteins with conventional and alternative signal peptide cleavage. Most notably, I show alternative cleavage for a

subgroup of proteins that have a glutamine in the expected cleavage site at position +1 and a proline in the predicted cleavage site at position -1.

I present experimental MS evidences to retrain the SP prediction model, which is based on over 700 proteins with novel experimentally identified SP sites that match the anticipated cleavage site. Alternative cleavages – which are more prevalent in transmembrane proteins – imply additional processing at unique consensus sequences and can be used to extend the SP prediction model. I anticipate increased prediction capabilities for signal peptide cleavage from the algorithms, particularly for transmembrane proteins. To make our signal peptide cleavage meta-analysis data easily accessible, I devised a searchable database. A similar meta-analysis could be conducted in the future for more species in order to obtain stronger and more extensive experimental evidence for signal peptide cleavage in other organisms.

This work is currently unpublished and a manuscript is in preparation. I have conceived the project, designed the data-analysis and interpreted the data together with Felix Meissner. Furthermore, I have developed all algorithms used to study signal peptide cleavage and prepared a first draft of the manuscript, as well as the script for the interactive shiny app. We aim to collaborate with S. Brunak and J. Armenteros to integrate the results about alternative signal peptide cleavage into SignalP before publication.

Proteomics meta-analysis identifies distinct N-terminal processing requirements for signal peptides of functionally divergent secreted and transmembrane proteins

Annika Frauenstein & Felix Meissner

4.2.1 Abstract

Proteins with cell surface or extracellular functions contain N-terminal signal-peptides that are processed post-translationally by signal peptide peptidases (SPPases). As experimental evidences for cleavage positions are incomplete, computational algorithms are commonly used to predict the neo N-termini of processed proteins. Here, we exploit the unexpected observation that proteolytically generated neo-N-terminal peptides can be identified systematically in shotgun proteomics datasets through tailored mass spectral search algorithms. We reanalyzed more than 13,000 human and mouse mass spectrometry (MS)-based proteomics raw files and systematically compared MS-detected cleavage sites to predictions of SignalP. Our analysis reveals SPP cleavage of more than 1300 proteins, 700 of which without prior experimental validation. Intriguingly, in addition to predicted canonical SPP cleavages, we identify alternative cleavages within +6/-4 amino acids of the canonical predicted site on more than 600 proteins. Proteins with canonical and alternative cleavages are functionally distinct and can be attributed diverse signal peptide consensus sequences. We describe canonical signal peptide cleavage for cytokines and proteins involved in development, whereas alternative cleavage is observed for transmembrane proteins. Our proteomic meta-analysis experimentally verifies and further characterizes signal peptide cleavage for the first time, globally, and opens novel opportunities to further explore protein N-terminal functions.

4.2.2 Introduction

Proteases hydrolyze peptide bonds in an irreversible post-translational process, which frequently alters the protein's structure and function [232]. Approximately 3% of human protein coding genes are predicted to encode proteases, which are categorized by their catalytic mechanism [233, 234]. Signal peptide peptidases (SPPs) and signal peptide peptidase-like proteases (SPPLs) belong to the aspartyl proteases [249-251] and are intramembrane-cleaving proteases with a conserved Y/FD and GxGD motif in the catalytic center [247, 248]. The physiological importance of SPP/ SPPLs is highlighted by their evolutionary conserved presence in eukaryotes such as fungi, protozoa, plants, and animals [247-249]. In mammals, five members of the SPP/SPPL family (SPP, SPPL2a, SPPL2b, SPPL2c, and SPPL3) with different subcellular localizations are classified [252]. Typical SPP substrates include signal sequence-derived signal peptides [257], but other SPP/ SPPL substrates have recently been identified, including type II

transmembrane proteins without predicted signal sequences, such as TNF-a [250] and CD74 [258, 259]. Signal peptides are located at the N-terminus of nascent proteins (e.g. membrane proteins, secretory proteins) and are cleaved co-translationally in a multi-step process [260]. Signal peptides facilitate the ribosome nascent chain (RNC) targeting to the ER the translocation initiation across the ER membrane [311]. While signal peptides harbor diverse primary sequences, they are composed by a strict tripartite structure: (1) a cationic N-terminal region (N-domain), (2) a hydrophobic core region (H-domain) containing at least six non-hydrophilic amino acids, and (3) a polar C-terminal region containing the signal peptidase cleavage site [262]. Primary sequence variation of signal peptides can affect the efficiency of protein targeting, translocation, and signal sequence cleavage and thus impact protein expression and function [267, 268, 312]. Hereditary diseases with decreased protein expression (e.g., Classic Ehlers-Danlos Syndrome [263], Crigler-Najjar Disease [264], Hereditary Angioedema [265]) are associated with mutations in the signal peptides, specifically an injection of a polar amino acid into the h-region [266].

Bioinformatic methods have been used extensively to predict signal peptides and the subsequent protein subcellular localization [313]. This includes both (1) distinguishing signal peptide-containing proteins and non-secretory proteins and (2) predicting the apparent signal peptide cleavage location to determine the exact N-terminus of the protein. Prediction algorithms have evolved in tandem with methodological advances in bioinformatics, beginning with basic statistics and weight matrices and progressing to the use of artificial neural network algorithms, hidden Markov models, support vector machines, and deep and recurrent ANNs. To date, the majority of signal peptide cleavage sites have been predicted rather than experimentally confirmed. Experimental verification of signal peptide cleavage sites usually involves small scale single protein centered studies in combination with Edman degradation or mass spectrometry to elucidate the N-terminal primary structure of individual proteins [314-316].

Tailored experimental approaches based on the negative or positive biochemical enrichment of neo-N-terminal peptides [317, 318] have identified up to 100 signal peptide cleavage sites in single datasets [319, 320], however comprehensive proteomics investigations have so far not been conducted. As signal peptide cleavage is efficient and generates up to 10% of all neo-N-terminal peptides in cells, we hypothesized that signal peptide cleavage sites should be readily identifiable from standard shotgun proteomics datasets without prior enrichment of neo-N-terminal peptides.

Here, we devised a proteomic meta-analysis strategy, to identify signal peptide cleavage sites from standard label-free proteomics shotgun data deposited in the public domain. We re-analyzed more than 47 datasets, comprising more than 13,000 raw files of human and mouse origin and systematically

compared the outcome to SignalP prediction. We report spectral evidence for signal peptide cleavage sites in over 1300 proteins, as predicted by SignalP. This is the first experimental verification of the predicted canonical cleavage site for over 700 of these proteins. Proteins with canonical cleavages were enriched for cytokines, developmental proteins and proteins involved in differentiation. Interestingly, in addition to the canonical signal peptide cleavage sites, we identify novel cleavages on more than 600 proteins within +6/-4 amino acids of the canonical site, which we term 'alternative' cleavages. Alternative cleavages predominantly appeared on membrane, transmembrane, glycoproteins, isomerases and hydrolases, whereas alternative cleavages occurred on proteins containing immunoglobulin regions. Proteins with canonical signal peptide cleavage sites are more likely to contain a Sec/SPI signal peptide (SignalP 5.0) than proteins with alternative cleavages. 800 of the proteins harbor more than one cleavage site in the analyzed cleavage window. Systematic evaluation of cleavage site probability and measured Neo-N-terminal peptide intensity, revealed distinct consensus sequences for predicted versus alternative cleavages. In summary, our meta-analysis enables previously unprecedented experimental verification of signal peptide cleavage from proteomic datasets and reveals intriguing insights into signal peptide cleavage.

4.2.3 Results

Currently, less than a quarter of predicted human signal peptide cleavage sites are verified experimentally according to Uniprot, due to the absence of systematic strategies to detect them (see Fig.1A, Supplementary table 1). As efficient co-translational signal peptide cleavage [321] by SPPs generates proteolytic protein fragments, we hypothesized, that these might be detectable in proteomics datasets without prior enrichment (Fig. 1B). As a result of SPP cleavage and conventional protease-based (e.g., Trypsin) digestion used in shotgun proteomics, unique peptide fragments are generated, that can be identified by tailored database searches. Instead of a standard and specific search for tryptic proteolysis on both the N- and C-terminus of the peptides, we explored the a semi-specific search option on the N-terminal and a tryptic cleavage site (e.g. Trypsin: K/R) only on the C-terminus in MaxQuant [60, 62, 322].

We applied our strategy and analysed more than 13 000 rawfiles originating from 6 murine and 41 human datasets (in total: 47) deposited in the public domain. We detect canonical signal peptide cleavage sites in 1300 proteins and alternative cleavage on more than 600 proteins.

4.2.3.1 Localization of neo-N-terminal cleavage sites

Next, we studied the localization of neo-N-terminal peptides in relation to the predicted signal peptide cleavage site by SignalP 5.0. Most cleavages of neo-N-terminal peptides were localized directly at the

predicted canonical CS. Interestingly, there was also an accumulation of cleavages in the direct vicinity of the predicted cleavage site. Most of the not predicted cleavages (75% percentile of neo-N-terminal peptides) occurred in a window of –4 and +6 (-4/+6) amino acids of the predicted signal peptide CS (Figure 2A). As protease activity in cellular lysates during sample preparation may introduce stochastic proteolytic cleavages into proteins, we next evaluated the extend of basal proteolytic processing. Our analysis shows a low-level stochastic cleavage of all proteins, decreasing linearly with increasing amino acid position from the N-terminus (Figure 2B). Cleavages occurred independent of the acetylation status of the N-terminus. Interestingly, neo-N-terminal SP peptides and transit peptides – an n-terminal signal sequence targeting proteins to mitochondria were detected with a likelihood distinct from stochastic occurrence at their site of cleavage, confirming their non-random, enzymatically-dependent appearance. Taken together, our analysis shows specific and proteolytic cleavage patterns at/around the predicted CS that can be discriminated clearly from stochastic background processing.

4.2.3.2 Classification and prioritization of high-confidence neo-N-terminal signal peptide cleavages

In order to prioritize high-confidence neo-N-terminal signal peptide cleavages, we filtered data based on (1) neo-N-terminal peptide intensity and (2) quality of mass spectral evidences (See Supplementary figures S2A, S2B). Based on the rationale that SPP mediated SP cleavage would give rise to higher abundant neo-N-termini than potential sample preparation artefacts, we first excluded neo-N-terminal peptides below the 75th percentile of the normalized peptide intensity distribution. For this analysis, peptide intensities were normalized to the most abundant neo-N-terminal peptide in the allocated cleavage window. An exemption was made for neo-N-terminal peptides with high intensities that were identified in more than one dataset. In total, 572 out of 3688 (15%) potential cleavages were excluded from further analysis (see Supplementary figure S2C). Then, we filtered for high spectral quality of the neo-N-terminal peptides in the selected cleavage window (See Supplementary figure 2D). To achieve this, we used the MaxQuant derived “delta score”, which describes the peptide identification score difference between the best and second best spectral evidences with distinct amino acid sequences. We observed slightly higher delta scores for neo-N-terminal peptides in proteins with signal peptides (98 +/- 57) compared to all proteins (82 +/- 50). As expected, we report higher quality spectra due to the benefits of multiple identifications of the same neo-N-terminal peptide (118 +/- 74) in multiple datasets. In summary, by classifying the neo-N-terminal peptides by intensity and spectral quality, we exclude #572 (15%) potential stochastic peptide evidences and define #3116 high quality SP cleavages.

4.2.3.3 Characterization and categorization of neo-N-terminal peptides

In total, we identify 3510 proteins containing signal peptides according to uniprot. For 1315 proteins we provide spectral evidence for the predicted SPP cleavage site. For an additional 412 proteins we detect alternative cleavage sites +6/-4 aa of the canonical cleavage site (see Figure 3A). On average, we identify 82 CS per dataset and 43 CS per 100 analyzed rawfiles (Supplementary figure S3A). Comparison to previously verified signal peptide cleavages confirms experimental data, and illuminates the N-termini of SP proteins that were formerly only predicted (see Figure 3B). Interestingly, we detect proportionately more proteins with alternative cleavage in proteins with no previous experimental verification of the signal peptide cleavage site. This may suggest that the prediction algorithm does not capture alternative cleavages well.

Besides this, we investigated the number of signal peptide CS detected per protein (see Figure 3C). For most proteins (1070 out of 1712, 62%), we identify only one neo-N-terminal peptide, either at or in close proximity to the predicted cleavage site. Few proteins had two (324, 19%) or more (318, 18%) cleavage sites, which suggests repeated processing of the respective protein by SPPase or single but distinct cleavages. Based on the number and type of cleavages, we categorized the proteins in Category 1 (predicted canonical cleavages), Category 2 (single alternative cleavage), Category 3 (alternative cleavage on proteins with predicted cleavage), and Category 4 (multiple alternative cleavages) (see Figure 3D).

Next, we asked whether proteins with distinct cleavage patterns differ in biological functions. Annotation enrichment analyses revealed that canonical cleavages (Category 1) (see Figure 3E, supplementary table x) are enriched on cytokines (36 of 56 total) and proteins with developmental functions (53 of 88 total). This may suggest, that canonical cleavage sites are conserved, presumably due to evolutionary constraints involving the biological functions of the N-terminus, e.g. as reported for chemokine binding to their receptors. In contrast, Single alternative cleavages (Category 2) are enriched for transmembrane proteins –Type I transmembrane proteins – and receptors. This indicates, the extracellular structure of Type I transmembrane proteins may have an alternative composition. Proteins with immunoglobulin domains (including Bence Jones proteins (9 out of 10 total)), ER localization and transport functions display cleavage sites alternative cleavage on proteins with a canonical cleavage (Category 3). Interestingly, also 12 out of 14 detected isomerases are members of this group. Multiple alternative cleavages (Category 4) are enriched for hydrolases (19 out of 191). Examples for proteins associated with significantly changing annotations are depicted in Supplementary Figure S3B.

Taken together, these results indicate different biological functions for proteins with distinct SP cleavages. Interestingly, transmembrane proteins have a lower ratio of experimentally verified proteins (24%) in comparison to all proteins (34%) or cytokines (59%), that show cleavage sites as predicted (Supplementary figure S3C). This implies that additional training data, as provided by our meta-analysis, could result in improved prediction accuracies for proteins associated with transmembrane annotations.

4.2.3.4 Systematic comparison of cleavage positions by MS and SignalP predictions

Given the apparent discrepancy between experimentally detected and predicted cleavage sites, we next systematically compared the quantitative MS data with the SignalP 5.0 prediction scores. The signal peptide prediction algorithm (1) differentiates between signal peptide-containing proteins and non-secretory proteins; and (2) predicts the signal peptide cleavage position and thereby the protein's post-translational N-terminus. For the first 70 aa of each signal peptide containing protein, SignalP provides three probability scores per position: 1) a signal peptide (SP (Sec/SPI) probability), 2) a cleavage site by SPP (CS probability) and 3) the absence of a signal peptide, e.g. the region downstream of the cleavage site (OTHER probability). Compared to the detected canonical cleavage positions (Cat.1, 3), the detected alternative cleavage positions (Cat 2,4) had significantly lower cleavage site probabilities for containing a Sec/SPI signal peptide (Figure 4A, Supplementary figure S4A), SP(Sec/SPI) probabilities (Supplementary figure S4B) and higher OTHER probabilities (Supplementary figure S4C).

We next evaluated whether peptide intensities correlate to CS probability scores from Signal P and aid in prioritizing high-confidence experimental n-termini. Therefore, we plotted for each n-terminus the ratios of actual versus predicted CS probability and peptide intensity. Our analysis detects n-termini with high probability cleavage sites, with both low (Quadrant 1) or high (Quadrant 2) relative abundances (Figure 4C). Likewise, n-termini with low CS probabilities at the experimentally identified sites – for example alternative cleavage sites - display both lower (Quadrant 3) or similar (Quadrant 4) relative abundances (Figure 4D). This analysis enables us to identify alternative cleavage sites with high relative CS probability scores from signal P (Quadrant 1, 2).

Next, we asked whether proteins in the four different quadrants are enriched for functionally distinct classes as defined by the detected cleavage patterns (Category 1-4, Figure 3D) (See Figure 4E). Cytokines and proteins involved in neurogenesis, differentiation and development have canonical cleavage sites as predicted and these have high peptide intensities (Quadrant 2). Interestingly, for transmembrane, glycoproteins and receptors we detect a single cleavage, either at an alternative cleavage site with high intensity and high cleavage site probability or predicted as the major cleavage site by SignalP (Quadrant

2). Isomerases, hydrolases and ER associated proteins display multiple high probability cleavage sites, that are experimentally of similar abundance as the predicted cleavage site (Quadrant 2). The multiple detected cleavage sites of immunoglobulins generally show low cleavage site probability and have a broad abundance distribution. Taken together, our analysis reveals distinct functional protein families with signal peptides and differentiates them by mode and amount of experimental cleavage sites in the vicinity of the predicted SPP cleavage site.

4.2.3.5 Consensus sequences of the cleavage sites

As the primary structure of the signal peptide defines the SPP cleavage position and it's prediction, we analyzed the amino acid sequences around the cleavage sites. To this end, we use …. An important parameter to consider are the amino acids at the positions -1 and -3 of the predicted cleavage site, which are in case of Category 1 proteins (proteins with only one cleavage site) primarily composed of A, G, S on position -1 and A, G, S, C on position -3 (Figure 5A, Figure S5A). In comparison to proteins with canonical cleavages (Cat. 1), protein sequences with alternative cleavage in quadrant 2 are enriched for serine at position -1 (Cat. 2,3,4) and alanine at -3 (Cat.3,4) of the predicted cleavage site (Figure 5B). Inversely, the valine proportion at position -3 and glycine proportion at position -1 in the predicted sequence are reduced in all proteins with experimentally observed alternative cleavage in Quadrant 2 (Cat.2,3,4) (Figure 5B). In Quadrant 4 the aminoacids are more heterogeneously distributed along the categories. Common to all proteins with alternative cleavage, alanine on position -3 (Cat.2,3,4) is enriched while glycine at position -1 (Cat.2,3,4) of the predicted cleavage site is reduced (Supplementary figure S5B). Alternative cleavage is also observed for proteins with increased glutamine on position +1 of the predicted cleavage site both in Quadrants 2 and 4 (Figure 5B, S5B). Conversely, the experimental protein sequences aligned to the experimental cleavage site display weaker enrichments of amino acids at position -3 and -1 in comparison to the predicted cleavage sites. However, alanine is proportionally higher enriched at positions -3 and -1, glycine at position -1 and interestingly also proline is enriched at position -1 in Cat.2 alternative cleavage in comparison to cleavage as (Figure 5C, Supplementary figure S5C).

Another important parameter of signal peptide structure is the hydrophobic domain – often composed of leucine, alanine and valine - upstream of the cleavage site, which is involved crucially in the signal peptide translocation and cleavage [323, 324]. The length of the hydrophobic domain is less variable than amino acids on positions -3 and -1 in the predicted sequences of the proteins belonging to different categories and quadrants. However, apart from the alternative cleavage Type 2 proteins with high probability cleavage sites and intensity (Quadrant 2) the hydrophobic domain is shifted away from the cleavage site

for the experimentally observed cleavage sites in comparison to the predicted cleavage site and seems to be stretched (Figure 5A, Supplementary figure 5A).

Next, we were also compared the protein sequences surrounding the CS of transmembrane proteins with predicted or alternative cleavage. Transmembrane proteins with alternative cleavage harbor higher probabilities for G, T, S on position -3 of the predicted cleavage site (Figure 5D). A, V and L are enriched on the position -3 in the consensus sequences around the cleavage site of experimentally observed alternative cleavage (Figure 5E, Supplementary Figure S5D).

In summary, the consensus sequences for experimental alternative cleavage possess weaker enrichments of amino acids compared to the predicted sequence, which might explain, why the algorithm chooses the other site as the predicted cleavage site. Furthermore, the general diversity on positions -1, -3 and +1 of the predicted signal peptides between the different experimental protein categories pinpoints towards subsets of proteins with certain AA compositions harboring alternative signal peptide cleavage as opposed to the predicted cleavage.

4.2.4 Discussion

Usually, signal peptide cleavage is predicted computationally by algorithms like SignalP, as experimental approaches are often laborious and single protein based. Here, we introduced a proteomics meta-analysis workflow for dissecting signal peptide cleavage in already existing, published datasets by employing an innovative computational strategy to identify neo-N-terminal peptides derived from SPP cleavage. Thus, we integrated the spectral information from over 30 individual datasets spanning more than 13000 rawfiles to locate signal peptide cleavage sites. In total, we identified 1315 signal peptide cleavage sites in human and murine tissues as predicted by SignalP. This more than doubles the currently available experimental evidence for these two species. Therefore, the results of our meta-analysis can serve as a resource for the research community, but also the advancement of prediction algorithms. Notably, our analysis also identifies 412 alternative cleavage sites in the vicinity of the predicted cleavage site (-3/+7 aa).

One of the major issues is the discrimination between bona fide cleavage sites and random sample preparation or non- SPP enzymatic derived neo-N-terminal peptides. Our key observation, that neo-N-terminal peptides are increasing non-linearly around the predicted cleavage site corroborates the enzymatically catalyzed reaction, as we observe a linear increase for proteins with no predicted signal peptides. We carefully selected the cleavage window around the predicted cleavage site we consider for potential signal peptide cleavage. To further limit the number of non SPP derived neo-N-terminal peptides

we introduced a stringent quality control including the neo-N-terminal peptide intensity, the quality of the spectra, and the number of replicates per dataset. This analysis lead to the exclusion of 15% of the potential cleavage sites.

More 60% of the proteins with signal peptide cleavage sites harbor one cleavage site, but also multiple has been observed. This led to the categorization of the proteins by type and number of cleavage sites into four categories. An explanation for multiple cleavages per protein may be diverse biological functions (e.g. immunoglobulin domains are enriched in Cat. 3), do not require an exact N-terminus, and SPPases cleave more often. A second explanation may be that Category 3 and 4 proteins are cut by different SPPase subclasses than Category 1 and 2 proteins. Accordingly, Cat. 1 and 2 proteins would be predominantly cleaved by SPPases that produce one bona-fide cleavage site. An indication for cleavage by different subclasses of SPPases is the different consensus sequences for the Cat. 1-4.

Our analysis of functional annotations of proteins in different cleavage categories, revealed signal peptides of cytokines, growthfactors, developmental proteins and proteins involved in differentiation to be cleaved as predicted. This may indicate, that cleavage sites on these proteins are conserved, and SPPases produce only one N-terminus, presumably due to evolutionary constraints involving their biological function. Additionally, transmembrane proteins either possess one cleavage site as predicted or one alternative cleavage site. An alternative N-Terminus of these type I transmembrane proteins could lead to differences in the intracellular signal transduction cascades they are involved in.

Our analysis of the sequence motifs aligned to the predicted cleavage site further showed differences between conventional and alternative cleavage of the signal peptides. Most prominently, alternative cleavage is observed for a subgroup of proteins, that have Q on position +1 and P on position -1 of the predicted cleavage site. While the sequence conservation in the vicinity of the cleavage site is weaker for proteins with alternative cleavage than for their predicted counter-part weaker conservation of A on positions -3, -1 and Glycine -1, S on -1 to -3 of the predicted cleavage sites can be observed when aligning by the experimental cleavage site. This weaker enrichment of these amino-acids might be the reason for prediction algorithms to choose another site as the highest probable cleavage site.

In the future, a similar meta-analysis could be performed for additional species to achieve higher and more comprehensive experimental evidence for signal peptide cleavage also in other organisms. With our meta-analysis we provide the experimental MS-evidences to re-train the SP prediction model, which is based on more than 700 proteins with novel experimentally detected SP sites that match the predicted

cleavage site. The alternative cleavages – enriched particularly in transmembrane proteins – indicate additional processing at distinct consensus sequences and can be used to extend the SP prediction model. Particularly from transmembrane proteins, we expect advanced prediction abilities of the algorithms. To enable easy availability of our signal peptide cleavage meta-analysis we have made the data available online in form of an easy searchable database.

4.2.5 Figures and figure legends

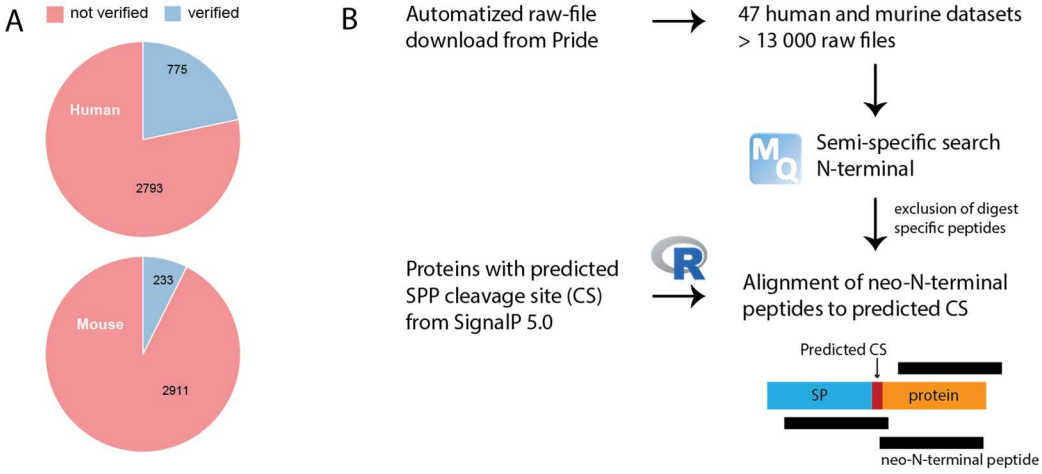


Figure 1: Meta-analysis strategy for the identification of signal peptide cleavages in shotgun proteomics data sets. (A) Number of human or murine proteins containing signal peptides according to Uniprot. Previously experimentally verified SPP cleavage sites as listed by Uniprot are shown in blue. Not experimentally verified SPP cleavage sites are shown in red. (B) Computational workflow including data download from Pride, semi-specific N-terminal spectral search to detect neo-N-terminal peptides, comparison to signal peptide cleavage prediction of SignalP 5.0, and alignment of the neo-N-terminal peptides to the predicted cleavage site.

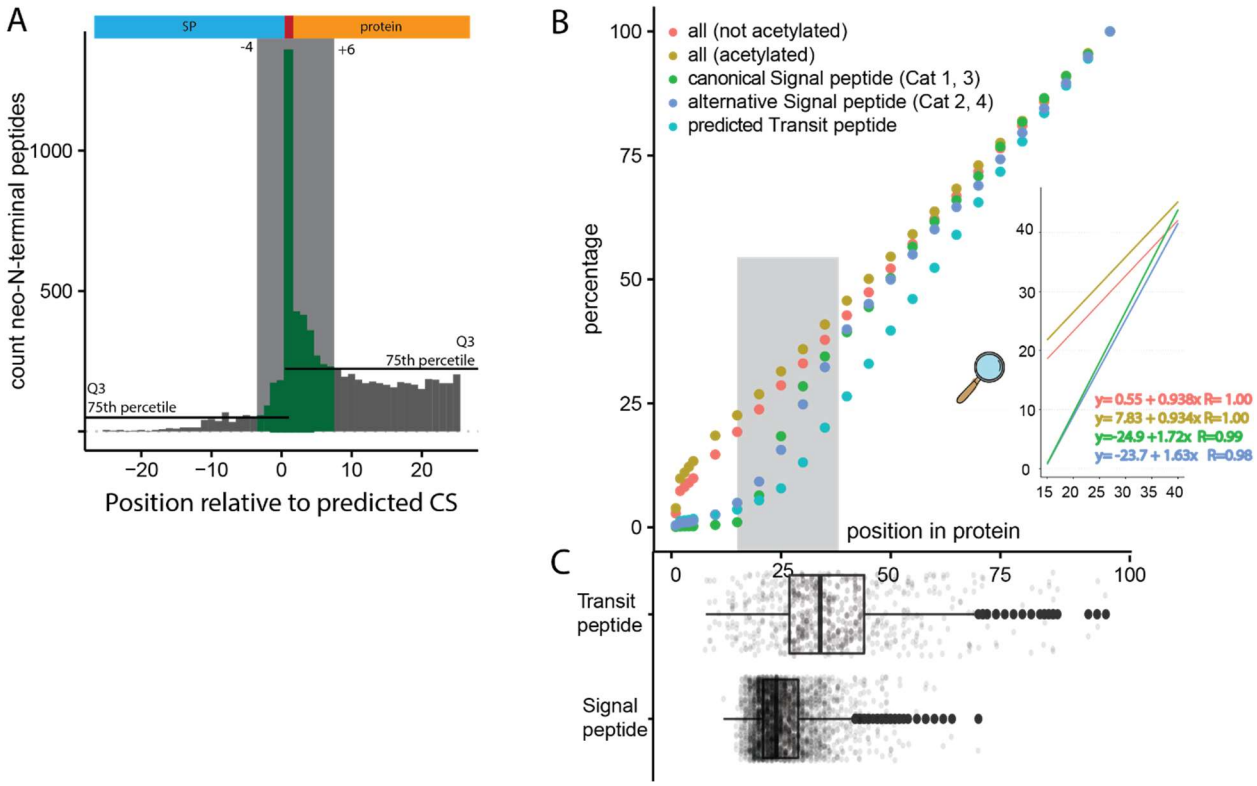


Figure 2: Localization and characterization of neo-N-terminal peptides (A) Occurrence of neo-N-terminal peptides relative to the predicted cleavage site (CS). Q3 (75th percentile) of the number of identified neo-N-termini was calculated for each amino acid position upstream and downstream of the predicted CS, respectively. Positions within the 75th percentile are colored in green. (B) Localization versus occurrence of Neo-N-terminal peptides within the first 100 AAs of all proteins with (fawn) or without acetylation (red), signal proteins with predicted (green, Categories 1 and 3) or alternative (blue, Categories 2 and 4) cleavage and proteins with Transit peptides with predicted (turquoise) or alternative (violet) cleavage. (C) Predicted localization of Transit or Signal peptides in the protein.

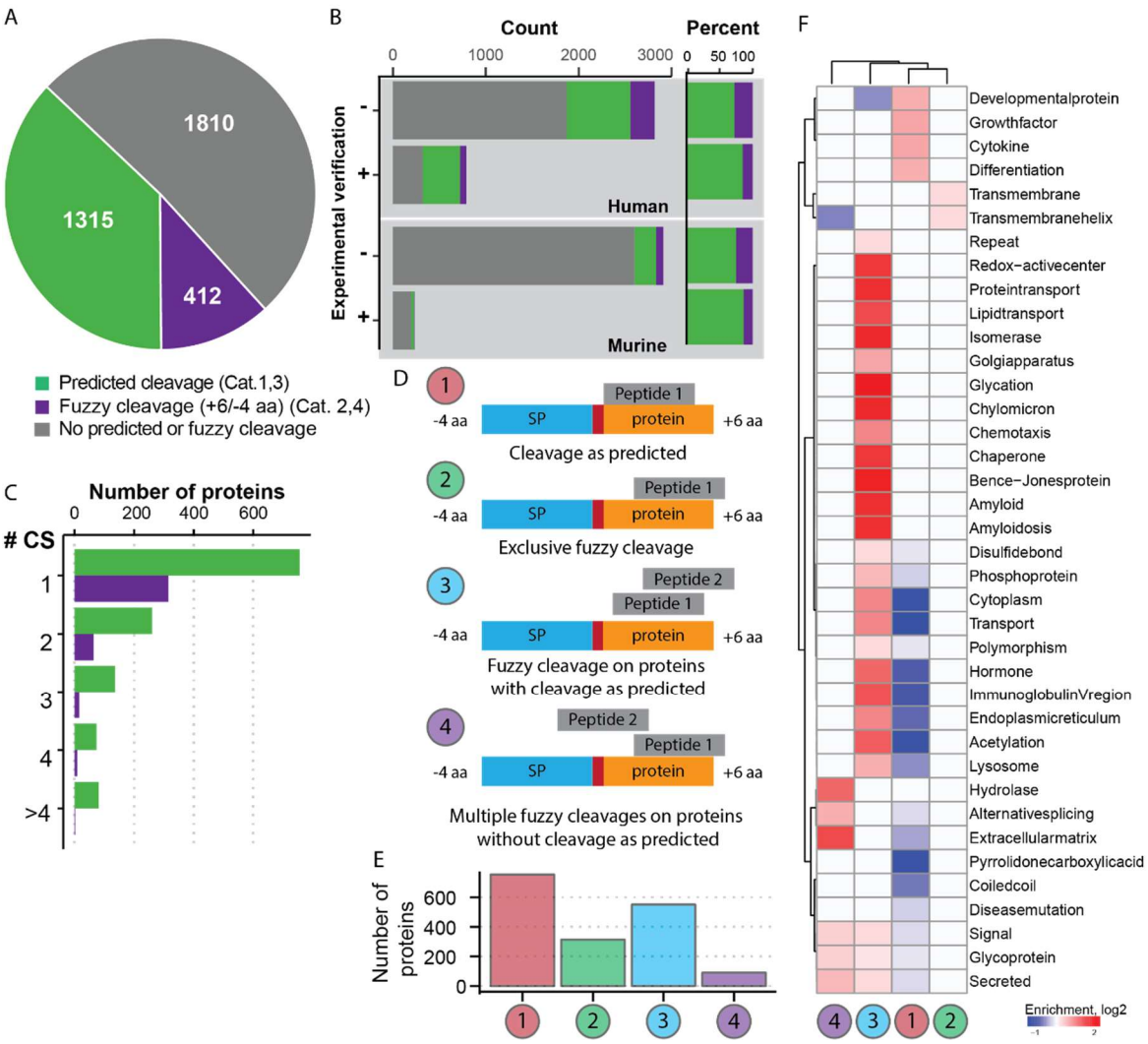


Figure 3: Categorization and functional classification of Neo-N-terminal peptides (A) Number of proteins (combination of mouse and human) with MS evidence for canonical cleavages (Cat. 1, 3: green), alternative cleavages (Cat. 2, 4: violet) or no cleavages (grey). (B) Number and percentage of proteins with MS evidence for canonical (green), alternative (violet) or no cleavage (grey) in experimentally verified or not verified SPs. (C) Bar plots showing the number of proteins with MS detected Neo-N-terminal peptides in the cleavage window (+6/-4 aa of the predicted CS) for proteins with canonical (green, Cat. 1 and 3) or alternative (violet, Cat. 2 and 4) cleavages. (D) Overview of possible cleavage categories: (Category 1) Proteins that exclusively harbor a canonical cleavage site as predicted by SignalP 5.0.; (Category 2) Proteins that exclusively harbor an 'alternative' cleavage site; (Category 3) Proteins that harbor a cleavage site as predicted by SignalP 5.0 plus an additional 'alternative' cleavage; (Category 4) Proteins harboring more than one 'alternative' cleavage. (E) Number of proteins per category. (F) Heatmap displaying enriched proteins of Category 1-4 based on a Fisher exact test (p-value< 0.01, intersection>3). Red corresponds to enrichment, blue to de-enrichment.

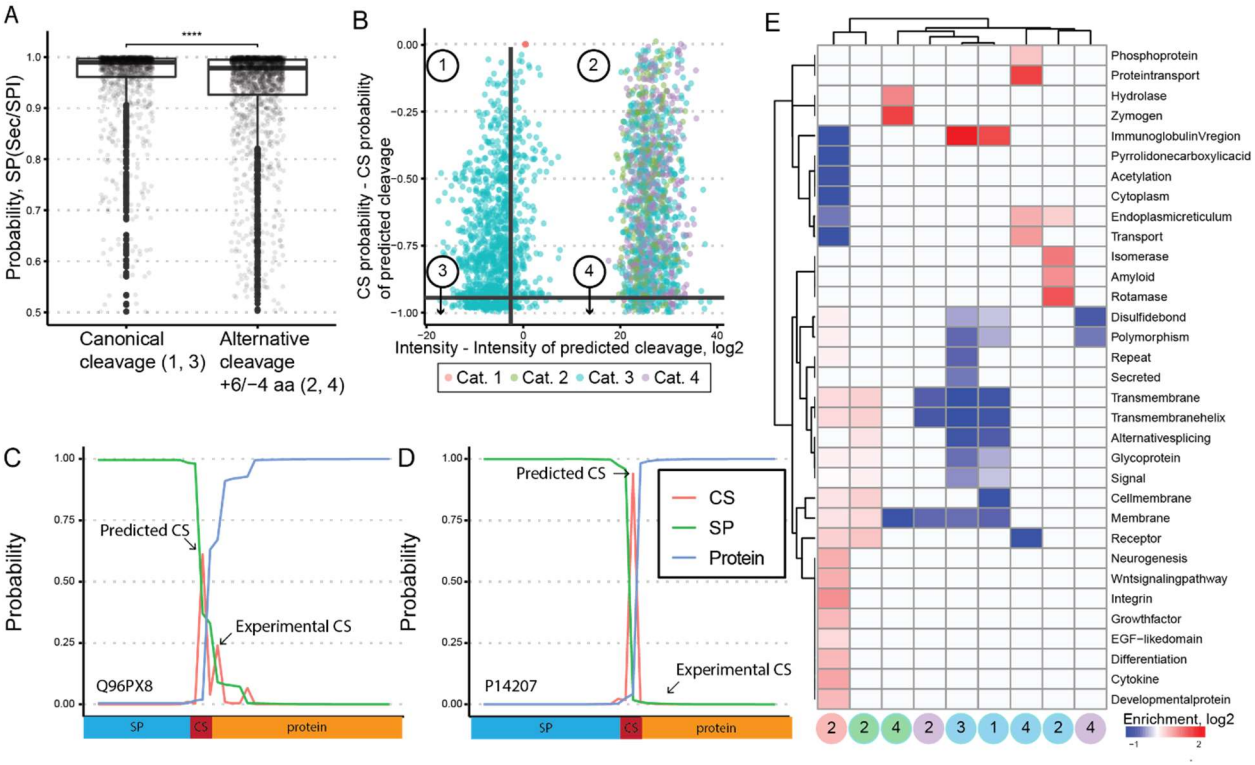


Figure 4: Systematic comparison of MS and SignalP scores and sequences (A) Probability of the protein harboring a signal peptide (Sec/SPI, SignalP 5.0) for all human proteins containing a signal peptide according to Uniprot. Dots show probabilities of individual canonical (Categories 1 and 3) or alternative (Categories 2 and 4) cleaved proteins. Two-sided T-test was performed between the two groups. ****: p-value < 0.0001. (B) Comparison of the delta cutting site (CS) probability (by SignalP 5.0 for each position of each protein) to the delta peptide Intensity (log2). Delta Cutting site probability and delta Intensity (log2) were calculated by subtracting the intensity/ CS probability of predicted site from the intensity/ CS of each site (= normalization to cleavage as predicted). The cleavage categories are colored in the four quadrants (1, 2, 3, 4) of the plot: red (cleavages of proteins with canonical cleavage, Category 1), blue (cleavages of proteins with canonical and alternative cleavages sites, Category 3), green (Cleavages on proteins with single alternative cleavage, Category 2), violet (cleavages of proteins with multiple alternative cleavage, Category 4). (C) Cleavage site (CS), Signal peptide (SP) and protein probability predicted by SignalP 5.0 for Q96PX8/ SLITRK1. Predicted and experimental cleavage sites are highlighted by an arrow. (D) Cleavage site (CS), Signal peptide (SP) and protein probability predicted by SignalP 5.0 for P14207/ FOLR. Predicted and experimental cleavage sites are highlighted by an arrow. (E) Hierarchical clustering displaying fisher exact test results (p-value < 0.01, intersection >3) on Uniprot Keywords for proteins within different quadrants from B. Red shows enrichment, blue de-enrichment. Annotations from proteins with caonical cleavage (Category 1, quadrant 2) and single alternative cleavage (Category 2, quadrant 2) are clustering together.

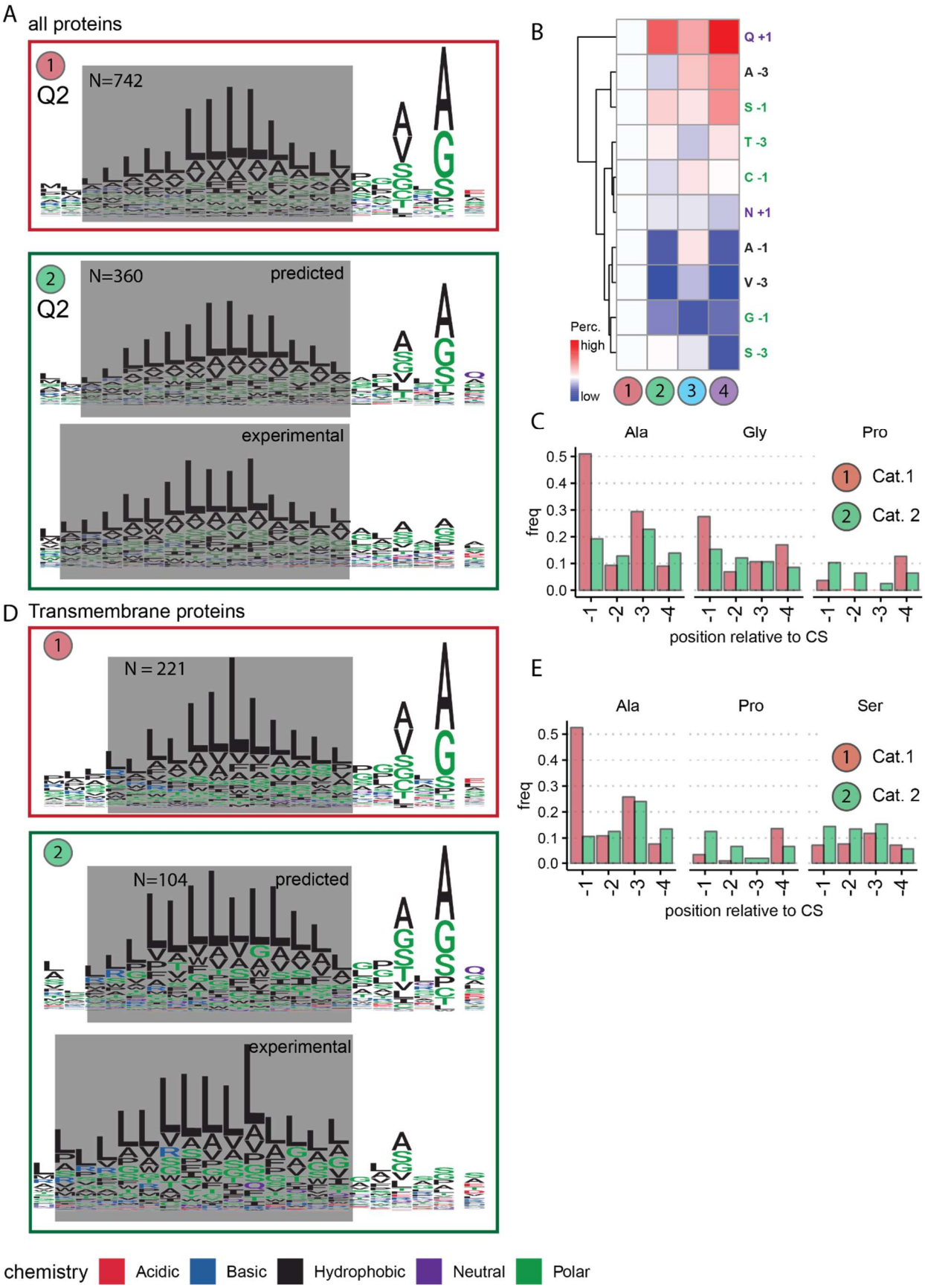
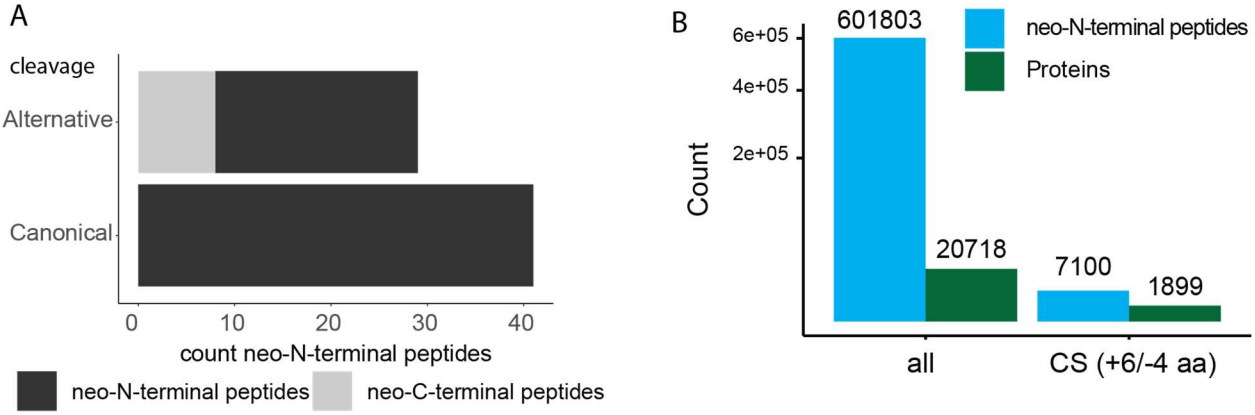
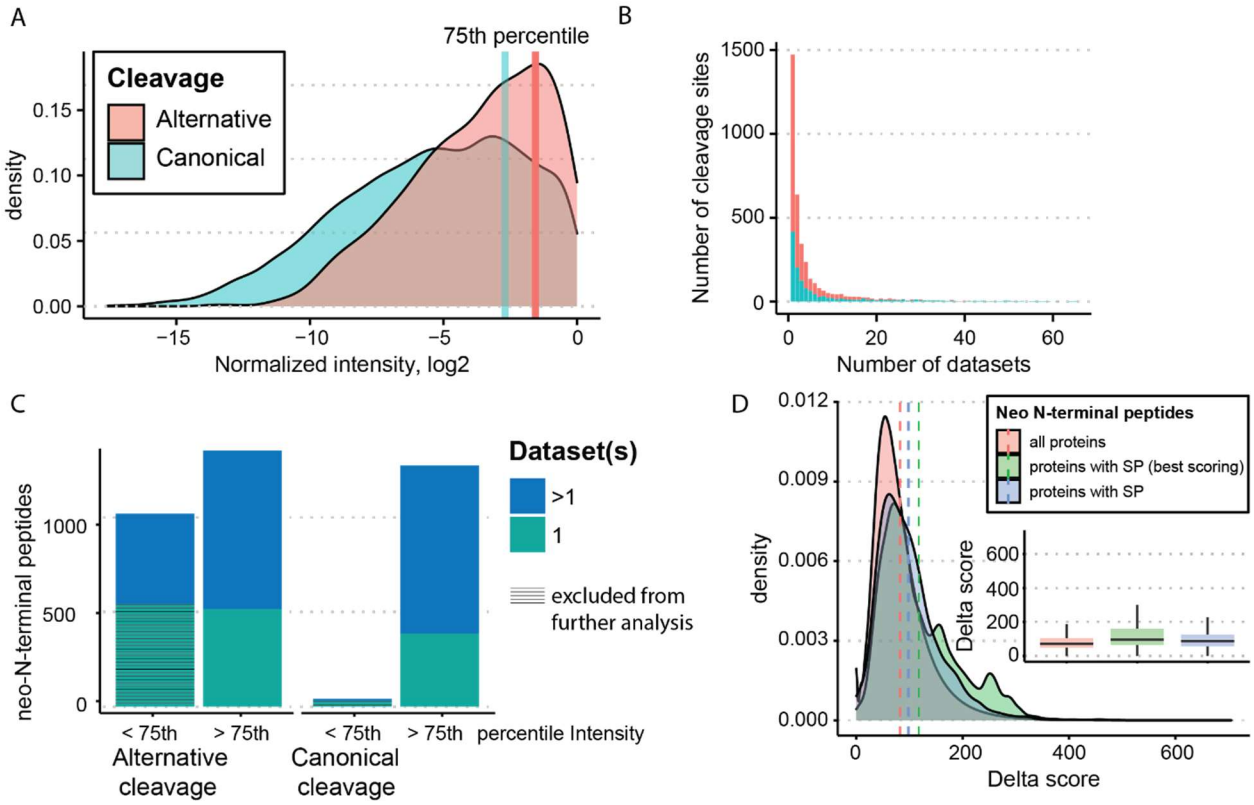

A
all proteins
1
Q2
N=742
2
Q2
N=360
predicted
experimental
B
Q +1
A -3
S -1
T -3
C -1
N +1
A -1
V -3
G -1
S -3
Perc.
high
low
1
2
3
4
C
Ala
Gly
Pro
freq
0.5
0.4
0.3
0.2
0.1
0.0
-1
-2
-3
-4
position relative to CS
1
Cat.1
2
Cat. 2
D
Transmembrane proteins
1
N = 221
2
N=104
predicted
experimental
E
Ala
Pro
Ser
freq
position relative to CS
1
Cat.1
2
Cat. 2
chemistry
Acidic
Basic
Hydrophobic
Neutral
Polar

Figure 5: Sequence analysis of canonical versus alternative cleavages. (A) Consensus sequence of all proteins with canonical cleavage only (Cat. 1, red) and single alternative cleavage (Cat. 2, green). Cleavage position is indicated by blank space. Motifs are sorted by quadrants (Q1, Q2) from Figure 4B. For proteins with alternative cleavage (Cat. 2), the motifs show the canonical and the experimental consensus sequence. N shows number of proteins corresponding to each sequence motif. Amino acids are colored by functional chemical groups. Grey boxes highlight the hydrophobic domain of the signal peptides. (B) Heatmap showing percentage differences of dominant amino acids of the canonical consensus sequence on indicated positions (-1, -3, +1) relative to the predicted cleavage site for different cleavage category of proteins with CS probability and intensity. Red is corresponding to higher, blue to lower percentages than the Cat. 1 consensus sequence. (C) Percentages of amino acids Ala, Gly, Pro for positions -1, -2, -3, -4 of the predicted (Cat. 1) or experimental (Cat. 2) cleavage site of all proteins. (D) Consensus sequence of transmembrane proteins with predicted cleavage only (Cat. 1, red) and single alternative cleavage (Cat. 2, green. For transmembrane proteins with alternative cleavage (Cat. 2) the motifs under predicted cleavage show the predicted and the experimental consensus sequence. Grey boxes highlight the hydrophobic domain of the signal peptides. (E) Percentages of amino acids Ala, Gly, Pro for positions -1, -2, -3, -4 of the predicted (Cat. 1) or experimental (Cat. 2) cleavage site of transmembrane proteins.

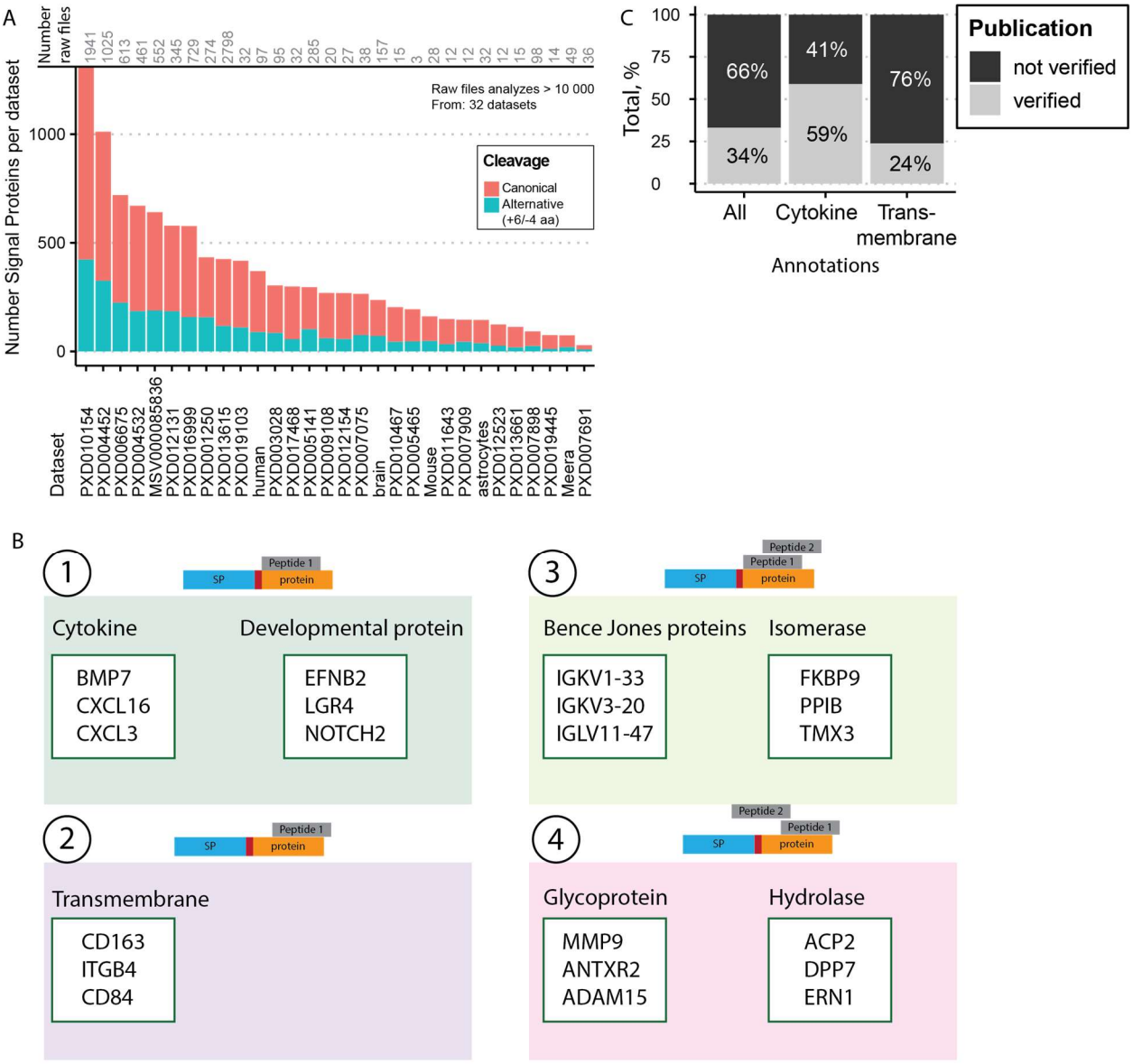
4.2.6 Supplementary figures and figure legends



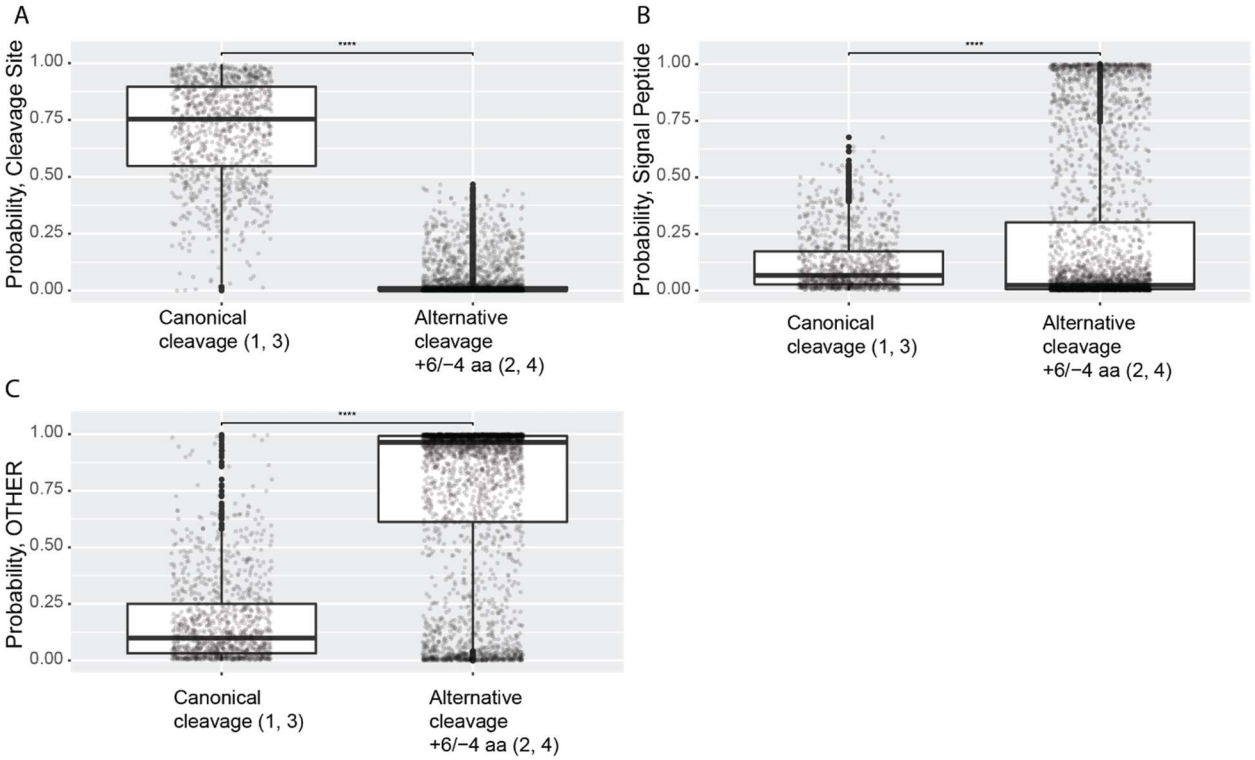
Supplementary Figure 1: (A) Number of neo-N-terminal peptides for canonical and alternative cleavage (+6/-4 aa) by N- or C-terminal semi-specific search algorithms. (B) Number of proteins and neo-N-terminal peptides of all proteins and proteins with neo-N-terminal peptides in the vicinity of the predicted CS (+6/-4 aa).



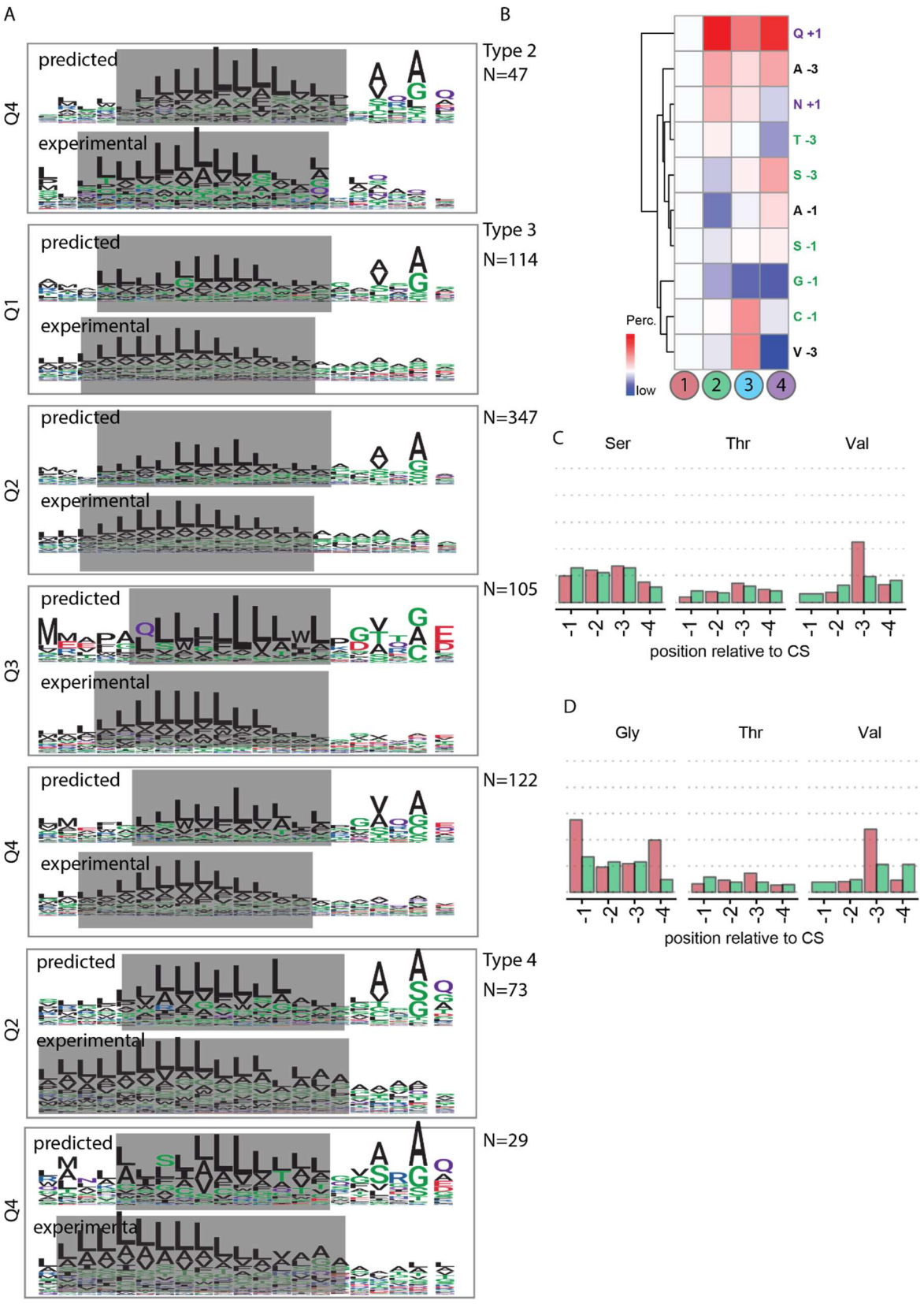
Supplementary figure 2: (A) Density distribution of normalized neo-N-terminal peptide intensities (log2) corresponding to the predicted CS (turquoise) and alternative cleavages (red) in the vicinity of the predicted cleavage site (-4/+6). Peptide intensities were normalized to the most abundant neo-N-terminal peptide in the cleavage window. Vertical lines denote the 75th percentile (Q3) of the neo-N-terminal peptide intensities. (D) Reproducibility of predicted (turquoise) and alternative (red) cleavage sites along different datasets. (C) Number of neo-N-terminal peptides with predicted or alternative cleavage dependent on the intensity threshold (75th percentile) and the reproducibility between datasets. Neo-N-terminal peptides below the 75th percentile and identification in only one dataset were excluded from further analysis (bars with stripes). (D) Density and box plots of the MaxQuant derived delta score (Score difference to the second-best identified peptide with a different amino acid sequence) of identified Neo-N-terminal peptides on all proteins (red) and proteins with predicted signal peptides (SP) (green and blue). Delta scores of Neo-N-terminal peptides on SP from all datasets are shown in blue, whereas the highest delta score of SPs across all datasets is shown in green. The latter has an increased median delta score (dashed lines) compared to SPs from all datasets.



Supplementary figure 3: (A) Overview of analyzed datasets. The number of identified signal proteins with either predicted (Category 1,3: turquoise) or alternative (Category 2, 4: red) cleavages are indicated. The number of raw files in each dataset is shown on top. (B) Protein examples of the four main cleavage categories described in Figure 3D. 3 proteins per annotation are listed. (C) Ratio of experimentally verified and not verified signal peptide cleavage sites in proteins with Uniprot keyword annotations “Cytokine”, “Transmembrane” or all identified proteins.



Supplementary figure 4: (A) Probability of the experimental cleavage site being a cleavage site for all human proteins identified with our meta-analysis. (B) Probability of the experimental cleavage site being localized on the signal peptide (Sec/SPI, SignalP 5.0) for all human proteins identified with our meta-analysis. (C) Probability of the experimental cleavage site being localized downstream of the predicted cleavage site for all human proteins identified with our meta-analysis. Dots show probabilities of individual experimental predicted (Categories 1 and 3) or alternative (Categories 2 and 4) cleavage sites. Two-sided T-test was performed between the two groups. ****: p-value < 0.0001.


A
predicted
experimental
Q4
Type 2
N=47
Q1
Type 3
N=114
Q2
N=347
Q3
N=105
Q4
N=122
Q2
Type 4
N=73
Q4
N=29
B
Q +1
A -3
N +1
T -3
S -3
A -1
S -1
G -1
C -1
V -3
Perc.
low
1
2
3
4
C
Ser
Thr
Val
-1
-2
-3
-4
position relative to CS
D
Gly
Thr
Val
position relative to CS
chemistry
Acidic
Basic
Hydrophobic
Neutral
Polar

Supplementary figure 5: (A) Consensus sequence of all proteins with single alternative cleavage (Cat. 2, green), alternative cleavage on proteins with predicted cleavage (Cat. 3, blue) and multiple alternative cleavage (Cat. 4, violet). Cleavage site position is indicated by blank space. Motifs are sorted by quadrants (Q1, Q2, Q3, Q4) from Figure 4B, if applicable. For proteins with alternative cleavage (Cat. 2, Cat. 3. Cat. 4) the motifs under predicted cleavage show the predicted and under experimental the experimental consensus sequence. N shows number of proteins corresponding to each sequence motif. Amino acids are colored into chemistry groups. Grey boxes highlight the hydrophobic domain of the signal peptides. (B) Heatmap showing differences in percentages of dominant amino acids of the predicted consensus sequence on certain positions (-1, -3, +1) relative to the predicted cleavage site for different cleavage category in Quadrant 4 from Figure 4B. Red is corresponding to higher, blue to lower percentages than the Cat. 1 consensus sequence. (C) Percentages of amino acids Ser, Thr, Val for positions -1, -2, -3, -4 of the predicted (Cat. 1, red) or experimental (Cat. 2, green) cleavage site of all proteins. (D) Percentages of amino acids Gly, Thr, Val for positions -1, -2, -3, -4 of the predicted (Cat. 1, red) or experimental (Cat. 2, green) cleavage site of transmembrane proteins.

4.2.7 Methods

4.2.7.1 Download datasets from pride

Proteomics datasets from species homo sapiens or mus musculus were selected. They were selected preferentially, if high number of raw files were present. Selected datasets were measured on QEPs, Q-HF, Q-HF-X, Fusion and Exploris. Selected datasets include digest with Trypsin, LysC, Chymotrypsin, GluC, LysN, ArgC and AspN. They were automatically downloaded via a FTP download script in R, that needs the html webaddress and the number of rawfiles of each dataset as an input.

4.2.7.2 MaxQuant analysis

MaxQuant software was used to analyze MS raw files. MS/MS spectra were searched against the human or murine Uniprot FASTA database and a common contaminants database (247 entries) by the Andromeda search engine [62]. Cysteine carbamidomethylation was set as a fixed modification, and N-terminal acetylation and methionine oxidation were set as variable modifications. Enzyme specificity was set to semi-specific N-terminal digest (enzyme depending on the dataset analyzed), with a maximum of two missed cleavages and a minimum peptide length of seven amino acids. FDR of 1% was applied at the peptide and protein level using a reverse database for target decoy. Peptide identification was performed with an allowed initial precursor mass deviation of up to 7 ppm and an allowed fragment mass deviation of 20 ppm and hits to the reverse database [62] were excluded from further analysis.

4.2.7.3 Cleavage site prediction and probabilities of signal peptides from SignalP

All human/murine proteins with possible signal peptides were identified via Uniprot (Download September 2020). Their FASTA files were compiled and analyzed in batches of 200 proteins via SignalP 5.0 (September 2020). Raw SignalP output files were compiled into one file and predicted cutting sites as well as cutting site probabilities of each position on all signal proteins were extracted. Cleavage site probabilities and the predicted cleavage site were matched to the positions of the semi-specific peptides of signal proteins.

4.2.7.4 Transit peptides

All human/ murine proteins with possible Transit peptides were identified via Uniprot (Download September 2020) and cleavage site predictions were matched to proteins via Uniprot IDs. If the location

of the semi-specific peptide was at the predicted cleavage site, the cleavage was called cleavage as predicted, whereas it was called alternative cleavage, if the location of the semi-specfic peptide was within +6/-4 AA of the predicted cleavage site.

4.2.7.5 Identification of predicted and alternative cleavage sites and protein categorization

Semi-specific peptides were filtered out of the dataset, while non semi-specific peptides derived from digests with Trypsin, LysC, Chymotrypsin, Gluc, LysN, ArgC and AspN were disregarded. Semi-specific peptides from proteins containing signal peptides were first matched to their respective protein sequence by their Uniprot ID, before the peptide was aligned to the protein sequence to identify the location of the semi-specific peptide. Neo-N-terminal originating from missed cleavage events, e.g. same N- but different C-terminus, were summarized into one cleavage site. Cleavage sites were aligned to the predicted cleavage site and the positions with the highest number of cleavage sites (70^{th} percentile, Q3) were identified, leading to an expanded cleavage window of -4/+6 amino acids around the predicted cleavage site. If the location of the semi-specific peptide was at the predicted cleavage site, the cleavage was called cleavage as predicted, whereas it was called alternative cleavage, if the location of the semi-specific peptide was within +6/- 4 aa of the predicted cleavage site.

Proteins with alternative or predicted cleavage were classified into four categories by number and type of cleavage:

Category	**Number of cleavages**	**Cleavage as predicted**	**Alternative cleavage**
Cat.1	1	1	NA
Cat.2	1	NA	1
Cat.3	>1	1	>=1
Cat.4	>1	NA	>1

4.2.7.6 Delta Score optimization

MaxQuant calculates for each peptide the Score difference (Delta Score) to the second best identified peptide with a different amino acid sequence. The Score is the andromeda score for the best associated MS/MS spectrum. The Delta Score of each peptide can be found in the msms outpt table. As multiple datasets were analyzed and also missed cleavages or different enzymatic digests can lead to repeated coverage of one cleavage site, the peptide with the highest Delta Score per position +6/-4 AA to the cleavage site was selected for further comparisons. Also, spectra depicted in the shiny app were selected based on highest delta score across datasets.

4.2.7.7 Normalization of peptide intensities and filtering by intensity

Per potential cleavage site within the cleavage window (+6/-4 aa to the predicted cleavage site) the semi-specific peptide with the maximum intensity of all raw files per dataset was identified. Logarithmized (log2) intensity of each semi-specific peptide was then normalized to the intensity of the semi-specific peptide with the maximum intensity individually for each dataset (Delta Intensity). Finally, the highest normalized intensity across datasets was selected for each cutting site. Peptides with a delta Intensity larger smaller than -2 were disregarded from further analysis, if they were only identified in one dataset and semi-specific peptides of similar intensity remained.

4.2.7.8 Normalization of peptide intensities from MQ and cleavage site probabilities SignalP5.0

Peptide within the cleavage window (+6/-4 aa to the predicted cleavage site) with maximum intensity of all raw files per dataset was identified. Peptide intensity of all potential alternative cleavages and predicted cleavage was then normalized to the intensity of the peptide containing the predicted cleavage site individually for each dataset. In case, no exact cleavage site has been identified, peptide intensity of the exact cleavage site was set to zero. Finally, the highest normalized intensity across datasets was selected for each cleavage site. Similarly, cleavage site probabilities of all cleavage sites extracted from SignalP 5.0 were normalized to the cleavage site probability of the predicted cleavage site.

4.2.7.9 Functional annotation of different cleavage categories with Fisher Exact test

Proteins were characterized based on number and type of cleavage into four categories. Fisher Exact test was performed in Perseus [325] with an p-value cutoff (p-value < 0.01) and a minimum of 3 proteins required at the intersection. Results were exported from Perseus and differences in annotation enrichment were visualized in R using the “heatmaply” package, by clustering columns and rows with Pearson correlation.

4.2.7.10 Analysis of the consensus sequence around the predicted or experimental cleavage sites

Proteins were aligned either at the predicted or experimental cleavage site from position -20 to +1 of the predictes/ experimental cleavage site. In case of shorter proteins, relevant positions were filled with blank spaces. Sequence logos were created in the R-environment using the “ggseqlogo” package with the parameters seq_type='aa' and method='bits'. For the heatmap of the predicted sequences to compare aa at different positions of cleavage categories, percentages of the individual aa were calculated and then compared to Cat. 1 by subtraction.

4.3 Project 3: Quantitative Proteomics of Secreted Proteins

Global studies of cellular secretomes are now possible due to advancements in experimental procedures, instrument performance, and computational analysis tools. Historically, secreted proteins have been quantified using antibody-based methods, such as enzyme-linked immunosorbent assays (ELISA), whose large-scale applicability is limited by their availability, specificity, and affordability. Our group has devised a first proteomics-based secretome workflow in 2013 [280]. This process allows the direct quantification and study of the secretome of activated immune cells by proteomic analysis. To explain all experimental and analytical procedures in detail and provide a universal guide for guide for experimental secretome investigations, I wrote a step-by-step [294] workflow for quantifying cellular protein secretion comprehensively using mass spectrometry-based proteomics. Briefly, *in vitro* or *ex vivo* secreted proteins are collected, digested by proteases, and the resultant peptide mixtures examined in single LC-MS/MS runs. The MaxQuant and Perseus computational environments are used for label-free quantification and bioinformatics analysis. This workflow enables the quantification of thousands of secreted proteins over a four-order-of-magnitude concentration range, allowing for the system-level analysis of secretory programs and the identification of proteins with unanticipated extracellular activities. The protocol discusses the pitfalls and alternative strategies experimental or analytical strategies the researcher can undertake. The entire protocol takes about two days, whereas MS measurements require less than 2h per sample.

We published the detailed protocol in Methods in Molecular Biology in 2018.

Chapter 14

Quantitative Proteomics of Secreted Proteins

Annika Frauenstein and Felix Meissner

Abstract

Secreted proteins such as cytokines, interleukins, growth factors, and hormones have pleiotropic functions and facilitate intercellular communication in organisms. Quantification of these proteins conventionally relies on antibody-based methods, i.e., enzyme-linked immunosorbent assays (ELISA), whose large-scale use is limited by availability, specificity, and affordability.

Here, we describe an experimental and bioinformatics workflow to comprehensively quantify cellular protein secretion by mass spectrometry. Secreted proteins are collected in vitro or ex vivo, digested with proteases and the resulting peptide mixtures are analyzed in single liquid chromatography–mass spectrometry (LC-MS/MS) runs. Label-free quantification and bioinformatics analysis is conducted in the MaxQuant and Perseus computational environment. Our workflow allows the quantification of thousands of secreted proteins spanning a concentration range of four orders of magnitude and permits the systems-level characterization of secretory programs as well as the discovery of proteins with unexpected extracellular functions.

Key words Mass spectrometry, Quantitative proteomics, Label-free quantification, Secretome, Secreted proteins, Cytokines, Interleukins, Interferons, Growth factors, Sample preparation

1 Introduction

Mass spectrometry-based proteomics is a powerful method for the comprehensive characterization of protein expression differences, protein-protein-interactions, and posttranslational modifications [1–3]. Advances in sample preparation, instrumentation, and computational analysis enable quantitative measurement of thousands of proteins in single-shot experiments [4]. Recently, we developed a sensitive label-free high-resolution mass spectrometry workflow to quantify secreted protein encompassing cytokines, growth factors, hormones, and other proteins with roles in intercellular communication [5]. Traditionally, secreted proteins are studied with antibody-based methods, such as enzyme-linked immunosorbent assays (ELISAs), focused on the detection of selected analytes. However, analyzing secretomes by mass spectrometry-based

Dominic De Nardo, Christine M. De Nardo (eds.), *Innate Immune Activation: Methods and Protocols*, Methods in Molecular Biology, vol. 1714, https://doi.org/10.1007/978-1-4939-7519-8_14,

proteomics enables the simultaneous identification and quantification of thousands of proteins. Furthermore, it allows a systems-level interpretation of secretory programs [6] and routes, as well as the discovery of proteins with unexpected extracellular functions or transcriptionally independent routes of cellular exit. Here, we describe an experimental and bioinformatics procedure for the comprehensive proteomics quantification of secreted proteins. Our workflow is sensitive, scalable, and compatible with the secretome analysis of primary cells. Correlation to antibody-based methods is high and quantitative differences up to four orders of magnitude are captured. Briefly, (1) secreted proteins are collected in vitro or ex vivo; (2) proteins are reduced and alkylated under denaturing conditions; (3) proteins are enzymatically digested; (4) peptides are separated by reversed-phase ultra-high pressure liquid chromatography (UHPLC) and analyzed with a high-resolution mass spectrometer; (5) label-free quantification and bioinformatics analysis is performed in the MaxQuant [7] and Perseus [8] environment.

2 Materials

Prepare all solutions using ultrapure water and analytical grade reagents (*see* **Note 1**). Prepare and store all reagents at room temperature (unless indicated otherwise).

2.1 Induction and Collection of Secreted Proteins

1. Cell culture medium: DMEM/RPMI without phenol red, 10% (v/v) heat-inactivated fetal bovine serum (FBS), 10 BmM HEPES, 1 mM pyruvate, 10 mM L-glutamine for primary dendritic cells, macrophages or monocytes (*see* **Note 2**).
2. Cell culture plates: 48-well, 12-well, or 6-well.
3. Water bath at 37 °C.
4. 18G needle.
5. Sterile filters: 0.22 µm.
6. 1.5 mL tubes.
7. Cell culture facility with incubator at 37 °C and 7% CO_2.

2.2 Sample Preparation for LC-MS/MS Analysis

1. Ice.
2. Octadecyl (C18) Solid Phase Extractions Disks (Empore).
3. Blunt-ended syringe.
4. Methanol (MeOH).
5. Modified Mini Centrifuge for stage tipping (Sonation).
6. Sonication device: we use Bioruptor Plus (Diagenode).
7. Sterile filters: 0.22 µm.

8. 1 M HEPES (pH 8.0): Weigh 11.9 g HEPES. Add water to a volume of 40 mL. Mix and adjust pH with NaOH. Add sterile water to a final volume of 50 mL and sterile filter.
9. 1 M NH_4HCO_3 (Ammoniumbicarbonat, ABC): Weigh 4.0 g NH_4HCO_3. Add sterile water to a final volume of 50 mL and sterile filter.
10. 4× Digestion buffer: 8 M Urea, 40 mM HEPES (pH 8.0). Always prepare fresh. Weigh 24.0 g Urea; Add 2 mL 1 M HEPES (pH 8.0) and make up to a volume of 40 mL with water. Mix and adjust to pH 8.0. Add sterile water to a final volume of 50 mL and sterile filter (*see* **Note 3**).
11. 1 M dithiotreitol (DTT): 1 M DTT, 50 mM ABC. Weigh 7.7 g DTT; add 2.5 mL 1 M ABC; Add sterile water to a final volume of 50 mL and sterile filter. Aliquot and store at −20 °C.
12. 0.55 M iodoacetamide (IAA): 0.55 M IAA, 50 mM ABC. Weigh 9.2 g IAA; add 2.5 mL 1 M ABC; Add sterile water to a final volume of 50 mL and sterile filter (0.22 µm). Aliquot and store at −20 °C.
13. 1 M Thiourea: 1 M Thiourea, 50 mM ABC. Weigh 3.8 g Thiourea; add 2.5 mL 1 M ABC; Add sterile water to a final volume of 50 mL and sterile filter. Aliquot and store at −20 °C.
14. 1 mM hydrochloric acid (HCl): Slowly add 4 µL HCl (37% w/w) to 12.5 mL water. Adjust the final volume to 50 mL with water.
15. Trypsin endopeptidase: 0.5 µg/µL Trypsin (Sigma) in 1 mM HCl. Dissolve 20 µg Trypsin in 40 µL 1 mM HCl. Store at −80 °C (*see* **Note 4**).
16. Lysyl endopeptidase (LysC): 0.5 µg/µL LysC (Wako) in 50 mM ABC. Prepare a 50 mM ABC solution by adding 50 µL 1 M ABC to 950 µL water. Dissolve 20 µg LysC in 40 µL 50 mM ABC. Store at −80 °C (*see* **Note 4**).
17. Stop solution: 6% (v/v), trifluoroacetic acid (TFA), 60% (v/v) acetonitrile. Slowly add 6 mL of TFA to 30 mL water. Add 60 mL acetonitrile. Adjust the final volume to 100 mL with water.
18. Buffer A: 0.5% (v/v) acetic acid. Slowly add 0.5 mL acetic acid to 90 mL water. Adjust the final volume to 100 mL with water.
19. Buffer B: 0.5% (v/v) acetic acid, 80% (v/v) acetonitrile. Slowly add 0.5 mL acetic acid to 15 mL water. Add 80 mL acetonitrile. Adjust the final volume to 100 mL with water.
20. Buffer A*: 0.5% (v/v) acetic acid, 0.1% (v/v) TFA, 2% (v/v) acetonitrile. Slowly add 0.5 mL acetic acid and 0.1 mL TFA to 90 mL water. Add 2 mL acetonitrile. Adjust the final volume to 100 mL with water.

21. Vacuum Concentrator: we use Vacuum Concentrator Plus (Eppendorf).

2.3 LC-MS/MS Analysis and Data Analysis

1. Chromatography column: 20–50 cm, 75 µm inner diameter, ReproSil-Pure C18-AQ 1.9 µm resin (Dr. Maisch GmbH).
2. Ultra-high pressure liquid chromatography (UHPLC), i.e., EASY-nLC 1000 ultra-high pressure system (Thermo Fisher Scientific).
3. A high-resolution mass spectrometer, i.e., Q Exactive, Q Exactive HF or HFX (Thermo Fisher Scientific).
4. A Nano-electrospray ion source, i.e., NanoFlex source (Thermo Fisher Scientific).
5. Column Oven (Sonation).
6. Running Buffer B: 0.1% (v/v) formic acid, 80% (v/v) acetonitrile. Slowly add 0.1 mL formic acid to 10 mL water. Add 80 mL acetonitrile. Adjust the final volume to 100 mL with water.
7. Download MaxQuant: http://www.coxdocs.org/doku.php?id=maxquant:common:download_and_installation
8. Download Perseus: www.coxdocs.org/doku.php?id=perseus:common:download_and_installation

3 Methods

Carry out all the procedures at room temperature unless specified otherwise.

3.1 Induction of Secreted Proteins

1. Start with cells in suspension (i.e., T cells) or adherent cells (i.e., bone marrow derived macrophages, BMDMs) already attached to the surface of tissue culture dishes. For considerations on the experimental setup, *see* **Notes 5** and **6**.
2. Wash the cells at least twice before commencing the experiment with pre-warmed serum-free cell culture medium without phenol red to dilute serum-proteins. Perform all the subsequent steps under serum-free conditions (*see* **Note 7**). We recommend a final cell density of 1×10^6 cells/mL (6 well plate: $2 \times 10,^6$ 12 well plate: $1 \times 10,^6$ 24 well plate: 0.4×10^6).
3. Activate the cells in medium without phenol red and incubate at 37 °C and 7% CO_2 for appropriate time-points (Fig. 1a, *see* **Notes 8** and **9**). For example, activate BMDMs with a Toll-like receptor (TLR) agonist such as the TLR4 ligand lipopolysaccharide (LPS) or leave them untreated as control for 6 h.

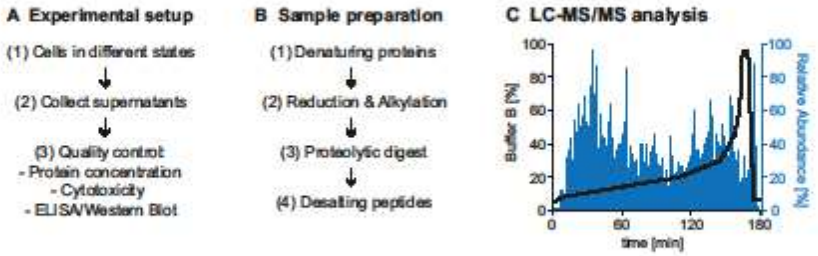


Fig. 1 Workflow for the LC-MS/MS-based analysis of secreted proteins: (**a**) Experimental workflow for collecting supernatants from activated cells including quality control for protein concentration, cell-viability, and cell responsiveness. (**b**) Sample preparation: proteins are reduced with DTT and alkylated with IAA under denaturing conditions before they are subjected to proteolytic digest with LysC and Trypsin and desalted on StageTips packed with C18 Solid Phase Extraction Disks. (**c**) Secretome LC-MS/MS analysis showing the gradient and relative abundance of the total ion current

4. At defined time-point(s) carefully remove supernatants with a syringe (<18G needle), and filter into 1.5 mL tubes (*see* **Note 10**).
5. Measure protein concentration with a method that does not interfere with the free amino acids of the medium (i.e., BCA-RAC assay according to the manufacturer's instructions). For downstream LC-MS/MS analysis, a minimum of 10 µg of protein in the supernatant is required.
6. Determine cell viability with a cytotoxicity assay (i.e., LDH release, trypan blue exclusion, etc.) and measure absolute amounts of selected proteins (i.e., cytokines) in supernatants with an antibody-based method (*see* **Notes 11** and **12**). Samples can be snap-frozen in liquid nitrogen and stored at −80 °C.

3.2 Secretome Digestion and Sample Preparation for LC-MS/MS Analysis

1. Denature proteins in the supernatants by adding 4× Digestion buffer to a final concentration of 2 M Urea, 10 mM HEPES (pH 8.0) per sample.
2. Sonicate the samples on ice for 15 min at maximum intensity [9–11] (*see* **Note 13**) (Fig. 1b).
3. Reduce disulfide bonds by adding 1 M DTT at a final concentration of 10 mM to the samples. Mix well by vortexing and incubate for 40 min at room temperature (*see* **Note 14**).
4. Alkylate reduced cysteines by adding 0.55 M IAA to a final concentration of 55 mM (*see* **Note 14**). Mix well and incubate for 40 min in the dark. Samples are incubated in the dark because iodoacetamide is light-sensitive. Excess iodoacetamide can be quenched by adding 1 M Thiourea to a final concentration of 100 mM.

5. Digest proteins by adding 0.5 µg LysC for 3 h and 0.5 µg Trypsin for at least 16 h. Inhibit proteases by acidifying the samples with Stop solution. (final concentration: 0.6% TFA, 6% acetonitrile) (*see* **Note 15**).
6. Prepare StageTips [9]: in short, stamp out a small disk from three layers of C18 Solid Phase Extraction Disks with a blunt-ended syringe and plunge into a 200 µL pipette tip. Equilibrate the StageTip by washing successively with 50 µL 100% MeOH, 50 µL Buffer B, and 100 µL Buffer A. In-between each wash step, centrifuge StageTips with a Modified Mini Centrifuge for 2 min at maximum speed, or until all fluid has been drained.
7. Desalt peptides by loading the entire digested and acidified supernatants onto equilibrated C18 StageTips (*see* **Note 16**).
8. Wash peptides with 200 µL Buffer A.
9. Centrifuge StageTips with a Sonation StageTip Centrifuge for 2 min at maximum speed.
10. Repeat **steps 8** and **9** of Subheading 2.3.
11. Elute peptides into PCR-tubes using 2 × 20 µL Buffer B.
12. Evaporate Buffer B in a pressure device for 30 min at 40 °C vacuum.
13. Resuspend peptides in 10 µL Buffer A*.

3.3 LC-MS/MS Analysis of Secretomes

1. Perform reverse-phase chromatography on a UHPLC coupled to a high-resolution mass spectrometer with a nano-electrospray ion source (Fig. 1c).
2. Load the peptides (1–2 µg) from the supernatants onto a C18-reversed phase column (*see* **Note 17**).
3. Separate the peptides with an almost linear gradient of 7–30% (v/v) Running Buffer B between 120 and 180 min (Fig. 1c) at a flow rate of 200–300 nL/min on 20–50 cm columns with a Column Oven temperature of 50–55 °C.
4. Perform LC-MS/MS analysis for an UHPLC coupled to a high-resolution mass spectrometer, i.e., QExactive HF as described by Scheltema et al. [12]: Acquire MS data with a data-dependent TOPX method to dynamically select the 10–15 most abundant precursor ions from the survey scan. Set the target value to 3×10^6 ions in a 200–1650 m/z range with 20 ms maximum injection time at a resolution of 70,000 at m/z 400. Isolate precursor ions with a 2.4 Th isolation window. Fragment precursor ions by higher-energy C-trap dissociation (HCD) with normalized collision energies of 27 eV. Acquire MS/MS scans with a target value of 1×10^6 and a maximum injection time of up to 100 ms, depending on

the overall peptide amount and complexity; however, duty cycle times should not exceed 1.5 s. Exclude unassigned precursor ion charge states and singly charged ions. To avoid repeated sequencing, exclude already selected sequenced ions dynamically for 20–30 s depending on the chromatography.

3.4 Analysis of Mass Spectra Using MaxQuant

1. Analyze raw data with the MaxQuant software searching peak lists with the Andromeda search engine [13] against the latest Uniprot FASTA database of the species used in your experiments combined with common contaminants (*see* **Note 18**).
2. Use standard MaxQuant settings [7]: in brief, set the maximum allowed mass deviation to 6 ppm for monoisotopic precursor ions and 0.5 Da for MS/MS peaks. Use Trypsin as the digestion mode (or change accordingly in case you used other proteases) and allow for a maximum amount of two missed cleavages. Set carbamidomethyl cysteine as a fixed modification and N-terminal acetylation, asparagine deamidation and methionine oxidation as variable modifications. False discovery rate (FDR) should be at 1% on both the peptide and protein levels.
3. Quantify proteins in MaxQuant using the built in XIC-based label-free quantification (LFQ) algorithm using fast LFQ [7, 14, 15].
4. Enable "match between runs" to transfer identifications between individual LC-MS runs.

3.5 Analysis of MaxQuant Output with Perseus

1. Analyze MaxQuant output files using standard settings in Perseus [8]. In short, upload the proteinGroups.txt file from the MaxQuant output folder (.../combined/txt/proteinGroups.txt) into Perseus. Select the LFQ intensity columns of your experiments as main columns (*see* **Note 19**).
2. Exclude contaminants, proteins identified by searching the reverse database and proteins only identified by a modified peptide by filtering rows based on categorical column (Filter rows → Filter rows based on categorical column).
3. Logarithmically transform dataset to obtain normally distributed LFQ intensities (Basic → Transform → *log2(x)*) (*see* **Note 20**).
4. Assign replicates of one experiment (i.e., treated, control, etc.) to one group by appending identical group names (Annotation rows → Categorical annotation rows) and filter out proteins not consistently identified and quantified (Filter rows → Filter rows based on valid values → Min. number of values: 2–3, Mode: In at least one group).

5. Replace missing values (i.e., NaN) with values from a normal distribution generated based on the measured data(Imputation → Replace missing values from a normal distribution) (*see* **Note 21**).
6. Check reproducibility between experiments by calculating the Pearson correlation (Basic → Column correlation). Visualize the correlation matrix with hierarchical clustering (Clustering/PCA → Hierarchical clustering) (*see* **Note 22**).
7. Perform a principal component analysis (Clustering/PCA → Principal component analysis) to evaluate biological versus experimental/technical variation.
8. Identify significant differences between secretomes from activated and non-activated cells or between different conditions, cellular states, genotypes, etc., by performing a statistical hypothesis test, i.e., unpaired T-test or ANOVA test with a permutation based FDR (*see* **Note 23**).
9. Visualize results by plotting *−log10 p-value* against *LFQ intensity fold-change* (Misc./Volcano plot/First group: activated samples, Second group: non-activated samples). For treated samples, Protein X (i.e., TNF upon LPS treatment) appears with a high enrichment fold change and low p-value on the upper right side of the plot (*see* **Notes 24** and **25**) (Fig. 2a).
10. To visualize significantly secreted proteins only, filter for proteins with significant differences between selected states (in the Volcano plot tab: export selection-reduce matrix) and plot them in a heat map (Clustering/PCA → Hierarchical clustering) (Fig. 2b).
11. Protein secretion data is often presented as fold-change secretion upon activation or fold-difference between states. For this purpose, data are normalized (i.e., subtract logarithmized LFQ intensities of non-activated from activated states, or calculate the difference between LFQ intensities of selected states). To control your secretome workflow, compare fold-secretion by LC-MS/MS calculated upon activation to absolute analyte amounts determined by antibody-based methods with a correlation analysis (Fig. 2c).
12. To visualize the LFQ intensity profiles of individual selected proteins across various biological conditions, as well as other proteins which display similar profiles and might behave similarly, select a reference protein, i.e., TNF, which is released in monocytes upon LPS treatment (Visualization → Profile plot → Reference profile → Euclidean → From selected profile: TNF → Refresh). Returning to the Profiles section sort the proteins by Euclidean distance to the selected reference protein. Proteins with distances closest to the selected protein show similar profiles (Fig. 2d).

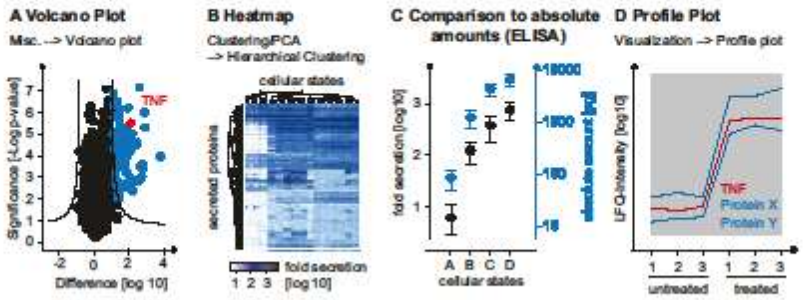


Fig. 2 Examples for data analysis of secretome experiments: (**a**) Volcano plot based on t-Test statistics showing the significance and fold-change difference of all proteins between samples. Significant outliers from t-Test statistics are controlled by a permutation-based FDR. (**b**) Perturbation matrix of multiple cellular states visualized by hierarchical clustering of differentially secreted proteins in a heatmap. (**c**) Comparison of fold secretion changes and absolute amounts of a selected protein (such as TNF) measured by ELISA to evaluate performance of secretome measurements. (**d**) Profile correlation analysis in Perseus to visualize different LFQ-intensity profiles over various conditions

4 Notes

1. Ultrapure water is prepared by purifying deionized water, to attain a resistivity of 18 MΩcm at 25 °C. Furthermore, all materials and reagents used for experiments should be free of polyethylene glycols or detergents as they interfere with downstream LC-MS/MS analysis.
2. Phenol red has a mass of 354.38 Da and interferes with LC-MS/MS analysis by generating a charged peak in the chromatogram.
3. An alkaline pH (pH 8–9) ensures the selective alkylation of cysteines.
4. Repeated freeze-thaw cycles reduce the proteolytic activity of Trypsin and LysC. Freeze-thaw one aliquot a maximum of five times.
5. Proteomic secretome experiments aim at determining quantitative differences between samples. Therefore, suitable controls for comparisons should be included, i.e., basal secretion from non-activated cells to evaluate activation-induced secretion. Depending on the biological question, comparisons of different agonists, cellular differentiation or activation states, genotypes or cell types may be valuable.
6. Perform all the experiments in at least triplicates to facilitate statistical analysis. Biological replicates instead of technical

replicates should be considered to capture experimental variances sufficiently.

7. The LC-MS/MS workflow requires serum-free conditions because high abundant serum proteins in supernatants result in low identification rates for secreted proteins due to dynamic range limitations: i.e., high abundant peptides from the serum hamper the sequencing of lower abundant peptides from secreted proteins. Serum-free conditions may affect cell-viability and responsiveness and we recommend prior assessment. For example, the cellular response upon activation in serum-free conditions can be evaluated with antibody-based methods in comparison to standard conditions containing serum.

   For cellular activation with proteins or ligands that require albumin or serum for function, pulse cells with the agonist in conditions optimal for activation. Before secretion is induced efficiently, i.e., 1–2 h upon activation of TLRs in bone marrow-derived dendritic cells (BMDCs) or macrophages (BMDMs)—wash the cells and replace activation medium with serum-free medium without phenol red as described.

8. To minimize basal protein secretion as well as intra- and intercellular feedback loops in cellular signaling cascades, selection of early time points upon activation may be advantageous.

9. The secretome workflow is compatible with different experimental setups. In our protocol, supernatants are collected in cell culture dishes; however, the cells can be differentiated or activated in vitro as well as in vivo. Secretomes from primary cells stimulated in vitro as well as from FACS/MACS-sorted cells directly isolated from humans or mice with defined treatments or pathologies may be particularly valuable.

   If the cells are activated in vivo, isolate cells and collect secreted proteins ex vivo in serum-free medium.

10. For large scale or tight time course experiments, removing detached cells and cellular debris by centrifugation (5 min at 350 × *g*) is an option.

11. Cell viability may be affected by serum-free conditions. Therefore, we recommend performing cytotoxicity assays such as the LDH assay or trypan blue exclusion assay. If cell viability is less than 90% in serum-free conditions, consider working in the presence of 0.1–1 µg/mL albumin or serum or a pulse strategy as outlined in **Note 9**. In case these options do not increase cell viability, protein-free medium may be worth considering (such as FreeStyle 293 Expression Medium from Thermo Fisher Scientific), although these media may require an additional acetone precipitation step.

12. Measure the absolute amount of selected secreted proteins, i.e., cytokines, with antibody-based methods. If absolute amounts of control cytokines such as TNF or IL6 are below 10–100 pg, increase cell numbers to adjust to the detection limit for LC-MS/MS instrumentation. As a reference, a few hundred thousand BMDCs and BMDMs activated with 100 ng/mL LPS for 6 h secrete sufficient amounts of proteins for a MS-based secretome analysis.
13. Urea is labile at higher temperatures and resulting cyanates can lead to protein modification (i.e., carbamylation of proteins). Make sure sonification is carried out on ice or while a cooling pump is running and perform this step in the cold room.
14. Reduction and alkylation is an important sample preparation step for LC-MS/MS analysis. Low alkylation efficiency of cysteines leads to decreased identification rates, because unmodified cysteines are oxidized in gas phase spontaneously or are not identified with standard search strategies.
15. To increase sequence coverage of proteins (i.e., number of peptides per protein) consider using other specific proteases (i.e., Chymotrypsin, GluC, etc.).
16. One C18 layer is able to bind 2–4 µg of digested proteins [16]. Do not load more than 15 µg of digested proteins per StageTip (three layers of C18 solid phase extractions disks) to avoid overloading and potential loss of more hydrophilic peptides.
17. Peptides should elute uniformly across the complete gradient and reach a total ion current (NL) of up to 2E10 on a QExactive HF (Fig. 2c).
18. MaxQuant includes a list of common contaminants in the database search, i.e., BSA, Trypsin, human collagens, etc. Thereby, incorrect peptide assignment (i.e., type I errors) from contaminating peptides is prevented [17].
19. Prior to Perseus analysis, open the summary.txt file from the MaxQuant output folder, i.e., in Excel: check if the number of peptides (in the column Peptides Sequences Identified) is comparable across samples. Also, the percentage of identified MS/MS should be above 30% in secretome experiments (Column: MS/MS Identified [%]).
20. A normally distributed dataset is a prerequisite for many statistical hypothesis tests. If your dataset is not normally distributed, check for problems with the biological experiment, sample preparation as well as LC-MS/MS system.
21. Missing values occur if proteins are of low abundance and thereby only stochastically sequenced by the mass spectrometer.

Replacing missing values with a fixed value (i.e., 0) leads to a bimodally distributed dataset (not compatible with many statistical tests). In Perseus you can, i.e., replace missing values by inserting random values from a Gaussian distribution centered around a downshifted median (imputation from normal distribution) [18].

22. A Pearson correlation >0.85 is expected for replicates of secretome measurements. For bioinformatics sample normalization in the MaxQuant environment, the LFQ quantification algorithm assumes comparable raw intensities (i.e., high reproducibility) for a large fraction of peptides. In case this criterion is not met (i.e., when comparing two conditions which differ significantly in cellular viability), protein quantification becomes inaccurate. Therefore, we also recommend confirming selected LC-MS/MS results, i.e., with antibody-based methods [5].
23. Proteomic datasets consist of thousands of measured data points leading to a multiple hypothesis testing problem when performing two-sample tests. A permutation-based false discovery rate (FDR) can be used to estimate the number of false positive significant hits [19]. Standard settings in Perseus using a FDR of 5% with S0 correction are suitable for a statistical analysis of secretome experiments.
24. Make sure that calculations for fold-difference of secretion are based on measured LFQ data and not imputed data.
25. Highly abundant intracellular proteins represent potential contaminants in cellular supernatants since their release may be a result of cell death. Correlation analysis of secretome to total protein expression data can indicate whether certain proteins might be present in secretomes due to cellular leakage.

References

1. Mann M, Nils a K, Nagaraj N et al (2013) The coming age of complete, accurate, and ubiquitous proteomes. Mol Cell 49:583–590
2. Larance M, Lamond AI (2015) Multidimensional proteomics for cell biology. Nat Rev Mol Cell Biol 16:269–280
3. Aebersold R, Mann M (2016) Mass-spectrometric exploration of proteome structure and function. Nature 537:347–355
4. Altelaar AFM, Munoz J, Heck AJR (2013) Next-generation proteomics: towards an integrative view of proteome dynamics. Nat Rev Genet 14:35–48
5. Meissner F, Scheltema RA, Mollenkopf H-J et al (2013) Direct proteomic quantification of the secretome of activated immune cells. Science 340:475–478
6. Rieckmann JC, Geiger R, Hornburg D, Wolf T, Kveler K, Jarrossay D, Sallusto F, Shen-Orr SS, Lanzavecchia A, Mann M, Meissner F (2017) Social network architecture of human immune cells unveiled by quantitative proteomics. Nat Immunol 18(5):583–593
7. Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.P.B.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26:1367–1372
8. Tyanova S, Temu T, Sinitcyn P et al (2016) The Perseus computational platform for com-

prehensive analysis of (prote)omics data. Nat Methods 13(9):731–740

9. Rappsilber J, Mann M, Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat Protoc 2:1896–1906
10. Wisniewski JR, Zougman A, Mann M (2009) Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. J Proteome Res 8:5674–5678
11. Wisniewski JR, Zougman A, Nagaraj N et al (2009) Universal sample preparation method for proteome analysis. Nat Methods 6:359–362
12. Scheltema RA, Hauschild J-P, Lange O et al (2014) The Q exactive HF, a benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field orbitrap analyzer. Mol Cell Proteomics 13:3698–3708
13. Cox J, Neuhauser N, Michalski A et al (2011) Andromeda: a peptide search engine integrated into the maxquant environment. J Proteome Res 10:1794–1805
14. Luber CA, Cox J, Lauterbach H et al (2010) Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. Immunity 32:279–289
15. Cox J, Hein MY, Luber CA et al (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics 13:2513–2526
16. Rappsilber J, Ishihama Y, Mann M (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal Chem 75:663–670
17. Griss J, Perez-Riverol Y, Lewis S et al (2016) Recognizing millions of consistently unidentified spectra across hundreds of shotgun proteomics datasets. Nat Methods 13:651–656
18. Liew AW-C, Law N-F, Yan H (2011) Missing value imputation for gene expression data: computational techniques to recover missing data from available information. Brief Bioinform 12:498–513
19. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci 98:5116–5121

4.4 Project 4: Quantitative and Dynamic Catalogs of Proteins Released during Apoptotic and Necroptotic Cell Death

At the level of individual effectors, TNF-induced cell death has been intensively explored. Numerous studies utilizing genetic mice models established the critical role of TNF-dependent apoptosis and necroptosis in inflammatory illness models. While necroptosis has long been regarded as a more inflammatory cell death route than apoptosis, few proteins have been identified to be secreted particularly by necroptotic cells and to be potentially harmful. Additionally, past research on the release of cytokines during necroptosis has yielded inconsistent results. Other recent publications in the field focused on the production of extracellular vesicles as a result of endosomal exocytosis during necroptosis, but especially on its effect on MLKL, the necroptosis terminal executor. Our study is the first to conduct a complete and comparative proteomic investigation of proteins secreted by cells undergoing apoptosis and necroptosis triggered by TNF.

Firstly, we achieved an unbiased perspective of the release of cytokines and other known immunogenic proteins following cell death using our systems-based quantitative approach. Secondly, proteins were identified released as a result of cell death type-specific pathways. Two mechanisms that are active during TNF-induced cell death were examined: (1) Shedding is occurring during necroptosis and involves the activation of metalloproteases called ADAMs, which then cleave receptors on the cell surface, resulting in the enrichment of receptor extracellular domains in supernatants. We present a systems-level picture of this mechanism and demonstrate that shedding occurs late in both TNF-induced necroptosis and apoptosis. (2) We detected an excess of lysosomal proteins in the supernatant of necroptotic cells prior to complete membrane permeabilization. Interestingly, we discovered mostly luminal lysosomal components in entire supernatants, but necroptotic cells' external vesicles included predominantly lysosomal membrane proteins. This observation, together with the detection of lysosomes at the plasma membrane, led us to conclude that necroptosis initiates lysosomal exocytosis.

To summarize, our study provides comprehensive, quantitative, and dynamic catalogs of proteins secreted during apoptotic and necroptotic cell death. This provides a foundation for examining the intricacy of biological processes that occur as a result of or are governed by these forms of cell death. Our findings indicate that a surprising variety of distinct proteins are secreted during cell death. Only a few of these directly contribute to inflammation, while others are more likely to work in more complex pathways. The project was a group internal collaboration with Maria Tanzer.

The corresponding article was published in Cell Reports in 2020 [326].

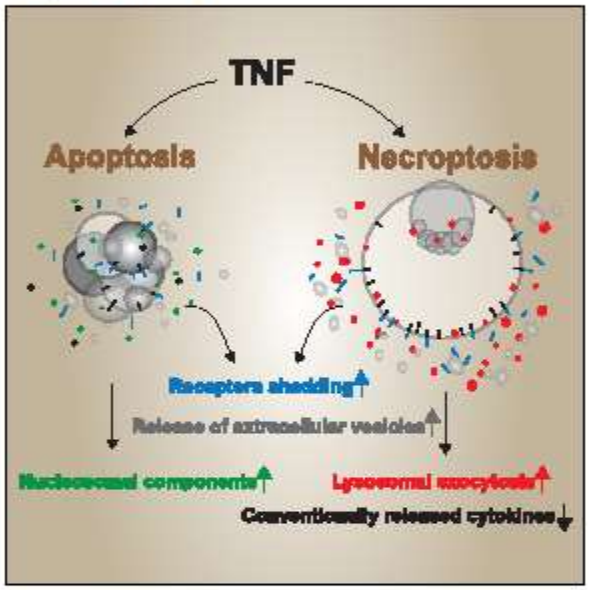
I have contributed the shedding analysis providing a tailored bioinformatics workflow in the R environment.

Cell Reports

Resource

Quantitative and Dynamic Catalogs of Proteins Released during Apoptotic and Necroptotic Cell Death

Graphical Abstract



Authors

Maria C. Tanzer, Annika Frauenstein, Che A. Stafford, Kshiti Phulphagar, Matthias Mann, Felix Meissner

Correspondence

mmann@biochem.mpg.de (M.M.), meissner@biochem.mpg.de (F.M.)

In Brief

Tanzer et al. provide a view of dynamically released proteins during TNF-induced apoptosis and necroptosis. The data highlight processes that are commonly and differentially regulated during these two different types of cell death.

Highlights

- The release of conventionally secreted cytokines is reduced in necroptotic cells
- Receptor shedding occurs during TNF-induced necroptosis and apoptosis
- Lysosomal components are released via lysosomal exocytosis during necroptosis

Tanzer et al., 2020, Cell Reports 30, 1260–1270
January 28, 2020 © 2019 The Authors.
https://doi.org/10.1016/j.celrep.2019.12.079

OPEN ACCESS
CellPress

Cell Reports
Resource

Quantitative and Dynamic Catalogs of Proteins Released during Apoptotic and Necroptotic Cell Death

Maria C. Tanzer,[1] Annika Frauenstein,[2] Che A. Stafford,[3] Kshiti Phulphagar,[2] Matthias Mann,[1,*] and Felix Meissner[2,4,*]
[1]Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany
[2]Experimental Systems Immunology, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany
[3]Gene Center and Department of Biochemistry, Ludwig-Maximilians-Universität, 81377 Munich, Germany
[4]Lead Contact
*Correspondence: mmann@biochem.mpg.de (M.M.), meissner@biochem.mpg.de (F.M.)
https://doi.org/10.1016/j.celrep.2019.12.079

SUMMARY

The inflammatory functions of the cytokine tumor necrosis factor (TNF) rely on its ability to induce cytokine production and to induce cell death. Caspase-dependent and caspase-independent pathways—apoptosis and necroptosis, respectively—regulate immunogenicity by the release of distinct sets of cellular proteins. To obtain an unbiased, systems-level understanding of this important process, we here applied mass spectrometry-based proteomics to dissect protein release during apoptosis and necroptosis. We report hundreds of proteins released from human myeloid cells in time course experiments. Both cell death types induce receptor shedding, but only apoptotic cells released nucleosome components. Conversely, necroptotic cells release lysosomal components by activating lysosomal exocytosis at early stages of necroptosis-induced membrane permeabilization and show reduced release of conventionally secreted cytokines.

INTRODUCTION

Tumor necrosis factor (TNF) is a major contributor to many inflammatory diseases, including psoriasis, rheumatoid arthritis, and inflammatory bowel disease (Bradley, 2008). This has initially been linked to TNF-induced production of various cytokines via nuclear factor κB (NF-κB) and mitogen-activated protein kinase (MAPK) pathway activation (Bradley, 2008). Subsequent reports, however, revealed a major role of TNF-induced cell death in inflammatory diseases (Rock and Kono, 2008). Binding of TNF to its receptor TNF-R1 activates the NF-κB and MAPK pathways, leading to cytokine production. Cellular inhibitor of apoptosis 1 and 2 (cIAP1 and cIAP2) regulate the activation of these pathways via their E3 ubiquitin ligase activity toward receptor-interacting serine/threonine-protein kinase 1 (RIPK1) (Bertrand et al., 2008). Inhibition of cIAPs using small molecules called Smac mimetics (SMs) results in decreased NF-κB signaling and caspase-8 activation (Arslan and Scheidereit, 2011; Vince et al., 2007; Varfolomeev et al., 2007). This triggers the activation of downstream caspases and initiates the apoptotic process, which involves nuclear fragmentation, blebbing, and cell shrinkage (Elmore, 2007). During the early stages of apoptosis, phosphatidylserine is exposed on the cell surface, which is recognized by macrophages that engulf dying cells (Fadok et al., 1992). Although released histones have been described to activate Toll-like receptors and induce cytokine production, apoptosis is conceptually considered an immunological silent death (Elliott and Ravichandran, 2010).

Caspase inhibition prevents apoptosis, but in many cell types it triggers an alternative cell death pathway termed necroptosis (Vercammen et al., 1998). Upon TNF signaling, caspase-8 regulates necroptosis by cleaving RIPK1 (Oberst et al., 2011; Pop et al., 2011). Caspase-8 inhibition leads to autophosphorylation and activation of RIPK1 and RIPK3. Active RIPK3 phosphorylates and activates the pseudokinase mixed lineage kinase domain-like protein (MLKL) (Murphy et al., 2013; Wang et al., 2014; Tanzer et al., 2017). Active MLKL translocates to the plasma membrane to induce membrane rupture (Hildebrand et al., 2014; Sun et al., 2012). Necroptosis is primarily thought of as an immunologically reactive process, because of its fast kinetics, the release of damage-associated molecular patterns (DAMPs) such as high mobility group box 1 protein (HMGB1) and ATP, as well as reduced macrophage engulfment (Kaczmarek et al., 2013).

Different modes of cell death can lead to significantly different physiological outcomes through the release of distinct molecules. For example, it was recently reported that necroptotic cell death in contrast to apoptotic cell death can drive a systemic immune response, leading to tumor regression (Snyder et al., 2019). The same group showed that cytokine mRNA translation continues at the endoplasmic reticulum (ER) during necroptosis even after plasma membrane rupture (Orozco et al., 2019). Another study reported that the cell death-dependent cytokine microenvironment determines the lineage commitment and thereby the harmfulness of hepatic cancer (Seehawer et al., 2018). Other studies compared the ability of cytokine and chemokine release by apoptotic and necroptotic cells (Kearney et al., 2015; Zhu et al., 2018). However, whereas Kearney et al. (2015) observed a decrease in cytokines released by necroptotic cells, Zhu et al. (2018) reported an activation of pro-inflammatory

cytokine gene expression instead. Together these studies underline the different properties of apoptotic and necroptotic cells on their surrounding microenvironment, which would now make it attractive to systematically investigate the released cellular contents.

Mass spectrometry-based proteomics has matured remarkably in recent years and now provides a comprehensive discovery tool to address diverse biological questions (Aebersold and Mann, 2016; Larance and Lamond, 2015). Here we set out to identify proteins released by apoptotic versus necroptotic cell death programs using a proteomics strategy we developed recently for the sensitive detection of secreted proteins (Meissner et al., 2013; Frauenstein and Meissner, 2018).

We induced TNF-mediated apoptosis and necroptosis in a lymphoma cell line and human primary macrophages. Proteomic analysis of supernatants and of enriched extracellular vesicles shed light on processes regulated during necroptosis and apoptosis and the cell death type-specific release of cytokines.

RESULTS

Differential Release of Proteins in Cells Undergoing TNF-Induced Apoptosis and Necroptosis

To define the inflammatory potential of apoptosis and necroptosis on a global scale, we set out to identify proteins released by cells undergoing TNF-mediated apoptosis or necroptosis. We induced both forms of cell death in the histiocytic lymphoma cell line U937 and in human primary macrophages, two cell types frequently used to investigate these cell death pathways (Tanzer et al., 2015; McComb et al., 2012) (Figure 1A). Apoptosis was triggered by TNF and the SM birinapant, while the further addition of the caspase inhibitor IDUN-6556 (IDN-6556) led to necroptosis. Both processes showed similar kinetics when comparing caspase cleavage (Figure 1B), a hallmark of apoptosis, with phosphorylation of MLKL (Figure 1C), a hallmark of necroptosis. During necroptosis, activated MLKL translocates to the plasma membrane and leads to immediate membrane permeabilization and propidium iodide intake (Hildebrand et al., 2014), whereas apoptotic cells retain plasma membrane integrity until the later stages when secondary necrosis occurs (Figure 1D; Figure S1A). A recent report showed that active caspase-8 can cleave gasdermin-D, a pore-forming protein that induces plasma membrane rupture during pyroptosis (Orning et al., 2018). We cannot exclude processing of gasdermin-D at later time points of apoptosis. Nevertheless, the slow kinetics and reduced level of propidium iodide intake (Figure 1D) and the delayed release of HMGB1 (Figures S2B and S2C) (a prominent marker of necroptosis and pyroptosis; Kaczmarek et al., 2013) are strong indications against a potential involvement of pyroptosis upon TNF and SM stimulation, at least before secondary necrosis.

To analyze proteins released by apoptotic and necroptotic cells, supernatants were harvested and digested, and the resulting peptides were directly analyzed by single-run liquid chromatography mass spectrometry (LC-MS/MS) coupled to high-resolution mass spectrometry on the quadrupole Orbitrap analyzer (STAR Methods; Figure 1A). Label-free quantification of the MS data and statistical analysis were performed using MaxQuant and Perseus (Cox and Mann, 2008; Tyanova et al., 2016). Overall, 3,507 proteins were identified from U937 supernatants, with an average of 2,058 protein groups per sample (Figure 1E). Principal-component analysis revealed partial stimulation and time-dependent separation of samples (Figures S1B and S1C). Necroptotic supernatants from later time points, when membrane permeabilization occurred, separated from the rest of the samples (Figures S1B and S1C). Consistently, the levels of hundreds of proteins were significantly changed in the supernatants of necroptotic cells, compared with control samples (Figure 1F). In contrast, this was the case for only 34 proteins in apoptosis. However, at a later time point apoptotic cells, which undergo secondary necrosis, also significantly released several hundreds of proteins (Figure S1D). The fold changes of all proteins measured across seven apoptotic and four necroptotic time points compared with respective controls are provided in Table S1. When we analyzed the proteome of the dying cells, the most apparent variation in protein levels occurred in necroptotic cells, including a reduction in cytosolic proteins, and an increase in mitochondrial proteins and proteins located at the ER (Figure S1E).

The Release of Conventionally Secreted Cytokines Is Significantly Reduced in Necroptotic Cells

TNF induces the production of a wide range of cytokines required to fight infection, and we were particularly interested in the effect of TNF-induced cell death on their release. CCL2 and IL8 were most strongly induced through TNF stimulation in U937 cells (Figures 2A and 2B), looking at all cytokines (retrieved from the keyword annotation "Cytokines"; STAR Methods). At later time points, levels of CCL2 and IL8 were diminished in supernatants of apoptotic cells, but this was not statistically significant. However, we observed a significant downregulation for both cytokines during necroptosis compared with TNF-only treatment. To confirm our mass spectrometry results, we used enzyme-linked immunosorbent assay (ELISA) to measure levels of CCL2 and HMGB1 in supernatants, which showed a similar trend (Figures S2A–S2C). CCL2 and IL8 contain an N-terminal signal peptide and are therefore part of the keyword annotation "Signal" (hereafter referred to as signal proteins). The signal peptide mediates conventional secretion through the ER-Golgi secretory pathway. Other conventionally secreted cytokines were also downregulated at later stages of necroptosis in U937 cells (Figure S2D) and primary macrophages (Figures 2C and 2D; Figure S2E). Levels of most non-conventionally released cytokines were primarily unaffected or increased in necroptotic supernatants because of membrane permeabilization (Figure S2F). AIMP-1 levels, for example, were enriched in a minor but significant way in necroptotic supernatants (Figure S2F).

We examined the release of all conventionally released proteins to test whether the secretion of the ER-Golgi network is generally affected by necroptosis. Surprisingly, we detected an increased release of signal proteins at 3 h of necroptosis treatment (Figure 2E; Figure S2G), while intracellular levels of

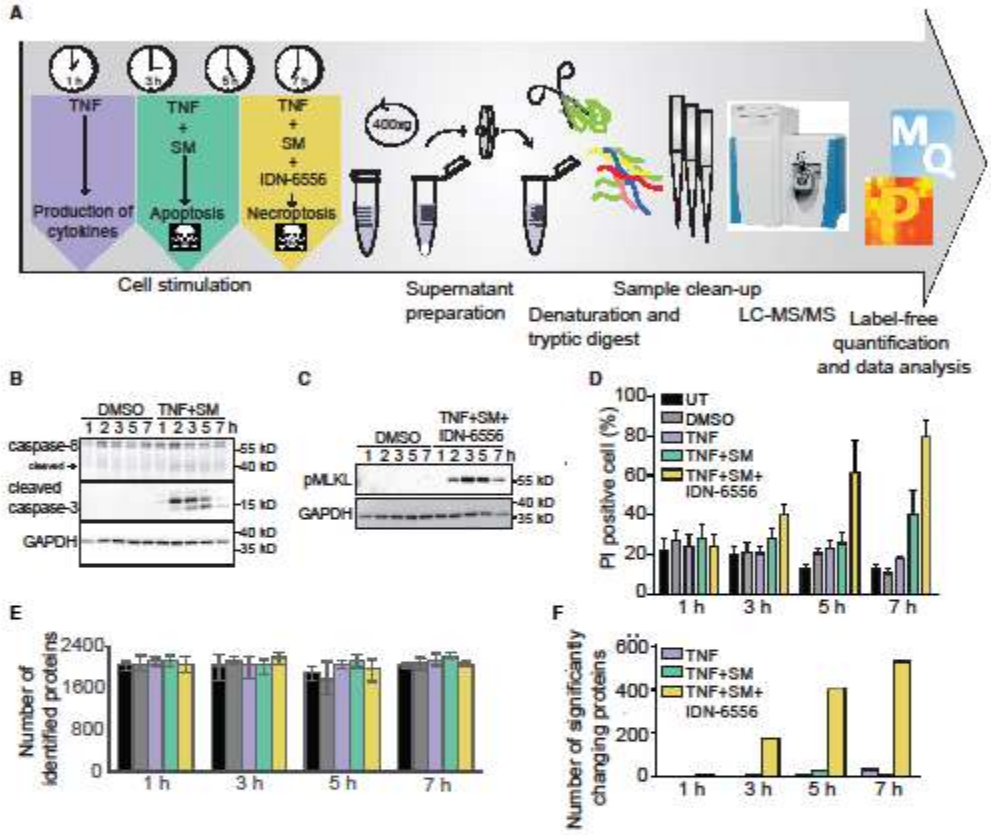


Figure 1. Supernatant Analysis of Cells Undergoing TNF-Induced Apoptosis and Necroptosis

(A) Schematic workflow of supernatant processing.

(B and C) Immunoblot of U937 cells treated with DMSO, TNF (30 ng/mL), and SM (birinapant, 250 nM) to induce apoptosis or with TNF, SM, and IDN-6556 (IDN-6556, 10 μM) to induce necroptosis for 1–7 h and blotted for the apoptotic markers cleaved caspase-8 and caspase-3 (B) and the necroptotic marker phosphorylated MLKL (C) and the loading control GAPDH.

(D) Cell death analysis by flow cytometry of propidium iodide-positive cells, which were treated as indicated over a time course of 1–7 h (±SEM, n = 3 or 4).

(E) Number of all proteins quantified in the supernatant of cells treated as indicated in (D).

(F) Number of significantly changing proteins in supernatants of cells treated as indicated (Student's t test: false discovery rate [FDR] = 0.05).

the same proteins were unaffected (Figure S2H). Proteins that are unconventionally secreted (assigned to the keyword annotation "Secreted", excluding signal proteins) were not changed in the supernatant of necroptotic cells (Figure 2E). The release of conventionally released proteins at 3 h of stimulation was most strongly inhibited in necroptotic cells by brefeldin A treatment (Figure 2F), which leads to disruption of the Golgi apparatus. This indicates that the ER-Golgi secretory pathway is activated at early stages of necroptosis.

However, at 5 (Figures S2I and S2J) and 7 h of stimulation (Figure 2G) we detected fewer signal proteins in supernatants of necroptotic cells compared with the control treated cells, while intracellular levels were again unaffected (Figure S2K). This corresponds to the missing impact of brefeldin A on the signal protein level in supernatants of necroptotic cells treated for 7 h (Figure 2H). To determine whether degradation of signal proteins occurs because of an increased release of proteases at later stages of necroptosis, we analyzed semi-tryptic peptides of those proteins. Semi-tryptic peptides are generated by proteolytic processing prior to the proteomic workflow. In necroptotic supernatants at 7 h of necroptosis induction semi-tryptic peptides of signal proteins are slightly

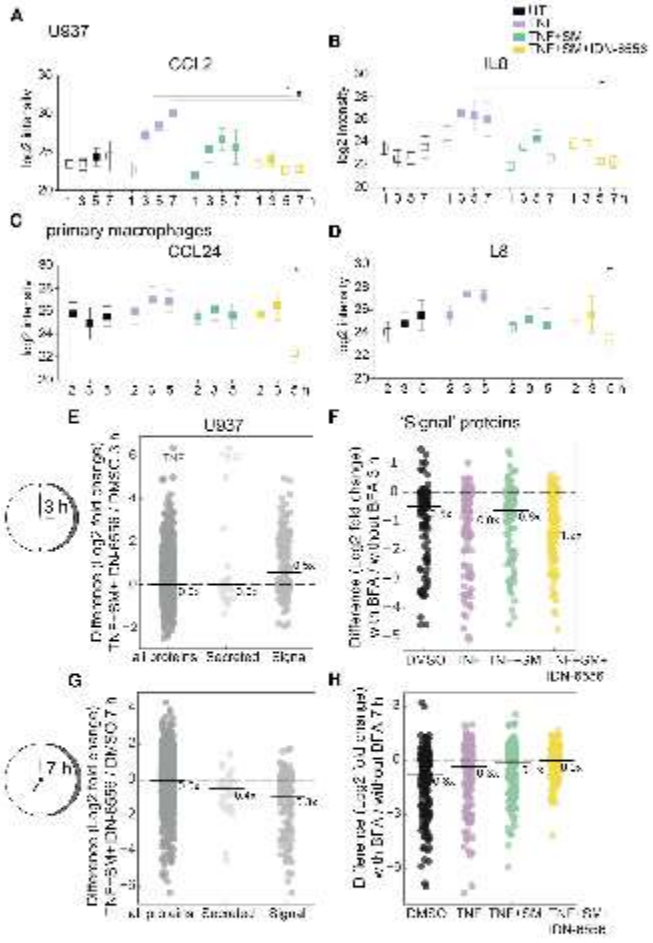


Figure 2. The Release of Conventionally Secreted Cytokines Is Reduced in Necroptotic Cells

(A–D) Graphs show the $\log_2$ intensity of CCL2 and IL8 in supernatants of U937 cells (A and B) and CCL24 and IL8 in supernatants of primary macrophages (C and D) measured using mass spectrometry. Cells were either left untreated (black squares) or treated with TNF (purple squares), TNF and SM (green squares), or TNF, SM, and IDN-6556 (yellow squares) over a time course of 1–7 h (for U937 cells) and 2–6 h (for primary macrophages) (±SEM, n = 3 or 4). Full squares represent two or more valid values for the data point. Asterisks show significant changes (Student's t test; FDR = 0.05).

(E and G) Swarm plot showing $\log_2$ fold change of all proteins (dark gray) and proteins assigned to the keyword annotation "Secreted" (all proteins annotated to the keyword annotation "Signal" were excluded, light gray, 1D annotation, not significant [n.s.], E, $p = 1.8 \times 10^{-4}$, G) and "Signal" (gray, $p = 2.1 \times 10^{-4}$, E, $p = 2.8 \times 10^{-21}$, G) released by necroptotic cells compared with the DMSO control at 3 h (E) and 7 h (G) of treatment. Taking the antilogarithm, we obtain fold changes of the mean protein distributions (x for fold). Fold changes above 0 represent enrichment and fold changes below 0 represent de-enrichment.

(F and H) Swarm plot showing $\log_2$ fold change of all "Signal" annotated proteins of U937 cells treated with brefeldin A (BFA; 5 μg/mL) and DMSO (black; $p = 1.1 \times 10^{-27}$, F, $p = 1.9 \times 10^{-6}$, H), TNF (purple; $p = 1.5 \times 10^{-36}$, F, n.s., H), TNF and SM (green; $p = 5.8 \times 10^{-26}$, F, n.s., H), or TNF, SM, and IDN-6556 (yellow; $p = 1.8 \times 10^{-26}$, F, n.s., H) compared with the same treatments without brefeldin A.

increased compared with the same peptides in necroptotic supernatants at 3 h of stimulation (Figure S2L). However, no semi-tryptic peptide of the conventionally released cytokines CCL2, IL8, and GRN could be detected. Necroptotic supernatants contained primarily semi-tryptic peptides of other proteins than signal proteins (Figure S2L), suggesting that the reduction of conventionally released proteins in supernatants of late-stage necroptotic cells is not only due to increased proteolytic processing. This indicates that the ER-Golgi secretory pathway is inhibited at later stages of necroptosis.

Although the inhibition of the ER-Golgi pathway at later stages of necroptosis prevents the release of conventionally released cytokines, TNF-induced cytokines CCL2 and IL8 levels were already reduced at 3 h of stimulation compared with TNF treatment alone. To test whether CCL2 is transcribed in response to TNF during necroptosis, we measured its mRNA levels and found a strong reduction compared with TNF-only treated cells after 3 and 7 h of stimulation (Figure S3M). The expression of other TNF-induced genes in U937 cells was also compromised. Of six proteins strongly upregulated in response to TNF treatment, we only found SOD2 being upregulated during necroptosis (Figure S3N). SQSTM1, NFKB2, ICAM1, PLAU, and JUND, which are all prominent targets of TNF, were not induced during necroptosis.

Receptor Shedding Occurs in Necroptotic and Apoptotic Cells

Comparing proteins released by necroptotic supernatants and apoptotic cells, we identified several proteins assigned to the keyword annotations "Receptor" as most enriched in both supernatants (Figures 3A–3C; Figure S3A; Table S2). This is unlikely to be a consequence of passive protein release caused by membrane damage, as necroptosis and apoptosis

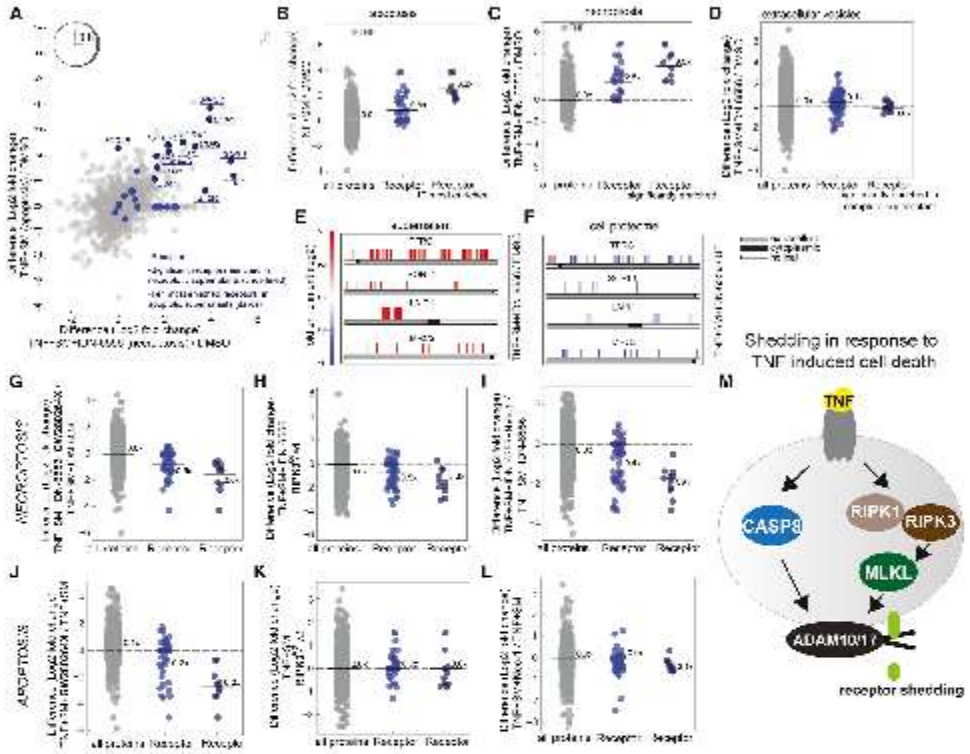


Figure 3. Necroptosis and Apoptosis Induce Receptor Shedding Mediated by ADAMs

(A) Scatterplot plotting the log_2 fold change of proteins released by necroptotic cells compared with the DMSO control against the fold change of proteins released by apoptotic cells compared with the DMSO control (3 h of treatment). Proteins assigned to the keyword annotation "Receptor" (blue), ten most enriched receptors in apoptotic supernatants (dark blue, italics), and most significant receptors in necroptotic supernatants (dark blue, underlined).

(B) Swarm plot showing log_2 fold change of all proteins (gray) and proteins assigned to the keyword annotation "Receptor" (blue; $p = 2.6 \times 10^{-6}$) released by apoptotic cells compared with the DMSO control at 3 h of treatment.

(C) Swarm plot showing log_2 fold change of all proteins (gray) and "Receptor" proteins (blue; $p = 1.3 \times 10^{-7}$) released by necroptotic cells compared with the DMSO control at 3 h of treatment.

(D) Swarm plot showing log_2 fold change of all proteins (gray), receptors (blue; n.s.), and significantly upregulated receptors in the supernatant of necroptotic cells (dark blue) present in extracellular vesicles of necroptotic cells compared with the DMSO control at 3 h of treatment.

(E and F) Matching of detected peptides by mass spectrometry to corresponding proteins using R shows peptide fold changes (log_2) in the supernatant (E) and the cellular proteome (F) of necroptotic U937 cells at 3 h of stimulation compared with DMSO-treated control cells. Selected proteins are the four most enriched receptors in the supernatant of necroptotic U937 cells compared with the DMSO control. Values are averages of n = 4.

(G) Swarm plot showing log_2 fold change of all proteins (gray), receptors (blue; $p = 6.8 \times 10^{-15}$), and significantly upregulated receptors in the supernatant of necroptotic cells (dark blue) present in supernatants of cells treated with TNF, SM, and IDN-6556 and the ADAM inhibitor GW280264X compared with cells treated with TNF, SM, and IDN-6556 for 3 h.

(H) Swarm plot showing log_2 fold change of all proteins (gray), receptors (blue; $p = 3.0 \times 10^{-3}$, FDR = 0.1), and significantly upregulated receptors in the supernatant of necroptotic cells (dark blue) present in supernatants of cells deficient for RIPK3 treated with TNF, SM, and IDN-6556 compared with wild-type cells treated the same for 3 h.

(I) Swarm plot showing log_2 fold change of all proteins (gray), receptors (blue; $p = 0.001$), and significantly upregulated receptors in the supernatant of necroptotic cells (dark blue) present in supernatants of cells treated with TNF, SM, IDN-6556, and the RIPK1 inhibitor necrostatin-1 (Nec-1) compared with cells treated with TNF, SM, and IDN-6556 for 3 h.

(J) Swarm plot showing log_2 fold change of all proteins (gray), receptors (blue; $p = 1.4 \times 10^{-6}$), and the ten most enriched receptors in the supernatants of apoptotic cells (dark blue) present in supernatants of cells treated with TNF, SM, and the ADAM inhibitor GW280264X compared with cells treated with TNF and SM for 3 h.

(legend continued on next page)

induction at 3 h precedes membrane permeabilization (Figure 1D). To test whether the detected receptors are instead released via exocytosis, we purified extracellular vesicles using serial centrifugation (STAR Methods). Proteomic analysis, however, showed no enrichment for receptors in extracellular vesicles of necroptotic and apoptotic cells compared with the control (Figure 3D; Figure S3B). Next, we sought to determine whether ectodomains are cleaved off the receptors and therefore enriched in necroptotic supernatants. Indeed, peptides derived from receptor ectodomains were preferentially released into the supernatant, whereas transmembrane or cytoplasmic domain peptides were either not detected or not enriched compared with supernatants of DMSO-treated cells (Figure 3E; Figure S3C). These cytoplasmic regions were readily detected when measuring the cellular proteome (Figure 3F; Figure S3D). A recent study showed that shedding during necroptosis contributes to inflammation and cell death in an MLKL-dependent manner (Cai et al., 2016). We show here that shedding also occurs to a lesser degree in apoptotic cells. Western blot analysis on the cell lysates did not detect phosphorylated/active MLKL at any time points during apoptosis induction, ruling out the possibility that necroptotic cells were present during apoptotic stimulation (Figure S4E).

Next, we tested whether this shedding is mediated via A disintegrin and metalloproteinase 10 (ADAM10) and 17, which are members of the disintegrins and metalloproteinase family (Cai et al., 2016). Indeed, specific inhibition of ADAM17 and ADAM10 with GW280264X reduced the release of receptors in necroptotic as well as apoptotic cells, while the release of proteins without receptor annotations remained unchanged (Figures 3G and 3J).

The lack of motif specificity for ADAM substrates complicates substrate prediction (Gooz, 2010), and we wished to investigate if there was a relationship between receptor abundance and receptor shedding. From the proteomic analysis of receptor levels in untreated cells, we conclude that the expression level itself has no impact on shedding (Figure S3F). Furthermore, the enrichment of shed receptors in the supernatant of necroptotic cells corresponded with their de-enrichment in these cells at 3 h after induction (Figure S3G).

Necroptosis-induced shedding was strongly inhibited in U937 cells deficient for RIPK3, the key activator of MLKL, or in U937 cells additionally treated with the RIPK1 inhibitor necrostatin-1 (Nec-1), where activation of MLKL cannot occur (Figures 3H and 3I; Figures S3H and S3I). In contrast to shedding during necroptosis, apoptosis-induced shedding was not prevented in RIPK3-deficient U937 cells or by addition of the RIPK1 inhibitor Nec-1 (Figures 3K and 3L). This indicates that TNF-induced shedding can be completely prevented only by blocking both caspase activation and the necroptotic pathway. Hence, we conclude that the shedding of receptors occurs during TNF-induced cell death and is not only restricted to necroptosis (Figure 3M).

Apoptotic Cells Exclusively Release Histones while Necroptotic Cells Release Lysosomal Proteins

We observed the release of apoptotic- and necroptotic-specific proteins. A group of proteins annotated as nucleosome components were most strongly enriched in supernatants of apoptotic cells (Figures 4A–4C; Figures S4A–S4C) at 3 h when DNA fragmentation and chromatin condensation were detected (Figures S4D and S4E). This was not the case in necroptosis, in which nucleosomes were de-enriched (Figure 4A; Figure S4F).

Necroptotic cells showed an increased release of lysosomal proteins after 3 h of necroptosis stimulation (Figures 4A and 4D; Figure S4G; Table S2), which was not evident during apoptosis. The release of lysosomal components was strongly reduced by the RIPK1 inhibitor Nec-1 and even more so in RIPK3-deficient U937 cells, suggesting a necroptosis-specific release pathway (Figures S4H and S4I).

Because many lysosomal proteins require cleavage for full activation, we examined whether the released lysosomal components are present in mature (cleaved) or premature (uncleaved) forms. Matching the peptide sequences to their proteins revealed that the majority were located within the mature forms of proteins released by necroptotic cells (Figure S4J). Consistently, we detected an increase of mature cathepsin B and cathepsin D in supernatants of necroptotic cells compared with DMSO-treated cells by immunoblotting (Figure S4K). Furthermore, we observed a significant increase of semi-tryptic peptides with comparable intensities to tryptic peptides of the same proteins in the supernatants of necroptotic cells (Figure S4L). Their near doubling (at 7 h) proves a dramatic increase of enzymatic activity in the supernatants of necroptotic cells. This is less likely due to increased ADAM activity, as they are active before and at 3 h of necroptosis induction. The numbers of all peptides identified across treatments did not significantly change (Figure S4M).

Membrane Permeabilization during Necroptosis Induces Lysosomal Exocytosis

Several studies have implicated the lysosomal machinery in necroptosis. The bursting of lysosomes within cells during necroptosis that preceded membrane permeabilization has been described (Vanden Berghe et al., 2010), and a role for MLKL in endosomal trafficking and increased vesicle formation has been reported (Yoon et al., 2017). Taking these studies into consideration, we asked whether lysosomal proteins were released via extracellular vesicles during exocytosis. Principal-component analysis revealed a clear separation of extracellular vesicle proteomes derived from apoptotic and necroptotic cells (Figure S4N). Consistent with previous reports, levels of MLKL and ESCRT proteins, which are involved in extracellular

(K) Swarm plot showing $\log_2$ fold change of all proteins (gray), receptors (blue; n.s.), and the ten most enriched receptors (dark blue) present in supernatants of cells deficient for RIPK3 treated with TNF and SM compared with wild-type cells treated the same for 3 h.
(L) Swarm plot showing $\log_2$ fold change of all proteins (gray), receptors (blue; $p = 2.0 \times 10^{-3}$, FDR = 0.1), and the ten most enriched receptors (dark blue) present in supernatants of cells treated with TNF, SM, and the RIPK1 inhibitor necrostatin-1 (Nec-1) compared with cells treated with TNF and SM for 3 hours.
(M) Scheme of receptor shedding downstream of apoptosis and necroptosis induction.

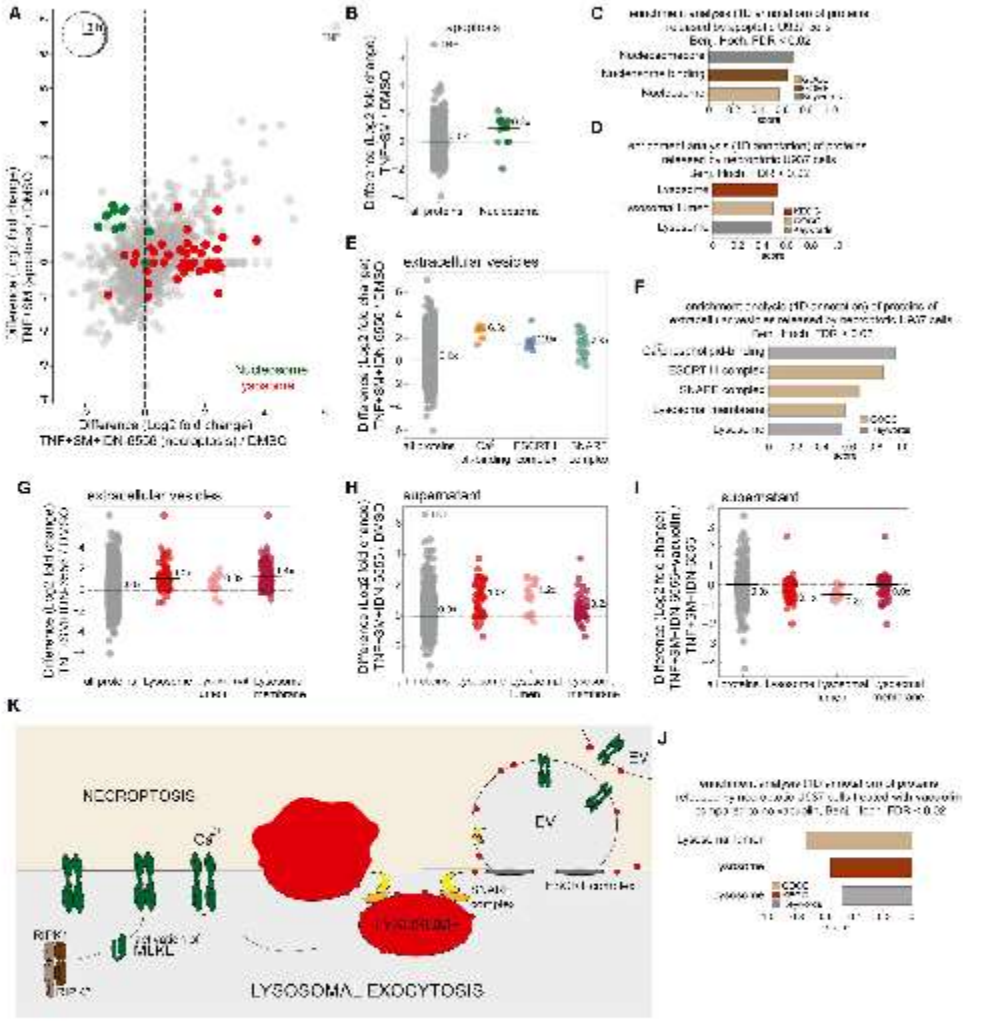


Figure 4. Lysosomal Proteins Are Released by Necroptotic Cells via Lysosomal Exocytosis

(A) Scatterplot plotting the $\log_2$ fold change of proteins released by necroptotic cells compared with the DMSO control against the fold change of proteins released by apoptotic cells compared with the DMSO control (3 h of treatment; "Nucleosome" [green], "Lysosome" [red]).

(B) Swarm plot showing $\log_2$ fold change of all proteins (gray) and proteins assigned to the GOCC annotation "Nucleosome" (green; $p = 1.2 \times 10^{-5}$) released by apoptotic cells compared with the DMSO control at 3 h of treatment.

(C and D) 1D annotation enrichment analysis of all proteins in the supernatant of apoptotic (C) and necroptotic (D) U937 cells treated for 3 h compared with the DMSO-treated control cells. Annotations including "Nucleosome" (C) and "Lysosome" (D) were selected and ranked from higher score to lower score (Benjamini-Hochberg FDR = 0.02) (STAR Methods). KEGG, Kyoto Encyclopedia of Genes and Genomes; GOCC, Gene Ontology cellular components; GOMF, Gene Ontology molecular function.

(E) Swarm plot showing $\log_2$ fold change of all proteins (gray) and proteins assigned to the annotation "Calcium/phospholipid binding" (keyword, yellow), the "ESCRTIII complex" (GOCC, blue), and "SNARE complex" (GOCC, green) present in extracellular vesicles of necroptotic cells compared with the DMSO control

(legend continued on next page)

vesicle formation and removal of MLKL from the plasma membrane, were increased in extracellular vesicles derived from necroptotic cells compared with control or apoptotic cells after 3 h of stimulation (Yoon et al., 2017; Gong et al., 2017) (Figures 4E and 4F; Figure S4O).

Interestingly, we observed a significant enrichment of lysosomal components in extracellular vesicles of necroptotic cells (Figures 4F and 4G; Figure S4O). Luminal lysosomal proteins were preferentially present in total supernatants, whereas membrane-bound ones were found mainly in purified extracellular vesicles (Figures 4H and 4G). Distribution of lysosomal membrane and lysosomal luminal proteins differed significantly ($p < 0.0001$). In classical endosomal exocytosis, luminal and membrane lysosomal proteins would be co-released in extracellular vesicles, which is not the case here.

An alternative process, termed lysosomal exocytosis, which occurs in response to plasma membrane permeabilization, for example, through ionomycin treatment or mechanical stress, has also been described in TNF-treated L929 cells, presumably to repair plasma membranes (Andrews, 2002; Reddy et al., 2001; Ono et al., 2001). In that process, the lysosomal membrane fuses with the plasma membrane, releasing its cargo into the extracellular space. This corresponds with our observation that luminal lysosomal proteins are enriched in the supernatant, and membrane lysosomal proteins are enriched in extracellular vesicles that also contain plasma membrane. Also the enrichment of calcium/phospholipid-binding and SNARE (SNAP receptor) proteins, which play an important role in vesicle fusion in response to increased intracellular free calcium upon membrane permeabilization, fits the model of lysosomal exocytosis. (Rao et al., 2004; Shen et al., 2016) (Figures 4E, 4F, and 4K; Figure S4O).

To determine the role of lysosomal exocytosis in the release of lysosomal components, we inhibited lysosomal exocytosis using vacuolin, which leads to homotypic fusion of endosomes and lysosomes. It blocks calcium-dependent exocytosis in a cell type-dependent manner (Cerny et al., 2004; Lu et al., 2014; Shaik et al., 2009). Treatment of necroptotic cells with vacuolin partially inhibited the release of lysosomal components in the total supernatant (Figures 4I and 4J).

DISCUSSION

Chronic inflammatory TNF-induced cell death is a major contributor to a range of inflammatory diseases, as dying cells expose immunogenic proteins normally found within the cell. Our proteomic approach provided a comprehensive and unbiased view on proteins released by apoptotic and necroptotic cells. As expected, many proteins were statistically significantly changed in supernatants of apoptotic cells compared with control. This was even more pronounced in necroptosis, particularly at later time points.

The release of pro-inflammatory cytokines contributes to the immunogenicity of cell death pathways. Whereas conventionally released cytokines were generally reduced in supernatants of necroptotic cells, non-conventionally released cytokines were primarily unchanged or increased compared with TNF alone or apoptosis. Plasma membrane permeabilization facilitates the unconventional release of cytokines. The reduction of conventionally released cytokines is due to the inhibition of the ER-Golgi trafficking pathway at later stages of necroptosis and the early plasma membrane permeabilization, which is most likely the reason for compromised transcription and expression of TNF-induced proteins. These findings do not conflict with the results of Orozco et al. (2019) regarding continuous cytokine mRNA translation despite plasma membrane permeabilization during necroptosis, because the cytokine production of stimulated living cells is expected to exceed that of necroptotic cells.

Other immunogenic proteins were differentially released by apoptotic and necroptotic cells, such as histones, which were increased only in supernatants of apoptotic cells. They are linked to diseases like lupus (Chen et al., 2014; Radic et al., 2004), chronic obstructive pulmonary disease (COPD), cerebral stroke, and sepsis. Conversely, HMGB1, a highly inflammatory molecule, was increased in the supernatants of necroptotic cells, as previously described (Magna and Pisetsky, 2014).

Besides the detection of proteins with known intercellular functions, our study provides insights into intracellular mechanisms that occur during TNF-induced cell death and that may also contribute to inflammation. For example, we observed an

at 3 h of treatment. The p value ($-\log_{10}$) of the 1D annotation enrichment for "Calcium/phospholipid binding" is 5.3×10^{-8}, for "ESCRTIII complex" is 2.9×10^{-5}, and for "SNARE complex" is 1.7×10^{-6}.

(F) 1D annotation enrichment analysis of all proteins of extracellular vesicles in the supernatant of necroptotic U937 cells treated for 3 h with TNF, SM, and IDN-6556 compared with the DMSO-treated control cells. Annotations are ranked from higher score to lower score (Benjamini-Hochberg FDR = 0.02).

(G) Swarm plot showing $\log_2$ fold change of all proteins (gray) and proteins assigned to the keyword annotations "Lysosome" (keywords, red), "Lysosomal lumen" (GOCC, nude), and "Lysosomal membrane" (GOCC, pink) in extracellular vesicles of necroptotic cells compared with the DMSO control at 3 h of treatment. The p value ($-\log_{10}$) of the 1D annotation enrichment for "Lysosome" is 3.8×10^{-23}, for "Lysosomal lumen" is 9.2×10^{-3} (FDR = 0.1), and for "Lysosomal membrane" is 2.7×10^{-39}.

(H) Swarm plot showing $\log_2$ fold change of the same annotation groups as in (G) released into the supernatant of necroptotic cells compared with the DMSO control at 3 h of treatment. The p value ($-\log_{10}$) of the 1D annotation enrichment for "Lysosome" is 1.5×10^{-10}, for "Lysosomal lumen" is 2.4×10^{-6}, and for "Lysosomal membrane" is 1.0×10^{-3}.

(I) Swarm plot showing $\log_2$ fold change of all proteins (gray) and lysosomal proteins (red) present in supernatants of cells treated with TNF, SM, IDN-6556, and the lysosomal exocytosis inhibitor vacuolin compared with cells treated with TNF, SM, and IDN-6556 for 3 h. The p value ($-\log_{10}$) of the 1D annotation enrichment for "Lysosome" is 9.8×10^{-12} and for "Lysosomal lumen" is 3.7×10^{-10}.

(J) 1D annotation enrichment analysis of all proteins in the supernatant of necroptotic U937 cells treated for 3 h with TNF, SM, IDN-6556, and vacuolin compared with the same treatment without vacuolin. Annotations containing "Lysosome" are selected and ranked from higher to lower score (Benjamini-Hochberg FDR = 0.02).

(K) Scheme of lysosomal exocytosis and formation of extracellular vesicle as response to plasma membrane permeabilization during necroptosis.

enrichment of receptor ectodomains in supernatants of necroptotic and apoptotic cells due to ADAM-mediated shedding. This is a physiologically important process: the cleavage of cell adhesion proteins promotes detachment of dying cells from surrounding tissue; receptor cleavage abrogates signal transduction into the cell as well as between cells by the generation of decoy receptors cleavage (Blaydon et al., 2011; Jones et al., 2016). Together, this switches signaling within dying cells to signaling between cells by the release of autocrine and paracrine signals. Shedding was already reported to be activated during necroptosis (Cai et al., 2016). Our data provide a catalog of proteins affected and demonstrate that it occurs at later stages of TNF-induced cell death and can be prevented only by inhibiting both apoptotic and necroptotic executioners. Further studies should delineate the intercellular activities of the distinct signals released by the different forms of cell death on neighboring cells.

A second unexpected group of proteins released by necroptotic cells belong to lysosomes. We conclude that this is caused by lysosomal fusion to the plasma membrane, a mechanism that results in the release of lysosomal cargo into the supernatant in response to membrane permeabilization, described as lysosomal exocytosis.

Lysosomal exocytosis was originally described as a plasma membrane repair mechanism in response to Ca^{2+} influx (Reddy et al., 2001). A study on the infection of pathogens with bacterial type III secretion system confirmed this role (Roy et al., 2004). During necroptosis the ESCRT machinery removes MLKL from the plasma membrane to prevent membrane permeabilization (Yoon et al., 2017; Gong et al., 2017). We speculate that lysosomal exocytosis is an additional and very last resort to protect the cell from death (Ono et al., 2001). Lysosomal exocytosis was also investigated as a defense mechanism to remove intracellular pathogens (Miao et al., 2015), but the effect of lysosomal enzymes on pathogens in the extracellular regions has not been studied in detail. Many lysosomal enzymes work best at acidic pH but may still retain some enzymatic activity in the extracellular space, which is sufficient to degrade the extracellular matrix (Fonović and Turk, 2014). This is consistent with our observation of increased semi-tryptic peptides in supernatants of necroptotic cells, which indicates increased proteolytic activity. These lysosomal enzymes may also affect cells in close proximity. Whether lysosomal exocytosis following necroptosis induction is just a consequence of cell membrane permeabilization or a last chance to rescue dying cells, and whether it significantly affects the environment, such as pathogens or neighboring cells, are exciting questions for future studies.

In conclusion, our study provides in-depth, quantitative, and dynamic catalogs of proteins differentially and commonly released during apoptotic and necroptotic cell death. This offers a solid basis for investigating the complexity of biological processes that are either incidental or regulated, harmful or functional during apoptosis and necroptosis. Our results demonstrate that an unexpectedly wide range of different proteins are released. Only some of these contribute directly to inflammation, whereas others are more likely to act in more complex ways that still need to be elucidated in detail in these very different modes of cell death.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Sample preparation for proteomic analysis and ELISA of supernatants
  - Preparation of extracellular vesicles
  - Western blotting
  - Cell death analysis
  - DNA laddering assay
  - qPCR
  - Chromatography and mass spectrometry
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2019.12.079.

ACKNOWLEDGMENTS

We thank James Murphy for critically reading the manuscript, Julie Stafford for proofreading, Georg H. Borner for his scientific input, and John Silke and Najoua Lalaoui (Walter and Eliza Hall Institute) for providing birinapant and compound A. Martin Spitaler, Giovanni Cardone, and Markus Ostler (Max Planck Institute of Biochemistry) helped with flow cytometry. We also thank all members of the Department of Proteomics and Signal Transduction at the Max Planck Institute of Biochemistry in Martinsried for help and discussions, and especially Bianca Spiettstoesser, Steven Dewitz, Igor Paron, Christian Deimi, and Gabriele Sowa for technical assistance. This work was supported by the Max Planck Society for the Advancement of Science Deutsche Forschungsgemeinschaft SFB914, and a Marie Sklodowska-Curie Actions Individual Fellowship to M.C.T.

AUTHOR CONTRIBUTIONS

M.C.T., M.M., and F.M. conceived the project and wrote the manuscript. M.M. and F.M. supervised the project. M.C.T. designed and performed all experiments and analyzed all data. A.F. helped with the preparation of human primary macrophages and modified the R script for peptide-to-protein matching. C.A.S. helped with western blotting and edited the manuscript. K.P. provided scientific input and edited the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 4, 2019
Revised: November 7, 2019
Accepted: December 19, 2019
Published: January 28, 2020

REFERENCES

Aebersold, R., and Mann, M. (2016). Mass-spectrometric exploration of proteome structure and function. Nature *537*, 347–355.

Andrews, N.W. (2002). Lysosomes and the plasma membrane: trypanosomes reveal a secret relationship. J. Cell Biol. *158*, 389–394.

Arslan, S.C., and Scheidereit, C. (2011). The prevalence of TNFα-induced necrosis over apoptosis is determined by TAK1-RIP1 interplay. PLoS ONE *6*, e25069.

Bertrand, M.J., Milutinovic, S., Dickson, K.M., Ho, W.C., Boudreault, A., Durkin, J., Gillard, J.W., Jaquith, J.B., Morris, S.J., and Barker, P.A. (2008). cIAP1 and cIAP2 facilitate cancer cell survival by functioning as E3 ligases that promote RIP1 ubiquitination. Mol. Cell *30*, 689–700.

Blaydon, D.C., Biancheri, P., Di, W.L., Plagnol, V., Cabral, R.M., Brooke, M.A., van Heel, D.A., Ruschendorf, F., Toynbee, M., Walne, A., et al. (2011). Inflammatory skin and bowel disease linked to ADAM17 deletion. N. Engl. J. Med. *365*, 1502–1508.

Bradley, J.R. (2008). TNF-mediated inflammatory disease. J. Pathol. *214*, 149–160.

Cai, Z., Zhang, A., Choksi, S., Li, W., Li, T., Zhang, X.M., and Liu, Z.G. (2016). Activation of cell-surface proteases promotes necroptosis, inflammation and cell migration. Cell Res. *26*, 886–900.

Cerny, J., Feng, Y., Yu, A., Miyake, K., Borgonovo, B., Klumperman, J., Meldolesi, J., McNeil, P.L., and Kirchhausen, T. (2004). The small chemical vacuolin-1 inhibits Ca(2+)-dependent lysosomal exocytosis but not cell resealing. EMBO Rep. *5*, 883–888.

Chen, R., Kang, R., Fan, X.G., and Tang, D. (2014). Release and activity of histone in diseases. Cell Death Dis. *5*, e1370.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Cox, J., and Mann, M. (2012). 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. BMC Bioinformatics *13* (Suppl *16*), S12.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. *10*, 1794–1805.

Elliott, M.R., and Ravichandran, K.S. (2010). Clearance of apoptotic cells: implications in health and disease. J. Cell Biol. *189*, 1059–1070.

Elmore, S. (2007). Apoptosis: a review of programmed cell death. Toxicol. Pathol. *35*, 495–516.

Fadok, V.A., Voelker, D.R., Campbell, P.A., Cohen, J.J., Bratton, D.L., and Henson, P.M. (1992). Exposure of phosphatidylserine on the surface of apoptotic lymphocytes triggers specific recognition and removal by macrophages. J. Immunol. *148*, 2207–2216.

Fonović, M., and Turk, B. (2014). Cysteine cathepsins and extracellular matrix degradation. Biochim. Biophys. Acta *1840*, 2560–2570.

Frauenstein, A., and Meissner, F. (2018). Quantitative proteomics of secreted proteins. Methods Mol. Biol. *1714*, 215–227.

Gong, Y.N., Guy, C., Olauson, H., Becker, J.U., Yang, M., Fitzgerald, P., Linkermann, A., and Green, D.R. (2017). ESCRT-III acts downstream of MLKL to regulate necroptotic cell death and its consequences. Cell *169*, 286–300.e16.

Gooz, M. (2010). ADAM-17: the enzyme that does it all. Crit. Rev. Biochem. Mol. Biol. *45*, 146–169.

Hildebrand, J.M., Tanzer, M.C., Lucet, I.S., Young, S.N., Spall, S.K., Sharma, P., Pierotti, C., Garnier, J.M., Dobson, R.C., Webb, A.I., et al. (2014). Activation of the pseudokinase MLKL unleashes the four-helix bundle domain to induce membrane localization and necroptotic cell death. Proc. Natl. Acad. Sci. U S A *111*, 15072–15077.

Jones, P., Côté, R.G., Cho, S.Y., Kile, S., Martens, L., Quinn, A.F., Thomeycroft, D., and Hermjakob, H. (2008). PRIDE: new developments and new datasets. Nucleic Acids Res. *36*, D878–D883.

Jones, J.C., Rustagi, S., and Dempsey, P.J. (2016). ADAM proteases and gastrointestinal function. Annu. Rev. Physiol. *78*, 243–276.

Kaczmarek, A., Vandenabeele, P., and Krysko, D.V. (2013). Necroptosis: the release of damage-associated molecular patterns and its physiological relevance. Immunity *38*, 209–223.

Kearney, C.J., Cullen, S.P., Tynan, G.A., Henry, C.M., Clancy, D., Lavelle, E.C., and Martin, S.J. (2015). Necroptosis suppresses inflammation via termination of TNF- or LPS-induced cytokine and chemokine production. Cell Death Differ. *22*, 1313–1327.

Kralj, M., Husnjak, K., Kozlovic, T., and Pavelić, J. (2003). Endogenous p21WAF1/CIP1 status predicts the response of human tumor cells to wild-type p53 and p21WAF1/CIP1 overexpression. Cancer Gene Ther. *10*, 457–467.

Larance, M., and Lamond, A.I. (2015). Multidimensional proteomics for cell biology. Nat. Rev. Mol. Cell Biol. *16*, 269–280.

Lu, Y., Dong, S., Hao, B., Li, C., Zhu, K., Guo, W., Wang, Q., Chaung, K.H., Wong, C.W., Wu, W.T., et al. (2014). Vacuolin-1 potently and reversibly inhibits autophagosome-lysosome fusion by activating RAB5A. Autophagy *10*, 1895–1905.

Magna, M., and Pisetsky, D.S. (2014). The role of HMGB1 in the pathogenesis of inflammatory and autoimmune diseases. Mol. Med. *20*, 138–146.

McComb, S., Cheung, H.H., Korneluk, R.G., Wang, S., Krishnan, L., and Sad, S. (2012). cIAP1 and cIAP2 limit macrophage necroptosis by inhibiting Rip1 and Rip3 activation. Cell Death Differ. *19*, 1791–1801.

Meissner, F., Scheltema, R.A., Mollenkopf, H.J., and Mann, M. (2013). Direct proteomic quantification of the secretome of activated immune cells. Science *340*, 475–478.

Miao, Y., Li, G., Zhang, X., Xu, H., and Abraham, S.N. (2015). A TRP channel senses lysosome neutralization by pathogens to trigger their expulsion. Cell *161*, 1306–1319.

Murphy, J.M., Czabotar, P.E., Hildebrand, J.M., Lucet, I.S., Zhang, J.G., Alvarez-Diaz, S., Lewis, R., Lalaoui, N., Metcalf, D., Webb, A.I., et al. (2013). The pseudokinase MLKL mediates necroptosis via a molecular switch mechanism. Immunity *39*, 443–453.

Nolte, H., MacVicar, T.D., Tellkamp, F., and Krüger, M. (2018). Instant Clue: a software suite for interactive data visualization and analysis. Sci. Rep. *8*, 12648.

Oberst, A., Dillon, C.P., Weinlich, R., McCormick, L.L., Fitzgerald, P., Pop, C., Hakem, R., Salvesen, G.S., and Green, D.R. (2011). Catalytic activity of the caspase-8-FLIP(L) complex inhibits RIPK3-dependent necrosis. Nature *471*, 363–367.

Ono, K., Wang, X., and Han, J. (2001). Resistance to tumor necrosis factor-induced cell death mediated by PMCA4 deficiency. Mol. Cell. Biol. *21*, 8278–8288.

Orning, P., Weng, D., Starheim, K., Ratner, D., Best, Z., Lee, B., Brooks, A., Xia, S., Wu, H., Kelliher, M.A., et al. (2018). Pathogen blockade of TAK1 triggers caspase-8-dependent cleavage of gasdermin D and cell death. Science *362*, 1064–1069.

Orozco, S.L., Daniels, B.P., Yatim, N., Messmer, M.N., Quarato, G., Chen-Harris, H., Cullen, S.P., Snyder, A.G., Ralli-Jain, P., Frase, S., et al. (2019). RIPK3 activation leads to cytokine synthesis that continues after loss of cell membrane integrity. Cell Rep. *28*, 2275–2287.e5.

Pop, C., Oberst, A., Drag, M., Van Raam, B.J., Riedl, S.J., Green, D.R., and Salvesen, G.S. (2011). FLIP(L) induces caspase 8 activity in the absence of interdomain caspase 8 cleavage and alters substrate specificity. Biochem. J. *433*, 447–457.

Radic, M., Marion, T., and Monestier, M. (2004). Nucleosomes are exposed at the cell surface in apoptosis. J. Immunol. *172*, 6692–6700.

Rao, S.K., Huynh, C., Proux-Gillardeaux, V., Galli, T., and Andrews, N.W. (2004). Identification of SNAREs involved in synaptotagmin VII-regulated lysosomal exocytosis. J. Biol. Chem. *279*, 20471–20479.

Reddy, A., Caler, E.V., and Andrews, N.W. (2001). Plasma membrane repair is mediated by Ca(2+)-regulated exocytosis of lysosomes. Cell *106*, 157–169.

Rock, K.L., and Kono, H. (2008). The inflammatory response to cell death. Annu. Rev. Pathol. *3*, 99–126.

Roy, D., Liston, D.R., Idone, V.J., Di, A., Nelson, D.J., Pujol, C., Blaska, J.B., Chakrabarti, S., and Andrews, N.W. (2004). A process for controlling

intracellular bacterial infections induced by membrane injury. Science *304*, 1515–1518.

Sedhawer, M., Heinzmann, F., D'Artista, L., Harbig, J., Roux, P.F., Hoenicke, L., Dang, H., Klotz, S., Robinson, L., Doré, G., et al. (2018). Necroptosis microenvironment directs lineage commitment in liver cancer. Nature *562*, 69–75.

Shaik, G.M., Dráberová, L., Haneberg, P., and Dráber, P. (2009). Vacuolin-1-modulated exocytosis and cell reseating in mast cells. Cell. Signal. *21*, 1337–1345.

Shen, Y.T., Gu, Y., Su, W.F., Zhong, J.F., Jin, Z.H., Gu, X.S., and Chen, G. (2016). Rab27b is involved in lysosomal exocytosis and proteolipid protein trafficking in oligodendrocytes. Neurosci. Bull. *32*, 331–340.

Snyder, A.G., Hubbard, N.W., Messmer, M.N., Kofman, S.B., Hagan, C.E., Orozco, S.L., Chiang, K., Daniels, B.P., Baker, D., and Oberst, A. (2019). Intratumoral activation of the necroptotic pathway components RIPK1 and RIPK3 potentiates antitumor immunity. Sci. Immunol. *4*, eaaw2004.

Sun, L., Wang, H., Wang, Z., He, S., Chen, S., Liao, D., Wang, L., Yan, J., Liu, W., Lei, X., and Wang, X. (2012). Mixed lineage kinase domain-like protein mediates necrosis signaling downstream of RIP3 kinase. Cell *148*, 213–227.

Tanzer, M.C., Tripaydonis, A., Webb, A.I., Young, S.N., Varghese, L.N., Hall, C., Alexander, W.S., Hildebrand, J.M., Silke, J., and Murphy, J.M. (2015). Necroptosis signalling is tuned by phosphorylation of MLKL residues outside the pseudokinase domain activation loop. Biochem. J. *471*, 255–265.

Tanzer, M.C., Khan, N., Rickard, J.A., Etemadi, N., Lalaoui, N., Spall, S.K., Hildebrand, J.M., Segal, D., Misaart, M., Chau, D., et al. (2017). Combination of IAP antagonist and IFNγ activates novel caspase-10- and RIPK1-dependent cell death pathways. Cell Death Differ. *24*, 481–491.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731–740.

Vanden Berghe, T., Vanlangenakker, N., Parthoens, E., Deckers, W., Devos, M., Festjens, N., Guerin, C.J., Brunk, U.T., Declercq, W., and Vandenabeele, P. (2010). Necroptosis, necrosis and secondary necrosis converge on similar cellular disintegration features. Cell Death Differ. *17*, 922–930.

Varfolomeev, E., Blankenship, J.W., Wayson, S.M., Fedorova, A.V., Kayagaki, N., Garg, P., Zobel, K., Dynek, J.N., Elliott, L.O., Wallweber, H.J., et al. (2007). IAP antagonists induce autoubiquitination of c-IAPs, NF-kappaB activation, and TNFalpha-dependent apoptosis. Cell *131*, 669–681.

Vercammen, D., Beyaert, R., Denecker, G., Goossens, V., Van Loo, G., Declercq, W., Grooten, J., Fiers, W., and Vandenabeele, P. (1998). Inhibition of caspases increases the sensitivity of L929 cells to necrosis mediated by tumor necrosis factor. J. Exp. Med. *187*, 1477–1485.

Vince, J.E., Wong, W.W., Khan, N., Feltham, R., Chau, D., Ahmed, A.U., Benetatos, C.A., Chunduru, S.K., Condon, S.M., McKinlay, M., et al. (2007). IAP antagonists target cIAP1 to induce TNFalpha-dependent apoptosis. Cell *131*, 682–693.

Wang, H., Sun, L., Su, L., Rizo, J., Liu, L., Wang, L.F., Wang, F.S., and Wang, X. (2014). Mixed lineage kinase domain-like protein MLKL causes necrotic membrane disruption upon phosphorylation by RIP3. Mol. Cell *54*, 133–146.

Yoon, S., Kovalenko, A., Bogdanov, K., and Wallach, D. (2017). MLKL, the protein that mediates necroptosis, also regulates endosomal trafficking and extracellular vesicle generation. Immunity *47*, 51–65.e7.

Zhu, K., Liang, W., Ma, Z., Xu, D., Cao, S., Lu, X., Liu, N., Shan, B., Qian, L., and Yuan, J. (2018). Necroptosis promotes cell-autonomous activation of proinflammatory cytokine gene expression. Cell Death Dis. *9*, 500.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-human caspase-8	MBL	Cat# M058-3, RRID: AB_590761
Anti-cleaved human caspase-3	Cell Signaling Technology	Cat# 9661, RRID: AB_2341188
Anti-human phospho (S358) MLKL	Abcam	Cat# ab187091, RRID: AB_2619685
Anti-human MLKL	Merck Millipore	Cat# MABC604
Anti-GAPDH	Cell Signaling Technology	Cat# 8884s, RRID: AB_11129865
Anti-human CATHEPSIN B	Calbiochem	Cat# IM27L, RRID: AB_2274848
Anti-human CATHEPSIN D	Abcam	Cat# ab75852, RRID: AB_1523267
Anti-rabbit IgG, HRP-linked	Cell Signaling Technology	Cat# 7074, RRID: AB_2099233
Anti-mouse IgG, HRP-linked	Cell Signaling Technology	Cat# 7076, RRID: AB_330924
Bacterial and Virus Strains		
XL1-Blue Competent Cells	Agilent Technologies	200249
Biological Samples		
Buffy Coats	Blood donations to the red cross: "Blutspendedienst des Bayerischen Roten Kreuzes gemeinnützige GmbH"	N/A
Chemicals, Peptides, and Recombinant Proteins		
Recombinant human TNF	ImmunoTools	11343017
Recombinant human GM-CSF	ImmunoTools	21173121
Birinapant	Selleckchem	S7015
Compound A	Gift from Prof. John Silke	N/A
IDN-6556 (Emricasan)	MedChem Express (MCE)	HY-10396
GW280264X	Adooq	AQ03632
Vacuolin	Sigma	673000-10mg
Necrostatin-1	Sigma	N9037-10mg
cOmplete, Mini Protease Inhibitor Cocktail	Sigma	4693159001
Propidium Iodide	Sigma	81845
Histopaque-1077	Sigma-Aldrich	10771
Heat Inactivated Fetal Bovine Serum	Invitrogen	10270106
Penicillin/Streptomycin	Invitrogen	15140122
RPMI medium	Invitrogen	61870044
Ponceau BS	Sigma	B6008-100 g
Blasticidin	Invivogen	Ant-bl-1
DME medium	Invitrogen	31966047
Serum- and phenolred free media	Thermo Fisher Scientific	11835063
Urea	Sigma	45129-500 g
Thiourea	Sigma	T8656-500 g
Trizma	Sigma	T1503-1kg
Dithiothreitol (DTT)	Sigma	D0632-100 g
Chloroacetamide (CAA)	Sigma	C0267-100 g
Iodoacetamide (IAA)	Sigma	I6125-100 g
Ammonium bicarbonate	Sigma	A6141
Trypsin	Sigma	T6567-1mg
Lys-C	Wako Chemicals	129-02541
DMSO	Sigma	D2650-100ml

(Continued on next page)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Acetone	Fisher Chemical	67-64-1
Acetonitrile	VWR	20048320
Trifluoroacetic acid	Merck	8082600100
Formic acid	Merck	1002641000
PBS	GIBCO	14190-094
Tween	Acros	233360010
EDTA	Sigma	03677-500 g
Skim milk powder	Roth	T145.3
BSA	Serva	11930.03
Glycerol	Sigma	G5516-1L
Sodium dodecyl sulfate (SDS)	Roth	CN30.3
Sodium deoxycholate (SDC)	Sigma	30970-100 g
Sodium chloride (NaCl)	VWR	27810.295
NuPAGE LDS Sample Buffer (4x)	Invitrogen	NP0007
Igepal	Sigma	I3021-500ml
GelRed nucleic acid stain	Biotium	41003-1
Agarose	Invitrogen	16500-500
RNase A	Thermo Fisher	EN0531
Proteinase K	Thermo Fisher	AM2546
Brefeldin A	Invitrogen	00-4506-51
Puromycin	Invivogen	Ant-pr-5
Polybrene	Sigma	107689
Critical Commercial Assays		
Monocyte isolation kit	Miltenyi Biotec	130-091-153
Human CCL2 Quantikine kit (ELISA)	R&D Systems	DCP00
Human HMGB1 Quantikine kit (ELISA)	Cloud-Clone	ABIN414391
SuperScript III	Invitrogen	11752-050
RNeasy Plus Mini Kit	QIAGEN	74134
QIAshredder	QIAGEN	79656
Deposited Data		
Raw Mass Spectrometry Data Files	This paper	ProteomeXchange Consortium via the PRIDE partner repository, with the dataset identifier PXD014966
Experimental Models: Cell Lines		
U937 cell line	ATCC	CRL-1593.2
293T cell line	ATCC	CRL-3216
Human primary macrophages made out of buffy coats	Blood donation to the Red cross	N/A
Oligonucleotides		
sgRNA 1 against RIPK3: gaattcgtgctgcgcctaga	Tanzer et al., 2017	Metabion
sgRNA 2 against RIPK3: cgcccccttggtgtccatcg	This study	Metabion
Recombinant DNA		
lentiCRISPR v2	Transomics	TELA1002
psPAX2	http://addgene.org/12260	Addgene #122260
pMD2.G	http://addgene.org/12259	Addgene #122259
Software and Algorithms		
MaxQuant	Cox and Mann, 2008, Version 1.5.0.38	https://www.biochem.mpg.de/5111795/maxquant
Perseus	Tyanova et al., 2016, Version 1.5.3.0	https://www.biochem.mpg.de/5111810/perseus

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
XCalibur	Thermo Scientific	https://www.thermofisher.com/order/catalog/product/OPTON-30487
Prism Graphpad	N/A	https://www.graphpad.com/scientific-software/prism/
Instant Clue	Nolte et al., 2018	http://www.instantclue.uni-koeln.de/
FlowJo	N/A	https://www.flowjo.com/
R	N/A	https://www.r-project.org/
Adobe Illustrator	N/A	https://www.adobe.com/de/products/illustrator.html?gclid=Cj0KCQiAj4bIBRC-ARIsAA4WaFIBQ4ih92FMEfFeoX9KzbZX3ANOyHPWvjWC_Dc8uqRkApS93wRL-0_MaAvthEALw_wcB&sdid=88X75SKP&mv=search&ef_id=Cj0KCQiAj4bIBRC-ARIsAA4WaFIBQ4ih92FMEfFeoX9KzbZX3ANOyHPWvjWC_Dc8uqRkApS93wRL-0_MaAvthEALw_wcB:G:s&s_kwcid=AL!3085!3!427727387414!e!g!!adobe%20illustrator
PRIDE	Jones et al., 2006	https://www.ebi.ac.uk/pride/archive/

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Felix Meissner (meissner@biochem.mpg.de). This study did not generate new unique materials and reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Experiments described in this study are performed with the U937 cell line and human primary macrophages. U937 were purchased at ATCC and cultured in RPMI supplemented with 10% FCS and Pen/Step. Human primary macrophages were generated by isolating PBMCs from buffy coats (generated from blood donations) using Histopaque-1077 (Sigma-Aldrich, 10771). Monocytes were isolated using a monocyte isolation kit (Miltenyi Biotec, 130-091-153), plated and 50 nM GM-CSF (ImmunoTools, 21173121) was added. Cells were left for 7 days to induce macrophage differentiation. All cells were stimulated and kept for the entire treatment duration in serum- and phenol-red free media (Thermo Fisher Scientific, 11835063).

METHOD DETAILS

Sample preparation for proteomic analysis and ELISA of supernatants

1.5×10^6 U937 cells or full 24 wells of primary human macrophages were treated with TNF (30 ng/ml) and the IAP inhibitor, birinapant (SM, 250 nM) (Compound A (2 μM) for human primary macrophages) to induce apoptosis or TNF and birinapant (SM, 250 nM) (Compound A (2 μM) for human primary macrophages) and the caspase inhibitor IDUN-6556 (IDN-6556/emricasan, 10 μM) to induce necroptosis, GW280264X (10 μM, pretreatment for 60 minutes before apoptosis and necroptosis induction), necrostatin-1 (Nec-1, 50 μM), vacuolin-1 (vacuolin, 10 μM) and brefeldin A (BFA, 3 μg/ml, 20 minutes before apoptosis and necroptosis induction). Control cells were treated with DMSO, TNF, birinapant and IDUN-6556 alone. For ELISA supernatants were taken and processed according to the manufacturer's instruction. For proteomic analysis supernatants were spun for 5 minutes at 500 x g to remove cells and filtered through 0.22 μm filters to remove cell debris. 8 M urea in 40 mM HEPES was added to bring the sample to a final concentration of 2.7 M urea, which was subsequently sonicated (Bioruptor) for 20 minutes. Proteins were reduced by the addition of 10 mM DTT and incubated for 30 minutes at room temperature. Proteins were alkylated by the addition of 55 mM iodoacetamide and incubated for 20 minutes at room temperature in the dark. 100 mM Thiourea was added before the addition of 1 μg lysC and trypsin and digested over night at room temperature. Enzyme activity was stopped by the addition of 2% ACN and 0.6% TFA and proteins were cleaned up on C18 StageTips (Frauenstein and Meissner, 2018).

Preparation of extracellular vesicles

12×10^6 U937 cells were treated for three hours with TNF and birinapant or TNF and birinapant and IDUN-6556 or with DMSO alone as a control. Cells were spun for 5 minutes at 500 x g. Supernatants were taken and spun again for 30 minutes at 10,000 x g. Supernatants were taken again and spun for 60 minutes at 100,000 x g to obtain extracellular vesicles. The pellets containing extracellular vesicles were washed with ice cold PBS and again centrifuged for 60 minutes at 100,000 x g. Extracellular vesicles were lysed in 8 M Urea with 50 mM Tris (pH8), reduced with 10 mM DTT, alkylated with 40 mM CAA and digested for 2 hours

with lysC (1 μg/sample) before 1:4 dilution with 50 mM ammonium bicarbonate and the addition of trypsin (1 μg/sample). Trypsin and lysC digestion occurred over night at room temperature and peptides were cleaned up by C18 stage tipping.

Western blotting

Two million cells were stimulated, washed in PBS and lysed in buffer (1% IGEPAL, 10% Glycerol, 2 mM EDTA, 50 mM Tris pH 7.5, 150 mM NaCl) supplemented with phosphatase- (Sigma-Aldrich, 4906845001) and protease inhibitors (Sigma-Aldrich, 4693159001). Lysates were kept on ice for 20 minutes and centrifuged at 16,100 x g for 15 minutes before the addition of 6 x SDS sample loading buffer (450 mM Tris-HCl, pH 8, 60% (v/v) glycerol, 12% (w/v) SDS, 0.02% (w/v) bromophenol blue, 600 mM DTT) to the supernatant, followed by boiling and sonication. For immunoblotting of supernatants ten million cells were stimulated in serum-free media, spun down and proteins in supernatants were precipitated with ice-cold Acetone (80% final concentration). The next day precipitates were spun down for 30 minutes at full speed and washed twice in 80% Acetone before addition of 6 x SDS sample loading buffer, boiling and sonication. Separation occurred on 12% Novex Tris-glycine gels (Thermo Fisher Scientific, XP00120BOX) and transferred onto PVDF membranes (Merck Millipore, IPVH00010) or Nitrocellulose membranes (Amersham, 10600002). Membranes were blocked in 5% milk and antibodies diluted in 2% BSA in PBST. Antibodies used for immunoblotting were as follows: anti-human caspase-8 (MBL, M058-3), anti-cleaved human caspase-3 (Cell Signaling, 9661), anti-human phospho (S358) MLKL (Abcam, ab187091), anti-human MLKL (Merck Millipore, MABC604), anti-GAPDH (Cell Signaling Technology, 8884s), anti-CATHEPSIN B (Calbiochem, IM27L) and anti-CATHEPSIN D (Abcam, ab75825).

Cell death analysis

1×10^5 U937 cells and human primary macrophages were plated in 24 well plates and treated with TNF (30 ng/ml), IDN-6556 (10 μM), necrostatin-1 (Nec-1, 50 μM), birinapant (SM, 250 nM) for U937 and Compound A (SM, 2 μM) for human primary macrophages. Cell death was measured by propidium iodide incorporation using flow cytometry (FACS Attune NxT, BD FACS Aria III) and analyzed using Graphpad Prism.

DNA laddering assay

DNA laddering assay was performed as described previously with some changes (Kalaj et al., 2003). After stimulation 2×10^6 cells were washed with PBS and lysed in 200 μl of lysis buffer (1% NP-40 in 20 mM EDTA, 50 mM Tris-HCl, pH 7.5), centrifuged at 3000 rpm for 5 min and supernatants were collected. SDS (1%) and RNase A (5 μg/ml) were added for 1 h at 56°C and proteinase K (2.5 μg/ml) was added for 1 h at 37°C. Afterward, ½ volume of ammonium acetate (stock 10 M), 2 volume of ice cold ethanol were added, followed by incubation overnight at −80°C. Samples were centrifuged the next day (14000 rpm) for 40 min at 4°C. Pellets were washed with 70% ethanol and dissolved in 20 μl of water. Equal amounts of DNA were loaded for each condition on a 2% agarose gel.

qPCR

RNA was isolated from 2×10^6 cells with the RNeasy Plus Mini kit (QIAGEN) and reversely transcribed with SuperScript III (Invitrogen). cDNA was amplified with SYBR Green on a Biorad C1000 Thermal Cycler. Primers used were CCL2 (CCCCAGTCACCTGCTGTTAT) and GAPDH (GTCTCCTCTGACTTCAACAGCG). Fold induction compared to untreated controls was calculated by the delta-delta CT method.

Chromatography and mass spectrometry

Samples were loaded onto 50-cm columns packed in-house with C18 1.9 μM ReproSil particles (Dr Maisch GmbH), with an EASY-nLC 1000 system (Thermo Fisher Scientific) coupled to the MS (Q Exactive HF, Thermo Fisher Scientific). A homemade column oven maintained column temperature at 60°C. Peptides were introduced onto the column with buffer A (0.1% Formic acid) and eluted with a 107.3-min gradient of 5 to 25% of buffer B (80% ACN, 0.1% Formic acid), both at a flow rate of 300 nl/min.

A data-dependent acquisition (TopN) MS method was used in which one full scan (300 to 1650 m/z, R = 60,000 at 200 m/z) at a target of 3×10^6 ions was first performed, followed by 15 data-dependent MS/MS scans with higher energy collisional dissociation [target 10^5 ions, max ion fill time 55 ms, isolation window 1.4 m/z, normalized collision energy 27%, R = 15,000 at 200 m/z]. Dynamic exclusion of 20 s and apex trigger (4 to 7 s) was enabled.

QUANTIFICATION AND STATISTICAL ANALYSIS

MS raw files were processed by the MaxQuant software version 1.5.0.38 (Cox and Mann, 2008) and fragments lists were searched against the human Uniprot Reference Proteome without isoforms (August 2015) by the Andromeda search engine (Cox et al., 2011) with cysteine carbamidomethylation as a fixed modification and N-terminal acetylation and methionine oxidations as variable modifications. We set the false discovery rate (FDR) to 0.01 at the peptide and protein levels and specified a minimum length of 7 amino acids for peptides. Enzyme specificity was set as C-terminal to Arginine and Lysine as expected using Trypsin and LysC as proteases and set as semi-tryptic for semi-tryptic peptide analysis, and a maximum of two missed cleavages.

All bioinformatics analyzes were done with the Perseus software (version 1.5.3.0) (Tyanova et al., 2016) of the MaxQuant computational platform. Quantified proteins are filtered for at least 75% of valid values among three or four biological replicates in at least one condition. Missing values were imputed and significantly up- or downregulated proteins were determined by multiple-sample test (FDR = 0.05) and Student's t test (two-sided), (FDR = 0.05). n represents replicates of the same cell line stimulated separately. Further statistical details of experiments can be found in the figure legends.

The 1D annotation enrichment analysis detects whether expression values of proteins belonging to an enrichment term (here we used: keywords, GOCC, GOMF, GOBP and KEGG name) show a systematic enrichment or de-enrichment compared to the distribution of all expression values (Cox and Mann, 2012).

Swarm plots were created using the software Instant Clue (Nolte et al., 2018).

DATA AND CODE AVAILABILITY

The MS-based proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository and are available via ProteomeXchange with identifier (PXD014966) (Jones et al., 2008).

4.5 Project 5: Proteomics reveals distinct mechanisms regulating the release of cytokines and alarmins during pyroptosis

Injured cells emit danger signals, warning the host of impending cell death. Major intercellular signaling molecules such as cytokines and alarmins activate neighboring cells and initiate inflammatory responses. Pyroptosis is the most proinflammatory form of programmed cell death; yet, aside from interleukins (Il)1, Il18, and the alarmin Hmgb 1, little is known about the production of proinflammatory proteins during this process. As pyroptosis needs the activation of inflammasomes via two signals, the composition of the released pro-inflammatory cocktail of proteins is extremely complex. It includes conventionally and unconventionally secreted proteins, as well as proteins released passively as a result of cell death. Proteins depart cells by a variety of pathways, including the plasma membrane's Gasdermin pores, cell lysis, the ER-Golgi pathway, extracellular vesicles, and receptor shedding.

Using our secretomics workflows, we developed an experimental strategy to address two long-standing questions: 1) which proteins are released during pyroptosis, and 2) via which cellular exit pathways are these proteins released. The critical conceptual contribution we make is to combine pharmacological, biochemical, and genetic methods with mass spectrometry-based secretomics to interrogate distinct phases of inflammasome activation and protein release, allowing us to deconstruct this complicated secretory program molecularly.

Our global and temporal resolved research demonstrates that the majority of proteins are released passively as a result of cell lysis. Specifically, low molecular weight proteins such as Il1b, alarmins, lysosomal proteases, and mitochondrial proteins are released by Gasdermin pores regardless of cell lysis. For the first time, we identify the release of several cytokines and alarmins, including members of the Mif, Aimp1, S100, and Galectin families that have not previously been associated with pyroptosis. Our investigation surprisingly found a protein exit route that is ER-Golgi reliant, but not extracellular vesicle dependent.

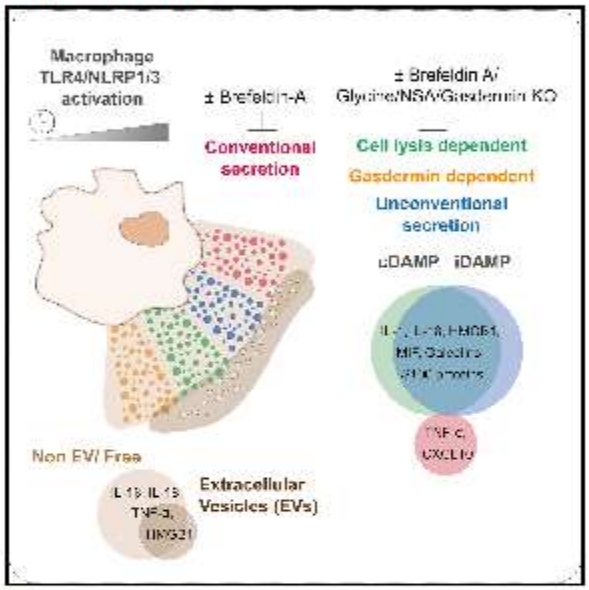
The project was a group-internal collaboration with Kshiti Phulphagar. The corresponding article was published in Cell Reports in 2021. I have contributed the shedding analysis.

Cell Reports

Resource

Proteomics reveals distinct mechanisms regulating the release of cytokines and alarmins during pyroptosis

Graphical abstract



Authors

Kshiti Phulphagar, Lars I. Kühn, Stefan Ebner, Annika Frauenstein, Jonathan J. Swietlik, Jan Rieckmann, Felix Meissner

Correspondence

meissner@biochem.mpg.de

In brief

Using mass-spectrometry-based proteomics in combination with pharmacological and genetic perturbation, Phulphagar et al. dissect protein release during macrophage pyroptosis. The study defines proteins exiting cells through active secretion pathways, extracellular vesicles, organellar damage, and gasdermin pores and identifies the gasdermin-mediated release of multiple alarmins.

Highlights

- Mass spectrometry resource for inflammasome-activated protein release in macrophages
- Inhibition of inflammasome assembly and protein release by ER-Golgi disruption
- Dissection of protein release by gasdermin, organelle damage, and extracellular vesicles
- Identification of released constitutive and induced alarmins

Phulphagar et al., 2021, Cell Reports 34, 108826
March 9, 2021 © 2021 The Author(s).
https://doi.org/10.1016/j.celrep.2021.108826

Cell Reports

CellPress
OPEN ACCESS

Resource

Proteomics reveals distinct mechanisms regulating the release of cytokines and alarmins during pyroptosis

Kshiti Phulphagar,[1,2,3] Lars I. Kühn,[1,3] Stefan Ebner,[1] Annika Frauenstein,[1] Jonathan J. Swietlik,[1] Jan Rieckmann,[1] and Felix Meissner[1,2,4,*]

[1]Experimental Systems Immunology Laboratory, Max Planck Institute of Biochemistry, Martinsried, Germany
[2]Institute of Innate Immunity, Department of Systems Immunology and Proteomics, Medical Faculty, University of Bonn, Bonn, Germany
[3]These authors contributed equally
[4]Lead contact
*Correspondence: meissner@biochem.mpg.de
https://doi.org/10.1016/j.celrep.2021.108826

SUMMARY

A major pathway for proinflammatory protein release by macrophages is inflammasome-mediated pyroptotic cell death. As conventional secretion, unconventional secretion, and cell death are executed simultaneously, however, the cellular mechanisms regulating this complex paracrine program remain incompletely understood. Here, we devise a quantitative proteomics strategy to define the cellular exit route for each protein by pharmacological and genetic dissection of cellular checkpoints regulating protein release. We report the release of hundreds of proteins during pyroptosis, predominantly due to cell lysis. They comprise constitutively expressed and transcriptionally induced proteins derived from the cytoplasm and specific intracellular organelles. Many low-molecular-weight proteins including the cytokine interleukin-1β, alarmins, and lysosomal-cargo proteins exit cells in the absence of cell lysis. Cytokines and alarmins are released in an endoplasmic reticulum (ER)-Golgi-dependent manner as free proteins rather than by extracellular vesicles. Our work provides an experimental framework for the dissection of cellular exit pathways and a resource for pyroptotic protein release.

INTRODUCTION

Secreted proteins such as cytokines, chemokines, and growth factors have various functions in the initiation and resolution of inflammation and the regulation of effector cells (Arango Duque and Descoteaux, 2014). A major pathway contributing to protein secretion in macrophages involves a highly pro-inflammatory form of cell death, known as pyroptosis. Pyroptosis plays a central role in the regulation of immune responses during infection and tissue injury (Bergsbaken et al., 2009). Efficient execution of pyroptosis requires the activation of intracellular signaling complexes termed inflammasomes and is regulated tightly in two steps.

Inflammasomes are comprised of activated cytosolic receptors such as certain nucleotide-binding oligomerization domain-like receptors (NLRs), an adaptor protein ASC, and a downstream effector protease known as Caspase-1.

For activation, a first signal induces transcription and expression of some inflammasome proteins such as the NLR family pyrin domain containing protein 3 (NLRP3), whereas some other NLRs such as NLRP1 are expressed constitutively. *In vitro*, this "priming" signal is commonly provided by ligation of innate immune receptors such as Toll-like receptor 4 (TLR4) through the Gram-negative bacterial cell membrane component lipopolysaccharide (LPS) (Broz and Dixit, 2016). LPS also induces numerous cytokines and chemokines including interleukin 1 beta (IL-1β), tumor necrosis factor alpha (TNF-α), and C-X-C motif chemokine ligand 10 (CXCL10), of which most are directly released conventionally through the endoplasmic reticulum (ER)-Golgi secretory pathway (Kaisho and Akira, 2000; Stow et al., 2009).

A second signal triggers inflammasome assembly through the activation of NLRs. Low cytosolic potassium levels activate NLRP3, whereas removal of the N terminus activates NLRP1 (Mitchell et al., 2019; Swanson et al., 2019). NLRs can then recruit the adaptor protein ASC and activate pro-inflammatory caspases such as caspase-1 that cleaves its substrates IL-1β, IL-18, and the gasdermin protein family member Gasdermin-D (GSDMD) into biologically active forms. The N-terminal fragment of GSDMD forms pores in the plasma membrane (PM) and eventually causes cell lysis (Ding et al., 2016; Kayagaki et al., 2015; Liu et al., 2016). GSDMD thereby facilitates the release of proteins lacking a signal peptide such as the cytokines IL-1β and IL-18 and proteins or "alarmins" derived from damaged cellular organelles such as high mobility group box 1 protein HMGB1, which also exhibits paracrine activity (Evavold et al., 2018; Heilig et al., 2018; Lamkanfi et al., 2010). As this protein export pathway does not involve active protein transport through the secretory pathway, it is broadly referred to as "unconventional secretion"

(Lopez-Castejon and Brough, 2011; Ng and Tang, 2016; Nickel and Rabouille, 2009).

Proteins released during inflammasome-activated pyroptosis therefore comprise conventionally (e.g., TNF-α), and unconventionally (e.g., IL-1β) released cytokines as well as alarmins (e.g., HMGB1) and other proteins leaking from the cytosol and damaged intracellular organelles. Unlike conventionally secreted proteins, which can be computationally predicted based on the primary amino acid sequence of their signal peptide, unconventionally released proteins can be identified with high confidence only experimentally by assessment of their release and the accompanied accumulation in cellular supernatants (Nickel and Rabouille, 2009). As IL-1- and IL-18-independent effects of inflammasome-mediated inflammatory programs have been described in different pathologies, the presence of novel immunomodulatory proteins such as alarmins can be hypothesized to be released during pyroptosis (Gong et al., 2020; Lammert et al., 2020).

In addition to translocation through GSDMD pores or cell lysis, various other export mechanisms have also been implicated in unconventional secretion such as secretion by lysosomes, extracellular vesicles (EVs), autophagosomes, and PM-derived microvesicles (Andrei et al., 1999; Baroja-Mazo et al., 2013; Brough et al., 2017; Dupont et al., 2011; Lopez-Castejon and Brough, 2011; Martin-Sánchez et al., 2016; Monteleone et al., 2015; Rubartelli et al., 1990; Tapia et al., 2019; Zhang et al., 2015). The identification of the predominant mechanisms of release for each protein can therefore enable the development of strategies to specifically interfere with distinct export pathways by abrogating the release of defined sets of proteins and thereby altering responses of neighboring cells or the entire organism to pyroptosis.

Mass spectrometry (MS)-based proteomics has matured remarkably over the last years and now provides a powerful set of tools for the systematic and unsupervised examinations of cellular processes (Aebersold and Mann, 2016; Larance and Lamond, 2015). Building on our recently developed MS-based approaches for the comprehensive analysis of intercellular immune signals, we set out to devise an experimental strategy to determine the inflammatory program of pyroptotic macrophages in its entirety and define the exit route for each protein (Meissner et al., 2013; Rieckmann et al., 2017; Tanzer et al., 2020). Our strategy comprises genetic knockouts (KOs) and knockins of inflammasome proteins to specifically inhibit or provide only one of the two signals required for full inflammasome activation. We pharmacologically interfere with ER-Golgi protein transport or passive cell lysis and assess the contribution of membrane-enclosed vesicles and receptor shedding to pyroptotic protein release, with the latter having implications in necroptosis—a distinct form of inflammatory cell death (Cai et al., 2016; Tanzer et al., 2020). Thereby, we determine for each protein whether transcription, caspase activation, active secretion, organellar damage, pore formation, or passive cell lysis is required for its release during pyroptosis.

RESULTS

Quantitative proteomics accurately determines protein release

We activated conventional protein secretion through one signal (TLR4 with LPS) and unconventional secretion through two signals (TLR4 with LPS plus NLRP3 with ATP or Nigericin) in murine bone-marrow-derived and human-monocyte-derived macrophages (BMDMs and HMDMs, respectively) and analyzed the secretory programs by label-free, high-resolution MS in a single-shot liquid chromatography-tandem MS (LC-MS/MS) format by using a Quadrupole-Orbitrap instrument (Figures 1A, S1A, and S1B; Kelstrup et al., 2018; Scheltema et al., 2014). Fold changes of proteins released in mouse and human macrophages upon TLR4 or TLR4 plus NLRP3 activation compared to untreated controls are indicated in Table S1. We identified on average almost 3,000 protein groups in cellular supernatants at a peptide and protein false discovery rate (FDR) of 1% (Figures S2A and 2B). We compared MS- to antibody-based (ELISA) protein quantification using TNF-α and IL-1β as reference proteins for conventional and unconventional secretion pathways (Figures 1B–1G and S2G). Both quantification methods showed similar secretion profiles for the two proteins (Figures 1B–1G), with an average correlation of 0.7 for TNF-α (Figure S2G) and 0.95 for IL-1β (Figure S2G). As expected, TNF-α was secreted after TLR4 activation regardless of the presence of a second activation signal (Figures 1B and 1E). In contrast, IL-1β was only secreted after priming the cells through TLR4, followed by activation of NLRP3 (Figures 1C and 1F). We further determined cell death by lactate dehydrogenase (LDH) release (Figure 1D) as well as by the summed MS signal of all proteins annotated as cytosolic (Figure 1G) and obtained a correlation of 0.85 between the two methods (Figure S2G). Our analysis shows that our method accurately quantifies conventional and unconventional protein secretion as well as cell death in activated macrophages.

Kinetics of conventional and unconventional protein release

To define pyroptotically released proteins, we assigned proteins to either (1) conventionally secreted or (2) unconventionally secreted proteins by comparing kinetics of secretomes from different experimental conditions (Figure 1).

First, we defined conventionally secreted proteins by comparing the secretome of TLR4-activated cells to unstimulated cells (Figure 1H). We identified 36 significantly released proteins, of which 30 exhibit a signal peptide and signal anchor. Among these proteins are known conventionally secreted, inflammatory mediators including cytokines (e.g., TNF-α and IL-6) and chemokines (e.g., CXCL10), confirming results of published work (Meissner et al., 2013). Functional enrichment analysis using annotations from the UniProt Keywords database revealed common terms of conventionally released inflammatory mediators such as "Cytokine," and "Inflammatory response" (Figure S3A).

Second, we defined unconventionally secreted proteins by the combined activation of TLR4 and NLRP3. Under these conditions, cellular supernatants contain conventionally secreted proteins due to the TLR4 priming signal and unconventionally released proteins comprising of proteins released by GSDMD pores or due to cell lysis (Figure 1I). This group of proteins contained prototypic unconventionally secreted proteins such as IL-1β, IL-1α, IL-18, and HMGB1 but excluded prototypic conventionally exported proteins like TNF-α and Cxcl10 (Figure 1I). Enrichment analysis of unconventionally secreted proteins using UniProt Keywords and Pfam protein domains unveiled the presence of diverse protein classes

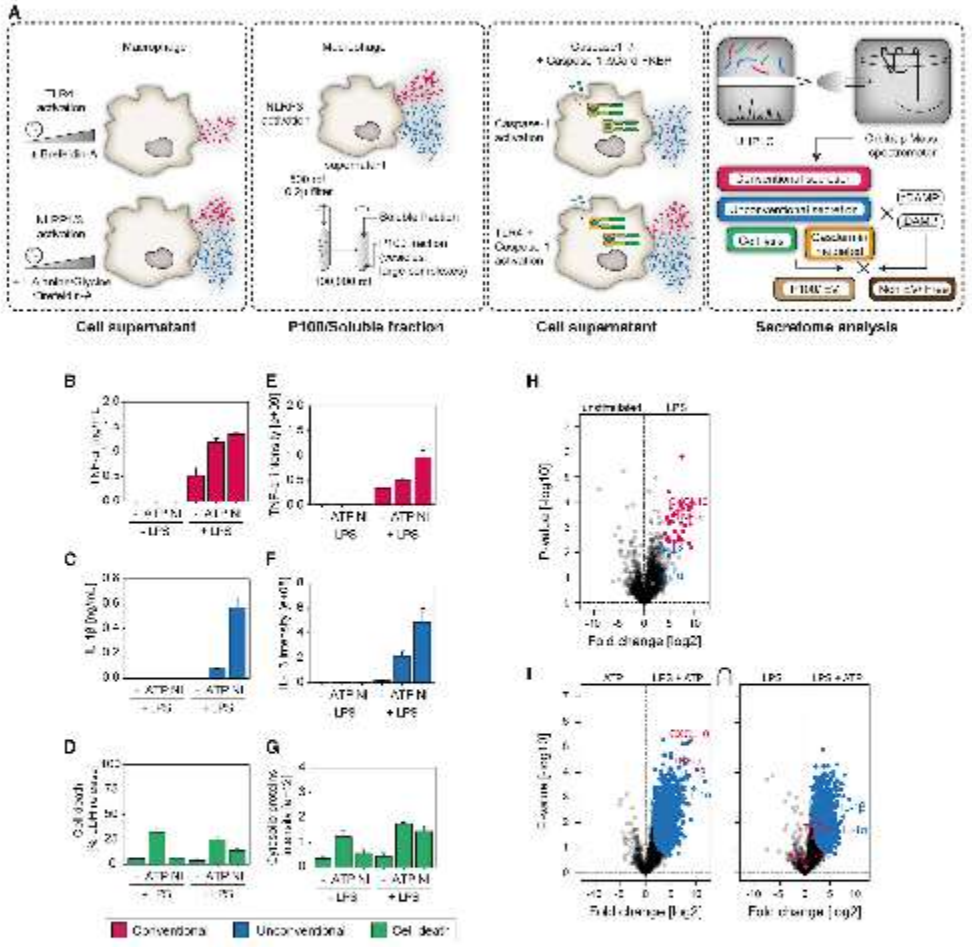


Figure 1. Quantitative proteomics workflow to identify proteins released by macrophage pyroptosis

(A) Schematic illustration of the experimental approach to distinguish between different modalities of protein release from cells, including conventional, unconventional, cell lysis dependent and independent, soluble or extracellular vesicle associated, dissection of released induced or constitutive alarmins, or damage-associated molecular patterns (DAMPs). Workflow consisting of macrophage activation, collection and processing of supernatants, sample preparation, LC-MS/MS analysis, and data processing.

(B–D) Quantification of TNF-α (B) and IL-1β (C) secretion by ELISA or cell death by LDH release (D) after TLR4 and NLRP3 activation in mouse macrophages. Data are represented as mean ± SD of three biological replicates.

(E–G) MS raw intensity-based quantification after TLR4 and NLRP3 activation in mouse macrophages of TNF-α (E) and IL-1β (F) secretion and cell death (G) measured by the summed MS raw intensity of all proteins annotated as "cytoplasm," "cytoplasmic part," "intracellular," or "intracellular part" by GOCC. Data are represented as mean ± SD of three biological replicates.

(H and I) Volcano plots showing secreted proteins from differentially activated BMDMs. Pairwise secretome comparisons to define conventionally (H) and unconventionally (I) released proteins. (H) Conventionally released proteins upon TLR4 activation compared to untreated controls (magenta).

(I) Unconventionally released proteins upon TLR4 plus NLRP3 compared to TLR4 as well as to only NLRP3-activated cells (blue). Significant proteins were determined with a parametric two-tailed Welch's t test (FDR < 0.05, $S_0 = 1$).

See also Figures S1–S4 and Table S1.

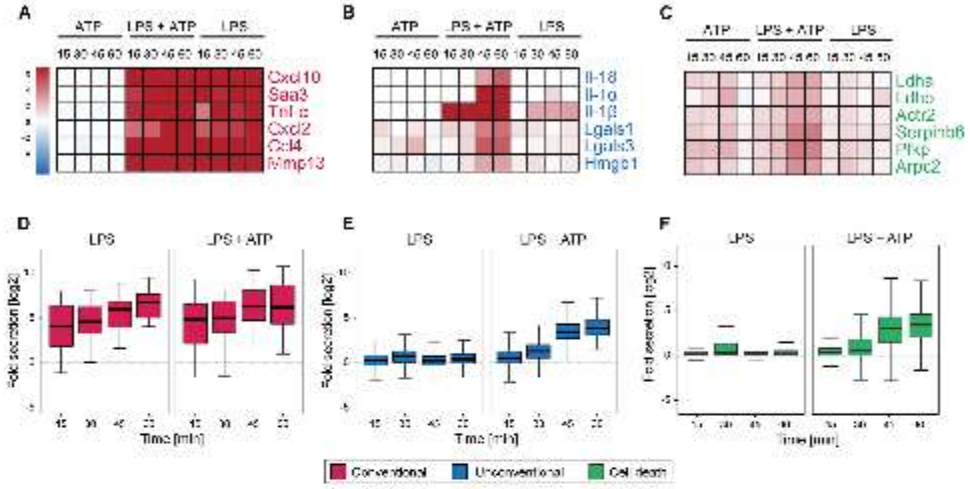


Figure 2. Proteomic analysis of pyroptosis-mediated conventional and unconventional protein secretion

(A–C) Heatmaps with secretion profiles of selected proteins known to be released via conventional unconventional programs.

(A) Cytokines containing a signal peptide and significantly released from cells by conventional secretion upon TLR4 activation with LPS.

(B) Cytokines/alarmins without a signal peptide significantly released from cells by unconventional secretion upon pyroptosis induction by TLR4 and NLRP3 activation with LPS and ATP.

(C) Selected cytoplasmic proteins, including subunits of the LDH complex.

(D–F) Fold secretion over time of conventional (D; magenta), unconventional (E; blue), or proteins annotated as "cytoplasm," "cytoplasmic part," "intracellular," or "intracellular part" by GOCC (F; green) by BMDMs activated for TLR4 with LPS and TLR4 plus ATP for NLRP3.

See also Figures S1, S2, and S4 and Table S1.

and structural domains such as Cofilin-actin de-polymerization factor (Cofilin-ADF) and Lin11, Isl-1 & Mec-3 (LIM), ADP ribosylation factor (ARF), and the caspase activation and recruitment domain (CARD), which is present for example in caspase-1 and ASC, two proteins involved in the formation of the NLRP3 inflammasome (Figure S3B). The abundance of conventionally secreted proteins increased gradually over time in response to TLR4 activation (Figures 2A and 2D), whereas unconventional protein release increased steadily in response only to TLR4 plus NLRP3 and was detectable already at 15 min after activation with a peak at 45–60 min (Figures 2B and 2E). We detected IL-1 release early upon inflammasome activation at 15 min, whereas other pyroptotically released cytokines and damage-associated molecular patterns (DAMPs) like IL-18 and HMGB1, respectively, with a similar size as IL-1β peaked later at 30–45 min (Figures 2B and 2E; Table S1).

Inhibition of NLRP3 with its small-molecule antagonist glyburide or KO of the downstream effector Caspase-1 abrogated the unconventional secretion completely (Figures S3D, S3E, and S4A–S4C; Swanson et al., 2019). Despite multiple comparisons, thousands of proteins were identified to be released with similar kinetics to IL1 and IL-18, impairing the discovery of potential novel alarmins. We therefore set out to define protein release pathways by assessing the role of the ER-Golgi compartment, extracellular vesicles, damaged intracellular organelles, PM pores, or cell lysis during pyroptosis.

The ER-Golgi compartment is required for inflammasome formation and unconventional protein release

Although the role of the ER-Golgi route conventional secretion has been well described, its contribution to unconventionally released cytokines such as IL-1β remains controversial and incompletely understood (Rubartelli et al., 1990; Zhang et al., 2017). We used the fungal metabolite Brefeldin A (BFA) as well as the small molecule Golgicide A (GCA), of which both prevent protein transport from the ER to the Golgi and disrupt the function of ER-Golgi secretory pathway (Chardin and McCormick, 1999; Sáenz et al., 2009). BFA and GCA perturb ER-Golgi trafficking by specifically inhibiting the guanine nucleotide exchange factors (GEFs) that regulate the formation of secretory vesicles. Although BFA targets several GEFs, including Golgi-specific BFA resistance factor 1 (GBF1), BFA-inhibited GEF 1 (BIG1), and BIG2, GCA is highly specific toward GBF1 (Sáenz et al., 2009). As expected, treatment of LPS-stimulated BMDMs with BFA inhibited the secretion of TNF-α and all other proteins we defined as conventionally secreted in a dose-dependent manner (Figures 3A and 3D). BFA also inhibited NLRP3-mediated release of IL-1β and cell death (Figures 3B, 3C, and 3E), in keeping with previous reports with a similar dose dependency (Zhang et al., 2017). However, we also observed a substantial reduction in IL-1β secretion by BFA treatment for inflammasome sensors, including baculoviral inhibitor of apoptosis domain repeat-

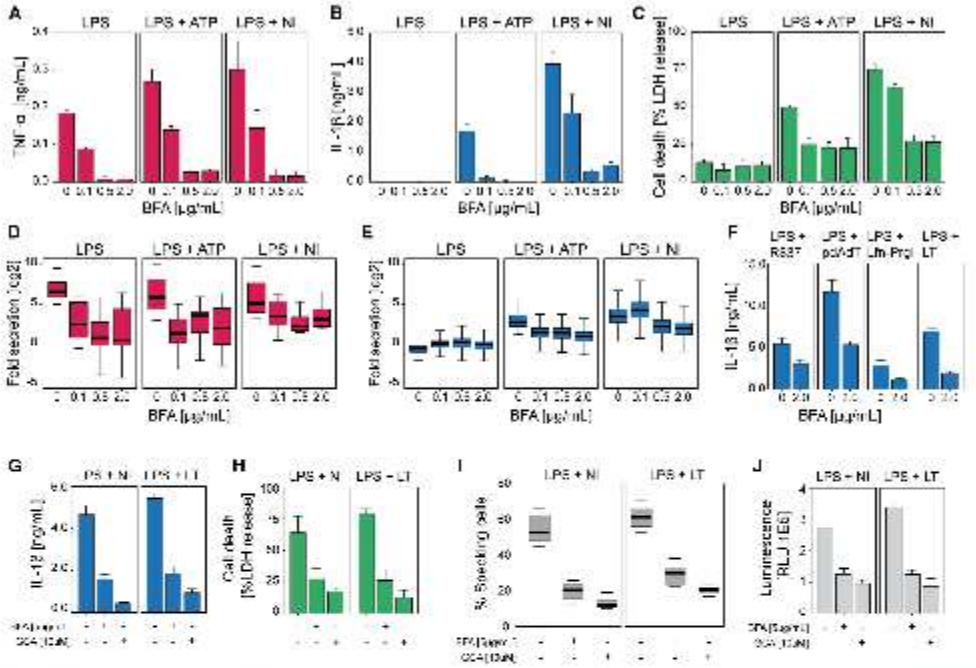


Figure 3. Conventional secretion, inflammasome formation, and unconventional secretion are inhibited by pharmacological ER-Golgi disruption

(A–C) ELISA of TNF-α (A), IL-1β (B), and LDH released (C) in supernatants of TLR4- or NLRP3-activated BMDMs in the presence of the indicated concentrations of BFA. Nl, Nigericin

(D and E) Fold secretion of conventionally (D; magenta) and unconventionally (E; blue) released proteins in TLR4- and TLR4 plus NLRP3-activated BMDMs in the presence of the indicated concentrations of BFA.

(F) IL-1β levels as determined by ELISA in supernatants of NLRP3-, AIM2-NLRC4-, and NLRP1-activated BMDMs in the presence of the indicated concentrations of BFA.

(G–H) IL-1β levels as determined by ELISA (G) and LDH release (H) in supernatants of NLRP3- and NLRP1-activated BMDMs in the presence of the indicated concentrations of BFA or GCA.

(I–J) Quantification of NLRP3 or NLRP1 inflammasome assembly in the presence or absence of the indicated concentrations or BFA or GCA by evaluating Asc speck formation by using fluorescent microscopy. Data are compiled from three biological replicates, each with 100–200 quantified cells.

(J) Measurement of Caspase-1 activity by luminescence, in which aminoluciferin generated by cleavage of its substrate Z-WEHD-aminoluciferin results in light production in the presence of the luciferase enzyme. Data are represented as mean ± SD of three biological replicates.

See also Figures S4 and S5 and Table S1.

containing protein 1 (NAIP1) activated with its ligand the bacterial type III secretion system component PrgI and absent in melanoma 2 protein (AIM2) activation with repetitive synthetic double-stranded DNA sequence of poly(dA-dT) (Figure 3F; Broz and Dixit, 2016). GCA induced a similar reduction in IL-1β secretion and cell death in BMDMs activated for both NLRP1 and NLRP3 (Figures 3G and 3H). Notably, we treated cells with BFA or GCA after the first priming signal, suggesting that inhibition occurs at the level of the inflammasome sensor or the execution of pyroptosis but not TLR4. Interestingly, both BFA and GCA substantially reduced Asc speck formation as well as Caspase-1 activation (Figures 3I, 3J, and S5A–S5C). Thus, our data show that ER-Golgi perturbation with BFA and GCA inhibits unconventional protein release by interfering with inflammasome formation, regardless of the tested inflammasome sensor.

Leakage of specific cellular organelles during pyroptosis

We next asked whether pyroptosis leads to the release of subcellular organelles and their contents into the extracellular environment. To assess the leakage of proteins from organelles, we inspected the subcellular origin of all significantly released proteins. We observed an enrichment of cytosolic and lysosomal proteins and, to a lesser degree, endosomal, mitochondrial,

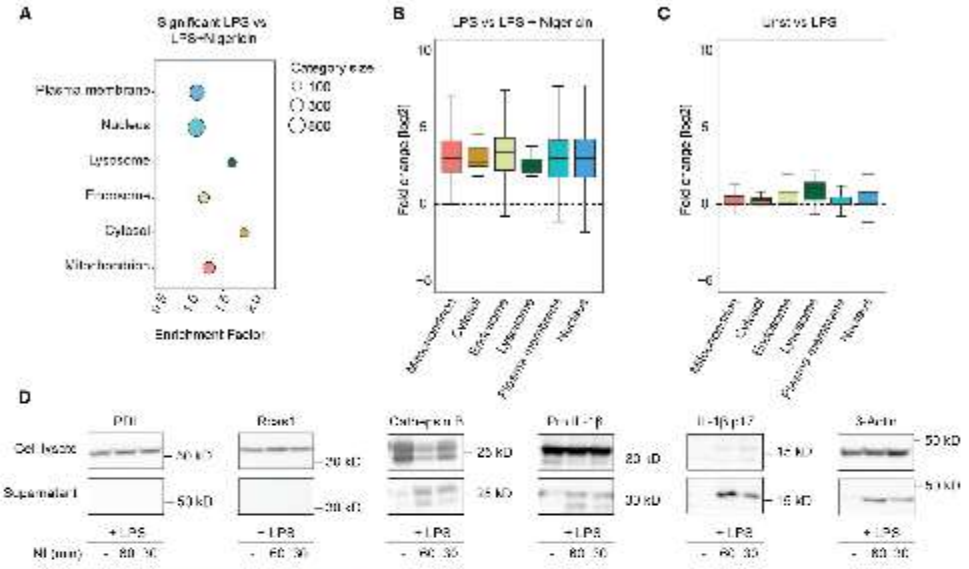


Figure 4. Extracellular leakage of organellar proteins

(A) Balloon plot of organelle marker enrichment analysis on the unconventional secretion signature after TLR4 and NLRP3 activation in mouse macrophages by using GOCC slim terms. Enrichment factors were calculated by Fisher's exact test (annotations with intersection sizes of > 15 and $p < 0.002$ are shown). Category sizes are represented by balloon size.

(B and C) Boxplots indicating the fold changes of all detected proteins belonging to indicated organelles in the unconventional (B) and conventional (C) secretomes.

(D) Western blots of cell lysates or supernatants after TLR4 or TLR4 and NLRP3 activation for ER, Golgi, and lysosomal markers PDI, Roast, and Cathepsin B, respectively; or pro IL-1β, IL-1β p17 fragment, and the housekeeping protein β-Actin. Immunoblotting data are representative of two biological replicates. See also Figure S3.

nuclear, and PM proteins (Figures 4A and 4B). Interestingly, proteins from the ER and Golgi were not enriched. In agreement with our MS analysis, immunoblotting confirmed the release of the luminal lysosomal marker Cathepsin B and the absence of ER lumen marker protein disulfide isomerase (PDI) and Golgi marker receptor binding cancer antigen expressed on SiSo cells (RCAS1) in pyroptotic supernatants (Figure 4D). Our analysis therefore suggests that pyroptosis induces the release of proteins from specific subcellular localizations into the extracellular environment, whereas other organelles are retained inside the cell corpse.

We and others have described the shedding of receptors during programmed cell death by necroptosis (Tanzer et al., 2020). To determine whether pyroptosis is also accompanied by receptor shedding, we assessed whether cell surface receptors are enriched in pyroptotic cell supernatants. We detected 62 receptors, of which 15 were significantly released. However, examination of the quantified peptides revealed that both cytoplasmic and extracellular domains of these receptors were released to a comparable extent (Figure S5A). These results suggest the absence of an active shedding process involving proteolytic cleavage of receptors, as this would be accompanied by an increase of the extracellular domains of receptors and not the corresponding cytoplasmic domain. We therefore surmise that the presence of receptors and other PM-localized proteins in pyroptotic supernatants can be ascribed primarily to the release of PM fragments or micro-vesicles.

Most cytokines are secreted in a soluble form and not in EVs

EVs have been implicated in programmed cell death and also in NLRP3-mediated release of both pro-form and mature IL-1β and IL-18 (Cypryk et al., 2018). To determine the contribution of this protein release route, we enriched EVs from LPS-activated or pyroptotic supernatants by using differential centrifugation as described previously (Lobb et al., 2015; Table S1). The majority of cytokines including IL-1β and IL-18 were not detected in the EV-enriched P100 fraction, indicating that they are predominantly secreted in a non-EV-bound, soluble form (Figures 5A–5C). Some cytokines like TNF-α and PF4 as well alarmins like HMGB1 were, however, also present in the P100 fraction, suggesting that they might be released in soluble forms as well as in membrane-associated or EV-incorporated forms (Figure 5B). As both, low- and high-abundant cytokines were equally

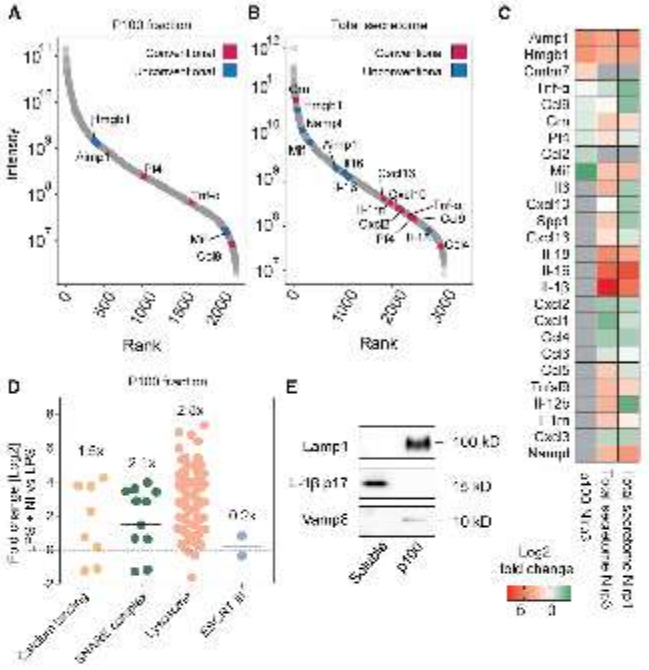


Figure 5. Majority of cytokines are released as free proteins and not in EVs

(A and B) Intensity-based ranking of cytokines identified in either P100 (A) or total secretome (B) of NLRP3-activated murine macrophages.

(C) Heatmap of fold secretion of conventionally and unconventionally released cytokines by murine macrophages in the total secretome after NLRP1/3 activation or P100/EV-associated secretome after NLRP3 activation.

(C) Swarm plots showing log2 fold change of proteins assigned to the annotation "Calcium/phospholipid binding," "SNARE complex," "Lysosome" and "ESCRTIII complex" present in P100/EV fraction of pyroptotic cells compared with TLR4-activated controls.

(D) Presence of lysosomal marker Lamp1, IL-1βp17 fragment, and the SNARE complex member Vamp8 in either the soluble or P100 fraction after NLRP3 activation evaluated by immunoblotting.

See also Figures S1, S3, and S6 and Table S1.

detected in the P100 fraction, our analysis is not biased toward the detection of high-abundant cytokines (Figure 5B). We conclude that secretory processes, in which cytokines and alarmins are packaged into vesicles and released in such a membrane-encapsulated form, contribute only minimally to their release during pyroptosis.

We also observed an enrichment of proteins in the categories "Lysosome," "SNARE complex," and "Calcium/PI binding," which we had also detected previously in vesicles released by cells undergoing necroptosis—another form of lytic cell death (Figure 5D; Tanzer et al., 2020). Both membrane as well as luminal lysosomal proteins were enriched to a comparable degree in the P100 fraction and secretome (Figure S3F). Conversely, cytosolic proteins were largely depleted in this fraction (Figures 5B–5D). Our data therefore suggest that endosomal and lysosomal proteins are released in a vesicular form during pyroptosis. We confirmed the absence of IL-1β or the presence of the lysosomal marker LAMP1 and SNARE complex member VAMP8 in the P100 fraction by immunoblotting (Figure 5E).

Unconventional release of low-MW, lysosomal, and mitochondrial proteins proceeds by a GSDMD-mediated but cell-lysis-independent pathway

Multiple recent studies indicate that during pyroptosis, low-molecular-weight (MW) proteins including IL-1β are released through PM GSDMD pores independently of late-stage cell lysis (DiPeso et al., 2017; Evavold et al., 2018; Heilig et al., 2018). To assess the contribution of GSMDM pores versus cell lysis to protein release, we induced pyroptosis in the presence of osmoprotectants L-alanine or glycine, which have been shown to block cell lysis but not the formation of GSDMD pores (Fink and Cookson, 2006; Loomis et al., 2019). Both L-alanine and glycine inhibited cell lysis that was measured by the release of the LDH complex of approximately 147 kD but not IL-1β (17 kD) or any conventionally secreted cytokine like TNF-α (26 kD), in keeping with previous reports (Figure S4D). We further evaluated protein release in GSDMD KO iBMDMs or wild-type (WT) BMDMs treated with the GSDMD inhibitor necrosulfonamide A (NSA), confirming that cell lysis as well as IL-1β release was inhibited, whereas conventional secretion of TNF-α was not (Figures S4E and S4F; Evavold et al., 2018; Heilig et al., 2018; Rathkey et al., 2018). At the global level, the correlation between all proteins released unconventionally versus the subset released in an osmoprotection-independent manner was mediocre (0.36) (Figures 6A–6C), whereas conventionally versus unconventionally released proteins showed a negative correlation (−0.16), indicating the independence of these two protein export pathways (Figures 6B and 6C).

Our proteomics experiments reveal that sensitivity to osmoprotection depends on the MW of a protein, as the release of low-MW proteins was inhibited far less than high-MW proteins (Figure 6D; Table S1). Osmoprotection barely reduced the release of low-MW proteins, such as IL-1β, the alarmins Galectin 3 (LGALS3) and HMGB1, or cytosolic proteins such as glutaredoxin 5 (GLRX5), whereas GSDMD KO or inhibition completely blocked the release of low- as well as high-MW proteins (Figures 6E and 6F). Surprisingly, a Gene Ontology term cellular component (GOCC) enrichment analysis revealed that proteins derived from lysosomes and, to a lesser extent, from mitochondria

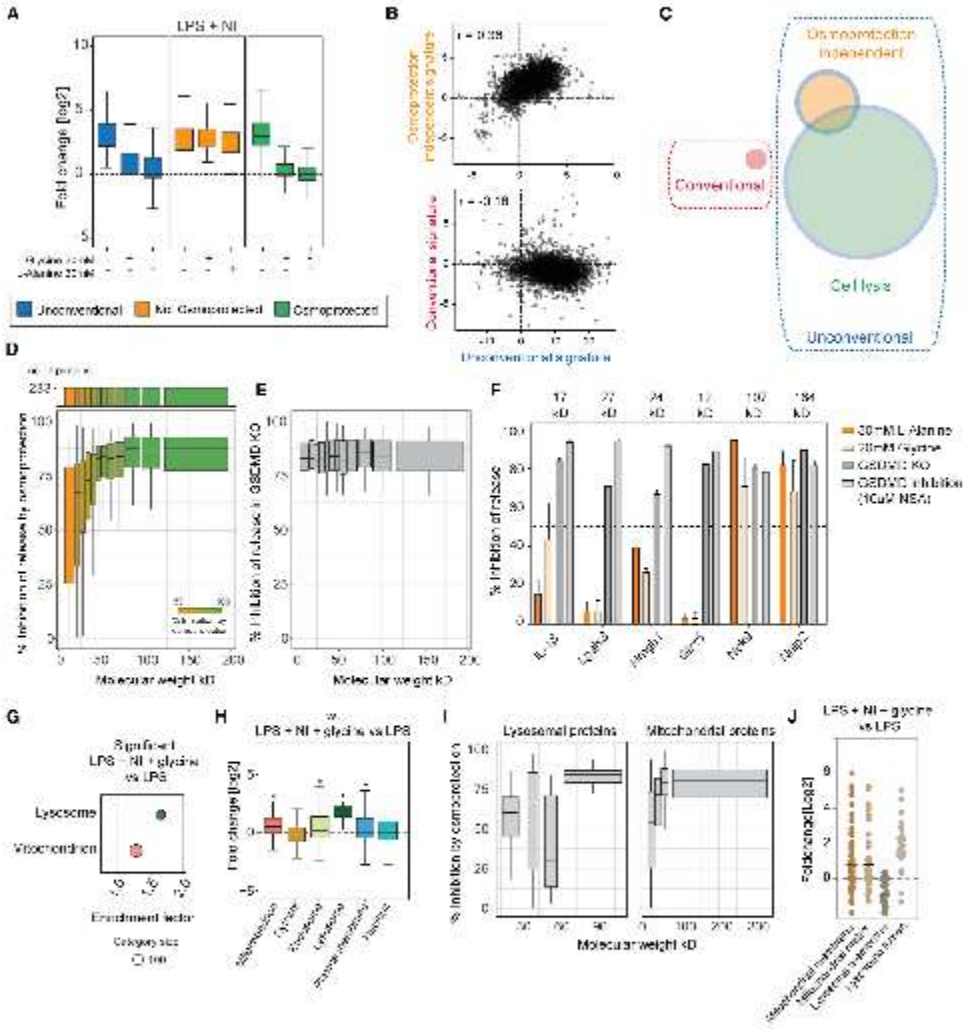


Figure 6. Release of low-MW, mitochondrial, and lysosomal proteins is dependent on GSDMD but proceeds independently of cell lysis

(A) Fold secretion of all unconventional, osmoprotection-independent (orange), and cell-lysis-dependent (green) protein release in TLR4-plus NLRP3-activated BMDMs in the presence of the indicated concentrations of L-Alanine or Glycine.

(B) Correlations of t test differences used to define the conventional, unconventional, osmoprotection, or cell lysis-independent and -dependent signatures. Conventional signature, LPS versus unstimulated; unconventional signature, t test differences LPS + Nl versus LPS; and osmoprotection/cell lysis independent, LPS + Nl + 30 mM L-alanine/20 mM glycine versus LPS.

(C) Venn diagram of proteins released conventionally, unconventionally, and osmoprotection independently or passively by cell death.

(D–F) Effect of osmoprotection or knockout of GSDMD on unconventionally released proteins with different MWs. Boxplot indicating the extent of inhibition by osmoprotection (D) or GSDMD knockout (E) to unconventionally released proteins binned according to increasing MW. Number of proteins per bin are indicated. Bar plots of selected proteins indicate the degree of inhibition of their release by osmoprotection, GSDMD knockout, or inhibition by NSA (F).

(legend continued on next page)

contribute to the osmoprotection-independent protein release (Figures 6G and 6H). Further analyses revealed that only lysosomal lumen but not membrane proteins are released independent of their MW, presumably as vesicular cargo (Figures 6I and 6J). In contrast, we detected both mitochondrial matrix and mitochondrial membrane proteins, which are sensitive to osmoprotection depending on their MW, arguing against a significant contribution of released mitochondrial vesicles under these conditions (Figure 6J). We therefore conclude that a GSDMD-dependent, cell-lysis-independent pathway facilitates the release of many constitutively expressed small-MW proteins derived from the cytoplasm and other cellular organelles, for which extracellular functions have not been described yet.

The majority of pyroptotically released alarmins do not require induction

As the two-step activation comprising priming and NLR activation executes a convoluted secretory program, we considered experimental systems, in which a single trigger is sufficient for inflammasome activation. Caspase 1/11 double knockout (DKO) iBMDMs reconstituted with Caspase-1 fused to the FK506 binding protein (FKBP) dimerization domain (Casp1ΔCard-FKBP) provides an elegant solution, as Caspase1 can be dimerized and activated by AP20187, leading to pyroptosis (MacCorkle et al., 1998; Rühl et al., 2018). Notably, activation of Caspase-1 is independent of priming and thereby enables the dissection of proteins released with and without priming. As expected, IL-1β is released only upon priming and AP20187 treatment, whereas conventional secretion, for example of TNF-α, requires priming but is unaffected by AP20187 (Figures 7A and 7B). Cell death, however, is executed only upon addition of AP20187, independent of priming (Figure 7C). This experimental setup therefore enables the distinction of two classes of proteins—constitutive and induced damage-associated molecular patterns (c- or iDAMPs) (Yatim et al., 2017). The release of the majority of unconventionally released proteins occurs irrespective of priming (Figure 7D). The alarmin HMGB1 is one such protein and can therefore be classified as a cDAMP. Our proteomic investigations reveal that pyroptotic cell death serves as an exit route also for other cDAMPs, such as macrophage migration inhibitory factor (MIF), mminoacyl TRNA Synthetase Complex Interacting Multifunctional Protein 1 (AIMP1), galectins, and S100 families (Figure 7D; Table S1). We confirmed the release of MIF and AIMP1 by ELISA and immunoblot analysis (Figures 7G and 7H). As most cDAMPs, such as MIF or LGALS 1, 3, and 9, have a low MW, they are released independently of cell lysis, whereas others, such as AIMP1 or nicotinamide phosphoribosyltransferase (NAMPT), are predominantly released due to cell lysis. The cellular mechanisms regulating the exit of major protein classes are shown in Figure 7I and Table S1.

In contrast, IL-1β is released only during pyroptosis after priming and therefore is a paradigm iDAMP (Figure 7E). IL-18 shares some features of both classes, and even though priming boosts its expression and release, a smaller amount is constitutively expressed and therefore released independently of priming (Figures 7E and 7F; Marshall et al., 1999). Our analysis reveals that IL-1β is unique in its regulation and release kinetics, as it is among the top 5 most highly induced proteins upon priming and thereby differs from most other pyroptotically released proteins. (Figure S7A). The majority of strongly induced proteins include conventionally released cytokines that do not require pyroptosis for release (Figure S7B). Only a handful of other proteins not associated with extracellular functions, such as NLRP3, a few Ifit protein family members, and Oasl and Gbp family proteins, require priming for their induction and release by pyroptosis.

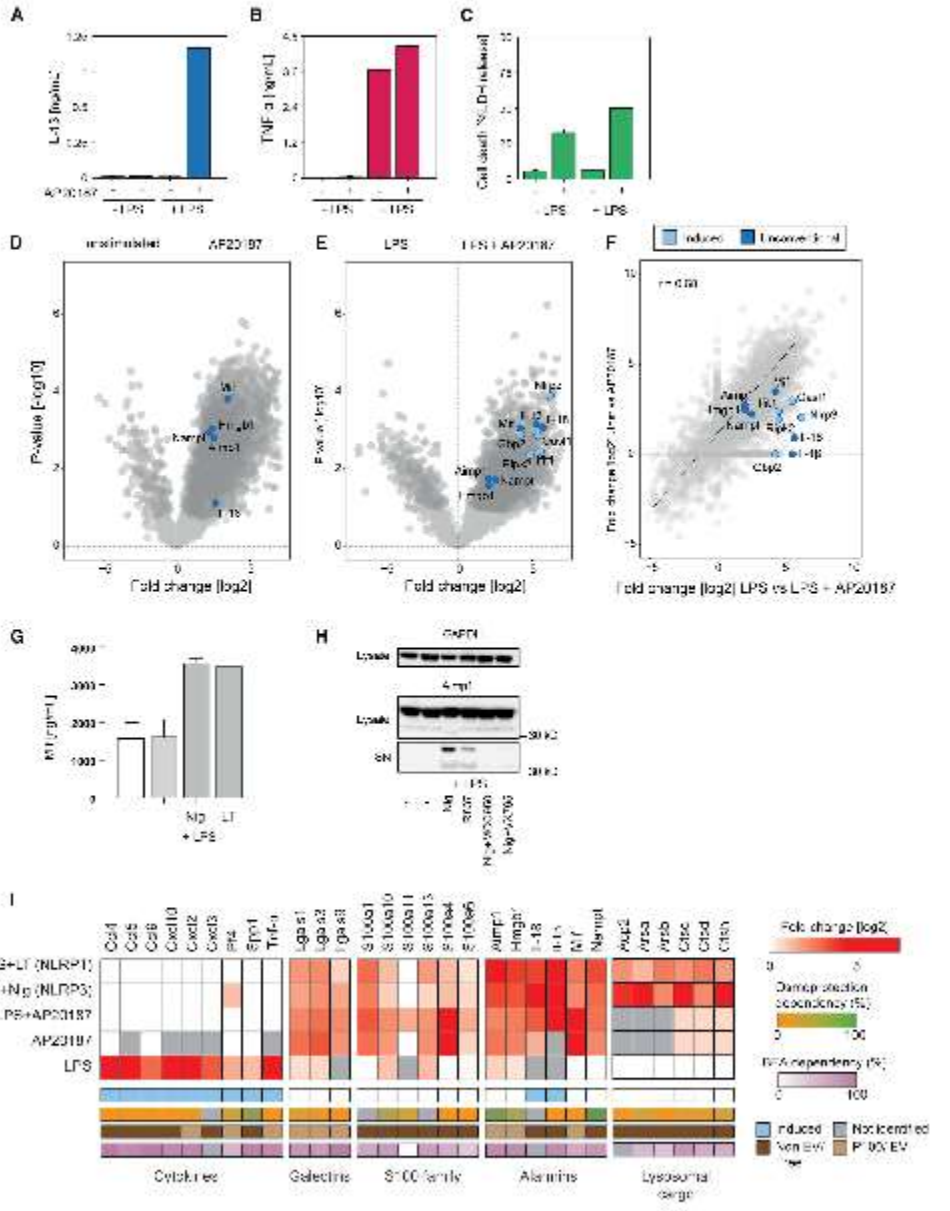
According to our global analysis, the cytokines IL-1β and IL-18 are unique in terms of their regulation in multiple steps, such as priming, proteolytic activation, MW, and PM-pore-mediated release. We did not detect any other proteins requiring a comparable complex regulation for release. This underscores the tight regulation of IL-1β and IL-18 activity, as they require hours to unfold their powerful inflammatory potential on neighboring cells or the whole organism. In contrast, a low MW and PM pore formation are sufficient for the release of most alarmins, which can be quickly released into the extracellular space and activate neighboring cells in a matter of minutes.

DISCUSSION

Activated macrophages control immune responses by releasing proteins, such as pro-inflammatory cytokines and alarmins. Variations in the secretory composition results in vastly different physiological responses of neighboring cells and the whole organism (Arango Duque and Descoteaux, 2014). We developed a generic and unbiased secretomics strategy based on pharmacological, genetic, and biochemical dissection of cellular protein release pathways. We applied it to analyze complex proinflammatory programs in macrophages upon activation of the following two major pathways of protein release: TLR4 activation and inflammasome-mediated pyroptotic cell death.

In addition to known cytokines and alarmins, we observed the unconventional release of thousands of proteins with diverse intracellular functions and localization during pyroptotic cell death. Our analytical framework facilitates the discovery of proteins released from sub-cellular organelles, including

(G) Organelle marker enrichment analysis of cell-lysis-independent unconventional release. Proteins significantly released upon NLRP3 activation in the presence of osmoprotection (LPS + Nl + 30 mM L-alanine/20 mM glycine versus LPS) were defined to be cell lysis independent. Significant proteins were determined with a parametric two-tailed Welch's t test (FDR < 0.05, S_0 = 1). Enrichment factors were calculated by Fisher's exact test as described previously (annotations with intersection sizes of >15 and p < 0.002 are shown). Category sizes are represented by balloon size.
(H) Boxplots indicating the fold changes of all detected proteins belonging to indicated organelles in the unconventional secretomes in the presence of glycine.
(I) Effect of osmoprotection on mitochondrial and lysosomal proteins of different MWs. Boxplots indicating the extent of inhibition by osmoprotection to unconventionally released mitochondrial and lysosomal proteins binned according to increasing MW.
(J) Swarm plots showing log2 fold change of proteins assigned to the indicated annotation released during osmoprotection of pyroptotic cells compared with TLR4 activated controls.
See also Figure S4 and Table S1.



(legend on next page)

intracellular proteins with moonlighting or unexpected extracellular functions (Hernández et al., 2014; Jeffery, 1999). Conventional and unconventional secretory programs appear conserved across species based on the observed high correlation of mouse and human macrophage secretomes (Figure S4F).

We report that ER-Golgi disruption with BFA and GCA inhibits unconventional protein release and pyroptosis by interfering with inflammasome formation. Our data suggest that this effect is not only specific to NLRP3 as reported by Zhang et al. (2017) but also independent of the inflammasome sensor involved, although cell type and duration of BFA treatment affects cellular phenotypes. For example, IL-1β release was reported to be independent from the ER-Golgi compartment in LPS-treated human monocytes (Rubartelli et al., 1990). Moreover, long-term ER stress by BFA presumably activates the unconventional release of IL1β, whereas acute ER stress inhibits it (Bronner et al., 2015; Menu et al., 2012).

We show that EVs do not significantly contribute to cytokine and alarmin release during macrophage pyroptosis. It is interesting to speculate that incorporated or associated proteins might signal over longer durations and distances or upon intracellular release by membrane fusion with recipient cells.

We identified the presence of a large number of lysosomal cargo proteins, including many proteases, in the extracellular space during pyroptosis. Lysosomal proteases have been detected in the extracellular environment during various biological processes, including necroptosis and extensive TLR4 activation (Meissner et al., 2013; Mohamed and Sloane, 2006; Tanzer et al., 2020). Lysosomal exocytosis has been described as a membrane repair mechanism triggered by calcium influx (Bergsbaken et al., 2011; Tanzer et al., 2020). However, it remains unresolved how lysosomal content is released upon TLR4 activation in the absence of pore formation, PM damage, or cell death and whether this involves calcium fluxes. The presence of lysosomal proteases in the extracellular environment has been shown to play a role in tissue regeneration or tumor metastasis by remodeling the extracellular matrix (ECM) (Mohamed and Sloane, 2006; Olson and Joyce, 2015), and further studies should evaluate how these processes are affected by different forms of cell death.

Osmoprotection with the amino acids L-alanine and glycine has been described to delay pyroptotic cell lysis but not GSDMD pore formation in the PM. Under these conditions, protein release takes places through GSDMD pores only. Our data show that although cell lysis facilitates the unconventional release of the majority of proteins, the release of a subset of low-MW proteins, namely, cytokines/DAMPs including IL-1β, occurs independently of cell death by a GSDMD-dependent pathway, confirming previous reports (Evavold et al., 2018; Heilig et al., 2018). Some low-MW proteins, however, are retained in the cell due to osmoprotection. This may be explained by their physiological assembly into higher order structures or large protein complexes.

We show that lysosomal cargo and some mitochondrial proteins are released independently of cell lysis. The osmoprotection-independent extracellular presence of lysosomal lumen proteins with diverse MWs suggests an involvement of a vesicular release route. In contrast, the osmoprotection-sensitive release of both mitochondrial matrix and membrane proteins suggests that mitochondrial proteins enter the cytosol due to organellar damage upon inflammasome activation, followed by their exit along with cytosolic proteins by gasdermin pores. It will be interesting to elucidate the role of vesicular protein release from different cellular compartments during later stages of cell lysis, after pore formation.

PM repair mechanisms activated by membrane damage may further fine-tune paracrine programs by favoring the release of lysosomal cargo and small-MW proteins during sub-lytic inflammasome activation (Rühl et al., 2018).

Different forms of cell death use distinct mechanisms to control their paracrine programs. Pyroptosis is unique in its fast release of cDAMPs and slow release of pro-inflammatory cytokines IL-1β and IL-18. The execution of necroptosis, in contrast, is slower and accompanied by increased receptor shedding, which is not detected in pyroptosis, and decreased cytokine production. Whether distinct sets of proteins exit cells through GSDMD versus the necroptotic terminal effector-mixed lineage kinase domain-like (MLKL)-mediated PM perturbation remains to be determined (Petrie et al., 2017; Tanzer et al., 2020). Furthermore, the cell type specificity of protein exit mechanisms and accompanying paracrine programs are exciting topics for future investigations.

In conclusion, we have devised an MS-based experimental strategy to define the exit routes for each cellular protein. We demonstrated the utility of our approach by identifying multiple alarmins with known extracellular signaling capabilities to be released via cell-lysis-dependent and -independent pathways of pyroptosis. Our work serves as a valuable resource for the identification of proteins with as yet undescribed extracellular functions. Future studies should reveal how the differentially regulated sets of proteins direct paracrine immune responses under inflammatory conditions involving inflammasomes.

Figure 7. Dissection of constitutively expressed versus TLR-activation-induced cytokines and DAMPs

(A–C) Release of TNF-α (A) and IL-1β (B) as measured by ELISA or LDH (C) from supernatants of Caspase 1/11 DKO immortalized murine macrophages expressing an engineered caspase-1 system (Casp1ΔCard-FKBP), allowing its controlled dimerization, activation, and pyroptosis by AP20187 (AP) in the presence or absence of TLR4 and AP20187 activation. Data are represented as mean ± SD of three biological replicates.

(D and E) Volcano plots showing secreted proteins from differentially activated engineered Casp1ΔCard-FKBP-expressing cells. Pairwise secretome comparisons are shown between untreated and AP20187-activated cells to define constitutively released DAMPs and cytokines (D) TLR4 or TLR4 and AP20187-activated cells (E) to define differentially released cytokines and DAMPs that require induction by TLR4 signaling.

(F) Scatterplot of fold changes of released proteins between untreated and AP20187-activated cells versus TLR4 or TLR4 and AP20187-activated cells. Conventionally and unconventionally released cytokines are indicated by red and blue circles, respectively.

(G and H) Quantification of pyroptotic release of MIF (G) and Almp1 (H) by ELISA and western blot, respectively.

(I) Heatmap of indicated conventionally secreted cytokines and pyroptotically released cDAMPs and IDAMPs or cytokines with mode of release as soluble or vesicle associated; osmoprotection or BFA dependency are indicated with the respective color coding.

See also Figures S1, S4, and S7 and Table S1.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Isolation and culture of murine bone marrow-derived macrophages
  - Generation of human monocyte-derived macrophages
  - Activation of BMDMs and human macrophages
  - Generation of cell lysates
  - Caspase 1 activity and cytotoxicity (LDH) assays
  - Enzyme-linked immunosorbent assay (ELISA)
  - Western blot
  - Immunostaining and fluorescent microscopy
  - P100 fraction/Extracellular vesicle isolation
  - Sample preparation for mass spectrometry
  - LC-MS/MS
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - LC-MS/MS data analysis
  - Data preparation, quality control, and copy number estimation
  - Bioinformatics analysis of protein secretion signatures and receptor shedding

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.108826.

ACKNOWLEDGMENTS

We thank Matthias Mann for providing LC-MS/MS infrastructure and support. We thank Stevan Dewitz, Korbinian Mayr, Mario Oroshi, Igor Paron, Christian Delmi, Bianca Spletstoeser, and Gaby Sowa for excellent technical assistance. Light microscopy was performed in the MPIB Imaging Facility. We thank Petr Broz for the GSDMD KO and Casp1/11 DKO Casp1ΔCard-FKBP iBMDMs. We thank all members of the Meissner, Zychlinsky, and Mann groups for helpful discussions. This work was funded by the Max Planck Society for the Advancement of Science and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project-IDs 165054396 (SFB 914), 360372040 (SFB 1335), and 408885537 (TRR 274). One or more of the authors of this paper self-identifies as a member of the LGBTQ+ community. The author list of this paper includes contributors from the location where the research was conducted who participated in the data collection, design, analysis, and/or interpretation of the work.

AUTHOR CONTRIBUTIONS

F.M. conceived the study. K.P. and F.M. designed experiments. K.P. performed experiments with help from S.E. and J.J.S. L.J.K. contributed experiments and analyses to Figures 1, 3, and S1–S4. K.P. analyzed data with suggestions from F.M. A.F. performed the receptor shedding analysis. K.P. and F.M. interpreted the data and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 24, 2020
Revised: November 25, 2020
Accepted: February 16, 2021
Published: March 9, 2021

REFERENCES

Aebersold, R., and Mann, M. (2016). Mass-spectrometric exploration of proteome structure and function. Nature *537*, 347–355.

Andrei, C., Dazzi, C., Lotti, L., Torrisi, M.R., Chimini, G., and Rubartelli, A. (1999). The secretory route of the leaderless protein interleukin 1beta involves exocytosis of endolysosome-related vesicles. Mol. Biol. Cell *10*, 1463–1475.

Arango Duque, G., and Descoteaux, A. (2014). Macrophage cytokines: Involvement in immunity and infectious diseases. Front. Immunol. *5*, 491.

Baroja-Mazo, A., Compan, V., Martín-Sánchez, F., Tapia-Abellán, A., Couillin, I., and Pelegrín, P. (2019). Early endosome autoantigen 1 regulates IL-1β release upon caspase-1 activation independently of gasdermin D membrane permeabilization. Sci. Rep. *9*, 5788.

Bergsbaken, T., Fink, S.L., and Cookson, B.T. (2009). Pyroptosis: host cell death and inflammation. Nat. Rev. Microbiol. *7*, 99–109.

Bergsbaken, T., Fink, S.L., den Hartigh, A.B., Loomis, W.P., and Cookson, B.T. (2011). Coordinated host responses during pyroptosis: caspase-1-dependent lysosome exocytosis and inflammatory cytokine maturation. J. Immunol. *187*, 2748–2754.

Boucher, D., Monteleone, M., Coll, R.C., Chen, K.W., Ross, C.M., Teo, J.L., Gomez, G.A., Holley, C.L., Bierschenk, D., Stacey, K.J., et al. (2018). Caspase-1 self-cleavage is an intrinsic mechanism to terminate inflammasome activity. J. Exp. Med. *215*, 827–840.

Bronner, D.N., Abuaita, B.H., Chen, X., Fitzgerald, K.A., Noñez, G., He, Y., Yin, X.M., and O'Riordan, M.X. (2015). Endoplasmic Reticulum Stress Activates the Inflammasome via NLRP3- and Caspase-2-Driven Mitochondrial Damage. Immunity *43*, 451–462.

Brough, D., Pelegrín, P., and Nicke, B. (2017). An emerging case for membrane pore formation as a common mechanism for the unconventional secretion of FGF2 and IL-1β. J. Cell Sci. *130*, 3197–3202.

Broz, P., and Dixit, V.M. (2016). Inflammasomes: mechanism of assembly, regulation and signaling. Nat. Rev. Immunol. *16*, 407–420.

Cai, Z., Zhang, A., Choksi, S., Li, W., Li, T., Zhang, X.M., and Liu, Z.G. (2016). Activation of cell-surface proteases promotes necroptosis, inflammation and cell migration. Cell Res. *26*, 886–900.

Chardin, P., and McCormick, F. (1999). Brefeldin A: the advantage of being uncompetitive. Cell *97*, 153–155.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. *10*, 1794–1805.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell. Proteomics *13*, 2513–2526.

Cypryk, W., Nyman, T.A., and Matikainen, S. (2018). From Inflammasome to Exosome-Does Extracellular Vesicle Secretion Constitute an Inflammasome-Dependent Immune Response? Front. Immunol. *9*, 2188.

Ding, J., Wang, K., Liu, W., She, Y., Sun, Q., Shi, J., Sun, H., Wang, D.C., and Shao, F. (2016). Pore-forming activity and structural autoinhibition of the gasdermin family. Nature *535*, 111–116.

DiPeso, L., Ji, D.X., Vance, R.E., and Price, J.V. (2017). Cell death and cell lysis are separable events during pyroptosis. Cell Death Discov. *3*, 17070.

Dupont, N., Jiang, S., Pilli, M., Ornatowski, W., Bhattacharya, D., and Deretic, V. (2011). Autophagy-based unconventional secretory pathway for extracellular delivery of IL-1β. EMBO J. *30*, 4701–4711.

Evavold, C.L., Ruan, J., Tan, Y., Xia, S., Wu, H., and Kagan, J.C. (2018). The Pore-Forming Protein Gasdermin D Regulates Interleukin-1 Secretion from Living Macrophages. Immunity *48*, 35–44.e6.

Fink, S.L., and Cookson, B.T. (2006). Caspase-1-dependent pore formation during pyroptosis leads to osmotic lysis of infected host macrophages. Cell. Microbiol. *8*, 1812–1825.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., et al. (2016). The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. *44*, D279–D285.

Gong, T., Liu, L., Jiang, W., and Zhou, R. (2020). DAMP-sensing receptors in sterile inflammation and inflammatory diseases. Nat. Rev. Immunol. *20*, 95–112.

Heilig, R., Dick, M.S., Sborgi, L., Meunier, E., Hiller, S., and Broz, P. (2018). The Gasdermin-D pore acts as a conduit for IL-1β secretion in mice. Eur. J. Immunol. *48*, 584–592.

Hernández, S., Ferragut, G., Amela, I., Perez-Pons, J., Piñol, J., Mozo-Villarías, A., Cedano, J., and Querol, E. (2014). MultitaskProtDB: a database of multitasking proteins. Nucleic Acids Res. *42*, D517–D520.

Jeffery, C.J. (1999). Moonlighting proteins. Trends Biochem. Sci. *24*, 8–11.

Kaisho, T., and Akira, S. (2000). Critical roles of Toll-like receptors in host defense. Crit. Rev. Immunol. *20*, 393–405.

Kayagaki, N., Stowe, I.B., Lee, B.L., O'Rourke, K., Anderson, K., Warming, S., Cuellar, T., Haley, B., Roose-Girma, M., Phung, Q.T., et al. (2015). Caspase-11 cleaves gasdermin D for non-canonical inflammasome signalling. Nature *526*, 666–671.

Kayagaki, N., Warming, S., Lamkanfi, M., Walle, L.V., Louie, S., Dong, J., Newton, K., Qu, Y., Liu, J., Heldens, S., et al. (2011). Non-canonical inflammasome activation targets caspase-11. Nature *479*, 117–121.

Kelstrup, C.D., Bekker-Jensen, D.B., Arrey, T.N., Hogrebe, A., Harder, A., and Olsen, J.V. (2018). Performance Evaluation of the Q Exactive HF-X for Shotgun Proteomics. J. Proteome Res. *17*, 727–738.

Lamkanfi, M., Sarkar, A., Vande Walle, L., Vitari, A.C., Amer, A.O., Wewers, M.D., Tracey, K.J., Kanneganti, T.-D., and Dixit, V.M. (2010). Inflammasome-dependent release of the alarmin HMGB1 in endotoxemia. J. Immunol. *185*, 4385–4392.

Lammert, C.R., Frost, E.L., Bellinger, C.E., Bolte, A.C., McKee, C.A., Hurt, M.E., Paysour, M.J., Ennerfelt, H.E., and Lukens, J.R. (2020). AIM2 inflammasome surveillance of DNA damage shapes neurodevelopment. Nature *580*, 647–652.

Lapaine, M., and Lamond, A.I. (2015). Multidimensional proteomics for cell biology. Nat. Rev. Mol. Cell Biol. *16*, 269–280.

Liu, X., Zhang, Z., Ruan, J., Pan, Y., Magupalli, V.G., Wu, H., and Lieberman, J. (2016). Inflammasome-activated gasdermin D causes pyroptosis by forming membrane pores. Nature *535*, 153–158.

Lobb, R.J., Becker, M., Wen, S.W., Wong, C.S., Wiegmans, A.P., Leimgruber, A., and Möller, A. (2015). Optimized exosome isolation protocol for cell culture supernatant and human plasma. J. Extracell. Vesicles *4*, 27031.

Loomis, W.P., den Hartigh, A.B., Cookson, B.T., and Fink, S.L. (2019). Diverse small molecules prevent macrophage lysis during pyroptosis. Cell Death Dis. *10*, 326.

Lopez-Castejon, G., and Brough, D. (2011). Understanding the mechanism of IL-1β secretion. Cytokine Growth Factor Rev. *22*, 189–195.

MacCorkle, R.A., Freeman, K.W., and Spencer, D.M. (1998). Synthetic activation of caspases: artificial death switches. Proc. Natl. Acad. Sci. USA *95*, 3655–3660.

Marshall, J.D., Abbate-Armitrano, M., Chehimi, S.S., Murphy, M., Olsen, H., and Trinchieri, G. (1999). Regulation of human IL-18 mRNA expression. Clin. Immunol. *90*, 15–21.

Martín-Sánchez, F., Diamond, C., Zeitler, M., Gomez, A.I., Baroja-Mazo, A., Bagnall, J., Spiller, D., White, M., Daniels, M.J.D., Mortellaro, A., et al. (2016). Inflammasome-dependent IL-1β release depends upon membrane permeabilisation. Cell Death Differ. *23*, 1219–1231.

Meissner, F., Schaltema, R.A., Mollenkopf, H.J., and Mann, M. (2013). Direct proteomic quantification of the secretome of activated immune cells. Science *340*, 475–478.

Menu, P., Mayor, A., Zhou, R., Tardivel, A., Ichijo, H., Mori, K., and Tschopp, J. (2012). ER stress activates the NLRP3 inflammasome via an UPR-independent pathway. Cell Death Dis. *3*, e261.

Mitchell, P.S., Sandstrom, A., and Vance, R.E. (2019). The NLRP1 inflammasome: new mechanistic insights and unresolved mysteries. Curr. Opin. Immunol. *60*, 37–45.

Mohamed, M.M., and Sloane, B.F. (2006). Cysteine cathepsins: multifunctional enzymes in cancer. Nat. Rev. Cancer *6*, 764–775.

Monteleone, M., Stow, J.L., and Schroder, K. (2015). Mechanisms of unconventional secretion of IL-1 family cytokines. Cytokine *74*, 213–218.

Nagaraj, N., Kulak, N.A., Cox, J., Neuhauser, N., Mayr, K., Hoerning, O., Vorm, O., and Mann, M. (2012). System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. Mol. Cell Proteomics *11*, M111.013722.

Ng, F., and Tang, B.L. (2016). Unconventional Protein Secretion in Animal Cells. Unconventional Protein Secretion (Springer New York), pp. 31–46.

Nickel, W., and Rabouille, C. (2009). Mechanisms of regulated unconventional protein secretion. Nat. Rev. Mol. Cell Biol. *10*, 148–155.

Olson, O.C., and Joyce, J.A. (2015). Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. Nat. Rev. Cancer *15*, 712–729.

Petrie, E.J., Hildebrand, J.M., and Murphy, J.M. (2017). Insane in the membrane: a structural perspective of MLKL function in necroptosis. Immunol. Cell Biol. *95*, 152–159.

Rathkey, J.K., Zhao, J., Liu, Z., Chen, Y., Yang, J., Kondolf, H.C., Benson, B.L., Chirieleison, S.M., Huang, A.Y., Dubyak, G.R., et al. (2018). Chemical disruption of the pyroptotic pore-forming protein gasdermin D inhibits inflammatory cell death and sepsis. Sci. Immunol. *3*, eaat2738.

Rieckmann, J.C., Geiger, R., Hornburg, D., Wolf, T., Kveler, K., Jarrossay, D., Sallusto, F., Shen-Orr, S.S., Lanzavecchia, A., Mann, M., and Meissner, F. (2017). Social network architecture of human immune cells unveiled by quantitative proteomics. Nat. Immunol. *18*, 583–593.

Rubartelli, A., Cozzolino, F., Talio, M., and Sitia, R. (1990). A novel secretory pathway for interleukin-1 beta, a protein lacking a signal sequence. EMBO J. *9*, 1503–1510.

Rühl, S., Shkarina, K., Demarco, B., Heilig, R., Santos, J.C., and Broz, P. (2018). ESCRT-dependent membrane repair negatively regulates pyroptosis downstream of GSDMD activation. Science *362*, 956–960.

Sáenz, J.B., Sun, W.J., Chang, J.W., Li, J., Bursulaya, B., Gray, N.S., and Haslam, D.B. (2009). Golgicide A reveals essential roles for GBF1 in Golgi assembly and function. Nat. Chem. Biol. *5*, 157–165.

Schaltema, R.A., and Mann, M. (2012). SprayQc: a real-time LC-MS/MS quality monitoring system to maximize uptime using off the shelf components. J. Proteome Res. *11*, 3458–3466.

Schaltema, R.A., Hauschild, J.-P., Lange, O., Hornburg, D., Denisov, E., Damoc, E., Kuehn, A., Makarov, A., and Mann, M. (2014). The Q Exactive HF, a Benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field Orbitrap analyzer. Mol. Cell. Proteomics *13*, 3698–3708.

Stow, J.L., Low, P.C., Offenhäuser, C., and Sangermani, D. (2009). Cytokine secretion in macrophages and other cells: pathways and mediators. Immunobiology *214*, 601–612.

Stutz, A., Horvath, G.L., Monks, B.G., and Latz, E. (2013). ASC speck formation as a readout for inflammasome activation. Methods Mol. Biol. *1040*, 91–101.

Swanson, K.V., Deng, M., and Ting, J.P.Y. (2019). The NLRP3 inflammasome: molecular activation and regulation to therapeutics. Nat. Rev. Immunol. *19*, 477–489.

Tanzer, M.C., Frauenstein, A., Stafford, C.A., Phulphagar, K., Mann, M., and Meissner, F. (2020). Quantitative and Dynamic Catalogs of Proteins Released during Apoptotic and Necroptotic Cell Death. Cell Rep. *30*, 1260–1270.e5.

Tapia, V.S., Daniels, M.J.D., Palazón-Riquelme, P., Dewhurst, M., Luheshi, N.M., Rivers-Auty, J., Green, J., Redondo-Castro, E., Kaldis, P., Lopez-Castejon, G., and Brough, D. (2019). The three cytokines IL-1β, IL-18, and IL-1α share related but distinct secretory routes. J. Biol. Chem. *294*, 8325–8335.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. USA *98*, 5116–5121.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731–740.

Weischenfeldt, J., and Porse, B. (2008). Bone Marrow-Derived Macrophages (BMM): Isolation and Applications. CSH Protoc. *2008*, pdb.prot5080-pdb.prot5080.

Yatim, N., Cullen, S., and Albert, M.L. (2017). Dying cells actively regulate adaptive immune responses. Nat. Rev. Immunol. *17*, 262–275.

Zhang, M., Kenny, S.J., Ge, L., Xu, K., and Schekman, R. (2015). Translocation of interleukin-1β into a vesicle intermediate in autophagy-mediated secretion. eLife *4*, e11205.

Zhang, Z., Meszaros, G., He, W.T., Xu, Y., de Fatima Magliarelli, H., Mailly, L., Mihlan, M., Liu, Y., Puig Gámez, M., Goginashvili, A., et al. (2017). Protein kinase D at the Golgi controls NLRP3 inflammasome activation. J. Exp. Med. *214*, 2671–2693.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-mouse Aimp1	Novus Biologicals	Cat#NBP2-27206SS
Anti-GAPDH	Cell Signaling Technology	Cat#8884s; RRID: AB_11129865
Anti-PDI	Cell Signaling Technology	Cat#3501; RRID: AB_2156433
Anti-Rcas1	Cell Signaling Technology	Cat#12290; RRID: AB_2736985
Anti-Ctsb	Cell Signaling Technology	Cat#31718; RRID: AB_2687580
Anti-β-Actin	Cell Signaling Technology	Cat#8457; RRID: AB_10950489
Anti-Vamp8	Cell Signaling Technology	Cat#13060; RRID: AB_2798103
Anti-Lamp1	Cell Signaling Technology	Cat#3243; RRID: AB_2134478
Anti-Asc	Sigma-Aldrich	04-147; RRID: AB_1977033
F(ab')2-AF647	Thermo Fisher Scientific	A21237; AB_1500743
HRP-linked anti mouse IgG	Cyvita	Cat#NXA931; RRID: AB_772209
HRP-linked anti rabbit IgG	Cyvita	Cat#NA934; RRID: AB_772206
Biological samples		
Buffy Coats	Blood donations to the red cross	"Blutspendedienst des Bayerischen Roten Kreuzes gemeinnützige GmbH"
Chemicals, peptides, and recombinant proteins		
Lipopolysaccharide (LPS) from Salmonella typhimurium S-form	Enzo Life Sciences	ALX-581-013-L002
Nigericin sodium salt from Streptomyces hygroscopicus	Thermo Fisher Scientific	N1495
Adenosine 5'-triphosphate disodium salt hydrate (ATP)	Sigma	Cat#7699
Lethal factor	List Biologicals	Cat#172B
Protective antigen	List Biologicals	Cat#171E
AP20187/B/B homodimerizor	Takara Biotech	Cat#635058
Brefeldin A	Cell Signaling Technology	Cat#9972
Golgicide A	Sigma-Aldrich	G0923
Glybenclamide/Glyburide	Novus Biologicals	Cat#NBP2-30141
Draq5	Thermo Fisher Scientific	Cat#62252
L-Alanine	Sigma-Aldrich	A7469
Glycine	Sigma-Aldrich	G7126
hM-CSF	Cell Signaling Technology	Cat#8929LC
Red blood cell lysis buffer	Sigma-Aldrich	R7757
Histopaque-1077	Sigma-Aldrich	10771
Heat inactivated Fetal Bovine Serum	Invitrogen	10270106
Penicillin/Streptomycin	Invitrogen	15140122
RPMI medium	Invitrogen	61870044
DME medium	Invitrogen	31966047
Phenol red free DMEM	Thermo Fisher Scientific	31053044
Urea	Sigma	45128-500 g
Thiourea	Sigma	T8656-500 g
Trizma	Sigma	T1503-1kg
Dithiothreitol (DTT)	Sigma	D0632-100 g

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chloroacetamide (CAA)	Sigma	C0267-100 g
Iodoacetamide (IAA)	Sigma	I6125-100 g
Ammonium bicarbonate	Sigma	A6141
Trypsin	Sigma	T6567-1mg
Lys-C	Wako Chemicals	Cat#129-02541
DMSO	Sigma	D2650-100ml
Acetone	Fisher Chemical	Cat#67-64-1
Acetonitrile	VWR	Cat#20048320
Trifluoroacetic acid	Merck	Cat#8082600100
Formic acid	Merck	Cat#1002641000
PBS	GIBCO	Cat#14190-094
Tween	Acros	Cat#233360010
EDTA	Sigma	03677-500 g
BSA	Serva	Cat#11930.03
Glycerol	Sigma	G5516-1L
Sodium dodecyl sulfate (SDS)	Roth	Cat#CN30.3
Sodium deoxycholate (SDC)	Sigma	30970-100 g
Sodium chloride (NaCl)	VWR	Cat#27810.295
NuPAGE LDS Sample Buffer (4x)	Invitrogen	NP0007
Puromycin	Invivogen	Ant-pr-5
Critical commercial assays		
CytoTox 96 Non-Radioactive Cytotoxicity Assay	Promega	G1780
Human Tnf ELISA kit	R&D Systems	Cat#DY210
Human IL1b ELISA kit	R&D Systems	Cat#DY201
Mouse Tnf ELISA kit	R&D Systems	Cat#DY410
Mouse IL1b ELISA kit	R&D Systems	Cat#DY401
Mouse Mif ELISA kit	R&D Systems	Cat#DY1978
Caspase-Glo® 1 Inflammasome Assay	Promega	G9951
Human Monocyte Isolation Kit II	Miltenyi Biotec	Cat#130-091-153
Deposited data		
Raw Mass Spectrometry Data Files	This paper	ProteomeXchange Consortium via the PRIDE partner repository, with the dataset identifier PXD018659
Experimental models: organisms/strains		
C57BL/6J wild type mice (WT)	The Jackson Laboratory	Cat#000664
Casp1/11 DKO, strain B6.129S2-Casp1tm1SeshCasp11del/J	BASF	N/A
129S4/SvJae	The Jackson Laboratory	Cat#009104
Experimental models: cell lines		
GSDMD –/– iBMDMs	Dr. Petr Broz	Heilig et al., 2018
Casp1/11 DKO iBMDMs reconstituted with Casp1ΔCard-FKBP	Dr. Petr Broz	Boucher et al., 2018
L929 cells	ATCC	ATCC CCL-1
Software and algorithms		
MaxQuant	Cox and Mann, 2008; Version 1.5.0.38	https://www.biochem.mpg.de/5111795/maxquant
Perseus	Tyanova et al., 2016; Version 1.5.3.0	https://www.biochem.mpg.de/5111810/perseus

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
XCalibur	Thermo Scientific	https://www.thermofisher.com/order/catalog/product/OPTON-30487
R	N/A	https://www.r-project.org/
ImageJ	N/A	https://imagej.nih.gov/ij/
Adobe Illustrator	N/A	https://www.adobe.com/de/products/illustrator.html
PRIDE	N/A	https://www.ebi.ac.uk/pride/archive/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Felix Meissner (meissner@biochem.mpg.de).

Materials availability

This study did not generate new unique materials and reagents.

Data and code availability

The datasets generated in this study are available via ProteomeXchange with identifier PXD018659.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Experiments described in this study are performed with the mouse and human primary macrophages or immortalized mouse macrophages.

C57BL/6J wild-type mice (WT) and 129S4/SvJae mice (WT) were obtained from The Jackson Laboratory. Caspase-1/11 double knockout mice (Casp1/11 DKO, strain B6.129S2-Casp1^{tm1Sesh}Casp11^{del}/J) were kindly provided by BASF. This strain was originally generated as Casp1 deficient mice but is effectively Casp1/11 double deficient due to an additional mutation in the Casp11 gene locus in the 129S2 background used for ES cell generation (Kayagaki et al., 2011). Both mouse strains were housed under specific-pathogen-free (SPF) conditions on a 12-hour light/dark cycle in the animal facility of the Max Planck Institute for Infection Biology OR Biochemistry. 10-12 week old male mice were sacrificed by cervical dislocation and directly used for bone marrow-derived macrophage preparation. Animal experiments were performed according to the German Animal Protection Law.

Human primary macrophages were generated by isolating PBMCs from buffy coats (generated from blood donations to the Red Cross or Blutspendedienst des Bayerischen Roten Kreuzes gemeinnützige GmbH) using Histopaque-1077 (Sigma-Aldrich, 10771). Gender or age of donors is not disclosed by the Red Cross. Monocytes were isolated using a monocyte isolation kit (Miltenyi Biotec, 130-091-153). Immortalized GSDMD KO and Casp1/11 DKO iBMDMs reconstituted with Casp1ΔCard-FKBP were a kind gift from Petr Broz (Boucher et al., 2018; Heilig et al., 2018).

METHOD DETAILS

Isolation and culture of murine bone marrow-derived macrophages

Bone marrow-derived macrophages (BMDMs) were prepared as described elsewhere (Weischenfeldt and Porse, 2008). In brief, bone marrow was collected from the femurs and tibiae of 8-12 weeks old male mice and filtered through a 70 μm nylon mesh filter. 5×10^6 to 1×10^7 bone marrow cells were plated on sterile, non-tissue culture treated Petri plates for a period of 7 days in macrophage differentiation medium DMEM supplemented with 10% (v/v) FCS and 20% (v/v) M-CSF-containing medium. Medium was replenished on day 3 of culture. M-CSF-conditioned medium was collected from L-929 cells. Cells were lifted from plates by incubating in cold PBS and re-plated for experiments.

Immortalized Casp1/11 DKO BMDMs reconstituted with Casp1ΔCard-FKBP were cultured in macrophage differentiation medium supplemented with 10ug/ml Puromycin.

Generation of human monocyte-derived macrophages

Blood was drawn with S-Monovette 9ml K2E-Gel (#02.1333.001, Sarstedt) from healthy volunteers according to the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report. Donors provided informed consent and all samples were collected with approval from the ethics committee. Blood was diluted 1:1 with PBS and centrifuged over Histopaque-1077 (#10771, Sigma) at 400 g for 40 min. After centrifugation, the peripheral blood mononuclear cell (PBMC) fraction was

collected and washed three times with PBS followed by centrifugation at 200 g for 15 min to remove platelets. Red blood cells were lysed with RBC lysis buffer (Sigma) for 3 min. PBMCs were washed again with PBS supplemented with 0.5% (w/v) bovine serum albumin (BSA) and 2 mM EDTA. Monocytes were isolated by negative selection using the human Monocyte Isolation Kit II (#130-091-153, Miltenyi Biotec) according to the manufacturer's instructions. For the generation of macrophages, 2×10^7 isolated monocytes were plated on sterile, non-tissue culture treated Petri plates for a period of 7 days in RPMI 1640 supplemented with 10% (v/v) FCS, and 50 ng/mL human M-CSF (#8929LC, Cell Signaling). Medium was refreshed on day 3 of culture. On day 7, cells were lifted from plates by incubating in cold PBS and re-plated for experiments.

Activation of BMDMs and human macrophages

All cells were grown, maintained and activated at 37°C and 5% CO_2 in a water-jacketed incubator. 1×10^6 WT or *Casp1/11* DKO murine macrophages were plated per 12-well in DMEM, 10% (v/v) FCS (DMEM complete medium) one day prior to stimulation. All BMDMs were primed with 200 ng/mL LPS from *Salmonella* Typhimurium for 2 h or left untreated and subsequently washed three times with serum- and phenol red-free DMEM supplemented with 2 mM L-glutamine (DMEM minimal medium). For the time course experiment (Dataset 1), cells were preincubated with 0 or 100 μM glyburide (#NBP2-30141, Novus Biologicals) for 15-30min before 2 mM ATP (#7699, Sigma) or 10 μM nigericin (N1495, Invitrogen) were added for 15, 30, 45, and 60 min. Lethal factor and protective antigen were added at 1ug/mL. To inhibit conventional protein secretion (Dataset 3), 0-2 μg/mL brefeldin A were added for the last 15-30 min of priming. After washing 0-2ug/ml BFA along with 5 mM ATP or 15 μM nigericin were added for 30-45 min. For osmoprotection experiments (Dataset 2), cells were incubated with 20-30mM L-alanine or Glycine in DMEM minimal medium for 1 h after priming and washing, before 5 mM ATP or 15 μM nigericin were added for 30-45 min. AP20187 was added at 100nM for 160min.

Human macrophages (Datasets 5 and 6) were seeded at a density of 1×10^6 cells per 12-well in RPMI 1640 containing 10% (v/v) FCS, 2 mM L-glutamine, MEM NEAA, 100 U/mL penicillin, 100 μg/mL streptomycin (RPMI complete medium) one day before activation. Cells were primed with 200 ng/mL LPS for 2 h in complete medium. 1 μg/mL Brefeldin A was added for the last 30 min of priming before cells were washed three times with serum- and phenol red-free RPMI 1640 supplemented with 2 mM HEPES and 2 mM L-glutamine (RPMI minimal medium). After washing, cells were incubated with 1 μg/mL brefeldin A and 2 mM ATP or 10 μM nigericin for 30-45 min.

All activation experiments were performed in three technical replicates. Cell supernatants were centrifuged at 400 g for 5 min or sterile filtered to remove cell debris and directly used for LDH assay or stored at −80°C until ELISA and sample preparation. After removal of supernatants, untreated and LPS treated samples were used for the generation of total cell lysates (Datasets 4 and 6).

Generation of cell lysates

To generate total lysates for mass spectrometry, cells were washed with PBS and resuspended in 8 M urea in 40 mM HEPES pH 8 and frozen at −80°C until sample preparation.

Caspase 1 activity and cytotoxicity (LDH) assays

Caspase 1 activity was measured using the Caspase-Glo® 1 Inflammasome Assay as per the manufacturer's instructions. Cell death was measured by means of lactate dehydrogenase (LDH) release into the cell supernatant using the CytoTox 96 Non-Radioactive Cytotoxicity Assay (#G1780, Promega). In brief, 50 μL cell supernatant were incubated with 50 μL LDH substrate and incubated in the dark at 37°C for 15 to 30 min. The enzymatic reaction was stopped by adding 50 μL stop solution. Whole cell lysates generated from unstimulated cells incubated with medium supplemented with 1.2% (v/v) Triton X-100 for 30 min served as a reference for maximal cell death. Cell death was calculated as follows: (LDH release/LDH whole cell lysate) × 100.

Enzyme-linked immunosorbent assay (ELISA)

Human and mouse IL-1β, TNF-α and MIF in cell supernatants were measured with ELISA Kits purchased from R and D Systems (#DY201, #DY401, #DY210, #DY410, #DY1978) according to the manufacturer's instructions.

Western blot

Samples were prepared by lysing cells in lysis buffer (1% Triton X-100, 50mM Tris-Hcl pH 7.5, 150mM NaCl. Nuclei were removed by centrifugation (10 min, 1000 × g). Proteins from cell supernatants were precipitated with acetone (final concentration 80%) o/n at −20C. Protein pellets obtained after centrifugation at 3900 x g were solubilized in 2% SDS, 50mM Tris-HCl pH 7.5. Protein concentration was normalized after measurement with a bicinchoninic acid assay (BCA; Thermo Fisher Scientific). Samples were reduced and denatured by adding NuPAGE LDS Sample Buffer (4 ×) and NuPAGE Sample Reducing Agent (10 ×) (both from Thermo Fisher) and heating at 85 °C for 10 min. Proteins were separated by 4%–12% SDS-PAGE in precast gels (Novex; Invitrogen) with MOPS buffer (Novex; Invitrogen). Proteins were transferred onto Immobilon-FL PVDF membranes (Millipore) and nonspecific binding was blocked with 5% non-fat milk in Tris-buffered saline with 0.1% Tween-20 (TBST) for 1 h, followed by overnight incubation with specific primary antibodies in 5% BSA in TBST as per the manufacturer's instructions.

The following primary antibodies were used: PDI (1:1000, CST #3501), Rcas1 (1:1000, CST #12290), Ctsb (1:1000, CST #31718), β-Actin (1:1000, CST #8457), IL-1β (R and D systems, #DY401 ELISA detection antibody 1:50), Aimp1 (NBP2-27206SS), Gapdh (1:1000, CST #8884s), Vamp8 (1:100, CST #13060), Lamp1 (1:100, CST #3243). Membranes were washed 3X and incubated in

HRP linked anti-mouse (Cytiva #NXA931) or rabbit IgG antibodies (Cytiva #NA934) in 5% BSA in TBS-T for 1 h at room temperature (RT) followed by washing 3X in TBS-T

Immunostaining and fluorescent microscopy

BMDMs were fixed with 4% formaldehyde and stained for ASC (Millipore 2EI-7, Thermo Fisher F(ab')2-AF647) as per the manufacturer's instructions. Nuclei were counterstained using Draq5. Microscopy was performed in the MPIB imaging facility. Cells were imaged using a GE Deltavision Elite (Cytiva Life Sciences) epifluorescence microscope. Images were quantified as described previously using ImageJ (Stutz et al., 2013).

P100 fraction/Extracellular vesicle isolation

2e7 bone marrow macrophages were treated for three hours with LPS followed by 45min with Nigericin or ethanol a control. Supernatants were collected and spun for 5 minutes at 500 x g. followed by another spin 30 minutes at 10,000 x g. Supernatants were concentrated 5X using 30kD cellulose filters and the retentates were collected and spun for 60 minutes at 100,000 x g to obtain extracellular vesicle P100 fraction. The pellets were washed with ice cold PBS and again centrifuged for 60 minutes at 100,000 x g. The resulting pellet was lysed in 8 M Urea with 50 mM Tris (pH8), reduced with 10 mM DTT, alkylated with 40 mM CAA and digested for 2 hours with lysC (1 µg/sample) before 1:4 dilution with 50 mM ammonium bicarbonate and the addition of trypsin (1 µg/sample). Trypsin and lysC digestion occurred over night at RT and peptides were cleaned up by C18 stage tipping.

Sample preparation for mass spectrometry

Total lysates were diluted to a final urea concentration of 2 M and sonicated on ice for 15 min (level 5, Bioruptor, Diagenode). Cell supernatants (400 µL each) were denatured with 2 M urea in 10 mM HEPES pH 8. Proteins of both sample types were reduced with 10 mM dithiotreitol for 30 min at RT followed by alkylation with 55 mM iodoacetamide for 20 min at RT in the dark. Remaining iodoacetamide was quenched with 100 mM thiourea. Proteins were digested with 1 µg LysC (#129-02541, Wako Chemicals) at RT for 3 h and 1 µg trypsin (#T6567, Sigma) at RT overnight. Protein digestion was stopped with 0.6% (v/v) trifluoroacetic acid and 2% (v/v) acetonitrile before peptides were loaded onto reversed phase C18 StageTips (#2215, 3M™ Empore™, IVA Analysentechnik). Supernatants and 50 µg of total lysates were loaded onto the C18 StageTips. Peptides were desalted using 0.5% (v/v) acetic acid and subsequently eluted from the C18 StageTips with 50 µL 80% (v/v) acetonitrile in 0.5% (v/v) acetic acid. After concentrating and drying in a SpeedVac (Thermo Scientific), peptides were resuspended in 10 µL 2% (v/v) acetonitrile, 0.1% (v/v) trifluoroacetic acid in 0.5% (v/v) acetic acid and stored at −20°C until mass spectrometric analysis.

LC-MS/MS

Peptide mixtures were analyzed in a single-run liquid chromatography mass spectrometry (LC-MS/MS) format (Nagaraj et al., 2012). Each peptide mixture was loaded onto a C18-reversed phase column (20 cm long for supernatants and 50 cm long for total lysates, 75 µm inner diameter) and separated with a non-linear gradient of 2 - 60% buffer B (80% (v/v) acetonitrile in 0.1% (v/v) formic acid) at a flow rate of 250 nL/min over 107 min for supernatants and 180 min for total lysates using a nanoflow UHPLC instrument (Easy-nLC 1200, Thermo Scientific). Chromatography columns (#TSP075375, Composite Metal Service Ltd.) were packed at the MPI of Biochemistry with ReproSil-Pur 120 C18-AQ 1.9 µm resin (#r119.aq., Dr. Maisch GmbH) in methanol. Chromatography and column oven (Sonation GmbH) temperature were controlled and monitored in real-time with SprayQC (Scheltema and Mann, 2012). Column oven temeprature was set to 40°C for supernatants and 55°C for total lysates. Separated peptides were analyzed on a benchtop quadrupole-Orbitrap instrument (Q Exactive HF/HFx mass spectrometer, Thermo Scientific) with a nanoelectrospray ion source (Thermo Scientific), which was coupled on-line to the liquid chromatography instrument.

The mass spectrometer was operated in a data dependent mode with a survey scan range of 300-1650 m/z and a resolution of 60,000 at m/z 200. Up to the 10 most abundant precursor ions with charge states 2 to 5 were isolated for higher-energy collisional dissociation (HCD) with Thomson (Th) isolation windows of 1.4 m/z for secretomes and 1.8 m/z for total proteomes. Normalized collision energies (NCE) for HCD were 26 (proteomes) and 27 (secretomes), respectively. Fragmentation spectra were acquired with a resolution of 15,000 at m/z 200. Dynamic exclusion duration of sequenced peptides was set to 20 s (secretomes) or 30 s (proteomes) to reduce repeated peptide sequencing. Maximum ion injection times were 20 ms for the full MS scan and 80 ms (secretomes) or 55 ms (proteomes) for the MS/MS scan. Automatic gain control (AGC, ion target values) was set to 3e6 for the survey and 1e5 for the MS/MS scan. MS data were acquired as described previously using the Xcalibur software (Thermo Scientific) (Meissner et al., 2013).

QUANTIFICATION AND STATISTICAL ANALYSIS

LC-MS/MS data analysis

Protein identification and quantification from MS raw files was performed separately for the murine and human datasets using the computational proteomics platform MaxQuant (software version 1.5.5.2 or 1.6.2.1) (Cox and Mann, 2008). Murine and human MS/MS spectra were searched against the respective UniProt FASTA databases and a common contaminant database by the implemented Andromeda search engine (Cox et al., 2011). Secretomes and proteomes of all murine and all human datasets, respectively, were analyzed together. To avoid matching between different sample types, secretomes were set to fraction 1 and parameter

group 0 and total lysates to fraction 3 and parameter group 1. The used digestion mode was Trypsin/P with a minimum peptide length of 7 amino acids and a maximum of two missed cleavages. Cysteine carbamidomethylation was set as fixed and methionine oxidation and N-terminal acetylation as variable modifications. False discovery rates (FDR) were 1% at the peptide and protein level. Peptide identification was performed with an allowed initial precursor mass deviation up to 4.5 ppm and an allowed fragment mass deviation of 20 ppm. Protein identifications required one unique or razor peptide. For nonlinear retention time alignment of all samples, the "Match between runs" option of MaxQuant was used. The time windows for matching peptide identifications across different samples and to search for the best alignment function were set to 0.7 min and 20 min. For label-free quantification (LFQ) via MaxLFQ, a minimum ratio count of 1 was used and Fast LFQ enabled with a minimum of 3 and an average of 6 neighbors (Cox et al., 2014).

Data preparation, quality control, and copy number estimation

Data was analyzed with the Perseus computational platform (version 1.5.5.5) and the statistical software environment R and R Studio, respectively (Tyanova et al., 2016). The integrated graphics device or the package "ggplot2" in R Studio were used to visualize data. Proteins matching to the reversed or contaminants database as well as peptides only identified by side modification were excluded from the analysis. Data were additionally filtered to contain at least two valid values per protein identification in at least one group of replicates. The number of identified protein groups per condition and the Pearson correlation between technical replicates were determined and subsequently missing values were replaced. Missing values were imputed separately for secretomes and proteomes by a Gaussian distribution with a 30% width relative to the standard deviation of the measured values and a downshift of the mean by 1.8 standard deviations of the valid data. To assess the accuracy of the mass spectrometry approach, MS raw intensities and LFQ values were correlated (Pearson) to ELISA values for TNF-α and IL-1β using pairwise complete observations. By analogy, the Pearson correlation of LDH release and the summed MS raw or LFQ intensities of all proteins annotated as "cytoplasm," "cytoplasmic part," "intracellular," or "intracellular part" (cytosolic proteins) by the Gene Ontology (GO) term cellular component (GOCC) was calculated. Of note, the label-free quantification technology MaxLFQ that in addition to peptide intensity information involves various normalization steps of the LC MS/MS runs to make protein amounts between different samples more comparable showed overall a much weaker correlation to the alternative (antibody and enzyme-based) quantification methods (Figures S2G–S2J). One reason being that the assumed requirements for MaxLFQ calculation, e.g., that protein compositions do not show major differences across different samples, are not fulfilled in our datasets. Hence, protein amounts were determined by label-free quantification using the summed raw intensities of the MS1 signal of each eluting peptide for all further analyses.

Bioinformatics analysis of protein secretion signatures and receptor shedding

To investigate differences in conventional and unconventional protein secretion from activated macrophages on a global scale we pairwise compared secretomes of differentially treated macrophages by calculating the ratios of individual protein abundances. A parametric two-tailed welch's t test with a permutation-based false discovery rate (FDR) of 5% and a S_0 parameter of 1 was employed to identify proteins that significantly differed in abundance (Tusher et al., 2001). Secretory signatures were functionally characterized by annotation enrichment analysis using annotations from the UniProt Keywords and Pfam protein domain databases (Finn et al., 2016). Enrichment factors were calculated with Fisher's exact test. All annotations with intersection sizes greater than 3 and p values less than 0.00025 were considered.

To analyze pyroptosis dependent receptor shedding, proteins were prefiltered by Uniprot Keywords 'Receptors' and their corresponding peptides were annotated "extracellular" and "intracellular" via Uniprot Keywords, respectively. Peptide intensities were logarithmized and normalized to LPS primed untreated conditions. A two-tailed t test was performed between extracellular and intracellular peptides of all receptors (Figure S6A) and selected receptors (Figures S6B and S6C) and significance was denoted by asterisks (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$).

4.6 Project 6: Environmental arginine controls multinuclear giant cell metabolism and formation

In joint illnesses such as rheumatoid arthritis (RA), Osteoclast mediated abnormal bone resorption results in bone deterioration. Osteoclast formation is governed by cytokines present in the surrounding environment, such as receptor activator of NF-B ligand (RANKL), and may also be influenced by external nutrients. Our study in mice showed, that decrease of the amino acid arginine inhibits osteoclastogenesis and ameliorates arthritis. We evaluated the effects of a recombinant modified form of arginase 1 (recArg1) – a previously shown activator of macrophages in autoimmunity - *in vitro* and in mice models of inflammatory arthritis to determine whether depletion of systemic arginine could alter cellular metabolism in bones and slow the course of inflammatory arthritis. We demonstrated that RANKL-mediated osteoclastogenesis requires extracellular arginine by integrating arthritis mice models with different omics methods (transcriptomics, proteomics, and metabolomics) in the presence and absence of recArg1. Systemic arginine restriction enhanced results in a variety of mouse arthritic models, particularly when osteoclast-mediated bone degradation was considered. Arginine deficiency inhibited RANKL's transcriptional and metabolic activities, resulting in metabolic quiescence in osteoclast precursor cells. Restriction of arginine had reversible effects on osteoclast development, and arginine precursors could compensate for their absence.

This paper was published in Nature communications in 2020 [327] and was a collaboration with Gernot Schabbauer of the Medical University of Vienna.

I contributed the proteomics sample preparation, MS analysis and data analysis.

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-14285-1 OPEN

Environmental arginine controls multinuclear giant cell metabolism and formation

Julia S. Brunner[1,2], Loan Vulliard[3], Melanie Hofmann[1,2], Markus Kieler[1,2], Alexander Lercher[3], Andrea Vogel[1,2], Marion Russier[4], Johanna B. Brüggenthies[4], Martina Kerndl[1,2], Victoria Saferding[5], Birgit Niederreiter[5], Alexandra Junza[6,7], Annika Frauenstein[4], Carina Scholtysek[8], Yohei Mikami[9,10], Kristaps Klavins[3], Gerhard Krönke[8], Andreas Bergthaler[3], John J. O'Shea[9], Thomas Weichhart[11], Felix Meissner[4], Josef S. Smolen[5], Paul Cheng[12], Oscar Yanes[6,7], Jörg Menche[3], Peter J. Murray[4], Omar Sharif[1,2], Stephan Blüml[2,5,13*] & Gernot Schabbauer[1,2,13*]

Multinucleated giant cells (MGCs) are implicated in many diseases including schistosomiasis, sarcoidosis and arthritis. MGC generation is energy intensive to enforce membrane fusion and cytoplasmic expansion. Using receptor activator of nuclear factor kappa-B ligand (RANKL) induced osteoclastogenesis to model MGC formation, here we report RANKL cellular programming requires extracellular arginine. Systemic arginine restriction improves outcome in multiple murine arthritis models and its removal induces preosteoclast metabolic quiescence, associated with impaired tricarboxylic acid (TCA) cycle function and metabolite induction. Effects of arginine deprivation on osteoclastogenesis are independent of mTORC1 activity or global transcriptional and translational inhibition. Arginine scarcity also dampens generation of IL-4 induced MGCs. Strikingly, in extracellular arginine absence, both cell types display flexibility as their formation can be restored with select arginine precursors. These data establish how environmental amino acids control the metabolic fate of polykaryons and suggest metabolic ways to manipulate MGC-associated pathologies and bone remodelling.

[1] Institute for Vascular Biology, Centre for Physiology and Pharmacology, Medical University Vienna, 1090 Vienna, Austria. [2] Christian Doppler Laboratory for Arginine Metabolism in Rheumatoid Arthritis and Multiple Sclerosis, 1090 Vienna, Austria. [3] CeMM Research Centre for Molecular Medicine of the Austrian Academy of Sciences, 1090 Vienna, Austria. [4] Max Planck Institute of Biochemistry, 82152 Martinsried, Germany. [5] Division of Rheumatology, Department of Internal Medicine III, Medical University of Vienna, 1090 Vienna, Austria. [6] CIBER of Diabetes and Associated Metabolic Diseases (CIBERDEM), 28029 Madrid, Spain. [7] Metabolomics Platform, IISPV, Department of Electronic Engineering, Universitat Rovira i Virgili, 43204 Tarragona, Spain. [8] Department of Internal Medicine 3, Rheumatology and Immunology, Friedrich-Alexander-University Erlangen-Nürnberg (FAU) and Universitätsklinikum Erlangen, 91054 Erlangen, Germany. [9] Molecular Immunology and Inflammation Branch, NIAMS, National Institutes of Health, Bethesda, MD, Bethesda, MD 20892, USA. [10] Division of Gastroenterology and Hepatology, Department of Internal Medicine, Keio University School of Medicine, Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan. [11] Center of Pathobiochemistry and Genetics, Institute of Medical Genetics, Medical University of Vienna, 1090 Vienna, Austria. [12] Bio Cancer Treatment International Ltd., 999077 Hong Kong, China. [13] These authors contributed equally: S. Blüml, G. Schabbauer. *email: stephan.blueml@meduniwien.ac.at; gernot.schabbauer@meduniwien.ac.at

Exacerbated bone resorption by tissue resident osteoclasts underlies the pathogenesis of multiple chronic inflammatory joint diseases, which are significant burdens to human health[1]. Tartrate-resistant acid phosphatase (TRAP)-positive multinucleated osteoclast generation from myeloid progenitors relies on cytokines, most importantly receptor activator of nuclear factor kappa-B ligand (RANKL) and macrophage colony stimulating factor (M-CSF)[2]. However, depending on the tissue, the transcriptional landscape of resident cells is also dictated by their surroundings, with changes in extracellular microenvironment through amino acid (AA) supplementation/restriction regulating immune cell plasticity[3–6]. A hallmark of this plasticity is the fine tuning of cellular metabolism that allows immune cells, especially macrophages to dynamically adjust their metabolism according to available extracellular energy sources. Such metabolic flexibility in response to changing environmental conditions represents an important paradigm in innate immune cell biology[4,6,7]. Analogously, cancer cells are able to adjust their metabolism in the face of nutrient shortage by rewiring their metabolism to efficiently utilize alternative metabolites, thereby sustaining their proliferation[8]. Nonetheless, compensation cannot occur for every nutrient as enzymes that deplete specific AAs can be utilized to limit auxotrophic cancer cell growth. For example, L-asparaginase is used in acute lymphatic leukaemia treatment, but its therapeutic efficacy is limited by side effects[9]. Recent clinical studies employed pegylated recombinant arginase 1 (recArg1 or BCT-100) to deplete arginine in various cancers, yet impacts of arginine availability on other disease settings remains relatively unexplored[10,11]. We reasoned that manipulating environmental arginine could alter local cellular metabolism within bone, possibly through impacts on tissue resident osteoclasts and thus change the trajectory of arthritis-induced bone loss.

To test this hypothesis, we used recArg1 to study the effects of arginine depletion in arthritis and osteoclastogenesis. RecArg1 exerts beneficial effects in murine arthritis by dampening osteoclast metabolism and differentiation. In arginine scarcity, osteoclasts and IL-4-induced multinucleated giant cells (MGCs) display flexibility in their development, as their formation can be restored by supplementation with select arginine precursors. Our results highlight the importance of environmental arginine in MGC development and show therapeutic effects of systemic extracellular arginine depletion in murine arthritis.

Results

RecArg1 exerts beneficial effects in murine arthritis. To investigate if systemic arginine manipulation modulated bone diseases, we first used recArg1 in serum transfer arthritis (K/BxN), which is characterized by RANKL-induced bone erosion (Supplementary Fig. 1a)[12]. Arginine depletion during disease caused improved clinical outcomes as evaluated by increased grip strength associated with reduced paw swelling (total score) and weight loss. Histological analysis revealed decreased formation of osteoclasts in arthritic paws, whereas we found negligible effects of arginine bioavailability on joint inflammation (Fig. 1a). Using a different arthritis model, we found arginine was equally important in tumour necrosis factor (TNF)-dependent arthritis ($hTNF^{Tg/+}$, Supplementary Fig. 1b)[13]. Its restriction improved clinical disease parameters, independent of weight, and led to decreased formation of bone resorbing osteoclasts, suggesting a predominant effect of recArg1 treatment on osteoclast formation in arthritis (Fig. 1b). Finally, we systemically depleted arginine in collagen-induced arthritis (CIA), a model also dependent on adaptive immunity (Supplementary Fig. 1c)[14]. In vivo imaging revealed attenuated osteoclast-specific paw cathepsin K (Ctsk) amounts during arginine restriction, accompanied by decreased clinical scores, unconnected from weight gain (Fig. 1c). Local *Ctsk*, *Acp5* and *Tnf* transcripts increased in paws of ill versus healthy animals, confirming that the disease was linked to elevated osteoclastogenesis and inflammation. Although arginine depletion decreased *Ctsk* and *Acp5* mRNA expression, no effects were observed on *Tnf* mRNA, indicating that arginine depletion had a preferential effect on the transcriptional program of osteoclastogenesis rather than a global effect on inflammation (Supplementary Fig. 1d). Concordantly, while a secondary collagen challenge ex vivo induced robust splenocyte proliferation in diseased versus healthy animals, we detected no difference in recArg1-treated versus sham-treated arthritic animals (Supplementary Fig. 1e). Systemic arginine amounts in healthy animals were 200–300 µM, slightly above adult humans[15] and as expected, depleted by recArg1 during disease (Fig. 1d). This was associated with increases in ornithine and other AAs (Supplementary Fig. 1f). Of note, in CIA, myeloid populations including osteoclast precursors were unaffected by arginine restriction, as was the systemic RANKL/OPG ratio in all models studied (Supplementary Fig. 1g–i), suggesting that decreased osteoclast numbers found in arthritis were due to differentiation of osteoclasts. Collectively, we concluded that recArg1 exerts beneficial effects in murine arthritis, likely by dampening osteoclastogenesis.

To validate a potential cross-talk between the arginine pathway and bone degrading osteoclasts in patients suffering from rheumatoid arthritis (RA), we evaluated patients with respect to their osteoclast activity, grouping them into those suffering from erosive and non-erosive RA. Serum crosslaps, collagen degradation fragments produced from bone resorbing osteoclasts, confirmed their enhanced activity in erosive versus non-erosive patients (Fig. 1e). Versus healthy controls, arginine serum amounts exhibited a tendency to be enhanced in humans with erosive RA (Fig. 1f). In these patients, arginase 1 (ARG1) was significantly increased versus non-erosive patients and positively correlated with patient erosion scores (Fig. 1g). Together these data suggest of a physiological role of ARG1 in limiting osteoclast activity by arginine removal and illustrate an importance of the arginine pathway as well for human bone degradation in RA.

Extracellular arginine removal attenuates osteoclastogenesis. To investigate the molecular and cellular mechanisms associated with arginine restriction in bone disease, we next dissected osteoclastogenesis induced with M-CSF/RANKL in the presence or absence of recArg1 using different systems-type approaches. For simplicity, we refer to M-CSF + RANKL treatment as RANKL and M-CSF + RANKL + recArg1 treatment as recArg1, unless otherwise indicated (Fig. 2a). The control RANKL transcriptome consisted of 464 mRNA expression changes with 249 being down- and 215 upregulated. As expected, the latter included hallmark osteoclastic genes; *Fos*, *Nfatc1*, *Acp5* and *Jdp2*[16] and KEGG analysis showed enrichment for osteoclast differentiation (Supplementary Fig. 2a, b). At the mRNA level, perturbations in arginine biosynthesis and arginine and proline metabolism suggested an osteoclast-dependent importance for arginine metabolism (Fig. 2b). Indeed, recArg1 completely blocked murine osteoclastogenesis, an effect abolished by enzyme denaturation and independent of effects on cell viability and also inhibited human osteoclastogenesis (Fig. 2c–f). Treatment with recArg1 affected RANKL-induced gene expression already after 24 h, preceding RANKL effects on osteoclast fusion that mainly occurred at 72–96 h (Fig. 2e, Supplementary Figs. 2c and 3a), indicating recArg1 affects osteoclastogenesis by selectively reducing RANKL-dependent gene expression. These data suggest an important link between RANKL signalling and arginine metabolism and that depletion of extracellular arginine upon recArg1

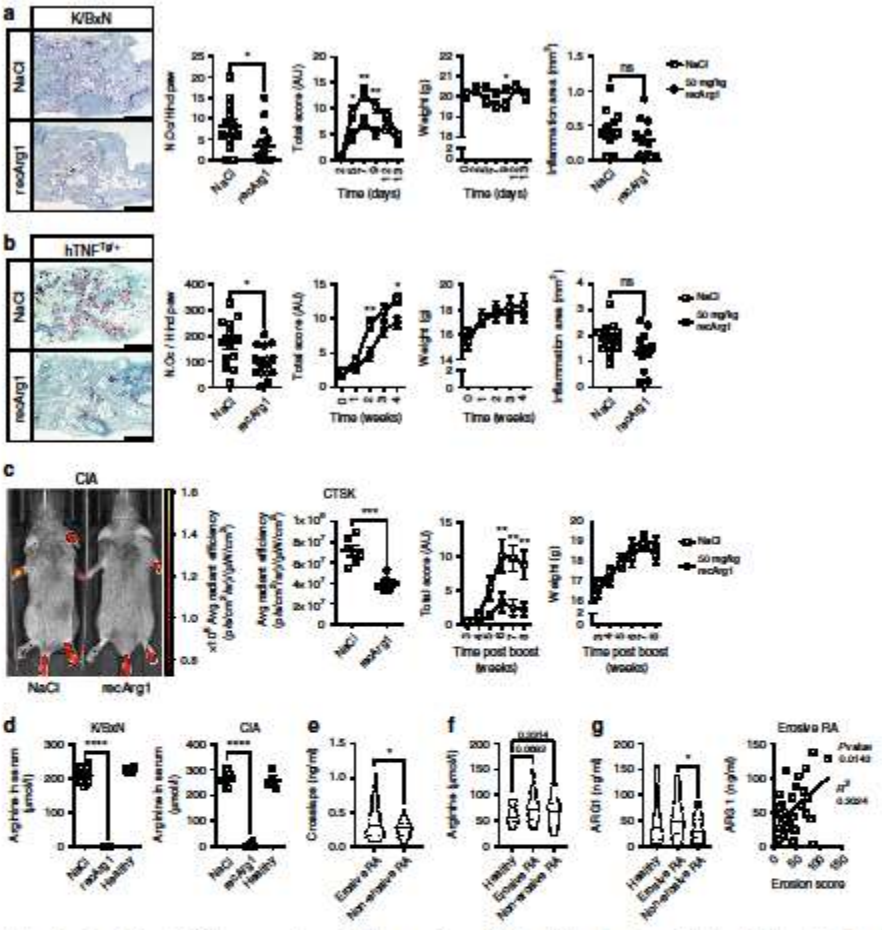


Fig. 1 Recombinant arginase 1 (recArg1) improves outcome in diverse murine arthritis models and arginase 1 is elevated in erosive RA patients. **a**, **b** Paw histology, osteoclast numbers per hind paw (N. Oc), total scores, weight and histology inflammation area of mice suffering from serum transfer arthritis (**a** K/BxN, NaCl $n = 12$, recArg1 $n = 13$ animals) or the hTNF$^{tg/+}$ mouse model (**b** NaCl $n = 13$, recArg1 $n = 14$ animals). Scale bar represents 1 mm. **c** Ctsk IVIS and quantification, total scores and weights of mice suffering from collagen-induced arthritis (CIA, NaCl $n = 13$, recArg1 $n = 14$ animals, Ctsk IVIS $n = 6$ animals). **d** RecArg1 depletes serum arginine; K/BxN ($n = 13$ animals per group), CIA (NaCl $n = 13$, recArg1 $n = 14$ animals). NaCl in **a**–**d** represents saline vehicle control group. **e** Serum crosslaps in the indicated patient groups (erosive RA $n = 30$, non-erosive RA $n = 29$ patients). **f** Serum arginine levels of all patient groups (healthy $n = 19$, erosive RA $n = 29$, non-erosive RA $n = 30$ patients). **g** Arginase 1 levels in all patient groups and correlation between arginase 1 and erosion score of patients suffering from erosive RA (healthy $n = 19$, erosive RA $n = 29$, non-erosive RA $n = 30$ patients). Data are mean ± SEM, $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, unpaired *t*-test (**a**–**c**, **e**–**g**), one-way ANOVA (**d**) and two-way ANOVA post-hoc pairwise comparisons with Bonferroni correction (**a**–**c**), linear regression (**g**). Source data are provided as a Source Data file.

treatment results in dynamic transcriptional changes that attenuate osteoclastogenesis.

To evaluate the importance of intracellular arginine degradation by cellular ARG1 during RANKL-induced osteoclast differentiation, we used the Tie2-cre system to selectively delete *Arg1* in hematopoietic osteoclast precursors[17]. Deficiency of cellular ARG1 within myeloid precursors did not affect osteoclast differentiation, indicating environmental decreases in extracellular arginine mediated by recArg1 exhibit distinct functions with respect to osteoclast differentiation versus those mediated by cellular ARG1 (Fig. 2g–h). Underscoring the negligible effects observed of conditional *Arg1* deletion on osteoclastogenesis, we observed persistent downregulation of *Arg1* transcript and protein levels post RANKL treatment during osteoclastogenesis of wildtype preosteoclasts (Fig. 2h–i, Supplementary Fig. 2a). To further test the specificity of extracellular arginine in RANKL signalling and osteoclastogenesis, we next treated innate immune cells with recArg1 during their differentiation in an identical manner to osteoclasts. RecArg1 exerted minor influences on macrophage and dendritic cell differentiation (Supplementary

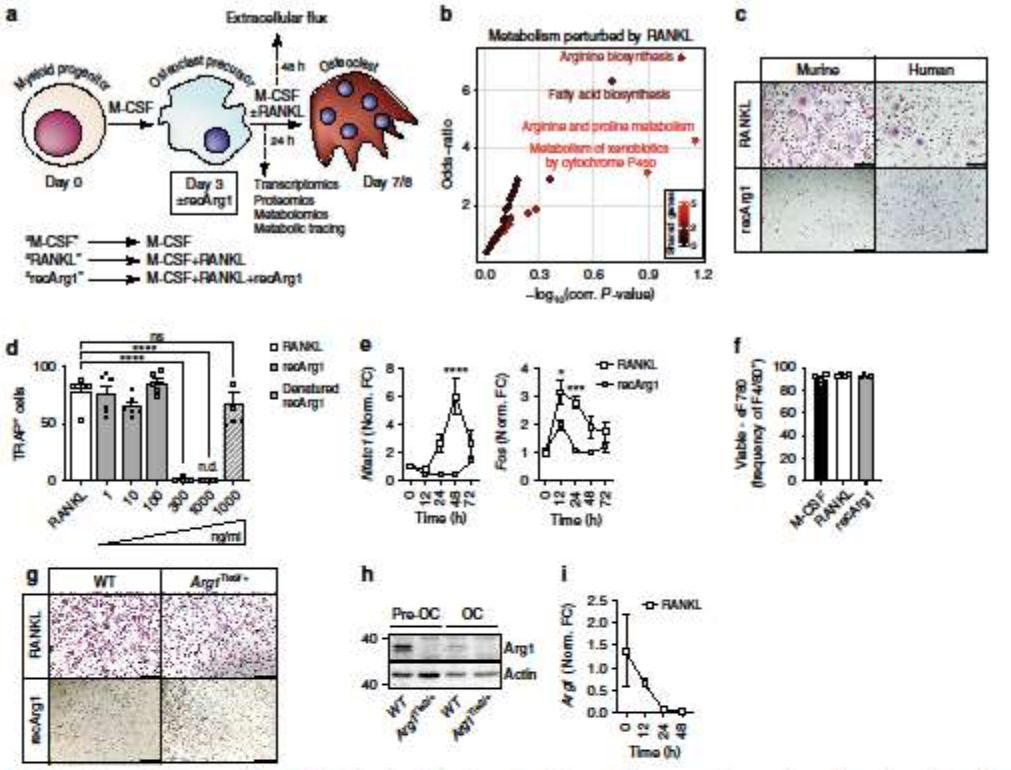


Fig. 2 Extracellular arginine is essential for RANKL-induced osteoclastogenesis. a Schematic of systems-wide approaches performed to understand the effect of recArg1 on RANKL dependent osteoclastogenesis. **b** KEGG metabolic enrichment of differentially expressed genes upon RANKL treatment. **c** Representative TRAP stainings depicting that recArg1 abolishes murine and human RANKL-induced osteoclastogenesis. **d** Quantification of murine data depicted in **c** ($n = 6$). **e** qRT-PCR time course of *Nfatc1* and *Fos* ($n = 3$). **f** Quantification of viable cells 24 h post RANKL/recArg1 treatment ($n = 4$). **g**, **h** Osteoclastogenesis is unchanged by *Arg1* deficiency. TRAP stainings (**g**) and Western blots of ARG1 in preosteoclasts (Pre-OC) and osteoclasts (OC) (**h**). **i** qRT-PCR time course of *Arg1* ($n = 3$). Data are mean ± SEM, $^*P < 0.05$, $^{***}P < 0.001$, $^{****}P < 0.0001$, one-way ANOVA (**d**) and two-way ANOVA post-hoc pairwise comparisons with Bonferroni correction (**e**). Scale bar represents 200 µm (**c**, **g**). Source data are provided as a Source Data file.

Fig. 4), suggesting differential environmental arginine requirements for M-CSF/GM-CSF versus RANKL signalling, the latter of which is well described to be important for multinucleated osteoclast formation.

RecArg1 counteracts RANKL cellular programmes. Up to now, we utilized exogenous recArg1 to deplete extracellular arginine to block osteoclastogenesis. We next determined the relative contributions of selective arginine depletion versus other potential effects of recArg1 including ornithine and urea generation, which are the products of the arginase reaction[6]. We compared preosteoclasts treated with RANKL (RANKL) versus recArg1 (RANKL/Arg-Depletion) or cultured in Arg-Free media (RANKL/Arg-Starvation). As controls, we re-supplemented arginine into Arg-Free media (RANKL/Arg-Rescue) and included M-CSF-treated preosteoclasts (M-CSF) starved of arginine (M-CSF/Arg-Starvation) (Fig. 3a). Plotting similarity between transcriptomic samples using multi-dimensional scaling (MDS), we found Arg-Starvation altered RANKL-induced effects, while the M-CSF transcriptome was minimally affected, confirming arginine requirements for RANKL signalling. RANKL/Arg-Depletion demonstrated more pronounced segregation from RANKL than RANKL/Arg-Starvation samples, whilst both profiles were distinct from M-CSF conditions (Fig. 3b and Supplementary Fig. 5a). Nonetheless, when we plotted RANKL-dependent gene expression in Arg-Depleted versus Arg-Starved cells, we observed highly correlated signatures (Fig. 3c). While recArg1 had the strongest effect on the RANKL transcriptome, most genes (119) significantly modified by Arg-Starvation (150) were shared with Arg-Depletion (420), although genes specific for both were identified (Supplementary Fig. 5b). Thus, we concluded that recArg1 largely acts on RANKL-induced MGC generation via arginine depletion, rather than via the products of the Arg1 reaction.

To dissect the cellular and metabolic changes caused by arginine restriction during osteoclast development, we next used a proteomic approach to probe whether Arg-Depletion in preosteoclasts caused an altered proteome (Fig. 3a, Supplementary Fig. 5c). We did not observe a global translational decrease upon Arg-Starvation or Arg-Depletion, and indeed, many

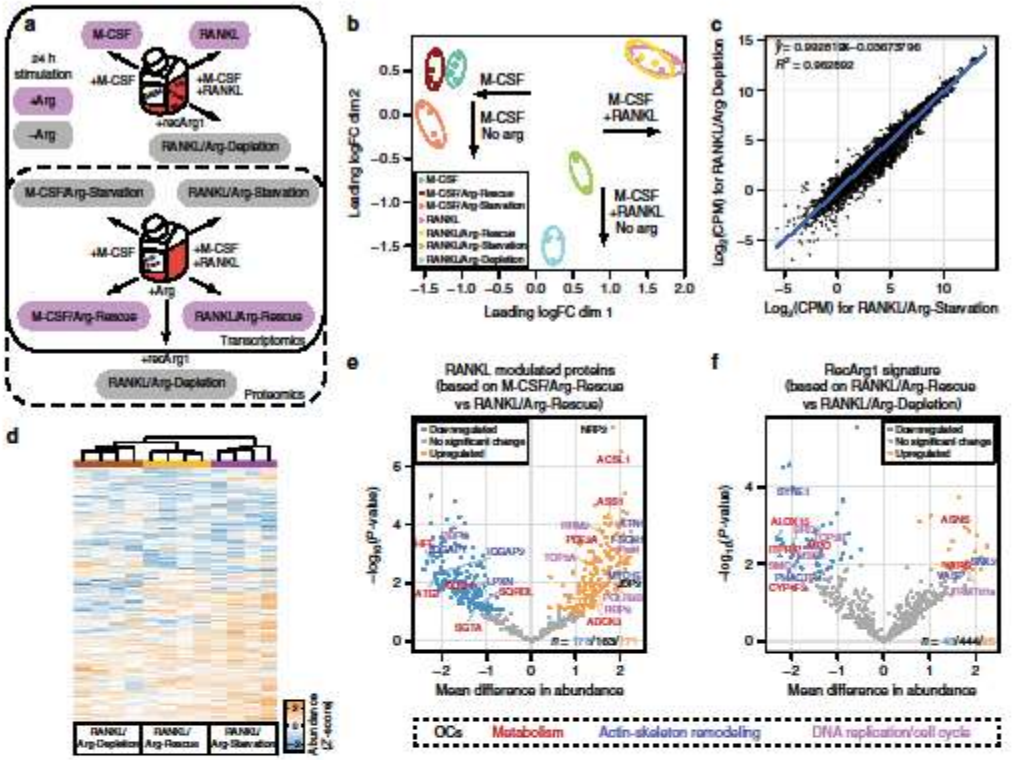


Fig. 3 Arginine presence specifically sustains RANKL gene and protein expression. a Workflow dissecting recArg1 specificity on RANKL signalling. Arg-sufficient (pink) and Arg-deficient (grey) conditions in arginine sufficient (aMEM) or deficient media (Arg-Free) depicted. **b** MDS of transcriptomic datasets in **a**, keeping all genes expressed in at least three of the samples. **c** Linear regression of average gene expression ($n = 4$) for RANKL/Arg-Depletion against RANKL/Arg-Starvation. **d** Heatmap of proteomics data in **a**, showing *Z*-score of abundance level per protein across all conditions (RANKL-conditions shown only, full heatmap in Supplementary Fig. 5d). **e** Volcano plot showing difference between M-CSF/Arg-Rescue versus RANKL/Arg-Rescue ($n = 4$). Highlighted are proteins modulated by RANKL in an arginine-dependent manner. **f** Volcano plot showing difference between RANKL/Arg-Rescue versus RANKL/Arg-Depletion ($n = 4$). Highlighted are proteins uniquely influenced by recArg1.

proteins were induced under low arginine conditions (Fig. 3d, Supplementary Fig. 5d). RANKL addition controlled expression of 349 out of 512 identified proteins with increases/decreases in 171/178 proteins, respectively (Fig. 3e). Only 43 RANKL-induced proteins were counteracted by recArg1 treatment, including those associated with osteoclastogenesis as we expected, as well as proteins involved in the actin cytoskeleton, DNA replication and the cell cycle (Fig. 3f). Thus, as found for the transcriptome, Arg-Depletion had a selective effect on proteins regulated by the RANKL pathway. Examining cell cycle status revealed that while arginine was required for RANKL-induced G_2/M cell cycle progression that preceded cell fusion, its withdrawal was equivalent to M-CSF- treated cells, suggesting diminished cellular proliferation (Supplementary Fig. 3a, b). Importantly, although arginine depletion through recArg1 modulated unique proteins, many mediated similar biological processes to the RANKL signature (Fig. 3f). Taken together, these data allowed us to conclude that RANKL cellular programming relies on extracellular arginine, as recArg1 largely counteracted the RANKL-dependent transcriptome and proteome.

Osteoclast oxidative metabolism is arginine dependent. Arginine withdrawal has often been associated with inhibition of mTORC1 signalling, a key regulator of cellular proliferation and translation[18]. Concordant with minor effects of arginine deprivation on preosteoclast translation, M-CSF/RANKL mediated mTOR phosphorylation and activation of downstream signalling pathways (pS6K, 4E-BP1) but was largely unaltered by recArg1 (Fig. 4a). In addition, genetic deletion of mTOR signalling components, *Tsc2* and *Rictor*[19], an arginine sensor important for mTORC1 responses (*Slc38a9*)[18,20] or components of the integrated stress response including the known AA sensor GCN2 (*Eif2ak4*) and one of its downstream transcriptional mediators (*Ddit3*)[21], could not bypass the effects of arginine withdrawal on osteoclastogenesis (Fig. 4b).

Having excluded roles of key arginine sensing pathways, we next probed the effect of arginine on osteoclast metabolism. Integrative network analysis of arginine effects on the RANKL signature, unifying transcriptomic and proteomic data, suggested that oxidative metabolism represented an important metabolic hub linked to arginine presence (Supplementary Fig. 6a and Supplementary Table 3). Indeed, we observed RANKL-induced

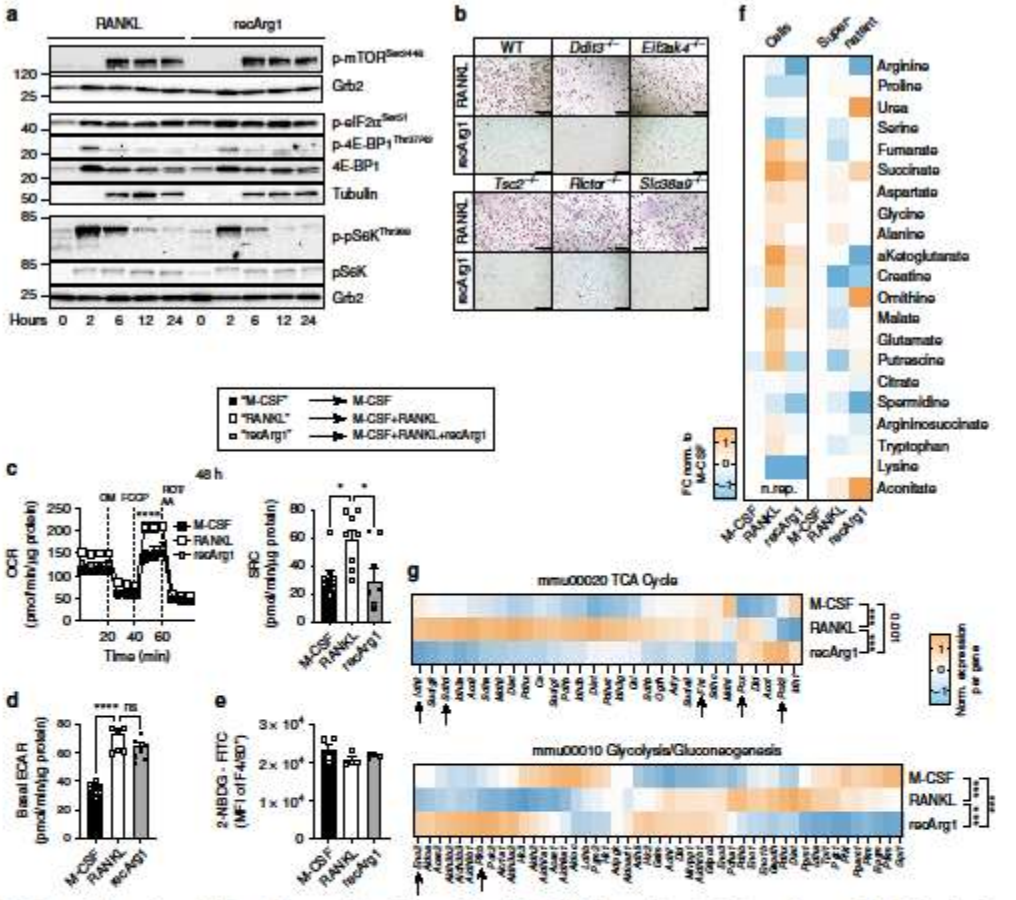


Fig. 4 Dynamic changes in metabolism to decreased extracellular arginine during RANKL-dependent osteoclastogenesis occur mTOR independently. a RecArg1 exerts negligible effects on RANKL-mediated mTOR-related pathways. **b** Deficiency in mTOR/AA-starvation components cannot rescue effects of Arg-Depletion. Scale bar represents 200 µm. **c–e** Oxygen consumption rate (OCR), spare respiratory capacity (SRC) and basal glycolytic rate (ECAR) of preosteoclasts ± recArg1 (**c**, **d** $n = 8$). Negligible effects of arginine deprivation on preosteoclast 2-NBDG (glucose) uptake (**e** $n = 4$). All data 48 h post stimulation. **f** RANKL differentially changed metabolites intra-/extracellularly ± recArg1. Data normalized against M-CSF (left) and represent mean ($n = 4$, RANKL intracellular $n = 3$; see Supplementary Fig. 7c, d) **g** Transcriptional profiles of selected KEGG pathways. TCA cycle and serine/purine biosynthetic enzymes indicated. Data represent mean ($n = 4$). Significance levels according to empirical permutation test, empirical *P*-values reported. Data are mean ± SEM, *$P < 0.05$, ****$P < 0.0001$, one-way ANOVA post-hoc pairwise comparisons with Bonferroni correction (**c–e**) and two-way ANOVA post Sidak's multiple comparisons test (**c**). Source data are provided as a Source Data file.

bursts in preosteoclast spare respiratory capacity at 48 h post stimulation, as previously described[22]. These RANKL-mediated increases were counteracted by recArg1, in a manner that was independent of changes in mitochondrial mass (Fig. 4c and Supplementary Fig. 7a, b)[23]. Moreover, versus the M-CSF control, RANKL markedly induced preosteoclast glycolytic rates, independent of changes in glucose uptake (Fig. 4d, e). Elevated preosteoclast glycolysis under RANKL treatment was independent of arginine as no decrease was observed with recArg1 addition (M-CSF+RANKL+recArg1) (Fig. 4d, e). Using a metabolomic approach, we observed RANKL-dependent induction of key intracellular tricarboxylic acid (TCA) cycle intermediates (e.g. malate, fumarate, succinate, α-ketoglutarate) 24 h post stimulation and only subtle changes in intracellular AAs (e.g. arginine, glycine, alanine) (Fig. 4f, Supplementary Fig. 7c). Oxygen consumption rates (OCRs) 24 h post RANKL treatment were unchanged compared to M-CSF, indicating that RANKL-induced changes in TCA metabolites preceded the observed boost in OXPHOS (Supplementary Fig. 7e).

Arginine is an integral part of the urea cycle, which is connected to the TCA cycle via a shunt through fumarate, linking both cycles[24]. RANKL-induced TCA metabolites were in agreement with robust TCA cycle enzymatic transcription, e.g. those converting isocitrate to α-ketoglutarate (Idh2), succinate to fumarate (Sdhd)

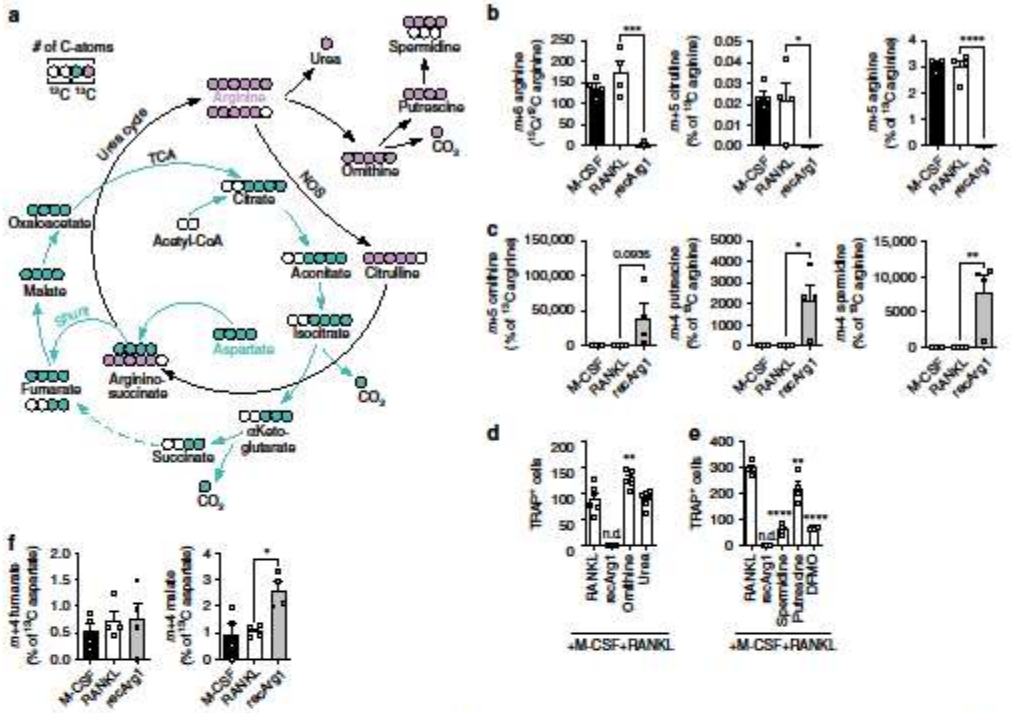


Fig. 5 Metabolic tracing of arginine and aspartate reveal arginine withdrawal instigates a dysregulated TCA cycle. a Schematic of the fate of the ^{13}C-labelled atoms deriving from arginine (pink) or aspartate (turquoise). **b** Ratio of intracellular $^{13}C_6$-labelled arginine to unlabelled $^{12}C_6$ arginine pool and percent $m+5$ citrulline and $m+5$ arginine normalized to the labelled $^{13}C_6$ arginine pool ($n=4$). **c** Intracellular $m+5$ ornithine, $m+4$ putrescine and $m+4$ spermidine normalized to $^{13}C_6$-labelled arginine input ($n=4$). **d** Excess of recArg1 degradation products ornithine and urea do not impact osteoclastogenesis ($n=6$). **e** Effect of extracellular polyamines and polyamine synthesis inhibitor difluoromethylornithine (DFMO) on osteoclastogenesis ($n=4$). **f** Intracellular $m+4$ fumarate and $m+4$ malate normalized to $^{13}C_4$-labelled aspartate input ($n=4$). Data are mean ± SEM, $^*P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, $^{****}P<0.0001$, one-way ANOVA post-hoc pairwise comparisons with Bonferroni correction (**b**–**f**). Source data are provided as a Source Data file.

and fumarate to malate (Fh1)[24] (Fig. 4g). RecArg1 further led to extracellular arginine depletion via conversion to ornithine and urea and reversed the intracellular build-up of malate, fumarate, succinate and α-ketoglutarate induced by RANKL (Fig. 4f and Supplementary Fig. 7c, d), suggesting arginine depletion was linked to a rewired TCA cycle. Indeed, recArg1 counteracted specific RANKL-induced TCA enzyme expression, while upregulating serine and purine biosynthetic enzymes (Pck2, Eno2, Phgdh) (Fig. 4g, Supplementary Fig. 2f)[8]. Consistent with the notion of diverted biosynthetic fluxes upon arginine depletion, increased expression of the pyruvate anabolic genes *Glud1* and *Gpt2* occurred (Supplementary Fig. 2d). Collectively, our data show that extracellular arginine is required for successful RANKL-dependent TCA cycle upregulation and consecutive increases in oxidative phosphorylation. Our data further suggest that either arginine turnover impacts the TCA cycle, or that the increases in OXPHOS by RANKL are linked to arginine-dependent transcription of TCA cycle enzymes and related metabolite induction or a combination of both.

RecArg1 instigates a dysregulated TCA cycle in osteoclasts. To distinguish between the possibilities noted above, we next performed ^{13}C-isotopologue labelling experiments. As the carbon backbone of arginine cannot serve as a direct source for TCA cycle intermediates[25], we supplied $^{13}C_6$ arginine in combination with $^{13}C_4$ aspartate in the first 24 h of osteoclastogenesis and measured intracellular isotopologue distribution via liquid chromatography–mass spectrometry (LC–MS) analysis. In this way, we could assess the contribution of each carbon source to both the urea and the TCA cycle simultaneously, as aspartate can be readily converted through argininosuccinate into fumarate via the TCA-Urea cycle shunt (Fig. 5a). $^{13}C_6$ arginine was readily taken up by both M-CSF- and RANKL-treated cells, as we detected 100–200 times more intracellular labelled arginine in relation to unlabelled arginine. In line with negligible increases in arginine transporter transcription (*Slc7a1*, *Slc7a2*) (Supplementary Fig. 2e)[26], we observed no difference in $^{13}C_6$-labelled arginine uptake upon RANKL treatment post 24 h. Both M-CSF and RANKL treated cells engaged in arginine recycling as indicated by $m+5$ labelling in both citrulline and arginine (Fig. 5b). Notably, as arginine was readily converted to ornithine and urea by recArg1 in the supernatant, we could barely detect labelled intracellular $^{13}C_6$ arginine upon recArg1 treatment (Figs. 4f, 5b, Supplementary Fig. 7d). Relative to the low amounts of

intracellular $^{13}C_6$ arginine, under arginine depletion, we detected high quantities of $^{13}C_5$-labelled ornithine and its downstream metabolites putrescine and spermidine, because the enhanced ornithine generated by recArg1 was further catabolized to form polyamines and came from the extracellular $^{13}C_6$ arginine supplied (Fig. 5c). Suggestive of uptake and/or synthesis from $m+5$ ornithine occurring under ^{13}C labelling, total intracellular amounts of the polyamines putrescine and spermidine were decreased upon recArg1 treatment in comparison to RANKL (Figs. 4f, 5c, Supplementary Fig. 7c). RNA sequencing analysis demonstrated that similar to macrophages, ornithine transcarbamylase (OTC) was not expressed in preosteoclasts[27] (see deposited data), therefore ornithine cannot be converted to citrulline in these cells. Together, these data indicate that recArg1 treatment converts all extracellular arginine to ornithine and the polyamines as demonstrated by the lack of $m+5$ citrulline and $m+5$ arginine (Fig. 5b, c).

Excess of extracellular ornithine and urea did not have any consequences on osteoclastogenesis, proving that the recArg1-instigated blocks in osteoclastogenesis were not due to increases in these metabolites (Fig. 5d). It was previously shown that extracellular polyamines negatively regulate osteoclastogenesis[28]. However, although recArg1 led to a shift in putrescine and spermidine amounts, their addition at excess (500 µM putrescine and 10 µM spermidine) during osteoclastogenesis caused an approximately 30 and 75% decrease in TRAP-positive cells, respectively. Thus, although accumulation of polyamines upon recArg1 might negatively impact osteoclastogenesis, their increased synthesis cannot fully account for the complete absence of TRAP-positive cells observed under arginine deprivation (Fig. 5e). As we wanted to test if inhibitory effects of recArg1 treatment on osteoclastogenesis were mediated by decreased intracellular polyamine synthesis (Fig. 4f, Supplementary Fig. 7c), we incubated cells with alpha-difluoromethylornithine (DFMO), thereby blocking conversion of ornithine to putrescine[29]. Despite DFMO reducing osteoclastogenesis, multinucleated mature osteoclasts were still observed (Fig. 5e). Overall, these data suggest an importance for polyamines in the context of osteoclastogenesis. However, the effects of recArg1 cannot be solely accounted by deregulated polyamine metabolism.

To prove an operating urea–TCA cycle shunt in preosteoclasts, we simultaneously performed a $^{13}C_4$ aspartate tracing. Suggesting the aspartate-argininosuccinate–fumarate shunt functioned in preosteoclasts independent of arginine presence, we observed equal $m+4$ fumarate labelling in all conditions (Fig. 5f). Notably, as fumarate was $m+4$ labelled, we concluded the carbon source stemmed from argininosuccinate, as direct conversion from aspartate to oxaloacetate would have yielded $m+2$ fumarate (Fig. 5a). Consistent with arginine deprivation globally downregulating TCA cycle enzymes (Fig. 4g), we observed higher abundance of $m+4$ malate relative to $^{13}C_4$ aspartate input following recArg1 treatment compared to M-CSF and RANKL control conditions (Fig. 5f). Although arginine and its catabolic products cannot directly enter the TCA cycle, the metabolic tracing data utilizing fully labelled aspartate strengthen the notion that arginine presence is required for a functioning TCA cycle.

AA precursors compensate for AA absence in osteoclastogenesis. So far we showed that the arginine-dependent dysregulated TCA cycle-initiated cell cycle arrest was not associated with known properties of quiescence including decreased glycolysis, reduced translation rates or mTORC1 deactivation (Figs. 3, 4 and Supplementary Fig. 3b). This suggested a discordant/distinct type of preosteoclast metabolic quiescence, different from that of T cells or hematopoietic stem cells[30–32]. We noted an in-between state of arginine-restricted preosteoclasts between the M-CSF and the M-CSF/RANKL signature (Fig. 3b), raising the possibility that arginine depletion resulted in a 'poised' preosteoclast state that was associated with a transient reversible osteoclastogenesis-block. Indeed, osteoclast formation could be restarted even after prolonged arginine starvation by arginine resupplementation and even exhibited a tendency towards being increased relative to controls (Supplementary Fig. 3c, d). Notably, the Arg-Starvation-mediated inhibition of polykaryons was restored by argininosuccinate and citrulline supplementation, but not by ornithine and proline, indicating that select arginine precursor metabolites bypass the dependence of arginine for RANKL signalling (Fig. 6a). α-Ketoglutarate was not able to rescue arginine absence (Supplementary Fig. 3e), suggesting that arginine scarcity could not be compensated by direct TCA cycle refuelling but only by AA derivatives. Together, dynamic metabolic adaptation to environmental arginine deficiency sustains cellular fitness until intermediates are re-supplied and RANKL-dependent cellular differentiation programs continued. In line with this finding, in vivo recArg1 administration hindered but did not completely abolish osteoclastogenesis, suggesting that arginine intermediates can compensate for arginine scarcity in vivo (Fig. 1).

To evaluate whether other AAs were essential for RANKL cellular programming, we next cultured preosteoclasts in media selectively devoid of AAs and demonstrated commonalities in metabolic blockage during osteoclastogenesis to a nutrient-limited environment, especially regarding essential AAs (Fig. 6b, Supplementary Fig. 8, Supplementary Table 1). As α-ketoisocaproate, ketoisoleucine and phenylpyruvate are intermediates of the osteoclastogenesis essential AAs leucine, isoleucine and phenylalanine, respectively, and can enter the TCA cycle eventually via acetyl-CoA[33], we next evaluated their compensatory potential for the aforementioned AAs. In line with TCA/urea cycle intermediates citrulline and argininosuccinate compensating for arginine deficiency (Fig. 6a), non-proteogenic metabolites α-ketoisocaproate, ketoisoleucine and phenylpyruvate could rescue defects in osteoclastogenesis that occurred in the absence of their related AA (Fig. 6c). Therefore, these data suggest that osteoclasts can adjust to the absence of environmental AAs by utilizing their respective precursors or intermediates for successful osteoclastogenesis.

IL-4 MGCs and osteoclasts share arginine requirements. Noting that mononuclear innate immune cell in vitro differentiation efficiency is independent of arginine absence (Supplementary Fig. 4) and that RANKL induces giant cell formation, we next examined whether the concept of arginine reliance holds true for other MGCs. IL-4 induces myeloid cell multinucleation[34], resulting in MGCs that are distinct from osteoclasts but share common characteristics[35]. Strikingly, extracellular arginine removal using recArg1 prevented IL-4-generated MGCs (Fig. 7a). Identical to the unremarkable requirements of intracellular ARG1 deficiency for RANKL-induced osteoclastogenesis, IL-4 robustly induced MGCs in the absence of *Arg1* (Fig. 7b). Using RT-PCR, we confirmed the previously observed effect of recArg1 on the expression of select RANKL-modulated metabolic enzymes *Pck2* and *Sdhd* (Figs. 5g, 7c). RANKL- or IL-4-dependent regulation of *Sdhd* was further associated with increased expression of the MGC signature gene *Dcstamp*. Notably, recArg1 treatment exhibited identical effects on both enzymes in both cell types, inducing *Pck2* and downregulating *Sdhd* (Fig. 7c). Further, arginine starvation abolished IL-4-induced MGC formation and this was compensated by the same urea cycle intermediates in MGCs as in osteoclasts (Figs. 6a and 7d, e) and associated with improved OXPHOS (Fig. 7f).

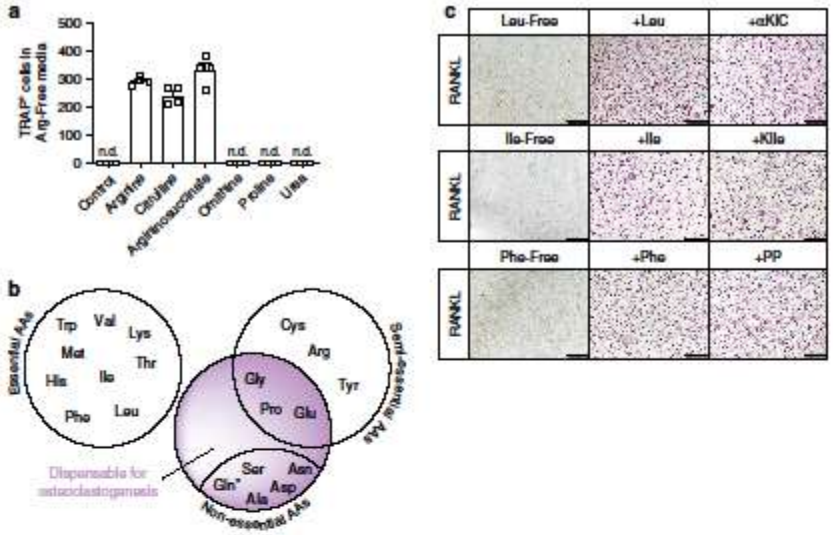


Fig. 6 Intermediates or precursors of amino acids including arginine can compensate for amino acid absence in osteoclastogenesis. a Urea cycle metabolites can/cannot rescue arginine requirements in osteoclastogenesis in Arg-Free media ($n = 4$). **b** Schematic demonstrating which amino acids are essential (white) versus non-essential (pink) for osteoclastogenesis. Star relates to TRAP-positivity with decreased occurrence of MGCs ($n = 4$). **c** Immediate intermediates of leucine, isoleucine and phenylalanine degradation compensate for the lack of these amino acids and rescue osteoclastogenesis. Representative TRAP stainings are shown ($n = 4$). αKIC alpha-ketoisocaproate, Ile isoleucine, KIle ketoisoleucine, Leu leucine, Phe phenylalanine, PP phenylpyruvate. Scale bar represents 200 µm. Data are mean ± SEM. Source data are provided as a Source Data file.

Discussion

Formation of MGCs such as osteoclasts is complex with previous studies identifying a range of essential molecules, ranging from surface receptors such as RANK to co-stimulatory molecules and transcription factors[2,16,35]. Here, we demonstrate that in addition to those factors, the extracellular availability of the AA arginine is critical for RANKL and IL-4 induced differentiation of myeloid precursor cells. While previous studies highlighted an importance for cellular ARG1 in negatively regulating osteoclastogenesis[36], here we demonstrate that extracellular arginine is essential for RANKL-induced metabolism. We found extracellular arginine presence is critical for RANKL to elicit cellular programs and metabolic changes, especially related to cellular respiration. Arginine withdrawal during RANKL signalling was intrinsically linked to downregulation of TCA cycle enzymes and metabolites. Metabolic tracing experiments further confirmed that a dysregulated TCA cycle under arginine withdrawal occurred, associated with malate accumulation derived from labelled aspartate. These changes occurred early during osteoclastogenesis and preceded arginine-dependent dampening of RANKL-induced bursts in cellular respiration. Together, we demonstrate that in preosteoclasts extracellular arginine presence is required for efficient TCA function in response to RANKL stimulation. Akin to osteoclasts, IL-4 induces the formation of MGCs[34]. These cell types displayed clear similarities in arginine-dependent oxidative metabolism, strengthening the notion of arginine requirements to meet MGC energy demands. Equivalent to osteoclasts, IL-4 MGCs were able to form by utilizing arginine precursors to adapt to arginine scarcity. In contrast to MGC differentiation, the presence of arginine was not required for the terminal differentiation of other myeloid cells, such as bone-marrow-derived macrophages or dendritic cells, suggesting specificity for polykaryon formation.

Interestingly, mTOR is a central metabolic integrator, whose role in osteoclastogenesis remains controversial[37,38]. Importantly, although numerous studies in transformed cells suggested absolute extracellular arginine requirements for mTOR activation[18], our data indicate that arginine is not required to sustain mTOR signalling during RANKL-dependent osteoclastogenesis. Indeed, upon arginine deprivation, we did not observe a reduction in translation rates or diminished glycolysis, but rather a type of preosteoclast metabolic quiescence associated with cell cycle arrest and distinct to that previously described in other cell types such as hematopoietic stem cells[31]. The effects of arginine deprivation on osteoclastogenesis were fully reversible upon arginine resupplementation and independent of increases in recArg1 degradation products, such as ornithine and urea. Notably, arginine absence could be compensated by its precursors citrulline and argininosuccinate, suggesting requirements of a functioning arginine recycling machinery during osteoclast development.

Although our data demonstrate the crucial role of extracellular arginine in osteoclastogenesis, absence of other AAs phenocopied the blocks observed under arginine absence. Thus, osteoclasts most likely couple extracellular AA amounts to their developmental program, as scarcity in single essential AAs could be compensated by osteoclast intrinsic re-synthesis through supplementation of its derivatives or precursors. While here we established TCA cycle dysregulation upon arginine withdrawal, we cannot be certain if this is true for the remaining osteoclastogenesis essential AAs identified in this study. Nonetheless, given osteoclastogenesis is an energy demanding process, it is plausible that similar mechanisms of energy shifts under AA starvation are in place as observed under arginine withdrawal. A plausible model for the effects of essential AAs on the selective metabolic

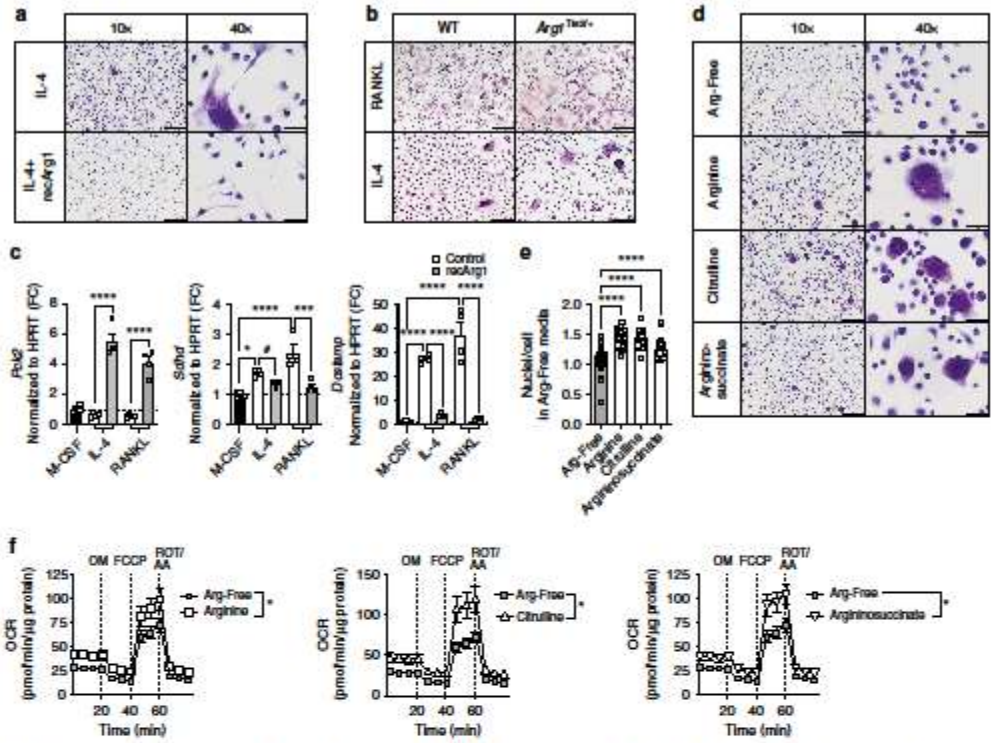


Fig. 7 Osteoclasts and IL-4-induced multinucleated giant cells (MGCs) share mechanisms of metabolic adaptation in nutrient scarcity. **a** Representative H&E stainings of MGCs ± recArg1. **b** Representative H&E stainings of MGCs from *Arg1*$^{Tie2/+}$ and WT littermate controls. **c** qRT-PCR of *Pdk3*, *Sdhd* and *Dcstamp* in IL-4- (M-CSF + IL-4) and RANKL (M-CSF + RANKL)-induced MGCs ± recArg1. All data correspond to day 7, # corresponds to $P = 0.01$ assessed by *t*-test ($n = 4$). **d**, **e** Representative H&E stainings (**d**) and quantifications (**e**) of nuclei per cell of IL-4-generated MGCs in the absence of arginine (Arg-Free Media), re-supplemented with arginine, citrulline or argininosuccinate. Data represent cells counted in 50 random frames ($n = 50$). **f** IL-4-induced MGC oxygen consumption rate (OCR) in the absence of arginine (Arg-Free Media), re-supplemented with arginine, citrulline and argininosuccinate. Significance was calculated by unpaired *t*-test between area under the curve (AUC) of indicated conditions (Arg-Free $n = 10$, arginine $n = 12$, citrulline $n = 11$, argininosuccinate $n = 11$). Data are mean ± SEM, $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, one-way ANOVA Tukey post-hoc test and *t*-test (**c**), one-way ANOVA post-hoc pairwise comparisons with Bonferroni correction (**e**) and two-way ANOVA post Sidak's multiple comparisons test (**f**). Scale bar represents 200 µm (10×), 50 µm (40×) (**a**, **d**) and 200 µm (RANKL) and 100 µm (IL-4) (**b**). Source data are provided as a Source Data file.

'poised' state of osteoclastogenesis is that osteoclast progenitors 'integrate' the lack of a key AA(s), which then triggers a pathway that causes a selective block in the TCA cycle. For example, AA-starved osteoclast progenitors could induce the downregulation or degradation of key central metabolic enzymes, linking the pathways. In this model, the effects of low AAs would be rapidly reversible upon AA restoration and osteoclast development could proceed, as we observed. Several examples of such reversible metabolite signalling are known, including oxygen and iron sensing[39,40]. Clearly, such an 'information transfer system' would need to be independent from the mTOR and GCN2 pathways as we found. Nevertheless, a 'sensor' must integrate information from all the AAs described herein that block osteoclastogenesis, or individual sensors for each AA may exist. However, such a model needs to also account for the 'rescue' effect of each AA precursor such as citrulline. Ultimately, we propose that AA presence is required to sustain oxidative metabolism, which is essential for multinucleation, and that this process may involve yet to be discovered metabolic regulatory pathways. Together, our data establish how availability of environmental AAs, especially arginine, control polykaryon developmental programs and metabolism. They further imply therapeutic strategies for arginine depletion in MGC-mediated diseases.

Methods

Human studies. Human patient serum was used freshly for osteoclastogenesis assays or processed and stored until analysis according to standard operating procedures by the MedUni Wien Biobank, a central facility included in a certified quality management system[41]. All study subjects provided informed consent. This study was approved by the local ethics committee of the Medical University Vienna (EK #559/2005). Arginine and ARG1 levels were assessed using ELISA (Arginine: Immundiagnostik AG #K7733; human ARG1: Antibody Online Gmbh #ABIN1113579).

Animals and arthritis models. Female wildtype animals (C57BL/6J, RRID: IMSR_JAX:000664 and DBA/1J, RRID: IMSR_JAX:000670) used for in vivo arthritis models were purchased from Charles River Laboratories. Serum-transfer arthritis was induced by intraperitoneal application of 150 µl of K/BxN serum on

day 0 and day 2 in C57BL/6J and animals were harvested post 13 days treatment[42]. Human tumour necrosis factor ($hTNF^{Tg/+}$, RRID: MGI:3053718) transgenic animals were identified from tail tissue PCR using hTNF^{Tg} primers (5′-TACCC CCTCCTTCAGACACC-3′ and 5′-GCCCTTCATAATATCCCCCA-3′) and were used for experiments at 5–6 weeks of age. Mice were harvested after 5 weeks of treatment. Female wildtypes (DBA/1J) were immunized subcutaneously with 50 µg of chicken type II collagen (Sigma-Aldrich #C9301) in 50 µl of H_2O, emulsified in 50 µl of Freund's complete adjuvant (Sigma-Aldrich #F5881) that was enriched with 10 µg/ml *Mycobacterium tuberculosis* (Difco/BD Biosciences #H37Ra), on day 1 and day 21. Mice were harvested after 8 weeks of treatment. Treatment regimen comprised a dose of 50 mg/kg recArg1 or NaCl administered intra-peritoneally twice weekly, starting 1 day before disease induction or at 5–6 weeks of age regarding $hTNF^{Tg/+}$ mice. $hTNF^{Tg/+}$ mice were a kind gift of the George Kollias[13]. Arginine levels were assessed using metabolomics or via ELISA (see below). Mouse TNFSF11/RANKL (Boster #EK0843) and mouse TNFRSF11B/osteoprotegerin (abcam #ab203365) serum levels were measured using ELISA. All animal procedures were approved by the local ethics committee of the Medical University Vienna (BMWFW-66.009/0013-V/3b/2019 and BMWFW-66.009/0227-WF/V/3b/2017) and were conducted in strict accordance with Austrian law.

Clinical assessment of murine arthritis. Clinical signs of arthritis are described by a well-established semi-quantitative double-blind score system[43]. Swelling per paw was recorded: 0 to 3 (0 = no swelling; 1 = mild swelling of toes/ankle; 2 = moderate swelling of toes/ankle; 3 = severe swelling of toes/ankle). Grip strength per paw was assessed on a wire mesh: 0 to 3 (0 = normal grip strength; 1 = mildly reduced grip strength; 2 = moderately reduced grip strength; 3 = severely reduced grip strength). Total score was calculated by combining scores of swelling and grip strength of all four paws. Inclusion criteria for the $hTNF^{Tg/+}$ control group was a combined clinical score of over 5, at 10 weeks of age.

Histological analyses. Hind paws were fixed in 4.5% formalin for 6 h and then decalcified in 14% pH 7.2 EDTA/ammonium hydroxide buffer (Sigma #318604) at 4 °C until bones were pliable. Afterwards, 2 µm decalcified paraffin-embedded sections were prepared and stained with haematoxylin and eosin (H&E) and TRAP (#387A, Leukocyte Phosphatase Staining Kit, Sigma Diagnostics). In brief, slides were stained for 10 min with 1:5 diluted Meyer's hemalum (Merck #1.09249.0500), rinsed with distilled water, differentiated in 1% HCl ethanol and rinsed again for 10 min. Afterwards, slides were stained in eosin working solution (300 ml 1% Eosin Sigma #318906, 600 ml distilled water, 0.1 ml acetic acid 100%) for 15 min. Slides were then rinsed with distilled water, 96% ethanol, *n*-Butyl-acetate and mounted. Regarding TRAP staining, slides were stained for 1 h at 37 °C with the TRAP staining solution (250 µl naphthol AS-BI phosphoric acid, 1000 µl acetate solution, 500 µl tartrate solution, 45 ml distilled water; Sigma-Aldrich #387A) protected from light. Afterwards, slides were developed for 2 min at 37 °C with a mix of 250 µl of Fast garnet GBC base solution and 250 µl of sodium nitrite solution. Nuclei were stained with Meyer's hemalum as described above and slides mounted in Aquatex (Merck #108562). Sections were analysed using an Axioskop 2 microscope (Carl Zeiss MicroImaging) and Osteomeasure Analysis System (OsteoMetrics) to quantify the areas of inflammation, erosion and osteoclast numbers.

Ctsk in vivo imaging. Cat K 680 FAST (Perkin Elmer #NEV11000) was reconstituted and 100 µl per animal injected intravenously according to the manufacturer's guidelines. Ctsk intensity (Avg Radiant Efficiency) was determined by IVIS® Lumina Series III (Perkin Elmer) 24 h post injection.

In vitro restimulation of spleen cells with collagen. Spleen cells were harvested, passed through a nylon mesh, and cultured at a density of 2×10^6 cells/ml in RPMI 1640 (L-glutamine, 10% FCS, penicillin/streptomycin, β2-mercaptoethanol) and stimulated with 100 µg/ml of chicken type II collagen (Sigma-Aldrich #C9301) for 72 h. During the last 18 h of culture, cells were incubated with 1 µCi/well of ^{3}H-thymidine, to quantify proliferation using a scintillation counter (Beckman).

Osteoclastogenesis and staining. Wildtype animals (C57BL/6J) used as bone marrow donors were bred at the Medical University Vienna. Hematopoietic stem cells of the bone marrow were isolated and cultured in complete MEMα (Gibco #32561037) containing 5% Pen-Strep (Gibco #15140122) and 10% foetal calf serum (FCS, Gibco #10082147) supplemented with 100 ng/ml M-CSF (R&D Systems #416). After 3 days, cells were harvested, plated and cultured in Full-MEMα supplemented with 30 ng/ml M-CSF and/or 50 ng/ml RANKL (R&D Systems #462) for another 3–4 days including one medium change on day 6. Osteoclasts were defined as multinucleated cells (≥3 nuclei) with TRAP positivity (Sigma-Aldrich #387A). RecArg1 was used in vitro at 1 µg/ml starting at day 3 of osteoclastogenesis, unless otherwise stated. Arg-Free and AA-free media (both Gibco custom MEMα based on #32561037) was supplemented with dialyzed FCS (Gibco #26400036) and AAs and used starting day 3 of osteoclastogenesis. Preosteoclasts in tryptophan and cysteine-free medium were cultured in 5% dialyzed FCS to control for serum-bound AAs[44] (Supplementary Table 1). 2-(difluoromethyl) ornithine (DFMO) was a kind gift from the Wester Lab (MUSC) and used at a concentration of 2.5 mM. $Tsc2^{fl/fl}$ and $Rictor^{fl/fl}$ LysMcre$^{+/-}$ have previously been described[45,46]. *Eif2ak4*$^{-/-}$ (B6.129S6-*Eif2ak4*$^{tm1.2Dmor}$/J) and *Atg7*$^{fl/fl}$*Tie2*$^{+/-}$ bone marrow was provided by the Murray Lab[17,44]. *Ddit3*$^{-/-}$ (B6.129S(Cg)-*Ddit3*$^{tm1.1Dron}$/J) animals were acquired from the Jackson Laboratories. *Slc38a9*$^{-/-}$ (unpublished) bone marrow was a kind gift from the Superti-Furga Lab[20]. Human monocytes were isolated from whole blood using Histopaque-1077 and Histopaque-1119 (Sigma #10771 and #11191) and CD14$^+$ cells sorted using CD14-PerCP-Cy5.5 (eBioscience #45-0149, 61D3). Monocytes were resuspended in complete MEMα medium (5% Pen-Strep, 10% FCS) with 25 ng/ml M-CSF. After overnight incubation with M-CSF, human RANKL (R&D Systems #390) was added at 25 ng/ml for another 3–4 days including one medium change on day 6. Osteoclasts were defined as multinucleated cells (≥3 nuclei) with TRAP positivity. Blood from healthy donors included in this study was provided by the Biobank for patients with rheumatic diseases and were approved by the Clinical Research Ethics Committees of the Medical University Vienna (EK Nr. 559/2005). An informed consent was obtained from all subjects.

Giant cell formation and staining. Bone marrow cells were cultured for 3 days in Full-MEMα containing 15% v/v L929-supernatant on 6-well non-treated plastic plates. Adherent macrophages were detached using 1× PBS (10 mM EDTA). Cells were plated in Permanox® 8-well chamber slides in Full-MEMα containing 15 ng/ml IL-4 (R&D Systems #404) and described stimuli for another 3 days, after which the medium was renewed. Three days later, cells were stained using H&E. H&E-stained slides were analysed using an Axioskop 2 microscope (Carl Zeiss MicroImaging). For nuclei counts, cells were washed with PBS and stained (0.5% PFA, 0.1% Triton X-100, 0.1% Hoechst 33342, 0.15% Phalloidin AlexaFluor488 in 1× PBS) for 20 min in the dark. Afterwards, cells were washed again and kept in PBS for imaging. Twenty-five random images per well were taken and analysed using Cell Profiler[47].

Macrophage and dendritic cell differentiation. Macrophage or dendritic cells were differentiated from hematopoietic stem cells cultured in presence of 30 ng/ml M-CSF (macrophages) or 20 ng/ml granulocyte macrophage colony-stimulating factor (GM-CSF, R&D #215 DCs) and 5 ng/ml IL-4 (dendritic cells), with complete medium changes on days 3 and 6. On day 7, cells were harvested and analysed using flow cytometry.

Quantitative PCR. Total RNA was extracted from paws and cells using RNA isolation kits (QIAGEN RNeasy Peqlab TriFast). Reverse-transcription was performed using commercially available kits (Applied Biosystems). SYBR Green Supermix (Bio-Rad Laboratories) was used for the qPCR reaction. Postamplification melting curve analysis and water controls were included to ensure absence of primer dimers. To obtain sample-specific ΔCt values, normalization to hypoxanthine phosphoribosyltransferase 1 (*Hprt*) within each sample was performed. Data are shown as fold change, where $2^{-\Delta\Delta Ct}$ values were calculated (ΔΔCt = ΔCt treatment − ΔCt control). RT-PCR was performed using the following primers: *Hprt*: 5′-CGCAGTCCCAGCGTCGTG-3′ and 5′-CCATCTCCTTCATGACATC TCGAG-3′; *Mstn1*: 5′-GACAGACATCCGGAGGAAGA-3′ and 5′-AGCCTTCT CCACGAAAATGA-3′; *Ctsk*: 5′-GGAAGAAGACTCACCAGAAGC-3′ and 5′-GT CATATAGCCGCCTCCACAG-3′; *Acp5*: 5′-ACAGCCCCCACTCCCACCCT-3′ and 5′-TCAGGGTCTGGGTCTCCTTGG-3′; *Tnf*: 5′-CCACCACGCTCTTCTG TCTAC-3′ and 5′-AGGGTCTGGGCCATAGAACT-3′; *Rn*: 5′-AGCCCAGACCT GCAGTGGCT-3′ and 5′-GCGCTCTGCCTCCTGACACG-3′; *Mt-ND2*: 5′-AGGG ATCCCACTGCACATAG-3′ and 5′-SCTCCTCATGCCCCTATGAAA-3′; *Arg1*: 5′-GGAAACGCCAATGAAGAGCTG-3′ and 5′-GCTTCCAACTGCCAGACTGT-3′; *Dcstamp*: 5′-TCCTCCATGAACAAACAGTTCCAA-3′ and 5′-AGACGTGGTT TAGGAATGCAGCTC-3′; *Sdhd*: 5′-TGGTCAGACCCGCTTATGTG-3′ and 5′-GGTCCAGTGGAGAGATGCAG-3′; *Pck2*: 5′-TGCCAGGCTGGAAAGTGGAG TGT-3′ and 5′-GCAACCCCAAAGAAGCCGTTCTCA-3′.

Western blot. Cells were washed with ice-cold PBS, scraped and centrifuged at 400 × *g* for 5 min at 4 °C and the resulting pellet was lysed with lysis buffer ((20 mM Hepes pH 7.4, 400 mM NaCl, 25% v/v glycerol, 1 mM EDTA, 0.5 mM NaF, 0.5 mM Na_3VO_4, 0.5 mM DTT) supplemented with Triton-X, PMSF, PIM and RPI shortly before use). The homogenate was cleared by centrifugation at 4 °C for 10 min at 16,000 × *g* and the supernatant containing the protein fraction recovered. Protein concentration in the supernatant was determined using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific #23225). A total of 15 µg of proteins were resolved by SDS-PAGE and transferred to PVDF membranes (GE Healthcare #10600023). Membranes were blocked with 1% Western Blocking Reagent (Roche #11096176001) and incubated with primary antibodies at 4 °C overnight. The following antibodies were used: p-p56K T389 (Cell Signaling #9234, 108D2, 1:1000), 4EBP1 (Cell Signaling #9452, 1:500), p-mTOR Ser2448 (Cell Signaling #5536, D9C2, 1:1000), total S6K (Cell Signaling #2708, 49D7, 1:1000), Grb2 (BD #610112, 81, 1:1000), p-eIF2a XP Ser51 (Cell Signaling #3398, D9G8, 1:1000), p-4EBP1 Thr37/46 (Cell Signaling #2855, 236B4, 1:500), Tubulin (Cell Signaling #3873, DM1A, 1:1000), ARG1 (Merck #ABS535, 1:000) and Actin (Sigma Aldrich #A2066, 1:1000). Incubated membranes were washed three times for 5 min with 0.5% Western Blocking Reagent and probed with the appropriate anti-IgG-horseradish peroxidase-linked (HRP) secondary antibody (GE Healthcare #NA934,

anti-rabbit IgG, 1:20,000; #NA931, anti-mouse IgG, 1:20,000; Promega #G135A, anti-chicken IgY, 1:20,000). Antigen-specific binding of antibodies was detected with SuperSignal West Femto and Pico Kits (Thermo Scientific #34095 and #34577).

Metabolism assays. OCR and extracellular acidification rate (ECAR) measurements were performed on a Seahorse XFe96 Analyzer (Agilent) using the Seahorse XF Cell Mito Stress test kit (Agilent #103015-100) according to the manufacturer's instructions. In brief, 150,000 preosteoclasts were seeded per well and treated with respective stimuli. During the measurements, oligomycin (1 µM), carbonyl cyanide-*p*-trifluoromethoxyphenylhydrazone (FCCP, 1 µM) and rotenone/antimycin A (500 nM) were subsequently injected. Raw data were analyzed using Wave Desktop Software (Agilent, version 2.6.1) and exported and graphed in GraphPad Prism 8 (GraphPad Software).

Flow cytometry. Plated preosteoclasts were harvested in Accutase solution (Sigma Aldrich #A6964) and later stained for DAPI (Biolegend #422801, 1:200), fixable viability dye (Invitrogen #65-0865-14, 1:2000), MitoTracker Green (Thermo Fisher #M7514, 100 nM) and F4/80-BV421 (BioLegend #123131, BM8, 1:80) according to the manufacturer's instructions. Bone marrow dendritic cells and macrophages were FC blocked with TruStain fcX (BioLegend #101320, 1:200) and surface stained for CD45.2-APC (TONBO #20-0454, 104, 1:80), CD11c-APC-eF780 (eBioscience #47-0114, N418, 1:50), CD11b-PeCy7 (eBioscience #25-0112, Clone M1/70, 1:400), MHC Class II-PE (TONBO #50-5321, MS/114.15.2, 1:80) and F4/80-BV421. For in vivo assessment of osteoclast precursors, washed splenocytes were stained for CD45.2-APC, F4/80-FITC (BioLegend #123108, BM8, 1:80), CD11b-PeCy7 and GR-1-PE (Ly-6G/Ly-6C, BioLegend #108408, RB6-8C5, 1:80). Cells were acquired and analysed using CytoFLEX S Flow Cytometry (Beckman Coulter), CytExpert (Version 2.0) and FlowJo (Version 10, LLC) software. MFI represents mean fluorescence intensity for Fig. 4e.

Bioinformatics analysis. All bioinformatics analyses and visualizations have been conducted using R version 3.5.1, ggplot2 3.1.0 and heatmaply 0.15.2 (ref. [48]). Network analysis relied on OmicsIntegrator 0.3.1 (ref. [49]) version 1.0 of the python implementation of Infomap[50] and Cytoscape 3.6.1 (ref. [51]). All the code and software parameters used for this study are provided through Jupyter notebooks on a Github repository.

Transcriptomics. Total RNA was prepared from 1×10^6 preosteoclasts using TRIzol Reagent (Thermo Fisher Scientific, #15596026). A total of 200 ng of total RNA was subsequently utilized for RNA-Seq library preparation by using TruSeq SR RNA sample prep kit (Illumina, #FC-122–1001), following the manufacturer's protocol. The libraries were sequenced for 50 cycles (single read) with a HiSeq 2000 (Illumina). Raw sequencing data were processed with CASAVA 1.8.2 to generate FastQ files. Sequence reads were mapped onto the mouse genome build mm9 using TopHat 2.0 (ref. [52]). From the RNA-Seq alignments files, read counts have been obtained for mm9 using Rsubread's featureCount[53], with and without allowing reads mapping multiple genes and counting overlapping reads for multiple genes. The results were similar using both methods. For instance, the Jaccard coefficient was equal to 0.837 when looking at genes differentially expressed between RANKL and RANKL/Arg-Depletion conditions with and without allowing for multi-mapping reads. Therefore, of the 469 genes found to be differentially expressed with or without read multimapping, 393 were common. The differential expression analysis was thus first based solely on counts obtained allowing multimapping reads (Figs. 2h, 3h, c, 4g and Supplementary Fig. 2). Genes were filtered to keep only those with $\log(CPM + 0.25) > 0$ in at least three samples. Library sizes were scaled, values were transformed and variance corrected in order to fit a linear model explaining the read abundances using the voom method[54] before identifying differentially expressed genes between conditions, with a log2 fold-change higher than 1 for a false discovery rate (FDR) of 5% using limma[55] and edgeR[56]. Hierarchical clustering with complete linkage of the z-score-transformed expression values for the 40 genes with highest difference between RANKL and RANKL/Arg-Depletion segregated all samples by experimental conditions. The gene sets identified were tested for enrichment using the clusterProfiler library[57] for the different levels of the Gene Ontology (GO)[58,59] and KEGG[60,61] in *Mus musculus*. Moreover, genes were filtered on membership to metabolism (mmu09100) or KEGG pathways of interest (mmu00010, mmu00020, mmu00220, mmu00330) using the REST API provided. Pairwise comparisons in expression levels between conditions were assessed using a permutation test, with the null hypothesis that there were no associations between the expression of any gene of the pathway and the conditions compared. Samples were ranked on their expression level for each gene in the pathway and the absolute difference between the sum of the ranks of the four replicates in each condition were summed over all genes, defining the test statistic. The association between samples and conditions was randomized 12,000 times and the same value computed, resulting in an empirical *P*-value by comparing the observed statistic to the simulated distribution. These values were reported after Bonferroni correction. The intersection of genes differentially expressed both with and without counting multimapping reads was then used to generate robust gene sets for the multi-omics integration analysis.

Proteomics. Cell pellets were lysed in 8 M urea, 10 mM HEPES (pH 8), 10 mM DTT and sonicated at 4 °C for 15 min (level 5, Bioruptor, Diagenode). Alkylation of reduced cysteines was performed in the dark for 30 min with 55 mM iodacetamide (IAA) followed by a two-step proteolytic digestion. Samples were digested at 21–24 °C with LysC (1:50, w/w, Wako) for 3 h. Cell lysates were adjusted to 2 M Urea with 50 mM ammoniumbicarbonate and then both cell lysates and supernatants were digested with trypsin (1:50, w/w, Promega) at 21–24 °C overnight. The resulting peptide mixtures were acidified and loaded on C18 StageTips (EmporeTM, IVA-Analysentechnik). Peptides were eluted with 80% acetonitrile (ACN), dried using a SpeedVac centrifuge, and resuspended in 2% ACN, 0.1% trifluoroacetic acid (TFA) and 0.5% acetic acid. Chemicals were purchased from Sigma-Aldrich unless stated otherwise. For ultra-high pressure LC-MS, peptides were separated on an EASY-nLC 1200 HPLC system (Thermo Fisher Scientific) coupled online to the Q Exactive HF-X mass spectrometer via a nanoelectrospray source (Thermo Fisher Scientific)[62]. Peptides were loaded in buffer A (0.5% formic acid) on in house packed columns (75-µm inner diameter, 50 cm length and 1.9 µm C18 particles from Dr. Maisch GmbH, Germany). Peptides were eluted with a nonlinear 170-min gradient of 5–60% buffer B (80% ACN, 0.5% formic acid) at a flow rate of 300 nl/min and a column temperature of 55 °C. The Q Exactive HF was operated in a data-dependent mode with a survey scan range of 300–1650 *m/z* and a resolution of 60,000–120,000 at *m/z* 200. Up to the ten most abundant isotope patterns with a charge > 1 were isolated with a 1.4 Thomson (Th) isolation window and subjected to higher-energy collisional dissociation (HCD) fragmentation at a normalized collision energy of 27. Fragmentation spectra were acquired with a resolution of 15,000 at *m/z* 200. Dynamic exclusion of sequenced peptides was set to 30 s to reduce repeated peptide sequencing. Thresholds for ion injection time and ion target values were set to 20 ms and 3E6 for the survey scans and 60 ms and 1E5 for the MS/MS scans, respectively. Data were acquired using the Xcalibur software (Thermo Scientific). MaxQuant software (version 1.5.3.2) was used to analyse MS raw files[63]. MS/MS spectra were searched against the human Uniprot FASTA database (Version July 2015, 91645 entries) and a common contaminants database (247 entries) by the Andromeda search engine[64]. Cysteine carbamidomethylation was applied as fixed and N-terminal acetylation, deamidation at NQ, and methionine oxidation as variable modifications. Enzyme specificity was set to trypsin with a maximum of 2 missed cleavages and a minimum peptide length of 7 AAs. An FDR of 1% was applied at the peptide and protein level. Peptide identification was performed with an allowed initial precursor mass deviation of up to 7 ppm and an allowed fragment mass deviation of 20 ppm. Nonlinear retention time alignment of all measured samples was performed in MaxQuant. Peptide identifications were matched across all samples within a time window of 1 min of the aligned retention times. A library for 'match between runs' in MaxQuant was built from duplicate and additional single shot MS runs from MACS enriched cell types. Protein identification required at least 1 'unique peptide' in MaxQuant. A minimum ratio count of 1 was required for valid quantification events via MaxQuant's Label Free Quantification algorithm (MaxLFQ). Downstream bioinformatic analysis was conducted in the Perseus environment[65]. Each quantified protein was required to be identified in three out of four replicates of at least one condition. Protein LFQ intensities were logarithmised to the base 2 and missing values imputed from a random normal distribution centred on the detection limit. Data were imported into Perseus, filtered to keep only proteins where the coefficient of variation of overall abundance was greater than the coefficients of variation in every condition, to select for proteins that consistently vary between conditions but not between replicates, and converted to *z*-scores. Average abundances among different conditions were compared by a two-sided Student's *t*-test (permutation-based FDR = 0.05, $s0 = 0.05$) and used to select proteins with differential abundance between conditions. As in the transcriptomic analysis, hierarchical clustering with complete linkage of the *z*-score-transformed data segregated the samples by experimental conditions.

Network integrative analysis. OmicsIntegrator was used to integrate transcriptomic and proteomic changes, by addressing the underlying molecular mechanisms of the effects of recArgI and the arginine-dependent effects of RANKL[49]. This tool circumvents the low overlap between proteomic and transcriptomic hits (Supplementary Fig. 6b), by inferring transcription factors likely to explain the transcriptional changes and searching for minimal sets of connected genes including these transcription factors and abundantly changed proteins. The first set of proteins and genes, describing the effect of recArgI, were defined by significantly changed abundance or expression respectively between RANKL/Arg-Rescue and RANKL/Arg-Depletion for proteomics data, and between RANKL and RANKL/Arg-Depletion for transcriptomics data. The second set, describing the arginine-dependent effect of RANKL, was defined as the exclusive disjunction of proteins or genes with changes in abundance or expression between RANKL-positive and RANKL-negative conditions in the presence and absence of arginine. By doing so, we selected for genes differentially expressed between M-CSF and RANKL or M-CSF/Arg-Rescue and RANKL/Arg-Rescue but not between M-CSF/Arg-Starvation and RANKL/Arg-Starvation, and proteins with different abundance between M-CSF/Arg-Rescue and RANKL/Arg-Rescue but not between M-CSF/Arg-Starvation and RANKL/Arg-Starvation. The protein interactions used were mouse–mouse interactions from BioGRID Release 3.4.160 (ref. [66]), weighted by directness and experimental strength of the assay (Supplementary Table 2). The

garnet tool was used to identify transcription factors explaining transcriptomic variability, converted to up-to-date murine identifiers and integrated with hit proteins via the Forest tool. The overlap between both resulting Steiner trees was contextualized by randomly selecting 2000 connected subgraphs of matched size on the BioGRID network and examining the corresponding shared nodes. Infomap algorithm and its python implementation[50] were used to cluster nodes into modules (two-level clustering) with a scaling factor of the link flows allowing a limited number of communities (--markov-time 2.5), which were visualized with the map generator[67]. Major network structures were identified according to coding theory, by looking at which grouping of the nodes allows the minimal coding of trajectories of random walks on the network (Supplementary Table 3). The information flow inside and between these groups therefore represents how tightly connected nodes are inside and between these groups. The resulting communities were analyzed for enrichment in GO Biological processes.

Metabolomics and $^{13}C_6$ arginine tracing. Cells were cultured for 24 h in the presence of fully labelled arginine ($^{13}C_6$ arginine, Sigma Aldrich). Medium was collected and 10 µl of ^{13}C-glycerol (150 µl/ml) was added as internal standard. Cells were scrapped, collected and frozen. Cell pellets were resuspended with 300 µl of cold methanol/water (8:1, v/v) containing ^{13}C-glycerol (5 µl/ml) as internal standard. Metabolites were extracted with three rounds of liquid N_2 immersion and sonication, followed by 1 h in ice before centrifugation at 22,800 × *g* (10 min at 4 °C). Medium was collected and 10 µl of ^{13}C-glycerol (150 µl/ml) was added as internal standard. The medium samples were lyophilized and resuspended in 500 µl of cold methanol/water (8:1, v/v). After vortexing, samples were left 1 h in ice and centrifuged 10 min at 22,800 × *g* (4 °C). Metabolite extraction of cells and medium was split into two aliquots of 250 µl (cells) or 400 µl (medium) for gas chromatography–MS (GC–MS), and 40 µl (both) for LC–MS analysis. For GC–MS analysis, samples were dried under a stream of N_2 gas and lyophilized before chemical derivatization with 40 µl of methoxyamine in pyridine (30 µg/ml) for 45 min at 60 °C. Samples were also silylated using 25 µl of *N*-methyl-*N*-trimethylsilyltrifluoroacetamide with 1% trimethylchlorosilane (Thermo Fisher Scientific) for 30 min at 60 °C to increase volatility of metabolites. A 7890A GC system coupled to a 7000 QqQ mass spectrometer (Agilent Technologies) was used for isotopologue determination. Derivatized samples were injected (1 µl) in the gas chromatograph system with a split inlet equipped with a J&W Scientific HP-5ms stationary phase column (30 m × 0.25 mm i.d., 0.1 µm film, Agilent Technologies). Helium was used as a carrier gas. The temperature gradient was the following: from 70 to 150 °C at a heating rate of 5 °C/min, from 150 to 250 °C at 10 °C/min and from 250 to 325 °C at 50 °C/min. Metabolites were ionized using positive chemical ionization (CI) with isobutene as reagent gas. Mass spectral data on the 7000 QqQ were acquired in scan mode monitoring selected ion clusters of the different metabolites (Supplementary Table 4). For LC–MS analysis, samples were analysed using an UHPLC system coupled to a 6490 QqQ mass spectrometer (Agilent Technologies) to determine arginine. Cells and medium extracts were injected (5 µl and 1 µl, respectively) and metabolites separated using an ACQUITY UPLC HSS T3 column (1.8 µm, 2.1 × 150 mm, Waters). The mobile phases used for the separation of the metabolites were A: water with 0.1% formic acid and B: ACN with 0.1% formic acid. The chromatographic gradient was isocratic for 2 min at 100% A, and from minute 2 to 3 decreased to 10% A. From minute 3 to 4, the percentage of A raised again to 100% and finally the column was equilibrated at 100% A until 8 min. Flow rate was 0.3 ml/min. The QqQ mass spectrometer worked in MRM mode using the transitions 181 → 74 (CE:20V) and 181 → 121 (CE:8V) to determine labelled Arg. The electrospray ionization source (ESI) worked in positive mode.

$^{13}C_6$ arginine and $^{13}C_4$ aspartate tracings. Cells were cultured for 24 h in the presence of fully labelled arginine ($^{13}C_6$ L-arginine, Cambridge Isotope Laboratories #CLM-2265) and fully labelled aspartate ($^{13}C_4$ L-aspartic acid, Cambridge Isotope Laboratories #CLM-1801). Cell extracts where centrifuged for 10 min at 5000 × *g*. The supernatant was collected and dried using nitrogen evaporator. The samples were reconstituted in 50 µl of methanol, centrifuged for 10 min at 1000 × *g* and supernatant was used for LC–MS analysis. For AA tracing, a Vanquish UHPLC system (Thermo Scientific) coupled to an Orbitrap Fusion Lumos (Thermo Scientific) mass spectrometer was used for the LC–MS analysis. The chromatographic separation for samples was carried out on an ACQUITY UPLC BEH Amide, 1.7 µm, 2.1 × 100 mm analytical column (Waters) equipped with a VanGuard: BEH Amide, 2.1 × 5 mm pre-column (Waters). The column was maintained at a temperature of 40 °C and 2 µl sample were injected per run. The mobile phase A was 0.15% formic acid (v/v) in water and mobile phase B was 0.15% formic acid (v/v) in 85% ACN (v/v) with 10 mM ammonium formate. The gradient elution with a flow rate 0.4 ml/min was performed with a total analysis time of 17 min. The mass spectrometer was operated in a positive electrospray ionization mode: spray voltage 3.5 kV; sheath gas flow rate 60 arb; auxiliary gas flow rate 20 arb; capillary temperature 285 °C. For the analysis a full MS scan mode with a scan range *m/z* 50–250, resolution 500,000; AGC target 2e5 and a maximum injection time 50 ms was applied. The data processing was performed with the TraceFinder 4.1 software (Thermo Scientific). For TCA cycle tracing, a Vanquish UHPLC system (Thermo Scientific) coupled to an Orbitrap Fusion Lumos (Thermo Scientific) mass spectrometer was used for the LC–MS analysis. The chromatographic separation for samples was carried out on an ACQUITY HSS T3, 1.8 µm, 2.1 × 100 mm analytical column (Waters) equipped with a VanGuard HSS T3, 2.1 × 5 mm pre-column (Waters). The column was maintained at a temperature of 40 °C and 2 µl of sample was injected per run. The mobile phase A was 0.1% formic acid (v/v) in water and mobile phase B was 0.1% formic acid (v/v) in methanol. The gradient elution with a flow rate 0.5 ml/min was performed with a total analysis time of 10 min. The mass spectrometer was operated both in positive and negative electrospray ionization mode: spray voltage for positive mode 3.5 kV and 3.0 kV for negative mode; sheath gas flow rate 60 arb; auxiliary gas flow rate 20 arb; capillary temperature 285 °C. For the analysis a full MS scan mode with a scan range *m/z* 80–400, resolution 500,000; AGC target 2e5 and a maximum injection time 50 ms was applied. The data processing was performed with the TraceFinder 4.1 software (Thermo Scientific).

Serum metabolomics. Serum concentrations of arginine were assessed using an arginine ELISA Kit (#K 207733, IDK). For serum samples, 10 µl of sample was placed on a 96-well hydrophobic filter plate and mixed with 10 µl of an isotopically labelled internal standard mixture. A total of 300 µl of methanol was added and the plate was shaken for 20 min at 450 rpm. Afterwards, the sample extract was collected in a 96-well plate by centrifuging the filter plate for 5 min at 500 × *g*. The sample extracts were used for LC–MS analysis. A Vanquish UHPLC system (Thermo Scientific) coupled with an Orbitrap Q Exactive (Thermo Scientific) mass spectrometer was used for the LC–MS analysis. The chromatographic separation for samples was carried out on an ACQUITY UPLC BEH Amide, 1.7 µm, 2.1 × 100 mm analytical column (Waters) equipped with a VanGuard: BEH C18, 2.1 × 5 mm pre-column (Waters). The column was maintained at a temperature of 40 °C and the sample injection volume was 2 µl. The mobile phase A was 0.15% formic acid (v/v) in water and mobile phase B was 0.15% formic acid (v/v) in 85% ACN (v/v) with 10 mM ammonium formate. The gradient elution with a flow rate 0.4 ml/min was performed with a total analysis time of 17 min. The Orbitrap Q Exactive (Thermo Scientific) mass spectrometer was operated in an electrospray ionization positive mode, spray voltage 3.5 kV, aux gas heater temperature 400 °C, capillary temperature 350 °C, aux gas flow rate 12. The metabolites of interest were analysed using a full MS scan mode, scan range *m/z* 50–400, resolution 35,000, AGC target 1e6, maximum IT 50 ms. The Trace Finder 4.1 software (Thermo Scientific) was used for the data processing. Seven-point linear calibration curves with internal standardization and 1/*x* weighing were constructed for the quantification of metabolites.

Image processing. Images were cropped and processed using Adobe Photoshop CS6, adjusting brightness and contrast. Original uncropped Western Blot images are provided in the source data file.

Statistics. Statistical analysis was performed using a two-tailed *t*-test for two groups, an ordinary one-way ANOVA followed by Bonferroni's multiple comparisons test or Tukey post-hoc test for multiple groups and a two-way ANOVA followed by Bonferroni's multiple comparisons test for curve analysis, unless otherwise stated. Statistical outliers for patient data in Fig. 1 have been excluded based on alpha = 0.05 on Prism 8 software (GraphPad, La Jolla, CA). In vitro data are representative of at least two repeats, while transcriptomics, proteomics and metabolomics correspond to biological replicates. Statistical significance is indicated by *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.001. All error bars indicate ± SEM.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

RNA-Seq data have been deposited in Gene Expression Omnibus (GEO) under accession number GSE125101. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[68] partner repository with the dataset identifier PXD012405. The source data underlying Figs. 1–2 and 4–7 and Supplementary Figs. 1–4 and 7 are provided as a Source Data file or available from the corresponding authors upon reasonable request.

Code availability

All code used in the proteomics and transcriptomics differential analyses and subsequent integration is available on GitHub (https://doi.org/10.5281/zenodo.2541644).

Received: 22 March 2019; Accepted: 16 December 2019;
Published online: 22 January 2020

References

1. McInnes, I. B. & Schett, G. The pathogenesis of rheumatoid arthritis. *N. Engl. J. Med.* 365, 2205–2219 (2011).

2. Kong, Y. Y. et al. OPGL is a key regulator of osteoclastogenesis, lymphocyte development and lymph-node organogenesis. *Nature* 397, 315–323 (1999).
3. Lavin, Y. et al. Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. *Cell* 159, 1312–1326 (2014).
4. O'Neill, L. A. & Pearce, E. J. Immunometabolism governs dendritic cell and macrophage function. *J. Exp. Med.* 213, 15–23 (2016).
5. Geiger, R. et al. L-arginine modulates T cell metabolism and enhances survival and anti-tumor activity. *Cell* 167, 829–842.e813 (2016).
6. Murray, P. J. Amino acid autotrophy as a system of immunological control nodes. *Nat. Immunol.* 17, 132 (2016).
7. Caputa, G., Castoldi, A. & Pearce, E. J. Metabolic adaptations of tissue-resident immune cells. *Nat. Immunol.* 20, 793–801 (2019).
8. Vincent, E. E. et al. Mitochondrial phosphoenolpyruvate carboxykinase regulates metabolic adaptation and enables glucose-independent tumor growth. *Mol. Cell* 60, 195–207 (2015).
9. Ali, U. et al. L-asparaginase as a critical component to combat Acute Lymphoblastic Leukaemia (ALL): a novel approach to target ALL. *Eur. J. Pharmacol.* 771, 199–210 (2016).
10. Yau, T. et al. Preliminary efficacy, safety, pharmacokinetics, pharmacodynamics and quality of life study of pegylated recombinant human arginase 1 in patients with advanced hepatocellular carcinoma. *Invest. New Drugs* 33, 496–504 (2015).
11. Qiu, F., Huang, J. & Sui, M. Targeting arginine metabolism pathway to treat arginine-dependent cancers. *Cancer Lett.* 364, 1–7 (2015).
12. Ditzel, H. J. The K/BxN mouse: a model of human inflammatory arthritis. *Trends Mol. Med.* 10, 40–45 (2004).
13. Keffer, J. et al. Transgenic mice expressing human tumour necrosis factor: a predictive genetic model of arthritis. *EMBO J.* 10, 4025–4031 (1991).
14. Courtenay, J. S., Dallman, M. J., Dayan, A. D., Martin, A. & Mosedale, B. Immunisation against heterologous type II collagen induces arthritis in mice. *Nature* 283, 666–668 (1980).
15. Morris, S. M. Jr. Arginine metabolism: boundaries of our knowledge. *J. Nutr.* 137, 1602s–1609s (2007).
16. Matsuyama, K. et al. The transcription factor Jdp2 controls bone homeostasis and antibacterial immunity by regulating osteoclast and neutrophil differentiation. *Immunity* 37, 1034–1046 (2012).
17. El Kasmi, K. C. et al. Toll-like receptor-induced arginase 1 in macrophages thwarts effective immunity against intracellular pathogens. *Nat. Immunol.* 9, 1399–1406 (2008).
18. Wang, S. et al. The amino acid transporter SLC38A9 is a key component of a lysosomal membrane complex that signals arginine sufficiency to mTORC1. *Science* 347, 188–194 (2015).
19. Weichhart, T., Hengstschläger, M. & Linke, M. Regulation of innate immune cell function by mTOR. *Nat. Rev. Immunol.* 15, 599–614 (2015).
20. Rebsamen, M. et al. SLC38A9 is a component of the lysosomal amino acid sensing machinery that controls mTORC1. *Nature* 519, 477–481 (2015).
21. Pakos-Zebrucka, K. et al. The integrated stress response. *EMBO Rep.* 17, 1374–1395 (2016).
22. Zeng, R., Faccio, R. & Novack, D. V. Alternative NF-kappaB regulates RANKL-induced osteoclast differentiation and mitochondrial biogenesis via independent mechanisms. *J. Bone Miner. Res.* 30, 2287–2299 (2015).
23. Ishii, K. A. et al. Coordination of PGC-1beta and iron uptake in mitochondrial biogenesis and osteoclast activation. *Nat. Med.* 15, 259–266 (2009).
24. 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).
25. Shambaugh, G. E. 3rd Urea biosynthesis I. The urea cycle and relationships to the citric acid cycle. *Am. J. Clin. Nutr.* 30, 2083–2087 (1977).
26. Broer, S. & Broer, A. Amino acid homeostasis and signalling in mammalian cells and organisms. *Biochem. J.* 474, 1935–1963 (2017).
27. Qualls, J. E. et al. Sustained generation of nitric oxide and control of mycobacterial infection requires argininosuccinate synthase 1. *Cell Host Microbe* 12, 313–323 (2012).
28. Yeon, J. T. et al. Natural polyamines inhibit the migration of preosteoclasts by attenuating Ca2+-PYK2-Src-NFATc1 signaling pathways. *Amino Acids* 46, 2605–2614 (2014).
29. Metcalf, B. W. et al. Catalytic irreversible inhibition of mammalian ornithine decarboxylase (E.C.4.1.1.17) by substrate and product analogs. *J. Am. Chem. Soc.* 100, 2551–2553 (1978).
30. Fletcher, M. et al. l-Arginine depletion blunts antitumor T-cell responses by inducing myeloid-derived suppressor cells. *Cancer Res.* 75, 275–283 (2014).
31. Chen, C. et al. TSC-mTOR maintains quiescence and function of hematopoietic stem cells by repressing mitochondrial biogenesis and reactive oxygen species. *J. Exp. Med.* 205, 2397–2408 (2008).
32. Buck, M. D., Sowell, R. T., Kaech, S. M. & Pearce, E. L. Metabolic Instruction of Immunity. *Cell* 169, 570–586 (2017).
33. Manoli, I. & Venditti, C. P. Disorders of branched chain amino acid metabolism. *Transl. Sci. Rare Dis.* 1, 91–110 (2016).
34. Milde, R. et al. Multinucleated giant cells are specialized for complement-mediated phagocytosis and large target destruction. *Cell Rep.* 13, 1937–1948 (2015).
35. Pereira, M. et al. Common signalling pathways in macrophage and osteoclast multinucleation. *J. Cell Sci.* 131, jcs216267 (2018).
36. Yeon, J. T., Choi, S. W. & Kim, S. H. Arginase 1 is a negative regulator of osteoclast differentiation. *Amino Acids* 48, 559–565 (2016).
37. Cejka, D. et al. Mammalian target of rapamycin signaling is crucial for joint destruction in experimental arthritis and is activated in osteoclasts from patients with rheumatoid arthritis. *Arthritis Rheumatol.* 62, 2294–2302 (2010).
38. Huynh, H. & Wan, Y. mTORC1 impedes osteoclast differentiation via calcineurin and NFATc1. *Commun. Biol.* 1, 29 (2018).
39. Salahudeen, A. A. et al. An E3 ligase possessing an iron-responsive hemerythrin domain is a regulator of iron homeostasis. *Science* 326, 722–726 (2009).
40. Lendahl, U., Lee, K. L., Yang, H. & Poellinger, L. Generating specificity and diversity in the transcriptional response to hypoxia. *Nat. Rev. Genet.* 10, 821–832 (2009).
41. Haslacher, H et al. Usage Data and Scientific Impact of the Prospectively Established Fluid Bioresources at the Hospital-Based MedUni Wien Biobank. *Biopreservation and biobanking* 16, 477–482 (2018).
42. Bluml, S. et al. Essential role of microRNA-155 in the pathogenesis of autoimmune arthritis in mice. *Arthritis Rheumatol.* 63, 1281–1288 (2011).
43. Bluml, S. et al. Loss of phosphatase and tensin homolog (PTEN) in myeloid cells controls inflammatory bone destruction by regulating the osteoclastogenic potential of myeloid cells. *Ann. Rheum. Dis.* 74, 227–233 (2015).
44. Van de Velde, L. A. et al. Stress kinase GCN2 controls the proliferative fitness and trafficking of cytotoxic T cells independent of environmental amino acid sensing. *Cell Rep.* 17, 2247–2258 (2016).
45. Copp, J., Manning, G. & Hunter, T. TORC-specific phosphorylation of mTOR: phospho-Ser2481 is a marker for intact mTORC2. *Cancer Res.* 69, 1821–1827 (2009).
46. Linke, M. et al. Chronic signaling via the metabolic checkpoint kinase mTORC1 induces macrophage granuloma formation and marks sarcoidosis progression. *Nat. Immunol.* 18, 293–302 (2017).
47. Carpenter, A. E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* 7, R100 (2006).
48. Galili, T., O'Callaghan, A., Sidi, J. & Sievert, C. heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics* 34, 1600–1602 (2018).
49. Tuncbag, N. et al. Network-based interpretation of diverse high-throughput datasets through the omics integrator software package. *PLoS Comput. Biol.* 12, e1004879 (2016).
50. Rosvall, M. & Bergstrom, C. T. Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems. *PLoS ONE* 6, e18209 (2011).
51. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504 (2003).
52. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36 (2013).
53. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* 41, e108 (2013).
54. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 15, R29 (2014).
55. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47 (2015).
56. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140 (2010).
57. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287 (2012).
58. Ashburner, M. et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29 (2000).
59. Consortium, T. G. O. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res.* 45, D331–D338 (2017).
60. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30 (2000).
61. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–D361 (2017).
62. Kelstrup, C. D. et al. Performance evaluation of the Q exactive HF-X for shotgun proteomics. *J. Proteome Res.* 17, 727–738 (2018).

63. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372 (2008).
64. Cox, J. et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* 10, 1794–1805 (2011).
65. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* 13, 731–740 (2016).
66. Chatr-Aryamontri, A. et al. The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* 45, D369–D379 (2017).
67. Bohlin L., Edler D., Lancichinetti A., Rosvall M. Community Detection and Visualization of Networks with the Map Equation Framework. In: Ding Y., Rousseau R., Wolfram D. (eds) *Measuring Scholarly Impact*. Springer, Cham 3–34 (2014)
68. Vizcaíno, J. A. et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* 44, 11033 (2016).

Acknowledgements

We thank Tetyana Shvets, Hannah Paar, Lucia Quesada Garrido, Mario Kuttke, Maximilian Kugler and Manuel Salzmann for technical assistance. We thank Giulio Superti-Furga for the *Slc35a9*$^{-/-}$ mice and George Kollias for providing the hTNF^{tg} mice. This research was funded by the Austrian Society for Rheumatology (ÖGR) (J.S.B), the FWF (G.S.: 30026, 31106; O.S.: 31568) and the Christian Doppler laboratory for Arginine Metabolism in Rheumatoid Arthritis and Multiple Sclerosis (G.S.). J.S.B., A.V. and A.L. were rewarded a DOC fellowship by the Austrian Academy of Sciences. A.B. received funding from the European Research Council (ERC) under the European Union's Seventh Framework Program and Horizon 2020 research and innovation program (grant agreement No. 677006, CMIL). M.R., J.B.B. and P.J.M. received funding from the Max Planck Gesellschaft.

Author contributions

J.S.B., P.J.M., O.S., S.B. and G.S. conceived and designed the study. J.S.B., M.H., M.Ki., A.L., A.V., M.Ke., V.S., B.N., A.J., A.F., C.S., Y.M., K.K., M.R., J.B.B. and O.Y. performed the experiments. L.V. analyzed transcriptomic and proteomic data and performed integrated network analysis. Y.M. performed RNA sequencing. A.F. performed proteomics. K.K., O.Y. and A.J. performed metabolomic analysis. J.S.B., L.V., A.L. and A.J. analyzed the data. G.K., A.B., J.J.O., T.W., F.M., J.S.S., P.C., J.M., O.Y. and P.J.M. provided key resources. J.S.B., L.V., O.S., S.B. and G.S. wrote the manuscript. All authors read, revised and approved the final manuscript.

Competing interests

The authors declare the following competing interests: P.C. is the founder of Biocancer Treatment International Ltd. P.C., G.S. and S.B. are listed as inventors on a patent (US9789169B2) covering recArg1/BCT-100. The remaining authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-14285-1.

Correspondence and requests for materials should be addressed to S.B. or G.S.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

5 Discussion

Nearly two decades ago, knowledge gained from the human genome sequencing project sparked great hopes and expectations for identifying disease-causing genetic elements. However, environmental alterations have a huge impact on post-transcriptional and post-translational regulation of the genome and lead to distinct phenotypes. Today, proteomics using mass spectrometry has been recognized as the most successful technique in proteomic research. Recent research has demonstrated that alterations in the proteome correlate strongly with disease symptoms and, in many cases, surpass genomic and transcriptomic-based techniques. Thus, proteomics is a potent method for comprehensively evaluating disorders by enabling the analysis of protein expression, post-translational modification and molecular interactions. Instrumentation and MS-based methods are dynamically evolving and allow researchers to continuously push new boundaries. More than 1000 proteins can be analyzed from individual cells due to advances in sample processing, LC performance and MS instrumentation. Advances in liquid chromatography allow the robust in-depth analysis of thousands of samples on a single chromatographic column without loss of performance.

5.1 Interaction proteomics

The majority of the proteome data produced in this thesis were obtained by an elaborate AE-LC-MS/MS study and provide significant information regarding the physical interactions and post-translational modifications in human immune cells.

I described MIP-APMS, a method for interrogating functional signal transduction networks in intracellular signaling pathways that combines simplified cell line generation and proteomics. We quantified over 370 PPIs and 80 PTMs in human monocytes following receptor activation or drug treatment. Our approach identified molecular links between PTMs and PPIs, as well as protein subnetworks that regulate cellular programs regulated by site-specific PTMs. Site-specific modification of amino acids to be phosphorylated or isgylated on TRAF2 resulted in divergent interactomes and altered cellular physiology for mutant proteins. Thus, structural insights into the interfaces between protein complexes and critical PTMs for protein complex stabilization can be elucidated. MIP-APMS experiments with temporal resolution enable the clarification of co-regulations at several biochemical layers, hence increasing our understanding of the molecular linkages that exist between the successive processes of signal transduction. By enhancing temporal resolution further, it may become able to unravel the causal relationships between PTM and PPIS regulation in greater detail.

Historically, AE- and AP-LC-MS/MS investigations have required ectopic expression and affinity tags fused to the protein of interest in order to enrich and identify interacting proteins, as well as unravel dynamic processes in protein-complex formation. However, studies using AP-LC-MS/MS frequently generate a large number of false-positive results, which frequently results in misinterpretation of the acquired interaction data. Fusion of an affinity tag to the protein of interest can increase false-negative and false-positive rates by interfering with the structure and function of the protein, mis-localizing the protein within the cell, destabilizing real interactions, and identifying proteins with a high affinity for the molecular tag utilized. Additionally, ectopic expression of the target protein can have a profound effect on the cellular proteome. Thus, validation and optimization of AP-LC-MS/MS procedures and statistical analysis are essential to ensure the highest possible quality of PPI data. In the future, CRISPR-Cas9-mediated genomic editing tools, like prime editing, may be used to directly introduce the affinity label into the genome to forego ectopic gene-expression with all it's disadvantages. Alternatively, antibody-based capture techniques (CO-IP-MS/MS) and proximity-dependent approaches have been used in the past. The CO-IP-MS/MS technique enables the identification of protein interactions in primary cells under physiological circumstances. However, the antibody's availability, affinity, and specificity must be carefully examined. Additionally, each antibody's binding conditions require time-consuming optimization, limiting its application to less extensive PPI research.

Often, cell lysis and buffer conditions have a major impact on the sensitivity and specificity of AP- and CO-IP-MS/MS procedures, frequently resulting in the loss of transient interactions. We optimized our MIP-APMS workflow to also capture these transient interactions. Alternatively, proximity-dependent techniques, such as BioID or APEX, rely on enzymes being fused to the target protein and covalently modify proteins in close proximity. This permits the detection of transient protein-protein interactions, however due to the broad labelling radius also non-direct interacting proteins are identified. A different approach is size-exclusion chromatography coupled to mass spectrometry (SEC-MS), which derives PPI information from co-eluting protein profiles without ectopic tagging of the protein. Thousands of protein interactions can be investigated in a single experiment. However, it is limited in resolution by the separation efficiency of the analytical size-exclusion chromatography.

Alternatively, by incorporating chemical crosslinking techniques [328, 329], transient interactors could be further stabilized. Combining APMS with structural information about the bait protein and its interactome may identify different interfaces between protein complexes that are altered by site-specific covalent modification or pharmacological action. Integrating protein crosslinking with PTM status, e.g. of wild-type

vs. mutant variant proteins of interest, would be particularly fascinating since it would allow for the resolution of altered structural interaction surfaces in addition to differential PTM-dependent PPIs.

Currently, high amounts of measurement time are needed for the conduction of an APMS experiment. Using data-independent collection methodologies and short LC gradients, the overall measuring time per sample at a comparable proteome depth may be further lowered in the future [57, 330]. The reproducibility, precision, and accuracy of quantification of changed peptides can be further enhanced by adopting isobaric labeling techniques rather than LFQ [111, 331].

5.2 Post-translational modifications

While mass spectrometry-based proteomics has significantly advanced our understanding of specific PTMs like phosphorylation and ubiquitinylation, this trend has been mostly limited to modifications with known compositions. Due to the fact that conventional bottom-up MS relies on database searches to identify peptides and modified peptides, it is fundamentally incapable of detecting modifications with an unknown mass. I used an open search mode in my thesis, utilizing a variety of PTM search algorithms – MaxQuant, MS-fragger, and PEAKS-based taggraph – to enable the unbiased characterization of PTMs on specific target proteins in AP-MS experiments without prior enrichment. De novo sequencing speed will become even more critical when high-resolution and high-volume mass spectrometers become more commonly available. Additionally, enrichment approaches remain the ideal alternative for performing extremely detailed analyses of selected PTMs, although we anticipate that this requirement will significantly decrease as instrumentation sensitivity improves. We demonstrate in Project 1 that, given high quality MS/MS spectra, open search enables the measurement of otherwise un-enrichable, but abundant PTMs.

Because PTMs on proteins are typically substoichiometric, their detection requires particular enrichment to aid in MS identification. As a result, analysis of a large number of additional extremely interesting PTMs has lagged much behind. PTMs can often be identified without enrichment, albeit not to the same depth, due to the higher scanning speed and dynamic range of contemporary MS apparatus. We have demonstrated this in Project 2, where we studied proteolysis by signal peptide peptidases giving rise to new N-termini of proteins with signal peptides in human and mice. This post-translational protein modification is so abundant, that it can be directly studied without further enrichment. We envision, that similar studies will be carried out in the future to further enlarge the landscape of experimentally verified signal peptide cleavage.

Post-translational modification is the primary source of protein variety. To obtain a better understanding of our organism's physiological/pathophysiological processes and to develop higher-quality and more efficient biopharmaceutical products and diagnostic techniques, additional knowledge in the fields of proteomics and PTM mapping is required. Proteomic techniques based on mass spectrometry have emerged as a powerful tool for screening and characterizing PTMs. Although current technology is unable of providing a comprehensive picture of the changed proteome, future proteomics will build on continuing advancements in mass spectrometry-based proteomic methods.

6 References

1. Consortium, E.P., *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. **489**(7414): p. 57-74.
2. Li, X. and J.L. Manley, *Cotranscriptional processes and their influence on genome stability.* Genes & Development, 2006. **20**(14): p. 1838-1847.
3. Hershey, J.W.B., N. Sonenberg, and M.B. Mathews, *Principles of translational control: an overview.* Cold Spring Harbor perspectives in biology, 2012. **4**(12): p. a011528.
4. Aebersold, R., et al., *How many human proteoforms are there?* Nature chemical biology, 2018. **14**(3): p. 206-214.
5. Ponomarenko, E.A., et al., *The Size of the Human Proteome: The Width and Depth.* Int J Anal Chem, 2016. **2016**: p. 7436849.
6. Wang, D., et al., *A deep proteome and transcriptome abundance atlas of 29 healthy human tissues.* Molecular Systems Biology, 2019. **15**(2): p. e8503.
7. Adhikari, S., et al., *A high-stringency blueprint of the human proteome.* Nature Communications, 2020. **11**(1): p. 5301.
8. Link, A.J., et al., *Direct analysis of protein complexes using mass spectrometry.* Nature Biotechnology, 1999. **17**(7): p. 676-682.
9. Wolters, D.A., M.P. Washburn, and J.R. Yates, *An Automated Multidimensional Protein Identification Technology for Shotgun Proteomics.* Analytical Chemistry, 2001. **73**(23): p. 5683-5690.
10. Zhang, Y., et al., *Protein analysis by shotgun/bottom-up proteomics.* Chemical reviews, 2013. **113**(4): p. 2343-2394.
11. Olsen, J.V., S.-E. Ong, and M. Mann, *Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues*.* Molecular & Cellular Proteomics, 2004. **3**(6): p. 608-614.
12. Vistain, L.F. and S. Tay, *Single-Cell Proteomics.* Trends in Biochemical Sciences, 2021.
13. Woods, A.G., et al., *Mass Spectrometry for Proteomics-Based Investigation*, in *Advancements of Mass Spectrometry in Biomedical Research*, A.G. Woods and C.C. Darie, Editors. 2019, Springer International Publishing: Cham. p. 1-26.
14. Lermyte, F., et al., *Top or Middle? Up or Down? Toward a Standard Lexicon for Protein Top-Down and Allied Mass Spectrometry Approaches.* Journal of the American Society for Mass Spectrometry, 2019. **30**(7): p. 1149-1157.
15. Leney, A.C. and A.J.R. Heck, *Native Mass Spectrometry: What is in the Name?* Journal of the American Society for Mass Spectrometry, 2017. **28**(1): p. 5-13.
16. Zhang, J., et al., *Top-down quantitative proteomics identified phosphorylation of cardiac troponin I as a candidate biomarker for chronic heart failure.* Journal of proteome research, 2011. **10**(9): p. 4054-4065.
17. Ge, Y., et al., *Top-down high-resolution mass spectrometry of cardiac myosin binding protein C revealed that truncation alters protein phosphorylation state.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(31): p. 12658-12663.
18. Ntai, I., et al., *Precise characterization of KRAS4b proteoforms in human colorectal cells and tumors reveals mutation/modification cross-talk.* Proceedings of the National Academy of Sciences of the United States of America, 2018. **115**(16): p. 4140-4145.
19. McCool, E.N., et al., *Deep Top-Down Proteomics Using Capillary Zone Electrophoresis-Tandem Mass Spectrometry: Identification of 5700 Proteoforms from the Escherichia coli Proteome.* Analytical chemistry, 2018. **90**(9): p. 5529-5533.
20. Brown, K.A., et al., *Top-down proteomics: challenges, innovations, and applications in basic and clinical research.* Expert Review of Proteomics, 2020. **17**(10): p. 719-733.

21. Zhang, X., *Less is More: Membrane Protein Digestion Beyond Urea-Trypsin Solution for Next-level Proteomics.* Molecular & cellular proteomics : MCP, 2015. **14**(9): p. 2441-2453.
22. Baez, N.O.D., J.A. Reisz, and C.M. Furdui, *Mass spectrometry in studies of protein thiol chemistry and signaling: opportunities and caveats.* Free radical biology & medicine, 2015. **80**: p. 191-211.
23. Cleland, W.W., *Dithiothreitol, a New Protective Reagent for SH Groups*.* Biochemistry, 1964. **3**(4): p. 480-482.
24. Getz, E.B., et al., *A Comparison between the Sulfhydryl Reductants Tris(2-carboxyethyl)phosphine and Dithiothreitol for Use in Protein Biochemistry.* Analytical Biochemistry, 1999. **273**(1): p. 73-80.
25. Tsiatsiani, L. and A.J.R. Heck, *Proteomics beyond trypsin.* The FEBS Journal, 2015. **282**(14): p. 2612-2626.
26. Rappsilber, J., Y. Ishihama, and M. Mann, *Stop and Go Extraction Tips for Matrix-Assisted Laser Desorption/Ionization, Nanoelectrospray, and LC/MS Sample Pretreatment in Proteomics.* Analytical Chemistry, 2003. **75**(3): p. 663-670.
27. Kulak, N.A., et al., *Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells.* Nature Methods, 2014. **11**(3): p. 319-324.
28. Hughes, C.S., et al., *Ultrasensitive proteome analysis using paramagnetic bead technology.* Molecular Systems Biology, 2014. **10**(10): p. 757.
29. Berger, S.T., et al., *MStern Blotting-High Throughput Polyvinylidene Fluoride (PVDF) Membrane-Based Proteomic Sample Preparation for 96-Well Plates.* Molecular & cellular proteomics : MCP, 2015. **14**(10): p. 2814-2823.
30. Dai, J., et al., *Protein Phosphorylation and Expression Profiling by Yin-Yang Multidimensional Liquid Chromatography (Yin-Yang MDLC) Mass Spectrometry.* Journal of Proteome Research, 2007. **6**(1): p. 250-262.
31. Motoyama, A., et al., *Anion and Cation Mixed-Bed Ion Exchange for Enhanced Multidimensional Separations of Peptides and Phosphopeptides.* Analytical Chemistry, 2007. **79**(10): p. 3623-3634.
32. Zhou, H., et al., *Toward a Comprehensive Characterization of a Human Cancer Cell Phosphoproteome.* Journal of Proteome Research, 2013. **12**(1): p. 260-271.
33. Kulak, N.A., P.E. Geyer, and M. Mann, *Loss-less Nano-fractionator for High Sensitivity, High Coverage Proteomics.* Molecular & Cellular Proteomics : MCP, 2017. **16**(4): p. 694-705.
34. Solovyeva, E.M., et al., *FractionOptimizer: a method for optimal peptide fractionation in bottom-up proteomics.* Analytical and Bioanalytical Chemistry, 2018. **410**(16): p. 3827-3833.
35. Aguilar, M.-I., *Reversed-Phase High-Performance Liquid Chromatography*, in *HPLC of Peptides and Proteins: Methods and Protocols*, M.-I. Aguilar, Editor. 2004, Springer New York: Totowa, NJ. p. 9-22.
36. Köcher, T., R. Swart, and K. Mechtler, *Ultra-High-Pressure RPLC Hyphenated to an LTQ-Orbitrap Velos Reveals a Linear Relation between Peak Capacity and Number of Identified Peptides.* Analytical Chemistry, 2011. **83**(7): p. 2699-2704.
37. Shen, Y., et al., *Packed Capillary Reversed-Phase Liquid Chromatography with High-Performance Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for Proteomics.* Analytical Chemistry, 2001. **73**(8): p. 1766-1775.
38. Wilson, I.D., et al., *High Resolution "Ultra Performance" Liquid Chromatography Coupled to oa-TOF Mass Spectrometry as a Tool for Differential Metabolic Pathway Profiling in Functional Genomic Studies.* Journal of Proteome Research, 2005. **4**(2): p. 591-598.
39. Thakur, S.S., et al., *Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation.* Molecular & cellular proteomics : MCP, 2011. **10**(8): p. M110.003699-M110.003699.

40. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64.
41. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.
42. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons.* Analytical Chemistry, 1988. **60**(20): p. 2299-2301.
43. Haag, A.M., *Mass Analyzers and Mass Spectrometers*, in *Modern Proteomics – Sample Preparation, Analysis and Practical Applications*, H. Mirzaei and M. Carrasco, Editors. 2016, Springer International Publishing: Cham. p. 157-169.
44. Miller, P.E. and M.B. Denton, *The quadrupole mass filter: Basic operating concepts.* Journal of Chemical Education, 1986. **63**(7): p. 617.
45. Stafford, G.C., et al., *Recent improvements in and analytical applications of advanced ion trap technology.* International Journal of Mass Spectrometry and Ion Processes, 1984. **60**(1): p. 85-98.
46. Mamyrin, B.A., *Time-of-flight mass spectrometry (concepts, achievements, and prospects).* International Journal of Mass Spectrometry, 2001. **206**(3): p. 251-266.
47. Comisarow, M.B. and A.G. Marshall, *Fourier transform ion cyclotron resonance spectroscopy.* Chemical Physics Letters, 1974. **25**(2): p. 282-283.
48. Comisarow, M.B. and A.G. Marshall, *Frequency-sweep fourier transform ion cyclotron resonance spectroscopy.* Chemical Physics Letters, 1974. **26**(4): p. 489-490.
49. Hu, Q., et al., *The Orbitrap: a new mass spectrometer.* Journal of Mass Spectrometry, 2005. **40**(4): p. 430-443.
50. Scigelova, M., et al., *Fourier Transform Mass Spectrometry.* Molecular & Cellular Proteomics, 2011. **10**(7): p. M111.009431.
51. Mann, M. and N.L. Kelleher, *Precision proteomics: The case for high resolution and high mass accuracy.* Proceedings of the National Academy of Sciences, 2008. **105**(47): p. 18132.
52. Roepstorff, P. and J. Fohlman, *Letter to the editors.* Biomedical Mass Spectrometry, 1984. **11**(11): p. 601-601.
53. Röst, H.L., et al., *OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data.* Nature Biotechnology, 2014. **32**(3): p. 219-223.
54. Kelstrup, C.D., et al., *Performance Evaluation of the Q Exactive HF-X for Shotgun Proteomics.* Journal of Proteome Research, 2018. **17**(1): p. 727-738.
55. Chapman, J.D., D.R. Goodlett, and C.D. Masselon, *Multiplexed and data-independent tandem mass spectrometry for global proteome profiling.* Mass Spectrometry Reviews, 2014. **33**(6): p. 452-470.
56. Ludwig, C., et al., *Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial.* Molecular systems biology, 2018. **14**(8): p. e8126-e8126.
57. Bruderer, R., et al., *Optimization of Experimental Parameters in Data-Independent Mass Spectrometry Significantly Increases Depth and Reproducibility of Results.* Molecular & Cellular Proteomics : MCP, 2017. **16**(12): p. 2296-2309.
58. Bilbao, A., et al., *Processing strategies and software solutions for data-independent acquisition in mass spectrometry.* PROTEOMICS, 2015. **15**(5-6): p. 964-980.
59. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data.* Electrophoresis, 1999. **20**(18): p. 3551-67.
60. Cox, J., et al., *Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment.* Journal of Proteome Research, 2011. **10**(4): p. 1794-1805.

61. Eng, J.K., A.L. McCormack, and J.R. Yates, *An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database.* J Am Soc Mass Spectrom, 1994. **5**(11): p. 976-89.
62. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.* Nat Biotech, 2008. **26**(12): p. 1367-1372.
63. Elias, J.E. and S.P. Gygi, *Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry.* Nat Methods, 2007. **4**(3): p. 207-14.
64. Nesvizhskii, A.I., *A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics.* J Proteomics, 2010. **73**(11): p. 2092-123.
65. Muth, T. and B.Y. Renard, *Evaluating de novo sequencing in proteomics: already an accurate alternative to database-driven peptide identification?* Brief Bioinform, 2018. **19**(5): p. 954-970.
66. Doellinger, J., L. Schaade, and A. Nitsche, *Comparison of the Cowpox Virus and Vaccinia Virus Mature Virion Proteome: Analysis of the Species- and Strain-Specific Proteome.* PLoS One, 2015. **10**(11): p. e0141527.
67. Hettich, R.L., et al., *Metaproteomics: harnessing the power of high performance mass spectrometry to identify the suite of proteins that control metabolic activities in microbial communities.* Anal Chem, 2013. **85**(9): p. 4203-14.
68. Zhu, Y., et al., *SpliceVista, a tool for splice variant identification and visualization in shotgun proteomics data.* Mol Cell Proteomics, 2014. **13**(6): p. 1552-62.
69. Chick, J.M., et al., *A mass-tolerant database search identifies a large proportion of unassigned spectra in shotgun proteomics as modified peptides.* Nat Biotechnol, 2015. **33**(7): p. 743-9.
70. Tran, N.H., et al., *Complete De Novo Assembly of Monoclonal Antibody Sequences.* Sci Rep, 2016. **6**: p. 31730.
71. Guthals, A., et al., *De Novo MS/MS Sequencing of Native Human Antibodies.* J Proteome Res, 2017. **16**(1): p. 45-54.
72. Sakurai, T., et al., *PAAS 3: A computer program to determine probable sequence of peptides from mass spectrometric data.* Biomedical Mass Spectrometry, 1984. **11**(8): p. 396-399.
73. Hamm, C.W., W.E. Wilson, and D.J. Harvan, *Peptide sequencing program.* Comput Appl Biosci, 1986. **2**(2): p. 115-8.
74. Johnson, R.S. and K. Biemann, *Computer program (SEQPEP) to aid in the interpretation of high-energy collision tandem mass spectra of peptides.* Biomed Environ Mass Spectrom, 1989. **18**(11): p. 945-57.
75. Dancík, V., et al., *De novo peptide sequencing via tandem mass spectrometry.* J Comput Biol, 1999. **6**(3-4): p. 327-42.
76. Taylor, J.A. and R.S. Johnson, *Sequence database searches via de novo peptide sequencing by tandem mass spectrometry.* Rapid Commun Mass Spectrom, 1997. **11**(9): p. 1067-75.
77. Lu, B. and T. Chen, *Algorithms for de novo peptide sequencing using tandem mass spectrometry.* Drug Discovery Today: BIOSILICO, 2004. **2**(2): p. 85-90.
78. Frank, A. and P. Pevzner, *PepNovo: de novo peptide sequencing via probabilistic network modeling.* Anal Chem, 2005. **77**(4): p. 964-73.
79. Chi, H., et al., *pNovo+: de novo peptide sequencing using complementary HCD and ETD tandem mass spectra.* J Proteome Res, 2013. **12**(2): p. 615-25.
80. Ma, B., *Novor: real-time peptide de novo sequencing software.* J Am Soc Mass Spectrom, 2015. **26**(11): p. 1885-94.
81. Ma, B., et al., *PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry.* Rapid Communications in Mass Spectrometry, 2003. **17**(20): p. 2337-2342.

82. Ankney, J.A., A. Muneer, and X. Chen, *Relative and Absolute Quantitation in Mass Spectrometry–Based Proteomics.* Annual Review of Analytical Chemistry, 2016. **11**(1): p. 49-77.
83. Rabilloud, T., *Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains.* Proteomics, 2002. **2**(1): p. 3-10.
84. Ong, S.E. and M. Mann, *A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC).* Nat Protoc, 2006. **1**(6): p. 2650-60.
85. Mann, M., *Functional and quantitative proteomics using SILAC.* Nat Rev Mol Cell Biol, 2006. **7**(12): p. 952-8.
86. Krüger, M., et al., *SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function.* Cell, 2008. **134**(2): p. 353-64.
87. Sury, M.D., J.X. Chen, and M. Selbach, *The SILAC fly allows for accurate protein quantification in vivo.* Mol Cell Proteomics, 2010. **9**(10): p. 2173-83.
88. Matthes, A., K. Köhl, and W.X. Schulze, *SILAC and alternatives in studying cellular proteomes of plants.* Methods Mol Biol, 2014. **1188**: p. 65-83.
89. Tyanova, S., M. Mann, and J. Cox, *MaxQuant for in-depth analysis of large SILAC datasets.* Methods Mol Biol, 2014. **1188**: p. 351-64.
90. Kratchmarova, I., et al., *Mechanism of Divergent Growth Factor Effects in Mesenchymal Stem Cell Differentiation.* Science, 2005. **308**(5727): p. 1472.
91. Gruhler, A., et al., *Quantitative Phosphoproteomics Applied to the Yeast Pheromone Signaling Pathway*S.* Molecular & Cellular Proteomics, 2005. **4**(3): p. 310-327.
92. Geiger, T., et al., *Super-SILAC mix for quantitative proteomics of human tumor tissue.* Nat Methods, 2010. **7**(5): p. 383-5.
93. Zeiler, M., et al., *A Protein Epitope Signature Tag (PrEST) library allows SILAC-based absolute quantification and multiplexed determination of protein copy numbers in cell lines.* Mol Cell Proteomics, 2012. **11**(3): p. O111.009613.
94. Kettenbach, A.N., J. Rush, and S.A. Gerber, *Absolute quantification of protein and post-translational modification abundance with stable isotope–labeled synthetic peptides.* Nature Protocols, 2011. **6**(2): p. 175-186.
95. Gerber, S.A., et al., *Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS.* Proceedings of the National Academy of Sciences, 2003. **100**(12): p. 6940.
96. Beynon, R.J., et al., *Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides.* Nature Methods, 2005. **2**(8): p. 587-589.
97. Thompson, A., et al., *Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS.* Analytical Chemistry, 2003. **75**(8): p. 1895-1904.
98. Ross, P.L., et al., *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents.* Mol Cell Proteomics, 2004. **3**(12): p. 1154-69.
99. DeSouza, L.V., et al., *Endometrial Carcinoma Biomarker Discovery and Verification Using Differentially Tagged Clinical Samples with Multidimensional Liquid Chromatography and Tandem Mass Spectrometry*.* Molecular & Cellular Proteomics, 2007. **6**(7): p. 1170-1182.
100. Ow, S.Y., et al., *Minimising iTRAQ ratio compression through understanding LC-MS elution dependence and high-resolution HILIC fractionation.* PROTEOMICS, 2011. **11**(11): p. 2341-2346.
101. Rauniyar, N. and J.R. Yates, *Isobaric Labeling-Based Relative Quantification in Shotgun Proteomics.* Journal of Proteome Research, 2014. **13**(12): p. 5293-5309.
102. Altelaar, A.F.M., et al., *Benchmarking stable isotope labeling based quantitative proteomics.* Journal of Proteomics, 2013. **88**: p. 14-26.

103. Ma, Y., et al., *iTRAQ-based quantitative analysis of cancer-derived secretory proteome reveals TPM2 as a potential diagnostic biomarker of colorectal cancer.* Frontiers of Medicine, 2016. **10**(3): p. 278-285.
104. Xin, Q.-L., et al., *Quantitative Proteomic Analysis of Mosquito C6/36 Cells Reveals Host Proteins Involved in Zika Virus Infection.* Journal of Virology, 2017. **91**(12): p. e00554-17.
105. Wang, P., et al., *Tau interactome mapping based identification of Otub1 as Tau deubiquitinase involved in accumulation of pathological Tau forms in vitro and in vivo.* Acta Neuropathologica, 2017. **133**(5): p. 731-749.
106. Dayon, L., et al., *Relative Quantification of Proteins in Human Cerebrospinal Fluids by MS/MS Using 6-Plex Isobaric Tags.* Analytical Chemistry, 2008. **80**(8): p. 2921-2931.
107. Werner, T., et al., *Ion coalescence of neutron encoded TMT 10-plex reporter ions.* Anal Chem, 2014. **86**(7): p. 3594-601.
108. Karp, N.A., et al., *Addressing Accuracy and Precision Issues in iTRAQ Quantitation*.* Molecular & Cellular Proteomics, 2010. **9**(9): p. 1885-1897.
109. Ow, S.Y., et al., *iTRAQ Underestimation in Simple and Complex Mixtures: "The Good, the Bad and the Ugly".* Journal of Proteome Research, 2009. **8**(11): p. 5347-5355.
110. Ting, L., et al., *MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics.* Nat Methods, 2011. **8**(11): p. 937-40.
111. Virreira Winter, S., et al., *EASI-tag enables accurate multiplexed and interference-free MS2-based proteome quantification.* Nature Methods, 2018. **15**(7): p. 527-530.
112. Old, W.M., et al., *Comparison of Label-free Methods for Quantifying Human Proteins by Shotgun Proteomics*S.* Molecular & Cellular Proteomics, 2005. **4**(10): p. 1487-1502.
113. Rai, A.J., et al., *HUPO Plasma Proteome Project specimen collection and handling: towards the standardization of parameters for plasma proteome samples.* Proteomics, 2005. **5**(13): p. 3262-77.
114. Luk, V.N. and A.R. Wheeler, *A digital microfluidic approach to proteomic sample processing.* Anal Chem, 2009. **81**(11): p. 4524-30.
115. Chelius, D. and P.V. Bondarenko, *Quantitative Profiling of Proteins in Complex Mixtures Using Liquid Chromatography and Mass Spectrometry.* Journal of Proteome Research, 2002. **1**(4): p. 317-323.
116. Kumar, A. and M. Snyder, *Protein complexes take the bait.* Nature, 2002. **415**(6868): p. 123-124.
117. Fields, S. and O. Song, *A novel genetic system to detect protein-protein interactions.* Nature, 1989. **340**(6230): p. 245-6.
118. Chien, C.T., et al., *The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest.* Proc Natl Acad Sci U S A, 1991. **88**(21): p. 9578-82.
119. Michnick, S.W., et al., *Detection of protein-protein interactions by protein fragment complementation strategies.* Methods Enzymol, 2000. **328**: p. 208-30.
120. Barrios-Rodiles, M., et al., *High-throughput mapping of a dynamic signaling network in mammalian cells.* Science, 2005. **307**(5715): p. 1621-5.
121. Selvin, P.R., *Fluorescence resonance energy transfer.* Methods Enzymol, 1995. **246**: p. 300-34.
122. Yugandhar, K., S. Gupta, and H. Yu, *Inferring Protein-Protein Interaction Networks From Mass Spectrometry-Based Proteomic Approaches: A Mini-Review.* Computational and structural biotechnology journal, 2019. **17**: p. 805-811.
123. Iacobucci, I., et al., *From classical to new generation approaches: An excursus of -omics methods for investigation of protein-protein interaction networks.* Journal of Proteomics, 2021. **230**: p. 103990.

124. Keilhauer, E.C., M.Y. Hein, and M. Mann, *Accurate Protein Complex Retrieval by Affinity Enrichment Mass Spectrometry (AE-MS) Rather than Affinity Purification Mass Spectrometry (AP-MS).* Molecular & Cellular Proteomics : MCP, 2015. **14**(1): p. 120-135.
125. Gingras, A.C., et al., *Analysis of protein complexes using mass spectrometry.* Nat Rev Mol Cell Biol, 2007. **8**(8): p. 645-54.
126. Selbach, M. and M. Mann, *Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK).* Nat Methods, 2006. **3**(12): p. 981-3.
127. Malovannaya, A., et al., *Analysis of the human endogenous coregulator complexome.* Cell, 2011. **145**(5): p. 787-99.
128. Dunham, W.H., M. Mullin, and A.-C. Gingras, *Affinity-purification coupled to mass spectrometry: Basic principles and strategies.* PROTEOMICS, 2012. **12**(10): p. 1576-1590.
129. Dekker, C., et al., *The interaction network of the chaperonin CCT.* Embo j, 2008. **27**(13): p. 1827-39.
130. Evan, G.I., et al., *Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product.* Mol Cell Biol, 1985. **5**(12): p. 3610-6.
131. Hopp, T.P., et al., *A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification.* Bio/Technology, 1988. **6**(10): p. 1204-1210.
132. Field, J., et al., *Purification of a RAS-responsive adenylyl cyclase complex from Saccharomyces cerevisiae by use of an epitope addition method.* Molecular and cellular biology, 1988. **8**(5): p. 2159-2165.
133. Hochuli, E., et al., *Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent.* Bio/Technology, 1988. **6**(11): p. 1321-1325.
134. Schmidt, T.G. and A. Skerra, *The random peptide library-assisted engineering of a C-terminal affinity peptide, useful for the detection and purification of a functional Ig Fv fragment.* Protein Eng, 1993. **6**(1): p. 109-22.
135. Smith, D.B. and K.S. Johnson, *Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase.* Gene, 1988. **67**(1): p. 31-40.
136. Chalfie, M., et al., *Green fluorescent protein as a marker for gene expression.* Science, 1994. **263**(5148): p. 802-5.
137. Uhlén, M., et al., *Gene fusion vectors based on the gene for staphylococcal protein A.* Gene, 1983. **23**(3): p. 369-78.
138. Gingras, A.-C., et al., *Analysis of protein complexes using mass spectrometry.* Nature Reviews Molecular Cell Biology, 2007. **8**(8): p. 645-654.
139. Goel, A., et al., *Relative position of the hexahistidine tag effects binding properties of a tumor-associated single-chain Fv construct.* Biochim Biophys Acta, 2000. **1523**(1): p. 13-20.
140. Hofemeister, H., et al., *Recombineering, transfection, Western, IP and ChIP methods for protein tagging via gene targeting or BAC transgenesis.* Methods, 2011. **53**(4): p. 437-52.
141. Rumlová, M., et al., *Comparison of classical and affinity purification techniques of Mason-Pfizer monkey virus capsid protein: the alteration of the product by an affinity tag.* Protein Expr Purif, 2001. **23**(1): p. 75-83.
142. Sowa, M.E., et al., *Defining the human deubiquitinating enzyme interaction landscape.* Cell, 2009. **138**(2): p. 389-403.
143. Goudreault, M., et al., *A PP2A phosphatase high density interaction network identifies a novel striatin-interacting phosphatase and kinase complex linked to the cerebral cavernous malformation 3 (CCM3) protein.* Mol Cell Proteomics, 2009. **8**(1): p. 157-71.
144. Mak, A.B., et al., *A lentiviral functional proteomics approach identifies chromatin remodeling complexes important for the induction of pluripotency.* Mol Cell Proteomics, 2010. **9**(5): p. 811-23.

145. Sardiu, M.E., et al., *Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics.* Proc Natl Acad Sci U S A, 2008. **105**(5): p. 1454-9.
146. Jansen, G., et al., *Drag&Drop cloning in yeast.* Gene, 2005. **344**: p. 43-51.
147. Jeronimo, C., et al., *Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme.* Mol Cell, 2007. **27**(2): p. 262-74.
148. Glatter, T., et al., *An integrated workflow for charting the human interaction proteome: insights into the PP2A system.* Molecular Systems Biology, 2009. **5**(1): p. 237.
149. Torres, J.Z., J.J. Miller, and P.K. Jackson, *High-throughput generation of tagged stable cell lines for proteomic analysis.* Proteomics, 2009. **9**(10): p. 2888-91.
150. Kittler, R., et al., *RNA interference rescue by bacterial artificial chromosome transgenesis in mammalian tissue culture cells.* Proc Natl Acad Sci U S A, 2005. **102**(7): p. 2396-401.
151. Hubner, N.C. and M. Mann, *Extracting gene function from protein-protein interactions using Quantitative BAC InteraCtomics (QUBIC).* Methods, 2011. **53**(4): p. 453-9.
152. Hein, Marco Y., et al., *A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances.* Cell, 2015. **163**(3): p. 712-723.
153. Anzalone, A.V., et al., *Search-and-replace genome editing without double-strand breaks or donor DNA.* Nature, 2019. **576**(7785): p. 149-157.
154. Chen, G.I., et al., *PP4R4/KIAA1622 forms a novel stable cytosolic complex with phosphoprotein phosphatase 4.* J Biol Chem, 2008. **283**(43): p. 29273-84.
155. Strambio-de-Castillia, C., et al., *A method for the rapid and efficient elution of native affinity-purified protein A tagged complexes.* J Proteome Res, 2005. **4**(6): p. 2250-6.
156. Cristea, I.M., et al., *Fluorescent proteins as proteomic probes.* Mol Cell Proteomics, 2005. **4**(12): p. 1933-41.
157. Oeffinger, M., et al., *Comprehensive analysis of diverse ribonucleoprotein complexes.* Nat Methods, 2007. **4**(11): p. 951-6.
158. Rigaut, G., et al., *A generic protein purification method for protein complex characterization and proteome exploration.* Nat Biotechnol, 1999. **17**(10): p. 1030-2.
159. Gavin, A.C., et al., *Proteome survey reveals modularity of the yeast cell machinery.* Nature, 2006. **440**(7084): p. 631-6.
160. Krogan, N.J., et al., *Global landscape of protein complexes in the yeast Saccharomyces cerevisiae.* Nature, 2006. **440**: p. 637.
161. Gloeckner, C.J., et al., *A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes.* Proteomics, 2007. **7**(23): p. 4228-34.
162. Bürckstümmer, T., et al., *An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells.* Nat Methods, 2006. **3**(12): p. 1013-9.
163. Park, J., et al., *The impact of cellular networks on disease comorbidity.* Mol Syst Biol, 2009. **5**: p. 262.
164. Behrends, C., et al., *Network organization of the human autophagy system.* Nature, 2010. **466**(7302): p. 68-76.
165. Choi, H., et al., *Analyzing protein-protein interactions from affinity purification-mass spectrometry data with SAINT.* Current protocols in bioinformatics, 2012. **Chapter 8**: p. Unit8.15-Unit8.15.
166. Mellacheruvu, D., et al., *The CRAPome: a contaminant repository for affinity purification–mass spectrometry data.* Nature Methods, 2013. **10**: p. 730.
167. Acuner Ozbabacan, S.E., et al., *Transient protein–protein interactions.* Protein Engineering, Design and Selection, 2011. **24**(9): p. 635-648.

168. Kaake, R.M., X. Wang, and L. Huang, *Profiling of protein interaction networks of protein complexes using affinity purification and quantitative mass spectrometry.* Mol Cell Proteomics, 2010. **9**(8): p. 1650-65.
169. Agou, F. and M. Véron, *In Vivo Protein Cross-Linking*, in *Protein-Protein Interactions: Methods and Applications*, C.L. Meyerkord and H. Fu, Editors. 2015, Springer New York: New York, NY. p. 391-405.
170. Tagwerker, C., et al., *A tandem affinity tag for two-step purification under fully denaturing conditions: application in ubiquitin profiling and protein complex identification combined with in vivocross-linking.* Mol Cell Proteomics, 2006. **5**(4): p. 737-48.
171. Guerrero, C., et al., *An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network.* Mol Cell Proteomics, 2006. **5**(2): p. 366-78.
172. Vasilescu, J., X. Guo, and J. Kast, *Identification of protein-protein interactions using in vivo cross-linking and mass spectrometry.* Proteomics, 2004. **4**(12): p. 3845-54.
173. Suchanek, M., A. Radzikowska, and C. Thiele, *Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells.* Nat Methods, 2005. **2**(4): p. 261-7.
174. Meunier, L., et al., *A subset of chaperones and folding enzymes form multiprotein complexes in endoplasmic reticulum to bind nascent proteins.* Molecular biology of the cell, 2002. **13**(12): p. 4456-4469.
175. Wong, S.S. and L.-J.C. Wong, *Chemical crosslinking and the stabilization of proteins and enzymes.* Enzyme and Microbial Technology, 1992. **14**(11): p. 866-874.
176. Fux, A., et al., *Chemical Cross-Linking Enables Drafting ClpXP Proximity Maps and Taking Snapshots of In Situ Interaction Networks.* Cell Chemical Biology, 2019. **26**(1): p. 48-59.e7.
177. Roux, K.J., et al., *A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells.* J Cell Biol, 2012. **196**(6): p. 801-10.
178. Hung, V., et al., *Proteomic Mapping of the Human Mitochondrial Intermembrane Space in Live Cells via Ratiometric APEX Tagging.* Molecular Cell, 2014. **55**(2): p. 332-341.
179. Rees, J.S., et al., *Selective Proteomic Proximity Labeling Assay Using Tyramide (SPPLAT): A Quantitative Method for the Proteomic Analysis of Localized Membrane-Bound Protein Clusters.* Current Protocols in Protein Science, 2017. **88**(1): p. 19.27.1-19.27.18.
180. Honke, K. and N. Kotani, *Identification of Cell-Surface Molecular Interactions under Living Conditions by Using the Enzyme-Mediated Activation of Radical Sources (EMARS) Method.* Sensors, 2012. **12**(12).
181. Geri, J.B., et al., *Microenvironment mapping via Dexter energy transfer on immune cells.* Science, 2020. **367**(6482): p. 1091-1097.
182. Li, P., et al., *Proximity Labeling of Interacting Proteins: Application of BioID as a Discovery Tool.* PROTEOMICS, 2017. **17**(20): p. 1700002.
183. Savitski, M.F. and M.M. Savitski, *Unbiased detection of posttranslational modifications using mass spectrometry.* Methods Mol Biol, 2010. **673**: p. 203-10.
184. Tarrant, M.K. and P.A. Cole, *The chemical biology of protein phosphorylation.* Annu Rev Biochem, 2009. **78**: p. 797-825.
185. Salazar, C. and T. Höfer, *Multisite protein phosphorylation--from molecular mechanisms to kinetic models.* Febs j, 2009. **276**(12): p. 3177-98.
186. Moura, M. and C. Conde, *Phosphatases in Mitosis: Roles and Regulation.* Biomolecules, 2019. **9**(2).
187. Carpy, A., et al., *Absolute proteome and phosphoproteome dynamics during the cell cycle of Schizosaccharomyces pombe (Fission Yeast).* Mol Cell Proteomics, 2014. **13**(8): p. 1925-36.

188. Ptacek, J., et al., *Global analysis of protein phosphorylation in yeast.* Nature, 2005. **438**(7068): p. 679-84.
189. Blume-Jensen, P. and T. Hunter, *Oncogenic kinase signalling.* Nature, 2001. **411**(6835): p. 355-65.
190. Cohen, P., *The role of protein phosphorylation in human health and disease. The Sir Hans Krebs Medal Lecture.* Eur J Biochem, 2001. **268**(19): p. 5001-10.
191. Cohen, P., *Protein kinases--the major drug targets of the twenty-first century?* Nat Rev Drug Discov, 2002. **1**(4): p. 309-15.
192. Olow, A., et al., *An Atlas of the Human Kinome Reveals the Mutational Landscape Underlying Dysregulated Phosphorylation Cascades in Cancer.* Cancer Res, 2016. **76**(7): p. 1733-45.
193. Bhullar, K.S., et al., *Kinase-targeted cancer therapies: progress, challenges and future directions.* Mol Cancer, 2018. **17**(1): p. 48.
194. Olsen, J.V., et al., *Global, in vivo, and site-specific phosphorylation dynamics in signaling networks.* Cell, 2006. **127**(3): p. 635-48.
195. Besant, P.G. and P.V. Attwood, *Mammalian histidine kinases.* Biochim Biophys Acta, 2005. **1754**(1-2): p. 281-90.
196. Fuhs, S.R., et al., *Monoclonal 1- and 3-Phosphohistidine Antibodies: New Tools to Study Histidine Phosphorylation.* Cell, 2015. **162**(1): p. 198-210.
197. Hardman, G., et al., *Strong anion exchange-mediated phosphoproteomics reveals extensive human non-canonical phosphorylation.* Embo j, 2019. **38**(21): p. e100847.
198. Needham, E.J., et al., *Illuminating the dark phosphoproteome.* Sci Signal, 2019. **12**(565).
199. Humphrey, S.J., S.B. Azimifar, and M. Mann, *High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics.* Nature Biotechnology, 2015. **33**: p. 990.
200. Bekker-Jensen, D.B., et al., *Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries.* Nat Commun, 2020. **11**(1): p. 787.
201. Gaberc-Porekar, V. and V. Menart, *Perspectives of immobilized-metal affinity chromatography.* J Biochem Biophys Methods, 2001. **49**(1-3): p. 335-60.
202. Thingholm, T.E. and M.R. Larsen, *The use of titanium dioxide micro-columns to selectively isolate phosphopeptides from proteolytic digests.* Methods Mol Biol, 2009. **527**: p. 57-66, xi.
203. Thingholm, T.E., et al., *Highly selective enrichment of phosphorylated peptides using titanium dioxide.* Nat Protoc, 2006. **1**(4): p. 1929-35.
204. Larsen, M.R., et al., *Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns.* Mol Cell Proteomics, 2005. **4**(7): p. 873-86.
205. Pinkse, M.W., et al., *Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns.* Anal Chem, 2004. **76**(14): p. 3935-43.
206. Aryal, U.K., D.J. Olson, and A.R. Ross, *Optimization of immobilized gallium (III) ion affinity chromatography for selective binding and recovery of phosphopeptides from protein digests.* J Biomol Tech, 2008. **19**(5): p. 296-310.
207. Gates, M.B., K.B. Tomer, and L.J. Deterding, *Comparison of metal and metal oxide media for phosphopeptide enrichment prior to mass spectrometric analyses.* J Am Soc Mass Spectrom, 2010. **21**(10): p. 1649-59.
208. Pandey, A., et al., *Analysis of receptor signaling pathways by mass spectrometry: identification of vav-2 as a substrate of the epidermal and platelet-derived growth factor receptors.* Proc Natl Acad Sci U S A, 2000. **97**(1): p. 179-84.
209. Haas, A.L., et al., *Interferon induces a 15-kilodalton protein exhibiting marked homology to ubiquitin.* Journal of Biological Chemistry, 1987. **262**(23): p. 11315-11323.

210. Perng, Y.C. and D.J. Lenschow, *ISG15 in antiviral immunity and beyond.* Nat Rev Microbiol, 2018. **16**(7): p. 423-439.
211. Pitha-Rowe, I., et al., *Retinoid target genes in acute promyelocytic leukemia.* Leukemia, 2003. **17**(9): p. 1723-30.
212. Kim, K.I., et al., *Enhanced antibacterial potential in UBP43-deficient mice against Salmonella typhimurium infection by up-regulating type I IFN signaling.* J Immunol, 2005. **175**(2): p. 847-54.
213. Jeon, Y.J., J.H. Park, and C.H. Chung, *Interferon-Stimulated Gene 15 in the Control of Cellular Responses to Genotoxic Stress.* Mol Cells, 2017. **40**(2): p. 83-89.
214. Loeb, K.R. and A.L. Haas, *The interferon-inducible 15-kDa ubiquitin homolog conjugates to intracellular proteins.* 1992. **267**(11): p. 7806-13.
215. Skaug, B. and Z.J. Chen, *Emerging role of ISG15 in antiviral immunity.* Cell, 2010. **143**(2): p. 187-90.
216. Shi, H.X., et al., *Positive regulation of interferon regulatory factor 3 activation by Herc5 via ISG15 modification.* Mol Cell Biol, 2010. **30**(10): p. 2424-36.
217. Sanyal, S., et al., *Type I interferon imposes a TSG101/ISG15 checkpoint at the Golgi for glycoprotein trafficking during influenza virus infection.* Cell Host Microbe, 2013. **14**(5): p. 510-21.
218. dos Santos, P.F. and D.S. Mansur, *Beyond ISGlylation: Functions of Free Intracellular and Extracellular ISG15.* Journal of Interferon & Cytokine Research, 2017. **37**(6): p. 246-253.
219. Werneke, S.W., et al., *ISG15 is critical in the control of Chikungunya virus infection independent of UbE1L mediated conjugation.* PLoS Pathog, 2011. **7**(10): p. e1002322.
220. Bogunovic, D., et al., *Mycobacterial Disease and Impaired IFN-γ Immunity in Humans with Inherited ISG15 Deficiency.* Science, 2012. **337**(6102): p. 1684.
221. Swaim, C.D., et al., *Extracellular ISG15 Signals Cytokine Secretion through the LFA-1 Integrin Receptor.* Mol Cell, 2017. **68**(3): p. 581-590.e5.
222. D'Cunha, J., et al., *In vitro and in vivo secretion of human ISG15, an IFN-induced immunomodulatory cytokine.* J Immunol, 1996. **157**(9): p. 4100-8.
223. Zhang, D. and D.-E. Zhang, *Interferon-stimulated gene 15 and the protein ISGylation system.* Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research, 2011. **31**(1): p. 119-130.
224. Wagner, S.A., et al., *A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles.* Molecular & cellular proteomics : MCP, 2011. **10**(10): p. M111.013284-M111.013284.
225. Elia, A.E., et al., *Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response.* Mol Cell, 2015. **59**(5): p. 867-81.
226. Rose, C.M., et al., *Highly Multiplexed Quantitative Mass Spectrometry Analysis of Ubiquitylomes.* Cell Syst, 2016. **3**(4): p. 395-403.e4.
227. Kim, W., et al., *Systematic and quantitative assessment of the ubiquitin-modified proteome.* Molecular cell, 2011. **44**(2): p. 325-340.
228. Giannakopoulos, N.V., et al., *Proteomic identification of proteins conjugated to ISG15 in mouse and human cells.* Biochem Biophys Res Commun, 2005. **336**(2): p. 496-506.
229. Zhao, C., et al., *Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways.* Proc Natl Acad Sci U S A, 2005. **102**(29): p. 10200-5.
230. Wong, J.J., et al., *HERC5 is an IFN-induced HECT-type E3 protein ligase that mediates type I IFN-induced ISGylation of protein targets.* Proc Natl Acad Sci U S A, 2006. **103**(28): p. 10735-40.
231. Zhang, Y., et al., *The in vivo ISGylome links ISG15 to metabolic pathways and autophagy upon Listeria monocytogenes infection.* Nature Communications, 2019. **10**(1): p. 5383.

232. Canbay, V. and U. auf dem Keller, *New strategies to identify protease substrates.* Current Opinion in Chemical Biology, 2021. **60**: p. 89-96.
233. 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, 2018. **46**(D1): p. D624-D632.
234. Turk, B., D. Turk, and V. Turk, *Protease signalling: the cutting edge.* The EMBO Journal, 2012. **31**(7): p. 1630-1643.
235. Bhagwat, S.R., K. Hajela, and A. Kumar, *Proteolysis to Identify Protease Substrates: Cleave to Decipher.* PROTEOMICS, 2018. **18**(13): p. 1800011.
236. Pérez-Silva, J.G., et al., *The Degradome database: expanding roles of mammalian proteases in life and disease.* Nucleic Acids Res, 2016. **44**(D1): p. D351-5.
237. Death, A.K., et al., *High glucose alters matrix metalloproteinase expression in two key vascular cells: potential impact on atherosclerosis in diabetes.* Atherosclerosis, 2003. **168**(2): p. 263-9.
238. Markmann, A., et al., *Down-regulation of calpain 9 is linked to hypertensive heart and kidney disease.* Cell Physiol Biochem, 2005. **15**(1-4): p. 109-16.
239. Talanian, R.V., et al., *Substrate specificities of caspase family proteases.* J Biol Chem, 1997. **272**(15): p. 9677-82.
240. Mahrus, S., et al., *Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini.* Cell, 2008. **134**(5): p. 866-76.
241. Gevaert, K., et al., *Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides.* Nat Biotechnol, 2003. **21**(5): p. 566-9.
242. Yoshihara, H.A.I., S. Mahrus, and J.A. Wells, *Tags for labeling protein N-termini with subtiligase for proteomics.* Bioorganic & Medicinal Chemistry Letters, 2008. **18**(22): p. 6000-6003.
243. Weeks, A.M. and J.A. Wells, *N-Terminal Modification of Proteins with Subtiligase Specificity Variants.* Current Protocols in Chemical Biology, 2020. **12**(1): p. e79.
244. Griswold, A.R., et al., *A Chemical Strategy for Protease Substrate Profiling.* Cell Chemical Biology, 2019. **26**(6): p. 901-907.e6.
245. Shin, S., et al., *Development of Multiplexed Immuno-N-Terminomics to Reveal the Landscape of Proteolytic Processing in Early Embryogenesis of Drosophila melanogaster.* Analytical Chemistry, 2020. **92**(7): p. 4926-4934.
246. Rogers, L.D. and C.M. Overall, *Proteolytic Post-translational Modification of Proteins: Proteomic Tools and Methodology*.* Molecular & Cellular Proteomics, 2013. **12**(12): p. 3532-3542.
247. Ponting, C.P., et al., *Identification of a novel family of presenilin homologues.* Human Molecular Genetics, 2002. **11**(9): p. 1037-1044.
248. Grigorenko, A.P., et al., *Novel Class of Polytopic Proteins with Domains Associated with Putative Protease Activity.* Biochemistry (Moscow), 2002. **67**(7): p. 826-834.
249. Weihofen, A., et al., *Identification of Signal Peptide Peptidase, a Presenilin-Type Aspartic Protease.* Science, 2002. **296**(5576): p. 2215.
250. Fluhrer, R., et al., *A γ-secretase-like intramembrane cleavage of TNFα by the GxGD aspartyl protease SPPL2b.* Nature Cell Biology, 2006. **8**(8): p. 894-896.
251. Friedmann, E., et al., *SPPL2a and SPPL2b promote intramembrane proteolysis of TNFα in activated dendritic cells to trigger IL-12 production.* Nature Cell Biology, 2006. **8**(8): p. 843-848.
252. Fluhrer, R., H. Steiner, and C. Haass, *Intramembrane Proteolysis by Signal Peptide Peptidases: A Comparative Discussion of GXGD-type Aspartyl Proteases*.* Journal of Biological Chemistry, 2009. **284**(21): p. 13975-13979.
253. Friedmann, E., et al., *Consensus analysis of signal peptide peptidase and homologous human aspartic proteases reveals opposite topology of catalytic domains compared with presenilins.* J Biol Chem, 2004. **279**(49): p. 50790-8.

254. Friedmann, E., et al., *SPPL2a and SPPL2b promote intramembrane proteolysis of TNFalpha in activated dendritic cells to trigger IL-12 production.* Nat Cell Biol, 2006. **8**(8): p. 843-8.
255. Behnke, J., et al., *Signal-peptide-peptidase-like 2a (SPPL2a) is targeted to lysosomes/late endosomes by a tyrosine motif in its C-terminal tail.* FEBS Lett, 2011. **585**(19): p. 2951-7.
256. Niemeyer, J., et al., *The intramembrane protease SPPL2c promotes male germ cell development by cleaving phospholamban.* EMBO Rep, 2019. **20**(3).
257. Weihofen, A., et al., *Release of signal peptide fragments into the cytosol requires cleavage in the transmembrane region by a protease activity that is specifically blocked by a novel cysteine protease inhibitor.* (0021-9258 (Print)).
258. Bergmann, H., et al., *B cell survival, surface BCR and BAFFR expression, CD74 metabolism, and CD8− dendritic cells require the intramembrane endopeptidase SPPL2A.* Journal of Experimental Medicine, 2012. **210**(1): p. 31-40.
259. Beisner, D.R., et al., *The intramembrane protease Sppl2a is required for B cell and DC development and survival via cleavage of the invariant chain.* Journal of Experimental Medicine, 2012. **210**(1): p. 23-30.
260. Lumangtad, L.A. and T.W. Bell, *The signal peptide as a new target for drug design.* Bioorganic & medicinal chemistry letters, 2020. **30**(10): p. 127115-127115.
261. Voorhees, R.M. and R.S. Hegde, *Structure of the Sec61 channel opened by a signal sequence.* Science, 2016. **351**(6268): p. 88-91.
262. von Heijne, G., *Signal sequences. The limits of variation.* J Mol Biol, 1985. **184**(1): p. 99-105.
263. Symoens, S., et al., *COL5A1 signal peptide mutations interfere with protein secretion and cause classic Ehlers-Danlos syndrome.* Hum Mutat, 2009. **30**(2): p. E395-403.
264. Seppen, J., et al., *A mutation which disrupts the hydrophobic core of the signal peptide of bilirubin UDP-glucuronosyltransferase, an endoplasmic reticulum membrane protein, causes Crigler-Najjar type II.* FEBS Lett, 1996. **390**(3): p. 294-8.
265. Bafunno, V., et al., *Mutational spectrum of the c1 inhibitor gene in a cohort of Italian patients with hereditary angioedema: description of nine novel mutations.* Ann Hum Genet, 2014. **78**(2): p. 73-82.
266. Laurila, K. and M. Vihinen, *Prediction of disease-related mutations affecting protein localization.* BMC Genomics, 2009. **10**: p. 122.
267. Levine, C.G., et al., *The efficiency of protein compartmentalization into the secretory pathway.* Mol Biol Cell, 2005. **16**(1): p. 279-91.
268. Kang, S.W., et al., *Substrate-specific translocational attenuation during ER stress defines a pre-emptive quality control pathway.* Cell, 2006. **127**(5): p. 999-1013.
269. Piersma, D., et al., *A common polymorphism renders the luteinizing hormone receptor protein more active by improving signal peptide function and predicts adverse outcome in breast cancer patients.* J Clin Endocrinol Metab, 2006. **91**(4): p. 1470-6.
270. Nielsen, H., et al., *A Brief History of Protein Sorting Prediction.* The protein journal, 2019. **38**(3): p. 200-216.
271. Ladunga, I., et al., *Improving signal peptide prediction accuracy by simulated neural network.* Comput Appl Biosci, 1991. **7**(4): p. 485-7.
272. Nielsen, H., et al., *Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites.* Protein Eng, 1997. **10**(1): p. 1-6.
273. Nielsen, H. and A. Krogh, *Prediction of signal peptides and signal anchors by a hidden Markov model.* Proc Int Conf Intell Syst Mol Biol, 1998. **6**: p. 122-30.
274. Vert, J.P., *Support vector machine prediction of signal peptide cleavage site using a new class of kernels for strings.* Pac Symp Biocomput, 2002: p. 649-60.

275. Almagro Armenteros, J.J., et al., *SignalP 5.0 improves signal peptide predictions using deep neural networks.* Nat Biotechnol, 2019. **37**(4): p. 420-423.
276. Choo, K.H., T.W. Tan, and S. Ranganathan, *A comprehensive assessment of N-terminal signal peptides prediction methods.* BMC Bioinformatics, 2009. **10**(15): p. S2.
277. Petersen, T.N., et al., *SignalP 4.0: discriminating signal peptides from transmembrane regions.* Nature Methods, 2011. **8**(10): p. 785-786.
278. Stastna, M. and J.E. Van Eyk, *Secreted proteins as a fundamental source for biomarker discovery.* Proteomics, 2012. **12**(4-5): p. 722-35.
279. Farhan, H. and C. Rabouille, *Signalling to and from the secretory pathway.* J Cell Sci, 2011. **124**(Pt 2): p. 171-80.
280. Meissner, F., et al., *Direct proteomic quantification of the secretome of activated immune cells.* Science, 2013. **340**(6131): p. 475-8.
281. Robinson, M.W. and B. Connolly, *Proteomic analysis of the excretory-secretory proteins of the Trichinella spiralis L1 larva, a nematode parasite of skeletal muscle.* Proteomics, 2005. **5**(17): p. 4525-32.
282. Obenauf, A.C., et al., *Therapy-induced tumour secretomes promote resistance and tumour progression.* Nature, 2015. **520**(7547): p. 368-72.
283. Rieckmann, J.C., et al., *Social network architecture of human immune cells unveiled by quantitative proteomics.* Nature Immunology, 2017. **18**: p. 583.
284. Roca-Rivada, A., et al., *CILAIR-Based Secretome Analysis of Obese Visceral and Subcutaneous Adipose Tissues Reveals Distinctive ECM Remodeling and Inflammation Mediators.* Sci Rep, 2015. **5**: p. 12214.
285. Cox, T.R., et al., *The hypoxic cancer secretome induces pre-metastatic bone lesions through lysyl oxidase.* Nature, 2015. **522**(7554): p. 106-110.
286. Blanco, M.A., et al., *Global secretome analysis identifies novel mediators of bone metastasis.* Cell Res, 2012. **22**(9): p. 1339-55.
287. Jin, L., et al., *Differential secretome analysis reveals CST6 as a suppressor of breast cancer bone metastasis.* Cell Res, 2012. **22**(9): p. 1356-73.
288. Kaur, A., et al., *sFRP2 in the aged microenvironment drives melanoma metastasis and therapy resistance.* Nature, 2016. **532**(7598): p. 250-4.
289. Mukherjee, P. and S. Mani, *Methodologies to decipher the cell secretome.* Biochim Biophys Acta, 2013. **1834**(11): p. 2226-32.
290. Ranganath, S.H., et al., *Harnessing the mesenchymal stem cell secretome for the treatment of cardiovascular disease.* Cell Stem Cell, 2012. **10**(3): p. 244-58.
291. Bárcena, C., et al., *Angiogenin secretion from hepatoma cells activates hepatic stellate cells to amplify a self-sustained cycle promoting liver cancer.* Sci Rep, 2015. **5**: p. 7916.
292. Weng, Y., et al., *In-Depth Proteomic Quantification of Cell Secretome in Serum-Containing Conditioned Medium.* Anal Chem, 2016. **88**(9): p. 4971-8.
293. Eichelbaum, K., et al., *Selective enrichment of newly synthesized proteins for quantitative secretome analysis.* Nat Biotechnol, 2012. **30**(10): p. 984-90.
294. Frauenstein, A. and F. Meissner, *Quantitative Proteomics of Secreted Proteins.* Methods Mol Biol, 2018. **1714**: p. 215-227.
295. Akira, S., S. Uematsu, and O. Takeuchi, *Pathogen recognition and innate immunity.* Cell, 2006. **124**(4): p. 783-801.
296. Janeway, C.A., Jr. and R. Medzhitov, *Innate immune recognition.* Annu Rev Immunol, 2002. **20**: p. 197-216.
297. Celhar, T., R. Magalhães, and A.M. Fairhurst, *TLR7 and TLR9 in SLE: when sensing self goes wrong.* Immunol Res, 2012. **53**(1-3): p. 58-77.

298. Kawai, T. and S. Akira, *The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors.* Nat Immunol, 2010. **11**(5): p. 373-84.
299. Satoh, T. and S. Akira, *Toll-Like Receptor Signaling and Its Inducible Proteins.* 2016. **4**(6).
300. Vogl, T., et al., *Mrp8 and Mrp14 are endogenous activators of Toll-like receptor 4, promoting lethal, endotoxin-induced shock.* Nat Med, 2007. **13**(9): p. 1042-9.
301. Pouliot, P., et al., *Myeloid-related proteins rapidly modulate macrophage nitric oxide production during innate immune response.* J Immunol, 2008. **181**(5): p. 3595-601.
302. Suganami, T., et al., *Attenuation of obesity-induced adipose tissue inflammation in C3H/HeJ mice carrying a Toll-like receptor 4 mutation.* Biochem Biophys Res Commun, 2007. **354**(1): p. 45-9.
303. Suganami, T., et al., *Role of the Toll-like receptor 4/NF-kappaB pathway in saturated fatty acid-induced inflammatory changes in the interaction between adipocytes and macrophages.* Arterioscler Thromb Vasc Biol, 2007. **27**(1): p. 84-91.
304. Zanoni, I., et al., *CD14 controls the LPS-induced endocytosis of Toll-like receptor 4.* Cell, 2011. **147**(4): p. 868-80.
305. Lin, S.C., Y.C. Lo, and H. Wu, *Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling.* Nature, 2010. **465**(7300): p. 885-90.
306. Kollewe, C., et al., *Sequential autophosphorylation steps in the interleukin-1 receptor-associated kinase-1 regulate its availability as an adapter in interleukin-1 signaling.* J Biol Chem, 2004. **279**(7): p. 5227-36.
307. Jiang, Z., et al., *Interleukin-1 (IL-1) receptor-associated kinase-dependent IL-1-induced signaling complexes phosphorylate TAK1 and TAB2 at the plasma membrane and activate TAK1 in the cytosol.* Mol Cell Biol, 2002. **22**(20): p. 7158-67.
308. Ajibade, A.A., H.Y. Wang, and R.F. Wang, *Cell type-specific function of TAK1 in innate immune signaling.* Trends Immunol, 2013. **34**(7): p. 307-16.
309. Chen, Z.J., *Ubiquitination in signaling to and activation of IKK.* Immunol Rev, 2012. **246**(1): p. 95-106.
310. Frauenstein, A., et al., *Identification of covalent modifications regulating immune signaling complex composition and phenotype.* Mol Syst Biol, 2021. **17**(7): p. e10125.
311. Blobel G Fau - Dobberstein, B. and B. Dobberstein, *Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma.* (0021-9525 (Print)).
312. Kim, S.J., et al., *Signal Sequences Control Gating of the Protein Translocation Channel in a Substrate-Specific Manner.* Developmental Cell, 2002. **2**(2): p. 207-217.
313. Zhang, W.-X., X. Pan, and H.-B. Shen, *Signal-3L 3.0: Improving Signal Peptide Prediction through Combining Attention Deep Learning with Window-Based Scoring.* Journal of Chemical Information and Modeling, 2020. **60**(7): p. 3679-3686.
314. Zhang, Z. and W.J. Henzel, *Signal peptide prediction based on analysis of experimentally verified cleavage sites.* Protein science : a publication of the Protein Society, 2004. **13**(10): p. 2819-2824.
315. Schechter, I., *Partial amino acid sequence of the precursor of immunoglobulin light chain programmed by messenger RNA in vitro.* Science, 1975. **188**(4184): p. 160-2.
316. Ying, H. and H. Liu, *Identification of an alternative signal peptide cleavage site of mouse monoclonal antibodies by mass spectrometry.* Immunol Lett, 2007. **111**(1): p. 66-8.
317. Huesgen, P.F. and C.M. Overall, *N- and C-terminal degradomics: new approaches to reveal biological roles for plant proteases from substrate identification.* Physiol Plant, 2012. **145**(1): p. 5-17.
318. Hartmann, E.M. and J. Armengaud, *N-terminomics and proteogenomics, getting off to a good start.* Proteomics, 2014. **14**(23-24): p. 2637-46.

319. Van Damme, P., et al., *N-terminal proteomics and ribosome profiling provide a comprehensive view of the alternative translation initiation landscape in mice and men.* Molecular & cellular proteomics : MCP, 2014. **13**(5): p. 1245-1261.
320. Eckhard, U., et al., *The Human Dental Pulp Proteome and N-Terminome: Levering the Unexplored Potential of Semitryptic Peptides Enriched by TAILS to Identify Missing Proteins in the Human Proteome Project in Underexplored Tissues.* Journal of Proteome Research, 2015. **14**(9): p. 3568-3582.
321. Jain, R.G., S.L. Rusch, and D.A. Kendall, *Signal peptide cleavage regions. Functional limits on length and topological implications.* J Biol Chem, 1994. **269**(23): p. 16305-10.
322. Cox, J., et al., *Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ.* Molecular & cellular proteomics : MCP, 2014. **13**(9): p. 2513-2526.
323. Chou, M.M. and D.A. Kendall, *Polymeric sequences reveal a functional interrelationship between hydrophobicity and length of signal peptides.* J Biol Chem, 1990. **265**(5): p. 2873-80.
324. Nilsson, I., P. Whitley, and G. von Heijne, *The COOH-terminal ends of internal signal and signal-anchor sequences are positioned differently in the ER translocase.* J Cell Biol, 1994. **126**(5): p. 1127-32.
325. Tyanova, S., et al., *The Perseus computational platform for comprehensive analysis of (prote)omics data.* Nat Meth, 2016. **13**(9): p. 731-740.
326. Tanzer, M.C., et al., *Quantitative and Dynamic Catalogs of Proteins Released during Apoptotic and Necroptotic Cell Death.* Cell Rep, 2020. **30**(4): p. 1260-1270.e5.
327. Brunner, J.S., et al., *Environmental arginine controls multinuclear giant cell metabolism and formation.* Nature Communications, 2020. **11**(1): p. 431.
328. Holding, A.N., *XL-MS: Protein cross-linking coupled with mass spectrometry.* Methods, 2015. **89**: p. 54-63.
329. Liu, F., et al., *Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry.* Nature Methods, 2015. **12**(12): p. 1179-1184.
330. Bache, N., et al., *A novel LC system embeds analytes in pre-formed gradients for rapid, ultra-robust proteomics.* Molecular & Cellular Proteomics, 2018: p. mcp.TIR118.000853.
331. Hogrebe, A., et al., *Benchmarking common quantification strategies for large-scale phosphoproteomics.* Nature Communications, 2018. **9**(1): p. 1045.

7 Acknowledgement

I would like to thank those who supported me and contributed to this work:

Matthias Mann for patience, advice, support, and inspiration. Great lab retreats, and happy hours.

Felix Meissner for mentoring, freedom to experiment, patience, and support.

Meera Phulphagar, Franziska Bruening and Maria Tanzer for friendship and support, and inspiration.

Sen Office (Daniel Hornburg, Jan Rieckmann, Jonathan Swietlik, Stefan Ebner, Jingyuan Cheng) for friendship, fun work atmosphere, support and great discussions.

Korbinian Mayr, Igor Paron, Mario Oroshi and Gaby Sowa for mass spec and technical support.

Alison Dalfovo, Theresa Schneider and Ute Sackers for administrative support.

Mann Lab for a great atmosphere with many inspiring characters.

IMPRS coordination office for support and organization of many PhD retreats, seminars, and BBQs.

My family for unconditional support.

8 Addendum

This article was published in the Trillium Immunologie Heft 1/2020 Kardiovaskuläre Immunologie.

Makrophagen: Am Anfang war die Fresszelle

Annika Frauenstein & Felix Meissner

Makrophagen sind Zellen mit außerordentlich vielfältigen Funktionen im Organismus. Neben hochspezialisierten Aufgaben, wie der Immunantwort gegenüber Pathogenen, erfüllen sie auch generische Aufgaben bei der Gewebeentwicklung und -homöostase. In diesem Artikel stellen wir aktuelle Trends in der Makrophagenforschung vor. Wir diskutieren, wie das phänotypische Spektrum von Gewebemakrophagen durch den ontogenetischen Ursprung und die Integration räumlicher und zeitlicher Impulse bestimmt wird, wie akzessorische Makrophagenfunktionen zu Pathologien wie Krebs und Atherosklerose beitragen, und wie innovative hochauflösende Technologien zur Aufklärung der komplexen Makrophagenbiologie beitragen können.

Schlüsselwörter: Makrophage, Fresszelle

Es war einmal eine Fresszelle

Der Vater der zellulären Immunologie, Ilya Metchnikoff, entdeckte Makrophagen im späten neunzehnten Jahrhundert, und nannte sie aufgrund ihrer phagozytotischen Natur „Riesenfresszelle“ (griechisch: μακρός groß, φαγεῖν essen, Abb. 1).

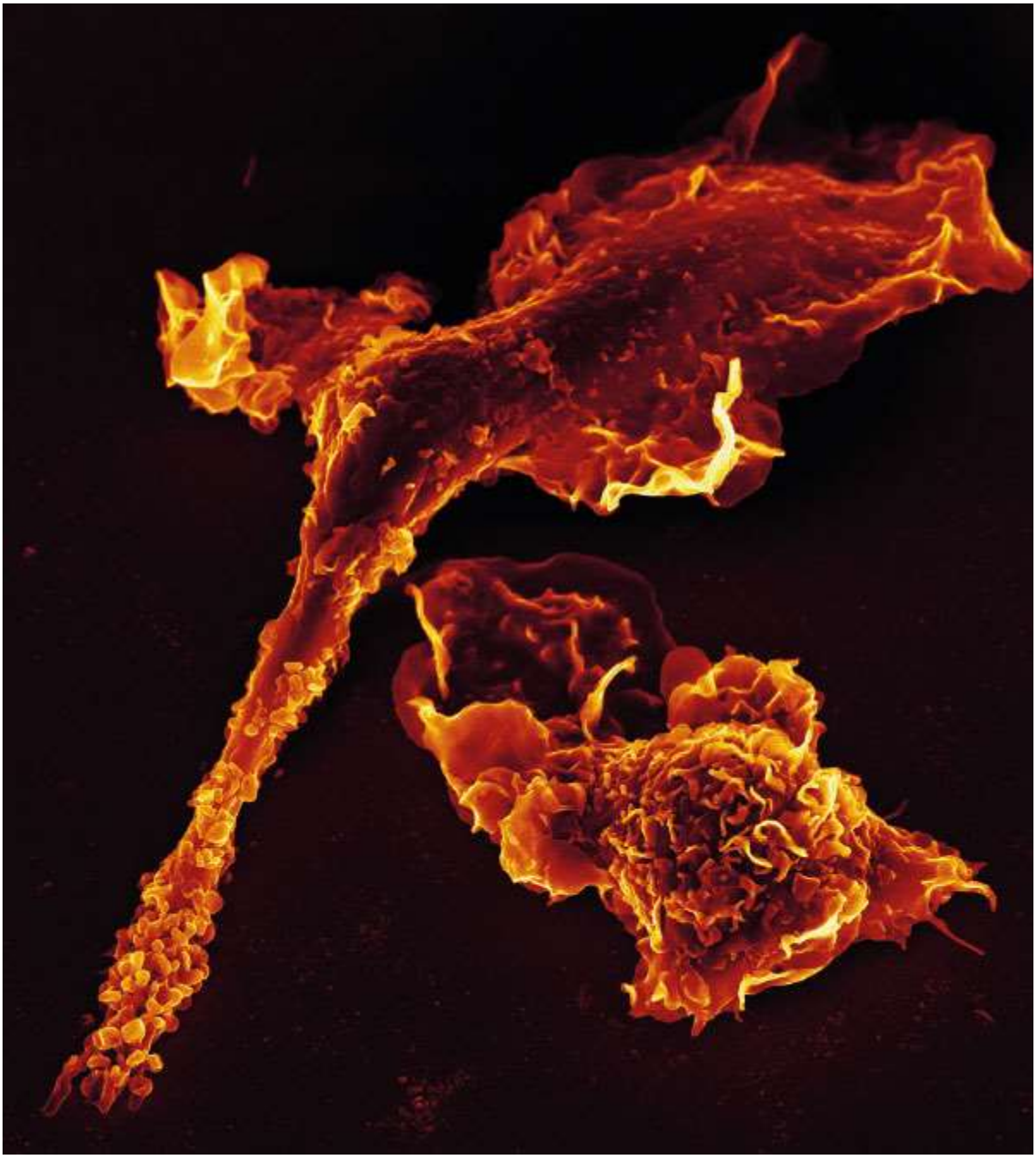


Abb. 1: Rasterelektronenmikroskopische Aufnahme von Makrophagen (von Volker Brinkmann).

Mittels Intravitalmikroskopie gelang es ihm, Makrophagenschwärme in Seesternlarven und infizierten Wasserflöhen zu visualisieren. Er antizipierte bereits ihre Bedeutung für das Immunsystem als wichtige Immuneffektoren –zuständig für Pathogenbeseitigung und Wundheilung [1]. Seitdem wurden viele weitere Funktionen von Makrophagen beobachtet, wie z. B. das Entfernen apoptotischer Zellen und die Remodellierung der extrazellulären Matrix. In den letzten Jahren sind nun überraschende neue Aspekte über die Biologie und Funktion von Makrophagen im systemischen Stoffwechsel, der Gewebehomöostase und -entwicklung entdeckt worden [2, 3].

In diesem Artikel schildern wir die wandelnde Sicht auf die Funktion von Makrophagen im historischen Kontext. Wir diskutieren spannende aktuelle Trends in der Makrophagenforschung, und wie neue und hochauflösende Techniken zu deren Aufklärung beitragen können.

Wachposten im Gewebe

Makrophagen wird schon sehr lange eine wichtige Rolle bei der Gewebeüberwachung nachgesagt. Für die Detektion sowohl von mikrobiellen Produkten (MAMPs – Mikroben-assoziierte molekulare Muster) als auch von „Gefahrensignalen“ des Wirts, die auf Veränderungen der Homöostase hindeuten (DAMPs – Gefahr-assoziierte molekulare Strukturen), sind Makrophagen mit keimbahnkodierten Mustererkennungsrezeptoren (PRRs) ausgestattet. Janeway schlug vor, dass diese PRRs für die Initiierung der adaptiven Immunität wichtig sind [4]. Wegweisende Arbeiten von Lemaître und Hoffmann führten dann zur Identifizierung der ersten PRRs in Fliegen, die ohne funktionelles Toll-Protein anfällig für Pilzinfektionen waren [5]. Seitdem wurden viele oberflächliche und zytosolische PRRs in Säugetierzellen identifiziert, die unterschiedlichste molekulare Strukturen und homöostatische Veränderungen erkennen [6]. Die Entdeckung eines makromolekularen Komplexes von Tschopp, „Inflammasom“ getauft, der entzündliche Immunpathologie auslösen kann, leitete später eine Renaissance der Forschungsgebiete der angeborenen Immunität und des Zelltods ein [7, 8]. Insbesondere der PRR Nlrp3 stellte sich als Schlüsselsignalweg für die Detektion steriler Entzündung und metabolischen Stresses heraus. Nun wird dieser Signalweg nicht nur mit vielen chronisch entzündlichen und metabolischen Erkrankungen in Zusammenhang gebracht [9], sondern hat auch das Interesse an der Entwicklung von Therapeutika zur Behandlung diverser entzündlicher Erkrankungen geweckt [10]. Neben PRRs nutzen Makrophagen eine Vielzahl anderer Rezeptoren und Signalwege, um Veränderungen von z. B. Nährstoff-, Metabolit-, Sauerstoffbedingungen und der Gewebemikroumgebung zu detektieren [11–13]. Die Komplexität des sensorischen Systems von Makrophagen ist bis jetzt noch nicht umfassend aufgeklärt und ein spannendes Gebiet aktiver Forschung.

Makrophagen-Polarisation

Makrophagen produzieren sowohl antimikrobielle reaktive Sauerstoff- und Stickstoffspezies, um Krankheitserreger zu bekämpfen, als auch interzelluläre Botenstoffe, um angeborene und adaptive Immunantworten, bzw. den Umbau und die Reparatur von Gewebe zu koordinieren [14]. Ein dichotomes Konzept der Makrophagenaktivierung hat viele Jahre geholfen, gegensätzliche Funktionen dieser Zellen zu beschreiben. Während „klassisch aktivierte” M1-Makrophagen durch PRR und Interferon induziert werden, um eine Immunantwort gegenüber Bakterien und intrazellulären Pathogenen auszulösen, werden „alternativ aktivierte“ M2-Makrophagen vornehmlich bei Asthma, Allergien und Wurmbefall durch Th2-Antworten und Zytokine wie IL-4 oder IL-13 gebildet [15, 16]. Allerdings existieren z. B. auch Makrophagen, die M2-ähnlich polarisiert sind und die einen wichtigen, aber noch nicht genau geklärten Beitrag beim Heilungsprozess leisten und potenziell bei Transplantationen und Operationen relevant sein könnten. Dieses und viele andere Beispiele illustrieren, dass sich Makrophagenfunktionen besser durch komplexere Konzepte erklären lassen, die von einem kontinuierlichen Spektrum der Makrophagenaktivierung ausgehen. Es hat sich gezeigt, dass Makrophagen nicht nur ein außergewöhnlich großes Repertoire an verschiedenen Phänotypen annehmen können, sondern

auf unterschiedliche zeitliche und räumliche Signale dynamisch ihre physiologischen Funktionen wechseln [17]. Aktuelle Modelle beschreiben, wie durch ein Zusammenspiel von epigenetischen Zell-intrinsischen und Aktivierungs-abhängigen Netzwerken von Transkriptionsfaktoren spezifische Phänotypen reguliert werden [18, 19]. Unvollständig verstanden ist, wie unterschiedliche Signalkaskaden die Vielzahl von Eingangssignalen integrieren und den Rahmen funktionaler Plastizität vorgeben. In diesem Zusammenhang ist spannend, inwiefern sich die regulatorischen Mechanismen unterscheiden, die zur Etablierung von „angeborenem Gedächtnis“ im Gegensatz zu „trainierter Immunität“ führen – zwei paradigmatische Konsequenzen systemischer Entzündung, die sich durch Unfähigkeit versus besondere Fähigkeit zur erneuten zellulären Aktivierung auszeichnen [20, 21].

Makrophagen-Ontogenese

Ähnlich wie bei der Aktivierung von Makrophagen hat sich auch unser Wissen über die Ontogenese rasch weiterentwickelt. Makrophagen sind in fast allen Geweben vorhanden und nicht nur für die Homöostase, sondern auch für die Entwicklung essenziell [22]. Entgegen des initialen Modells von Van Furth, gemäß dessen Gewebe-residente Makrophagen nur aus Blutmonozyten rekrutiert werden (mononukleäres phagozytotisches Konzept) [23], wissen wir nun, dass diese aus drei unterschiedlichen Quellen stammen: nämlich aus Dottersack, fetaler Leber und hämatopoetischen Stammzellen (HSC) im Knochenmark.
Fate Mapping und Pulsmarkierungsstudien von Vorläuferzellen haben wesentlich dazu beigetragen, eine genetische Basis für die Makrophagenentwicklung zu etablieren. So wurde eine frühe hämatogene Welle im Dottersack identifiziert, die zumindest teilweise unabhängig vom Transkriptionsfaktor Myb ist und auch erythromyeloide Vorläuferzellen (EMP) generiert [24, 25]. Diese Zellen führen entweder direkt zu primitiven Dottersack-Makrophagen oder wandern in die fetale Leber und erzeugen fetale Monozyten, die das embryonale Gewebe besiedeln und dort zu Makrophagen differenzieren [26]. Alternativ können Makrophagen durch eine Myb-abhängige Hämatopoese entstehen, die in der fetalen Leber beginnt, hämatopoetische Stammzellen bildet und im Erwachsenenalter fortbesteht [27].
Interessant ist, dass der Ursprung von Makrophagen im adulten Gewebe stark variiert. Während einige Gewebe hauptsächlich von aus EMP stammenden Makrophagen besiedelt werden (wie z. B. Langerhans-Zellen in der Haut und Mikroglia im Gehirn [28, 29]), werden andere Gewebe hauptsächlich mit aus HSC stammenden Makrophagen bevölkert (wie z. B. im Herz, Haut und Darm [30–32]). Die Infiltration durch von Monozyten abstammende Makrophagen könnte daher auf einen höheren „Verschleiß“ eines Organs hinweisen. Kurioserweise können sich in der gleichen Gewebeumgebung die Funktionen von aus Monozyten abstammenden und Gewebe-residenten Makrophagen stark unterscheiden. Daher kann der Makrophagenursprung für die Pathogenese von Krankheiten eine wichtige Rolle spielen, wie z. B. während experimenteller Autoimmunenzephalitis [33]. Da sich verfügbare Studien auf Mausmodelle beschränken, bleibt es jedoch unklar, ob Gewebemakrophagen bei Erwachsenen tatsächlich nicht durch von adulten Monozyten stammende Makrophagen ersetzt werden, insbesondere während des Alterns oder über längere Zeiträume wiederholter Belastungen.
Eine spannende Beobachtung ist, dass aus Dottersäcken hervorgehende Makrophagen (und vermutlich fetale und Leber-abgeleitete Makrophagen) durch Selbsterneuerung über die gesamte Lebensdauer aufrechterhalten werden können [34]. Jedoch bleibt bis jetzt ungeklärt, ob alle oder

einige wenige gewebsresidente Makrophagen in den Zellzyklus eintreten können, und sich selbst durch asymmetrische Zellteilung (ähnlich wie Stammzellen) erneuern.

Funktionsspektrum von Makrophagen

Trotz unterschiedlicher Ontogenität können Makrophagen im Gewebe ähnliche Funktionen haben. Alveolarmakrophagen aus der fetalen Leber und Mikroglia aus dem Dottersack üben z. B. während der Homöostase immunsuppressive Funktionen aus und regulieren durch die Deaktivierung autoreaktiver T-Zellen die Aktivierungsschwelle für T-Zell-vermittelte Immunreaktionen [35]. Lokal produzierte sogenannte „Gewebeidentitätssignale“ erklären diese Makrophagenfunktionen in Abhängigkeit von Stoffwechsel, Nährstoff- und Mikrobiotaexposition. So wird z. B. die Differenzierung von Mikroglia und Peritonealmakrophagen als Reaktion auf lokale Produktion von TGF-β im Gehirn bzw. Retinsäure im Omentum gesteuert [36, 37]. Interessant ist, dass sich Makrophagen an eine sich sehr stark ändernde lokale Homöostase im Laufe des Lebens anpassen können, wie z. B. während der Entwicklung des erwachsenen Gehirns oder des Darms vor und nach der perinatalen Besiedlung durch Mikroben [38, 39].
Makrophagen als akzessorische (Zubehörs)-Zellen zu betrachten, kann konzeptionell zum Verständnis der besonderen Plastizität von Phänotyp und Funktion in Geweben beitragen [11]. Durch zelluläre „Arbeitsteilung“ werden generische Aufgaben, wie die Entfernung apoptotischer Zellen, hochspezialisierte Aufgaben (wie Knochenresorption durch Knochenmakrophagen/Osteoklasten, Recycling von Tensiden durch Lungenalveolarmakrophagen oder Bereitstellung von neurotrophen Faktoren durch Mikroglia [40–42]), aber auch „auf Abruf“-Funktionen, wie immunologische und Reparaturprozesse an Gewebemakrophagen, abdelegiert. Dementsprechend können Makrophagen auch als akzessorische Zellen z. B. für Tumoren verstanden werden, die in diesem Fall tumorfördernde Prozesse unterstützen [43, 44]. Neben Krebs gilt eine entscheidende Rolle von Makrophagen in diversen Pathologien wie Atherosklerose [45], Osteoporose, Fettleibigkeit und Typ-2-Diabetes [46, 47] und Fibrose [48] als erwiesen und macht bestimmte Populationen von Makrophagen mit spezifischen Funktionen therapeutisch interessant.

Hochauflösende methodische Ansätze zur Analyse von Makrophagen

Es besteht ein wachsender Bedarf an der Entwicklung von Konzepten, die den Gesamtphänotyp von Makrophagen als Hierarchie aus Abstammung, Identitäts- und „auf Abruf“-Funktionsmodulen beschreiben (Abb. 2).

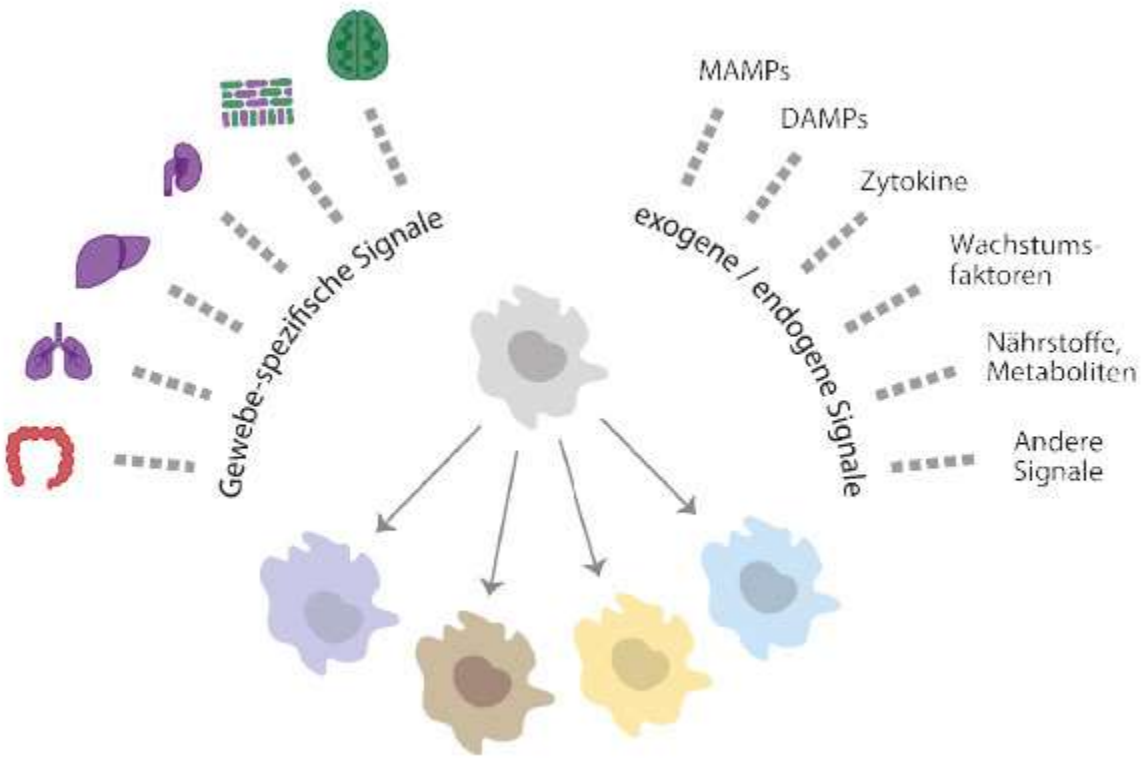


Abb. 2: MultidimensionaleMakrophagenaktivierungDie Makrophagenphysiologie wird dynamisch von Ursprung, Gewebe und sowohl endogenen als auch exogenen Signalen bestimmt. Der überwiegende Makrophagenursprung der individuellen Gewebe ist farblich markiert: Primitive Dottersack-Makrophagen grün, Makrophagen aus fötalen Monozyten lila und Makrophagen aus adulten Monozyten rot.

Um das Zusammenspiel vielzahliger Faktoren zur Differenzierung, Erhaltung und Aktivierung von Gewebe-spezifischen Makrophagenpopulationen zu verstehen, sind innovative hochauflösende Techniken besonders gut geeignet.
Zur Erforschung zellulärer Heterogenität wird die zukünftige Forschung von der Verwendung von genetischen Modellen profitieren, die präziseres genetisches und zeitliches Fate Mapping von unterschiedlichen Vorläuferpopulationen erlauben. In diesem Zusammenhang sind hochdimensionale Zelltechnologien wie mikrofluidische Genom-, Epigenom- und Transkriptomanalysen [49] attraktiv, da diese – im Gegensatz zu klassischen Methoden (wie der Durchflusszytometrie) – für jede Zelle eine große Anzahl von Genen analysieren, und eine unvoreingenommene Bestimmung von Zellpopulationen möglich machen. Methoden, die eine genaue Bestimmung des Immunphänotyps ermöglichen, so wie CyTOF, aber auch Durchflusszytometrie oder Immunhistochemie, werden zur Evaluation funktioneller Aspekte und Effektormolekülkombinationen in heterogenen Makrophagenpopulationen beitragen [50, 51].
Die Orchestrierung der Gewebehomöostase und entzündlicher Prozesse ist aufgrund der komplexen Verwendung vieler pleiotroper interzellulärer Botenstoffe noch unvollständig verstanden. Auf Massenspektrometrie basierende analytische Verfahren können hier einen wesentlichen Beitrag leisten, unterschiedliche Biomolekülklassen, wie z. B. Proteine, Metaboliten und Lipide, umfassend zu identifizieren. In Kombination mit funktionellen

pharmakologischen und genetischen Strategien können so gegenseitige phänotypische Abhängigkeiten von Makrophagen und Gewebe, wie z. B. „Gewebeidentitätssignale“ die Makrophagenphänotypen diktieren, bzw. Gewebefunktionen durch die Makrophagenphänotypen diktiert werden, bestimmt werden [52]. Ähnlich wie Einzelzellsequenzierungsmethoden die Gewebeheterogenität auf Nukleinsäureebene auflösen, werden die zu erwartenden Entwicklungen Massenspektrometrie-basierter Technologien in Zukunft umfassende Analysen vieler physiologisch relevanter Biomolekülklassen in seltenen Primärzelltypen oder sogar Einzelzellen ermöglichen [53, 54].

Literatur

1. Cavaillon, J.M., The historical milestones in the understanding of leukocyte biology initiated by Elie Metchnikoff. J Leukoc Biol, 2011. 90(3): p. 413-24.
2. Wynn, T.A., A. Chawla, and J.W. Pollard, Macrophage biology in development, homeostasis and disease. Nature, 2013. 496(7446): p. 445-55.
3. Gordon, S., A. Pluddemann, and F. Martinez Estrada, Macrophage heterogeneity in tissues: phenotypic diversity and functions. Immunol Rev, 2014. 262(1): p. 36-55.
4. Janeway, C.A., Jr., Approaching the asymptote? Evolution and revolution in immunology. Cold Spring Harb Symp Quant Biol, 1989. 54 Pt 1: p. 1-13.
5. Lemaitre, B., et al., The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults. Cell, 1996. 86(6): p. 973-83.
6. Takeuchi, O. and S. Akira, Pattern recognition receptors and inflammation. Cell, 2010. 140(6): p. 805-20.
7. Broz, P. and V.M. Dixit, Inflammasomes: mechanism of assembly, regulation and signalling. Nat Rev Immunol, 2016. 16(7): p. 407-20.
8. Martinon, F., K. Burns, and J. Tschopp, The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. Mol Cell, 2002. 10(2): p. 417-26.
9. Robbins, G.R., H. Wen, and J.P. Ting, Inflammasomes and metabolic disorders: old genes in modern diseases. Mol Cell, 2014. 54(2): p. 297-308.
10. Mangan, M.S.J., et al., Targeting the NLRP3 inflammasome in inflammatory diseases. Nat Rev Drug Discov, 2018. 17(8): p. 588-606.
11. Okabe, Y. and R. Medzhitov, Tissue biology perspective on macrophages. Nat Immunol, 2016. 17(1): p. 9-17.
12. El Kasmi, K.C. and K.R. Stenmark, Contribution of metabolic reprogramming to macrophage plasticity and function. Seminars in Immunology, 2015. 27(4): p. 267-275.
13. O'Neill, L.A. and E.J. Pearce, Immunometabolism governs dendritic cell and macrophage function. J Exp Med, 2016. 213(1): p. 15-23.
14. Mantovani, A., et al., The chemokine system in diverse forms of macrophage activation and polarization. Trends Immunol, 2004. 25(12): p. 677-86.
15. Martinez, F.O. and S. Gordon, The M1 and M2 paradigm of macrophage activation: time for reassessment. F1000prime reports, 2014. 6: p. 13-13.
16. Murray, P.J., et al., Macrophage activation and polarization: nomenclature and experimental guidelines. Immunity, 2014. 41(1): p. 14-20.
17. Ginhoux, F., et al., New insights into the multidimensional concept of macrophage ontogeny, activation and function. Nature Immunology, 2016. 17(1): p. 34-40.
18. Gosselin, D., et al., Environment drives selection and function of enhancers controlling

tissue-specific macrophage identities. Cell, 2014. 159(6): p. 1327-40.
19. Ostuni, R., et al., Latent enhancers activated by stimulation in differentiated cells. Cell, 2013. 152(1-2): p. 157-71.
20. Smale, S.T., A. Tarakhovsky, and G. Natoli, Chromatin contributions to the regulation of innate immunity. Annu Rev Immunol, 2014. 32: p. 489-511.
21. Netea, M.G., et al., Trained immunity: A program of innate immune memory in health and disease. Science, 2016. 352(6284): p. aaf1098.
22. Epelman, S., K.J. Lavine, and G.J. Randolph, Origin and functions of tissue macrophages. Immunity, 2014. 41(1): p. 21-35.
23. van Furth, R. and Z.A. Cohn, The origin and kinetics of mononuclear phagocytes. J Exp Med, 1968. 128(3): p. 415-35.
24. Gomez Perdiguero, E., et al., Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. Nature, 2015. 518(7540): p. 547-51.
25. Schulz, C., et al., A lineage of myeloid cells independent of Myb and hematopoietic stem cells. Science, 2012. 336(6077): p. 86-90.
26. Hoeffel, G., et al., C-Myb(+) erythro-myeloid progenitor-derived fetal monocytes give rise to adult tissue-resident macrophages. Immunity, 2015. 42(4): p. 665-78.
27. Orkin, S.H. and L.I. Zon, Hematopoiesis: an evolving paradigm for stem cell biology. Cell, 2008. 132(4): p. 631-44.
28. Ginhoux, F. and S. Jung, Monocytes and macrophages: developmental pathways and tissue homeostasis. Nat Rev Immunol, 2014. 14(6): p. 392-404.
29. Merad, M., et al., Langerhans cells renew in the skin throughout life under steady-state conditions. Nat Immunol, 2002. 3(12): p. 1135-41.
30. Molawi, K., et al., Progressive replacement of embryo-derived cardiac macrophages with age. J Exp Med, 2014. 211(11): p. 2151-8.
31. Varol, C., A. Mildner, and S. Jung, Macrophages: development and tissue specialization. Annu Rev Immunol, 2015. 33: p. 643-75.
32. Bain, C.C., et al., Constant replenishment from circulating monocytes maintains the macrophage pool in the intestine of adult mice. Nat Immunol, 2014. 15(10): p. 929-937.
33. Ajami, B., et al., Infiltrating monocytes trigger EAE progression, but do not contribute to the resident microglia pool. Nat Neurosci, 2011. 14(9): p. 1142-9.
34. Sieweke, M.H. and J.E. Allen, Beyond stem cells: self-renewal of differentiated macrophages. Science, 2013. 342(6161): p. 1242974.
35. Holt, P.G., Alveolar macrophages. III. Studies on the mechanisms of inhibition of T-cell proliferation. Immunology, 1979. 37(2): p. 437-45.
36. Butovsky, O., et al., Identification of a unique TGF-beta-dependent molecular and functional signature in microglia. Nat Neurosci, 2014. 17(1): p. 131-43.
37. Okabe, Y. and R. Medzhitov, Tissue-specific signals control reversible program of localization and functional polarization of macrophages. Cell, 2014. 157(4): p. 832-44.
38. Ueno, M. and T. Yamashita, Bidirectional tuning of microglia in the developing brain: from neurogenesis to neural circuit formation. Curr Opin Neurobiol, 2014. 27: p. 8-15.
39. Fulde, M. and M.W. Hornef, Maturation of the enteric mucosal innate immune system during the postnatal period. Immunol Rev, 2014. 260(1): p. 21-34.
40. Pollard, J.W., Trophic macrophages in development and disease. Nat Rev Immunol, 2009. 9(4): p. 259-70.
41. Boyle, W.J., W.S. Simonet, and D.L. Lacey, Osteoclast differentiation and activation.

Nature, 2003. 423(6937): p. 337-42.
42. Hussell, T. and T.J. Bell, Alveolar macrophages: plasticity in a tissue-specific context. Nat Rev Immunol, 2014. 14(2): p. 81-93.
43. Vitale, I., et al., Macrophages and Metabolism in the Tumor Microenvironment. Cell Metab, 2019. 30(1): p. 36-50.
44. Noy, R. and J.W. Pollard, Tumor-associated macrophages: from mechanisms to therapy. Immunity, 2014. 41(1): p. 49-61.
45. Moore, K.J. and I. Tabas, Macrophages in the pathogenesis of atherosclerosis. Cell, 2011. 145(3): p. 341-55.
46. Chawla, A., K.D. Nguyen, and Y.P. Goh, Macrophage-mediated inflammation in metabolic disease. Nat Rev Immunol, 2011. 11(11): p. 738-49.
47. Olefsky, J.M. and C.K. Glass, Macrophages, inflammation, and insulin resistance. Annu Rev Physiol, 2010. 72: p. 219-46.
48. Wynn, T.A. and L. Barron, Macrophages: master regulators of inflammation and fibrosis. Semin Liver Dis, 2010. 30(3): p. 245-57.
49. Klein, A.M., et al., Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell, 2015. 161(5): p. 1187-1201.
50. Becher, B., et al., High-dimensional analysis of the murine myeloid cell system. Nat Immunol, 2014. 15(12): p. 1181-9.
51. Bendall, S.C., et al., Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science, 2011. 332(6030): p. 687-96.
52. Rieckmann, J.C., et al., Social network architecture of human immune cells unveiled by quantitative proteomics. Nat Immunol, 2017. 18(5): p. 583-593.
53. Zhang, L. and A. Vertes, Single-Cell Mass Spectrometry Approaches to Explore Cellular Heterogeneity. Angew Chem Int Ed Engl, 2018. 57(17): p. 4466-4477.
54. Marx, V., A dream of single-cell proteomics. Nat Methods, 2019. 16(9): p. 809-812.