



OPEN ACCESS

EDITED BY
Magdalena Plebanski,
RMIT University, Australia

REVIEWED BY
Zissis C. Chroneos,
The Pennsylvania State University,
United States
Hester Doyle,
Yale University, United States

*CORRESPONDENCE
Tue Bjerg Bennike
✉ tbe@hst.aau.dk

RECEIVED 07 July 2023

ACCEPTED 26 September 2023

PUBLISHED 06 October 2023

CITATION

Bennike TB (2023) Advances in proteomics: characterization of the innate immune system after birth and during inflammation. *Front. Immunol.* 14:1254948. doi: 10.3389/fimmu.2023.1254948

COPYRIGHT

© 2023 Bennike. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advances in proteomics: characterization of the innate immune system after birth and during inflammation

Tue Bjerg Bennike*

Medical Microbiology and Immunology, Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

Proteomics is the characterization of the protein composition, the proteome, of a biological sample. It involves the large-scale identification and quantification of proteins, peptides, and post-translational modifications. This review focuses on recent developments in mass spectrometry-based proteomics and provides an overview of available methods for sample preparation to study the innate immune system. Recent advancements in the proteomics workflows, including sample preparation, have significantly improved the sensitivity and proteome coverage of biological samples including the technically difficult blood plasma. Proteomics is often applied in immunology and has been used to characterize the levels of innate immune system components after perturbations such as birth or during chronic inflammatory diseases like rheumatoid arthritis (RA) and inflammatory bowel disease (IBD). In cancers, the tumor microenvironment may generate chronic inflammation and release cytokines to the circulation. In these situations, the innate immune system undergoes profound and long-lasting changes, the large-scale characterization of which may increase our biological understanding and help identify components with translational potential for guiding diagnosis and treatment decisions. With the ongoing technical development, proteomics will likely continue to provide increasing insights into complex biological processes and their implications for health and disease. Integrating proteomics with other omics data and utilizing multi-omics approaches have been demonstrated to give additional valuable insights into biological systems.

KEYWORDS

plasma, mass spectrometry, complement system, neutrophils, neutrophil extracellular traps, ontogeny, rheumatoid arthritis, inflammatory bowel disease

1 Introduction

Proteins are products of our transcribed and translated genome and are the effector molecules for the majority of cellular functions. Proteomics is the large-scale identification and relative quantification of proteins, peptides, and post-translational modifications

(PTMs), and can cover the study of all expressed proteins by an organism, i.e. the proteome. Liquid chromatography mass spectrometry (LC-MS/MS) remains the most widely used core technology in proteomics for large-scale characterization of the proteome. The versatile method can analyze a wide variety of sample types and provide comprehensive characterization of the proteome. In the context of this review, "proteomics" refers to identification and relative quantification of peptides and proteins by LC-MS/MS.

The proteomics workflows have undergone continuous advancements, especially in terms of sample preparation and speed and sensitivity of the platforms largely driven by the technical development and novel MS data acquisition and sample preparation methods. In the early 2010s, the typical sample preparation prior to analysis took one to two days. Proteins in the thousands could be measured in tissue samples in the course of 3-4 hours of LC-MS/MS analysis time (1-3). Today, parallel sample preparation in 96-well plates enable rapid processing of hundreds of samples in a matter of hours rather than days supporting large scale studies (4, 5). The analysis of more than 1,000 plasma samples have made it possible to study the effect of weight loss and maintenance, and associate it with a reduction of biological pathways related to inflammation and activation of lipid metabolism (6, 7). In tissue samples, over 10,000 proteins can be identified in less than two hours (8, 9). As an example, the methodology has been applied to identified altered protein levels during testis cancer and link downregulation of the complement system, likely to evade the innate immune response, with the pathogenesis of the disease (8). Identification of cancer-specific biomarkers, e.g. proteins which leak from the tumor microenvironment to the blood, holds great potential for noninvasive diagnostics.

Advancements in proteomics have found broad applications, amongst others in the field of immunology to study the innate immune system humoral and cellular components in health and disease. Traditionally, 'humoral' described antibody-mediated adaptive immunity. However, the term has now been expanded to include the soluble components of both the innate and adaptive immune system (10). This review highlights recent developments in proteomics sample preparation protocols and MS methods and provides examples of recent discoveries within the innate immune system in health and disease driven by proteomics. The aim is to provide the reader with an overview of currently available methods and proteomics capabilities to study the innate immune system.

Section 2 provides a brief overview of humoral components of the innate immune system in the blood and neutrophil granulocytes, including neutrophil extracellular traps (NETs) and the formation by NETosis. The levels of humoral components change during early development, infections, and during inflammatory diseases, and may provide a perspective on ongoing innate immune system processes and identify biological targets with translational potential for guiding diagnosis and treatment decisions.

Section 3 describes the main technical challenges in plasma proteomics, and the principle of the LC-MS/MS proteomics workflow. The section emphasizes advancements in workflows aimed at increasing the coverage of the plasma proteome to

amongst others study the humoral innate immune system components.

Section 4 highlights recent applications of plasma proteomics to study the developmental trajectory in levels of the humoral innate immune system components during the critical first week after birth, which plays a vital role in newborn immunity (11).

Finally, section 5 discusses the role of NETs in inflammatory diseases and its potential as a therapeutic target. NETs are generated by the most abundant innate immune cell, the neutrophils, and are increasingly recognized as a major contributor to the onset and persistence of prevalent inflammatory diseases (12, 13), including rheumatoid arthritis (RA) and inflammatory bowel disease (IBD) (3), and have been associated with cancers (14).

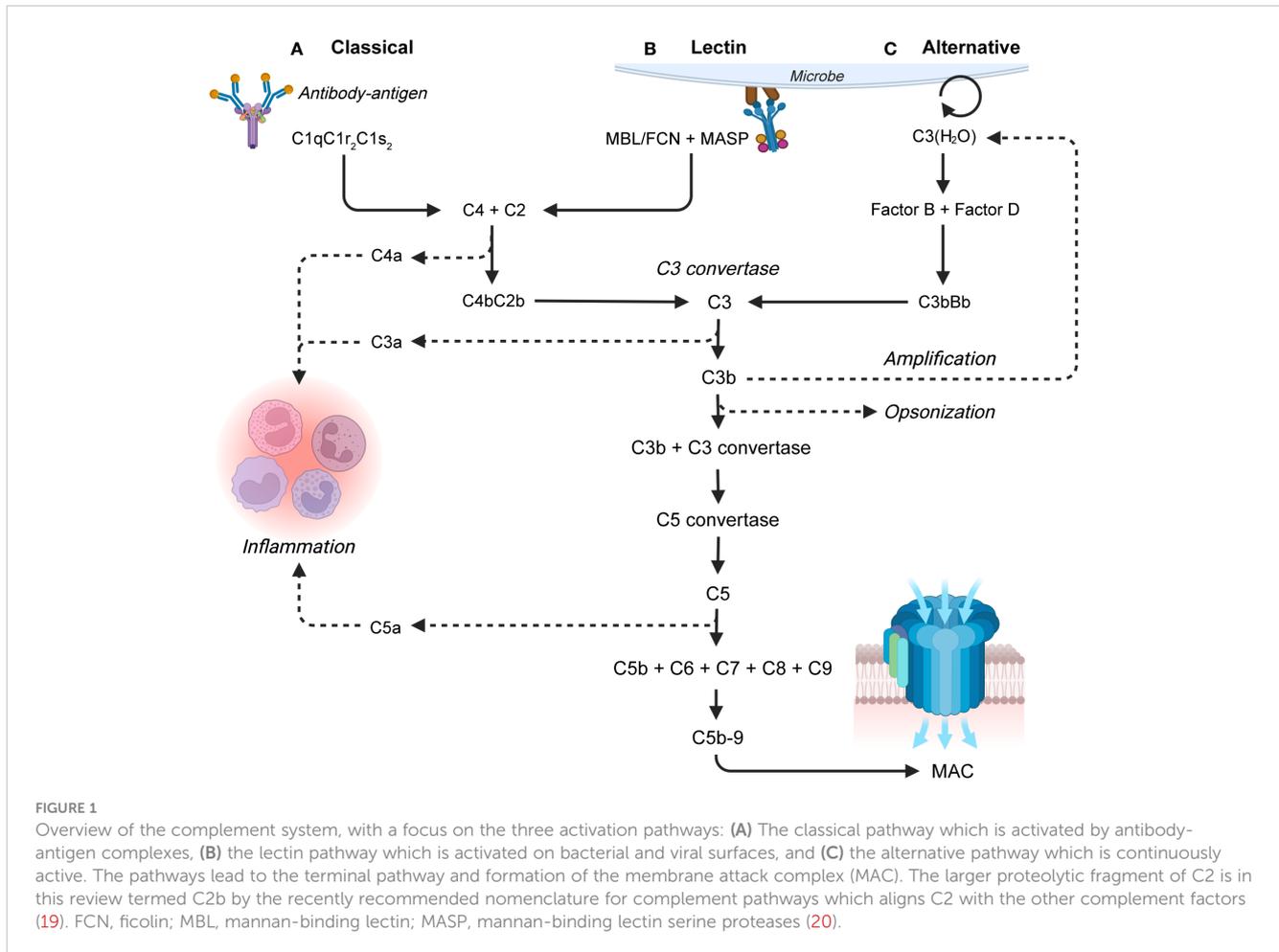
2 A brief overview of the innate immune system

The innate immune system comprises cellular components (including neutrophils and monocytes/macrophages) and humoral components (including the complement system and acute phase proteins). It is the first line of defense against infections. Breach of the physical/chemical barrier or invasion of pathogens will initiate an innate immune response within minutes to hours. The rapid response is facilitated by the persistent availability of the innate immune components in blood, which are continuously synthesized irrespective of any pathogen encounter (15). Consequently, the pattern recognition receptors (PRRs) in the innate immune system have an affinity to conserved pathogen-associated molecular patterns (PAMPs), and re-exposure to a given pathogen results in the same innate immune response. This is in contrast to the adaptive immune response which is modulated upon re-exposure (16, 17). Most of the humoral innate immune system components are protein-based and can be studied by proteomics to characterize the innate immune response in diseases.

2.1 The complement system

The complement system is part of the first innate immune response against invading pathogens and include more than 30 proteins circulating in the plasma, in addition to complement receptors on immune cells including neutrophils and monocytes/macrophages. Similar to the coagulation system in blood, the complement system utilizes serine proteases which are sequentially activated in amplification cascade reactions thereby allowing binding of few PAMPs by PRRs to initiate a vast response (18).

The complement system can be activated by three major pathways. The classical pathway (Figure 1A) can be initiated by binding of complement component (C)1q to antigen-antibody immune complexes, consisting of multiple proximate immunoglobulin (Ig)G or few IgM molecules (21, 22). The lectin pathway (Figure 1B) can be initiated by binding of ficolins (FCN), mannose-binding lectin (MBL), and collectins to carbohydrates or



N-acetyl-glucosamine exposing groups found on the surface of many bacteria and viruses (23, 24). Activation of the classical and lectin pathways lead to cleavage of C4 by the C1-complex (C1qC1r₂C1s₂) or mannan-binding lectin serine proteases (MASP) 1 and 2, respectively. C4 is cleaved into the anaphylatoxin C4a and a larger C4b fragment. C4b binding of C2 enables cleavage into a C2a fragment and a larger proteolytic C2b, which remains bound to C4b and form a C3-convertase (C4bC2b) (25, 26). The alternative pathway (Figure 1C) is continuously active in low levels due to hydrolysis of the C3-thioester group resulting in formation of C3(H₂O) with a "C3b-like" activity. Binding of factor B (27) and cleavage of the complex by factor D form a C3-convertase (C3bBb). The C3bBb-complex can be stabilized by binding to properdin (also known as complement factor P, CFP), the only known complement activation enhancer (28).

All three activation pathways result in cleavage of C3 into the anaphylatoxin C3a and the opsonin C3b. In addition to spontaneous activation, the alternative pathway can amplify complement activation by producing additional alternative C3 convertases using C3b generated by either pathway (28). High local concentrations of C3b initiates the terminal pathway, where C3b associates with a C3 convertase and forms a C5 convertase complex. The C5 convertase initiates the formation and assembly of the membrane attack complex (MAC), by cleaving C5 into the

anaphylatoxin C5a and C5b. C5b recruits C6, C7, and C8, and induces C9 polymerization, thereby generating a C5b-C9 transmembrane MAC pore structure (29). The MAC causes bacterial lysis capable of eliminating certain pathogens (30). In addition to MAC formation and opsonization, complement induces several pro-inflammatory effects many of which are induced by the complement anaphylatoxins (C3a, C4a, and C5a). The effects include chemotaxis, synthesis of reactive oxygen species (ROS) in neutrophils and monocytes/macrophages, and histamine release from mast cells and vasodilation (29, 31).

2.2 Neutrophils and neutrophil extracellular traps

The immunological actions by the complement system works in synergy with a variety of immune cells, including neutrophils and monocytes/macrophages. Neutrophils represent the most abundant circulating immune cells and the first cells recruited from blood. The cells are recruited by cytokines and anaphylatoxins to the site of inflammation where the immunological actions include secretion of inflammatory proteins, ROS, and phagocytosis (32). Stimulation by C5a causes neutrophils to release C3, complement factor B (CFB), CFP, and enhance local complement system activation (33).

Neutrophils may also release NETs in response to large pathogens (34) or specific stimuli including gram-positive and gram-negative bacteria, fungi, lipopolysaccharide (LPS), and IL-8 (35).

NETs were first described in 2004 and are web-like structures with a backbone of decondensed chromatin fibers coated with antimicrobial proteins (36). NETs are important for immunity and are released in a cell process termed NETosis. Initiation results in activation of NADPH-oxidase (NOX2). NOX2 generates intracellular ROS which increases the nucleus membrane permeability, leading to activation and translocation of granular neutrophil elastase to the nucleus. In the nucleus, neutrophil elastase partially degrades specific histones resulting in chromatin decondensation, enhanced by myeloperoxidase (MPO). Increased intracellular calcium levels during NETosis activate the enzyme peptidyl deiminase (PAD) 4. PAD4 localizes to the cell nucleus and catalyzes the deimination of arginine amino acids into citrullines on histone H3 proteins (37). Citrullination reduces the positive charge of proteins (38) and relaxes the chromatin binding by histone H3, thereby causing further chromatin decondensation. The neutrophil nuclear envelope disassembles and nuclear decondensed chromatin mixes with cytoplasmic and granule components (39). Finally, the plasma membrane permeabilizes, and NETs coated with proteases and antimicrobial peptides and proteins expand into the extracellular space (40, 41). Different processes of NETosis exist, where not all result in the immediate death of the neutrophil (42). The NETs trap and neutralize pathogens, and neutrophil elastase has been found responsible for the majority of the proteolytic activity (43). The anti-microbial effect is enhanced by NET-bound CFP, produced by neutrophils, which enables NETs to augment complement activation and induce MAC formation via the alternative pathway (33). In this way, neutrophils and NETs work in synergy with the complement system to eliminate pathogens.

2.3 Acute phase proteins

Several of the pro-inflammatory cytokines and chemokines that are released during infection and inflammation stimulate hepatic synthesis and circulatory secretion of acute phase proteins (APP). APP are defined as proteins where the serum concentration changes by >25% during inflammation (44). Interleukin (IL)-6 is widely viewed as the major inducer of APP synthesis, but APP-synthesis can also be stimulated by e.g. IL-1, tumor necrosis factor alpha (TNF- α), interferon gamma (IFN- γ), and transforming growth factor beta (TGF- β) (45). The positively regulated APP increase during infection and inflammation and most are involved in the innate immune response. These include complement components (C3, C4, C9, and factor B), serum amyloid A (SAA1), C-reactive protein (CRP), and transport proteins including haptoglobin (HP) (44).

SAA1 and CRP are among the APP with the highest increase during infection (46). SAA1 has a wide range of inflammatory functions for priming neutrophils and increasing the production of ROS. CRP binds peptides on bacterial surfaces and can activate complement by C1q binding thereby contributing to inflammation

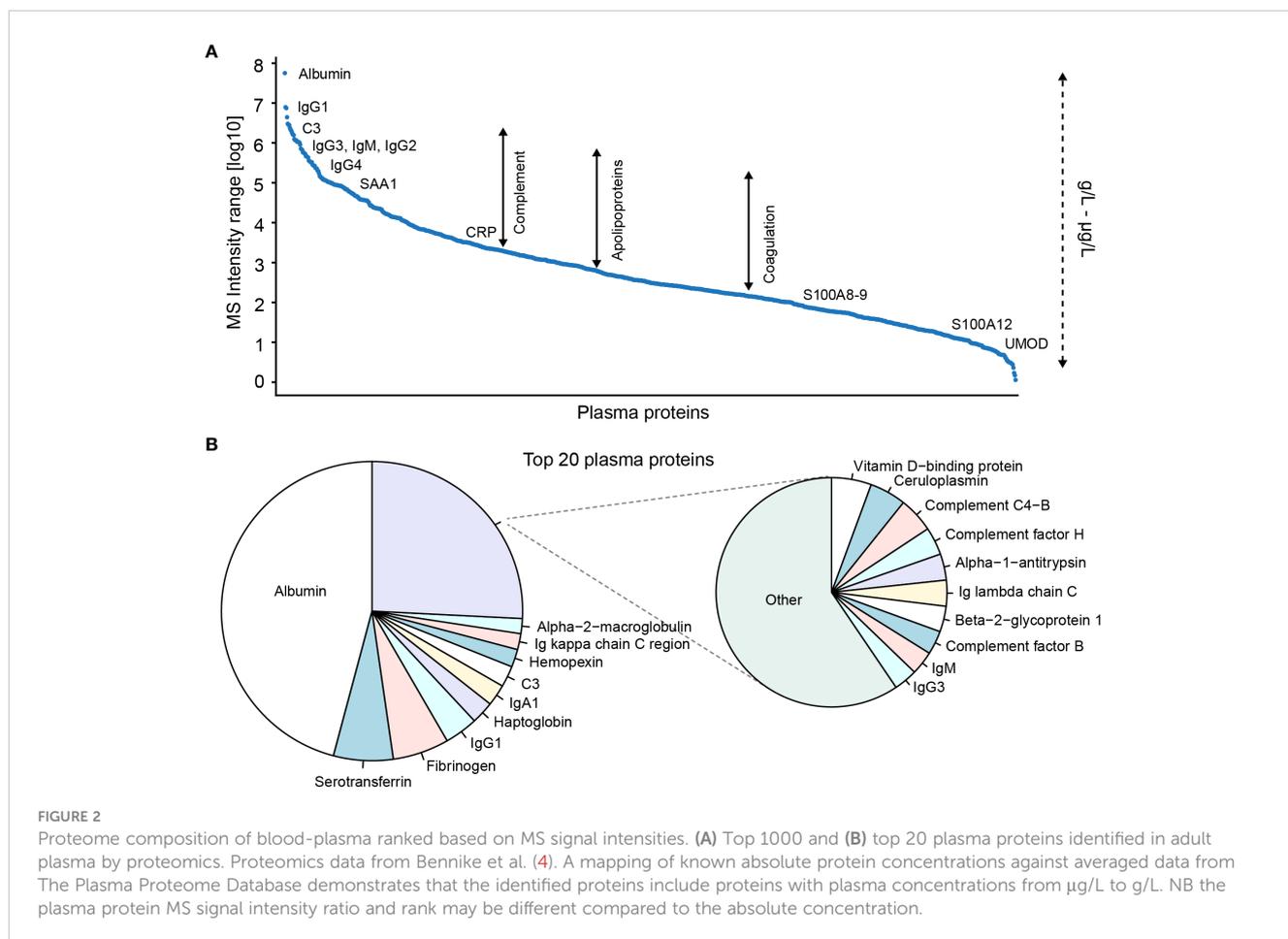
(47, 48). Additionally, CRP on bacterial surfaces can bind to receptors for the Fc portion of IgG and perform opsonization of pathogens (49), and stimulate generation of additional pro-inflammatory cytokines that enhance the inflammatory response (47, 48). CRP blood levels are today widely used as a non-specific but clinically useful marker for infection (50).

HP is a scavenger protein which binds free hemoglobin in plasma which is an oxidizing agent causing tissue damage. In addition, bacteria can obtain iron from the heme group to increase proliferation. HP binding of hemoglobin reduces the available iron for bacterial growth and HP can bind CD11b/CD18 integrins on neutrophils and regulate the activity of the immune system, thereby asserting an anti-bacterial effect. Thus, HP has both antioxidant, antimicrobial, and anti-inflammatory properties (32, 51), and increased plasma levels have been reported during infection and inflammation (52) and in newborns in the days after birth (53).

3 Plasma proteomics to study the humoral innate immune system

Blood-plasma is an easily obtainable material with a high protein concentration of 50-70 g/L (54). Over 5,000 proteins in plasma have been identified by proteomics (55, 56) demonstrating an impressive complexity of the plasma proteome. The proteins include the majority of the innate immune system components and signaling molecules that circulate in the blood (Figure 2A). Monitoring the levels of these components may be used to characterize ongoing innate immune reactions and inflammatory processes in inflammatory diseases. As an example, Bennike et al. have applied plasma proteomics to study the relative levels of the complement system components in the first week after birth (53). In addition, inflammatory processes are linked with cancer (57), and proteomics identification of cancer biomarkers with diagnostic and prognostic value in plasma holds vast potential to guide diagnosis and therapy decisions (57).

Proteomics thereby seems promising as a methodology to identify novel blood-based markers with translational potential for diagnostics and prognostics. However, relatively few biomarkers identified by proteomics have progressed to be used in clinical practice (58). This is likely a consequence of the protein composition of plasma, and plasma-derived body fluids such as synovial fluid, which presents several technical challenges for plasma proteomics, resulting in a lack of sensitive, reproducible, and high-throughput compatible proteomics workflows. The concentration range of the plasma proteins covers a vast dynamic range of 11-12 orders of magnitude (54), by far exceeding the dynamic range of modern mass spectrometers (59). In addition, the distribution of the protein concentrations is highly skewed, with the 22 most abundant proteins, including albumin, Ig's, and C3, constituting 99% of the total protein amount (54). The relatively few highly abundant proteins hinder the proteomics-based detection and quantification of the lower abundant proteins (Figure 2B), effectively reducing the number of detectable and quantifiable plasma proteins (4).



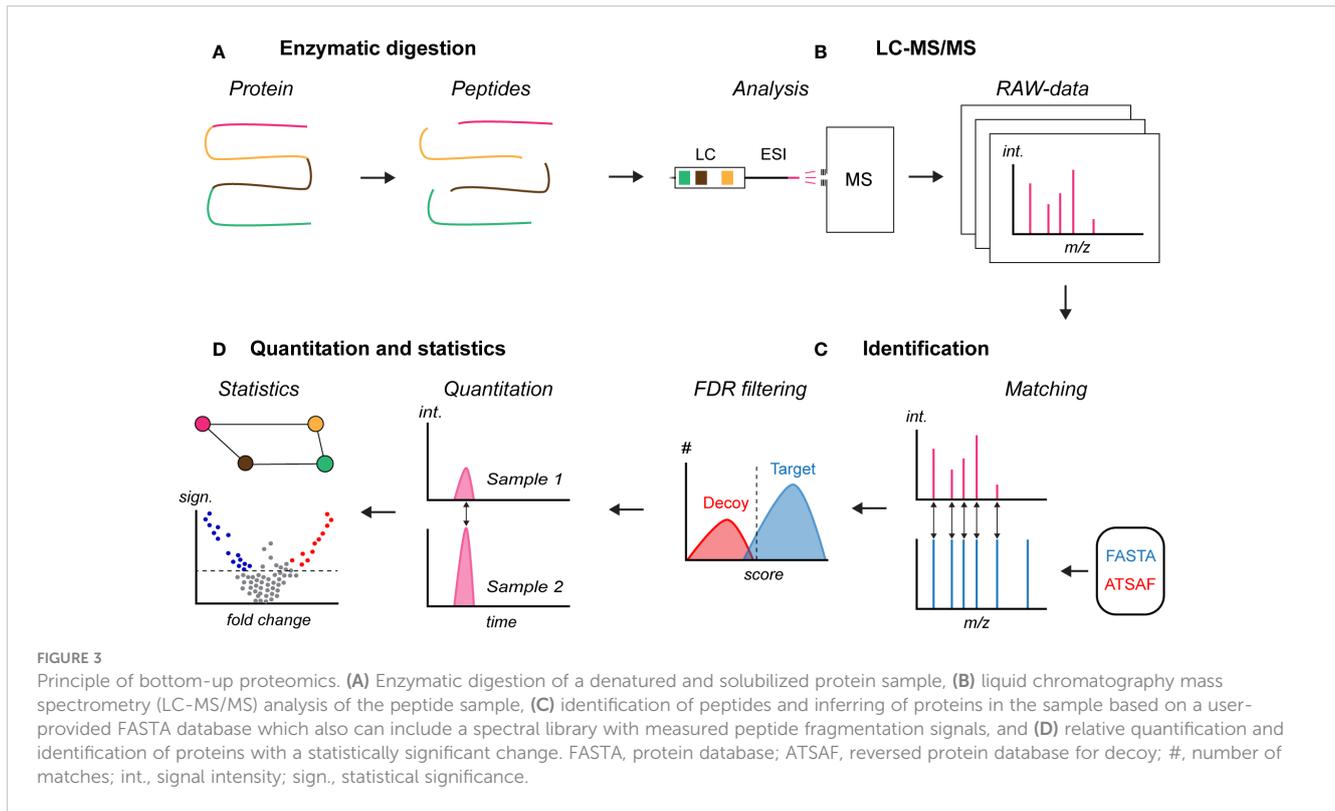
This section discusses the widely used bottom-up proteomics workflow and addresses recent developments with a focus on improving the coverage of the plasma proteome (Figure 3, will be used throughout this section). Bottom-up proteomics refers to the analysis of samples where the proteins have been enzymatically proteolyzed into peptides prior to LC-MS/MS analysis (Figure 3A) (60). While mass spectrometers can measure the mass of intact proteins, analysis on the peptide level has several advantages over "top-down" proteomics, including an increased efficiency in obtaining amino acid sequence information. This is especially true for complex protein mixtures, e.g. from biological samples (61).

3.1 Sample preparation for bottom-up proteomics

A vital part of experimental designs is randomization of the sample order. Sample handling and processing can introduce differences between the samples. If the samples are batched (grouped) and processed condition-wise without randomization, the unintentionally introduced technical variance can be inseparable from the biological variance of interest. Such batch effects can bias the interpretation of the results which can lead to false conclusions. In proteomics, systematic differences leading to batch effects can be introduced at several steps, including during

sample handling, sample preparation, and LC-MS/MS analysis where the performance often changes over time due to amongst others column wear, drift in calibration, temperature fluctuations, and contaminations. A complete randomization of the sample order can, however, create imbalanced designs by mere chance, where all samples of a given condition are batched. This can be overcome by block randomization, where the samples are batched into blocks according to a known grouping. While recommended, block randomization has yet to find widespread usage by the proteomics society (62).

The first step in preparing a sample for bottom-up proteomics in a generic workflow is solubilization of the proteins and inhibition of *in vivo* protease activity. Plasma and plasma-derived body fluids can often readily be mixed with a lysis buffer containing detergents to solubilize the proteins. Homogenization of solid samples, such as tissues or cells, is critical to ensure optimal proteome coverage, and can be done mechanically in a lysis buffer. To prevent unintended proteins degradation by *in vivo* proteases, and ensure a later efficient enzymatic digestion, the proteins are denatured by reducing the cysteine disulfide bonds with dithiothreitol (DTT) or tris(2-carboxyethyl)phosphine (TCEP), which often is included in the lysis buffer. The denaturation can be enhanced by heating. To prevent reformation of the disulfide bonds which would complicate the downstream data-analysis, reduction is followed by alkylation, commonly with iodoacetamide (IAA) (63). The denatured and



solubilized protein sample is subsequently proteolyzed into peptides. Commonly the protease trypsin is used which cleaves the peptide bond on the C-terminal side after arginine and lysine. The enzymatic activity of trypsin is usually high, but can be reduced, e.g. in the presence of a proline on the C-terminal side of arginine or lysine residues (64) or upon modifications at or close to the cleavage site, such as, citrullination of arginine as shown by Bennike et al. (38). Following digestion, the peptide-sample can be purified from salts and contaminants using C18 columns, either offline or automatically in combination with the LC-MS/MS analysis (63). No universally applicable method for protein digestion has been developed, and the recent development in available methods are discussed in the following sections.

3.1.1 In-solution digestion

In-solution is arguably the simplest digestion method, where all chemicals and components are sequentially added to a tube. The usage of a single tube minimize protein loss caused by adsorption to the tube and tip walls. However, given the absence of buffer exchange or wash steps all chemicals must be compatible with the digestion conditions and downstream analysis, and C18 purification is often required to remove contaminants from the peptide product. In addition, the enzymatic digestion efficiency is often reduced leading to a decreased number of identified peptides and proteins as shown by Bennike et al. and others (65, 66).

3.1.2 In-gel digestion

The in-gel digestion protocol was introduced by Shevchenko et al. in 1996 (67) and still has widespread usage due to its many advantages. Briefly explained, the homogenized and solubilized

protein sample is pre-separated by one- or two-dimensional gel electrophoresis on polyacrylamide gels, which efficiently remove low molecular weight impurities. The purification enables the use of detergents and buffer components which are incompatible with downstream analysis. The gels are stained, often by silver-staining or Coomassie brilliant blue, bands of interest are excised from the gel, enzymatically digested, and analyzed by LC-MS/MS (67, 68). In-gel digestion used to be a cornerstone for proteomics research as a well-established and robust method. However, in spite of the many advantages of in-gel digestion, the method has several drawbacks. These include a relatively laborious and time-consuming protocol (5) with a high risk of keratin contamination from the researcher and a lower reproducibility in protein identifications (65, 69), compared to other methods.

3.1.3 Gel-free digestion methods for proteomics

A paradigm shift was the introduction of sample preparation and digestion for proteomic analyses using spin filters termed filter-aided sample preparation (FASP), which overcomes many of the limitations from the in-gel and in-solution protocols. The FASP method was described by Manza *et al* in 2005 (69), fully realized by Wiśniewski et al. in 2009 (70), and further improved by Leon et al. in 2013 (66).

Briefly explained, solubilized proteins are transferred to a molecular cutoff ultrafiltration unit. All reactions are done on the filter which facilitates buffer exchanges and washes. The volume during digestion can be kept at a minimum which increases the protease concentration and digestion efficiency. Compared to the in-gel digestion, the hands-on and total sample preparation time is reduced, and the method has a lower risk of keratin contamination

from the researcher. As shown by Bennike et al. and others, FASP has a higher reproducibility in protein identifications, compared to in-gel and higher number of identified proteins as compared to in-solution (65, 69). The FASP protocol has been extensively used to characterize the innate immune system in health and diseases for a variety of sample-types, including plasma, plasma-derived synovial fluid (38, 65), and tissue (3).

A drawback of FASP is the relatively large loss of sample material to the FASP filters. While plasma studies using blood drawn with syringes are not sample limited, studies of finger- or heel-prick samples, e.g. of newborn blood or isolated immune cells, may benefit from protocols requiring less material. To improve the proteome coverage of sample-limited studies, the in-StageTip (iST) method by Kulak et al. (71) and suspension trap (S-trap) method by Zougman et al. (72) were developed and published in 2014. Rather than using the larger spin-filters in FASP, iST employs a C18 disk inserted into a pipette tip container and S-trap uses several disks of different filters to retain proteins and peptides and facilitate C18 purification. Another widely used method by Hughes et al. (73, 74) termed single-pot solid-phase-enhanced sample preparation (SP3) relies on binding the proteins and peptides on the hydrophilic surface of carboxylate-coated magnetic beads. Washes and buffer exchanges are performed while retaining the beads using a magnet. The usage of small disks or beads compared to the relatively large FASP spin-filters minimizes sample loss. A direct comparison of iST, SP3, and FASP showed similar performances for 20 μg of starting material, although handling time was reduced with iST and SP3 (75). However, below 10 μg of starting material, the number of identified proteins and the quantitative reproducibility decrease for FASP. In contrast, iST and SP3 provide high proteome coverage down to 1 μg of starting material (75). Geyer et al. (76) performed a proof-of-concept study using iST with 5 μL blood collected from a single finger prick and identified 285 proteins across ten individuals, which could be beneficial where normal blood-draw by syringe may be challenging, e.g. in studies of newborns. However, the protocol has yet to find wide usage.

3.1.4 Parallel and automated digestion methods

An increasing number of proteomics studies include hundreds to thousands of samples (7, 77, 78). In translational and clinical proteomics with large cohorts, the multi-tier plate is the preferred format for sample processing and storage. The plates facilitate the usage of multi-channel pipettes and are more easily automated (5).

A 96-well plate format of the FASP protocol was published in 2013 (79). However, the several centrifugation steps needed for liquid transfer during buffer changes could only be performed at 2,200 \times g which, compared to the original FASP protocol with 14,000 \times g, prolonging all centrifugation steps from tens of minutes to three to four hours. To decrease the sample preparation time while keeping the required amount of sample low, the MStern digest method was developed and published by Berger et al. in 2015 (5), and modified for plasma by Bennike et al. in 2018 (4, 80). The MStern protocol employs large pore polyvinylidene difluoride (PVDF) membranes in a 96-well plate for protein retention, thereby utilizing the protein charge interactions known from western blotting and vacuum for the liquid transfer steps (5). The

MStern protocol is a simple and fast method and has e.g. been applied by Boada et al. (81) to study host response pathways during virus infections, by Bennike et al. (4) to study the impact of surgery on the serum proteome, and by Bennike et al. (53, 82) to study the development of the complement system over the first week after birth. Similarly, a 96-well plate format of the SP3 protocol has been developed and was in 2020 automated (83). The MStern protocol results in a higher ratio of tryptic missed cleaves which may increase the repeatability in protein identifications across samples as additional peptides can be generated from each protein (5). However, the missed cleavages may be problematic e.g. when studying the PTM citrullination. The high efficiency of trypsin to cleave after arginine amino acids, is inhibited by arginine citrullination, and as demonstrated by Bennike et al. (38), a tryptic missed cleavage after arginine can be used to support identification of citrullinated peptides. However, due to the higher number of peptides with missed cleavages with the MStern protocol, a higher ratio of peptides falsely annotated as citrullinated can be expected with this protocol. Although this has yet to be experimentally demonstrated, studies of citrullination will likely benefit from the iST or SP3 protocols. A comparison between the iST, SP3 and MStern digest protocols found the highest number of identified proteins with iST and SP3 (84). However, the study used urine as material for the evaluation, and a comparison between the 96-well plate digest protocols remains to be performed using plasma which would be beneficial for large-scale plasma proteomics studies.

3.2 Proteomics analysis using LC-MS/MS

Following protein digestion, the generated peptide sample is analyzed by LC-MS/MS (Figure 3B). In short, the peptides are introduced onto a reversed-phase LC column. The column contains a stationary phase of hydrophobic C18 material, which binds the peptides. A mobile phase is passed through the column consisting of water mixed with acetonitrile acidified with formic acid. By gradually increasing the acetonitrile-to-water ratio from approximately 2% to 40% over minutes to hours, the hydrophobicity of the mobile phase can steadily be increased. At a given hydrophobicity, which is dependent on the peptide sequences and modifications, the peptides will pass from the stationary phase into the mobile phase and be eluted from the column. The eluting peptides are ionized and introduced into the mass spectrometer using electrospray ionization (ESI) (85), and the mass spectrometer measures the mass-to-charge ratio (m/z) of the peptides and/or peptide fragments and the signal intensities (61, 86). Lighter ions or ions with a higher charge will have a lower m/z value, while heavier ions or ions with a lower charge will have a higher m/z value. The signal intensity is dependent on the concentration, but also the properties of the peptides, including amino acid sequence and modifications. Resultingly, two peptides with distinct amino acid sequence can yield different signal intensities despite having the same concentration. However, the relative signal intensity of a given peptide correlates well with the absolute concentration, which is used for relative quantification in

proteomics. Based on the m/z 's and signal intensities, the plot of which is termed a mass spectrum, the ions can be separated. Due to the naturally occurring C^{13} isotopes, the mass spectrum will contain m/z signals from a given peptide with only C^{12} , with one C^{13} , two C^{13} , and so forth, thereby generating an isotopic distribution on the mass spectrum with a mass difference of one neutron (1.0087 Da). Based on the measured m/z difference in the isotopic distribution, the charge and mass of the peptide can be calculated (61, 86).

Peptide signals selected for identification are isolated by the mass spectrometer, and fragmented by applying energy, typically in a collision-induced dissociation (CID) process. The mechanisms of CID reactions are incompletely understood, but preferentially the C-N peptide bonds are broken. The m/z signals from the resulting peptide fragments are dependent on the peptide amino acid sequence and modifications, and the signals are measured by the mass spectrometer (61, 86). From the measured peptide fragments, the peptide sequences are determined from which the proteins in the sample can be inferred (87).

Several acquisition methods exist for the selection of peptides for fragmentation and identification. The two widely used methods data dependent acquisition (DDA) and data independent acquisition (DIA) are discussed in the following sections.

3.2.1 Data dependent acquisition

DDA is arguably the most applied acquisition method in proteomics. However, it is expected that DIA will supersede DDA in the near future. In DDA, the m/z signals of peptides eluting from the column are determined in a mass spectrum termed a MS1. Typically, an m/z range of 400-1600 is monitored, yielding coverage of most peptides generated by trypsin from human samples. Based on the MS1, top N peptides (typically 5-20) are selected for identification based on intensity (Figure 4A). One at a time, the selected peptides are isolated, fragmented, and the fragment masses are determined in a mass spectrum termed a MS2. The top N MS2-

spectra are preceded by another MS1 within 1-3 seconds, and the cycle is repeated throughout the analysis which can last tens of minutes to a few hours per sample (88).

The data produced by modern mass spectrometers is a collection of MS-spectra which can be tens of spectra per second of analysis time. The identification of peptides and proteins from database searching was in 1994 automated with SEQUEST, in a landmark publication by Eng, McCormack, and Yates (89). Since then, a number of search engines have been developed, including Mascot, and Andromeda (90) integrated in the widely used MaxQuant software (91). The MS2 m/z signals are matched (Figure 3C) to predicted values based on a user-provided protein database which allows for identifying the peptide and protein sequences. Resultingly, proteins which are not present in the database will not be identified even if they are present in the sample. Entire proteomes are often included in the search and uniprot.org is a widely used resource. The protein databases have been established in sequencing studies which have enabled the field of proteomics (63).

Although widely used, DDA is less efficient when analyzing samples with few high abundant proteins such as plasma. During elution of high abundant peptides, the MS1 mass spectrum, which detects the resulting peptides eluting from the LC column in the entire monitored m/z range, will be dominated by individual peptide signals. This reduces the relative signal intensity of all other co-eluting peptides. High abundant peptides may thereby prevent the selection of low-abundant peptides for fragmentation and sequencing in MS2, effectively hindering the identification of lower abundant proteins in plasma. Consequently, during elution of high-abundant peptides the absolute sensitivity of the mass spectrometer is reduced yielding a lower number of identifiable proteins (59).

3.2.2 Data independent acquisition

Several alternatives to DDA have been developed of which DIA (also known as SWATH-MS) (92) arguably is the most widely used

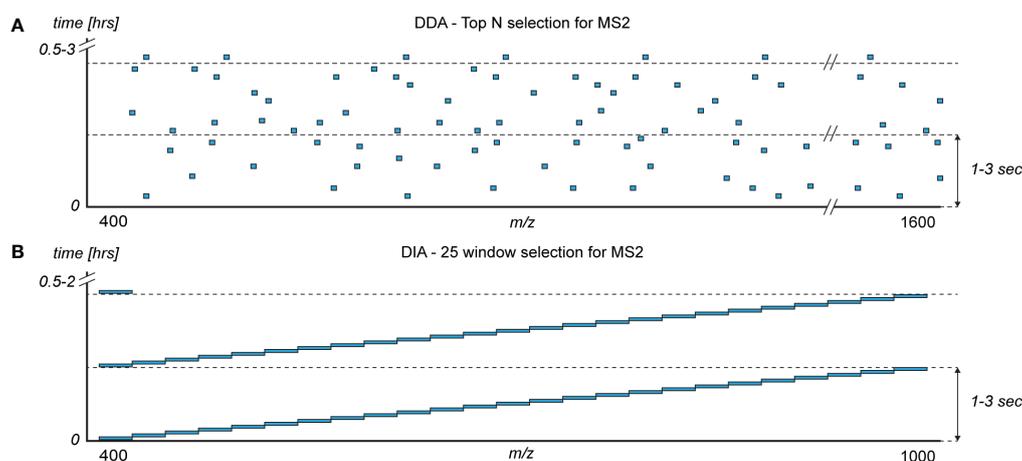


FIGURE 4

The monitored mass-to-charge (m/z) range (x-axis) of the intact peptides (MS1) during a sample analysis (y-axis) with the two commonly used mass spectrometer acquisition methods for selecting peptides for fragmentation (MS2): (A) data dependent acquisition (DDA) and (B) data independent acquisition (DIA). Blue boxes represent the typical distribution of m/z windows for selecting peptides for fragmentation and identification.

alternative. DIA is based on dividing the monitored m/z range into a number of fixed windows (Figure 4B), typically with a width of m/z 20–40. A number of different DIA window acquisition schemes exists which are compared elsewhere (93). Commonly, the monitored m/z -range is reduced in DIA compared to DDA to ensure a fast and sensitive method. However, the impact on the detectable proteins is minimal as e.g. m/z 375–1200 covers 95% of the peptides generated by trypsin from the top 1,000 plasma proteins (4). The mass spectrometer slides through the windows sequentially in a fixed manner, isolating and co-fragmenting all peptides in a given m/z window (Figure 4B). Resultingly, during elution of high-abundant peptides only the m/z windows containing the high-abundant eluting peptides have reduced sensitivity, which increases the number of identifiable proteins in plasma (8). Added, the non-deterministic nature of DIA decreases the technical variance in protein identification and quantification (4, 94).

The MS2 mass spectrum obtained with DIA contain multiplexed fragment signals from several peptides, in contrast to MS2 mass spectrum obtained with DDA which mostly contain signals from individual peptides. The multiplexed DIA-data complicates the downstream peptide identification and quantitation (92, 95), and has been a key limitation of DIA until recently. However, over the past decade a number of strategies and software solutions have become available for inferring identifications and relative quantifications from DIA data enabling the methodology (93, 96). Examples include DIA-NN, that exploits deep neural networks for the processing of DIA data (97) and Spectronaut by Biognosys. Recent support of DIA in the widely used LC-MS/MS data analysis software MaxQuant with the incorporation of MaxDIA (98), underlines the development and increasing popularity of DIA. From DIA MS2 mass spectrum, peptide sequences are determined by comparison to previously measured MS2 mass spectrum of individual peptides, collected in project-specific spectral libraries. The spectral libraries are typically established using DDA of pooled samples, and the library size can be expanded by analyzing depleted and/or fractionated plasma, which is discussed in section 3.4. As demonstrated by Bennike et al. (4), expanded spectral libraries allows for analyzing plasma samples using DIA, while making use of depleted and/or fractionated samples in the spectral library. The methodology allows for doubling the number of identifiable plasma proteins and increase the quantitative accuracy as compared to DDA, thereby monitoring a larger fraction of the humoral innate immune system components. Expanded spectral libraries can in this way be used to increase the number of identifiable proteins in the DIA analysis up to a certain point, where reduced benefits have been reported of very large spectral libraries (93).

The recent technical development now allows for inferring peptide identifications directly from DIA data based in spectral library-free analysis, thereby only relying on a user-provided protein database. Library-free analysis is supported by both DIA-NN, Spectronaut, and MaxDIA (93, 97–99). Eventually the need for DDA-generated spectral libraries may be removed altogether. However, library-free analysis requires additional computational time and resources which may pose a limitation for large studies. In

addition, the ability of spectral library-free methods to identify novel features remains to be investigated, which speculatively could pose a problem for identification of e.g. novel PTMs.

3.2.3 Identification and filtration

To ensure reliable identifications across various platforms, sample types, and protein databases, identified peptides are filtered using a target-decoy competition approach (Figure 3C) (100). In short, the protein database is scrambled by reversing the amino acid sequences and used as a decoy. The measured MS2-spectra are matched against both the target and decoy database. All peptide matches are given a score based amongst others on the number of m/z signals and accuracy, with higher scores representing more confident matches. Matches in the decoy database are expected to occur at random and not be present in the sample, and on average they will have lower scores. A minimum score is then determined where commonly 1% of the list of identifications are decoys, termed the global false discovery rate (FDR). All matches with the minimum score or higher in the target protein database are accepted as being true matches, yielding a list of identifications where a maximum of 1% is expected to be false identifications. The method is largely independent of the score used by the search engine or the proteomics workflow and sample (101, 102). However, although widely used as the default method, the target-decoy approach only controls the average proportion of false discoveries (100), and as shown, additional filtering of proteins may be needed (65).

3.2.4 Quantitation and quality control

Following identification, peptides and proteins can be relatively quantified based on the correlation between the peptide signal intensity and concentration in the sample (Figure 3D). The quantification strategies in proteomics can roughly be divided into label-free approaches (including spectral counting, MS1-based, and MS2-based) and stable isotope labeling approaches (including SILAC, iTRAQ, and TMT). Briefly explained, in label-free approaches each sample is analyzed separately by LC-MS/MS and the relative peptide quantities are determined by comparing the signal intensities across the different samples (or with spectral counting the number of MS2 spectra). In labelling approaches, each sample is given a unique chemical or isotopic tag prior to LC-MS/MS analysis. The samples are then mixed and analyzed by LC-MS/MS which, amongst others, reduces the number of samples to be analyzed. The relative peptide quantities in the samples are determined based on the signal intensity of the unique mass-different tags. In both approaches, the peptide quantities are subsequently used to calculate the parent protein abundances. A discussion of quantitation strategies is beyond the scope of this review, but have been extensively compared and discussed elsewhere (103, 104).

Following quantitation, a crucial step is quality control of the data. This is commonly done by visually inspecting the LC-MS/MS raw-data and investigating quality control samples which should be run periodically to ensure system performance. High intensity signals from contaminants, low peptide intensities, and low number of identifications can indicate issues with e.g. sample

processing or the LC-MS/MS analysis. In addition, unsupervised statistical methods such as principal components analysis (PCA) are often applied to identify outlying samples and identify batch effects, e.g. by investigating if the data is dependent on the sample processing order. Outlying samples should be critically evaluated and potentially removed following proper argumentation. If batch effects are present, attempts can computationally be made to minimize the impact, e.g. using the R package ComBat (105), given that the experimental design has been made with an appropriate randomization.

Quantitation and quality control may be proceeded by identification of features with a statistically significant difference between two or more groups (e.g. sick *versus* healthy) (Figure 3D). Depending on the number of quantified proteins or peptides, the number of performed statistical tests can be vast. Correction for multiple hypothesis testing, by adjusting p-values using e.g. Benjamini-Hochberg (106) or permutation-based FDR control (107), is often paramount to control for false positives (type I errors).

3.3 Statistical workflows to identify differentiating proteins

A common and critical step in many proteomics studies is the identification of peptides and proteins with a statistically significant difference between the groups. The t-tests remain widely used in proteomics, presumably due to the robustness, simplicity, and current integration into recognized and widely used data analysis tools, including Perseus (108). However, while highly proven, t-test based approaches do not allow for taking confounding factors such as sex and replicates into account, which often necessitates the merging or subdividing of the dataset, which may have an adverse outcome on the conclusions. Several alternative statistical methods have been proposed to the t-test (109, 110) including linear models (LM) and the extended linear mixed effect models. LM allows for including confounding factors, such as age, sex, and batches. The compatibility with more complex experimental designs can remove the need to subdivide and merge the dataset, thereby allowing for analyzing data from all samples in one analysis. Correlations between samples due to confounding factors or repeated measures can be modelled, which may remove the need to merge e.g. technical replicates prior to analysis. LM-approaches thereby enable the evaluation of flexible hypotheses and interactions beyond a direct comparison between groups (111). LM has been demonstrated to have a higher statistical power to identify proteins with statistically significantly changed levels. This is especially the case with smaller sample sizes ($n = 5-15$ samples), whereas the statistical power often is more similar to the t-test with higher number of samples ($n = 50-100$) (109, 110). Resultingly, proteomics studies may often benefit from alternative analytical strategies to the t-test (109–113). The increased biological insights which may be obtained have been demonstrated by the reanalysis of proteomics data with LM approaches, which originally was performed using t-tests. Kirov et al. (114) reanalyzed proteomics data from a study of colon tissue samples from UC patients and

controls. Using LM rather than t-tests, additional proteins could be associated with UC. The analysis made it possible to associate degradation of the extracellular matrix in the colon mucosa with UC, and identify neutrophils and NETs as the likely drivers of the degradation. The increased statistical power in the reanalysis to identify additional differentiating proteins compared to the t-test-based analysis likely arose from the technical replicates which did not need merging in the LM-based analysis.

Given the robustness and current implementation into existing tools, the t-test remains easy to apply, whereas the application of LM-approaches to proteomics data mostly remains confined to the R framework. Several packages have been developed in R for analyzing proteomics data using additional statistical methods, including MSstats, Proteus, limma, and lme4, which are compared elsewhere (115). However, most R-packages require some scripting experience. Although the statistical analysis of proteomics data increasingly is moving to R, the lack of implementation in the analytical tools commonly used by the proteomics research community, including Perseus (108), likely represents an obstacle for wider adaptation and may explain the retained popularity of t-test based approaches. Incorporation of alternatives to the t-test into "click-based" proteomics data analysis software which is commonly used by the proteomics society, including Perseus, may help maximize the biological insights which can be obtained from proteomics studies.

3.4 Workflows to study the plasma proteome

Proteomics have been applied to measure the levels of humoral innate immune system components in plasma, to identify ongoing immune processes in health and disease. Identification of innate immune system markers in plasma holds great potential to noninvasively guide diagnosis and treatment decisions in inflammatory diseases and cancers. However, as outlined plasma proteomics remains technically challenging due to the proteome composition (Figure 2). Various workflows for plasma proteomics have been developed, which will be discussed in this section. In addition, a number of practical considerations have been published by Geyer et al to avoid systematic biases in plasma proteomics studies (116).

3.4.1 Depletion of high abundant proteins

Depletion of high abundant proteins can reduce the dynamic range of concentrations and increase the coverage of the plasma proteome. The targeted depletion of high abundant plasma proteins with antibodies was introduced in 2003 (117), and often depletion of the top 7, 14, or 20 most abundant plasma proteins are applied (56). However, as demonstrated antibody-based depletion schemes typically result in a modest increase in number of protein identification by 25-50% while increasing the required hands-on time during sample preparation (4, 118–120), and care should be taken when interpreting the levels of the proteins targeted for depletion. Non-targeted proteins may also unintentionally bind to the depletion material and be partially depleted (120). In addition,

the physiological function of albumin as a protein carrier causes albumin depletion to result in unintentional co-depletion of many other non-targeted proteins, including clinically useful biomarkers (121). However, other studies report of relatively few non-targeted proteins to be depleted using top 7 and 14 approaches (118). In sum, the impact of unintended protein depletion when depleting high abundant plasma proteins remains debated, and depletion strategies remain commonly applied in plasma proteomics to increase the proteome coverage. Validation of findings in non-depleted samples e.g. using antibody-based approaches is recommended, and care should especially be taken when interpreting the levels of the proteins targeted for depletion.

3.4.2 Preanalytical fractionation

Preanalytical fractionation also allows for increasing the proteome coverage of plasma samples by expanding the LC-MS/MS analysis time (4, 58). The sample is fractionated, and each fraction is subsequently analyzed by LC-MS/MS, from which the term prefractionation arises. Prefractionation can be performed before digestion on the protein level or after on the peptide level (58). An example of protein-level prefractionation is the polyacrylamide gels from in-gel digestion, and on the peptide level high pH C18 and strong cation exchange (STX) (122). Using extensive prefractionation techniques, Garay-Baquero et al. (55) identified more than 5000 plasma proteins spanning 11 orders of magnitude in concentration and Keshishian et al. (56) identified 5300 plasma proteins using prefractionation in combination with depletion. The studies demonstrate the impressive complexity of plasma and underline the potential for plasma proteomics in characterizing health and disease. Several plasma workflows with prefractionation have therefore been developed (58). However, prefractionation techniques increase the hands-on time for sample preparation and LC-MS/MS analysis by several factors (4), which may be infeasible in large-scale proteomics studies (4, 56). Accordingly, the study by Garay-Baquero et al. only included 21 samples and Keshishian et al. four samples.

3.4.3 Workflows for large cohorts

Focusing on workflows which are compatible with larger cohorts, in 2016 Geyer et al. (7) reported in a landmark publication on the analysis of nearly 1,300 non-depleted non-prefractionated plasma samples using DDA. By incorporating proteomics data in a spectral library-like manner, 437 proteins could on average be identified per individual. The large number of samples enabled a mapping of the effects of losing weight on the plasma proteome, and included reduced levels of CRP and SAA1, and a reduction in other markers of systemic low-grade inflammation in response to weight loss.

In 2018, Bennike et al. (4) developed a 96-well plate workflow based on MStern digest and DIA with an extended spectral library, which from 1 μ L non-depleted plasma allows for monitoring on average 439 plasma proteins per sample. A proof of concept was generated on samples from 62 individuals demonstrated an improved proteome coverage and quantitative accuracy of DIA compared to DDA, and the methodology covered all main complement system components. The methodology was used to

assess the impact of total pancreatectomy with islet autotransplantation (TPIAT), a specialized surgical procedure performed to relieve the pain of severe chronic pancreatitis. In TPIAT, the pancreas is removed and the insulin-producing islet cells are isolated, and transplanted into the liver to retain endocrine production. The study of 96 samples collected before and a year after surgery demonstrated that TPIAT reduced the levels of many apolipoproteins and IgM. In 2019 Bruderer et al. (6) used a similar workflow and analyzed 1508 non-depleted plasma samples using DIA, and identified 408 proteins per participant. The dataset enabled characterization of the effects of weight loss and weight maintenance on the proteome, revealing a reduction in low-grade inflammation.

In 2023, Viode et al. (78) developed the PerCA workflow, for acid-based precipitation and depletion of high abundant plasma proteins. The perCA workflow enabled, for the first time, the preparation and analysis of more than 3,000 depleted plasma samples identifying 576 plasma proteins on average per individual using DDA, and 1,343 protein using DIA. The increase in identifications with DIA vs DDA is in accordance with the reported increase in identifications reported by Bennike et al. (4). The vast number of samples enabled detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)-derived proteins in COVID-19 patient plasma, and the study correlated the SARS-CoV-2 protein abundance with the respiratory status of the patients.

The recent development in plasma proteomics workflows now allows for monitoring more than 1,300 plasma proteins in cohorts with thousands of samples. The increased number of monitorable plasma proteins allows for studying an even greater part of the innate immune system's humoral components. This is of high relevance in studies of health and disease, e.g. the immune system development and inflammatory diseases (123). In cancer, the tumor microenvironment stimulates secretion of pro-inflammatory cytokines, which may inhibit or promote cancer proliferation and metastasis through inflammatory processes, e.g. by secretion of the neutrophil chemoattractant IL-8 that also stimulates NETosis (124). The cytokines are detectable in circulation and thereby holds great potential as noninvasive diagnostics and prognostic cancer biomarkers, which has yet to be fully explored (57). The usage of DIA in combination with acid-based precipitation and depletion of high abundant plasma proteins has set a new standard for the feasible number of analyzable samples. Plasma samples should ideally be analyzed using DIA-based approaches, in combination with LM-approaches for identifying statistically significantly regulated proteins. Synovial fluid is a plasma filtrate and by combining proteomics analysis of synovial fluid with transcriptomics of the synovial membrane, Bennike et al. (65) have demonstrated that the majority of the synovial fluid proteins are plasma derived, with a small fraction secreted directly to the fluid. Given the similar proteome composition of high-abundant proteins in plasma and plasma-derived synovial fluid as demonstrated by Bennike et al. (65), studies of synovial fluid and other plasma-derived body-fluids may likely also benefit from proteomics workflows optimized for plasma. While current workflows providing coverage of the majority of the innate

immune system humoral components in workflows compatible with thousands of samples, the concentration of low abundant signaling molecules including cytokines remains below the limit of detection (4, 38, 65, 76). Proteomics workflows enabling determining the levels of cytokines, many of which are APP inducing, would likely cause a paradigm shift, and enable characterization of disease specific cellular signaling and innate immune reactions in inflammatory diseases and cancers.

4 Ontogeny of the complement system components in the first week after birth

The innate immune system plays a significant role in providing immunity for newborns (125–127). Plasma proteomics workflows have recently been applied to study the early developmental trajectory of the innate immune system humoral components in the critical first weeks after birth (53, 82). In the mature immune system, a key aspect of immunity to bacterial infections are circulating antibodies against polysaccharides on pathogenic bacterial surfaces. However, coming from a tightly regulated *in utero* environment with tolerance-inducing immune suppression, the newborn immune system is in a uniquely developing state and is naïve to most pathogens. Lacking an efficient adaptive immune response, in the first six months after birth newborns largely rely on the innate immune system for immunity. The immune protection is aided by maternal IgG that is transferred from the mother across the placenta and circulate in the fetus and newborn plasma in the first months after birth until degraded (125–127). In addition, maternal secretory IgA (sIgA) represents over 90% of the Ig-content in human milk and early colostrum, and postnatally sIgA adds to protection against infections, primarily in the intestinal epithelium and do not enter circulation in substantial amounts (128–130).

As outlined in the introduction, the complement system is a vital part of the innate immune system response, working in combination with immune cells. The importance of the complement system for newborn immunity to pathogens is increasingly studied (53, 82, 131). However, few plasma proteomics studies have been performed of the innate immune system development in the critical first week after birth, likely due to the difficulty in obtaining relevant samples and the analytical challenges in analyzing small sample volumes (125, 132, 133). Given the central role of the innate immune system for newborn immunity, this section discusses recent proteomics findings in the development of innate immune system humoral components circulating in plasma in the first week after birth.

4.1 Complement system components levels at birth

To enable a functional interpretation of the relative changes to complement components (Figure 1) which are provided by plasma

proteomics, complement levels in newborns relative to adults will first be described.

The complement components are synthesized early in fetal life, mainly by hepatocytes with a few exceptions including C1q and C7 being secreted mainly from neutrophils and monocytes/macrophages (134). The concentration of most complement components are reduced, with term newborns ranging from 10% to 80% relative to adults (131, 135, 136). A notable exception is C7 from the MAC which is near adult levels (131, 136). Newborn levels of MBL from the lectin activation pathway (Figure 1B) and C3 are disputed, with reported levels ranging from near adult levels to approximately 50% (131, 135, 137). The overall deficiency in complement level are even more pronounced in premature newborns (131, 136, 138), and in premature newborns with patent ductus arteriosus, reduced levels of C8 have been shown (139). Resultingly, newborns and in particular premature newborns have a decreased efficiency in bacterial lysis and opsonization, leading to reduced immunity and increased susceptibility to infections (131).

4.2 Complement system component levels the first week after birth

As outlined, proteomics can characterize the main complement system components in a plasma sample, and a proteomic study by Bjelosevic et al. (140) has shown that the plasma profiles of neonates are strikingly dissimilar to that of children and adults.

Focusing on the critical first week after birth, a study by Lee et al. (82) analyzed blood from newborns collected on the day of birth and a sample on day 1, day 3 and day 7. The samples were analyzed by proteomics, transcriptomics, metabolomics, and single-cell immunophenotyping. Using "multi-omics" integration the study found that the developmental profiles overall follow a robust trajectory in the first week after birth, even in populations from different geographical areas. The findings underline that after birth the innate immune system undergoes vast, but robust, developmental changes within the first week after birth.

Focusing on the complement system, a re-analysis of the proteomics data by Bennike et al. (53) using LM-approaches enabled a characterization of the developmental trajectory of all main complement components. The study confirmed that levels of complement components mainly related to the classical activation and terminal pathways increase over the first week after birth. Namely, plasma levels of the PRR in the classical activation pathway (Figure 1A) C1q, and the other components in the C1-complex (C1r, and C1s) increase significantly as early as day 3. In contrast, plasma levels of the PRR in the lectin activation pathway (Figure 1B), FCN-2 and FCN-3 decreased over the first week after birth. Lacking a distinct PRR, the alternative activation pathway (Figure 1C) is more difficult to characterize, but levels of CFB increased as early as day 1. In contrast, levels of the central alternative complement enhancer CFP was decreased at day 7. Given the relative deficiency in most complement components, the findings could indicate that the development of the classical activation pathway is the first to be initiated postnatally (53).

Supporting these findings, a proteomics study by Grumach et al. (138) found no change to MBL nor CFP levels at day 5.

Focusing downstream of the complement PRRs, the study by Bennike et al. (53) demonstrated that plasma levels of C2 and C4 in the classical and lectin activation pathways increased in the first week after birth and as early as 24 hours after birth. In addition, plasma levels of main complement components in the MAC (C5, C6, C8, and C9) increased whereas plasma levels of C7 decreased. C7 is often the rate-limiting factor in MAC formation and is at relatively high levels at birth (131, 140). The observed increase may indicate the development of components in the classical activation pathway, the terminal pathway, and MAC towards functional levels in the first week after birth.

4.3 The classical complement system pathway early after birth

The apparent increase in main components in the classical activation pathway and the terminal pathway in the first week after birth goes against the previous notion that the entire complement system develops as a single unit (141). However, the findings are supported by an increased functional activity of the classical complement pathway in plasma from 5 days old newborns compared to at birth in cord blood (138). The difference in activity was even more pronounced for preterm neonates (< 34 week) where no significant functional activity of the classical pathway could be measured in cord blood but functional activity could be measured at day 5 (138), thereby aligning with the trajectory of the complement component plasma levels.

Speculatively, these findings could indicate a key role for the classical complement pathway in neonatal immunity. The main target molecules of the classical complement pathway PRR are IgG and IgM. IgG is the most abundant Ig in circulation in newborns, infants, and adults (142). Newborn IgG levels are above adult levels, due to active maternal transfer across the placenta mediated by the neonatal Fc receptor (143). IgG exists in subclasses IgG1-4, where IgG1, and IgG3 have the highest classical complement activation potential (22, 144). Nearly mirroring the activation potential, the placental transfer efficiency has been reported to be IgG1>IgG3>IgG4=IgG2 (145) making IgG1 the main maternally transferred Ig-subclass over the placenta and the most abundant in fetal and maternal circulation (142). Bennike et al. (53) have integrated proteomics and transcriptomics data and demonstrated an increased endogenous IgG1 synthesis in newborn whole blood as early as three days after birth. However, the synthesis rates remained insufficient to maintain the high IgG1 plasma levels at birth, caused by the efficient placental transfer of maternal IgG1. Accordingly, IgG levels reportedly decline over the first week after birth (53) and until five months of age (146). In contrast, IgM cannot transverse the placental membrane and newborn levels are low relative to adults (146). Bennike et al. reported that increased plasma IgM levels and endogenous synthesis can be detected in the whole blood cells within the first week after birth (53).

The studies demonstrate the increased plasma concentration of nearly all main components of the classical- and terminal complement pathways early from birth. Speculatively, increasing the classical- and terminal complement pathways to functional plasma concentrations early from birth, would enable newborns to initiate a humoral innate immune response against pathogens that are recognized, for instance, by maternal IgG1. This would highlight the importance of the classical complement pathway for newborn immunity, which evolutionarily is the most recently developed complement pathway compared to the lectin- and alternative pathway (147). Arguing against this hypothesis, higher plasma levels of activation product specific to the alternative pathway has been reported in 6 hour old newborns with early onset infection, compared to the activation product from the classical pathway (148). This could indicate increased activity of the alternative pathway in neonates. However, impaired regulation by the low newborn levels of CFH and CFI, which only increase significantly after the first week after birth (53, 138), may also explain the findings. Additional analysis of the complement levels at a higher time resolution and plasma proteome coverage, using recently developed plasma proteomics workflows, may enable additional insights into the development of the newborn innate immune system in the critical first week after birth.

5 Neutrophil extracellular traps in inflammatory diseases

The formation of NETs by neutrophil NETosis represents an important innate immune response to pathogens. In addition, NETs have been correlated to cancer progression and metastasis, and IL-8 secreted by the tumor microenvironment has been found to recruit neutrophils and induce NETosis (124). The role of neutrophils in cancers both as pro- and anti-tumor activity inducing entities is becoming increasingly apparent (14, 123). NET enzymes may degrade the extracellular matrix, in some circumstances inhibiting tumor growth and in others supporting metastasis (14), and NETs have been reported to catch circulating cancer cells and thereby promote metastasis (149). NETs may thereby exhibit both a pro- and anti-tumor activity (14). Proteomics has been applied to study the protein coating on NETs. A study by Chapman et al. has identified 197 proteins on NETs and found that the protein coating is vastly depending on the type of stimulus (150), supplementing the list of 29 NET-associated proteins previously reported by O'Donoghue et al. (43). Citrullinated histone H3 bound to cell-free DNA is today considered a hallmark of most NETs, in addition to neutrophil elastase and MPO (151, 152). However, no NET-unique protein markers nor PTMs have been identified, and proteomics identification is often supported by immunostaining and cell-free DNA measurements. While being an important innate immune reaction providing immunity, NET-formation has also been associated with the pathogenesis of several immune-related diseases, including the prevalent RA (37) and IBD (3). The recent findings will be discussed in this section.

5.1 Rheumatoid arthritis and neutrophil extracellular traps

RA is the most common systemic inflammatory joint disease, affecting 0.5%–1% of the population worldwide. The disease is characterized by autoantibodies and immune complexes, leading to neutrophil infiltration of the synovial fluid in the joints and can ultimately result in cartilage damage and joint destruction. In RA, the first described autoantibody was rheumatoid factor (RF) which has affinity to the Fc part of IgG (153, 154). PAD4 citrullination has been associated with RA (151) and anti-citrullinated protein antibodies (ACPA) have been identified in 50–70% of RA patients at diagnosis, with a specificity over 90%. The ACPA have affinity for amongst others citrullinated fibrinogen-alpha, vimentin, and actin (155, 156), and ACPA targeting of citrullinated proteins is thought to drive the inflammation in ACPA positive RA (157).

As shown by several proteomics studies, ACPA targets are present in the synovial fluid from RA patients (37, 38, 158). However, the initial tolerance breach and production of ACPA targeting autoantibodies in RA can be detected years before disease onset, and may take place outside the joints (153). In addition to the synovial joints, several organs have been suggested as the tissue where the initial triggering of tolerance breach in RA occurs, based on the proteomic identification of ACPA targets. Using proteomics, Ytterberg et al. (159) identified joint ACPA targets in the synovial and lung tissue, suggesting the lungs as a site of tolerance breach. Accordingly, the environmental risk factor with the highest impact for RA is smoking, and smoking has been reported to increase PAD2 expression in the lung tissue and citrullination in bronchoalveolar lavage cells (160). However, also mucosal inflammatory processes have been associated with local ACPA production in serum positive individuals without clinical RA (161), and using proteomics Bennike et al. (162) have identified ACPA targets in the colon of RA patients and controls. Given the findings and the heterogeneity and systemic nature of RA, a single pathogenic origin is likely too simplistic to cover the complexity of ACPA positive RA.

Eventually, the disease progresses, and immune cells are stimulated to infiltrate the synovial membrane and synovial fluid, where several studies have associated NETs with RA. Proteomics studies by Birkelund et al. (158) and Spengler et al. (37) have demonstrated that NETs can be detected in the synovial fluid during RA. A proteomics study by Chapman et al. (150) has demonstrated that NETs generated by neutrophils from healthy controls and from RA patients are coated with ACPA targets, including citrullinated histone H3. Supporting this, a proteomics study by Spengler et al. (37) demonstrated that neutrophils isolated from RA synovial fluid express high levels of PADs compared to osteoarthritis (OA) patients, and that NETs in RA synovial fluid are coated with PAD2 and PAD4 enzymes and citrullinated proteins. In addition, the level of cell-free DNA was increased in RA consistent with increased levels of NETs in the synovial fluid. In agreement, the proteomics study by Birkelund et al. (158) found increased levels of NET associated proteins in synovial fluid RA compared to spondylarthritis (SpA) patients, and a correlation was found between cell-free DNA levels and NET proteins.

Speculatively, the increased levels of NETs in RA synovial fluid could be more than the passive result of neutrophil infiltration and activation, but rather have a central part in RA by actively driving the disease through the generation and presenting of ACPAs targets. The increased levels of NETs, and NET-bound ACPA targets within the joint would maintain the ACPA induced inflammation in RA. In addition, PADs bound to NETs could citrullinate proximate proteins which would represent additional ACPA targets within the joint, thereby further maintaining inflammation (37). Supporting this, autoantibodies in the serum of RA patients has been found to react with NETs (163), which highlight the involvement of NETs in the etiology of RA. Accordingly, the therapeutic inhibition of NETosis and increase clearance of NETs are being investigated in inflammatory diseases and cancer, and is discussed in section 5.3.

5.2 Inflammatory bowel disease and neutrophil extracellular traps

NETs have also been associated with the disease etiology of IBD (3), a disorder which affects more than 7 million people worldwide with an increasing incidence (164). IBD comprise several chronic disorders of the gastrointestinal tract, where Crohn's disease (CD) and ulcerative colitis (UC) are the most prevalent. IBD result in a reduced quality of life and life expectancy for the patients, and a significant economic burden for society due to lost labor and direct health care expenses (165, 166).

A hallmark of IBD is neutrophil infiltration and inflammation of the intestinal mucosa. The phenotype varies considerably between patients and the etiology and pathogenesis of IBD remains incompletely explained (166). However, there seems to be agreement on the following hypothesis of the IBD etiology; "an aberrant immune response against the gut microbiota, triggered by environmental factors in a genetically susceptible host" (167).

An increasing number of proteomics studies have investigated the IBD colon mucosa using LC-MS/MS based techniques and associated NETs with IBD. First reported in 2015, Bennike et al. (3) observed NETs in the UC colon mucosa, detected using proteomics and visualized in the colon tissue with confocal microscopy, thereby associating NETs with IBD. The abundance of the NET protein S100A9 was found to correlate with the colon inflammation (3). A reanalysis of the dataset by Kirov et al. (114) using LM-approaches identified five-times the number of proteins with significantly different levels in the colon mucosa. The increased amounts of data, allowed for identifying degradation of the colon mucosa extracellular matrix as a key element in UC, including lower levels of elastin and collagen. The degradation is likely driven by the increase of several proteases, including neutrophil collagenase and the NET associated neutrophil elastase. Consistent with the protease-driven extracellular matrix degradation, a comparison to published transcriptomics datasets found increased mRNA levels of elastin and collagen. This could be indicative of a transcriptional compensation, which although compensatory is insufficient to maintain the protein levels and thereby prevent degradation of the extracellular matrix. The findings demonstrate that degradation

of the extracellular matrix is part of the UC pathology, and may be driven in part by neutrophils and NETs (114).

The findings have been supported by additional immunohistochemical studies. In 2019, the presence of NETs in UC mucosal tissue was confirmed by Dinallo et al. (168) and in 2020 Li et al. (169) identified NETs in CD, in addition to UC tissues, and increased plasma levels of cell-free DNA. The finding of NETs in CD was confirmed in 2022 by Schroder et al. (170). The studies also confirmed the correlation between NET abundance and degree of inflammation, further supporting the involvement of NETs in the IBD etiology. Dinallo et al. (168) demonstrated that circulating neutrophils isolated from UC patients produce NETs in response to TNF- α stimulation, and reduced expression of NET-related proteins and diminished NET formation were seen in patients receiving successful treatment with anti-TNF- α (168). The findings indicate a role for NETs in sustaining mucosal inflammation in UC and that part of the anti-inflammatory effect of successful IBD treatment could be through reducing NETosis.

5.3 Neutrophil extracellular traps as a therapeutic target

NETs have been found to correlate with inflammation and suggested to have an active role in driving chronic inflammation (171). Given the association between NETs and inflammatory diseases, NET targeted therapies are under consideration focused on improving NET clearance or inhibiting NET formation.

During infections, NETs have been found to persist for several days (34). Although the aspects of mechanisms that clear NETs remain incompletely described (172), the degradation by deoxyribonuclease (DNase) I (34) and endocytosis by macrophages have been identified (173). In inflammatory diseases it has been suggested that DNase I is inhibited by anti-NET antibodies, which may block access to the NETs, and cause impairment of NET degradation which results in sustained inflammation (174). Speculatively, this could be relevant in a subset of inflammatory diseases where autoantibodies play a central part, including RA. Supporting this, increased NET degradation through administration of DNase I or inhibition of NETosis decreased cytokine levels in a DSS colitis mouse model of IBD (169). A proteomics study by Fisher et al. (175) found high levels of NETs in of sputum from severe SARS-CoV-2 induced COVID-19 patients, which was confirmed by microscopy. Treatment with DNase reduced the NETs levels and inflammation in the airways, and plasma proteomics revealed a lower systemic inflammation. The findings indicate that NETs may contribute to the respiratory failure in severe COVID-19 patients and NETs targeting therapies using DNase may be efficient. However, following removal of chromatin from the NETs with DNase, active neutrophil elastase and histones persist and may cause inflammation until cleared through unknown mechanisms (176). General inhibition of PADs may therefore be a more efficient strategy, and inhibition of NETosis by Cl-amidine have been shown to reduce the severity of collagen-induced arthritis in mouse models (177, 178). Although the studies did not directly investigate the

impact on NET formation, the findings suggest that inhibiting NETs by inhibiting PADs may represent a therapeutic target for the treatment of RA and IBD. Given the pro-tumor activity exhibited by NETs in some tumors, the potential of NET targeting are also investigated as means of therapy, where e.g. Cools-Lartigue et al have shown in a mouse model that treatment with DNase may reduce metastasis slightly (149).

Based on the findings, NETs are currently being investigated as a therapeutic target in IBD (179), RA (179, 180), and cancers (57). Inhibition of NET production and proteolytic activity in PAD4 and neutrophil elastase-deficient mouse models have been found to prevent collateral host tissue damage by NETs (176). While efforts are being made to produce an isotype specific and potent PAD4 inhibitor, currently no PAD4 inhibitor is ready for clinical use (151). Additional characterization of NET formation and clearance may help guide studies focused on novel therapeutic approaches to improve NET clearance or inhibiting NET formation for treatment of various inflammatory diseases and cancers. In addition, identification of innate immune system components, including NETs, released to circulation, with predictive value for diagnosis and prognosis may help guide therapy decisions.

6 Conclusion and future directions

Recent advances within proteomics have enabled the increasingly comprehensive characterizing of the proteome, while reducing the analysis and sample preparation time, thus facilitating the analysis of larger cohorts. The development is especially apparent within the technically challenging analysis of blood-plasma, where recently developed workflows enable the monitoring of more than a thousand proteins in more than three thousand samples. Additional studies of plasma, and samples with similar challenges such as synovial fluid, may benefit significantly from DIA-based methodologies as the default analytical strategy. The development of plasma proteomics workflows compatible with large cohorts, with sufficient sensitivity to determine the levels of low abundant cytokines, would likely cause a paradigm shift and enable characterization of disease specific cellular signaling and innate immune reactions. Non-invasive identification of biomarkers with diagnostic and prognostic value holds great promise for guiding treatment decisions for inflammatory diseases and cancers. The potential of integrating data from several different omics including proteomics, transcriptomics, and metabolomics in multi-omics approaches to obtain unique biological insights, have been demonstrated in several studies, highlighting their value in understanding the innate immune system in health and disease. The recently developed plasma proteomics workflows hold great potential for characterizing the innate immune system in health and disease. Additional studies of the critical early phase after birth where the innate immune system undergoes vast changes will likely further increase our understanding of the early innate immune system development. Similarly, the association between NETs and several inflammatory diseases and cancer is becoming increasingly evident. A promising avenue under investigation is the development of therapies that either inhibit NET formation or enhance NET clearance.

Additional studies into the formation, composition, function, and clearance of NETs, which likely exhibit greater heterogeneity than currently recognized, seem promising in advancing the development of novel NET targeting strategies. These approaches hold potential for mitigating the detrimental effects of NETs and reducing inflammation associated with these diseases. The continued development of proteomics platforms and workflows will continue to offer valuable insights into the proteome and advance our understanding of health and disease with translational perspectives for diagnostics and therapy.

Author contributions

TB: Conceptualization, Data curation, Visualization, Writing – original draft, Writing – review & editing.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Scheltema RA, Hauschild J-P, Lange O, Hornburg D, Denisov E, Damoc E, et al. 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* (2014) 13:3698–708. doi: 10.1074/mcp.M114.043489
- Cox J, Mann M. Quantitative, high-resolution proteomics for data-driven systems biology. *Annu Rev Biochem* (2011) 80:273–99. doi: 10.1146/annurev-biochem-061308-093216
- Bennike TB, Carlsen TG, Ellingsen T, Bonderup OK, Glerup H, Bøgested M, et al. Neutrophil extracellular traps in ulcerative colitis: A proteome analysis of intestinal biopsies. *Inflammatory Bowel Dis* (2015) 21:2052–67. doi: 10.1097/MIB.0000000000000460
- Bennike TB, Bellin MD, Xuan Y, Stensballe A, Møller FT, Beilman GJ, et al. A cost-effective high-throughput plasma and serum proteomics workflow enables mapping of the molecular impact of total pancreatectomy with islet autotransplantation. *J Proteome Res* (2018) 17:1983–92. doi: 10.1021/acs.jproteome.8b00111
- Berger ST, Ahmed S, Muntel J, Cuevas Polo N, Bachur R, Kentsis A, et al. MStern blotting—high throughput polyvinylidene fluoride (PVDF) membrane-based proteomic sample preparation for 96-well plates. *Mol Cell Proteomics* (2015) 14:2814–23. doi: 10.1074/mcp.O115.049650
- Bruderer R, Muntel J, Müller S, Bernhardt OM, Gandhi T, Cominetti O, et al. Analysis of 1508 plasma samples by capillary-flow data-independent acquisition profiles proteomics of weight loss and maintenance[S]. *Mol Cell Proteomics* (2019) 18:1242–54. doi: 10.1074/mcp.RA118.001288
- Geyer PE, Wewer Albrechtsen NJ, Tyanova S, Grassl N, Iepsen EW, Lundgren J, et al. Proteomics reveals the effects of sustained weight loss on the human plasma proteome. *Mol Syst Biol* (2016) 12:901. doi: 10.15252/msb.20167357
- Muntel J, Gandhi T, Verbeke L, Bernhardt OM, Treiber T, Bruderer R, et al. Surpassing 10 000 identified and quantified proteins in a single run by optimizing current LC-MS instrumentation and data analysis strategy. *Mol Omics* (2019) 15:348–60. doi: 10.1039/C9MO00082H
- Kawashima Y, Nagai H, Konno R, Ishikawa M, Nakajima D, Sato H, et al. Single-shot 10K proteome approach: over 10,000 protein identifications by data-independent acquisition-based single-shot proteomics with ion mobility spectrometry. *J Proteome Res* (2022) 21:1418–27. doi: 10.1021/acs.jproteome.2c00023
- Mantovani A, Garlanda C. Humoral innate immunity and acute-phase proteins. *New Engl J Med* (2023) 388:439–52. doi: 10.1056/NEJMra2206346
- Yu JC, Khodadadi H, Malik A, Davidson B, Salles É da SL, Bhatia J, et al. Innate immunity of neonates and infants. *Front Immunol* (2018) 9:1759. doi: 10.3389/fimmu.2018.01759
- Lang KS, Burov A, Kurrer M, Lang PA, Recher M. The role of the innate immune response in autoimmune disease. *J Autoimmun* (2007) 29:206–12. doi: 10.1016/j.jaut.2007.07.018
- Zouali M, La Cava A. Editorial: innate immunity pathways in autoimmune diseases. *Front Immunol* (2019) 10:1245. doi: 10.3389/fimmu.2019.01245
- Masucci MT, Minopoli M, Del Vecchio S, Carriero MV. The emerging role of neutrophil extracellular traps (NETs) in tumor progression and metastasis. *Front Immunol* (2020) 11:1749. doi: 10.3389/fimmu.2020.01749
- Getz GS. Bridging the innate and adaptive immune systems. *J Lipid Res* (2005) 46:619–22. doi: 10.1194/jlr.E500002-JLR200
- Mackay IR, Rosen FS, Delves PJ, Roitt IM. The immune system: First of two parts. *New Engl J Med* (2000) 343:37–49. doi: 10.1056/NEJM200010123431520
- Mackay IR, Rosen FS, Delves PJ, Roitt IM. The immune system: Second of two parts. *New Engl J Med* (2000) 343:108–17. doi: 10.1056/NEJM200007133430207
- Ricklin D, Hajishengallis G, Yang K, Lambris JD. Complement - a key system for immune surveillance and homeostasis. *Nat Immunol* (2010) 11:785–97. doi: 10.1038/ni.1923
- Bohlsón SS, Garred P, Kemper C, Tenner AJ. Complement nomenclature—Deconvoluted. *Front Immunol* (2019) 10:1308. doi: 10.3389/fimmu.2019.01308
- Scientific Image and Illustration Software*. BioRender. Available at: <https://www.biorender.com/>.
- Gaboriaud C, Thielens NM, Gregory LA, Rossi V, Fontecilla-Camps JC, Arlaud GJ. Structure and activation of the C1 complex of complement: unraveling the puzzle. *Trends Immunol* (2004) 25:368–73. doi: 10.1016/j.it.2004.04.008
- Vidarsson G, Dekkers G, Rispens T. IgG subclasses and allotypes: from structure to effector functions. *Front Immunol* (2014) 5:520. doi: 10.3389/fimmu.2014.00520
- Matsushita M, Endo Y, Fujita T. Structural and functional overview of the lectin complement pathway: its molecular basis and physiological implication. *Arch Immunol Ther Exp* (2013) 61:273–83. doi: 10.1007/s00005-013-0229-y
- Kjaer TR, Thiel S, Andersen GR. Toward a structure-based comprehension of the lectin pathway of complement. *Mol Immunol* (2013) 56:222–31. doi: 10.1016/j.molimm.2013.05.220
- Thiel S, Vorup-Jensen T, Stover CM, Schwaebel W, Laursen SB, Poulsen K, et al. A second serine protease associated with mannan-binding lectin that activates complement. *Nature* (1997) 386:506–10. doi: 10.1038/386506a0
- Matsushita M, Thiel S, Jensenius JC, Terai I, Fujita T. Proteolytic activities of two types of mannan-binding lectin-associated serine protease. *J Immunol* (2000) 165:2637–42. doi: 10.4049/jimmunol.165.5.2637

Acknowledgments

I am grateful for guidance and inputs from Gunna Christiansen, Svend Birkelund, Hanno Steen, Vibeke Andersen, and Kathrine Bjerg Bennike.

Conflict of interest

The author declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

27. Lachmann PJ. Chapter 4 - the amplification loop of the complement pathways. In: Alt FW, editor. *Advances in Immunology. Advances in Immunology*. (2009) (Amsterdam: Academic Press), p. 115–49. doi: 10.1016/S0065-2776(08)04004-2
28. Fearon DT, Austen KF. Properdin: binding to C3b and stabilization of the C3b-dependent C3 convertase. *J Exp Med* (1975) 142:856–63. doi: 10.1084/jem.142.4.856
29. Merle NS, Church SE, Fremeaux-Bacchi V, Roumenina LT. Complement system part I – molecular mechanisms of activation and regulation. *Front Immunol* (2015) 6:262. doi: 10.3389/fimmu.2015.00262
30. Opstrup KV, Bennike TB, Christiansen G, Birkelund S. Complement killing of clinical *Klebsiella pneumoniae* isolates is serum concentration dependent. *Microbes Infection* (2023) 25:105074. doi: 10.1016/j.micinf.2022.105074
31. Bajic G, Degn SE, Thiel S, Andersen GR. Complement activation, regulation, and molecular basis for complement-related diseases. *EMBO J* (2015) 34:2735–57. doi: 10.15252/embj.201591881
32. Bouti P, Webbers SDS, Fagerholm SC, Alon R, Moser M, Matlung HL, et al. $\beta 2$ integrin signaling cascade in neutrophils: more than a single function. *Front Immunol* (2021) 11:619925. doi: 10.3389/fimmu.2020.619925
33. Yuen J, Pluthero FG, Douda DN, Riedl M, Cherry A, Ulanova M, et al. NETosing neutrophils activate complement both on their own NETs and bacteria via alternative and non-alternative pathways. *Front Immunol* (2016) 7:137. doi: 10.3389/fimmu.2016.00137
34. Branzk N, Lubojemska A, Hardison SE, Wang Q, Gutierrez MG, Brown GD, et al. Neutrophils sense microbe size and selectively release neutrophil extracellular traps in response to large pathogens. *Nat Immunol* (2014) 15:1017–25. doi: 10.1038/ni.2987
35. Brinkmann V. Neutrophil extracellular traps in the second decade. *J Innate Immun* (2018) 10:414–21. doi: 10.1159/000489829
36. Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, et al. Neutrophil extracellular traps kill bacteria. *Science* (2004) 303:1532–5. doi: 10.1126/science.1092385
37. Spengler J, Lugonja B, Ytterberg AJ, Zubarev RA, Creese AJ, Pearson MJ, et al. Release of active peptidyl arginine deiminases by neutrophils can explain production of extracellular citrullinated autoantigens in rheumatoid arthritis synovial fluid. *Arthritis Rheumatol* (2015) 67:3135–45. doi: 10.1002/art.39313
38. Bennike T, Lauridsen KB, Olesen MK, Andersen V, Birkelund S, Stensballe A. Optimizing the identification of citrullinated peptides by mass spectrometry: utilizing the inability of trypsin to cleave after citrullinated amino acids. *J Proteomics Bioinform* (2013) 6:12. doi: 10.4172/jpb.1000293
39. Fuchs TA, Abed U, Goosmann C, Hurwitz R, Schulze I, Wahn V, et al. Novel cell death program leads to neutrophil extracellular traps. *J Cell Biol* (2007) 176:231–41. doi: 10.1083/jcb.200606027
40. Papayannopoulos V, Metzler KD, Hakkim A, Zychlinsky A. Neutrophil elastase and myeloperoxidase regulate the formation of neutrophil extracellular traps. *J Cell Biol* (2010) 191:677–91. doi: 10.1083/jcb.201006052
41. Metzler KD, Fuchs TA, Nauseef WM, Reumaux D, Roesler J, Schulze I, et al. Myeloperoxidase is required for neutrophil extracellular trap formation: implications for innate immunity. *Blood* (2011) 117:953–9. doi: 10.1182/blood-2010-06-290171
42. Yipp BG, Petri B, Salina D, Jenne CN, Scott BNV, Zbytniuk LD, et al. Dynamic NETosis is carried out by live neutrophils in human and mouse bacterial abscesses and during severe gram-positive infection. *Nat Med* (2012) 18:1386–93. doi: 10.1038/nm.2847
43. O'Donoghue AJ, Jin Y, Knudsen GM, Perera NC, Jenne DE, Murphy JE, et al. Global substrate profiling of proteases in human neutrophil extracellular traps reveals consensus motif predominantly contributed by elastase. *PLoS One* (2013) 8:e75141. doi: 10.1371/journal.pone.0075141
44. Morley JJ, Kushner I. Serum C-reactive protein levels in disease. *Ann N Y Acad Sci* (1982) 389:406–18. doi: 10.1111/j.1749-6632.1982.tb22153.x
45. Castell JV, Gómez-Lechón MJ, David M, Andus T, Geiger T, Trullenque R, et al. Interleukin-6 is the major regulator of acute phase protein synthesis in adult human hepatocytes. *FEBS Lett* (1989) 242:237–9. doi: 10.1016/0014-5793(89)80476-4
46. Pepys MB, Baltz ML. Acute phase proteins with special reference to C-reactive protein and related proteins (pentaxins) and serum amyloid A protein. *Adv Immunol* (1983) 34:141–212. doi: 10.1016/s0065-2776(08)60379-x
47. Agrawal A, Volanakis JE. Probing the C1q-binding site on human C-reactive protein by site-directed mutagenesis. *J Immunol* (1994) 152:5404–10. doi: 10.4049/jimmunol.152.11.5404
48. Agrawal A, Shrive AK, Greenhough TJ, Volanakis JE. Topology and structure of the C1q-binding site on C-reactive protein1. *J Immunol* (2001) 166:3998–4004. doi: 10.4049/jimmunol.166.6.3998
49. Bodman-Smith KB, Melendez AJ, Campbell I, Harrison PT, Allen JM, Raynes JG. C-reactive protein-mediated phagocytosis and phospholipase D signalling through the high-affinity receptor for immunoglobulin G (Fc γ RI). *Immunology* (2002) 107:252–60. doi: 10.1046/j.1365-2567.2002.01481.x
50. Escadafal C, Incardona S, Fernandez-Carballo BL, Dittrich S. The good and the bad: using C reactive protein to distinguish bacterial from non-bacterial infection among febrile patients in low-resource settings. *BMJ Global Health* (2020) 5:e002396. doi: 10.1136/bmjgh-2020-002396
51. El Ghmati SM, Van Hoeyveld EM, Van Strijp JG, Ceuppens JL, Stevens EA. Identification of haptoglobin as an alternative ligand for CD11b/CD18. *J Immunol* (1996) 156:2542–52. doi: 10.4049/jimmunol.156.7.2542
52. MacKellar M, Vigerust DJ. Role of haptoglobin in health and disease: A focus on diabetes. *Clin Diabetes* (2016) 34:148–57. doi: 10.2337/diaclin.34.3.148
53. Bennike TB, Fatou B, Angelidou A, Diray-Arce J, Falsafi R, Ford R, et al. Preparing for life: plasma proteome changes and immune system development during the first week of human life. *Front Immunol* (2020) 11:578505. doi: 10.3389/fimmu.2020.578505
54. Anderson NL. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* (2002) 1:845–67. doi: 10.1074/mcp.R200007-MCP200
55. Garay-Baquero DJ, White CH, Walker NF, Tebruegge M, Schiff HF, Ugarte-Gil C, et al. Comprehensive plasma proteomic profiling reveals biomarkers for active tuberculosis. *JCI Insight* (2020) 5:e137427. doi: 10.1172/jci.insight.137427
56. Keshishian H, Burgess MW, Gillette MA, Mertins P, Clauser KR, Mani DR, et al. Multiplexed, quantitative workflow for sensitive biomarker discovery in plasma yields novel candidates for early myocardial injury. *Mol Cell Proteomics* (2015) 14:2375–93. doi: 10.1074/mcp.M114.046813
57. Kartikasari AER, Huertas CS, Mitchell A, Plebanski M. Tumor-induced inflammatory cytokines and the emerging diagnostic devices for cancer detection and prognosis. *Front Oncol* (2021) 11:692142. doi: 10.3389/fonc.2021.692142
58. Kaur G, Poljak A, Ali SA, Zhong L, Raftery MJ, Sachdev P. Extending the depth of human plasma proteome coverage using simple fractionation techniques. *J Proteome Res* (2021) 20:1261–79. doi: 10.1021/acs.jproteome.0c00670
59. Kaufmann A, Walker S. Comparison of linear intrascan and interscan dynamic ranges of Orbitrap and ion-mobility time-of-flight mass spectrometers. *Rapid Commun Mass Spectrometry* (2017) 31:1915–26. doi: 10.1002/rcm.7981
60. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* (2003) 422:198–207. doi: 10.1038/nature01511
61. Steen H, Mann M. The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol* (2004) 5:699–711. doi: 10.1038/nrm1468
62. Burger B, Vaudel M, Barsnes H. Importance of block randomization when designing proteomics experiments. *J Proteome Res* (2021) 20:122–8. doi: 10.1021/acs.jproteome.0c00536
63. Shuken SR. An introduction to mass spectrometry-based proteomics. *J Proteome Res* (2023) 22:2151–71. doi: 10.1021/acs.jproteome.2c00838
64. Yen C-Y, Russell S, Mendoza AM, Meyer-Arendt K, Sun S, Cios KJ, et al. Improving sensitivity in shotgun proteomics using a peptide-centric database with reduced complexity: protease cleavage and SCX elution rules from data mining of MS/MS spectra. *Anal Chem* (2006) 78:1071–84. doi: 10.1021/ac051127f
65. Bennike T, Ayturk U, Haslauer CM, Froehlich JW, Proffen BL, Barnaby O, et al. A normative study of the synovial fluid proteome from healthy porcine knee joints. *J Proteome Res* (2014) 13:4377–87. doi: 10.1021/pr500587x
66. Leon IR, Schwammle V, Jensen ON, Sprenger RR. Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. *Mol Cell Proteomics* (2013) 12:2992–3005. doi: 10.1074/mcp.M112.025585
67. Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal Chem* (1996) 68:850–8. doi: 10.1021/ac950914h
68. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* (2006) 1:2856–60. doi: 10.1038/nprot.2006.468
69. Manza LL, Stamer SL, Ham A-JL, Codreanu SG, Liebler DC. Sample preparation and digestion for proteomic analyses using spin filters. *Proteomics* (2005) 5:1742–5. doi: 10.1002/pmic.200401063
70. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nat Methods* (2009) 6:359–62. doi: 10.1038/nmeth.1322
71. Kulak NA, Pichler G, Paron I, Nagaraj N, Mann M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods* (2014) 11:319–24. doi: 10.1038/nmeth.2834
72. Zougman A, Selby PJ, Banks RE. Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *PROTEOMICS* (2014) 14:1006–0. doi: 10.1002/pmic.201300553
73. Hughes CS, Foehr S, Garfield DA, Furlong EE, Steinmetz LM, Krijgsveld J. Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol Syst Biol* (2014) 10:1–10. doi: 10.15252/msb.20145625
74. Hughes CS, Moggridge S, Müller T, Sorensen PH, Morin GB, Krijgsveld J. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* (2019) 14:68–85. doi: 10.1038/s41596-018-0082-x
75. Sielaff M, Kuharev J, Bohn T, Hahlbrock J, Bopp T, Tenzer S, et al. Evaluation of FASP, SP3, and iST protocols for proteomic sample preparation in the low microgram range. *J Proteome Res* (2017) 16:4060–72. doi: 10.1021/acs.jproteome.7b00433
76. Geyer PE, Kulak NA, Pichler G, Holdt LM, Teupser D, Mann M. Plasma proteome profiling to assess human health and disease. *Cell Syst* (2016) 2:185–95. doi: 10.1016/j.cels.2016.02.015

77. Niu L, Thiele M, Geyer PE, Rasmussen DN, Weibel HE, Santos A, et al. Noninvasive proteomic biomarkers for alcohol-related liver disease. *Nat Med* (2022) 28:1277–87. doi: 10.1038/s41591-022-01850-y
78. Viode A, van Zalm P, Smolen KK, Fatou B, Stevenson D, Jha M, et al. A simple, time- and cost-effective, high-throughput depletion strategy for deep plasma proteomics. *Sci Adv* (2023) 9:ead9717. doi: 10.1126/sciadv.adf9717
79. Switzer L, van Angeren J, Pinkse M, Kool J, Niessen WMA. A high-throughput sample preparation method for cellular proteomics using 96-well filter plates. *Proteomics* (2013) 13:2980–3. doi: 10.1002/pmic.201300080
80. Bennike TB, Steen H. High-throughput parallel proteomic sample preparation using 96-well polyvinylidene fluoride (PVDF) membranes and C18 purification plates. In: Greening DW, Simpson RJ, editors. *Serum/plasma Proteomics: Methods and Protocols. Methods in Molecular Biology*. New York, NY: Springer (2017). p. 395–402. doi: 10.1007/978-1-4939-7057-5_27
81. Boada P, Fatou B, Belperron AA, Sigdel TK, Smolen KK, Wurie Z, et al. Longitudinal serum proteomics analyses identify unique and overlapping host response pathways in Lyme disease and West Nile virus infection. *Front Immunol* (2022) 13:1012824. doi: 10.3389/fimmu.2022.1012824
82. Lee AH, Shannon CP, Amenyogbe N, Bennike TB, Diray-Arce J, Idoko OT, et al. Dynamic molecular changes during the first week of human life follow a robust developmental trajectory. *Nat Commun* (2019) 10:1092. doi: 10.1038/s41467-019-08794-x
83. Müller T, Kalxdorf M, Longuespée R, Kazdal DN, Stenzinger A, Krijgsveld J. Automated sample preparation with SP3 for low-input clinical proteomics. *Mol Syst Biol* (2020) 16:e9111. doi: 10.15252/msb.20199111
84. Ding H, Fazelinia H, Spruce LA, Weiss DA, Zderic SA, Seeholzer SH. Urine proteomics: evaluation of different sample preparation workflows for quantitative, reproducible, and improved depth of analysis. *J Proteome Res* (2020) 19:1857–62. doi: 10.1021/acs.jproteome.9b00772
85. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. *Science* (1989) 246:64–71. doi: 10.1126/science.2675315
86. Roepstorff P, Fohlman J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *BioMed Mass Spectrom* (1984) 11:601. doi: 10.1002/bms.1200111109
87. Hunt DF, Henderson RA, Shabanowitz J, Sakaguchi K, Michel H, Sevilir N, et al. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* (1992) 255:1261–3. doi: 10.1126/science.1546328
88. Kelstrup CD, Bekker-Jensen DB, Arrey TN, Hogrebe A, Harder A, Olsen JV. Performance evaluation of the Q exactive HF-X for shotgun proteomics. *J Proteome Res* (2018) 17:727–38. doi: 10.1021/acs.jproteome.7b00602
89. Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* (1994) 5:976–89. doi: 10.1016/1044-0305(94)80016-2
90. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: A peptide search engine integrated into the maxQuant environment. *J Proteome Res* (2011) 10:1794–805. doi: 10.1021/pr101065j
91. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed maxLFQ. *Mol Cell Proteomics* (2014) 13:2513–26. doi: 10.1074/mcp.M113.031591
92. Liu Y, Hüttenhain R, Surinova S, Gillet LCJ, Mouritsen J, Brunner R, et al. Quantitative measurements of N-linked glycoproteins in human plasma by SWATH-MS. *Proteomics* (2013) 13:1247–56. doi: 10.1002/pmic.201200417
93. Gotti C, Roux-Dalvai F, Joly-Beauparlant C, Mangnier L, Leclercq M, Droit A. Extensive and accurate benchmarking of DIA acquisition methods and software tools using a complex proteomic standard. *J Proteome Res* (2021) 20:4801–14. doi: 10.1021/acs.jproteome.1c00490
94. Muntel J, Xuan Y, Berger ST, Reiter L, Bachur R, Kentsis A, et al. Advancing urinary protein biomarker discovery by data-independent acquisition on a quadrupole-orbitrap mass spectrometer. *J Proteome Res* (2015) 14:4752–62. doi: 10.1021/acs.jproteome.5b00826
95. Ludwig C, Gillet L, Rosenberger G, Amon S, Collins BC, Aebersold R. Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol Syst Biol* (2018) 14:e8126. doi: 10.15252/msb.20178126
96. Li KW, Gonzalez-Lozano MA, Koopmans F, Smit AB. Recent developments in data independent acquisition (DIA) mass spectrometry: application of quantitative analysis of the brain proteome. *Front Mol Neurosci* (2020) 0:564446. doi: 10.3389/fnmol.2020.564446
97. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* (2020) 17:41–4. doi: 10.1038/s41592-019-0638-x
98. Sinitcyn P, Hamzeiy H, Salinas Soto F, Itzhak D, McCarthy F, Wichmann C, et al. MaxDIA enables library-based and library-free data-independent acquisition proteomics. *Nat Biotechnol* (2021) 39:1563–73. doi: 10.1038/s41587-021-00968-7
99. Pino LK, Just SC, MacCoss MJ, Searle BC. Acquiring and analyzing data independent acquisition proteomics experiments without spectrum libraries. *Mol Cell Proteomics* (2020) 19:1088–103. doi: 10.1074/mcp.P119.001913
100. Ebadi A, Freestone J, Noble WS, Keich U. Bridging the false discovery gap. *J Proteome Res* (2023) 22:2172–8. doi: 10.1021/acs.jproteome.3c00176
101. Käll L, Canterbury JD, Weston J, Noble WS, MacCoss MJ. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat Methods* (2007) 4:923–5. doi: 10.1038/nmeth1113
102. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* (2007) 4:207–14. doi: 10.1038/nmeth1019
103. Taverna D, Gaspari M. A critical comparison of three MS-based approaches for quantitative proteomics analysis. *J Mass Spectrometry* (2021) 56:e4669. doi: 10.1002/jms.4669
104. Li Z, Adams RM, Chourey K, Hurst GB, Hettich RL, Pan C. Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ orbitrap velos. *J Proteome Res* (2012) 11:1582–90. doi: 10.1021/pr200748h
105. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* (2007) 8:118–27. doi: 10.1093/biostatistics/kjx037
106. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Ser B (Methodological)* (1995) 57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
107. Camargo A, Azuaje F, Wang H, Zheng H. Permutation – based statistical tests for multiple hypotheses. *Source Code Biol Med* (2008) 3:15. doi: 10.1186/1751-0473-3-15
108. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* (2016) 13:731–40. doi: 10.1038/nmeth.3901
109. Jeanmougin M, de Reynies A, Marisa L, Paccard C, Nuel G, et al. Should we abandon the t-test in the analysis of gene expression microarray data: A comparison of variance modeling strategies. *PLoS One* (2010) 5:e12336. doi: 10.1371/journal.pone.0012336
110. Pursiheimo A, Vehmas AP, Afzal S, Suomi T, Chand T, Strauss L, et al. Optimization of statistical methods impact on quantitative proteomics data. *J Proteome Res* (2015) 14:4118–26. doi: 10.1021/acs.jproteome.5b00183
111. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res* (2015) 43:e47. doi: 10.1093/nar/gkv007
112. Daly DS, Anderson KK, Panisko EA, Purvine SO, Fang R, Monroe ME, et al. Mixed-effects statistical model for comparative LC-MS proteomics studies. *J Proteome Res* (2008) 7:1209–17. doi: 10.1021/pr070441i
113. Klann K, Münch C. PBLMM: Peptide-based linear mixed models for differential expression analysis of shotgun proteomics data. *J Cell Biochem* (2022) 123:691–6. doi: 10.1002/jcb.30225
114. Kirov S, Sasson A, Zhang C, Chasalow S, Dongre A, Steen H, et al. Degradation of the extracellular matrix is part of the pathology of ulcerative colitis. *Mol Omics* (2019) 15:67–76. doi: 10.1039/c8mo00239h
115. Bai M, Deng J, Dai C, Pfeuffer J, Sachsenberg T, Perez-Riverol Y. LFQ-based peptide and protein intensity differential expression analysis. *J Proteome Res* (2023) 22:2114–23. doi: 10.1021/acs.jproteome.2c00812
116. Geyer PE, Voytik E, Treit PV, Doll S, Kleinhempel A, Niu L, et al. Plasma Proteome Profiling to detect and avoid sample-related biases in biomarker studies. *EMBO Mol Med* (2019) 0:e10427. doi: 10.15252/emmm.201910427
117. Pieper R, Su Q, Gatlin CL, Huang S-T, Anderson NL, Steiner S. Multi-component immunoaffinity subtraction chromatography: an innovative step towards a comprehensive survey of the human plasma proteome. *Proteomics* (2003) 3:422–32. doi: 10.1002/pmic.200390057
118. Tu C, Rudnick PA, Martinez MY, Cheek KL, Stein SE, Slebos RJC, et al. Depletion of abundant plasma proteins and limitations of plasma proteomics. *J Proteome Res* (2010) 9:4982–91. doi: 10.1021/pr100646w
119. Smith MPW, Wood SL, Zougman A, Ho JTC, Peng J, Jackson D, et al. A systematic analysis of the effects of increasing degrees of serum immunodepletion in terms of depth of coverage and other key aspects in top-down and bottom-up proteomic analyses. *PROTEOMICS* (2011) 11:2222–35. doi: 10.1002/pmic.201100005
120. Roche S, Tiers L, Provansal M, Seveno M, Piva M-T, Jouin P, et al. Depletion of one, six, twelve or twenty major blood proteins before proteomic analysis: The more the better? *J Proteomics* (2009) 72:945–51. doi: 10.1016/j.jprot.2009.03.008
121. Zhou M, Lucas DA, Chan KC, Issaq HJ, Petricoin EF, Liotta LA, et al. An investigation into the human serum “interactome”. *ELECTROPHORESIS* (2004) 25:1289–98. doi: 10.1002/elps.200405866
122. Cao Z, Tang H-Y, Wang H, Liu Q, Speicher DW. Systematic comparison of fractionation methods for in-depth analysis of plasma proteomes. *J Proteome Res* (2012) 11:3090–100. doi: 10.1021/pr201068h
123. Furumaya C, Martinez-Sanz P, Bouti P, Kuijpers TW, Matlung HL. Plasticity in pro- and anti-tumor activity of neutrophils: shifting the balance. *Front Immunol* (2020) 11:2100. doi: 10.3389/fimmu.2020.02100
124. Alfaro C, Teixeira A, Oñate C, Pérez G, Sanmamed MF, Andueza MP, et al. Tumor-produced interleukin-8 attracts human myeloid-derived suppressor cells and

- elicits extrusion of neutrophil extracellular traps (NETs). *Clin Cancer Res* (2016) 22:3924–36. doi: 10.1158/1078-0432.CCR-15-2463
125. Kollmann TR, Kampmann B, Mazmanian SK, Marchant A, Levy O. Protecting the newborn and young infant from infectious diseases: lessons from immune ontogeny. *Immunity* (2017) 46:350–63. doi: 10.1016/j.immuni.2017.03.009
126. Balbus JM, Barouki R, Birnbaum LS, Etzel RA, Gluckman SPD, Grandjean P, et al. Early-life prevention of non-communicable diseases. *Lancet* (2013) 381:3–4. doi: 10.1016/S0140-6736(12)61609-2
127. Fouda GG, Martinez DR, Swamy GK, Permar SR. The Impact of IgG transplacental transfer on early life immunity. *Immunohorizons* (2018) 2:14–25. doi: 10.4049/immunohorizons.1700057
128. Van de Perre P. Transfer of antibody via mother's milk. *Vaccine* (2003) 21:3374–6. doi: 10.1016/S0264-410X(03)00336-0
129. Hanson LA. Breastfeeding provides passive and likely long-lasting active immunity. *Ann Allergy Asthma Immunol* (1998) 81:523–33. doi: 10.1016/S1081-1206(10)62704-4
130. Mantis NJ, Rol N, Corthésy B. Secretory IgA's complex roles in immunity and mucosal homeostasis in the gut. *Mucosal Immunol* (2011) 4:603–11. doi: 10.1038/mi.2011.41
131. McGreal EP, Hearne K, Spiller OB. Off to a slow start: Under-development of the complement system in term newborns is more substantial following premature birth. *Immunobiology* (2012) 217:176–86. doi: 10.1016/j.imbio.2011.07.027
132. Howie SRC. Blood sample volumes in child health research: review of safe limits. *Bull World Health Organ* (2011) 89:46–53. doi: 10.2471/BLT.10.080010
133. Jennewein MF, Butler AL, Alter G. Neonate-omics: charting the unknown immune response in early life. *Cell* (2018) 174:1051–3. doi: 10.1016/j.cell.2018.08.001
134. Morgan BP, Gasque P. Extrahepatic complement biosynthesis: where, when and why? *Clin Exp Immunol* (1997) 107:1–7. doi: 10.1046/j.1365-2249.1997.d01-890.x
135. Kilpatrick DC, Liston WA, Midgley PC. Mannan-binding protein in human umbilical cord blood. *Nat Immunol* (1996) 15:234–40.
136. Pettengill MA, van Haren SD, Levy O. Soluble mediators regulating immunity in early life. *Front Immunol* (2014) 5:457. doi: 10.3389/fimmu.2014.00457
137. Sallenbach S, Thiel S, Aebi C, Otth M, Bigler S, Jensenius JC, et al. Serum concentrations of lectin-pathway components in healthy neonates, children and adults: mannan-binding lectin (MBL), M-, L-, and H-ficolin, and MBL-associated serine protease-2 (MASP-2). *Pediatr Allergy Immunol* (2011) 22:424–30. doi: 10.1111/j.1399-3038.2010.01104.x
138. Grumach AS, Ceccon ME, Rutz R, Fertig A, Kirschfink M. Complement profile in neonates of different gestational ages. *Scandinavian J Immunol* (2014) 79:276–81. doi: 10.1111/sji.12154
139. Sellmer A, Henriksen TB, Palmfeldt J, Bech BH, Astono J, Bennike TB, et al. The patent ductus arteriosus in extremely preterm neonates is more than a hemodynamic challenge: new molecular insights. *Biomolecules* (2022) 12:1179. doi: 10.3390/biom12091179
140. Bjelosevic S, Pascovici D, Ping H, Karlaftis V, Zaw T, Song X, et al. Quantitative age-specific variability of plasma proteins in healthy neonates, children and adults*. *Mol Cell Proteomics* (2017) 16:924–35. doi: 10.1074/mcp.M116.066720
141. Strunk RC, Fenton LJ, Gaines JA. Alternative pathway of complement activation in full term and premature infants. *Pediatr Res* (1979) 13:641–3. doi: 10.1203/00006450-197905000-00013
142. Morell A, Terry WD, Waldmann TA. Metabolic properties of IgG subclasses in man. *J Clin Invest* (1970) 49:673–80. doi: 10.1172/JCI106279
143. Lozano NA, Lozano A, Marini V, Saranz RJ, Blumberg RS, Baker K, et al. Expression of FcRn receptor in placental tissue and its relationship with IgG levels in term and pre-term newborns. *Am J Reprod Immunol* (2018) 80:e12972. doi: 10.1111/ajr.12972
144. Bindon CI, Hale G, Brüggemann M, Waldmann H. Human monoclonal IgG isotypes differ in complement activating function at the level of C4 as well as C1q. *J Exp Med* (1988) 168:127–42. doi: 10.1084/jem.168.1.127
145. Clements T, Rice TF, Vamvakas G, Barnett S, Barnes M, Donaldson B, et al. Update on transplacental transfer of IgG subclasses: impact of maternal and fetal factors. *Front Immunol* (2020) 11:1920. doi: 10.3389/fimmu.2020.01920
146. Aksu G, Genel F, Koturoğlu G, Kurugöl Z, Kütükçüler N. Serum immunoglobulin (IgG, IgM, IgA) and IgG subclass concentrations in healthy children: a study using nephelometric technique. *Turk J Pediatr* (2006) 48:19–24.
147. Lu J, Kishore U. C1 complex: an adaptable proteolytic module for complement and non-complement functions. *Front Immunol* (2017) 8:592. doi: 10.3389/fimmu.2017.00592
148. Zilow G, Zilow EP, Burger R, Linderkamp O. Complement activation in newborn infants with early onset infection. *Pediatr Res* (1993) 34:199–203. doi: 10.1203/00006450-199308000-00020
149. Cools-Lartigue J, Spicer J, McDonald B, Gowing S, Chow S, Giannias B, et al. Neutrophil extracellular traps sequester circulating tumor cells and promote metastasis. *J Clin Invest* (2013) 123:3446–58. doi: 10.1172/JCI67484
150. Chapman EA, Lyon M, Simpson D, Mason D, Beynon RJ, Moots RJ, et al. Caught in a trap? Proteomic analysis of neutrophil extracellular traps in rheumatoid arthritis and systemic lupus erythematosus. *Front Immunol* (2019) 10:423. doi: 10.3389/fimmu.2019.00423
151. Liu X, Arfman T, Wichapong K, Reutelingsperger CPM, Voorberg J, Nicolaes GAF. PAD4 takes charge during neutrophil activation: Impact of PAD4 mediated NET formation on immune-mediated disease. *J Thromb Haemost* (2021) 19:1607–17. doi: 10.1111/jth.15313
152. Li M, Lin C, Leso A, Nefedova Y. Quantification of citrullinated histone H3 bound DNA for detection of neutrophil extracellular traps. *Cancers (Basel)* (2020) 12:3424. doi: 10.3390/cancers12113424
153. Weyand CM, Goronzy JJ. The immunology of rheumatoid arthritis. *Nat Immunol* (2021) 22:10–8. doi: 10.1038/s41590-020-00816-x
154. Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. *Lancet* (2016) 388:2023–38. doi: 10.1016/S0140-6736(16)30173-8
155. Darrach E, Rosen A, Giles JT, Andrade F. Peptidylarginine deiminase 2, 3 and 4 have distinct specificities against cellular substrates: novel insights into autoantigen selection in rheumatoid arthritis. *Ann Rheum Dis* (2012) 71:92–8. doi: 10.1136/ard.2011.151712
156. van Beers JJBC, Schwarte CM, Stammen-Vogelzangs J, Oosterink E, Božič B, Puijntjens GJM. The rheumatoid arthritis synovial fluid citrullinome reveals novel citrullinated epitopes in apolipoprotein E, myeloid nuclear differentiation antigen, and β -actin. *Arthritis Rheumatism* (2013) 65:69–80. doi: 10.1002/art.37720
157. Holmes CL, Shim D, Kernien J, Johnson CJ, Nett JE, Shelef MA. Insight into Neutrophil Extracellular Traps through Systematic Evaluation of Citrullination and Peptidylarginine Deiminases. *J Immunol* (2019) 2019:2160192. doi: 10.1155/2019/2160192
158. Birkelund S, Bennike TB, Kastaniegaard K, Lausen M, Poulsen TBG, Kragstrup TW, et al. Proteomic analysis of synovial fluid from rheumatic arthritis and spondyloarthritis patients. *Clin Proteomics* (2020) 17:29. doi: 10.1186/s12014-020-09292-9
159. Ytterberg AJ, Joshua V, Reynisdottir G, Tarasova NK, Rutishauser D, Ossipova E, et al. Shared immunological targets in the lungs and joints of patients with rheumatoid arthritis: identification and validation. *Ann Rheumatic Dis* (2015) 74:1772–7. doi: 10.1136/annrheumdis-2013-204912
160. Makrygiannakis D, Hermansson M, Ulfgrén A-K, Nicholas AP, Zandman AJW, Eklund A, et al. Smoking increases peptidylarginine deiminase 2 enzyme expression in human lungs and increases citrullination in BAL cells. *Ann Rheum Dis* (2008) 67:1488–92. doi: 10.1136/ard.2007.075192
161. Holers VM, Demoruelle MK, Kuhn KA, Buckner JH, Robinson WH, Okamoto Y, et al. Rheumatoid arthritis and the mucosal origins hypothesis: external protection becomes internal destruction. *Nat Rev Rheumatol* (2018) 14:542–57. doi: 10.1038/s41584-018-0070-0
162. Bennike TB, Ellingsen T, Glerup H, Bonderup OK, Carlsen TG, Meyer MK, et al. Proteome analysis of rheumatoid arthritis gut mucosa. *J Proteome Res* (2017) 16:346–54. doi: 10.1021/acs.jproteome.6b00598
163. Pratesi F, Dioni I, Tommasi C, Alcaro MC, Paolini I, Barbetti F, et al. Antibodies from patients with rheumatoid arthritis target citrullinated histone 4 contained in neutrophils extracellular traps. *Ann Rheum Dis* (2014) 73:1414–22. doi: 10.1136/annrheumdis-2012-202765
164. Alatab S, Sepanlou SG, Ikuta K, Vahedi H, Bisignano C, Safiri S, et al. The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* (2020) 5:17–30. doi: 10.1016/S2468-1253(19)30333-4
165. Kobayashi T, Siegmund B, Le Berre C, Wei SC, Ferrante M, Shen B, et al. Ulcerative colitis. *Nat Rev Dis Primers* (2020) 6:1–20. doi: 10.1038/s41572-020-0205-x
166. Chang JT. Pathophysiology of inflammatory bowel diseases. *N Engl J Med* (2020) 383:2652–64. doi: 10.1056/NEJMr2002697
167. Khan I, Ullah N, Zha L, Bai Y, Khan A, Zhao T, et al. Alteration of gut microbiota in inflammatory bowel disease (IBD): cause or consequence? IBD treatment targeting the gut microbiome. *Pathogens* (2019) 8:126. doi: 10.3390/pathogens8030126
168. Dinallo V, Marafini I, Di Fusco D, Laudisi F, Franzè E, Di Grazia A, et al. Neutrophil extracellular traps sustain inflammatory signals in ulcerative colitis. *J Crohns Colitis* (2019) 13:772–84. doi: 10.1093/ecco-jcc/jjy215
169. Li T, Wang C, Liu Y, Li B, Zhang W, Wang L, et al. Neutrophil extracellular traps induce intestinal damage and thrombotic tendency in inflammatory bowel disease. *J Crohn's Colitis* (2020) 14:240–53. doi: 10.1093/ecco-jcc/jjz132
170. Schroder AL, Chami B, Liu Y, Doyle CM, El Kazzi M, Ahlenstiel G, et al. Neutrophil extracellular trap density increases with increasing histopathological severity of crohn's disease. *Inflammation Bowel Dis* (2022) 28:586–98. doi: 10.1093/ibd/izab239
171. Yipp BG, Kubes P. NETosis: how vital is it? *Blood* (2013) 122:2784–94. doi: 10.1182/blood-2013-04-457671
172. Papayannopoulos V. Neutrophil extracellular traps in immunity and disease. *Nat Rev Immunol* (2018) 18:134–47. doi: 10.1038/nri.2017.105
173. Farrera C, Fadeel B. Macrophage clearance of neutrophil extracellular traps is a silent process. *J Immunol* (2013) 191:2647–56. doi: 10.4049/jimmunol.1300436
174. Hakkim A, Füllrohr BG, Amann K, Laube B, Abed UA, Brinkmann V, et al. Impairment of neutrophil extracellular trap degradation is associated with lupus nephritis. *Proc Natl Acad Sci U.S.A.* (2010) 107:9813–8. doi: 10.1073/pnas.0909927107
175. Fisher J, Mohanty T, Karlsson C, Khademi SMH, Malmström E, Frigyesi A, et al. Proteome profiling of recombinant DNase therapy in reducing NETs and aiding

recovery in COVID-19 patients. *Mol Cell Proteomics* (2021) 20:100113. doi: 10.1016/j.mcpro.2021.100113

176. Kolaczowska E, Jenne CN, Surewaard BGJ, Thanabalasuriar A, Lee W-Y, Sanz M-J, et al. Molecular mechanisms of NET formation and degradation revealed by intravital imaging in the liver vasculature. *Nat Commun* (2015) 6:6673. doi: 10.1038/ncomms7673

177. Chumanevich AA, Causey CP, Knuckley BA, Jones JE, Poudyal D, Chumanevich AP, et al. Suppression of colitis in mice by Cl-amidine: a novel peptidylarginine deiminase inhibitor. *Am J Physiol Gastrointest Liver Physiol* (2011) 300:G929–38. doi: 10.1152/ajpgi.00435.2010

178. Willis VC, Gizinski AM, Banda NK, Causey CP, Knuckley B, Cordova KN, et al. N- α -Benzoyl-N5-(2-Chloro-1-Iminoethyl)-L-Ornithine Amide, a

Suppression of colitis in mice by Cl-amidine, Reduces the Severity of Murine Collagen-Induced Arthritis. *J Immunol* (2011) 186:4396–404. doi: 10.4049/jimmunol.1001620

179. Drury B, Hardisty G, Gray RD, Ho G. Neutrophil extracellular traps in inflammatory bowel disease: pathogenic mechanisms and clinical translation. *Cell Mol Gastroenterol Hepatol* (2021) 12:321–33. doi: 10.1016/j.jcmgh.2021.03.002

180. Chirivi RGS, van Rosmalen JWG, van der Linden M, Euler M, Schmets G, Bogatkevich G, et al. Therapeutic ACPA inhibits NET formation: a potential therapy for neutrophil-mediated inflammatory diseases. *Cell Mol Immunol* (2021) 18:1528–44. doi: 10.1038/s41423-020-0381-3