A proteomics view at human circulating cells

Patrick Wijten

The research in this thesis was performed in the Biomolecular Mass Spectrometry and Proteomics Group at Utrecht University and performed within the framework of Center for Translational Molecular Medicine (CTMM)

The research was financially supported by the Dutch Heart Foundation and the Netherlands Proteomics Centre.

Printed by: ProefschriftMaken || www.proefschriftmaken.nl

ISBN: 978-94-6295-654-4

A proteomics view at human circulating cells

Een proteomics perspectief op humane circulerende cellen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. A. de Boer, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 14 juni 2017 des middags te 12.45 uur

door
Patrick Wijten

geboren op 29 augustus 1984 te Eindhoven

Promotor: Prof. dr. A. J. R. Heck
Copromotor: Dr. A. Scholten

Table of Contents

1.
General introduction

1.1 Circulating cells

We are continuously subjected to exposure of organisms that are either inhaled, swallowed or are inhabiting our mucosal surface or skin. Whether or not these organisms penetrate and cause disease is determined by both the pathogenicity of the organism with the virulence factors at its disposal and the host defense mechanisms integrity. The immune system is an extensive network of lymphoid organs, cells, humoral factors and cytokines. Immunity can be divided into two parts based on the speed and specificity of the reaction: the innate and the adaptive immune system. There is a lot of crosstalk between these forms of immune response. In contrast to the adaptive immune response that is antigen specific and adapts to a changing environment, the innate immune response is fixed and encompasses elements of the immune system that are required for immediate host defense, such as: granulocytes, monocytes, macrophages, complement system, cytokines and acute phase proteins[1]. This form of immune response is highly conserved even across the simplest animals, confirming its importance in survival[2]. Adaptive immunity on the other hand is a hallmark of the immunity of higher animals. The response consists of antigen-specific reactions through T lymphocytes and B lymphocytes. Contrary to the innate immune system where the reaction is rapid, the adaptive response is precise but takes anywhere between several days to weeks to develop. The adaptive response has memory, causing subsequent exposure to lead to a more rapid and vigorous response, yet still not immediate[3,4]. The innate response is typically involved in inflammation and is, therefore, associated with collateral damage to host tissues when hyperresponsiveness is induced of this type immunity. In the following paragraphs I take a closer look at the particular classes of immune cells that were the subject of investigation in the proteomics studies described in this thesis.

Platelets

Platelets are anucleate cells that originate from megakaryocytes in the bone marrow and are the smallest circulating cells in the blood. The megakaryocyte requires several steps from immature cells (Figure 1A) to form and release platelets (Figure 1E). Initially the cells undergo nuclear endomitosis, organelles synthesis and a dramatic cytoplasmic maturation and expansion while an array of microtubule emanating from the centrosomes is established (Figure 1B). Centrosomes disassemble and microtubules translocate to the cell cortex prior to the onset of proplatelet formation. Proplatelet formation then commences with the development of thick pseudopods (Figure 1C). The sliding of overlapping microtubules drives proplatelet elongation as the organelles are tracked into proplatelet ends where the nascent platelets assemble (Figure 1D). Finally the entire megakaryocyte cytoplasm is converted into a mass of proplatelets which are released from the cell. Eventually the nucleus is extruded from this mass of proplatelets and individual platelets are released from the proplatelet ends (Figure 1E)[5]. As stated before, platelets are anucleate and as such cannot synthesize new RNA molecules. However, translational activity has been shown in platelets in the past due to mRNA molecules being present within the platelet[6].

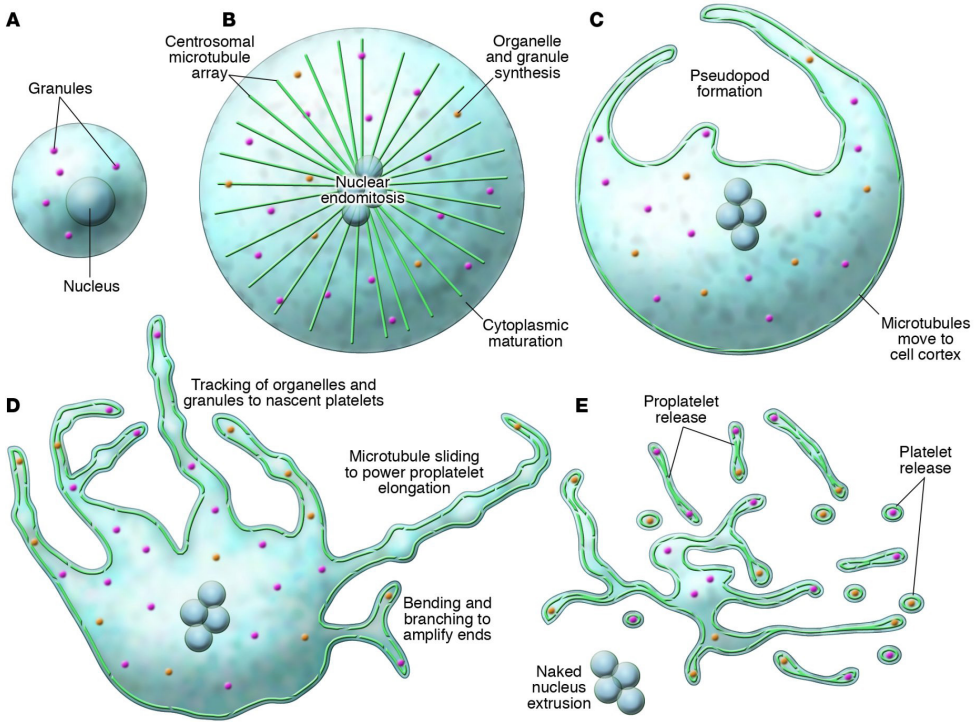

Figure 1: Overview of megakaryocyte production of platelets.

Patel *et al*[5].

Platlets are most known for their involvement in wound healing. When an injury of the vasculature occurs, the endothelial layer of a blood vessel becomes disrupted, exposing the extracellular matrix which in turn induces platelets to trigger a repair process[7]. This repair process is mediated by blood rheology. Blood flows at differential speed in a vessel, the center attaining the highest rate, as a result of this shear forces develop between the different layers of blood flow, the extent of which are dependent on flow rate and vascular radius[7,8]. Platelet function is directly dependent on the shear conditions[9,10] as injury in vessels with high shear rates require von Willebrand Factor (vWF) to slow down and establish firm adhesion[11], where it will attach to the exposed collagen of the extracellular matrix, causing it to unroll and allow the subsequent tethering to the surface of other platelets via their glycoprotein (GP)Ib receptors[8]. In vessels with relatively low shear the vWF-GPIb interaction is less important[12]. Platelets tethered to the extracellular matrix will become activated by means of collagen which activates GPVI and $\alpha_2\beta_1$[13], which in turn trigger the release of secondary activators such as the $P2Y_{12}$ receptor agonist adenosine diphosphate (ADP), the TP agonist thromboxane (TXA2) and serotonin[14-16]. At the same time the tissue factors exposed by the endothelium triggers the formation of thrombin, further activating the platelets via protease activated receptors PAR-1 and PAR-4[17]. The adhesion of platelets to the extracellular matrix is mediated by integrins such as collagen adhesion receptor $\alpha_2\beta_1$ and $\alpha_{IIb}\beta_3$. Platelet activation

triggers a conformational change in $\alpha_{IIb}\beta_3$ that leads to an increased affinity for fibrinogen, fibrin and vWF which allows for platelets to form an aggregate to seal the vessel wall injury[8].

Besides their best known role as cellular mediator of thrombosis, there is now a growing appreciation that platelets are key effector cells in inflammatory and immune processes. Platelets have the ability to secrete many molecules that are able to attract and interact with the innate and adaptive immune systems and as such they have a pivotal role in inflammatory and immune processes modulation[18]. Platelets contain granules that store a large quantity of molecules with specific immune function[18-21]. But also on the surface of the platelets some of the most important receptors are located known for their function in innate immune cells, namely the Toll-Like Receptors (TLR)[22-24]. Although a lot is known it is imperative to gain more comprehensive knowledge of all platelets have to offer.

Granulocytes

Granulocytes play essential roles during inflammation, both for clearing pathogens and for immune-regulation[25-29]. The 3 types of granulocytes (neutrophils, eosinophils and basophils) mature from the bone marrow from a common myeloid progenitor[28]. After being fully differentiated these granulocytes are released into the circulation where they remain in G_0 phase for the duration of their short life. Spontaneous apoptosis will occur after a maximum of approximately 5 days in absence of external stimuli[30]. Granulocytes migrate into tissues in response to chemo-attractants such as interleukin-8 (neutrophils) and eotaxin (eosinophils and basophils). Extravasation and pre-activation by cytokines and/or molecules derived from pathogens or damaged cells induce a fully primed phenotype that is essential for an optimal response upon engagement of micro-organisms. The killing is mediated by a wide arsenal of toxic effector molecules that are secreted secreted both into the phagolysosome (neutrophils) as well as into the extracellular space (all granulocytes) to subdue the invading pathogen. As these effector molecules are toxic to the host tissue the release of these molecules is finely tuned. Besides subduing pathogens in a direct fashion, granuloctyes also exhibit immunomodulatory properties as they produce and secrete a large number of different cytokines, chemokines and other pro-inflammatory mediators[25,26,29]. Besides the common aspects of the granulocyte types, there are clear differences in: morphology, relative abundance in different tissues, and killing mechanisms. The function of the different cells determine what effectors they secrete upon activation.

Neutrophils

A central feature of the innate reaction is the recruitment and activation of neutrophils at the site of infection to eradicate pathogens[31,32]. A wide variety of immunological functions have been attributed to neutrophils in recent years shifting perspective from neutrophils being general killer cells to a key modulator of the immune response[33,34]. Pertinent to its many functions are the different granule subsets that are mobilized upon stimulation that either

fuse with the cytoplasmic membrane or the phagosomal membrane. This results in exocytosis and/or expression of its soluble and membrane-bound proteins on the cell surface that are essential for neutrophil-endothelial interaction, extravasion, phagocytosis and elimination of micro-organisms[35]. Neutrophils contain an heterogeneous population of granules that are formed during the terminal stage granulopoietic differentiation[36]. The different granule subsets can be distinguished by their protein content[37]. This content is not dependent on soring of proteins to individual storage granules but rather by the stage of differentiation the proteins were synthesized, according to the targeting-by timing hypothesis[38]. During each stage of differentiation granules will be formed with proteins that are synthesized at that moment, this yields granules characteristic for each stage of development.

There also is a degree of overlap of proteins and granule surface markers during differentiation of neutrophils in the bone marrow. In total six different neutrophil precursors can be distinguished based on morphology, granules and marker proteins: myeloblasts, promyelocytes, myelocytes, metamyelocytes, band cells and segmented or mature polymorphonuclear leukocytes. The content of the resulting granules produced in each stage will differ from each other to varying degree resulting in 3 distinct granule subsets called azurophil granules (AGs) identifiable by myeloperoxidase, specific granules (SGs) identifiable by lactotransferrin or neutrophil gelatinase-associated lipocalin (NGAL) and gelatinase containing granules (GGs) identifiable by MMP-9. Besides the differences in content the capacity for mobilization in response to stimulus also differs, the later it was produced the more it is prone to exocytosis[39].

Azurophil granules (AGs) are formed early during granulopoiesis and as such least prone to exocytosis, this is caused by the absence of exocytosis related proteins on the membrane of AGs like vesicle-associated membrane protein (VAMP)-2 and synaptosome associated protein (SNAP)-25 that allow fusion with the plasma membrane, restricting the AGs to operate in the phagosome. This protects the environment from its highly toxic content as most AG proteins work on microbes in a direct fashion as a toxic agent such as elastase, defensins, lysozyme and bactericidal/permeability-increasing protein (BPI). Whereas other proteins change the environment such that the microbe can be indirectly eliminated. Examples are myeloperoxidase (MPO) and arginase 1 (ARG1). MPO produces HOCl from H_2O_2 that is produced by the NAPDH oxidase complex[40]. This process mainly takes place in the phagosome with which the AGs fuse after a pathogen is absorbed[41]. SGs and GGs are highly similar and differ from AGs in the sense that these granules are peroxidase negative. Characteristic for SGs is the presence of NGAL and absence of gelatinase, whereas the opposite is true for GGs. The proteins in SGs are more potent in their antimicrobial effect compared to GGs. NGAL is important in this aspect as it binds to bacterial sidophores that are responsible for iron trafficking from the host to the microbe or growth enhancement[42,43]. Human cathelicidin (hCAP)-18 is another protein that can be found in SGs in inactive form. In

turn, serine proteases stored in AGs are able to activate hCAP-18 once the neutrophil becomes activated allowing the resulting LL-37 peptide to exert its function of disrupting bacterial membranes[44]. Apart from gelatinase, the GG contains a matrix metalloproteinase 9 (MMP-9), which is capable of cleaving type IV collagen[45]. Besides collagen degradation MMP9 plays a role in the cleavage of several chemokines and as such acts as a regulator of the neutrophil response[46]. Another function of GGs is to prime the cell for an oxidative burst by means of upregulating cytochrome b558 which upon stimulation by a secondary stimulus forms an active complex[47]. Besides these granule subsets neutrophils also contain secretory vescicles (SV) which are formed by endocytosis at the final stage of granulopoiesis[48]. SVs are the most readily exocytoseable organelles of neutrophils responsible for providing the cell surface with membrane-bound proteins essential for extravasation[49-53]. SVs exert an immunological function due to the translocation of adhesion receptors such as complement receptor 1 (CR1) onto the plasma membrane, rather than excretion of immunological proteins from their lumen[54]. MMP-9, along with elastase is a regulatory factor in neutrophil migration across the basement membrane of endothelial cell layers.

Eosinophils

Eosinophils have been extensively characterized for several decades, elucidating multiple aspects of these cells[55-57]. Still, perspective on eosinophils function is undergoing changes[58]. For a long time it was believed that eosinophils promote immunity against helminthes, but this has been challenged by recent animal studies showing that eosinophils prevent inducible NO synthase production in macrophages and neutrophils essential for enhancing parasite clearing resulting in sustained chronic infection and ensure parasite survival in the host[59,60]. Similarly eosinophils are known to be part of the pathophysiology of asthma as they are recruited to and activated in lung tissue of asthma patients[61,62]. Recent findings suggest that the pathophysiology of asthma arise from dysregulated eosinophils in the airways. Aspecific activation of these cells may cause collateral damage to bronchial tissue. This is in part based on studies with asthma models in eosinophil deficient mice[63,64]. However, it is clear that eosinophils are also capable of sophisticated immune functions, illustrated by their nuanced degranulation responses to distinct stimuli[58,65,66]. This is further demonstrated by their ability to engage in complex interactions with a wide variety of other leukocytes such as T cells, T_H cells, B cells, dendritic cells and mast cells by means of secreting cell specific chemokines and other mediators[67-73]. The core proteins in eosinophilic granules are major basic protein (MBP), eosinophilic-derived neurotoxin (EDN), eosinophil cationic protein (ECP) and eosinophil peroxidase (EPX), each of which exists in a variety of isoforms. Distinct (non-) cytotoxic functions can be attributed to these proteins, such as the aforementioned anti-viral and anti-parasitic functions by permeabilization of membranes, ROS production and stimulation of other parts of the immune system like mast cell degranulation. Besides these core proteins, granules of eosinophils like the other granulocytes are a rich source of

bio-active lipids that include but are not limited to PAF, LTC4, PGE2[56]. These lipids are involved in migration and priming of an array of immune cells. They are bound to serglycan proteins for preservation to increase the magnitude of secretion[74]. A variety of cytokines and chemokines are both secreted from stored granules or synthesized ***de novo*** and subsequently secreted upon eosinophil activation. IL-4, IL-6, IL-10, IL-12, IL-13, IFN-γ and TNF-α are stored in granules and require specific stimulation[68] whereas proteins such as IL-2, IL-5, IL16, IL-18, TGF-α, TGF-β, CCL5, CCL11 are being produced ***de novo[75].***

Basophils

Unlike neutrophils and eosinophils, basophils are not phagocytosing cells nor are they involved in direct killing of micro-organisms. In fact, basophils were historically primarily associated with immediate hypersensitivity reactions. This was based on their cell surface expression of the high-affinity IgE receptor (Fcε R1) that is always bound to monomeric IgE. Crosslinking of surface IgE bound to its receptor results in the release histamine and other mediators important for atopy[76]. Basophils share this FcεR1 mediated activation with mast cells. In addition, basophils and mast cells share mechanisms, such as lipoxygenases and cyclooxygenases important for the synthesis of several immunological active mediators like leukotrienes and prostaglandins[1]. But contrary to mast cells which originate from monocytes and reside in the peripheral tissue after maturation, basophils are predominantly located in the circulation. Histamine is an important component found in basophilic granules which is commonly identified in tissue and circulation which considered to originate from the basophils that resides in that particular matrix[77,78]. Recently it has become apparent that these cells express a plethora of surface receptors that whilst ligated are capable to activate basophils to produce cytokines that in turn promote and regulate Th2 adaptive immune responses[25,79]. The reason for this relatively late recognition on the immune modulation process is at least in part due to having the lowest concentration of all white blood cells in circulation under normal circumstances. Amongst these cytokines that are also produced ***de novo*** are IL-4, IL-6, IL-9, IL-13 and granulocyte macrophage colony stimulating factor (GM-CSF)[80,81]. Similar to other granulocytes these proteins are bound to serglycan proteins to prevent degradation and increase the magnitude of secretion[74].

1.2 Proteomics

Introduction to proteomics

DNA is often called the blueprint of life as they encode for all genes that exert a function within a cell. Although DNA is the blueprint it is not the direct plan to construct the final product, this role is fulfilled by RNA molecules that are transcribed from the coding region of the DNA. Finally the RNA is being translated into chains of amino acids that once folded into a three-dimensional structure constitutes to a protein. These proteins act as key effectors directly involved in a plethora of cellular processes.

There are several ways of looking in a system-wide manner at the processes that go on in a cell to unravel their function and their underlying processes. One could look at the DNA content using a genomics approach and while this approach provides a wealth of information, it is mostly qualitative and does not provide a dynamic picture. Looking at the mRNA using a transcriptomics approach adds a more dynamic level of information, but as this is essentially a messenger, it does not exactly reflect the situation as encountered in a cell due to editing in transcription in the form of splicing for instance. To really monitor the prime acting biological molecules in the cell, the protein we need to use technologies that are the focus of this thesis: proteomics.

Analogic to genomics and transcriptomics, proteomics is a branch of analytical science focusing on elucidating both the qualitative and quantitative state of the complete set of proteins being either expressed or present in a defined entity of living matter such as organisms, organs, tissues or cells[82-84]. These proteins are to some extent dynamic in terms of abundance, isoforms and post-translational modifications, especially concerning different stimuli and disease states. This opposed to the genome and the transcriptome which are relatively static by comparison.

There are some additional challenges associated with proteomics compared to genomics and transcriptomics. First of all the scale of the potential data, where genomics analysis deals with roughly 20 000 genes, that number is on average double that amount for transcript variants[85] each of which can be transcribed into a protein which can be post-translationally modified in a vast amount of different ways, resulting into a grand total of likely more than one million protein isoforms. This is further complicated by localization, interaction and the dynamic range of these proteins that can vary depending on the matrix over 6 orders of magnitude.

This increase in sample complexity poses challenges on the analytical strategies to accurately identify and quantify the protein content of a given matrix. Not one method can cover all possible needs of a researcher to unravel their proteome of interest and as such a vast number of possible workflows exist. The common or at least dominant denominator of those approaches is the use of mass spectrometry[86-88].

Mass spectrometry based proteomics

The analytical methods can either be used to measure protein derived peptides in bottom-up proteomics[89] or measure intact proteins directly with top-down proteomics[90]. Bottom-up proteomics is the more widely implemented of the two and typically consists of 3 aspects: sample preparation, MS analysis and data analysis as illustrated in figure 2. The following sections will elaborate on the different aspects of experimental design and types of instrumentation commonly associated with it.

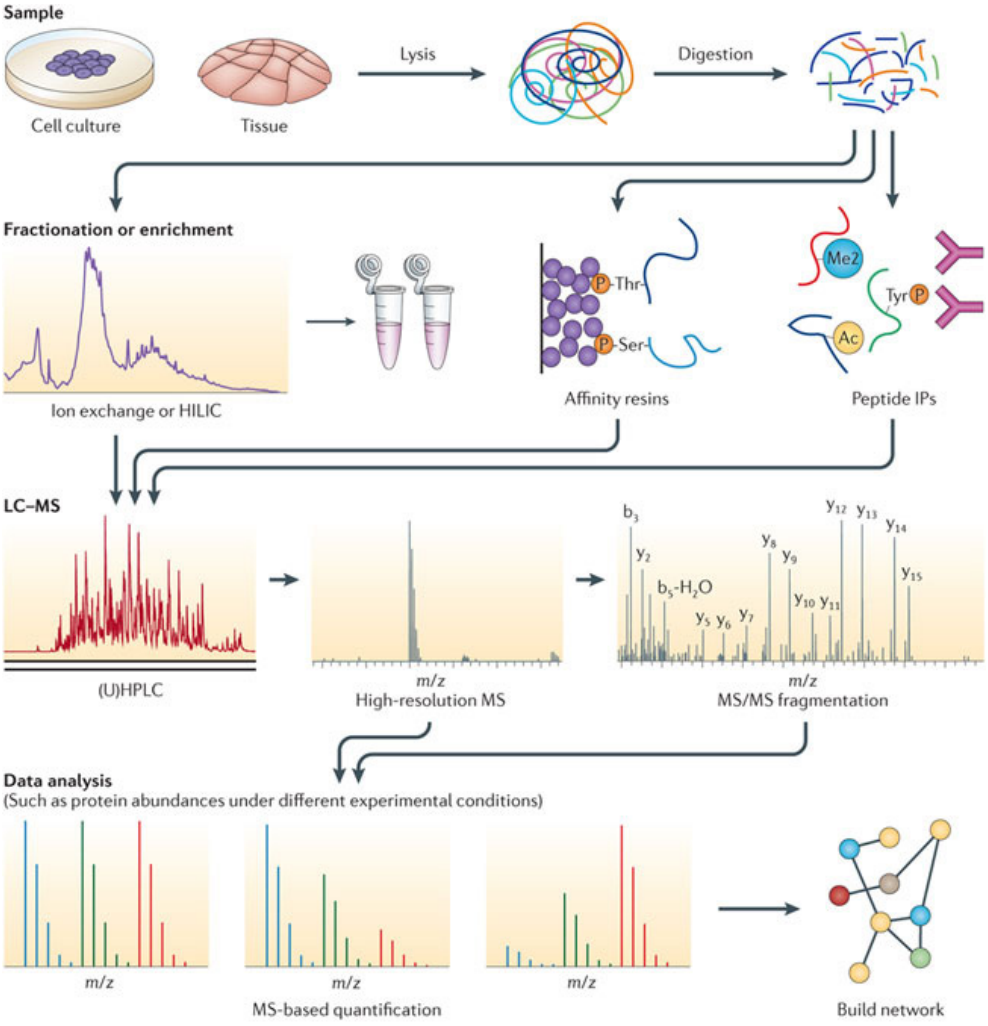

Figure 2: Typical proteomics workflow for bottom–up proteomics, requires adjustment according to the research question at hand.

Adapted from Altelaar *et al*[86].

Sample preparation

The proteome of a variety of sample matrices can be analyzed using mass spectrometry in a bottom-up approach. Several sample preparation steps are required in order to obtain peptides. Initially the proteins of interest need to be either extracted from the sample matrix which can be comprised of synthetic protein mixtures, biochemically enriched extracts such as organelle fractions; ***in vitro*** cultured cells; ***in vivo*** serum, circulating cell and tissue samples. Typically this can be achieved by cellular disruption using an appropriate lysis buffer and sonication followed by centrifugation. This method allows for cytosolic proteins to be recovered at high yield, in order to isolate membrane proteins for instance additional steps have to be added to the workflow to increase efficiency. In order to preserve the proteins from endogenous proteases, protease inhibitors need to be added to the lysis buffer. This is particular the case for the primary cells investigated in this thesis, that harbor numerous, abundant and highly active proteases. After protein extraction the proteins are required to be digested to peptides. In order to achieve this, the proteins are denatured, protein disulfide bonds are reduced and subsequently alkylated to prevent refolding of the proteins. This allows better access for the proteolytic enzymes to which the denatured proteins are subjected next. The proteolytic enzyme trypsin is most commonly used due to its specific cleavage at the C-terminus of arginine and lysine[91]. In order to increase digestion efficiency trypsin has been complemented with Lys-C amongst others with great success[92,93]. The use of trypsin ensures peptides of similar length and charge that are convenient for identification by the mass spectrometer.

Reducing sample complexity

After proteolytic digestion of the biological matrix of interest the entirety of the proteome is converted to peptides, hundreds of thousands of them to be precise. The sheer number of peptides is too complex to directly use for MS analysis as it sequences peptides one at a time. Therefore separation using LC is required to reduce sample complexity in order to present the peptides at a more manageable rate to the MS. The separation in liquid chromatography is based on the physicochemical properties of the peptide amino acid sequence. A commonly used option for separation is ion exchange and in particular strong cation exchange (SCX) the separation of which is based on the number of positive charges[94-96]. The problem with this method of separation is that besides a low resolving power, a salt gradient is used for elution which is not directly compatible with electrospray ionization (ESI), making it an “offline” separation as additional sample cleanup is required. The most common stationary phase used is reversed phase which separates peptides based on hydrophobicity[97] which has as superior resolving power[98] compared to the other available options and the eluent used is compatible with electrospray ionization (ESI) which allows for direct interfacing to the MS, thus being considered “online”. Considering the high orthogonality of SCX and RP, the former is often used as an initial step to pre-fractionate, the fractions of which are injected into an online reversed phase LC-MS setup[99].

Mass spectrometry and peptide sequencing

Mass spectrometry (MS) in proteomics is used to sequence peptides by first detecting the mass-to-charge ratio (m/z) of the analyte ions. In order to analyze the mass of an ion several steps are required. First the analyte is required to be ionized which happens in the ion source, next the ions are directed into the vacuum of the mass spectrometer where the ions are guided and separated by a series of manipulations by electric fields prior to detection by a mass analyzer of the resulting discrete ion packages of interest. To sequence these peptide ions a second stage of analysis is required, called tandem mass spectrometry (MS/MS). In this stage the peptide precursor ion is fragmented into daughter ions in order to generate a mass ladder of the peptide which are subsequently analyzed. The following paragraphs will elaborate on the processes and common MS instrumentation involved.

Ion Source

The peptide mixture as introduced to the mass spectrometer cannot be analyzed in its current form. In order for that to be possible the peptides are required to be in the gas phase prior to entering the vacuum, meaning the non-volatile hydrophobic peptides need to be simultaneously ionized and separated from the solvent they are introduced in without degradation. This process takes place in the ion source. Two suitable types of soft ionization techniques are commonly used in proteomics: matrix-assisted laser desorption/ionization (MALDI)[100] and electrospray ionization (ESI)[101].

In MALDI a matrix (e.g. α-cyano-4-hydroxycinnamic acid) in acidic environment is mixed with the analyte solution which upon solvent evaporation form matrix/analyte co-crystals on a metal target plate, depicted in figure 3a. These crystals are then irradiated by a pulsed laser, causing ablation and desorption of both analyte and matrix material. The matrix then facilitates protonation of the peptides generating predominantly singly charged $[M+H]^+$ species[102]. MALDI suffers from restriction to offline sample preparation and separation methods and a somewhat low laser shot to shot reproducibility.

As opposed to MALDI, in ESI ions are being produced from a liquid phase by applying a high voltage (1-3 kV) between the emitter and the mass spectrometer inlet. The result is a Taylor cone forming at the charged tip that yields small droplets that in turn burst into smaller drops due to Coulomb fission, this process continues for several cycles until the analyte is desolvated to the extent that gas phase ions are formed, as depicted in figure 3b[103,104]. This process is aided by the use of a heated transfer capillary. Using ESI multiply charged ions are generated depending on length and amino acid sequence, for tryptic peptides $z \geq 2$. ESI has become the ionization method of choice when it comes to high throughput proteomics by allowing the direct interfacing of LC separation with MS analysis and maximizing MS analysis time due to the continuous flow of ions[101]. ESI is the ionization method used in all experiments in this thesis.

Mass analyzers

In order to determine the mass of ions, a mass analyzer is required. The mass analyzer detects the mass to charge ratio of the analyte and it to do so a variety of types of mass analyzers operating on different principles are available. The most commonly used mass analyzers in proteomics are time-of-flight (TOF) which uses ion velocities or flight times, the quadrupole (Q) based on the stability of ion trajectories in oscillating electric fields, the linear ion trap (IT) and the orbitrap which both rely on the m/z resonance frequency. The choice of mass analyzer depends on the nature of the proteomics experiment, for the discovery stage detection speed and resolution are often the most important factors whereas in targeted

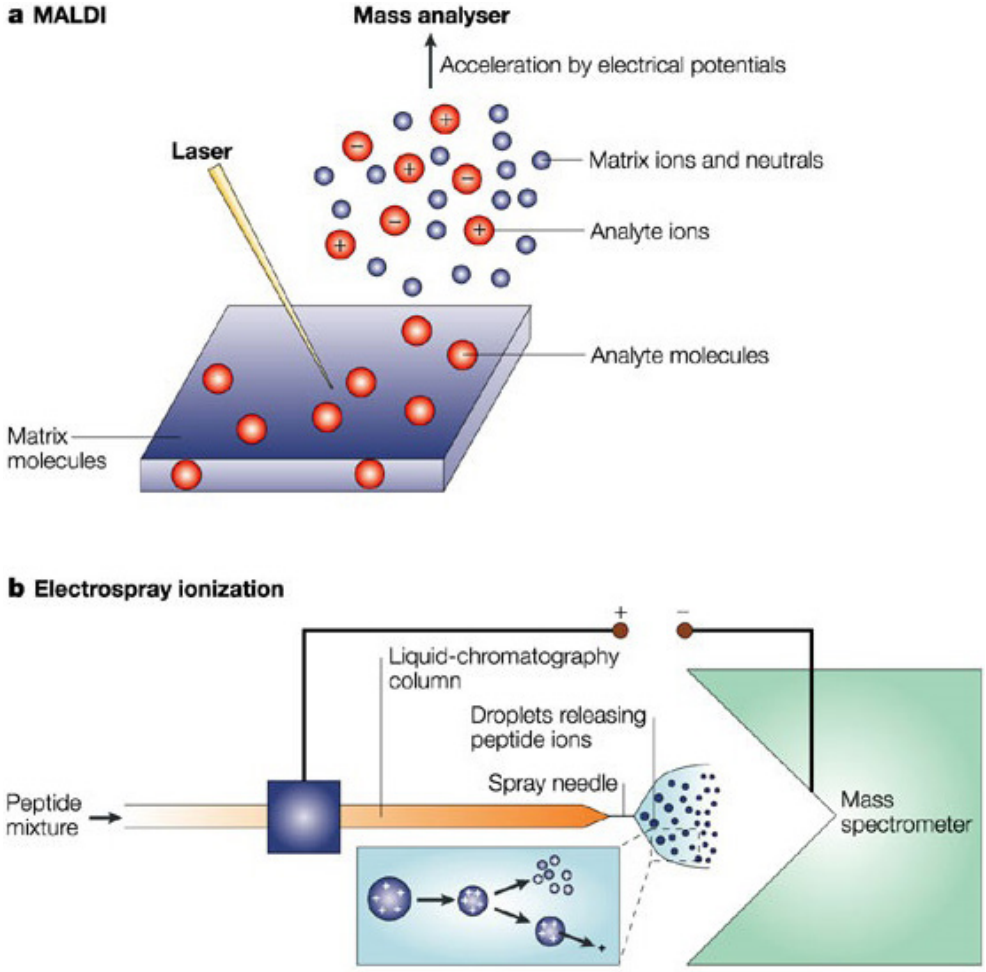

Figure 3: Soft ionization techniques most commonly used in high throughput proteomics experiments. (A) matrix–assisted laser desorption/ionization (MALDI) uses irradiation of matrix/analyte crystals to facilitate proton transfer from to the peptide. (B) electrospray ionization (ESI) induces desolvation through coulombic fission until gas phase peptide ions are obtained.

Adapted from Steen *et al*[87].

experiments selectivity, sensitivity and dynamic range will be prioritized. The sequential use of mass analyzers addressing the requirements of in particular discovery stage experiments is possible with the current range of hybrid instruments available[88]. These mass analyzers will be discussed in more detail in the following paragraphs.

Time-of-flight

The time-of-flight (TOF) analyzer determines an ion m/z ratio via a time measurement. Ion packages are initially static and start out from the same position where they are accelerated by an electric field of known strength[105]. All ions will thus have identical kinetic energy per charge state, the difference in velocity the ions progress is dependent its m/z ratio and is measured by the time it takes for it to reach the detector at a known distance. There is a possibility of differences in initial kinetic energy, to circumvent this problem a reflectron is commonly used. The reflectron uses a constant electrostatic field to reflect the ion beam towards the detector. If ions have more initial kinetic energy, they will penetrate deeper into the reflectron, prolonging the path to the detector[106]. The TOF MS is one of the fastest scanning instruments, able to produce a full scan in microseconds.

Quadrupole

This mass analyzer consists of four cylindrical or ideally hyperbolic shaped metal rods that are in perfect alignment in a parallel fashion[107,108]. A radio frequency (RF) voltage is applied on the rods in such a fashion that the potential is identical for opposing rods and 180° out of phase for adjacent rods, this yields a dynamic electrical field that causes ions to progress in a circular trajectory on the xy-plane for a specific m/z value. The trajectory of the ions will be stable within a certain m/z range and be preserved as they progress through the quadrupole in a corkscrew trajectory whereas ions with a m/z out of this range will have an unstable trajectory and will ultimately either be ejected from the quadrupole into the vacuum or collide with the rods, losing their charge and will not be able to reach the detector. In case only a RF voltage is applied the trajectory is stable for a broad range of m/z values, allowing to capture and focus the whole range of charged ions, this is used to guide ions to different sections of the mass spectrometer or for fragmentation purposes that will be discussed at a later stage. The discrimination on m/z range can be manipulated with specific sets of RF and DC voltages allowing different m/z ranges to reach either the next mass analyzer or the detector, turning the quadrupole into a mass filter. Different RF only configurations are possible, including a hexapole and octapole configuration that are used to maximize transmission efficiency of a broad m/z range.

Linear Ion Trap

The linear ion trap (LIT) or 2D trap is essentially a quadrupole analyzer with an electrode at each end[109,110]. These electrodes can be plates or short quadrupoles that are kept at a static electric potential. This configuration allows ions to be confined inside the rods by the quadrupolar field in the radial dimension and by the electric field of the closing electrodes

in the axial dimension. As opposed to quadrupoles in the ion trap all ions of different masses within a specified m/z are stored together and selectively and sequentially ejected according to their m/z to obtain a spectrum. Once ejected from a slit in 2 of the rods, the ions are collected at the detector for ion counting. In this type of mass analyzer a space-charge effect occurs, charged ions repel each other when in close proximity leading to destabilized ion trajectories[111]. In order to circumvent this effect the ions are collisionally cooled by the introduction of an inert gas such as helium, focusing the ions towards the center of the trap. The excess of kinetic energy is negated allowing for higher storage volumes in the LIT yielding an increased dynamic range. The helium can also be used to fragment the trapped ions. Another way of controlling the space-charge effect is to regulate the ion population inside the instrument by doing a scan of the ion population prior to the actual measurement and using that to automatically adjust the ion accumulation time in the LIT[112].

Orbitrap

The orbitrap is a recently developed mass analyzer that employs orbital trapping as method for m/z analysis and has gained popularity rapidly since its introduction. The orbitrap was invented by Makarov in 2000[113] based on the Kingdon trap, first described in 1923[114]. The electrostatic trap is comprised of a spindle shaped inner electrode and a barrel shaped outer electrode which is split into two halves. Ions are trapped by balancing the pull of the electrostatic attraction towards the inner electrode and the centrifugal forces. This causes the ions to be harmonically oscillating back and forward along the axis of the inner electrode. The frequency of the oscillations is inverse proportional to the square root of the m/z of the ion and are detected by the image current it induces that a differential amplifier connected to the two halves of the outer electrode records. Fast Fourier transformation then converts the recorded time-domain signal into a mass/charge spectrum[115]. The instrument is capable of achieving high resolution (>100 000 FMHW) and mass accuracy (<2 ppm) at LC separation time scale[116,117]. In a hybrid setting combined with the speed of linear ion traps and tandem MS functionality this mass analyzer has become one of the most frequently used mass analyzers in current proteomics analysis[118].

Hybrid instruments

As mentioned in the previous section, mass analyzers are used in hybrid setting, in fact most commercially MS instruments are composed of multiple mass analyzers. By using complimentary mass analyzers higher protein identification rates can be attained or specific aspects of the analysis can be enhanced. In this thesis the following hybrid instruments were used: the LTQ-Orbitrap Velos, the Q-Exactive and the Q-TOF.

Detectors

After separation of ions by mass analyzers other than orbitraps, the signal of the m/z value needs to be determined by the detector. The electron multiplier (EM) is one of the possible

detectors and is widespread in usage in mass spectrometers used in proteomics. Commonly the EM consists of a continuous dynode with a concave metal surface that catches ions that originate from Q, LIT or TOF mass analyzer. Upon hitting the surface an ion will trigger a cascade of electrons forming amplifying the signal exponentially, up to 10^7 as depicted in figure 4. This will result in a rapidly measurable current at the end of the dynode[108].

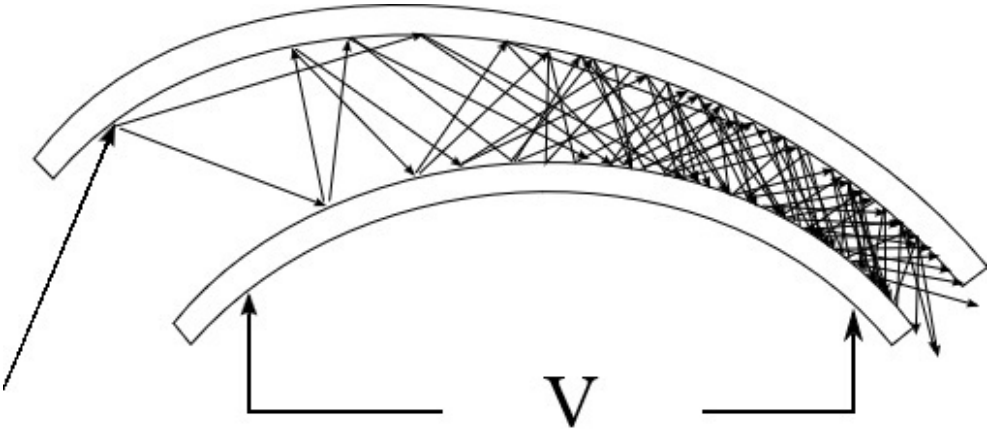

Figure 4: The electron multiplier detector principle, an electron multiplication cascade is triggered upon hitting the secondary emissive material, amplifying the initial signal up to 7 orders of magnitude.

Peptide fragmentation

In order to identify a peptide and consequently the protein it corresponds to, the m/z of a precursor ion alone is insufficient. For the next step in identifying the peptide through tandem MS, the precursor ion is fragmented into a range of product ions that yield a mass ladder that can be used to determine its identity. There are several types of low-energy fragmentation techniques available suitable for proteomics to dissociate a peptide at its backbone. Depending on how the energy is imparted onto the precursor ion it can fragment in several ways, the designation (Roepstorff-Fohlman-Biemann nomenclature) for the resulting fragment ions is illustrated in figure 5[119,120].

Collision-induced dissociation (CID) is the most frequently used fragmentation technique[121]. Protonated peptides are excited using an electric field to increase the available kinetic energy and are subjected to collisions with inert gas molecules: typically either helium, nitrogen, argon or xenon are present in the collision cell. The collisions with the gas lead to a conversion of kinetic to vibrational energy in the peptide ion, ultimately leading to fragmentation. The site of fragmentation is dictated by the proton affinity of the fragment as protons localized on the basic side chain can migrate along the backbone upon excitation with a preference for amines. The C-N bonds are weakened by the proximity of a proton to the backbone amino group and combined with the increase of internal energy peptide fragmentation

will take place, generating predominantly b- and y- ions (Figure 5), along with additional fragments such as immonium ions, neutral losses for example[122]. This type of fragmentation can be subcategorized into resonant excitation and beam-type CID[123]. Resonant excitation which is performed in an ion trap and referred to as IT-CID or tandem in time, where due to restricted space only a small amount of kinetic energy can be bestowed upon a peptide using resonance frequencies (typically a few eV), that combined with a higher pressure of helium molecules results in the slow accumulation of energy via multiple collisions leading to fragmentation in a millisecond scale. Beam-type CID takes place in dedicated collision cells in tandem quadrupoles (Q-TOF, QqQ,) or multipoles (HCD) and is referred to as tandem in space. Ions enter the collision cell filled with gas at a much higher initial kinetic energy by being accelerated in an electric field, allowing for a much higher amount of energy transfer in a shorter timescale. The energy transfer rate is typically several tens to hundreds of eV, converted to mainly internal energy by the initial collision with a larger gas molecule such as nitrogen or argon. The transfer of energy typically takes only microseconds to complete. This is in particular advantageous for PTM analysis.

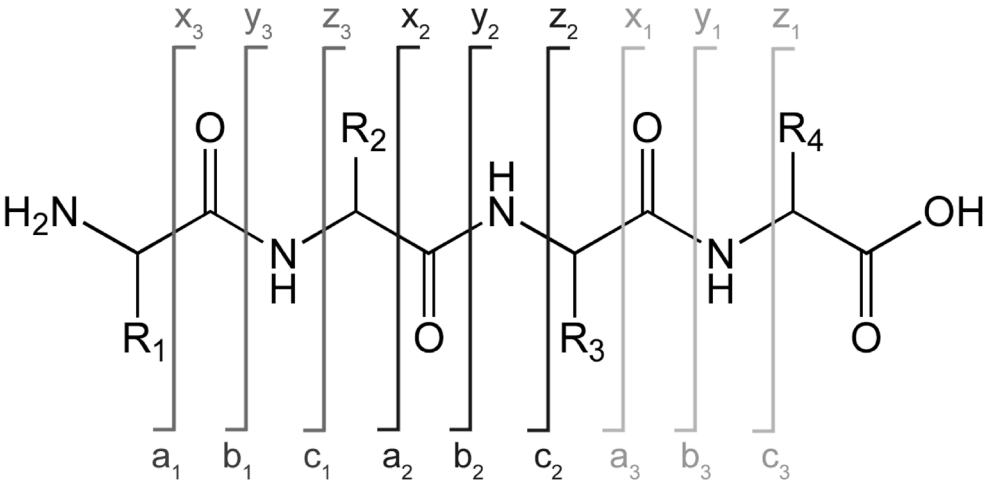

Figure 5: Roepstorff–Fohlman–Biemann nomenclature if peptide backbone fragments.

Adapted from Roepstorff *et al*[119].

Electron capture ad or transfer dissociation (ECD, ETD) are a more recently introduced fragmentation technique in which electrons are either directly captured (ECD) or transferred from a radical anion such as fluoranthene (ETD) to a peptide cation brought into close proximity within the ion trap[124,125]. As the positively charged multi-protonated peptide captures an electron, the backbone is fragmented at preferentially the N-Cα bond, yielding c- and z-type ions (Figure 5). ETD is less efficient for doubly charged cation peptides due to their low charge densities, it requires supplemental resonance excitation otherwise it

results in charge reduced and non-fragmented precursor ions[126,127]. ETD is considered to be complementary to CID which is less efficient on peptides with more than two positive charges. The strength of both fragmentation techniques was recently combined by a hybrid mass spectrometer as mentioned before in which an automated decision tree selects the optimal fragmentation method based on precursor charge and length[128,129].

Data analysis

As stated before MS/MS spectra contain information regarding peptide sequence and inferred from that the protein identity. Typically a single LC-MS experiment already generates several tens of thousands of spectra, a full scale discovery experiment consisting of multiple LC-MS runs yields a vast multitude of that, thus an automated method of data interpretation is required. For that purpose several database search algorithms have been developed, such as Mascot[130]D, Sequest[131], Andromeda[132], X!Tandem[133]. The initial step of these search algorithms is to generate a theoretical list of possible peptides and corresponding masses by doing an ***in silico*** digest of the proteome of interest. The user defined cleavage enzyme and number of possible peptide modifications mainly dictate the size of the peptide list. The ***in silico*** peptide masses are then compared to the acquired precursor masses within a specified mass tolerance, those yielding a positive match are subsequently analyzed on MS/MS level. For each of the matched peptides a theoretical MS/MS spectrum is generated according to known fragmentation rules which is then compared to the experimental MS/MS spectrum of the precursor and is scored for similarity based on number and type of fragment ions matched. The sequence with the highest score is considered to be the observed peptide sequence yielding a peptide spectrum match (PSM).

When matching a large number of spectra to all possible candidates in the database, there is a risk of random and thus incorrect matches. Although the odds of incorrect matches can be reduced to some extent by decreasing precursor and fragment ion mass tolerance which is attainable by using high accuracy mass analyzers, which in turn reduces the amount of possible sequence candidates, there is no way to get rid of false-positive identifications completely due to the huge amount of data. To account for these false positives a false discover rate (FDR) is estimated, for which several statistical models have been developed. One method to estimate the FDR is to perform a second search of the experimental data against a decoy database of either a reversed or scrambled sequences generated from the database the initial search was performed on[134,135]. All matches against this decoy database are to be considered false positive. The FDR is calculated by dividing the number of false hits by the total number of PSMs. It is common practice in the proteomics community to accept a FDR of 1%, nevertheless additional discriminating parameters such as minimum peptide score and minimal sequence length are used to increase the stringency of the dataset[136]. Applying these additional filters will often result

in the unwanted loss of true positives. Recently more sophisticated semi-supervised machine learning algorithms have been developed to reduce loss of information by including a number of other parameters, such as precursor and fragment mass error, peptide hydrophobicity and charge[137].

Quantitative proteomics

Everything described until now yields qualitative data about peptides and their corresponding proteins. However, when looking for a biomarker or trying to elucidate a pathway the difference is likely not as black and white as qualitative data can provide. There is more information that can be extracted from the data generated as described thus far or by adding one or more additional steps in the sample preparation protocol a wealth of quantitative data can be obtained. One could look at merely the signal intensity of the peptides in MS, but the intensity is dependent on more factors than simply their concentration. Physicochemical properties like solubility and the peptides ability to be ionized effect the intensity, but also peptide digestion efficiency earlier in the process result in a non-linear correlation between peptide intensity and protein concentration thus the intensity obtained for different peptides cannot be compared. In order for proteomics to be truly quantitative, a comparison is only possible between identical peptides. To enable quantitative proteomics, several strategies have been developed which can be divided into two categories, label-free and stable isotope labeling[138]. The label-free and most stable isotope labeling strategies are relative quantification methods mainly for comparing differences in protein abundance between several different conditions. One of the stable isotope approaches however is used for absolute quantification where protein copy numbers or concentration are determined by means of internal standard combined with a target method (Figure 6).

Label–free quantification

The most straight forward form of relative quantification is label-free quantification as it does not require any additional sample preparation steps and therefore suitable for all types of samples, contrary to the other quantification methods (Figure 6). Because each sample is analyzed separately, the number of samples that can be compared is infinite. This quantification method is less accurate compared to label-based approaches due to possible variations in sample preparation (e.g. digestion variability) and LC-MS conditions (e.g. shift in peptide retention time, variation in ionization). Consistency in sample preparation is of the utmost importance hence fractionation and enrichment steps are usually avoided unless proven to be very robust. The need for consistency negatively affects the proteome coverage. The high variability can be compensated by measuring several technical replicates. The quantification is based on either spectral counting[139] or comparing peptide signal intensity[140,141]. Spectral counting is defined as the total number of spectra identified for a protein which in term is used to infer relative protein abundance. A more elegant label-free method of quantification is to compare identical peptides from different samples. However,

this requires that besides the sample preparation also the LC-MS runs are to be stable and reproducible in terms of peak shape and retention time. The reproducibility is required as the aforementioned features are extracted, aligned using normalization and subsequently used for quantification[142,143]. A caveat of label-free proteomics is that it requires most replicate measurements in order to boost the confidence in the peptide and protein quantitation.

Isotope labeling quantification

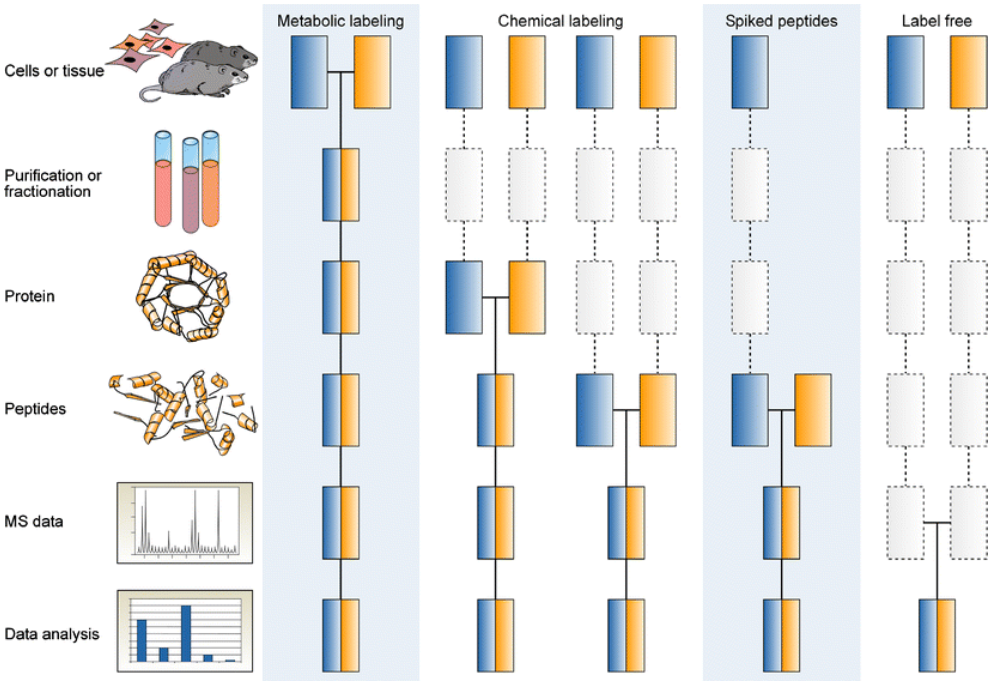

Figure 6: Quantitative mass spectrometry–based proteomics workflows. Boxes in blue and yellow represent two experimental conditions. Horizontal lines indicate when samples are combined and the dashed lines indicate points at which experimental variation and thus quantitation errors can occur.

Adapted from Ong *et al*[138].

Labeling strategies in quantitative mass spectrometry are based on the incorporation of different stable isotopes in or onto the peptides to distinguish between samples. The introduction of the stable isotopes can be done at both the protein and peptide level as a variety of methods have been developed. When analyzing identical peptide sequences from differently labeled samples, their chemical properties remain the same but will be distinguishable through the different masses of the labels. Additional steps are required before or during the proteomics workflow, but in doing so variation introduced in sample preparation will be equal for both samples. This will result in a more accurate determination of changes in proteome between the samples.

Metabolic labeling

Metabolic labeling is achieved by growing cells or organisms on a heavy isotope enriched nutrition source or by means of substituting a specific amino acid containing either ^{13}C and/or ^{15}N instead of the naturally abundant ^{12}C and ^{14}N isotopes[144,145]. The most popular way of metabolic labeling is stable isotope labeling by amino acids in cell culture (SILAC)[145]. This method is based on adding heavy isotopes to the media substituting their natural counterparts. The cells will go through its regular growth and division cycles until approximately 98% of their synthesized proteins contain these specific heavy isotopes. The choice of heavy amino acid to incorporate is important, as the cells will have to be auxotrophic for the amino acid of choice, an added constraint for proteomics is that the amino acid chosen should be present in all peptides. Considering the majority of proteomics experiments relies on the use of trypsin as proteolytic enzyme, isotope labeled lysine and arginine would be recommended as this will ensure each peptide with the exception of the C-terminal peptide of the protein will have the isotope incorporated. The advantage of incorporating the labels this early in the proteomics workflow is that samples can be mixed at a cellular level or immediately after lysis, thus introducing minimal variations and yielding the highest accuracy of protein ratio determination. However, maintaining the model system on expensive media can be both costly and time consuming, limiting its use. Furthermore, certain cell lines and organisms are unable to implement the labels and the number of samples that can be compare is typically limited to two or three. However, recently a novel approach was introduced which drastically extended the multiplexing capacity of metabolic labeling based on neutron encoding (NeuCode), introduced by Coon et al[146]. NeuCode exploits the millidalton range differences of stable isotopes (e.g. a peptide with six ^{13}C atoms and two ^{15}N atoms would result in an additional 8.0142 Da whereas a peptide containing eight ^{2}H atoms would result in an additional 8.0502 Da, the difference of 36 mDa between these 2 labels can be measured using a high-resolution mass spectrometer.

Chemical labeling

For aforementioned samples unsuitable for metabolic labeling such as human primary cells, chemical labeling provides a good alternative since it can be applied to every biological sample after sample lysis. A wide variety of isotopic labels can be introduced at either the protein or the peptide level. Detection of the chemical label is performed at the MS1 level for isotopic labels where a mass difference distinguishes between the labeled peptides, isobaric labels are detected at the MS2 level where a reporter ion is measured to yield a peptide ratio. Although the labels are introduced at a later stage in the proteomics workflow, it has been concluded in a recent study that this type of labeling has a similar accuracy compared to metabolic labeling[147].

Chemical labeling at the protein level can be done using isotope-coded affinity tag (ICAT)[148]. In this chemical labeling method the labeling reaction is implemented following protein reduction. The label consists of a thio-reactive group, an isotopic linker region and an acid cleavable biotin moiety. The thio-reactive group acts on the free reactive cysteines made available by reduction, the linker region which contains either regular or ^{13}C residues depending on the label used and a cleavable tag for affinity purification. Implementing the label at this stage allows for pooling the samples prior digestion which eliminates that variability. However, since only peptides containing cysteines are used, the number of peptides available for quantification is low. Quantification occurs on the MS1 level.

For labeling on the peptide level several strategies are available, most target the primary amine groups present lysines and free peptide/protein amino termini. One of these strategies is stable isotope dimethyl labeling[149,150]. This strategy involves the derivatization of amino groups using formaldehyde and cyanoborohydride with a different number of ^{13}C and deuterium isotopes to create dimethylated peptides each sample is apart a minimum of 4 Da and is measured at the MS1 level. The strategy is fast, cost-effective, has a high labeling efficiency and can be applied to any type of biological sample. There is however a potential disadvantage that has to be taken into account, the chemical properties of deuterium are slightly different compared to hydrogen and can result in a subtle shift in LC retention time which can have a detrimental effect on quantification accuracy.[151]

Isobaric chemical labels are a popular alternative for targeting primary amines. There are two commercially available variants of this type of chemical labeling: isobaric tags for relative and absolute quantification (iTRAQ)[152] and tandem mass tag (TMT)[153] allowing comparison of up to ten different samples[154]. Both strategies rely on the incorporation of a reporter moiety with differential mass and a balance group that combined have an identical mass for each label. The result is that peptides originating from different samples and differentially labeled have identical mass at the MS1 stage. The advantage is that the MS1 spectra are less complex and have higher precursor signals. Upon precursor fragmentation, the isobaric tags fall apart into different reporter ions (typically 1 Da apart) and neutral balance groups. The quantitative information is in these cases extracted from the reporter ions in the MS2 spectra, contrary to previously described strategies. A potential issue with this type quantification is that peptides with similar precursors masses are co-isolated and fragmented with the peptide of interest and thus have a detrimental effect on the quantification accuracy[155]. This can be overcome by using an MS3 method[156], which has its own drawback in terms of a longer duty cycle and lower sensitivity that in turn negatively affects the quantification rate. In order to perform this quantification method a mass spectrometer capable of measuring low m/z values is required in contrast to other quantification methods.

Absolute quantitation

There are different strategies available for the absolute quantification of protein copy numbers. The most common approach is the use of stable isotope labeled standards these are reference peptides with a known concentration and is commonly denoted in copy number per cell[157,158]. Contrary to the previous chemical labeling strategies, the sample is not modified but a labeled standard is introduced in each sample. It is essentially an improved variant of label-free quantification as the processing of the sample is similar. The ratio between the stable isotope labeled standard and the natural peptides is then used to improve the precision of quantification between different samples and thus yielding an improvement in overall accuracy. The heavy standard can be introduced in the form of peptides after digestion (commonly referred to as AQUA)[158] or before digestion using biosynthesized heavy proteotypic peptide concatamers (QconCat)[159] or even synthetic heavy full length proteins (PSAQ)[160]. The latter method being preferred as it minimizes variability in terms of digestion, solubility and protein enrichment. Since the synthesis of fully labeled proteins is still expensive, the labeled peptides are more frequently used. The throughput is less compared to other labeling strategies due to a lack of possible multiplexing.

1.3 Thesis outline

This thesis describes in the following chapters the proteome analysis of primary human circulating immune cells. The challenges and opportunities in biomarker research associated with the clinical application of proteomics are explored. A more fundamental question about the protein constituents of several of these cells is being addressed as well as the possibility to use proteomics to define their secreted factors.

Chapter two describes some of the challenges associated with an important clinical application of proteomics, the quest for protein biomarkers. The object of this study is to find a potential biomarker for coronary artery disease in blood platelets. Case and control subjects were compared by labeling a pool of each with stable isotopes. Using a stringent p-value several candidate biomarkers were identified, however upon closer inspection unavoidable and variable contamination originating from the patient matrix such as serum, erythrocyte and leukocyte components became apparent. Those contaminants were traced back to individual patients using label-free quantitation and revealed that it is imperative to interface the proteome one is studying with the proteome of possible contaminants to refine the list of potential biomarkers.

In chapter three the focus is on defining a subset of secreted proteins. A novel method for both identifying and quantifying the proteins secreted by blood platelets upon stimulation using stable isotope labeling was developed. The copy numbers of the platelet proteome were monitored in both resting state and stimulated state, yielding a remarkably small percentage of significantly released proteins. These released proteins span a concentration range of at least 5 orders as confirmed by ELISA.

Chapter four is comprised of a preliminary elucidation and comparison of the proteomes of neutrophil, eosinophil and basophil granulocytes in order to determine what separates these cell types at the proteome level from each other. For each of these cell types a protein signature has been established indicative of their diverging functions.

References

1. Parkin, J. & Cohen, B. An overview of the immune system. ***Lancet*** 357, 1777-1789, doi:10.1016/S0140-6736(00)04904-7 (2001).
2. Janeway, C. A., Jr. & Medzhitov, R. Innate immune recognition. ***Annu Rev Immunol*** 20, 197-216, doi:10.1146/annurev.immunol.20.083001.084359 (2002).
3. Delves, P. J. & Roitt, I. M. The immune system. First of two parts. ***N Engl J Med*** 343, 37-49, doi:10.1056/NEJM200007063430107 (2000).
4. Delves, P. J. & Roitt, I. M. The immune system. Second of two parts. ***N Engl J Med*** 343, 108-117, doi:10.1056/NEJM200007133430207 (2000).
5. Patel, S. R., Hartwig, J. H. & Italiano, J. E., Jr. The biogenesis of platelets from megakaryocyte proplatelets. ***J Clin Invest*** 115, 3348-3354, doi:10.1172/JCI26891 (2005).
6. McRedmond, J. P. ***et al.*** Integration of proteomics and genomics in platelets: a profile of platelet proteins and platelet-specific genes. ***Mol Cell Proteomics*** 3, 133-144, doi:10.1074/mcp.M300063-MCP200 (2004).
7. Jackson, S. P. The growing complexity of platelet aggregation. ***Blood*** 109, 5087-5095, doi:10.1182/blood-2006-12-027698 (2007).
8. Varga-Szabo, D., Pleines, I. & Nieswandt, B. Cell adhesion mechanisms in platelets. ***Arterioscler Thromb Vasc Biol*** 28, 403-412, doi:10.1161/ATVBAHA.107.150474 (2008).
9. Kroll, M. H., Hellums, J. D., McIntire, L. V., Schafer, A. I. & Moake, J. L. Platelets and shear stress. ***Blood*** 88, 1525-1541 (1996).
10. Nesbitt, W. S. ***et al.*** A shear gradient-dependent platelet aggregation mechanism drives thrombus formation. ***Nat Med*** 15, 665-673, doi:10.1038/nm.1955 (2009).
11. Weiss, H. J., Turitto, V. T. & Baumgartner, H. R. Effect of shear rate on platelet interaction with subendothelium in citrated and native blood. I. Shear rate--dependent decrease of adhesion in von Willebrand's disease and the Bernard-Soulier syndrome. ***J Lab Clin Med*** 92, 750-764 (1978).
12. Savage, B., Almus-Jacobs, F. & Ruggeri, Z. M. Specific synergy of multiple substrate-receptor interactions in platelet thrombus formation under flow. ***Cell*** 94, 657-666 (1998).
13. Morton, L. F., Hargreaves, P. G., Farndale, R. W., Young, R. D. & Barnes, M. J. Integrin alpha 2 beta 1-independent activation of platelets by simple collagen-like peptides: collagen tertiary (triple-helical) and quaternary (polymeric) structures are sufficient alone for alpha 2 beta 1-independent platelet reactivity. ***Biochem J*** 306 (Pt 2), 337-344 (1995).
14. De Clerck, F., Xhonneux, B., Leysen, J. & Janssen, P. A. Evidence for functional 5-HT2 receptor sites on human blood platelets. ***Biochem Pharmacol*** 33, 2807-2811 (1984).
15. Hechler, B. & Gachet, C. P2 receptors and platelet function. ***Purinergic Signal*** 7, 293-303, doi:10.1007/s11302-011-9247-6 (2011).
16. McNicol, A. & Israels, S. J. Platelet dense granules: structure, function and implications for haemostasis. ***Thromb Res*** 95, 1-18 (1999).
17. Coughlin, S. R. Thrombin signalling and protease-activated receptors. ***Nature*** 407, 258-264, doi:10.1038/35025229 (2000).

18. Morrell, C. N., Aggrey, A. A., Chapman, L. M. & Modjeski, K. L. Emerging roles for platelets as immune and inflammatory cells. ***Blood*** 123, 2759-2767, doi:10.1182/blood-2013-11-462432 (2014).
19. Del Conde, I., Cruz, M. A., Zhang, H., Lopez, J. A. & Afshar-Kharghan, V. Platelet activation leads to activation and propagation of the complement system. ***J Exp Med*** 201, 871-879, doi:10.1084/jem.20041497 (2005).
20. Czapiga, M., Kirk, A. D. & Lekstrom-Himes, J. Platelets deliver costimulatory signals to antigen-presenting cells: a potential bridge between injury and immune activation. ***Exp Hematol*** 32, 135-139, doi:10.1016/j.exphem.2003.11.004 (2004).
21. Silverstein, R. L. & Nachman, R. L. Thrombospondin binds to monocytes-macrophages and mediates platelet-monocyte adhesion. ***J Clin Invest*** 79, 867-874, doi:10.1172/JCI112896 (1987).
22. Clark, S. R. ***et al.*** Platelet TLR4 activates neutrophil extracellular traps to ensnare bacteria in septic blood. ***Nat Med*** 13, 463-469, doi:10.1038/nm1565 (2007).
23. Stahl, A. L. ***et al.*** Lipopolysaccharide from enterohemorrhagic Escherichia coli binds to platelets through TLR4 and CD62 and is detected on circulating platelets in patients with hemolytic uremic syndrome. ***Blood*** 108, 167-176, doi:10.1182/blood-2005-08-3219 (2006).
24. Semple, J. W., Aslam, R., Kim, M., Speck, E. R. & Freedman, J. Platelet-bound lipopolysaccharide enhances Fc receptor-mediated phagocytosis of IgG-opsonized platelets. ***Blood*** 109, 4803-4805, doi:10.1182/blood-2006-12-062695 (2007).
25. Karasuyama, H., Obata, K., Wada, T., Tsujimura, Y. & Mukai, K. Newly appreciated roles for basophils in allergy and protective immunity. ***Allergy*** 66, 1133-1141, doi:10.1111/j.1398-9995.2011.02613.x (2011).
26. Jacobsen, E. A., Taranova, A. G., Lee, N. A. & Lee, J. J. Eosinophils: singularly destructive effector cells or purveyors of immunoregulation? ***J Allergy Clin Immunol*** 119, 1313-1320, doi:10.1016/j.jaci.2007.03.043 (2007).
27. Brinkmann, V. ***et al.*** Neutrophil extracellular traps kill bacteria. ***Science*** 303, 1532-1535, doi:10.1126/science.1092385 (2004).
28. Geering, B., Stoeckle, C., Conus, S. & Simon, H. U. Living and dying for inflammation: neutrophils, eosinophils, basophils. ***Trends Immunol*** 34, 398-409, doi:10.1016/j.it.2013.04.002 (2013).
29. Mantovani, A., Cassatella, M. A., Costantini, C. & Jaillon, S. Neutrophils in the activation and regulation of innate and adaptive immunity. ***Nat Rev Immunol*** 11, 519-531, doi:10.1038/nri3024 (2011).
30. Pillay, J. ***et al.*** In vivo labeling with 2H2O reveals a human neutrophil lifespan of 5.4 days. ***Blood*** 116, 625-627, doi:10.1182/blood-2010-01-259028 (2010).
31. Witko-Sarsat, V., Rieu, P., Descamps-Latscha, B., Lesavre, P. & Halbwachs-Mecarelli, L. Neutrophils: molecules, functions and pathophysiological aspects. ***Lab Invest*** 80, 617-653 (2000).
32. Mayadas, T. N., Cullere, X. & Lowell, C. A. The multifaceted functions of neutrophils. ***Annu Rev Pathol*** 9, 181-218, doi:10.1146/annurev-pathol-020712-164023 (2014).
33. Mocsai, A. Diverse novel functions of neutrophils in immunity, inflammation, and beyond. ***J Exp Med*** 210, 1283-1299, doi:10.1084/jem.20122220 (2013).

34. Kolaczkowska, E. & Kubes, P. Neutrophil recruitment and function in health and inflammation. ***Nat Rev Immunol*** 13, 159-175, doi:10.1038/nri3399 (2013).
35. Rorvig, S., Ostergaard, O., Heegaard, N. H. & Borregaard, N. Proteome profiling of human neutrophil granule subsets, secretory vesicles, and cell membrane: correlation with transcriptome profiling of neutrophil precursors. ***J Leukoc Biol*** 94, 711-721, doi:10.1189/jlb.1212619 (2013).
36. Borregaard, N. Neutrophils, from marrow to microbes. ***Immunity*** 33, 657-670, doi:10.1016/j.immuni.2010.11.011 (2010).
37. Borregaard, N. & Cowland, J. B. Granules of the human neutrophilic polymorphonuclear leukocyte. ***Blood*** 89, 3503-3521 (1997).
38. Le Cabec, V., Cowland, J. B., Calafat, J. & Borregaard, N. Targeting of proteins to granule subsets is determined by timing and not by sorting: The specific granule protein NGAL is localized to azurophil granules when expressed in HL-60 cells. ***Proc Natl Acad Sci U S A*** 93, 6454-6457 (1996).
39. Sengelov, H. ***et al.*** Mobilization of granules and secretory vesicles during in vivo exudation of human neutrophils. ***J Immunol*** 154, 4157-4165 (1995).
40. Klebanoff, S. J., Kettle, A. J., Rosen, H., Winterbourn, C. C. & Nauseef, W. M. Myeloperoxidase: a front-line defender against phagocytosed microorganisms. ***J Leukoc Biol*** 93, 185-198, doi:10.1189/jlb.0712349 (2013).
41. Winterbourn, C. C. & Kettle, A. J. Redox reactions and microbial killing in the neutrophil phagosome. ***Antioxid Redox Signal*** 18, 642-660, doi:10.1089/ars.2012.4827 (2013).
42. Goetz, D. H. ***et al.*** The neutrophil lipocalin NGAL is a bacteriostatic agent that interferes with siderophore-mediated iron acquisition. ***Mol Cell*** 10, 1033-1043 (2002).
43. Flo, T. H. ***et al.*** Lipocalin 2 mediates an innate immune response to bacterial infection by sequestrating iron. ***Nature*** 432, 917-921, doi:10.1038/nature03104 (2004).
44. Gudmundsson, G. H. ***et al.*** The human gene FALL39 and processing of the cathelin precursor to the antibacterial peptide LL-37 in granulocytes. ***Eur J Biochem*** 238, 325-332 (1996).
45. Devarajan, P., Johnston, J. J., Ginsberg, S. S., Van Wart, H. E. & Berliner, N. Structure and expression of neutrophil gelatinase cDNA. Identity with type IV collagenase from HT1080 cells. ***J Biol Chem*** 267, 25228-25232 (1992).
46. Opdenakker, G. ***et al.*** Gelatinase B functions as regulator and effector in leukocyte biology. ***J Leukoc Biol*** 69, 851-859 (2001).
47. Mansfield, P. J., Hinkovska-Galcheva, V., Shayman, J. A. & Boxer, L. A. Granulocyte colony-stimulating factor primes NADPH oxidase in neutrophils through translocation of cytochrome b(558) by gelatinase-granule release. ***J Lab Clin Med*** 140, 9-16 (2002).
48. Borregaard, N. ***et al.*** Stimulus-dependent secretion of plasma proteins from human neutrophils. ***J Clin Invest*** 90, 86-96, doi:10.1172/JCI115860 (1992).
49. Sengelov, H., Kjeldsen, L. & Borregaard, N. Control of exocytosis in early neutrophil activation. ***J Immunol*** 150, 1535-1543 (1993).
50. Brumell, J. H. ***et al.*** Subcellular distribution of docking/fusion proteins in neutrophils, secretory cells with multiple exocytic compartments. ***J Immunol*** 155, 5750-5759 (1995).

51. Sengelov, H., Boulay, F., Kjeldsen, L. & Borregaard, N. Subcellular localization and translocation of the receptor for N-formylmethionyl-leucyl-phenylalanine in human neutrophils. ***Biochem J*** 299 (Pt 2), 473-479 (1994).
52. Borregaard, N., Miller, L. J. & Springer, T. A. Chemoattractant-regulated mobilization of a novel intracellular compartment in human neutrophils. ***Science*** 237, 1204-1206 (1987).
53. Hartl, D. ***et al.*** Infiltrated neutrophils acquire novel chemokine receptor expression and chemokine responsiveness in chronic inflammatory lung diseases. ***J Immunol*** 181, 8053-8067 (2008).
54. Sengelov, H., Kjeldsen, L., Kroeze, W., Berger, M. & Borregaard, N. Secretory vesicles are the intracellular reservoir of complement receptor 1 in human neutrophils. ***J Immunol*** 153, 804-810 (1994).
55. Blanchard, C. & Rothenberg, M. E. Biology of the eosinophil. ***Adv Immunol*** 101, 81-121, doi:10.1016/S0065-2776(08)01003-1 (2009).
56. Giembycz, M. A. & Lindsay, M. A. Pharmacology of the eosinophil. ***Pharmacol Rev*** 51, 213-340 (1999).
57. Rothenberg, M. E. & Hogan, S. P. The eosinophil. ***Annu Rev Immunol*** 24, 147-174, doi:10.1146/annurev.immunol.24.021605.090720 (2006).
58. Rosenberg, H. F., Dyer, K. D. & Foster, P. S. Eosinophils: changing perspectives in health and disease. ***Nat Rev Immunol*** 13, 9-22, doi:10.1038/nri3341 (2013).
59. Fabre, V. ***et al.*** Eosinophil deficiency compromises parasite survival in chronic nematode infection. ***Journal of immunology*** 182, 1577-1583 (2009).
60. Gebreselassie, N. G. ***et al.*** Eosinophils preserve parasitic nematode larvae by regulating local immunity. ***J Immunol*** 188, 417-425, doi:10.4049/jimmunol.1101980 (2012).
61. Jacobsen, E. A., Ochkur, S. I., Lee, N. A. & Lee, J. J. Eosinophils and asthma. ***Current allergy and asthma reports*** 7, 18-26 (2007).
62. Wegmann, M. Targeting eosinophil biology in asthma therapy. ***Am J Respir Cell Mol Biol*** 45, 667-674, doi:10.1165/rcmb.2011-0013TR (2011).
63. Lee, J. J. ***et al.*** Defining a link with asthma in mice congenitally deficient in eosinophils. ***Science*** 305, 1773-1776, doi:10.1126/science.1099472 (2004).
64. Yu, C. ***et al.*** Targeted deletion of a high-affinity GATA-binding site in the GATA-1 promoter leads to selective loss of the eosinophil lineage in vivo. ***The Journal of experimental medicine*** 195, 1387-1395 (2002).
65. Neves, J. S. ***et al.*** Eosinophil granules function extracellularly as receptor-mediated secretory organelles. ***Proceedings of the National Academy of Sciences of the United States of America*** 105, 18478-18483, doi:10.1073/pnas.0804547105 (2008).
66. Spencer, L. A. ***et al.*** Cytokine receptor-mediated trafficking of preformed IL-4 in eosinophils identifies an innate immune mechanism of cytokine secretion. ***Proceedings of the National Academy of Sciences of the United States of America*** 103, 3333-3338, doi:10.1073/pnas.0508946103 (2006).
67. Wang, H. B., Ghiran, I., Matthaei, K. & Weller, P. F. Airway eosinophils: allergic inflammation recruited professional antigen-presenting cells. ***Journal of immunology*** 179, 7585-7592 (2007).

68. Spencer, L. A. ***et al.*** Human eosinophils constitutively express multiple Th1, Th2, and immunoregulatory cytokines that are secreted rapidly and differentially. ***J Leukoc Biol*** 85, 117-123, doi:10.1189/jlb.0108058 (2009).
69. Wang, H. B. & Weller, P. F. Pivotal advance: eosinophils mediate early alum adjuvant-elicited B cell priming and IgM production. ***Journal of leukocyte biology*** 83, 817-821, doi:10.1189/jlb.0607392 (2008).
70. Yang, D. ***et al.*** Eosinophil-derived neurotoxin (EDN), an antimicrobial protein with chemotactic activities for dendritic cells. ***Blood*** 102, 3396-3403, doi:10.1182/blood-2003-01-0151 (2003).
71. Yang, D. ***et al.*** Eosinophil-derived neurotoxin acts as an alarmin to activate the TLR2-MyD88 signal pathway in dendritic cells and enhances Th2 immune responses. ***The Journal of experimental medicine*** 205, 79-90, doi:10.1084/jem.20062027 (2008).
72. Elishmereni, M. ***et al.*** Physical interactions between mast cells and eosinophils: a novel mechanism enhancing eosinophil survival in vitro. ***Allergy*** 66, 376-385, doi:10.1111/j.1398-9995.2010.02494.x (2011).
73. Liu, L. Y. ***et al.*** Generation of Th1 and Th2 chemokines by human eosinophils: evidence for a critical role of TNF-alpha. ***J Immunol*** 179, 4840-4848 (2007).
74. Kolset, S. O., Prydz, K. & Pejler, G. Intracellular proteoglycans. ***Biochem J*** 379, 217-227, doi:10.1042/BJ20031230 (2004).
75. Hogan, S. P. ***et al.*** Eosinophils: biological properties and role in health and disease. ***Clin Exp Allergy*** 38, 709-750, doi:10.1111/j.1365-2222.2008.02958.x (2008).
76. Durrani, S. R. ***et al.*** Innate immune responses to rhinovirus are reduced by the high-affinity IgE receptor in allergic asthmatic children. ***J Allergy Clin Immunol*** 130, 489-495, doi:10.1016/j.jaci.2012.05.023 (2012).
77. Stone, K. D., Prussin, C. & Metcalfe, D. D. IgE, mast cells, basophils, and eosinophils. ***J Allergy Clin Immunol*** 125, S73-80, doi:10.1016/j.jaci.2009.11.017 (2010).
78. Knol, E. F. & Olszewski, M. Basophils and mast cells: Underdog in immune regulation? ***Immunol Lett*** 138, 28-31, doi:10.1016/j.imlet.2011.02.012 (2011).
79. Cromheecke, J. L., Nguyen, K. T. & Huston, D. P. Emerging role of human basophil biology in health and disease. ***Curr Allergy Asthma Rep*** 14, 408, doi:10.1007/s11882-013-0408-2 (2014).
80. Knol, E. F. & Gibbs, B. F. Editorial: Basophil survival and immunomodulatory function are uniquely regulated by a novel MyD88-dependent pathway. ***J Leukoc Biol*** 86, 753-755, doi:10.1189/jlb.0409248 (2009).
81. Schroeder, J. T. Basophils: emerging roles in the pathogenesis of allergic disease. ***Immunol Rev*** 242, 144-160, doi:10.1111/j.1600-065X.2011.01023.x (2011).
82. James, P. Protein identification in the post-genome era: the rapid rise of proteomics. ***Q Rev Biophys*** 30, 279-331 (1997).
83. Pandey, A. & Mann, M. Proteomics to study genes and genomes. ***Nature*** 405, 837-846, doi:10.1038/35015709 (2000).
84. Wilkins, M. R. ***et al.*** From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. ***Biotechnology (N Y)*** 14, 61-65 (1996).

85. Lander, E. S. ***et al.*** Initial sequencing and analysis of the human genome. ***Nature*** 409, 860-921, doi:10.1038/35057062 (2001).
86. Altelaar, A. F., Munoz, J. & Heck, A. J. Next-generation proteomics: towards an integrative view of proteome dynamics. ***Nat Rev Genet*** 14, 35-48, doi:10.1038/nrg3356 (2013).
87. Steen, H. & Mann, M. The ABC's (and XYZ's) of peptide sequencing. ***Nat Rev Mol Cell Biol*** 5, 699-711, doi:10.1038/nrm1468 (2004).
88. Yates, J. R., Ruse, C. I. & Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. ***Annu Rev Biomed Eng*** 11, 49-79, doi:10.1146/annurev-bioeng-061008-124934 (2009).
89. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. ***Nature*** 422, 198-207, doi:10.1038/nature01511 (2003).
90. Tran, J. C. ***et al.*** Mapping intact protein isoforms in discovery mode using top-down proteomics. ***Nature*** 480, 254-258, doi:10.1038/nature10575 (2011).
91. Olsen, J. V., Ong, S. E. & Mann, M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. ***Mol Cell Proteomics*** 3, 608-614, doi:10.1074/mcp.T400003-MCP200 (2004).
92. Wisniewski, J. R. & Mann, M. Consecutive proteolytic digestion in an enzyme reactor increases depth of proteomic and phosphoproteomic analysis. ***Anal Chem*** 84, 2631-2637, doi:10.1021/ac300006b (2012).
93. Glatter, T. ***et al.*** Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion. ***J Proteome Res*** 11, 5145-5156, doi:10.1021/pr300273g (2012).
94. Isobe, T., Takayasu, T., Takai, N. & Okuyama, T. High-performance liquid chromatography of peptides on a macroreticular cation-exchange resin: application to peptide mapping of Bence--Jones proteins. ***Anal Biochem*** 122, 417-425 (1982).
95. Mant, C. T. & Hodges, R. S. Separation of peptides by strong cation-exchange high-performance liquid chromatography. ***J Chromatogr*** 327, 147-155 (1985).
96. Cachia, P. J., van Eyk, J. E., Chong, P. C. S., Taneja, A. & Hodges, R. S. Separation of basic peptides by cation-exchange high-performance liquid chromatography. ***Journal of Chromatography A*** 266, 651-659 (1983).
97. Gruber, K. A., Stein, S., Brink, L., Radhakrishnan, A. & Udenfriend, S. Fluorometric assay of vasopressin and oxytocin: a general approach to the assay of peptides in tissues. ***Proc Natl Acad Sci U S A*** 73, 1314-1318 (1976).
98. Shen, Y. ***et al.*** Automated 20 kpsi RPLC-MS and MS/MS with chromatographic peak capacities of 1000-1500 and capabilities in proteomics and metabolomics. ***Anal Chem*** 77, 3090-3100, doi:10.1021/ac0483062 (2005).
99. Delahunty, C. & Yates, J. R., 3rd. Protein identification using 2D-LC-MS/MS. ***Methods*** 35, 248-255, doi:10.1016/j.ymeth.2004.08.016 (2005).
100. Hillenkamp, F., Karas, M. & Holtkamp, D. K., P. Energy deposition in ultraviolet laser desorption mass spectrometry of biomolecules. ***International Journal of Mass Spectrometry and Ion Processes*** 69, 265-276 (1986).

101. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. ***Science*** 246, 64-71 (1989).
102. Cohen, S. L. & Chait, B. T. Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. ***Anal Chem*** 68, 31-37 (1996).
103. Gomez, A. & Tang, K. Charge and fission of droplets in electrostatic sprays. ***Physics of Fluids*** 6, 404-414 (1994).
104. Iribarne, J. V. & Thomson, B. A. On the evaporation of small ions from charged droplets. ***Journal of Chemical Physics*** 64, 2287-2295 (1976).
105. Wolff, M. M. S., W.E. A pulsed mass spectrometer with time dispersion. ***Review of Scientific Instruments*** 24, 616-617 (1953).
106. Grix, R., Kutscher, R. L., G., Grüner, U., Wollnik, H. & Matsuda, H. A time-of-flight mass analyzer with high resolving power. ***Rapid Communications in Mass Spectrometry*** 2, 83-85 (1988).
107. Wolfgang, P. S., H. Ein neues Massenspektrometer ohne Magnetfeld. ***Zeitschrift für Naturforschung A*** 8, 448-450 (1957).
108. de Hoffmann, E. S., V. Mass Spectrometry: Principles and Applications, 3rd Edition. ***John Wiley (distributor), Hoboken, N.J.*** (2007).
109. Hager, J. W. & Le Blanc, J. C. High-performance liquid chromatography-tandem mass spectrometry with a new quadrupole/linear ion trap instrument. ***J Chromatogr A*** 1020, 3-9 (2003).
110. Douglas, D. J., Frank, A. J. & Mao, D. Linear ion traps in mass spectrometry. ***Mass Spectrom Rev*** 24, 1-29, doi:10.1002/mas.20004 (2005).
111. Schwartz, J. C., Senko, M. W. & Syka, J. E. A two-dimensional quadrupole ion trap mass spectrometer. ***J Am Soc Mass Spectrom*** 13, 659-669, doi:10.1016/S1044-0305(02)00384-7 (2002).
112. Page, J. S. ***et al.*** Automatic gain control in mass spectrometry using a jet disrupter electrode in an electrodynamic ion funnel. ***J Am Soc Mass Spectrom*** 16, 244-253, doi:10.1016/j.jasms.2004.11.003 (2005).
113. Makarov, A. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. ***Anal Chem*** 72, 1156-1162 (2000).
114. Kingdon, K. H. A Method for the Neutralization of Electron Space Charge by Positive Ionization at Very Low Gas Pressures. ***Physical Review*** 21 (1923).
115. Senko, M. W., Canterbury, J. D., Guan, S. & Marshall, A. G. A high-performance modular data system for Fourier transform ion cyclotron resonance mass spectrometry. ***Rapid communications in mass spectrometry : RCM*** 10, 1839-1844, doi:10.1002/(SICI)1097-0231(199611)10:14<1839::AID-RCM718>3.0.CO;2-V (1996).
116. Michalski, A. ***et al.*** Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. ***Mol Cell Proteomics*** 11, O111 013698, doi:10.1074/mcp.O111.013698 (2012).
117. Olsen, J. V. ***et al.*** Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. ***Mol Cell Proteomics*** 4, 2010-2021, doi:10.1074/mcp.T500030-MCP200 (2005).
118. Makarov, A. ***et al.*** Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. ***Anal Chem*** 78, 2113-2120, doi:10.1021/ac0518811 (2006).

119. Roepstorff, P. & Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. ***Biomed Mass Spectrom*** 11, 601, doi:10.1002/bms.1200111109 (1984).
120. Biemann, K. Mass spectrometry of peptides and proteins. ***Annu Rev Biochem*** 61, 977-1010, doi:10.1146/annurev.bi.61.070192.004553 (1992).
121. McLuckey, S. A. Principles of collisional activation in analytical mass spectrometry. ***J Am Soc Mass Spectrom*** 3, 599-614, doi:10.1016/1044-0305(92)85001-Z (1992).
122. Paizs, B. & Suhai, S. Fragmentation pathways of protonated peptides. ***Mass Spectrom Rev*** 24, 508-548, doi:10.1002/mas.20024 (2005).
123. Sleno, L. & Volmer, D. A. Ion activation methods for tandem mass spectrometry. ***J Mass Spectrom*** 39, 1091-1112, doi:10.1002/jms.703 (2004).
124. Zubarev, R. A., Kelleher, N. L. & McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. ***J. Am. Chem. Soc*** 120 (1998).
125. Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J. & Hunt, D. F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. ***Proc Natl Acad Sci U S A*** 101, 9528-9533, doi:10.1073/pnas.0402700101 (2004).
126. Good, D. M., Wirtala, M., McAlister, G. C. & Coon, J. J. Performance characteristics of electron transfer dissociation mass spectrometry. ***Mol Cell Proteomics*** 6, 1942-1951, doi:10.1074/mcp.M700073-MCP200 (2007).
127. Pitteri, S. J., Chrisman, P. A. & McLuckey, S. A. Electron-transfer ion/ion reactions of doubly protonated peptides: effect of elevated bath gas temperature. ***Anal Chem*** 77, 5662-5669, doi:10.1021/ac050666h (2005).
128. Frese, C. K. ***et al.*** Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. ***J Proteome Res*** 10, 2377-2388, doi:10.1021/pr1011729 (2011).
129. Swaney, D. L., McAlister, G. C. & Coon, J. J. Decision tree-driven tandem mass spectrometry for shotgun proteomics. ***Nature methods*** 5, 959-964, doi:10.1038/nmeth.1260 (2008).
130. Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. ***Electrophoresis*** 20, 3551-3567, doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2 (1999).
131. Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. ***J Am Soc Mass Spectrom*** 5, 976-989, doi:10.1016/1044-0305(94)80016-2 (1994).
132. Cox, J. ***et al.*** Andromeda: a peptide search engine integrated into the MaxQuant environment. ***J Proteome Res*** 10, 1794-1805, doi:10.1021/pr101065j (2011).
133. Duncan, D. T., Craig, R. & Link, A. J. Parallel tandem: a program for parallel processing of tandem mass spectra using PVM or MPI and X!Tandem. ***J Proteome Res*** 4, 1842-1847, doi:10.1021/pr050058i (2005).
134. Higgs, R. E. ***et al.*** Estimating the statistical significance of peptide identifications from shotgun proteomics experiments. ***J Proteome Res*** 6, 1758-1767, doi:10.1021/pr0605320 (2007).
135. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. ***Nature methods*** 4, 207-214, doi:10.1038/nmeth1019 (2007).

136. Weatherly, D. B. ***et al.*** A Heuristic method for assigning a false-discovery rate for protein identifications from Mascot database search results. ***Mol Cell Proteomics*** 4, 762-772, doi:10.1074/mcp.M400215-MCP200 (2005).
137. Kall, L., Canterbury, J. D., Weston, J., Noble, W. S. & MacCoss, M. J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. ***Nature methods*** 4, 923-925, doi:10.1038/nmeth1113 (2007).
138. Ong, S. E. & Mann, M. Mass spectrometry-based proteomics turns quantitative. ***Nat Chem Biol*** 1, 252-262, doi:10.1038/nchembio736 (2005).
139. Liu, H., Sadygov, R. G. & Yates, J. R., 3rd. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. ***Anal Chem*** 76, 4193-4201, doi:10.1021/ac0498563 (2004).
140. Chelius, D. & Bondarenko, P. V. Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. ***J Proteome Res*** 1, 317-323 (2002).
141. Old, W. M. ***et al.*** Comparison of label-free methods for quantifying human proteins by shotgun proteomics. ***Mol Cell Proteomics*** 4, 1487-1502, doi:10.1074/mcp.M500084-MCP200 (2005).
142. Christin, C. ***et al.*** Time alignment algorithms based on selected mass traces for complex LC-MS data. ***J Proteome Res*** 9, 1483-1495, doi:10.1021/pr9010124 (2010).
143. Finney, G. L. ***et al.*** Label-free comparative analysis of proteomics mixtures using chromatographic alignment of high-resolution muLC-MS data. ***Anal Chem*** 80, 961-971, doi:10.1021/ac701649e (2008).
144. Gouw, J. W., Tops, B. B. & Krijgsveld, J. Metabolic labeling of model organisms using heavy nitrogen (15N). ***Methods Mol Biol*** 753, 29-42, doi:10.1007/978-1-61779-148-2_2 (2011).
145. Ong, S. E. ***et al.*** Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. ***Mol Cell Proteomics*** 1, 376-386 (2002).
146. Hebert, A. S. ***et al.*** Neutron-encoded mass signatures for multiplexed proteome quantification. ***Nature methods*** 10, 332-334, doi:10.1038/nmeth.2378 (2013).
147. Altelaar, A. F. ***et al.*** Benchmarking stable isotope labeling based quantitative proteomics. ***J Proteomics*** 88, 14-26, doi:10.1016/j.jprot.2012.10.009 (2013).
148. Gygi, S. P. ***et al.*** Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. ***Nat Biotechnol*** 17, 994-999, doi:10.1038/13690 (1999).
149. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. ***Nat Protoc*** 4, 484-494, doi:10.1038/nprot.2009.21 (2009).
150. Hsu, J. L., Huang, S. Y., Chow, N. H. & Chen, S. H. Stable-isotope dimethyl labeling for quantitative proteomics. ***Anal Chem*** 75, 6843-6852, doi:10.1021/ac0348625 (2003).
151. Zhang, R., Sioma, C. S., Thompson, R. A., Xiong, L. & Regnier, F. E. Controlling deuterium isotope effects in comparative proteomics. ***Anal Chem*** 74, 3662-3669 (2002).
152. Ross, P. L. ***et al.*** Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. ***Mol Cell Proteomics*** 3, 1154-1169, doi:10.1074/mcp.M400129-MCP200 (2004).

153. Thompson, A. ***et al.*** Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. ***Anal Chem*** 75, 1895-1904 (2003).
154. Werner, T. ***et al.*** Ion coalescence of neutron encoded TMT 10-plex reporter ions. ***Anal Chem*** 86, 3594-3601, doi:10.1021/ac500140s (2014).
155. Christoforou, A. L. & Lilley, K. S. Isobaric tagging approaches in quantitative proteomics: the ups and downs. ***Anal Bioanal Chem*** 404, 1029-1037, doi:10.1007/s00216-012-6012-9 (2012).
156. Ting, L., Rad, R., Gygi, S. P. & Haas, W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. ***Nature methods*** 8, 937-940, doi:10.1038/nmeth.1714 (2011).
157. Barr, J. R. ***et al.*** Isotope dilution--mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I. ***Clin Chem*** 42, 1676-1682 (1996).
158. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W. & Gygi, S. P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. ***Proc Natl Acad Sci U S A*** 100, 6940-6945, doi:10.1073/pnas.0832254100 (2003).
159. Pratt, J. M. ***et al.*** Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. ***Nat Protoc*** 1, 1029-1043, doi:10.1038/nprot.2006.129 (2006).
160. Brun, V. ***et al.*** Isotope-labeled protein standards: toward absolute quantitative proteomics. ***Mol Cell Proteomics*** 6, 2139-2149, doi:10.1074/mcp.M700163-MCP200 (2007).

2.

Biomarker discovery in circulating cells benefits from curating for prospective blood matrix contaminants

Patrick Wijten[1,2], Onno B. Bleijerveld[1,2,4], Salvatore Cappadona[1,2], Elizabeth McClellan[3], Philip G. de Groot[5], Imo Hoefer[4], Gerard Pasterkamp[4], Albert J.R. Heck[1,2]*, Arjen Scholten[1,2]*

Manuscript in preparation

[1]Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, the Netherlands.

[2]Netherlands Proteomics Centre, Padualaan 8, 3584 CH Utrecht, The Netherlands;

[3]Department of Bioinformatics, Erasmus Medical Centre, Rotterdam, The Netherlands;

[4]Department of Experimental Cardiology, University Medical Centre Utrecht, The Netherlands.

[5]Department of Clinical Chemistry and Haematology, University Medical Center Utrecht, Utrecht.

Abstract

Most protein based biomarker studies target accessible patient material such as plasma and urine, however analysis of these samples is severely hampered by the extreme high abundance of a few proteins. To overcome such unfavorable dynamic range issues, circulating cells (platelets, granulocytes, monocytes etc.), which also hold disease state specific biochemical information, are promising alternatives. A particular challenge is that all circulating cell isolation strategies suffer from potential contamination with other circulating cells and plasma. Here we explore a proteomics approach to identify potential biomarkers for coronary artery disease using carefully isolated platelets of patients and controls. High-end proteomics technology (2D-chromatography, stable isotope labeling, ETD and HCD-fragmentation) is applied on platelets from a small pooled cohort of carefully matched cases and controls. Out of the 2440 proteins quantified, 62 were initially revealed as significantly different (stringent p-value<0.0001). However upon closer inspection, also at individual patient level using label free quantitation, unavoidable and variable contamination with plasma, erythrocyte and leukocyte components became apparent. To deduce protein expression changes related to disease, we present how correlating the quantitative patient/control study with reported quantitative plasma, erythrocyte and leukocyte proteomes is beneficial to omit taking false positive 'novel biomarkers' into the verification phase. Our data reveal that, when contamination issues from other blood matrices are dealt with, circulating cells may provide a good biomarker source.

Introduction

Over the last decade the human plasma proteome has been probed extensively to extract biomarkers for predictive and/or prognostic purposes for a large variety of diseases [1], as it is accessible and routinely collected from patients and controls[2]. Furthermore, no other readily available biofluid has such intimacy with the body as plasma, as it contains tissue leakage and apoptosis products. From a proteome point of view, in plasma, protein abundances span at least 10 orders of magnitude in terms of concentration[3]. Mass spectrometric analysis is severely hampered by this unfavorable dynamic range since 20 proteins comprise 99% of the plasma proteome mass[3]. Despite the high potential, these issues make the mass spectrometry based exploration of biomarkers in plasma one of the most challenging tasks in the field. Illustrative is the overview of the pilot phase of the HUPO Plasma Proteome Project, in which 35 laboratories contributed data to only yield 3020 plasma proteins [4], whereas current efforts in cultured cells nowadays easily exceed the 10.000 mark[5, 6]. More recently the human plasma peptide atlas was released containing a stringently selected reliable set of almost 2000 human plasma proteins[7]. When performing routine analysis of plasma, typical LC-MS experiments result in the identification of only ~200 proteins in a single run[8]. When 2-dimensional peptide chromatography is used in combination with a sophisticated immunodepletion, this number can be raised to 700 [9]. This number is likely too low to find meaningful, often low-abundant, specific novel biomarkers, especially when performing analyses at individual patient level in larger cohorts. These low-abundant proteins typically present themselves in the lower ng/ml to pg/ml concentration range and are therefore lost in the dynamic range[10, 11].

An alternative for plasma is tissue, although this lacks the desired accessibility, or cannot be obtained in ample amounts. Even when patient tissues are relatively easily obtained, choosing the proper control tissue is often the next hurdle. Another challenge is the subsequent clinical implementation of tissue biomarkers. Questions arise, where and how to sample in a non-invasive way, or if this is not possible, whether and how to find back specific tissue biomarkers in plasma or other body fluids. Related to all these challenging issues, to date, only few tissue leakage markers have reached the clinic.

Atherosclerosis is an inflammatory process of the vascular wall, involving many different cell types, including circulating leukocytes[12]. It is initiated by local lipid deposits and leukocyte accumulation in the vessel wall. Histologically, atherosclerotic plaques are usually classified as either stable or unstable plaques [13]. The latter are more likely to rupture, resulting in an acute arterial occlusion, but both can lead to clinical symptoms such as chronic tissue ischemia. Because of the imminent risk of plaque rupture, which may lead to acute ischemic events like myocardial infarction or stroke, biomarkers to predict the likelihood of developing such an event are heavily sought[14-17]. However, the additive value of plasma biomarkers

on top of existing risk scores (e.g. Framingham Risk Score, FRS) so far remains minimal[18]. Therefore, new sources for biomarker discovery should be explored.

Over the past 10 years, some proteomics atherosclerosis biomarker studies were performed employing several different sample types. Early on, 2D-gel electrophoresis comparisons of coronary artery tissue from coronary artery disease patients and healthy individuals revealed the involvement of the usual suspects of inflammatory and stress response proteins such as Hsp20, Hsp27 and Superoxide dismutase[19,20]. Since atherosclerosis is defined by the aberrant interplay between circulating lipids, cells and their interaction with the endothelial of the vasculature, recent efforts have focused on finding biomarkers in vascular tissue (e.g. in atherosclerotic plaques). For instance, we have recently shown that circulating levels of the extracellular matrix protein osteopontin-1 are a good predictor of atherosclerosis risk [21].

In this study we explore an alternative strategy using another type of sample for biomarker discovery: circulating cells. For inflammatory or inflammation-like processes such as atherosclerosis, this is a logical and interesting source of biomarkers, not only because of the above described mechanistic involvement of these cells in atherosclerotic plaque growth and progression, but also because of the better accessibility of circulating cells as compared to vascular tissue, in particular in patients with coronary artery disease. Furthermore, recent studies have indicated the potential of circulating blood cells as biomarkers, either by using FACS analysis to identify specific leukocyte subpopulations as endothelial progenitor cells (EPC) [22], $CD14^{+}CD16^{+}$ monocytes [22] or $CD4^{+}CD28^{-}$ T-cells [23], or by platelet aggregation assays [24]. This indicates the high potential of circulating blood cells as sources of biomarkers for cardiovascular disease. Two proteomics studies have sought protein biomarkers for acute coronary syndrome in circulating monocytes and employed 2D-GE [25,26] which revealed ~1000 spots of which a handful were differential, including for instance protein S100A8, a well-studied general inflammation marker.

Given these prior results, we hypothesized that circulating cells may combine the best of both worlds to search for biomarkers since they are almost as accessible as plasma, but are thought to have much better dynamic range properties. Here we present a systematic, widely applicable analytical platform to discover and verify novel biomarkers for atherosclerosis in circulating cells. We show that when circulating cell proteomes are mined, taking variable contamination of frequently occurring blood matrix contaminants into account is essential.

Experimental procedures

Subjects

Ten male patients with clinical symptoms of coronary artery disease enrolled in the Circulating Cells study (see acknowledgement) were included for this study. All patients and controls provided written informed consent. Five patients with 1-vessel coronary artery disease and (sub)total stenosis of one coronary segment were selected based on the angiographic description. Five age- and sex-matched patients without significant coronary atherosclerosis served as control. Table 1 summarizes the baseline characteristics of the study subjects, a more detailed description of the clinical parameters can be found in supplementary table 1.

Table 1: Characteristics of Atherosclerosis Patients and Controls.

	(Sub)total Occlusion	Non-significant atherosclerosis	p-value
Age	58 ± 13.5	60 ± 12.1	0.834
BMI	27.85 ± 2.47	27.44 ± 5.05	0.465
Diabetes	0/5	1/5	1.000
Active smoking	1/5	1/5	1.000
Statin use	5/5	4/5	1.000
Platelet inhibitor use	4/5	4/5	1.000

Age and BMI are indicated with the mean and standard error, a Mann-Whitney U test was used for the p-values.

Isolation and Storage of Platelets

Isolation of the platelets was performed in line with other studies involving platelet proteomics [27, 28], with the exception that EDTA was used as anti-coagulant. EDTA blood samples were centrifuged for 15 minutes at 156 x g without brake at room temperature to obtain platelet rich plasma (PRP). The upper fractions of the resulting supernatant (PRP) were collected and transferred to new tubes in order to remove any contaminating leukocytes. The platelets were pelleted by centrifugation for 15 minutes at 330 g at room temperature without brake. The supernatant containing the plasma fraction was aspirated and stored whereas the platelet pellet was resuspended in 1 ml PBS. Following washing with PBS, the platelets were pelleted by centrifugation for 10 minutes at 330 g without brake at room temperature. After discarding the supernatant, the platelet pellet were resuspended in a minimal volume PBS and stored immediately at -80°C until analysis.

Sample preparation

The stored platelet samples were reconstituted in a buffer containing 100 mM Tris, 10 mM DTT, 2% SDS at pH 8.0 with Complete Mini protease inhibitor (Roche Diagnostics, Mannheim, Germany) and PhosSTOP phosphatase inhibitor cocktail (Roche Diagnostics). The cell suspension was subsequently subjected to ultra-sonication (3 times, 30 seconds of pulses at 100% amplitude, 80% interval were performed with 30 second intervals followed by 60 seconds of pulses at 100% amplitude, 100% interval). The remaining cell debris was removed by centrifugation at 1,000 g for 10 minutes at 4 °C. Two sample pools were generated, consisting of 25 μg protein per patient adding up to 125 μg protein per pool. Proteins were then reduced, alkylated and digested using the FASP approach, in which the buffer is exchanged to 8 M urea pH 8.0 in order to remove the SDS present in the sample, as described previously [29]. Digestion was performed for 4 hours with Lys-C (Wako, Richmond, VA, USA) after which the mixture was diluted 4-fold to 2M urea and digested with trypsin (Promega, Madison, WI, USA) at 37 °C overnight. Finally the sample was acidified with formic acid to a final concentration of 5%. Tryptic peptides were desalted using Sep-Pak C18 cartridges (Waters Corporation, Milford, MA). The peptides were subsequently labeled on-column with stable isotope dimethyl labeling as described previously[30, 31], the patients were labeled 'intermediate', whereas controls received the 'light' label. Labeling efficiency was checked by LC-MS/MS before mixing the pools in a 1:1 ratio.

Strong Cation Exchange (SCX) Chromatography

The samples were dried *in vacuo*, and re-suspended in 10% formic acid. SCX was performed on an Agilent 1100 HPLC system (Agilent Technologies, Waldbronn, Germany) consisting of a nanopump, a micro well plate autosampler, a multiple wavelength detector and a fraction collector. The columns used were a Opti Lynx C18, 40 μm, 100 Å, 2.1 mm x 15 mm (Optimize Technologies, Oregon, OR) for trapping and a Zorbax BioSCX-series II (Agilent Technologies, Waldbronn, Germany) 0.8 mm × 50 mm analytical column with 3.5 μm particles. A total of 250 μg of protein digest was loaded on the trapping column using a buffer containing 0.05% formic acid, pH 2.9 onto the trap column and subsequently eluted using a buffer containing 0.05% formic acid, 80% acetonitrile, pH 2.9 onto the analytical column. For SCX separation, a buffer containing 0.05% formic acid, 20% acetonitrile, pH 2.9 (buffer A) and a buffer containing 0.05% formic acid, 20% acetonitrile, pH 2.9 and 500 mM NaCl (buffer B) were used. After 7 minutes of equilibration with buffer A, the separation was performed by a nonlinear 65 min elution gradient: from 0-17 min, 0% buffer B; from 17-22 min, from 2-3% buffer B; from 22-24 min, 3-5% buffer B; from 24-32 min, 5-8% buffer B; from 32-40 min,8-20% buffer B; from 40-48 min, 20- 40% buffer B; from 48- 53 min, 40-90% buffer B. The column was subsequently washed for 3 min with 90% buffer B and finally equilibrated with 100% buffer A again for 9 mins. A total number of 50 SCX fractions were collected and dried in a vacuum centrifuge. Fractions 5-30 were reconstituted in 10% formic acid for further analysis.

Liquid chromatography and tandem mass spectrometry

LC–MS/MS was performed with a nano-LC coupled to an LTQ-Orbitrap Velos (Thermo Scientific, Bremen, Germany). The nano-LC consists of an Agilent 1200 series LC system equipped with a 20 mm ReproSil-Pur C18-AQ (Dr. Maisch GmbH, Ammerbuch, Germany) trapping column (packed in-house, i.d., 100 μm; resin, 5 μm) and a 400 mm ReproSil-Pur C18-AQ (Dr. Maisch GmbH, Ammerbuch, Germany) analytical column (packed in-house, i.d., 50 μm; resin, 3 μm) arranged in a vented-column configuration. The flow was passively split to 100 nl/min. A 3 hour elution profile consisting of: 0-10 min isocratic solvent A (0.1 M acetic acid) at 5 μL/min for sample trapping, followed by a gradient of 10.1-117 min, 10-25% solvent B (0.1 M acetic acid in 80% acetonitrile); 117.1-152 min, 25-50% solvent B; 152.1-154 min, 100% solvent B, 154.1-169, 0% solvent B. Nanospray was achieved using a distally coated fused silica emitter (made in-house, o.d. 375 μm; i.d. 20 μm) biased to 1.7 kV. The mass spectrometer was operated in the data-dependent mode to automatically switch between MS and MS/MS. The high resolution survey full scan was acquired in the orbitrap from m/z 350 to m/z 1500 with a resolution of 30.000 (FHMW) whereas the MS^2 scan is acquired at a resolution of 7500. The ten most intense precursors were isolated with an isolation width of 1.5 m/z and fragmented using a data-dependent decision tree method utilizing HCD, ETD-IT (ion trap read-out) and ETD-FT (orbitrap read-out) [12]. In brief, doubly charged peptides were always subjected to HCD fragmentation as well as triply charged peptides with an m/z >750 Th and are analyzed by the orbitrap. Triply charged peptides were fragmented by ETD when m/z < 750 Th. Analysis of ETD fragment ions was performed in the ion trap (ETD-IT). Quadruply charged ions were all fragmented with ETD and the fragments were analyzed either by ETD-IT (m/z >1000 Th) or the orbitrap (ETD-FT, m/z <1000 Th). Higher charged species were always analyzed by ETD-IT. The normalized collision energy for HCD was set to 35%. ETD reaction time was set to 50 ms for doubly charged precursors. Supplemental activation was enabled.

Data analysis: Identification and Quantitation

Peak lists were generated from the raw data files using the Proteome Discoverer software package version 1.3.339 (Thermo Scientific, Bremen, Germany). Peptide identification was performed by searching the peak list against a concatenated target-decoy database containing the human sequences in the Uniprot database (release 2010_12, 41032 sequences) supplemented with a common contaminants database using the Mascot search engine version 2.3 (Matrix Science, London, United Kingdom) via the Proteome Discoverer interface. The search parameters included the use of trypsin as proteolytic enzyme allowing up to a maximum of 2 missed cleavages. Carbamidomethylation of cysteines was set as a fixed modification whereas oxidation of methionines, the dimethyl "light" and "intermediate" labels on N-termini and lysine residues were set as variable modifications. Precursor mass tolerance was initially set at 50 ppm, while fragment mass tolerance was set at 0.6 Da for ETD-IT fragmentation and 0.05 Da for HCD-FT and ETD-FT fragmentation. Subsequently, the peptide identifications were filtered in Proteome Discoverer for true mass accuracy <4 ppm

and a score >20 until an FDR <1% at peptide level was achieved. The obtained data is publicly available in the repository Tranche (https://proteomecommons.org/) using the following hash code: aympzBVJaubCo1Y2m+wr6YDcm4FK5FLicJbsC9cS5bgSWcUmGjV0vUC8MjgFe/mSiLBE2wr4X7qHwJAdQ418qaNxSywAAAAAAAAR9A==
Peptide and protein quantification were performed using Proteome Discoverer (v.1.3.339) as described previously [33]. Protein ratios were normalized based on the median of protein mean ratios.

Spectral count ranking of platelet data and published proteomes.
To build an abundance map of the platelet proteome the amount of peptide spectral matches were normalized to the molecular weight of the protein (Fabb = ΣPSMs / Mw). Subsequently, the abundance in ppm is calculated by dividing a proteins Fabb by the sum of all Fabb's in the analysis multiplied by 106 as described earlier [34, 35]. The same was done with the 233 proteins in the erythrocyte dataset of Ringrose et al. and the 601 proteins in the leukocyte dataset of Raijmakers et al.[36, 37] both acquired and analyzed under similar conditions. The 50 most abundant plasma proteins were obtained from Tu et al.'s table of the analysis of crude plasma[8]. Abundances therein are largely based on several recent plasma proteome studies [38-40].

Selection of biomarker candidates for follow–up
The initial selection of biomarker candidates was performed by using a stringent 99.99% confidence interval. Therefore, protein ratios were binned according to their corresponding quantification count, i.e. their amount of quantified peptide pairs. Binning was performed in such a way that each bin contains at least 200 proteins and proteins with identical quantification counts can only be designated to one bin. This resulted in the following bins: 1, 2, 3, 4-5, 6-8, 9-13, 14-19, 20-36 and >36 quantification counts. The 99.99% confidence interval (3.89σ) is subsequently determined for each separate bin and a protein therein is considered significant if outside this interval. Next we removed common contaminants (e.g. histones, which cannot be found in anuclear platelets. Next we used a selection of additional criteria to deduce a proper set of potential biomarkers: (i) proteins which are quantified by at least 4 peptides were included and (ii) proteins found in the top 50 of either the erythrocyte, plasma or leukocyte proteome were excluded.

Analysis of Individual Patient Samples
The samples of individual patients were prepared as described above. Per patient, 100 μg protein lysate was digested and subsequently desalted using Sep-Pak C18 cartridges prior to a single run LC-MS/MS analysis of 90 minutes on an Agilent 1200 series HPLC-Chip LC system connected to an Agilent 6520 Q-TOF mass spectrometer (Agilent, Santa Clara, CA). In brief, an Agilent ProtID-Chip-43 (II) (G4240-62005), which has a 4 mm, 40 nl trapping column (Zorbax 300-C18, μm) and a 43 mm, i.d. 75 μm analytical column (Zorbax 300-C18, 5 μm) was used [41]. Trapping of peptides was performed at 4 μL/min using 2% FA in water, and analysis

was performed by switching the trap in-line with the analytical column and nanoflow pump, followed by a nonlinear gradient with solvent C (0.1% formic acid): 0-5 min, 0-10% solvent D (0.1% formic acid in 80% acetonitril) ; 5-65 min, 10-32% solvent D; 65-70 min, 32-40% solvent D; 70-73 min, 40-100% solvent D. The column was subsequently washed for 1.5 min with 10% solvent D and finally equilibrated with 100% solvent C for 15 mins.

The Q-TOF mass spectrometer was operated with the ADC at 2 GHz, "Extended resolution mode" in a data-dependent manner, automatically switching between MS and MS/MS, with a quad AMU setting of 400 m/z. Acquisition times were 500 ms for MS spectra (m/z 350-2400) and 750 ms for MS/MS spectra (m/z 59-3000). The two most intense ions (minimum intensity 1500) were selected for collision induced fragmentation with argon as a collision gas using a collision energy corresponding to 3 V per 100 m/z units. A 30 s dynamic exclusion window was used. The obtained data is publicly available in the repository Tranche (https://proteomecommons.org/) using the previously mentioned hash code.

Label–free Data Analysis

Label-free data analysis was performed by PVIEW [42]. The software takes centroided mzXML files as input and exports CSV files for further analysis of peptide and protein ratios. The original raw files were converted to the mzXML data format [43] using trapper v. 4.3.1 (Seattle Proteome Center/Institute for Systems Biology). MS/MS spectra were searched against a Uniprot human database supplemented with common contaminants (release 2010_12, 20516 sequences) which was automatically extended by the software with an equal number of reversed protein entries for target-decoy analysis, at a desired FDR<1%. The search parameters were as follows: trypsin as proteolytic enzyme, allowing up to a maximum of 2 missed cleavages. Carbamidomethylation of cysteines was set as a fixed modification whereas oxidation of methionines was set as variable modification. Peptide abundances were automatically normalized by a robust central tendency approach [44] to account for slightly differing protein amounts loaded. Features were only quantified when present in at least two replicates of two individuals in either patients or controls.

Results and discussion

Here we set out to develop a biomarker discovery platform for circulating cells. In a discovery and verification approach we focus on the differential platelet proteome of severe atherosclerosis patients and matched controls. We particularly focus on challenges associated with clinical sample collection associated with platelet isolation.

Evaluating contamination of the Platelet Proteome

In the discovery phase (Figure 1), we first performed an in-depth identification and quantification experiment using 2-dimensional chromatography (SCX and RP), stable

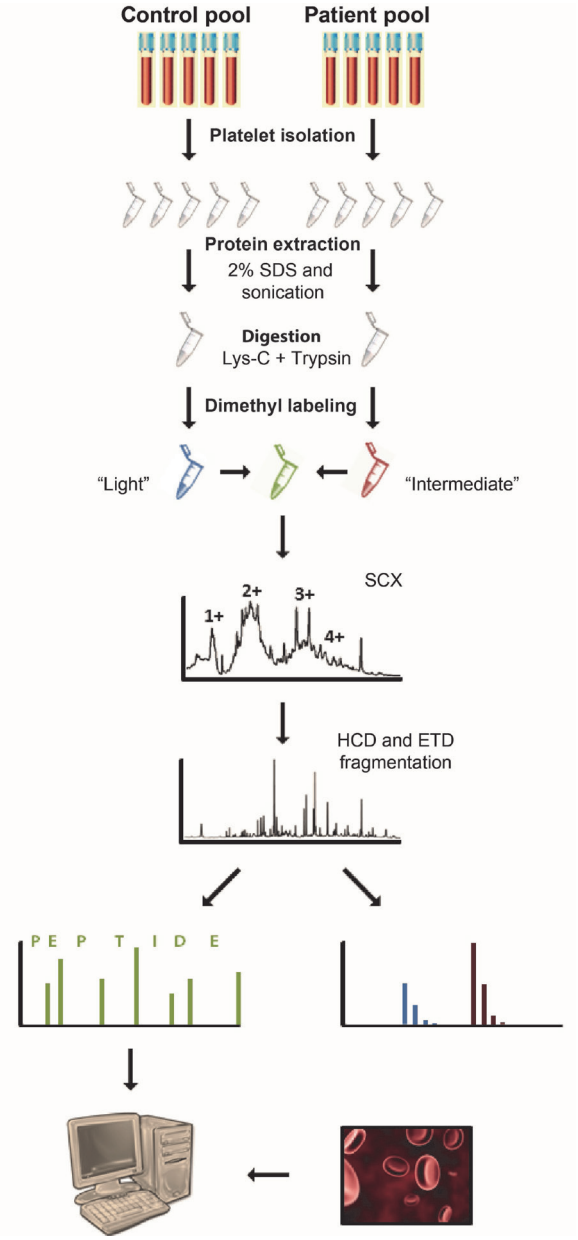

Figure 1: Experimental outline of the discovery phase.

Platelets of patients and controls were pooled (5 each) based on total protein content. After lysis and digestion with trypsin, samples were differentially labeled with stable isotopes using dimethylation and subsequently mixed. 2-dimensional chromatography with strong cation exchange (SCX) and reversed phase was used to reduce sample complexity prior to MS/MS analysis using a decision tree method with HCD and ETD for fragmentation. The obtained quantitative platelet proteome was interfaced with potential contaminating proteomes (erythrocytes, plasma and leukocytes) to ensure the identification of specific platelet biomarker candidates.

isotope dimethyl labeling[30, 31] and a decision tree peptide fragmentation schedule with ETD and HCD[32]. Due to the elaborateness of this platform, discovery cannot be performed easily on individual patients. Therefore, we chose to pool two stringently selected small cohorts (Table 1) consisting of patients with (sub)total coronary obstruction and patients without significant coronary atherosclerosis as controls. To minimize variability in this small discovery cohort, patients and controls were matched on several criteria (Table 1, supplemental table 1). This resulted in the identification of 2901 proteins (supplemental table 2). HCD contributed 13804, and ETD identified 7187 unique peptides resulting in 2627 (1015 only by HCD) and 1886 (274 only by ETD) unique proteins respectively. These results are as expected[32] and show the usefulness of such technology for biomarker discovery: increased identifications and quantifications without additional analysis time.

In biomarker discovery, sample uniformity is of the utmost importance in order to deduce changes in protein expression related to disease, rather than sample variability. When collecting circulating cells, contamination with plasma proteins and high abundant proteins from small amounts of other circulating cell types (e.g. erythrocytes and leukocytes) is extremely hard to avoid and should therefore be taken into account in the discovery, particularly when applying such in-depth approaches as presented here. In this way follow-up investments into the verification and validation phase are more warranted.

Platelet isolation was performed with the utmost care and according to well established protocols (see Materials and Methods). Only samples with a complete white cell pallet were selected. Pink or red cell pallets were omitted to avoid large erythrocyte contaminations. Despite these precautions, we could readily observe plasma, erythrocyte and leukocyte contaminants in our platelet proteome. Figure 2A shows that the erythrocyte proteome (233 proteins from Ringrose et al. [37]) and the 50 most abundant plasma proteins (based on a comprehensive concentration table from [8]) are identified in our platelet proteome. This is also the case for the leukocyte proteome which itself also suffers from contamination (601 proteins from Raijmakers et al.[36]). Based on the 105,000 identified MS/MS spectra, we built a quantitative map of the platelet proteome [34, 35] (in ppm, Figure 2B, supplemental table 2, see materials and methods for details). The same was done with the already available erythrocyte[37] and leukocyte[36] proteome data (supplemental tables 3 and 4). This analysis revealed that the 10 most abundant proteins from plasma, leukocytes and erythrocytes are well represented in the platelet proteome (~20,000 – 500 ppm, Figure 2B). When evaluated within the context of the entire obtained platelet proteome, the 10 most abundant erythrocyte, leukocyte and plasma proteins (supplemental table 3-5) comprise approximately 11% of the entire platelet protein population (Figure 2C). These data are based on qualitative data assessment alone and it should be taken into account that several of the abundant contaminants could also be expressed, or at least associated with platelets. An example is serum albumin, which is actively released from platelets.

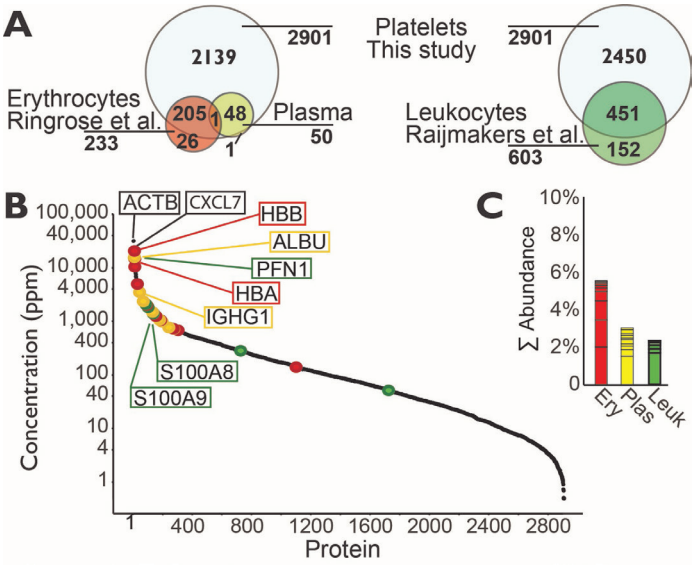

Figure 2: Contamination of the Platelet Proteome.

(**A**) The left Venn diagram displays the overlap in protein identification between the platelet proteome in this study and the erythrocyte proteome of Ringrose et al. 37 and the top 50 most abundant plasma proteins based on actual concentrations 8; the second Venn diagram displays the overlap between the platelet proteome in this study and the leukocyte proteome of Raijmakers et al. 41. (**B**) Protein abundance map of the platelet proteome in ppm (black). Highlighted are the top-10 of most abundant erythrocyte proteins (red dots), plasma proteins (yellow dots) and leukocyte proteins (green dots). Marked are the three most abundant proteins from each matrix: ACTB = cytoplasmic actin, CXCL7= platelet basic protein, HBA=hemoglobin-α, HBB = hemoglobin-ß, HBD = hemoglobin-δ, ALBU = serum albumin, APOA1 = Apolipoprotein A-I, IGHG1 = Ig gamma-1 chain C region, PFN1 = Profilin 1, S100A8 = Protein S100-A8 and S100A9 = Protein S100-A9 C. At total platelet proteome level the abundance of the top-10 erythrocyte proteins (supplemental table 3), top-10 leukocyte proteins (supplemental table 4) and top-10 plasma proteins (supplemental table 5) make up 6%, 2% and 3% of the total protein content of the entire "purified" platelet proteome respectively.

Nonetheless, we argue that contamination of clinically isolated platelets could cause sample inequality, even when present at low levels. This may lead to incorrectly assigned biomarkers or offsets in later quantitative analyses, particularly when these are related to total protein content. Therefore, quantitative data of the pooled samples (discovery phase), as well as at individual patient level (verification phase), are necessary to deduce such information and is evaluated below.

Discovery Phase

Of the 2901 identified platelet proteins we could differentially quantify 2440 (84%, supplemental table 2). Their ratios are displayed in figure 3 as a function of the amount of peptide pairs used for protein quantitation. From these data it is clear that the majority of the platelet proteome is consistent between patients and controls (>80% of proteins have a ratio

between 0.8 and 1.2). Although several candidate biomarkers are revealed at first glance, especially in the down-regulated part, we first superimposed the top-50 of the contaminating proteomes on these quantitative data (Figure 3). The difference in concentration as presented by Farrah et al. and Tu et al. between serum albumin (4.0×10^7 ng/mL) and proteins around rank 50 such as serum amyloid P-component and Complement C1q subcomponent subunit B (~4.0×10^4 ng/mL, ~5.5×10^4 ng/mL) is ~500-1000 fold[7, 8]. Considering that the top protein albumin constitutes to ~15000 ppm, then the effect of rank 50 proteins should be barely contributing. The same considerations were taken into account for the hemoglobins (~15000-20000 ppm in platelets) and the other erythrocyte proteins, as well as the S100 proteins from leukocytes (~1400 ppm in platelets). When looking at the erythrocyte top-50/platelet overlap, the effect of contamination on biomarker discovery is apparent since the hemoglobins (HBA, HBB), were all quantified with a patient/control ratio of ~0.4-0.5, as were 10 others (supplemental table 3). In contrast, several other abundant erythrocyte proteins appear with a ratio of around 1.0, indicative of their presence in both erythrocytes and platelets for example: PPIA, Peptidyl-prolyl cis-trans isomerase A, ranked the 8th most abundant erythrocyte protein and CAH2 (carbonic anhydrase 2, ranked 7th, supplemental table 2). The leukocyte top-50/platelet overlap displays a similar trend where a set of proteins such as protein S100A9 (rank 1 in leukocytes) and protein S100A8 (rank 2) also cluster at a patient/control ratio of ~0.4, whereas another set of leukocyte proteins

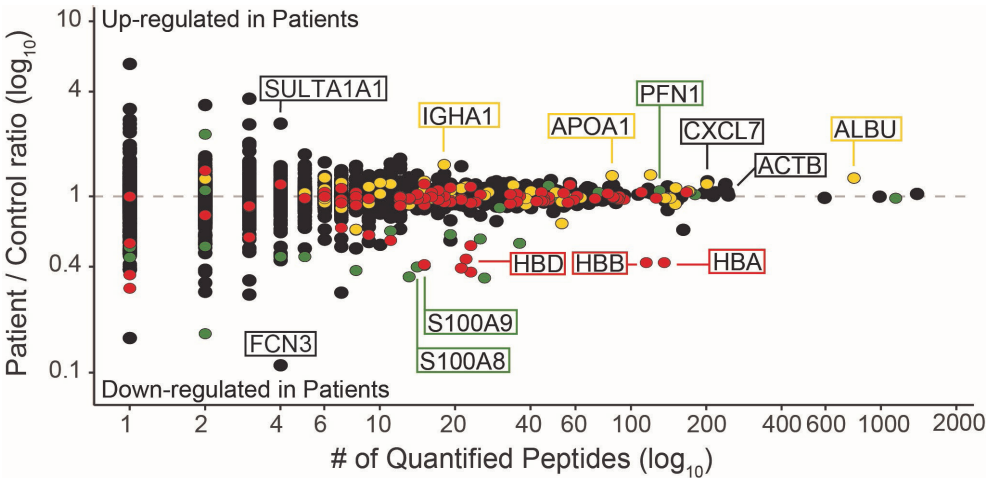

Figure 3: Differential proteome analysis of platelets from patients and controls.

The patient/control ratio (log10 scale) of the 2440 quantified proteins plotted against the number of quantified peptides for each protein shows high consistency of the expressed platelet proteomes of patients and controls down to 4 quantified peptides. The 50 most abundant erythrocyte (red, [37]), plasma (yellow, [8]) and leukocyte (green, [41]) proteins are highlighted. Platelet proteins in black, marked proteins as in figure 2 and FCN3 = Ficolin-3 and SULTA1A1 = Sulfotransferase 1A1.

such as plastin-2 and vimentin cluster at a patient/control ratio of 0.6. These two clusters may originate from two distinct subclasses of leukocytes. Other abundant leukocyte proteins appear with a ratio 1.0, indicative of their co-occurrence in platelets. The plasma contamination does not reveal itself as prominently and seems almost equal between patient and control samples. The ratio of serum albumin (ALBU), the most abundant plasma protein is 1.2 (Figure 3), as is the ratio of Apolipoprotein-AI (APOA1, plasma rank 5) and IGHA1 (Ig alpha-1 chain C region, plasma rank 4). From interfacing the quantitative data with the potential contaminating proteome we propose that excluding the top-50 serum, erythrocyte and leukocyte proteins is sufficient to significantly reduce false positive biomarkers in the verification candidate list.

Verification in individual patients

Although useful in determining which proteins not to pursue in a follow-up, the pooled samples above show only an average effect. Therefore, to further investigate the effect of the contaminants on individual patient samples, the platelet lysates of the individual patients and controls were analyzed with a label-free quantitative proteomics setup. For this we utilized a microfluidics chip-based reversed phase HPLC setup directly coupled to a QTOF mass spectrometer. In this way we take advantage of the very stable retention times of the chip based microfluidics device as presented previously [41]. All samples were run in triplicate and the PVIEW software package was used for alignment, identification and quantitation [42]. With this less in-depth setup we could quantify only the 151 most abundant platelet proteins in each patient (supplemental table 6).

First we evaluated how the observed contaminants were distributed over the different patients and controls (Figure 4A). This shows that the contamination pattern for plasma is high in control 1 and patient 3 (top panel), whereas it is almost absent in control 4. This pattern is entirely different for the erythrocyte contamination as monitored through the amount of hemoglobin isoforms (bottom panel), which are actually most pronounced in control 4. The observed patient/control ratio in the labeled-pool approach could be matched in the label free data:, e.g. HBA had ratios of 0.42 (labeling) and 0.43 (label-free) and HBD 0.44 and 0.52 (supplemental tables 2 and 6).

To verify the extent of contamination we compared the total ion intensities of contaminants and platelet proteins (Figure 4B). This revealed that in the best cases (control 5) total contaminant intensity was below 3-4% for either plasma, erythrocytes or leukocytes, whereas in the worst cases one of these could be as high as 15% (control 1, patient 3). We also evaluated the variability of platelet specific proteins over the entire intensity range of the label-free experiment ($7e^7$ down to $2e^5$, Figure 4C). These are all within a satisfactory 1.5 fold range, meaning that contamination at the levels found in our samples is not interfering with normalization and alignment of the label-free software.

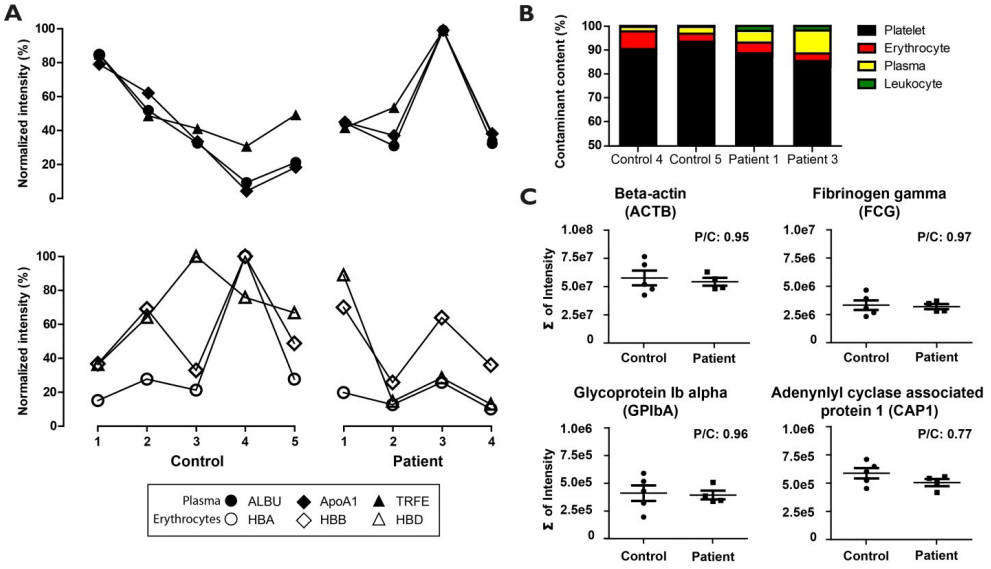

Figure 4: Label–free differential quantification in individual patients and controls.

(**A**) Normalized intensities (QTOF data) of three plasma proteins (top, black data points) and of three erythrocyte proteins (bottom panel, white data points). The top 3 peptides were used for quantification. Patient 5 was omitted due to irreproducible LC-MS behavior during analysis. Depicted are averages of triplicate analyses. The relative standard deviations (not depicted), excluding 5% outliers were within 70%, but more regularly within 30% (supplemental table 6). Protein naming as in figure 2 and 3, TRFE = Serotransferrin. (**B**) %-fraction of 'contaminant intensity' in selected samples show large variability in contamination. Displayed are the summed peptide intensities of platelet proteins (black), top-10 erythrocyte proteins (red), top-10 plasma proteins (yellow) and top-50 leukocyte proteins (green) as % of the total intensity in the run. (**C**) After normalization (for details see materials and methods), identified platelet proteins can be reliably quantified over 2.5 orders of magnitude in signal (7.5e7 (Beta-Actin) down to 2.0 e5 (Adenylyl cyclase associated protein 1).

Candidate Selection

Using all considerations outlined above, only a small subset of proteins remains as candidate biomarker for atherosclerosis. Using stringent criteria on the quantitative data: (i) ratio outside the 99.99% confidence interval and (ii) ≥4 peptides quantified leaves 37 proteins (Figure 5). Removal of 'regulated'contaminants reduces this by another 60% to 15 potential candidates. The strength of this approach is highlighted by the apparent regulation of the Calgranin proteins S100A8 and S100A9, earlier identified as biomarkers for coronary artery disease [25, 26]. Although these proteins of monocyte/macrophage origin reveal themselves as biomarkers in platelets, they are in fact differential leukocyte contamination of the platelet proteome. (for review, see [45]).

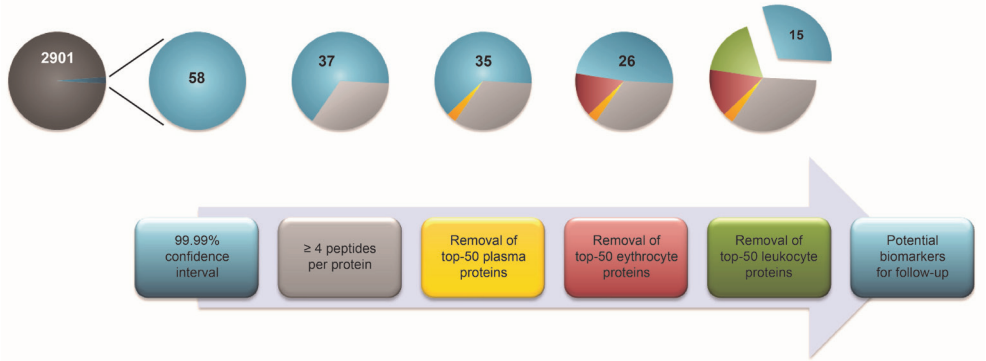

Figure 5: Biomarker candidate selection benefits from correlating obvious contaminating proteomes.

The amount of potential biomarkers identified in our large scale quantitative approach drops significantly when stringent criteria are applied (from left to right). First, the observed fold change was scrutinized in relation to the amount of observed data-points (number of quantified peptides), leaving 58 potential biomarkers. Proteins remaining differential within the very stringent 99.99% confidence interval were only taken forward when quantified with 4 or more peptides (37 candidates left). Subsequently, false positives due to contamination from the high abundant proteomes (top-50) of other circulating cell types were removed (2 for plasma, 9 for erythrocytes and 11 for leukocytes), leaving 15 potential candidates for follow-up (table 2).

Amongst the 15 candidates (Table 2), three proteins were found up-regulated in patients, 12 down-regulated. Most were found with close to 2-fold patient/control ratios, but Sulfotransferase 1A1 (SULFA1A1) was found 2.6-fold upregulated (Table 1, Figure 3). Sulfotransferase 1A1 is known to be present in platelets and a recent study implicated several mutations in this gene connected to outcomes of acute coronary syndrome patients [46]. From the 12 down-regulated proteins observed, two had high ratios, of which Ficolin-3 proved most prominent (9.1-fold). Interestingly, FCN3 gene variations cause loss or reduction in the plasma Ficolin-3 levels and are associated with immunodeficiency [47]. Also, low ficolin-3 levels in serum are associated with the severity and unfavorable outcome of acute ischemic stroke [48], matching very well with our results. In addition, we found Galectin-3-binding protein, LGALS3BP 3.6-fold down-regulated. Since it was not the purpose of this study, further verification and validation of these three putative biomarkers should provide further evidence for their involvement in atherosclerosis. Their relative low abundance in the platelet proteome would warrant the development of immune based assays.

Table 2: Final biomarker candidate list of significantly up– and downregulated proteins.

Accession	Rank	ppm	Description	MW [kDa]	Ratio P/C	# Quant. pep.
Upregulated						
P62942	348	571	Peptidyl-prolyl cis-trans isomerase FKBP1A,FKBP1A	11.9	1.64	12
P50225	533	400	Sulfotransferase 1A1,SULT1A1	34.1	2.6	4
P19823	1027	160	Inter-alpha-trypsin inhibitor heavy chain H2,ITIH2	106.4	1.48	21
Downregulated						
P07355	874	206	Annexin A2,ANXA2	38.6	0.53	10
Q15149	878	205	Plectin,PLEC	531.5	0.64	162
P80723	891	201	Brain acid soluble protein 1,BASP1	22.7	0.51	7
P21266	972	171	Glutathione S-transferase Mu 3,GSTM3	26.5	0.56	6
P26583	1083	142	High mobility group protein B2,HMGB2	24	0.51	4
P02730	1162	123	Band 3 anion transport protein,SLC4A1	101.7	0.56	19
P08133	1179	120	Annexin A6,ANXA6	75.8	0.55	9
O75636	1273	104	Ficolin-3,FCN3	32.9	0.11	4
Q14764	1383	88	Major vault protein,MVP	99.3	0.55	9
Q08380	1590	64	Galectin-3-binding protein,LGALS3BP	65.3	0.28	7
P33241	1615	61	Lymphocyte-specific protein 1,LSP1	37.2	0.45	4
Q16799	1895	36	Reticulon-1,RTN1	83.6	0.5	5

Potential biomarkers are sorted by decreasing protein concentration. Proteins with the most prominent P/C ratios (> 2.5-fold) are highlighted in bold. Rank= abundance rank in the obtained platelet proteome, ppm= relative protein concentration in the platelet proteome, MW= Molecular Weight of the protein, P= patients, C=controls, # Quant pep. = number of quantified peptide pairs.

Conclusion

The main purpose of this study was to build and evaluate a platform for biomarker discovery in isolated circulating cells. We evaluate the feasibility of these promising novel matrices for (atherosclerosis) biomarker discovery. Using platelets as a template, we readily reach a depth of proteome coverage of at least 3- to 4-fold beyond plasma. However, even minor contamination with other blood matrices represents a serious challenge when using isolated circulating cells at this depth level. We suggest the biomarker community to take the utmost care in the isolation of circulating cells and be suspicious of potential differential

contamination by other circulating matrices. It should never be assumed that samples can be 100% pure and homogeneously collected and prepared. Where possible, cell sorting should be used, although in our hands even there contaminations were visible (unpublished).

On a more positive note, we show that when cells are carefully selected and candidate biomarkers are evaluated with the described contamination biases in mind, a small, manageable candidate list can be generated for verification and ultimately validation. If contamination levels in the individual patient samples do not exceed 10-15% of the total proteome, circulating cells can be an interesting alternative for biomarker discovery, especially in inflammatory and cardiovascular diseases.

Acknowledgements
Mirjam Damen is kindly acknowledged for technical assistance. This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), project CIRCULATING CELLS (grant 01C-102), and supported by the Netherlands Heart Foundation. The Netherlands Proteomics Centre embedded in the Netherlands Genomics Initiative is kindly acknowledged for financial support. The Cardiovascular “Focus en Massa programma” at Utrecht University is acknowledged for additional financial support.

References

1. Hanash SM, Pitteri SJ, Faca VM. Mining the plasma proteome for cancer biomarkers. ***Nature***. 2008;452:571-579
2. Xu X, Veenstra TD. Analysis of biofluids for biomarker research. ***Proteomics. Clinical applications***. 2008;2:1403-1412
3. Anderson NL, Anderson NG. The human plasma proteome: History, character, and diagnostic prospects. ***Mol Cell Proteomics***. 2002;1:845-867
4. Omenn GS, States DJ, Adamski M, Blackwell TW, Menon R, Hermjakob H, Apweiler R, Haab BB, Simpson RJ, Eddes JS, Kapp EA, Moritz RL, Chan DW, Rai AJ, Admon A, Aebersold R, Eng J, Hancock WS, Hefta SA, Meyer H, Paik YK, Yoo JS, Ping P, Pounds J, Adkins J, Qian X, Wang R, Wasinger V, Wu CY, Zhao X, Zeng R, Archakov A, Tsugita A, Beer I, Pandey A, Pisano M, Andrews P, Tammen H, Speicher DW, Hanash SM. Overview of the hupo plasma proteome project: Results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. ***Proteomics***. 2005;5:3226-3245
5. Munoz J, Low TY, Kok YJ, Chin A, Frese CK, Ding V, Choo A, Heck AJ. The quantitative proteomes of human-induced pluripotent stem cells and embryonic stem cells. ***Mol Syst Biol***. 2011;7:550
6. Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, Paabo S, Mann M. Deep proteome and transcriptome mapping of a human cancer cell line. ***Mol Syst Biol***. 2011;7:548
7. Farrah T, Deutsch EW, Omenn GS, Campbell DS, Sun Z, Bletz JA, Mallick P, Katz JE, Malmstrom J, Ossola R, Watts JD, Lin B, Zhang H, Moritz RL, Aebersold R. A high-confidence human plasma proteome reference set with estimated concentrations in peptideatlas. ***Mol Cell Proteomics***. 2011;10:M110 006353
8. Tu C, Rudnick PA, Martinez MY, Cheek KL, Stein SE, Slebos RJ, Liebler DC. Depletion of abundant plasma proteins and limitations of plasma proteomics. ***J Proteome Res***. 2010;9:4982-4991
9. Qian WJ, Kaleta DT, Petritis BO, Jiang H, Liu T, Zhang X, Mottaz HM, Varnum SM, Camp DG, 2nd, Huang L, Fang X, Zhang WW, Smith RD. Enhanced detection of low abundance human plasma proteins using a tandem igy12-supermix immunoaffinity separation strategy. ***Mol Cell Proteomics***. 2008;7:1963-1973
10. Gong Y, Li X, Yang B, Ying W, Li D, Zhang Y, Dai S, Cai Y, Wang J, He F, Qian X. Different immunoaffinity fractionation strategies to characterize the human plasma proteome. ***J Proteome Res***. 2006;5:1379-1387
11. Surinova S, Schiess R, Huttenhain R, Cerciello F, Wollscheid B, Aebersold R. On the development of plasma protein biomarkers. ***J Proteome Res***. 2011;10:5-16
12. Hansson GK, Libby P. The immune response in atherosclerosis: A double-edged sword. ***Nat Rev Immunol***. 2006;6:508-519
13. Libby P. Changing concepts of atherogenesis. ***J Intern Med***. 2000;247:349-358
14. Duran MC, Martin-Ventura JL, Mohammed S, Barderas MG, Blanco-Colio LM, Mas S, Moral V, Ortega L, Tunon J, Jensen ON, Vivanco F, Egido J. Atorvastatin modulates the profile of proteins released by human atherosclerotic plaques. ***Eur J Pharmacol***. 2007;562:119-129

15. Martin-Ventura JL, Duran MC, Blanco-Colio LM, Meilhac O, Leclercq A, Michel JB, Jensen ON, Hernandez-Merida S, Tunon J, Vivanco F, Egido J. Identification by a differential proteomic approach of heat shock protein 27 as a potential marker of atherosclerosis. ***Circulation***. 2004;110:2216-2219
16. von Zur Muhlen C, Schiffer E, Zuerbig P, Kellmann M, Brasse M, Meert N, Vanholder RC, Dominiczak AF, Chen YC, Mischak H, Bode C, Peter K. Evaluation of urine proteome pattern analysis for its potential to reflect coronary artery atherosclerosis in symptomatic patients. ***J Proteome Res***. 2009;8:335-345
17. de la Cuesta F, Alvarez-Llamas G, Maroto AS, Donado A, Zubiri I, Posada M, Padial LR, Pinto AG, Barderas MG, Vivanco F. A proteomic focus on the alterations occurring at the human atherosclerotic coronary intima. ***Mol Cell Proteomics***. 2011;10:M110 003517
18. Wang TJ, Gona P, Larson MG, Tofler GH, Levy D, Newton-Cheh C, Jacques PF, Rifai N, Selhub J, Robins SJ, Benjamin EJ, D'Agostino RB, Vasan RS. Multiple biomarkers for the prediction of first major cardiovascular events and death. ***N Engl J Med***. 2006;355:2631-2639
19. Lepedda AJ, Cigliano A, Cherchi GM, Spirito R, Maggioni M, Carta F, Turrini F, Edelstein C, Scanu AM, Formato M. A proteomic approach to differentiate histologically classified stable and unstable plaques from human carotid arteries. ***Atherosclerosis***. 2009;203:112-118
20. You SA, Archacki SR, Angheloiu G, Moravec CS, Rao S, Kinter M, Topol EJ, Wang Q. Proteomic approach to coronary atherosclerosis shows ferritin light chain as a significant marker: Evidence consistent with iron hypothesis in atherosclerosis. ***Physiol Genomics***. 2003;13:25-30
21. de Kleijn DP, Moll FL, Hellings WE, Ozsarlak-Sozer G, de Bruin P, Doevendans PA, Vink A, Catanzariti LM, Schoneveld AH, Algra A, Daemen MJ, Biessen EA, de Jager W, Zhang H, de Vries JP, Falk E, Lim SK, van der Spek PJ, Sze SK, Pasterkamp G. Local atherosclerotic plaques are a source of prognostic biomarkers for adverse cardiovascular events. ***Arterioscler Thromb Vasc Biol***. 2010;30:612-619
22. Povsic TJ, Goldschmidt-Clermont PJ. Endothelial progenitor cells: Markers of vascular reparative capacity. ***Ther Adv Cardiovasc Dis***. 2008;2:199-213
23. Alber HF, Duftner C, Wanitschek M, Dorler J, Schirmer M, Suessenbacher A, Frick M, Dichtl W, Pachinger O, Weidinger F. Neopterin, cd4+cd28- lymphocytes and the extent and severity of coronary artery disease. ***Int J Cardiol***. 2009;135:27-35
24. Breet NJ, van Werkum JW, Bouman HJ, Kelder JC, Ruven HJ, Bal ET, Deneer VH, Harmsze AM, van der Heyden JA, Rensing BJ, Suttorp MJ, Hackeng CM, ten Berg JM. Comparison of platelet function tests in predicting clinical outcome in patients undergoing coronary stent implantation. ***Jama***. 2010;303:754-762
25. Barderas MG, Tunon J, Darde VM, De la Cuesta F, Duran MC, Jimenez-Nacher JJ, Tarin N, Lopez-Bescos L, Egido J, Vivanco F. Circulating human monocytes in the acute coronary syndrome express a characteristic proteomic profile. ***J Proteome Res***. 2007;6:876-886
26. Poduri A, Bahl A, Talwar KK, Khullar M. Proteomic analysis of circulating human monocytes in coronary artery disease. ***Mol Cell Biochem***. 2011
27. O'Neill EE, Brock CJ, von Kriegsheim AF, Pearce AC, Dwek RA, Watson SP, Hebestreit HF. Towards complete analysis of the platelet proteome. ***Proteomics***. 2002;2:288-305

28. Piersma SR, Broxterman HJ, Kapci M, de Haas RR, Hoekman K, Verheul HM, Jimenez CR. Proteomics of the trap-induced platelet releasate. ***J Proteomics***. 2009;72:91-109
29. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. ***Nat Methods***. 2009;6:359-362
30. Boersema PJ, Raijmakers R, Lemeer S, Mohammed S, Heck AJ. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. ***Nat Protoc***. 2009;4:484-494
31. Kovanich D, Cappadona S, Raijmakers R, Mohammed S, Scholten A, Heck AJ. Applications of stable isotope dimethyl labeling in quantitative proteomics. ***Anal Bioanal Chem***. 2012
32. Frese CK, Altelaar AF, Hennrich ML, Nolting D, Zeller M, Griep-Raming J, Heck AJ, Mohammed S. Improved peptide identification by targeted fragmentation using cid, hcd and etd on an ltq-orbitrap velos. ***J Proteome Res***. 2011;10:2377-2388
33. Scholten A, Mohammed S, Low TY, Zanivan S, van Veen TA, Delanghe B, Heck AJ. In-depth quantitative cardiac proteomics combining electron transfer dissociation and the metalloendopeptidase lys-n with the silac mouse. ***Mol Cell Proteomics***. 2011;10:O111 008474
34. Aye TT, Scholten A, Taouatas N, Varro A, Van Veen TA, Vos MA, Heck AJ. Proteome-wide protein concentrations in the human heart. ***Mol Biosyst***. 2010;6:1917-1927
35. Peng M, Taouatas N, Cappadona S, van Breukelen B, Mohammed S, Scholten A, Heck AJ. Protease bias in absolute protein quantitation. ***Nat Methods***. 2012;9:524-525
36. Raijmakers R, Heck AJ, Mohammed S. Assessing biological variation and protein processing in primary human leukocytes by automated multiplex stable isotope labeling coupled to 2 dimensional peptide separation. ***Mol Biosyst***. 2009;5:992-1003
37. Ringrose JH, van Solinge WW, Mohammed S, O'Flaherty MC, van Wijk R, Heck AJ, Slijper M. Highly efficient depletion strategy for the two most abundant erythrocyte soluble proteins improves proteome coverage dramatically. ***J Proteome Res***. 2008;7:3060-3063
38. Haab BB, Geierstanger BH, Michailidis G, Vitzthum F, Forrester S, Okon R, Saviranta P, Brinker A, Sorette M, Perlee L, Suresh S, Drwal G, Adkins JN, Omenn GS. Immunoassay and antibody microarray analysis of the hupo plasma proteome project reference specimens: Systematic variation between sample types and calibration of mass spectrometry data. ***Proteomics***. 2005;5:3278-3291
39. Qian WJ, Jacobs JM, Camp DG, 2nd, Monroe ME, Moore RJ, Gritsenko MA, Calvano SE, Lowry SF, Xiao W, Moldawer LL, Davis RW, Tompkins RG, Smith RD. Comparative proteome analyses of human plasma following in vivo lipopolysaccharide administration using multidimensional separations coupled with tandem mass spectrometry. ***Proteomics***. 2005;5:572-584
40. Wu SL, Choudhary G, Ramstrom M, Bergquist J, Hancock WS. Evaluation of shotgun sequencing for proteomic analysis of human plasma using hplc coupled with either ion trap or fourier transform mass spectrometry. ***J Proteome Res***. 2003;2:383-393
41. Raijmakers R, Kraiczek K, de Jong AP, Mohammed S, Heck AJ. Exploring the human leukocyte phosphoproteome using a microfluidic reversed-phase-tio2-reversed-phase high-performance liquid chromatography phosphochip coupled to a quadrupole time-of-flight mass spectrometer. ***Anal Chem***. 2010;82:824-832

42. Khan Z, Bloom JS, Garcia BA, Singh M, Kruglyak L. Protein quantification across hundreds of experimental conditions. ***Proc Natl Acad Sci U S A***. 2009;106:15544-15548
43. Pedrioli PG, Eng JK, Hubley R, Vogelzang M, Deutsch EW, Raught B, Pratt B, Nilsson E, Angeletti RH, Apweiler R, Cheung K, Costello CE, Hermjakob H, Huang S, Julian RK, Kapp E, McComb ME, Oliver SG, Omenn G, Paton NW, Simpson R, Smith R, Taylor CF, Zhu W, Aebersold R. A common open representation of mass spectrometry data and its application to proteomics research. ***Nat Biotechnol***. 2004;22:1459-1466
44. Callister SJ, Barry RC, Adkins JN, Johnson ET, Qian WJ, Webb-Robertson BJ, Smith RD, Lipton MS. Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. ***J Proteome Res***. 2006;5:277-286
45. Perera C, McNeil HP, Geczy CL. S100 calgranulins in inflammatory arthritis. ***Immunol Cell Biol***. 2010;88:41-49
46. O'Halloran AM, Patterson CC, Horan P, Maree A, Curtin R, Stanton A, McKeown PP, Shields DC. Genetic polymorphisms in platelet-related proteins and coronary artery disease: Investigation of candidate genes, including n-acetylgalactosaminyltransferase 4 (galnt4) and sulphotransferase 1a1/2 (sult1a1/2). ***J Thromb Thrombolysis***. 2009;27:175-184
47. Munthe-Fog L, Hummelshoj T, Honore C, Madsen HO, Permin H, Garred P. Immunodeficiency associated with fcn3 mutation and ficolin-3 deficiency. ***N Engl J Med***. 2009;360:2637-2644
48. Fust G, Munthe-Fog L, Illes Z, Szeplaki G, Molnar T, Pusch G, Hirschberg K, Szegedi R, Szeplaki Z, Prohaszka Z, Skjoedt MO, Garred P. Low ficolin-3 levels in early follow-up serum samples are associated with the severity and unfavorable outcome of acute ischemic stroke. ***J Neuroinflammation***. 2011;8:185

3.

High precision platelet releasate definition by quantitative reversed protein profiling

Patrick Wijten[1-2], Thijs van Holten[3], Liy Liy Woo[1-2], Onno B. Bleijerveld[1-2,4], Mark Roest[3], Albert J.R. Heck[1-2*] and Arjen Scholten[1-2*]

Adapted from:
Arteriosclerosis, Thrombosis and Vascular Biology
2013, 33, 7, 1635-1638

[1]Biomolecular Mass Spectrometry and Proteomics, Utrecht Institute for Pharmaceutical Sciences and Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands
[2]Netherlands Proteomics Centre, Utrecht, The Netherlands
[3]Department of Clinical Chemistry and Haematology, University Medical Centre Utrecht, Utrecht, The Netherlands
[4]Department of Experimental Cardiology, University Medical Centre Utrecht, Utrecht, The Netherlands.

Abstract

Platelet activation and subsequent protein release plays an important role in healthy hemostasis and inflammatory responses, yet the identity and quantity of proteins in the platelet releasate is still debated. Here, we present a reversed releasate proteomics approach to determine unambiguously and quantitatively proteins released from activated platelets. Isolated platelets were mock and fully stimulated after which the released proteins in the supernatant were removed. Using high-end proteomics technology (2D-chromatography, stable isotope labeling, ETD and HCD-fragmentation) allowed us to quantitatively discriminate the released proteins from uncontrolled lysis products. Monitoring the copy numbers of around 4500 platelet proteins, we observed that following stimulation via thrombin and collagen only 124 (<3%) proteins were significantly released ($p<0.05$). The released proteins span a concentration range of at least 5 orders, as confirmed by ELISA. The released proteins were highly enriched in secretion tags and contained all known factors at high concentrations (>100 ng/mL, e.g. Thrombospondin, von Willebrand factor and Platelet factor 4). Interestingly, in the lower concentration range of the releasate many novel factors were identified. Our reversed releasate dataset forms the first unambiguous, in depth repository for molecular factors released by platelets.

Introduction

It is widely accepted that activated platelets release proteins into the circulation via various mechanisms, of which the α-granule route and shedding are the most established. The protein constituents of this releasate are associated with inflammation, coagulation, (tumor) angiogenesis, cell growth and adhesion[1-5]. Although some of these factors are routinely assayed in the clinic, a comprehensive and unambiguous quantitative map of the platelet releasate is still at large. Previous releasate proteomics approaches relied on the isolation of α-granules from isolated platelets[6, 7] or the investigation of the isolated supernatant after *in vitro* platelet stimulation[8, 9]. However, these direct methods suffer from a high false positive rate due to contamination by uncontrolled platelet lysis. Here we take a reverse approach, monitoring quantitatively the concentration changes of all proteins in the platelets following platelet stimulation, whereby we assume that the released protein content should be detectable in the reduction of its level from the whole platelet proteome. For accuracy, our strategy uses stable isotope labeled quantitation (Figure 1A) to discriminate the released proteins from uncontrolled lysis products. In other words, most platelet proteins do not change except the ones that are significantly released upon activation.

Methods

Platelet Isolation and Stimulation

Venous blood was collected with an open system, anticoagulated with 3.2% tri-sodium citrate (Merck), from healthy volunteers after obtaining informed consent. Platelet rich plasma (PRP) was prepared by centrifugation of whole blood at 160g for 15 minutes room temperature, no brake. Washed platelets were prepared by adding PRP with ACD (8.5mM tri-sodium citrate, 7.1mM citric acid, 5.5mM D-glucose; final concentrations) and centrifugation for 15 minutes at 340g at room temperature, no brake. The platelet pellet was resuspended in Tris-buffer (145mM NaCl, 5mM KCl, 260nM NaH2PO4, 1mM MgSO4, 100mM Tris, 5.5mM D-glucose, pH 6.5) with 10ng/mL prostacyclin. Platelets were centrifuged for 15 minutes at 340g at room temperature and resuspended up to 200 x 109/L in Tris-buffer (pH 7.4). Platelets were not used until 30 minutes after isolation. To ensure platelets are in resting state the mean platelet volume (MPV) was measured during the isolation procedure. MPV increase for each of the donors was <1.2 fL indicating resting state was maintained. The washed platelets were split into equal halves for each replicate and stimulated using 5 µg/mL collagen and 1U/mL thrombin for 5 minutes, or mock treated and left at rest (Figure 1A). After stimulation, platelet suspensions were centrifuged at 4000g for 2 minutes, the supernatant was aspirated and the platelet pellet was snap-frozen in liquid nitrogen.

Sample preparation

The stored platelet pellets were thawed and centrifuged at 14,000 g for 10 min at 4 °C. The pellet was reconstituted in a buffer containing 100 mM Tris, 10 mM DTT, 2% SDS at pH 8.0 with Complete Mini protease inhibitor (Roche Diagnostics, Mannheim, Germany) and PhosSTOP phosphatase inhibitor cocktail (Roche Diagnostics). The cell suspension was subsequently subjected to ultra-sonication as follows: (samples were pulsed for 30 seconds at 100% amplitude, 80% interval followed by 30 seconds rest in 3 cycles; subsequently, a final pulse of 60 seconds at 100% amplitude, 100% interval was performed. The remaining cell debris was removed by centrifugation at 14,000 g for 5 minutes at room temperature. Proteins were then reduced, alkylated and digested using the FASP approach, in which the buffer is exchanged to 8 M urea pH 8.0 in order to remove the SDS present in the sample, as described previously[10]. Digestion was performed for 4 hours with Lys C (Wako, Richmond, VA, USA) after which the mixture was diluted 4-fold to 2M urea and digested with trypsin (Promega, Madison, WI, USA) at 37 °C overnight. Finally the sample was acidified with formic acid to a final concentration of 5%. Tryptic peptides were desalted using Sep-Pak C18 cartridges (Waters Corporation, Milford, MA). The peptides were subsequently labeled on-column with stable isotope dimethyl labeling as described previously[11], the resting state platelets were labeled 'light', whereas the stimulated platelets received the 'intermediate' label. Labeling efficiency was checked by LC-MS/MS before mixing the pools in a 1:1 ratio.

Strong Cation Exchange (SCX) Chromatography

The samples were dried ***in vacuo***, and re-suspended in 10% formic acid. SCX was performed on an Agilent 1100 HPLC system (Agilent Technologies, Waldbronn, Germany) consisting of a nanopump, a micro well plate autosampler, a multiple wavelength detector and a fraction collector. The columns used were a Opti Lynx C18, 40 µm, 100 Å, 2.1 mm x 15 mm (Optimize Technologies, Oregon, OR) for trapping and a Zorbax BioSCX-series II (Agilent Technologies, Waldbronn, Germany) 0.8 mm × 50 mm analytical column with 3.5 µm particles. A total of 250 µg of protein digest was loaded on the trapping column using a buffer containing 0.05% formic acid, pH 2.9 onto the trap column and subsequently eluted using a buffer containing 0.05% formic acid, 80% acetonitrile, pH 2.9 (buffer A) onto the analytical column. For SCX separation, buffer A and a buffer containing 0.05% formic acid, 80% acetonitrile, pH 2.9 and 500 mM NaCl (buffer B) were used. The separation was performed by a nonlinear 65 min elution gradient: from 0-17 min, 0% buffer B; from 17-22 min, from 2-3% buffer B; from 22-24 min, 3-5% buffer B; from 24-32 min, 5-8% buffer B; from 32-40 min, 8-20% buffer B; from 40-48 min, 20- 40% buffer B; from 48- 53 min, 40-90% buffer B. The column was subsequently washed for 3 min with 90% buffer B and finally equilibrated with 100% buffer A again for 9 mins. A total number of 50 SCX fractions were collected and dried in a vacuum centrifuge. Fractions 7-30 were reconstituted in 10% formic acid for further analysis.

Liquid chromatography and tandem mass spectrometry

LC–MS/MS was performed with a nano-LC coupled to an LTQ-Orbitrap Velos (Thermo Scientific, Bremen, Germany). The nano-LC consists of an Agilent 1200 series LC system equipped with a 20 mm ReproSil-Pur C18-AQ (Dr. Maisch GmbH, Ammerbuch, Germany) trapping column (packed in-house, i.d., 100 µm; resin, 5 µm) and a 400 mm ReproSil-Pur C18-AQ (Dr. Maisch GmbH, Ammerbuch, Germany) analytical column (packed in-house, i.d., 50 µm; resin, 3 µm) arranged in a vented-column configuration. The flow was passively split to 100 nl/min. A 3 hour elution profile consisting of: 0-10 min isocratic solvent A (0.1 M acetic acid) at 5 µL/min for sample trapping, followed by a gradient of 10.1-117 min, 10-25% solvent B (0.1 M acetic acid in 80% acetonitrile); 117.1-152 min, 25-50% solvent B; 152.1-154 min, 100% solvent B, 154.1-169, 0% solvent B. Nanospray was achieved using a distally coated fused silica emitter (made in-house, o.d. 375 µm; i.d. 20 µm) biased to 1.7 kV. The mass spectrometer was operated in the data-dependent mode to automatically switch between MS and MS/MS. The high resolution survey full scan was acquired in the orbitrap from m/z 350 to m/z 1500 with a resolution of 30.000 (FHMW) whereas the MS^2 scan is acquired at a resolution of 7500. The ten most intense precursors were isolated with an isolation width of 1.5 m/z and fragmented using a data-dependent decision tree method utilizing HCD, ETD-IT (ion trap read-out) and ETD-FT (orbitrap read-out)[12]. In brief, doubly charged peptides were always subjected to HCD fragmentation as well as triply charged peptides with an m/z >750 Th and are analyzed by the orbitrap. Triply charged peptides were fragmented by ETD when m/z < 750 Th. Analysis of ETD fragment ions was performed in the ion trap (ETD-IT). Quadruply charged ions were all fragmented with ETD and the fragments were analyzed either by ETD-IT (m/z >1000 Th) or the orbitrap (ETD-FT, m/z <1000 Th). Higher charged species were always analyzed by ETD-IT. The normalized collision energy for HCD was set to 35%. ETD reaction time was set to 50 ms for doubly charged precursors. Supplemental activation was enabled.

Data analysis: Identification and Quantitation

Peak lists were generated from the raw data files using the Proteome Discoverer software package version 1.3.339 (Thermo Scientific, Bremen, Germany). Peptide identification was performed by searching the individual peak lists (HCD, ETD-IT and ETD-FT) against a concatenated target-decoy database containing the human sequences in the Uniprot database (release 2012_06) supplemented with a common contaminants database using the Mascot search engine version 2.3 (Matrix Science, London, United Kingdom) via the Proteome Discoverer interface. The search parameters included the use of semitrypsin as proteolytic enzyme allowing up to a maximum of 2 missed cleavages. Carbamidomethylation of cysteines was set as a fixed modification whereas oxidation of methionines and the dimethyl “light” and “intermediate” labels on N-termini and lysine residues were set as variable modifications. Precursor mass tolerance was initially set at 50 ppm, while fragment mass tolerance was set at 0.6 Da for ETD-IT fragmentation and 0.05 Da for HCD

and ETD-FT fragmentation. Subsequently, the peptide identifications were filtered for true mass accuracy <4 ppm and an ion score of 40 until a False Discovery Rate (FDR) <1% at peptide level was achieved. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[13] with the dataset identifier PXD000072 and DOI 10.6019/PXD000072.

Peptide and protein quantification were performed using Proteome Discoverer as described previously[14]. Protein ratios were normalized per replicate based on the average ratio of the top 3 cytoskeletal platelet proteins based on the number of PSMs: Talin-1, Filamin-A and Myosin-9.

Releasate determination

For each biological replicate a 95% confidence interval was used to determine the outliers. The ratios were binned according to the number of quantification counts in such a fashion that each bin consists of at least 200 proteins and proteins with an identical quantification count can only be contributed to 1 bin. Proteins that were determined as outlier with a downregulated resting / activation ratio in at least 2 of 3 biological replicates are contributed to the releasate. The % release was subsequently calculated by (1 – the average ratio of the 3 replicates) x 100%, only proteins with a positive % were considered.

Release concentration calculation

The number of spectral counts for each protein was corrected for its respective molecular weight and served as a proxy for the protein abundance, as described previously by us and others[15-19]. Next, the proteome was calibrated using a set of proteins for which the copy numbers per platelet is known and interpolating the unknown proteins using a robust non-linear fit, similar to Burkhardt et al[15]. Assuming 200 million platelets per ml blood, the copy number per cell could be expressed as ng protein per ml blood by multiplying the copy number of a particular protein with the amount of platelets per ml blood and the molecular weight of the protein, divided by Avogadro's number. Finally we were able to calculate the amount of protein released into the bloodstream (if all platelets are assumed to be fully activated), by multiplying the ng/ml concentration of a particular protein with its respective % release as determined from the resting / activated ratio.

Enzyme linked immunosorbent assays (ELISA's) on releasate

The concentrations in the releasate of beta-thromboglobulin (β-TG, THBS1), platelet derived growth factor (PDGF)-B, platelet factor (PF4), chemokine (C-C motif) ligand 5 (RANTES, CCL5), and von Willebrand Factor (vWF) were determined with enzyme linked immunosorbent assays as follows. Washed platelets were stimulated with using 5 µg/mL collagen and 1U/mL thrombin for 5 minutes. Platelet suspensions were centrifuged at 4000g for 2 minutes and

the supernatant was collected. Nunc (Greiner) plates were coated with mouse anti-human β-TG (R&D, MAB393), mouse anti-human RANTES (R&D), recombinant PDGF-Rβ (R&D, DY220), mouse anti-human PF-4 (R&D, MAB7951), or rabbit anti-human vWF (DAKO, A0082). Standards for intrapolation of protein concentrations were prepared with normal pool serum, except for PDGF-BB that was prepared with recombinant protein. Plates were blocked with 1% bovine serum albumin, and subsequently incubated with supernatants and standards. Plates were washed with phosphate buffered saline (PBS) pH 7.4 with 0.05% Tween 20. Bound factors were detected with biotin coupled goat anti-human β-TG (R&D, BAF393), goat anti-human RANTES (R&D, AB287-NA), biotin coupled goat anti-human PDGF-BB (R&D, DY220), goat anti-human PF-4 (R&D, AF795), or horse radish peroxidase (HRP) conjugated rabbit anti-human vWF (DAKO, P0226). Plates were washed with PBS pH 7.4 with 0.05% Tween 20. Biotin coupled antibodies were detected with streptavidin-HRP (DAKO, P0397), goat anti-human antibodies with rabbit anti-goat HRP antibodies (DAKO, P0449).

Results and Discussion

Identity of the Platelet Releasate

Performing the reversed releasate quantitative proteomics experiments on three individuals resulted in the identification of 4375 unique proteins of which 2970 (68%) could be quantified in at least 2 out of 3 individuals (supplemental table I). First, we set out to identify the individual components within the releasate upon full platelet activation. Therefore, we evaluated the dimethyl intensity ratio between the light (resting) and intermediate (activated) labeled peptides in each individual (Figure 1).

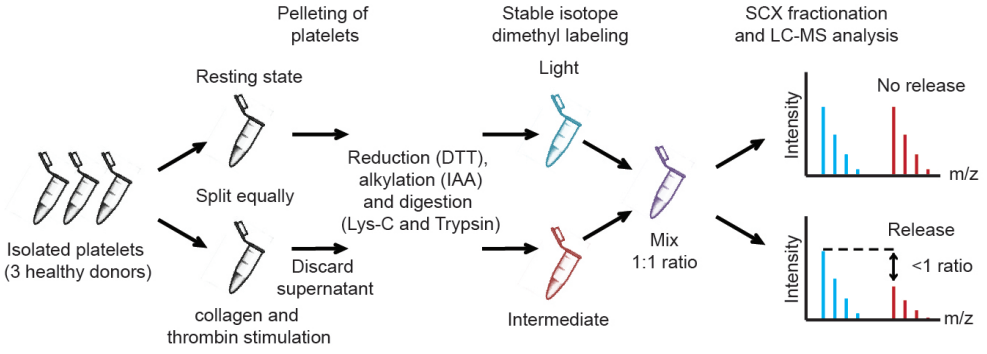

Figure 1: Flowchart depicting platelet reversed releasate analysis.

Platelets from three individual healthy donors were isolated and analyzed to exclude donor specific anomalies. Half of each platelet sample was left untreated, whereas the other half was stimulated with a combination of thrombin and collagen. Both samples were thoroughly washed and subsequently, the supernatant was removed. The platelets were lysed, digested with Lys-C and trypsin and differentially labeled with stable isotopes using dimethylation (light and intermediate), and subsequently mixed prior to LC MS/MS analysis.

As was hypothesized, the vast majority of proteins (>95%) had a resting/activated ratio close to 1.0 indicating that these proteins are not being released at a significant level of their copy number. These ratios were highly reproducible between replicates with a median RSD between ratios of 9.6% (Figure 2A, supplemental table I). In a typical proteomics experiment such as this one, the distribution of ratios is quite broad for proteins quantified with only one, or a few peptides, while the distribution of ratios becomes more discrete when more peptides are available for determining the exact ratio. Therefore, to identify proteins that are being released, a dynamic cut-off was used based on the number of peptides quantified. Using these statistical criteria described in the methods, 124 proteins were found to be significantly reduced in the activated platelets with copy number reductions between 25 and 90%. A selection of these known, and many novel releasate proteins is depicted in figure 2B alongside several matrix proteins, which did not show a reduction of their respective protein content (Table I).

Table 1: List of releasate proteins.

Accession	Genename	Description	Category	Release (%)	Concentration (ng/ml)	Releasate concentration (ng/ml)
P07996	THBS1	Thrombospondin-1	Cell and matrix interaction	74.9	3638.4	2261.21
P02775	PPBP	Platelet basic protein	Immune	86.8	14394.6	1473.93
P02776	PF4	Platelet factor 4	Immune	80.9	17575.7	1364.64
P02768	ALB	Serum albumin	Transport	83.7	2484.5	855.80
P02679	FGG	Fibrinogen gamma chain	Cell and matrix interaction	33.0	5607.1	667.44
P02675	FGB	Fibrinogen beta chain	Cell and matrix interaction	33.1	5430.7	614.39
P68871	HBB	Hemoglobin subunit beta	Transport	45.5	9121.1	513.41
P02671	FGA	Fibrinogen alpha chain	Cell and matrix interaction	34.7	1917.3	355.76
P02042	HBD	Hemoglobin subunit delta	Transport	46.2	5621.1	292.18
Q14315	FLNC	Filamin-C	Cell and matrix interaction	74.1	316.2	265.44
Q13201	MMRN1	Multimerin-1	Pro- and anticoagulation	31.4	1060.4	229.05
P69905	HBA1	Hemoglobin subunit alpha	Transport	46.7	3730.3	171.31
Q14766	LTBP1	Latent-transforming growth factor beta-binding protein 1	Growth factor	41.5	487.0	160.36
P04275	VWF	von Willebrand factor	Cell and matrix interaction	27.5	450.1	160.05
P09486	SPARC	SPARC	Transport	82.6	1034.2	146.81
P12259	F5	Coagulation factor V	Pro- and anticoagulation	36.3	391.7	145.68
P01857	IGHG1	Ig gamma-1 chain C region	Immune	81.7	854.5	120.26
P02787	TF	Serotransferrin	Transport	82.9	345.1	87.40
P02751	FN1	Fibronectin	Cell and matrix interaction	43.3	199.8	80.62
P01024	C3	Complement C3	Immune	83.3	153.5	80.39

Accession	Genename	Description	Category	Release (%)	Concentration (ng/ml)	Releasate concentration (ng/ml)
P01859	IGHG2	Ig gamma-2 chain C region	Immune	85.5	548.2	73.31
P10909	CLU	Clusterin	Other	36.2	807.3	72.33
P01860	IGHG3	Ig gamma-3 chain C region	Immune	76.0	511.0	68.83
P14543	NID1	Nidogen-1	Cell and matrix interaction	62.9	205.4	62.81
O43852	CALU	Calumenin	Transport	62.0	573.4	58.01
P05121	SERPINE1	Plasminogen activator inhibitor 1	Proteases and their inhibitors	80.6	385.6	56.75
P01137	TGFB1	Transforming growth factor beta-1	Growth factor	40.6	627.8	50.64
P07225	PROS1	Vitamin K-dependent protein S	Pro- and anticoagulation	79.2	217.1	46.61
P01834	IGKC	Ig kappa chain C region	Immune	80.1	992.8	45.42
P05067	APP	Amyloid beta A4 protein	Other	78.9	181.5	43.28
P01009	SERPINA1	Alpha-1-antitrypsin	Proteases and their inhibitors	73.3	311.1	41.35
P62805	HIST1H4A	Histone H4	Other	43.6	1263.4	32.39
P10124	SRGN	Serglycin	Proteases and their inhibitors	88.3	452.0	29.50
P02766	TTR	Transthyretin	Transport	65.2	602.7	27.72
P0CG05	IGLC2	Ig lambda-2 chain C regions	Immune	84.3	628.0	26.78
P00747	PLG	Plasminogen	Proteases and their inhibitors	69.7	97.9	18.92
Q14112	NID2	Nidogen-2	Cell and matrix interaction	58.0	70.3	17.69
P07093	SERPINE2	Glia-derived nexin	Proteases and their inhibitors	63.6	181.3	17.63
P07602	PSAP	Proactivator polypeptide	Other	40.8	201.4	16.95
P02647	APOA1	Apolipoprotein A-I	Transport	49.0	288.0	16.59
P02774	GC	Vitamin D-binding protein	Transport	83.1	117.2	16.39
P08603	CFH	Complement factor H	Immune	80.1	51.0	15.24
P02765	AHSG	Alpha-2-HS-glycoprotein	Transport	75.3	148.8	14.71
P04196	HRG	Histidine-rich glycoprotein	Cell and matrix interaction	74.7	104.2	14.39
P01033	TIMP1	Metalloproteinase inhibitor 1	Proteases and their inhibitors	81.9	183.7	12.14
P00915	CA1	Carbonic anhydrase 1	Other	50.4	214.9	11.26
P04114	APOB	Apolipoprotein B-100	Transport	71.0	13.8	10.34
P04217	A1BG	Alpha-1B-glycoprotein	Other	87.3	75.1	10.33
O00391	QSOX1	Sulfhydryl oxidase 1	Other	84.2	53.7	10.11
P02749	APOH	Beta-2-glycoprotein 1	Transport	87.8	97.2	10.00
P01008	SERPINC1	Antithrombin-III	Proteases and their inhibitors	75.0	84.3	9.88
P01876	IGHA1	Ig alpha-1 chain C region	Immune	63.7	127.1	9.86
P02790	HPX	Hemopexin	Transport	76.6	82.3	9.64
Q9UHQ9	CYB5R1	NADH-cytochrome b5 reductase 1	Other	37.1	208.0	9.40
P04004	VTN	Vitronectin	Cell and matrix interaction	49.1	111.0	9.31
P11233	RALA	Ras-related protein Ral-A	Other	48.0	225.7	9.27
Q15389	ANGPT1	Angiopoietin-1	Growth factor	67.2	80.2	9.12

Accession	Genename	Description	Category	Release (%)	Concentration (ng/ml)	Releasate concentration (ng/ml)
P29122	PCSK6	Proprotein convertase subtilisin/ kexin type 6	Other	31.3	80.0	7.84
P13501	RANTES (CCL5)	C-C motif chemokine 5	Immune	36.3	497.0	7.69
P01042	KNG1	Kininogen-1	Pro- and anticoagulation	71.6	54.2	7.59
P00738	HP	Haptoglobin	Transport	80.5	70.6	7.36
Q99969	RARRES2	Retinoic acid receptor responder protein 2	Other	74.3	152.4	7.07
P0C0L4	C4A	Complement C4-A	Immune	78.6	20.2	6.80
P01011	SERPINA3	Alpha-1-antichymotrypsin	Proteases and their inhibitors	75.0	67.0	6.79
Q15582	TGFBI	Transforming growth factor-beta-induced protein ig-h3	Growth factor	71.6	47.5	6.71
Q14624	ITIH4	Inter-alpha-trypsin inhibitor heavy chain H4	Proteases and their inhibitors	71.6	36.0	6.66
P49767	VEGFC	Vascular endothelial growth factor C	Growth factor	82.4	60.5	6.49
Q92896	GLG1	Golgi apparatus protein 1	Other	80.2	25.0	6.26
P00751	CFB	Complement factor B	Immune	78.7	35.2	5.90
P58215	LOXL3	Lysyl oxidase homolog 3	Other	72.8	38.4	5.88
P25311	AZGP1	Zinc-alpha-2-glycoprotein	Other	88.7	67.3	5.80
P01127	PDGFB	Platelet-derived growth factor subunit B	Growth factor	59.6	110.5	5.64
P16403	HIST1H1C	Histone H1.2	Other	50.5	149.4	5.38
Q9UBS4	DNAJB11	DnaJ homolog subfamily B member 11	Other	35.2	118.2	5.37
P03956	MMP1	Interstitial collagenase	Other	88.9	42.7	5.30
Q92734	TFG	Protein TFG	Transport	28.5	122.4	4.86
Q06481	APLP2	Amyloid-like protein 2	Other	77.2	28.5	4.56
Q08380	LGALS3BP	Galectin-3-binding protein	Cell and matrix interaction	55.2	46.1	4.37
P02763	ORM1	Alpha-1-acid glycoprotein 1	Other	84.3	75.4	4.34
Q96JJ7	TMX3	Protein disulfide-isomerase TMX3	Other	32.4	85.5	4.28
P19876	CXCL3	C-X-C motif chemokine 3	Immune	90.0	125.1	4.11
P02549	SPTA1	Spectrin alpha chain, erythrocyte	Cell and matrix interaction	30.6	20.9	4.00
Q8IUE6	HIST2H2AB	Histone H2A type 2-B	Other	46.1	177.4	3.96
Q16363	LAMA4	Laminin subunit alpha-4	Cell and matrix interaction	42.6	19.3	3.65
Q13627	DYRK1A	Dual specificity tyrosine-phosphorylation-regulated kinase 1A	Other	34.4	45.6	3.51
P02748	C9	Complement component C9	Immune	54.3	36.5	3.14
P02730	SLC4A1	Band 3 anion transport protein	Transport	49.2	26.1	3.06

Accession	Genename	Description	Category	Release (%)	Concentration (ng/ml)	Releasate concentration (ng/ml)
P61626	LYZ	Lysozyme C	Other	49.1	117.9	3.05
P33908	MAN1A1	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA	Other	65.6	26.7	3.01
P15907	ST6GAL1	Beta-galactoside alpha-2,6-sialyltransferase 1	Other	83.0	30.4	2.84
Q16706	MAN2A1	Alpha-mannosidase 2	Other	74.1	13.5	2.69
O60245	PCDH7	Protocadherin-7	Cell and matrix interaction	81.5	13.7	2.67
Q13103	SPP2	Secreted phosphoprotein 24	Proteases and their inhibitors	65.8	58.3	2.57
P02743	APCS	Serum amyloid P-component	Other	77.9	48.9	2.57
Q10471	GALNT2	Polypeptide N-acetylgalactosaminyltransferase 2	Other	31.6	46.6	2.51
Q12841	FSTL1	Follistatin-related protein 1	Other	52.2	50.7	2.48
P35625	TIMP3	Metalloproteinase inhibitor 3	Proteases and their inhibitors	63.2	58.7	2.48
P19652	ORM2	Alpha-1-acid glycoprotein 2	Other	86.6	45.1	2.41
Q16610	ECM1	Extracellular matrix protein 1	Growth factor	86.1	20.5	2.38
P01034	CST3	Cystatin-C	Proteases and their inhibitors	77.1	67.3	2.33
O14498	ISLR	Immunoglobulin superfamily containing leucine-rich repeat protein	Immune	60.6	30.8	2.08
P01871	IGHM	Ig mu chain C region	Immune	46.3	36.0	2.05
P56202	CTSW	Cathepsin W	Other	39.7	46.3	2.04
P00450	CP	Ceruloplasmin	Transport	72.5	11.6	2.04
P02753	RBP4	Retinol-binding protein 4	Other	89.5	38.5	2.01
O94923	GLCE	D-glucuronyl C5-epimerase	Cell and matrix interaction	35.4	32.9	2.00
P01596	KV104	Ig kappa chain V-I region CAR	Immune	76.9	75.7	1.98
P05452	CLEC3B	Tetranectin	Proteases and their inhibitors	85.9	39.3	1.94
P07357	C8A	Complement component C8 alpha chain	Immune	82.1	16.3	1.86
P05160	F13B	Coagulation factor XIII B chain	Pro- and anticoagulation	77.8	14.1	1.71
O94907	DKK1	Dickkopf-related protein 1	Other	85.8	24.7	1.41
P07358	C8B	Complement component C8 beta chain	Immune	89.6	10.6	1.24
Q06033	ITIH3	Inter-alpha-trypsin inhibitor heavy chain H3	Proteases and their inhibitors	44.8	12.4	1.12
O75167	PHACTR2	Phosphatase and actin regulator 2	Other	79.5	10.2	1.09
P13671	C6	Complement component C6	Immune	80.6	6.8	1.01
P03886	MT-ND1	NADH-ubiquinone oxidoreductase chain 1	Other	57.2	19.9	0.90
Q92520	FAM3C	Protein FAM3C	Other	70.5	21.6	0.84

Accession	Genename	Description	Category	Release (%)	Concentration (ng/ml)	Releasate concentration (ng/ml)
P49908	SEPP1	Selenoprotein P	Other	72.8	12.3	0.78
Q14353	GAMT	Guanidinoacetate N-methyltransferase	Other	65.2	20.2	0.77
Q96IY4	CPB2	Carboxypeptidase B2	Other	71.5	11.0	0.75
P01621	KV303	Ig kappa chain V-III region NG9 (Fragment)	Immune	80.9	33.1	0.70
P22105	TNXB	Tenascin-X	Cell and matrix interaction	55.9	1.9	0.68
P29622	SERPINA4	Kallistatin	Pro- and anticoagulation	62.9	11.0	0.65
Q7L3B6	CDC37L1	Hsp90 co-chaperone Cdc37-like 1	Other	42.0	13.7	0.46

The table includes Uniprot accession number; gene name; protein description; categorization for major platelet activation functions, % release, the interpolation of a non-linear robust fit of the ppm (ppm (each individual PSM corrected for MW normalized to the sum of all corrected PSMs; distributed over 1;000;000 ppm) with a set of proteins with known copy numbers converted to a concentration in ng/ml; the estimation of the concentration protein released per ml of blood.

The next step is to qualitatively compare the releasate determined in this study to several recent other studies. The proteome as determined in this study is highly comparable to a recently published platelet reference proteome (Figure 3A)[15], both in numbers as in the % of documented and/or predicted secreted proteins (Figure 3B). What is striking however, is that the releasate as determined in our study is remarkably small in relation to previous reports[8, 9]. These methods relied on a qualitative, rather than a quantitative method and are likely to contain many proteins originating from undesired platelet lysis. The strength of our novel method is reflected in the percentage of documented and/or predicted secreted proteins (based on SignalP[20] and/or Uniprot). Where the 124 candidates from our study convincingly show 80% meeting this criterium, the releasates from previous studies range from 30 to 60% (Figure 3B). When comparing all previous studies, the number of proteins attributed to the releasate goes up over time, while the % of documented/predicted secreted proteins in these releasates is decreasing. This suggests that the increased depth did not lead to the identification of more released proteins, but rather more false positives. Illustratively, all of the previous studies contained high abundant matrix proteins contaminating the releasate, such as Actin, Talin and others. This signifies that novel approaches are much-needed to identify novel releasate proteins, for instance by making use of relative quantitation by stable isotopes, as presented here.

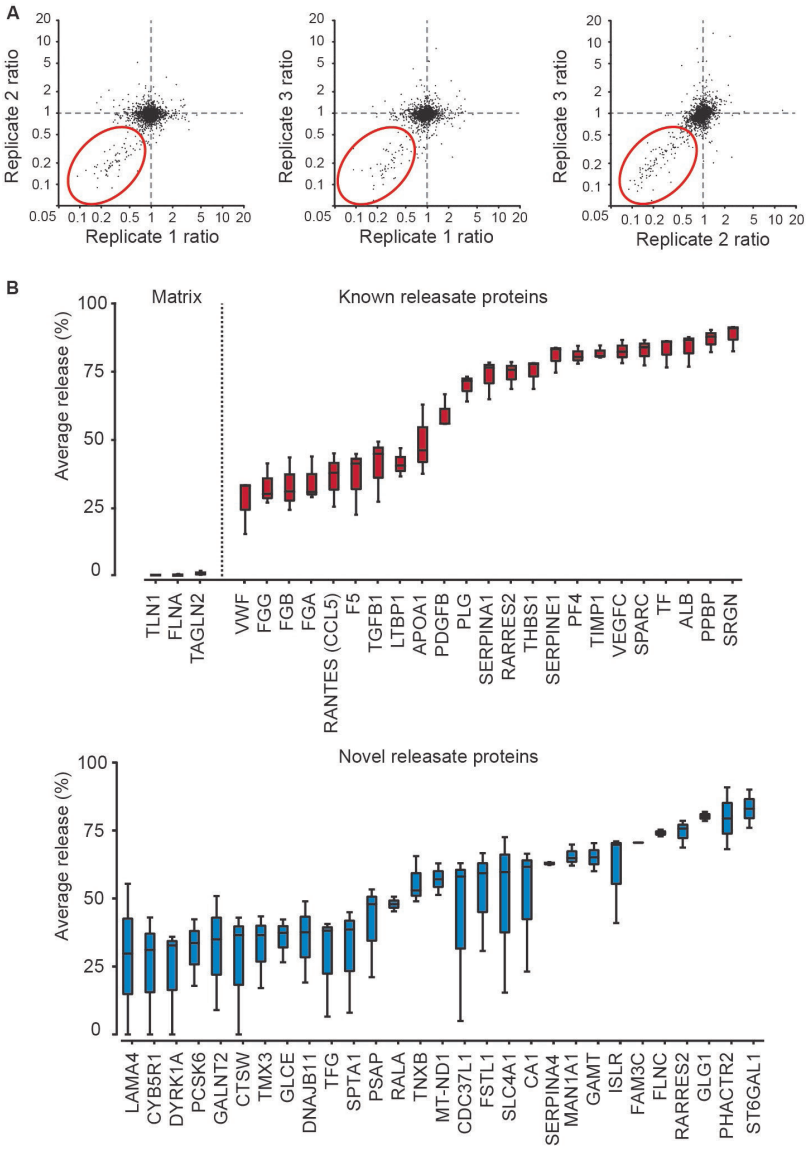

Figure 2: Quantitative profiling of the platelet reversed releasate.

(**A**) Reproducibility of the obtained quantitative release ratios. Correlation diagrams between the 3 biological replicates. The vast majority of the detected protein abundance ratios are close to 1 (high clustering of data points near the origin), the median coefficient of variation between the 3 replicas is 9.6% indicating the reproducibility of the measurements. Consistency within ratios of released proteins is revealed in the bottom left quadrant, as annotated by the red ovals. (**B**) Selection of quantified proteins either being strongly, weakly or not released upon platelet activation. Different releasate proteins release their copy number to a different extend as depicted by boxplots which indicate the percentage of release (average and standard deviation acquired from experiments with 3 donors). Displayed are: non-released matrix proteins (black), a selection of known releasate proteins (red) and novel releasate proteins (blue). All 124 released proteins are listed in supplemental table II.

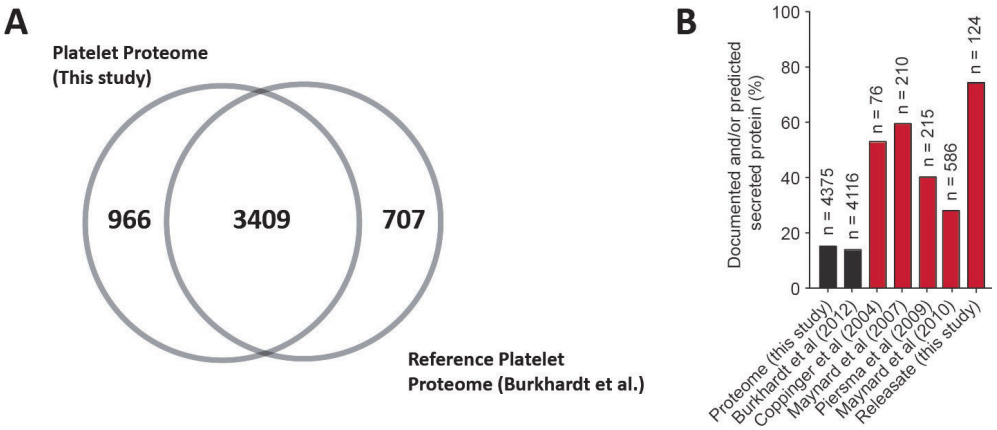

Figure 3: Qualitative comparison of our dataset with reference proteomes.

(A) Overlap in protein identifications between the platelet proteomes determined in this study (left) and recently by Burkhardt et al.[15] (right). Both studies used a similar LC-MS/MS approach leading to similar numbers of identifications with a good overlap in protein identifications. (**B**) The prevalence of signal/secretion tags (using SignalP[20] and Uniprot) is highest in our reversed releasate dataset, when compared to previously published releasate proteomes[6-9] (red) and the full platelet proteomes[15] (black) of Burkhardt et al.[15] and this study. The total number of non-redundant proteins identified in each of the releasate studies (n) is indicated above the bar.

As a benchmark for our approach, all currently known, important releasate proteins are present in our dataset such as Platelet factor 4 (PF4), RANTES (CCL5), von Willebrand Factor (vWF) etc. (Figure 2B). In order to have an indication as to whether the proteins that are attributed to the releasate are in fact feasible candidates, all of these proteins were subjected to a functional protein association network analysis (Figure 4). A large portion of the releasate proteins found in this study already have a known relationship with one or more other proteins and are clustered according to their similar function. Many of these proteins either have a direct or an indirect link to one of the benchmark proteins. The fibrinogen complex is visible in the light green cluster, apart from the fibrinogen subunits it contains factor XIIIb. Once fibrinogen is converted to insoluble fibrin strands, they are cross linked with factor XIII to form a blood clot[21]. The proteins in the orange cluster are proteins that are involved in inhibiting clot formation. Both these clusters represent some of the key functions of platelets. Although most proteins are connected and can be clustered according to function, there are also a lot of unconnected potentially novel proteins of which either the function has not been elucidated yet (e.g. Follistatin-related protein 1 indicated as FSTL1) or the protein's function is only elucidated in a different cell type (e.g. Tenascin X, indicated as TNXB) that is also plausible in platelets. These proteins may provide additional information about platelet functioning.

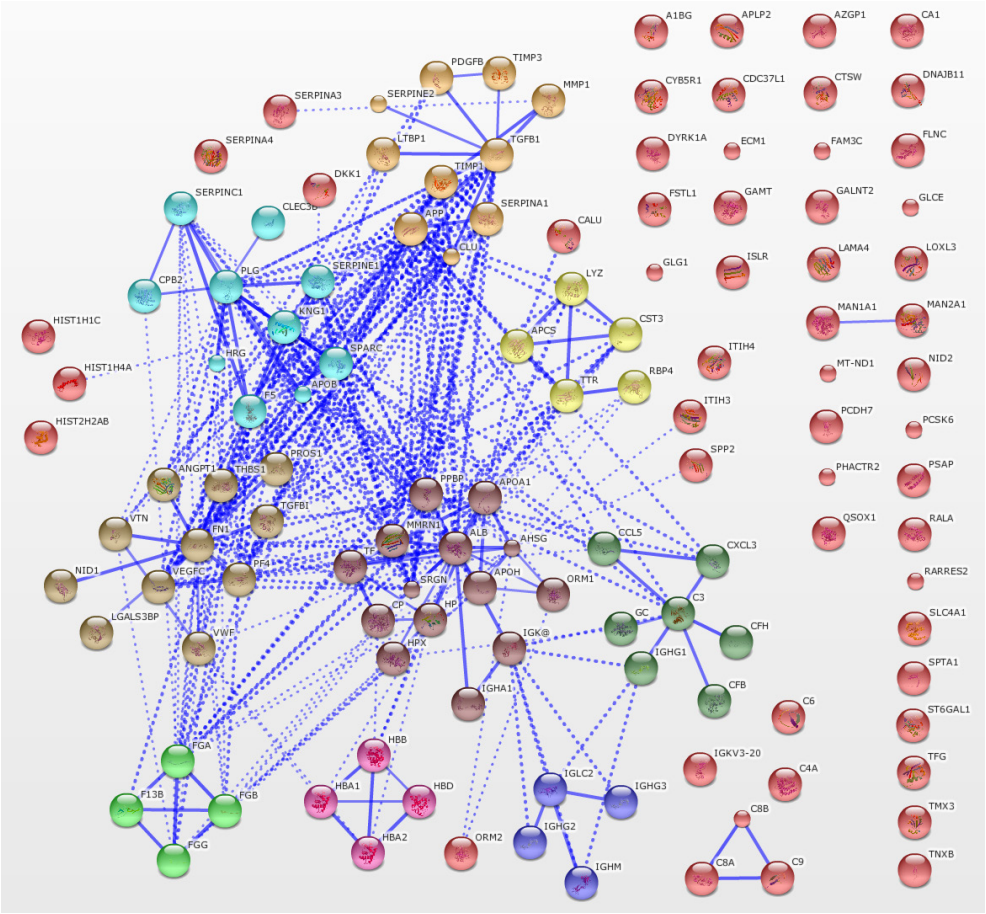

Figure 4: Network analysis of releasate proteins.

Network analysis of the releasate proteins using String version 9.05 in which the line thickness connecting the need indicate the confidence of the interaction. Proteins are clustered into 10 clusters using KMEANS based on the String global scores. The unconnected (red) nodes are candidate novel releasate proteins that require further evaluation.

Quantitation of the Releasate; Estimating the Concentration in Blood

An added advantage of our approach is that it also allows us to semi-quantify the absolute protein levels of the full platelet proteome[15-19], and consequently also of the released proteins. We used a robust non-linear fit of the molecular weight corrected spectral counts to a set of proteins with a known copy number per cell[15]. Overall, our platelet proteome compares both qualitatively (Figure 3A) and quantitatively (Figure 5) very well with the 4116 unique proteins identified by Burkhart et al. recently[15]. The qualitative overlap is about 80%, which is as good as can be expected in proteomics from a biological replicate. The quantitative comparison between both experiments yields a very good correlation (R^2=0.901), for virtually all proteins the maximum difference is about one order of magnitude however for the vast majority of proteins the copy numbers are much more similar, also in line with the expected accuracy of such an absolute quantitation experiment[17, 22].

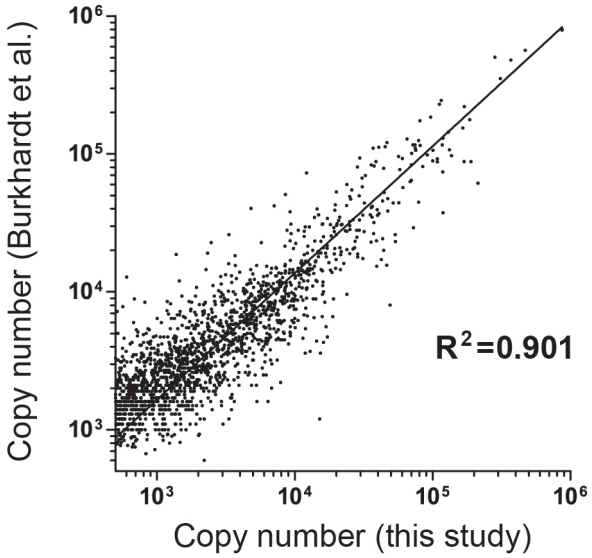

Figure 5: Qualitative and quantitative comparison of our dataset with a reference proteome.

The correlation between the copy number as determined in this study (x-axis) and that from the data of Burkhardt et al[15] (y-axis) shows strong correlation as represented by the Pearson coefficient (R^2).

If we now combine the absolute quantitation results with the determined release fraction (Figure 2B) for every of the 124 releasate proteins and assume an average number of $2x10^8$ platelets per milliliter blood, we can calculate the concentration of each protein in the releasate in ng/mL blood (see methods for details). After performing this calculation, we observed that the protein concentrations in the releasate span over 5 orders of magnitude, ranging from highly abundant (µg/ml) hallmark releasate proteins such as thrombospondin-1, platelet basic protein and PF4, to lower abundant (fg/ml) novel releasate proteins such as Tenascin-X and Kallistatin (Figure 4, supplemental table II). To assess whether these numbers

make sense we performed ELISA experiments for several releasate proteins within the supernatant of the stimulated platelets (Figure 6A). The calculated released concentrations based on MS-data match with the concentration determined by ELISA within a margin of 1 order, validating that our calculated proteomics-based protein abundance is a good proxy for the actual content that is released into the blood. Finally, we classified all the proteins of the releasate according to major functional categories associated with platelet activation (Figure 6B).

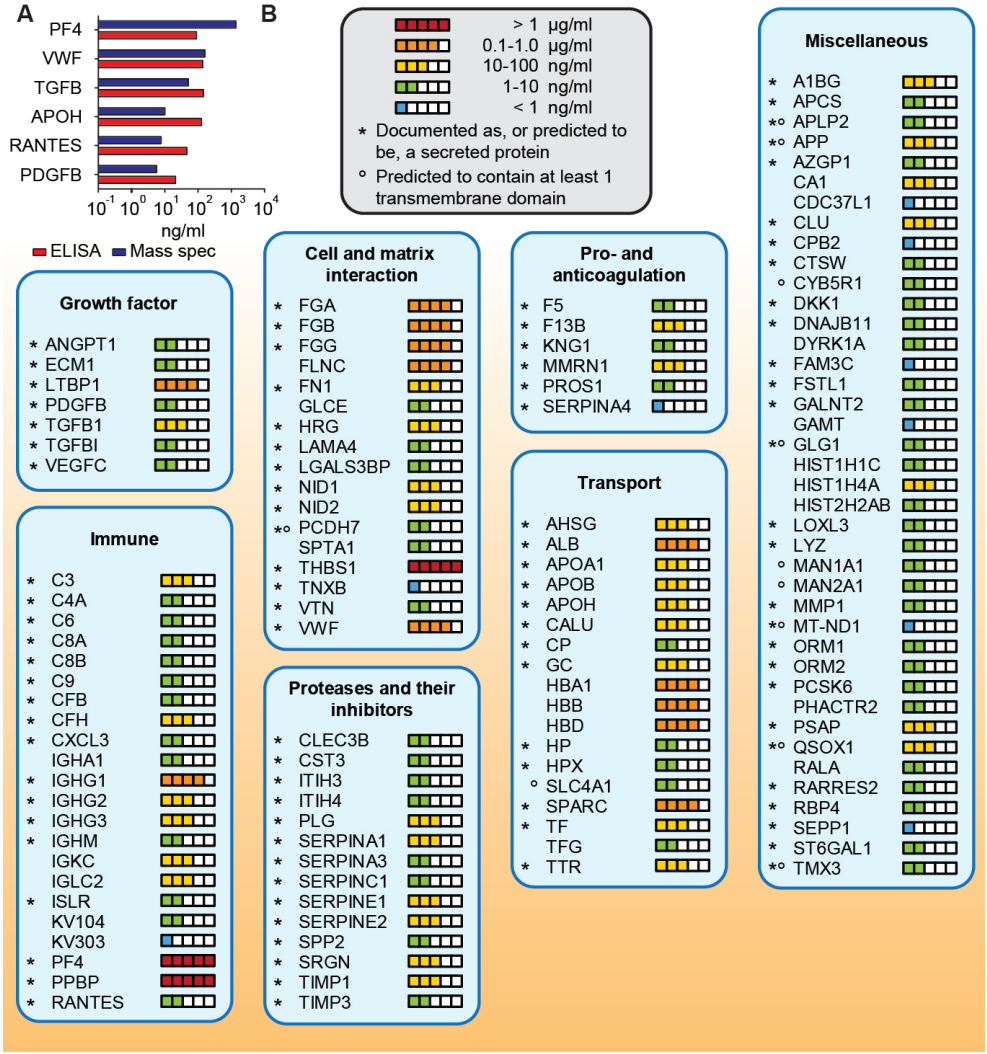

Figure 6: Quantitation and functional characterization of proteins in the platelet releasate.

(A) Protein concentration in the platelet releasate of a selected set of proteins in ng/ml as determined by MS-based absolute quantitation (blue), and by ELISA experiments on the supernatant (red). These data are in favorable agreement. (B) Functional categorization of the 124 platelet releasate proteins into the major functions known for platelet activation. The asterisk indicates a documented or predicted secreted protein, the ° indicates a putative shedded protein due to presence of at least one transmembrane domain predicted by hidden Markov model23. The colored boxes indicate the estimated amount of release.

Closer inspection revealed that proteins involved in cell and matrix interactions, which is associated with wound healing, were amongst the highly released proteins (~2 µg/mL). Similarly proteins belonging to the classes transport and immune are on average relatively more abundant than proteins belonging to the proteases and inhibitors, growth factor and pro- and anticoagulation classes as can be expected. Another phenomenon that was observed was that several proteins that were attributed to the releasate contain transmembrane domains and are likely the result of protein shedding, such as Amyloid-like protein 2 (APLP2) as indicated in Figure 6. For APLP2 it is known that this protein is shed as has been reported by Endres et al24. However, peptides that could prove that this was indeed the result of shedding are not found for this and all other potentially shedded proteins.

There are also a number of novel releasate proteins found in this study, as determined by String network analysis (Figure 2, Figure 4 and Figure 6B). For example: Follistatin-like protein 1 (FSTL1) has been found in osteoblasts and fibroblasts and characterized to play a role in the inflammatory pathway [25]; filamin C (FLNC) plays a role in the crosslinking of actin in muscle cells26; and tenascin-X (TNXB) binds to several collagen subtypes and plays a role in wound healing, as shown in patients with Ehlers-Danlos syndrome[27].

Our quantitative method is advantageous compared to the methods previously employed to determine the releasate with regard to accuracy as the change of finding false positives is reduced. However, like any other method there are drawbacks. In this case the ability to determine whether or not a protein is released is based on the statistical significance of a ratio compared to the distribution, if a protein is released only to a limited extend (in our case <25% of the copy number) the likeliness that it is considered an outlier is decreased as the protein ratio is either much closer to or disappears in the local distribution of ratios. Another class of releasate proteins that is difficult to identify with this method are the proteins that are released by shedding. This involves transmembrane proteins to be cleaved by a protease upon activation after which an extracellular fragment of the protein is released. The fact that the ratios of these proteins are much closer to the distribution makes their identification more difficult. In addition the shedded protein can only be identified by comparing each peptide ratio, as only the ratio of the peptides belonging to the shedded protein fragment will decrease upon release. The other peptide ratios of that protein will remain the same. The evidence of these shedding events is in the data generated by this method but the usability heavily depends on the sequence coverage in the shedded and unshedded part of the protein and manual verification.

In conclusion, we have developed a robust reversed releasate quantitative protein profiling method to more reliably identify and semi-quantify the platelet releasate with high accuracy. It covers an unprecedented depth of 5 orders of magnitude, exposing many novel molecular determinants of the platelet releasate to inspire further investigations into their roles.

Supplementary Information

Supplementary information can be accessed online on http://atvb.ahajournals.org.

Data Supplement 1 contains the supplemental methods for the article that are also incorporated in the text of this chapter.

Data Supplement 2 contains supporting information that was not in the article but is incorporated in this chapter.

Data Supplement 3 contains 2 supplemental tables:

Supplemental table I contains the full proteome as identified in this study.

Supplemental table II contains an extended version of Table I as found in this chapter.

Acknowledgements

This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine (http://www.ctmm.nl), project CIRCULATING CELLS (grant 01C-102), and supported by the Dutch Heart Foundation. The Netherlands Proteomics Centre embedded in the Netherlands Genomics Initiative is kindly acknowledged for financial support. The Cardiovascular “Focus en Massa” program at Utrecht University is acknowledged for additional financial support. Mirjam Damen is acknowledged for technical assistance.

References

1. Heemskerk JW, Bevers EM, Lindhout T. Platelet activation and blood coagulation. ***Thromb Haemost***. 2002;88:186-193
2. Li Z, Delaney MK, O'Brien KA, Du X. Signaling during platelet adhesion and activation. ***Arterioscler Thromb Vasc Biol.*** 2010;30:2341-2349
3. Semple JW, Italiano JE, Jr., Freedman J. Platelets and the immune continuum. ***Nat Rev Immunol.*** 2011;11:264-274
4. Peterson JE, Zurakowski D, Italiano JE, Jr., Michel LV, Connors S, Oenick M, D'Amato RJ, Klement GL, Folkman J. Vegf, pf4 and pdgf are elevated in platelets of colorectal cancer patients. ***Angiogenesis.*** 2012;15:265-273
5. de Groot PG, Urbanus RT, Roest M. Platelet interaction with the vessel wall. ***Handb Exp Pharmacol.*** 2012:87-110
6. Maynard DM, Heijnen HF, Gahl WA, Gunay-Aygun M. The alpha-granule proteome: Novel proteins in normal and ghost granules in gray platelet syndrome. ***J Thromb Haemost.*** 2010;8:1786-1796
7. Maynard DM, Heijnen HF, Horne MK, White JG, Gahl WA. Proteomic analysis of platelet alpha-granules using mass spectrometry. ***J Thromb Haemost.*** 2007;5:1945-1955
8. Coppinger JA, Cagney G, Toomey S, Kislinger T, Belton O, McRedmond JP, Cahill DJ, Emili A, Fitzgerald DJ, Maguire PB. Characterization of the proteins released from activated platelets leads to localization of novel platelet proteins in human atherosclerotic lesions. ***Blood.*** 2004;103:2096-2104
9. Piersma SR, Broxterman HJ, Kapci M, de Haas RR, Hoekman K, Verheul HM, Jimenez CR. Proteomics of the trap-induced platelet releasate. ***J Proteomics.*** 2009;72:91-109
10. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. ***Nat Methods.*** 2009;6:359-362
11. Boersema PJ, Raijmakers R, Lemeer S, Mohammed S, Heck AJ. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. ***Nat Protoc.*** 2009;4:484-494
12. Frese CK, Altelaar AF, Hennrich ML, Nolting D, Zeller M, Griep-Raming J, Heck AJ, Mohammed S. Improved peptide identification by targeted fragmentation using cid, hcd and etd on an ltq-orbitrap velos. ***J Proteome Res.*** 2011;10:2377-2388
13. Vizcaino JA, Cote R, Reisinger F, Barsnes H, Foster JM, Rameseder J, Hermjakob H, Martens L. The proteomics identifications database: 2010 update. ***Nucleic Acids Res***. 2010;38:D736-742
14. Scholten A, Mohammed S, Low TY, Zanivan S, van Veen TA, Delanghe B, Heck AJ.In-depth quantitative cardiac proteomics combining electron transfer dissociation and the metalloendopeptidase lys-n with the silac mouse. ***Mol Cell Proteomics.*** 2011;10:O111 008474
15. Burkhart JM, Vaudel M, Gambaryan S, Radau S, Walter U, Martens L, Geiger J, Sickmann A, Zahedi RP. The first comprehensive and quantitative analysis of human platelet protein composition allows the comparative analysis of structural and functional pathways. ***Blood***. 2012;120:e73-82

16. Malmstrom J, Beck M, Schmidt A, Lange V, Deutsch EW, Aebersold R. Proteome-wide cellular protein concentrations of the human pathogen leptospira interrogans. ***Nature.*** 2009;460:762-765
17. Peng M, Taouatas N, Cappadona S, van Breukelen B, Mohammed S, Scholten A, Heck AJ. Protease bias in absolute protein quantitation. ***Nat Methods.*** 2012;9:524-525
18. Lu P, Vogel C, Wang R, Yao X, Marcotte EM. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. ***Nat Biotechnol.*** 2007;25:117-124
19. Aye TT, Scholten A, Taouatas N, Varro A, Van Veen TA, Vos MA, Heck AJ. Proteome-wide protein concentrations in the human heart. ***Mol Biosyst.*** 2010;6:1917-1927
20. Petersen TN, Brunak S, von Heijne G, Nielsen H. Signalp 4.0: Discriminating signal peptides from transmembrane regions. ***Nat Methods.*** 2011;8:785-786
21. Muszbek L, Bereczky Z, Bagoly Z, Komaromi I, Katona E. Factor xiii: A coagulation factor with multiple plasmatic and cellular functions. ***Physiol Rev.*** 2011;91:931-972
22. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. Global quantification of mammalian gene expression control. ***Nature.*** 2011;473:337-342
23. Sonnhammer EL, von Heijne G, Krogh A. A hidden markov model for predicting transmembrane helices in protein sequences.***Proc Int Conf Intell Syst Mol Biol.*** 1998;6:175-182
24. Endres K, Postina R, Schroeder A, Mueller U, Fahrenholz F. Shedding of the amyloid precursor protein-like protein aplp2 by disintegrin-metalloproteinases. ***Febs J.*** 2005;272:5808-5820
25. Wilson DC, Marinov AD, Blair HC, Bushnell DS, Thompson SD, Chaly Y, Hirsch R. Follistatin-like protein 1 is a mesenchyme-derived inflammatory protein and may represent a biomarker for systemic-onset juvenile rheumatoid arthritis. ***Arthritis Rheum.*** 2010;62:2510-2516
26. Fujita M, Mitsuhashi H, Isogai S, Nakata T, Kawakami A, Nonaka I, Noguchi S, Hayashi YK, Nishino I, Kudo A. Filamin c plays an essential role in the maintenance of the structural integrity of cardiac and skeletal muscles, revealed by the medaka mutant zacro. ***Dev Biol.*** 2012;361:79-89
27. Burch GH, Gong Y, Liu W, Dettman RW, Curry CJ, Smith L, Miller WL, Bristow J. Tenascin-x deficiency is associated with ehlers-danlos syndrome. ***Nat Genet.*** 1997;17:104-108

4.

Defining proteome signatures of neutrophil, eosinophil and basophil granulocytes

Patrick Wijten[1-2], Tamar Tak[3], Leo Koenderman[3], Arjen Scholten[1-2*] and Albert J.R. Heck[1-2*]

[1]Biomolecular Mass Spectrometry and Proteomics, Utrecht Institute for Pharmaceutical Sciences and Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands

[2]Netherlands Proteomics Centre, Utrecht, The Netherlands

[3]Department of Respiratory Medicine, University Medical Centre Utrecht, Utrecht, The Netherlands

Abstract

Circulating neutrophils, eosinophils and basophils play essential roles during microbe induced and sterile inflammation. While these cell types have overlapping functions they certainly also have diverging roles. This has become quite apparent from studies focusing on a single process. However, no comparative differential characterization is available, for instance at the full proteome level, for at least the eosinophils and basophils, but also so far published neutrophil proteomes lack a depth that can be attained by current proteomics technologies. Using high-end mass spectrometry based proteomics technology (including 2D-chromatography, ETD and HCD peptide fragmentation), we here present comprehensive reference proteomes of these well-sorted three distinctive granulocytes. The depth of the proteome analysis allows us to identify and quantify 3596 neutrophil proteins, 4174 eosinophil proteins and 5012 basophil proteins, spanning in abundance typically a dynamic range of about 5 orders of magnitude. Although a large overlap in between the three proteomes is apparent, we used unbiased cluster analysis to define a set of signature proteins for each of these cell types. Subsequent GO term enrichment proved successful to identify key neutrophil proteins and many novel potential effectors. The same analysis proved less successful for eosinophils and basophils, nevertheless a wealth of information is available for both cell types for future studies where our data can serve as reference resource proteomes.

Introduction

Inflammation is a paramount component of the immune response, which is initiated as a response to either a real or perceived threat to tissue homeostasis[1]. Granulocytes are the key effector cells of several inflammatory responses and as such are recruited in large numbers. Granulocytes mature in the bone marrow and are released fully differentiated as either neutrophils, eosinophils or basophils into the circulation[2]. These mature granulocytes can be distinguished by differences in nuclear morphology, size, granule content and function. Apart from morphological and functional differences these cells also differ in abundance with neutrophils being the most abundant of the granulocytes with 2.0-7.5 x 10^9 cells per liter of blood, eosinophils at least one order of magnitude less present with 0.04-0.44 x 10^9 cells per liter of blood and basophils being the least abundant with 0.0-0.1 x 10^9 cells per liter of blood[1]. Terminally differentiated granulocytes are non-dividing cells (G_0 phase) for the duration of their relatively short lifespan[3-5]. The life span of circulating neutrophils in absence of extracellular stimuli is within 1-5 days[5]. Upon inflammation granulocytes migrate into tissues where they reside for prolonged periods of time[4]. Phagocytosis by resident macrophages (efferocytosis) in liver, spleen and bone marrow takes care of clearing these cells[6, 7]. Each of these subsets of granulocytes are triggered by specific stimuli originating from either pathogens or other immune cells to respond and are among the first cells to arrive at the inflammatory tissue.

Neutrophils are the first line of host defense against a variety of infectious pathogens including but not limited to bacteria, fungi and protozoa[8]. Upon infection neutrophils are recruited to the site of infection where they recognize and phagocytose microbes followed by killing the pathogens through cytotoxic mechanisms such as the production of reactive oxygen species, release of proteases and antimicrobial peptides and expulsion of their nuclear contents to form neutrophil extracellular traps[9-11]. Besides their role in the acute inflammatory response neutrophils are also able to communicate with macrophages, dendritic cells and cells of the adaptive immune system through cell-cell contract or soluble effectors[12-14].

The function of eosinophils in homeostasis is less clear compared to that of neutrophils[15]. For a long time it was believed that eosinophils promote immunity against helminthes, but this has been questioned by recent animal studies showing that eosinophils prevent inducible NO synthase production in macrophages and neutrophils essential for enhancing parasite clearing resulting in sustained chronic infection and ensure parasite survival in the host[16, 17]. However, it is clear that eosinophils play an important role in the pathophysiology of allergic diseases, such as allergic asthma where they are recruited to and activated in the lung tissue[18, 19]. Recent findings suggest that the collateral damage in the airways in asthma patients originates from dysregulation of tissue eosinophilis[20-22]. This hypothesis is supported in asthma models in eosinophil deficient mice[23, 24]. It is however clear that eosinophils are capable of sophisticated immune functions, shown by their nuanced degranulation

responses to distinct stimuli[15, 25, 26]. This is further demonstrated by their ability to engage in complex interactions with a wide variety of other leukocytes such as T cells, T_H cells, B cells, dendritic cells and mast cells[27-32].

Basophils were historically primarily associated with immediate hypersensitivity reactions. This was based on their cell surface expression of the high-affinity IgE receptor (Fcε R1) which upon crosslinking release histamine and other atopy mediators[33]. Now it has become apparent that these cells also express a plethora of surface receptors that whilst ligated are capable to activate basophils to produce cytokines that in turn promote and regulate Th2 adaptive immune responses[34].

Acquiring a comprehensive proteomic coverage of these cells can aid in the understanding of their diverse functions, confirm expression of transcripts and aid in finding altered pathways in disease. Extensive proteomic coverage is available for many cell lines, tissues and cells but for many hematopoietic cells like the granulocytes described in this study this coverage is limited or unavailable. From a proteomics perspective concerning neutrophils Zhou et al., Rørvig et al. and Tomazella et al. yielded the most data rich studies to date[35-37]. Tomazella et al. (~1200 proteins) analyzed both the detergent-insoluble and whole cell lysate fractions of resting neutrophils whereas Rørvig et al. (~1500 proteins) analyzed the subsets of the neutrophil granular proteome and Zhou et al. (~2300) analyzed whole cell lysates of neutrophils originating from several healthy donors and severe trauma patients. For eosinophils Straub et al and Wilkerson et al. yielded the most in depth proteomes[38, 39]. Straub[39] et al. did 2D-GE on eosinophils yielding 423 unique proteins. Wilkerson[38] et al. on the other hand employed an LC-MS workflow on whole cell lysate (~4400 proteins), which represents the most comprehensive proteomics dataset to date. As for the basophils, no significant proteomics studies have been published. Here we analyzed the proteomes of neutrophils, eosinophils and basophils from healthy donors using a LC-MS approach in order to develop reference proteomes and define a protein signature for each type of granulocyte.

Methods

Neutrophil, Eosinophil and Basophil Isolation

Cells were isolated from heparinized blood by lysing erythrocytes with 5:1 lysis buffer consisting of 150 mM NH_4Cl, 10 mM $KHCO_3$ and 0.1 mM NA_2EDTA in ddH_2O, at an osmolarity of 305 mOsm. Resulting whole leukocyte preparations were stained with CD41-FITC (clone VIPL3, Life Technologies, Carlsbad, CA, USA), CD16-Krome Orange (clone 3G8 Beckman Coulter Pasadena, CA, USA), CD193-Alexa647 (clone 83103, BD Biosciences San Jose, CA, USA), and CD9 (clone S32, a kind gift from Ing. D. Kanters) in PBS supplemented with 0.32% trisodium citrate and 10% human pasteurized plasma solution (Sanquin, Amsterdam,

the Netherlands). After two washes, anti-moue-IgM-PE (SouthernBiotech, Birmingham, AL, USA) was added as a secondary antibody for CD9. Cells were sorted on a FACSAria-III cell-sorter (BD Biosciences, San Jose, CA, USA) with the following gating strategy (figure 1): SSC^{high},$CD16^{high}$,$CD9^{negative}$ for neutrophils without bound platelets, eosinophils SSC^{high}, $CD16^{negative}$, $CD9^{high}$, $CD193^{positive}$, $CD41^{negative}$ and basophils SSC^{low}, $CD193^{positive}$, $CD9^{high}$, $CD41^{negative}$. Cells were collected and washed 3 times in protein-free PBS before snap-freezing and storage in liquid nitrogen.

Sample preparation

The stored neutrophils, eosinophils and basophils were thawed and centrifuged at 14,000 g for 10 min at 4 °C. The pellet was reconstituted in a buffer containing 100 mM Tris, 10 mM DTT, 2% SDS at pH 8.0 with Complete Mini protease inhibitor (Roche Diagnostics, Mannheim, Germany) and PhosSTOP phosphatase inhibitor cocktail (Roche Diagnostics). The cell suspension was subsequently subjected to ultra-sonication as follows: (samples were pulsed for 30 seconds at 100% amplitude, 80% interval followed by 30 seconds rest in 3 cycles; subsequently, a final pulse of 60 seconds at 100% amplitude, 100% interval was performed. The remaining cell debris was removed by centrifugation at 14,000 g for 5 minutes at room temperature. Proteins were then reduced, alkylated and digested using the FASP approach, in which the buffer is exchanged to 8 M urea pH 8.0 in order to remove the SDS present in the sample, as described previously[40]. Digestion was performed for 4 hours with LysC (Wako, Richmond, VA, USA) after which the mixture was diluted 4-fold to 2M urea and digested with trypsin (Promega, Madison, WI, USA) at 37 °C overnight. Finally the sample was acidified with formic acid to a final concentration of 5%. Tryptic peptides were desalted using Sep-Pak C18 cartridges (Waters Corporation, Milford, MA).

Strong Cation Exchange (SCX) Chromatography

The samples were dried *in vacuo*, and re-suspended in 10% formic acid. SCX was performed on an Agilent 1100 HPLC system (Agilent Technologies, Waldbronn, Germany) consisting of a nanopump, a micro well plate autosampler, a multiple wavelength detector and a fraction collector. The columns used were a Opti Lynx C18, 40 μm, 100 Å, 2.1 mm x 15 mm (Optimize Technologies, Oregon, OR) for trapping and a Zorbax BioSCX-series II (Agilent Technologies, Waldbronn, Germany) 0.8 mm × 50 mm analytical column with 3.5 μm particles. A total of 250 μg of protein digest was loaded on the trapping column using a buffer containing 0.05% formic acid, pH 2.9 onto the trap column and subsequently eluted using a buffer containing 0.05% formic acid, 80% acetonitrile, pH 2.9 (buffer A) onto the analytical column. For SCX separation, buffer A and a buffer containing 0.05% formic acid, 80% acetonitrile, pH 2.9 and 500 mM NaCl (buffer B) were used. The separation was performed by a nonlinear 65 min elution gradient: from 0-17 min, 0% buffer B; from 17-22 min, from 2-3% buffer B; from 22-24 min, 3-5% buffer B; from 24-32 min, 5-8% buffer B; from 32-40 min, 8-20% buffer B; from 40-48 min, 20- 40% buffer B; from 48- 53 min, 40-90%

buffer B. The column was subsequently washed for 3 min with 90% buffer B and finally equilibrated with 100% buffer A again for 9 mins. A total number of 50 SCX fractions were collected and dried in a vacuum centrifuge. Fractions 6-27 were reconstituted in 10% formic acid for further analysis.

Liquid chromatography and tandem mass spectrometry

LC–MS/MS was performed with a nano-UPLC coupled to LTQ-Orbitrap Q-Exactive (Thermo Fisher Scientific, Bremen, Germany). The nano-LC consists of an Proxeon EASY-nLC 1000 system (Thermo Scientific, Bremen, Germany) equipped with a 20 mm ReproSil-Pur C18-AQ (Dr. Maisch GmbH, Ammerbuch, Germany) trapping column (packed in-house, i.d., 100 μm; resin, 3 μm) and a 450 mm Poroshell 120 EC-C18 (Agilent Technologies, Little Fall, DE, USA) analytical column (packed in-house, i.d., 50 μm; resin, 2.7 μm) arranged in a vented-column configuration. Based on the intensity of the UV signal on SCX, the fractions were either submitted to a 3 hour elution profile consisting of solvent A (0.1 M acetic acid) and solvent B (0.1 M acetic acid in 80% acetonitrile) at 100 nl/min after sample trapping at nl/min, the gradient consists of 0-151 min, 7-30% solvent B, 151-154 min, 30-100% solvent B; 154-159 min, 100% solvent B, 159-160, 100-7% solvent B, followed by 160-180 min equilibrating at 7% B or a 90 minute elution profile consisting of solvent A (0.1 M acetic acid) and solvent B (0.1 M acetic acid in 80% acetonitrile) at 100 nl/min after sample trapping at nl/min, the gradient consists of 0-61 min, 7-30% solvent B, 61-64 min, 30-100% solvent B; 64-69 min, 100% solvent B, 69-70, 100-7% solvent B, followed by 70-90 min equilibrating at 7% B. Nanospray was achieved using a distally coated fused silica emitter (made in-house, o.d. 375 μm; i.d. 20 μm) biased to 1.7 kV. The mass spectrometer was operated in the data-dependent mode to automatically switch between MS and MS/MS. The high resolution survey full scan was acquired in the orbitrap from m/z 350 to m/z 1500 with a resolution of 35.000 (FHMW) whereas the MS^2 scan is acquired at a resolution of 17.500. The ten most intense precursors were isolated with an isolation width of 1.5 m/z and fragmented using HCD. The target ion setting was 3E6 for MS and 5E4 for MS/MS, with a maximum fill-time of 250 ms and 120 ms.

Data analysis: Identification and Quantitation

Peak lists were generated from the raw data files using the Proteome Discoverer software package version 1.4.0.288 (Thermo Scientific, Bremen, Germany). Peptide identification was performed by searching the peak lists against a concatenated target-decoy database containing the human sequences in the Swissprot database (release 2014_08) supplemented with a common contaminants database using the Mascot search engine version 2.4 (Matrix Science, London, United Kingdom) via the Proteome Discoverer interface. The search parameters included the use of trypsin as proteolytic enzyme allowing up to a maximum of 2 missed cleavages. Carbamidomethylation of cysteines was set as a fixed modification whereas oxidation of methionines was set as variable modification.

Precursor mass tolerance was initially set at 50 ppm, while fragment mass tolerance was set at 0.05 Da using HCD fragmentation. A target false discovery rate (FDR) of <1% at the peptide level was attained using the percolator algorithm[41]. Subsequently, the peptide identifications were filtered for an ion score of 20. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[42] with the dataset identifier PXD001716 and DOI 10.6019/PXD001716.

Peptide and protein quantification were performed using Proteome Discoverer as described previously[43].

Clustering of protein quantitation data

Proteins with missing ppm values were substituted by a ppm value 1 order of magnitude below the lowest determined ppm value. For each protein the fraction per subtype of the total granulocyte population was calculated which were subjected to Fuzzy C-means clustering using R with the Mfuzz plugin (2.3.1)[44]. In addition to assigning each protein into a cluster, Fuzzy clustering indicates how well each protein fits in each cluster. By choosing a minimal required membership value, outliers are excluded from analysis. The required number of clusters was determined by calculating the average centroid distance for clustering runs with 2 up to 40 clusters[45]. After 6 clusters, addition of extra clusters did little to improve clustering results. Thus the amount of clusters was set to 6.

Results and Discussion

Here we describe an in-depth proteome analysis of neutrophil, eosinophil and basophil granulocytes that were isolated by means of FACS sorting from a healthy donor using aforementioned gating strategy. Samples were re-analyzed after sort, which revealed a neutrophil purity of 99,8%, eosinophil purity of 98,9% and basophil purity of 94,8%, the contaminants were likely cells that did not survive the sort as microscopic evaluation of May-Grünwald-Giemsa stained cytospins revealed >99% purity (Figure 1). Each cell-type was subjected to the described proteomics workflow separately for label-free analysis, a technical replicate was included for each sample.

The correlation is the highest as expected between replicates ($\rho \geq 0.93$) for all granulocytes, whereas a lower correlation but still high similarity between different types of granulocytes (ρ: 0.76-0.84) is observed (Figure 2A). Subsequently replicates were combined in order to get a more comprehensive dataset. Correlation between neutrophils and eosinophils is ρ=0.83, between neutrophils and basophils ρ=0.78 and highest between eosinophils and basophils (ρ=0.86) which is intuitive as both originate from a common progenitor[2], on the other hand their functions are quite divergent.

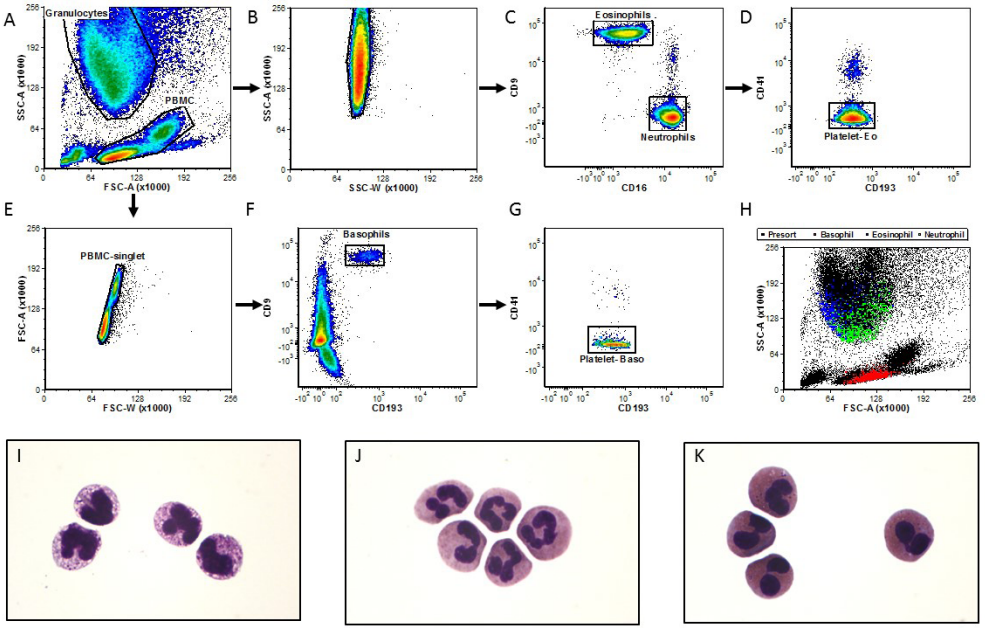

Figure 1: Gating strategy of granulocyte sort.

Granulocytes and PBMC's were separated on FSC/SSC (**A**). For both populations, doublets were excluded **(B** and **E**). Neutrophils and eosinophils were separated based on CD16 and CD9 (**C**). To prevent contamination by adherent platelets, only $CD9^{negative}$ neutrophils were sorted. Similarly, platelet-free eosinophils were sorted by excluding cells positive for platelet marker CD41 (**D**). Basophils were sorted based on CD9 and CD193 (**F**) with exclusion of $CD41^{positive}$ cells (**G**). Re-analysis of the sorted populations (**H**) reveals neutrophils, eosinophils and basophils on their expected locations in the scatterplot. Microscopic evaluation of sorted populations after May-Grünwald-Giemsa staining clearly show the distinct phenotypes of basophils (**I**), neutrophils (**J**) and eosinophils (**K**).

After combining the replicates the proteome of the neutrophils is comprised of 3596 proteins, the eosinophil proteome consists of 4174 proteins and the basophil proteome is encompasses 5012 proteins. The difference in size between these proteomes can be explained due to differences in dynamic range, a quarter of the protein content of the neutrophils is comprised of a mere 4 proteins (Neutrophil defensin 1, Protein S100-A8, Protein S100-A9 and Histone H4) while the same amount of protein content is comprised of 15-20 proteins for basophils and eosinophils (Figure 3A). The qualitative overlap between the 3 types of granulocytes is high with 3051 proteins being found in all datasets (Figure 3B). Considering their common lineage this might be expected.

The neutrophil and basophil datasets are the biggest generated to date, the eosinophil dataset is comparable to the biggest dataset generated. Some proteomics studies are available for neutrophils and eosinophils but to the best of our knowledge none of the sort exist for basophils. The existing neutrophil proteomes of Tomazella et al.[36], Rørvig et al.[35] and

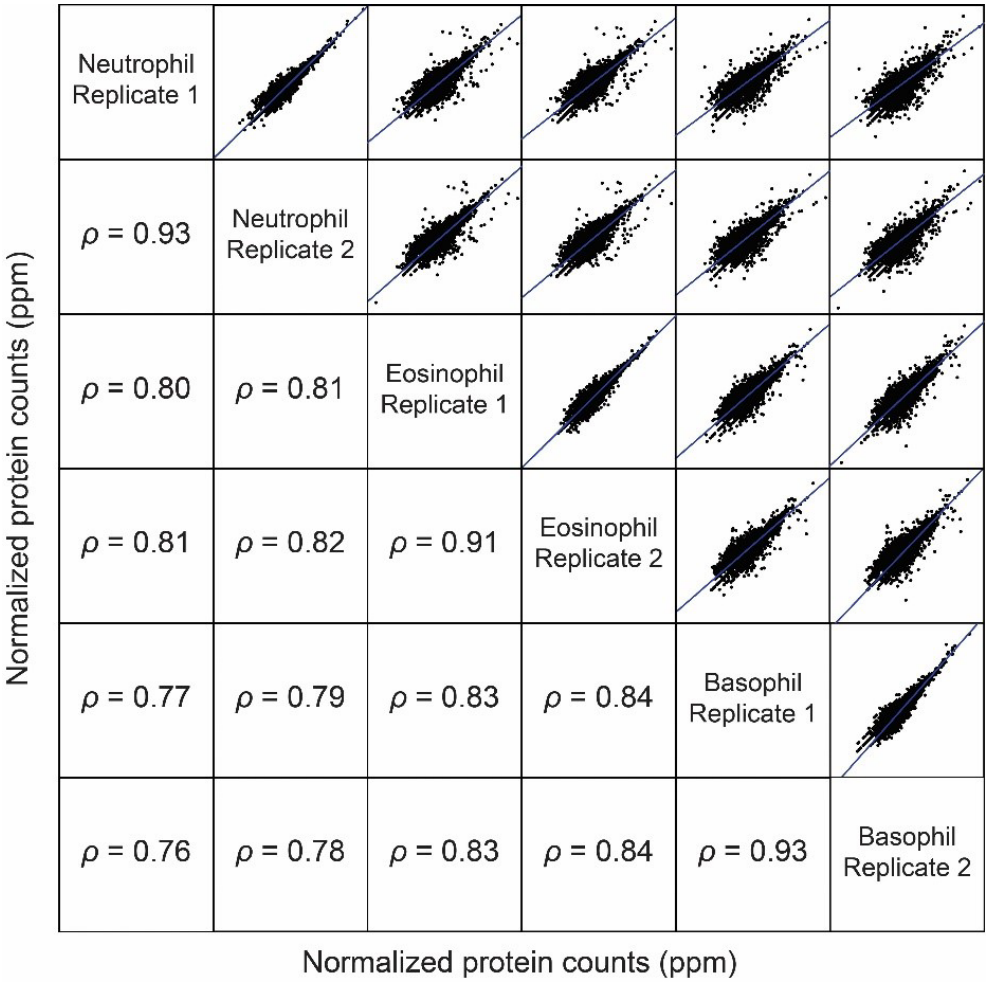

Figure 2: Scatter plot matrix and Spearman correlation of the 2 technical replicates of each granulocytes subtype.

A high Spearman correlation between the two technical replicates of neutrophil, basophil and eosinophil granulocytes was obtained based on the ppm normalized proteomes.

Zhou et al.[37] and eosinophil proteome of Straub et al.[39] and Wilkerson et al.[38] were compared to their corresponding proteomes generated in this study after conversion of accession numbers to uniprot IDs, only unique IDs and the primary ID per PSM were considered for comparison (Figure 4).

The overlap between the neutrophil proteomes is overall quite well when looking at Tomazella et al. (82%) Rørvig et al. (74%) and this study, the majority of proteins found in those studies are also found in this study. For Zhou et al. the overlap is much lower (20%). Considering that both

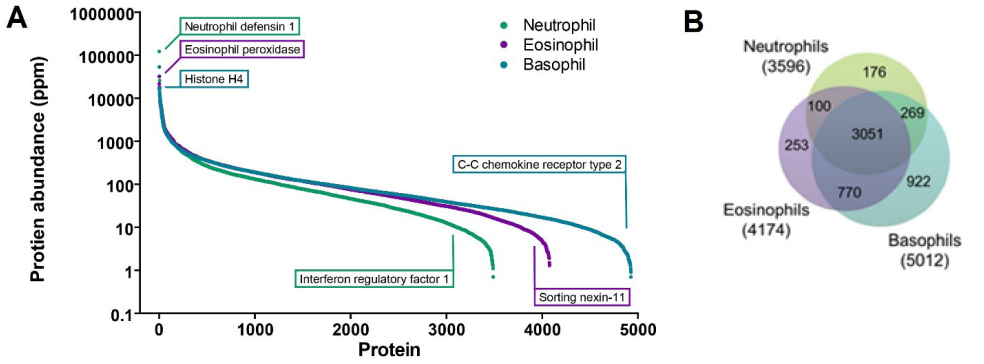

Figure 3A: Protein abundance map of the granulocyte subtypes.

Depicted are the neutrophil proteome (green), eosinophil proteome (purple) and the basophil proteome (blue) in ppm. Highlighted are the most abundant proteins, some hallmark proteins (if applicable) and the lowest abundant proteins per cell type. B) Venn diagram depicting the overlap in proteome between the subtypes of granulocytes. For each of the subtype of granulocytes the unique protein entries of the combined replicates are compared.

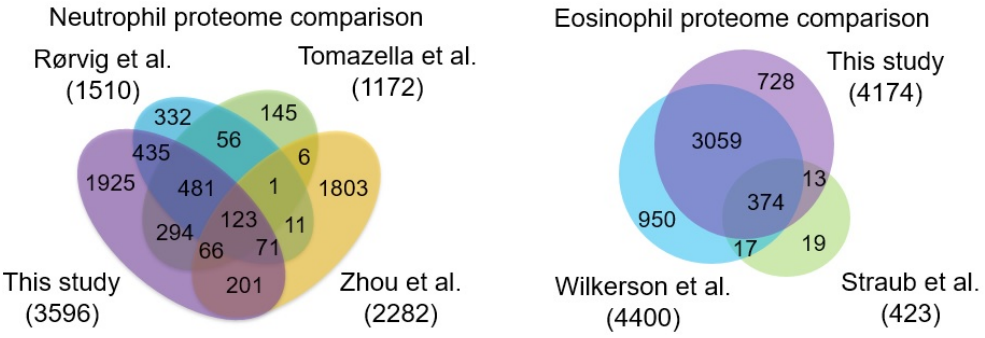

Figure 4: Venn diagram depicting the overlap in proteome between this and other studies.

The neutrophil proteome as determined in this study was compared to the neutrophil proteomes as described by Tomazella et al.[36], Rørvig et al.[35] and Zhou et al[37]. Similarly the eosinophil proteome as determined in this study was compared to the eosinophil proteome as described by Straub et al.[39] and Wilkerson et al.[38]. All comparisons were made based on unique Uniprot accession numbers.

Zhou et al. and this study analysed the whole cell lysate and this study was performed using a more advanced proteomics workflow one would expect a much larger degree of overlap. An explanation for this difference could be that Zhou et al. listed the no longer maintained IPI annotated proteins which had to be converted to Uniprot accessions for comparison, for each conversion 1 or more Uniprot accessions were generated, the first of which was taken for comparison which likely resulted in a less than ideal conversion. Ideally the data from Zhou et al. would have been re-analyzed using the search engine used in this study to ensure identical treatment in assigning accession numbers, however raw data was unavailable.

Abundant proteins

A first glance at the granulocyte subtypes is taken by looking at the most abundant proteins for each proteome (Table 1). This reveals several proteins known to be highly abundant in neutrophils as expected, such as neutrophil defensin 1 (DEF1)[46], protein S100-A8 and –A9 [47] as well as lysozyme C[48]. Similarly eosinophil peroxidase and eosinophil cationic protein[49] can be found in the eosinophil proteome. For basophils the high abundant proteins are not as clearly defined but it has a high abundant protein in common with eosinophils in galectin-10[50] as well as DEF1 which it has in common with neutrophils.

Another striking feature in the most abundant proteins in the 3 granulocytes is the presence of a lot of histones. Histones are known as the main protein components of chromatin on which the DNA is winded. However, recently a novel function of first neutrophils and later eosinophils

Table 1: The 20 most abundant proteins of neutrophils, eosinophils and basophils.

Rank	Neutrophil (49%)	Eosinophil (44%)	Basophil (47%)
1	Neutrophil defensin 1	Histone H4	Histone H4
2	Protein S100-A8	Galectin-10	Galectin-10
3	Protein S100-A9	Eosinophil peroxidase	Neutrophil defensin 1
4	Histone H4	Proteoglycan 3	Actin, cytoplasmic 1
5	Actin, cytoplasmic 1	Actin, cytoplasmic 1	Histone H2B type 1-K
6	Actin, aortic smooth muscle	Actin, aortic smooth muscle	Histone H2B type 2-F
7	Actin, alpha skeletal muscle	Actin, alpha skeletal muscle	Histone H2B type 1-M
8	Lysozyme C	Bone marrow proteoglycan	Histone H2B type 1-L
9	Lactotransferrin	Non-secretory ribonuclease	Histone H2A.Z
10	Histone H2B type 1-K	Eosinophil cationic protein	Histone H2A type 1-H
11	Histone H2B type 2-F	Protein S100-A8	Histone H2A type 1-B/E
12	Histone H2B type 1-L	Histone H2A.Z	Histone H3.1
13	Histone H2B type 1-M	Histone H2A type 1-H	Histone H2B type 1-J
14	Histone H2B type 1-J	Histone H2A type 1-B/E	Actin, aortic smooth muscle
15	Histone H2A.Z	Beta-actin-like protein 2	Actin, alpha skeletal muscle
16	Histone H2A type 1-H	Histone H2A type 2-B	Histone H2A type 2-B
17	Beta-actin-like protein 2	Histone H2B type 1-K	Vimentin
18	Histone H2A type 1-B/E	Profilin-1	Beta-actin-like protein 2
19	Cathepsin G	Histone H2B type 1-M	Profilin-1
20	Profilin-1	Histone H2B type 2-F	Histone H1.2

The top proteins in terms of protein content as determined by normalized molecular weight corrected spectral counts. In parenthesis is presented how much these top 20% proteins roughly contribute to the total proteome mass.

has been discovered in which these histones are part of novel antimicrobial mechanisms called neutrophil/eosinophil/basophil extracellular traps (NETs[9], EETs[51] and BETs[52]). The main protein components for NET's are histones H1, H2A, H2B, H3 and H4 along with neutrophil elastase, calprotectin (protein S100 A8/A9 complex)[53], myeloperoxidase, bactericidal/permeability-increasing protein[54], cathepsin G, lactotransferrin[55], matrix metalloproteinase-9, peptidoglycan recognizing proteins, pentraxin[56] and LL-37[57]. All of which are present in the neutrophil proteome and to varying degree in the eosinophil and basophil proteome. The protein constituents of the EETs and BETs have not been elucidated yet. Besides these components 3 types of actin are highly abundant; the cytoplasmic actin, the aortic smooth muscle actin and the alpha skeletal muscle actin. The cytoplasmic actin is to be expected, the other two were identified based on 1 unique peptide each, sharing 15 peptides with the cytoplasmic actin. Considering the shared peptides account for the vast majority of the detected signal it is likely that these peptides originate from the cytoplasmic actin.

Clustering and GO analysis

The most abundant proteins emphasize the common elements between the types of granulocytes (matrix, nucleus and potentially extracellular trap proteins), as over half of these proteins are abundantly found in all granulocytes. In an attempt to display the differences all proteins were clustered according to ppm ratio of the specific granulocyte subtype as fraction of the total granulocyte content as described in the methods. The optimal number of clusters was determined to be 6 as depicted in figure 5. Clusters 1-3 represent proteins from which probably a proteome signature can be extracted for neutrophils, eosinophils and basophils, respectively. Proteins in cluster 4 are low in neutrophils and relatively high in eosinophils and basophils. Clusters 5 and 6 are centered on a relative abundance of 0.3 with one of the cell types being a slight outlier, thus both clusters are indicative for proteins that are present in all granulocytes and similar in expression.

In order to determine whether there are clearly defined differences in inferred function between the proteomes, all protein clusters were subjected to gene ontology (GO) enrichment analysis using GOrilla[58]. Clusters were analysed for enriched on either biological processes, molecular functions or cellular components based on annotations found in literature. For GO enrichment analysis protein clusters are tested for enrichment of GOterms in that specific cluster with the whole granulocyte proteome as background dataset. Cluster 1, the neutrophils specific proteins illustrate enrichment in GO terms one might expect, the top 10 GO terms are listed in table 2. Amongst those GO terms there are many that are related to immunological functions, especially related to defence against bacteria and fungi. For example the proteins known to act in the defence response to bacteria found in our dataset include: neutrophil elastase (ELANE)[59], several lectins (CLEC4D, CLEC4E)[60], cathelicidin antimicrobial peptide (CAMP)[61], defensins[62] and many others. For defence against fungi identified proteins include amongst others

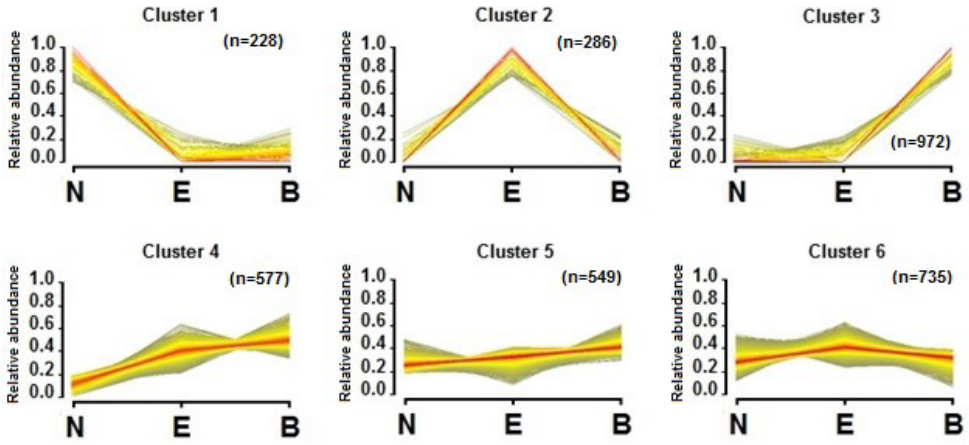

Figure 5: Soft clustering of all proteins identified.

Proteins were subjected to unsupervised clustering with the Fuzzy c-means algorithm which resulted in 6 patterns of dynamic changes.

lactotransferrin[63] and cathepsin G[64]. Furthermore components related to the membrane and receptor are also enriched, amongst others CD93[65], and more importantly CXCR2 which is also known as a interleukin-8 receptor; interleukin 8 induces chemotaxis in primarily neutrophils[66].

The same analysis was performed for cluster 2 of the eosinophil specific proteins and cluster 3 of the basophil specific proteins, but no relevant GO terms were enriched significantly. This absence of evidence for enriched terms does not imply the absence of function but has more to do with the inherent incompleteness of the Gene Ontology[67]. GO is a representation of the

Table 2: The most significant enriched GO terms in cluster 1.

GO Term	Description	FDR q-value	Enrichment	Type
GO:0006952	defense response	5.47E-20	3.75	Process
GO:0005576	extracellular region	7.65E-20	4.64	Component
GO:0006955	immune response	4.87E-14	3.58	Process
GO:0005615	extracellular space	2.02E-13	3.37	Component
GO:0031224	intrinsic component of membrane	1.17E-11	2.08	Component
GO:0016021	integral component of membrane	4.31E-10	2.02	Component
GO:0006954	inflammatory response	8.33E-09	4.14	Process
GO:0042742	defense response to bacterium	9.91E-09	5.95	Process
GO:0009617	response to bacterium	1.20E-08	5.56	Process
GO:0031012	extracellular matrix	1.87E-08	6.99	Component

These GO terms are common for all neutrophils.

current state of knowledge and as such is very dynamic. In order to find enriched GO terms it requires these terms to be annotated in the first place. For many of the proteins encountered in the eosinophil and basophil proteome there was no GO annotation available. But as was the case for neutrophils, the primary chemo attractant (eotaxin for eosinophils and basophils) receptor (CCR3) was found in both proteomes[68-70]. Clusters 4-6 are indicative of proteins common between the proteomes of these granulocytes and only yielded common aspects such as membrane-bound vesicle, focal adhesion and antigen processing and as such will not be further discussed.

Protein signatures

In order to still be able determine what sets these cell types apart, a protein signature per granulocyte subtype was defined according to the particular protein abundance in a proteome as a fraction of the total granulocyte abundance for that protein. A 4-fold cut-off compared to the fractional abundance in the other two proteomes was applied for a protein to be specific for a cell type, the resulting protein signatures of 30 proteins per cell-type are listed below (Table 3) and proteins relevant to immune functions will discussed.

Table 3: Protein signatures for neutrophils, eosinophils and basophils.

Rank	Neutrophil	Eosinophil	Basophil
1	Neutrophil defensin 1	Eosinophil peroxidase	Tryptase alpha/beta-1
2	Protein S100-A8	Proteoglycan 3	A-kinase anchor protein 12
3	Protein S100-A9	Eosinophil cationic protein	Neuroblast differentiation-associated protein AHNAK
4	Lysozyme C	Arachidonate 15-lipoxygenase	N(G),N(G)-dimethylarginine dimethylaminohydrolase 2
5	Lactotransferrin	Fructose-1,6-bisphosphatase 1	Synaptotagmin-like protein 3
6	Cathepsin G	Plasminogen activator inhibitor 2	Acylphosphatase-1
7	Neutrophil gelatinase-associated lipocalin	Retinoid-inducible serine carboxypeptidase	Costars family protein ABRACL
8	Cathelicidin antimicrobial peptide	Alpha-endosulfine	CB1 cannabinoid receptor-interacting protein 1
9	Annexin A3	Gamma-interferon-inducible lysosomal thiol reductase	Transcription elongation factor A protein 1
10	Neutrophil elastase	Perilipin-2	Protein NDRG1
11	Resistin	MARCKS-related protein	Cytochrome b5 type B
12	Matrix metalloproteinase-9	cAMP-regulated phosphoprotein 19	Hepatoma-derived growth factor-related protein 3
13	Azurocidin	Indoleamine 2,3-dioxygenase 1	Dipeptidyl peptidase 4

Rank	Neutrophil	Eosinophil	Basophil
14	Haptoglobin	Glutathione S-transferase Mu 3	Mast cell carboxypeptidase A
15	Erythrocyte band 7 integral membrane protein	Platelet-activating factor acetylhydrolase 2, cytoplasmic	ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 1
16	Leukotriene A-4 hydrolase	Fructose-1,6-bisphosphatase isozyme 2	Septin-1
17	Peptidoglycan recognition protein 1	CD81 antigen	Runt-related transcription factor 1
18	Neutrophil collagenase	Serine protease 33	V-type proton ATPase 16 kDa proteolipid subunit
19	Chitinase-3-like protein 1	Protein spinster homolog 3	15-hydroxyprostaglandin dehydrogenase [NAD(+)]
20	Chitotriosidase-1	Protein Hikeshi	High mobility group nucleosome-binding domain-containing protein 5
21	Arginase-1	Synaptosomal-associated protein 29	Deoxyribonuclease-2-alpha
22	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-10	Sialic acid-binding Ig-like lectin 8	Serine/threonine-protein kinase Nek7
23	Alpha-1-acid glycoprotein 1	CCAAT/enhancer-binding protein epsilon	NADH dehydrogenase [ubiquinone] iron-sulfur protein 5
24	Phospholipase B-like 1	Ribosomal protein S6 kinase alpha-2	Hematopoietic prostaglandin D synthase
25	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-7	EGF-like module-containing mucin-like hormone receptor-like 1	Small nuclear ribonucleoprotein E
26	Cysteine-rich secretory protein 3	Protein THEMIS2	Plasminogen receptor (KT)
27	Olfactomedin-4	TBC1 domain family member 5	GRB2-related adapter protein 2
28	Leukotriene-B(4) omega-hydroxylase 2	Ectonucleoside triphosphate diphosphohydrolase 1	Interleukin-3 receptor subunit alpha
29	Pentraxin-related protein PTX3	Sphingomyelin phosphodiesterase 3	Germinal center-associated signaling and motility-like protein
30	CD177 antigen	Nesprin-2	Cytochrome c oxidase copper chaperone

The top specific proteins in terms of protein content as determined by normalized molecular weight corrected spectral counts and at least a ratio of 4 compared to the other type of granulocytes.

Proteome signature of neutrophils

Many of the neutrophil specific proteins exhibit antimicrobial function as could be expected up front. One of the hallmark neutrophil effector proteins is Neutrophil defensin 1 (DEF1), which is localized in azurophil granules[71]. It exerts its antimicrobial activity by permeabilizing the plasma membrane of microbes[72, 73]. Despite DEF1 being a hallmark protein for neutrophils it was also detected to substantial degree in basophils, the abundance in neutrophils however is so much higher that it was indicated as a specific protein. Tschopp et al. described the occurrence of DEF1 in basophils by means of a Western blot before[74]. Like DEF1, Lysozyme C is found in azurophil granules, it primarily has a bacteriolytic function, responsible for cleaving peptidoglycans that are found in the cell walls of bacteria[71, 75.] Protein S100-A8 and Protein S100-A9 are also main components in neutrophils[48, 53], these small calcium binding proteins are located in the cytosol and are found at high levels in the extracellular milieu during inflammatory conditions as a result of neutrophil cell death[76]. The dimer S100-A8/A9 (calprotectin) is proposed to have a pro-inflammatory role by means of induction of chemotaxis and adhesion neutrophils to fibrinogen by means of macrophage-1 antigen (MAC1) upregulation[77]. Lactotransferrin also called lactoferrin is found in specific granules and exerts antimicrobial activity by sequestering free iron which in turn inhibits microbial growth[78]. Besides this indirect mode of action lactotransferrin is also reported to have a direct bactericidal activity by inducing the release of lipopolysaccharides from the bacterial outer membrane[79]. Annexin A3, like lactotransferrin is mostly found in specific granules[80], and is implicated to be involved in phagocytosis by means of facilitating granule fusion[81]. Cathelicidin antimicrobial peptide (CAMP) is known to bind to bacterial lipopolysaccharides, but also exert direct antimicrobial activity after it is processed by a serine protease dependent mechanism into dermcidin and LL-37 upon secretion into the phagosome[82]. Neutrophil gelatinase-associated lipocalin (NGAL) is the marker protein for specific granules, its anti-bacterial activity is mediated by binding to bacterial sidophores that are responsible for iron trafficking from the host to the microbe to enhance growth. By this activity NGAL inhibits growth of bacteria[83, 84]. Neutrophil elastase is involved in the hydrolysis of proteins including elastin. It inhibits C5a-dependent neutrophil enzyme release and chemotaxis and thereby down-regulating acute inflammation[85]. Cathepsin G is serine protease with trypsin/chymotrypsin-like specificity responsible for amongst others cleaving complement C3[86], which in turn plays a major role in subduing bacterial infections[87]. Besides previously mentioned activity, both cathepsin G and neutrophil elastase are also capable of inactivating the C5a receptor to modulate neutrophil function[88]. Neutrophil collagenase (MMP8) is a collagenase which can either be activated by an oxidant or by the previously mentioned Cathepsin G[89]. It is responsible for cleaving cytokines and in turn increase their chemotactic activity, CXCL5[90] and LIX[91] are amongst them. MMP8 is mostly active during phagocytosis[92]. Matrix metalloproteinase-9 (MMP-9, Gelatinase B) is the key marker of gelatinase granules, which is capable of cleaving type IV collagen[93]. Besides collagen degradation MMP9 plays a role in the cleavage of several chemokines, amongst others IL-8

which is truncated at the amino terminus resulting into a tenfold more potent chemokine involved in the feedback loop for neutrophil activation and chemotaxis and as such MMP-9 acts as a regulator of response[94]. Azurocidin is predominantly found in azurophil granules[95], it has a strong cytotoxic action which is limited to Gram-negative bacteria, this due to a strong affinity of the very basic N-terminal half for the negatively charged lipopolysaccharides that are unique to the Gram-negative bacterial outer envelope[96]. Arginase-1 (ARG1) is present in both azurophil granules[35, 97] and gelatinase granules[98]. ARG1 converts L-arginine to urea and L-ornithine, which depletes the phagosomal environment of L-arginine inhibiting growth of viruses, parasites and bacteria.

Proteome signature of eosinophils

Eosinophil specific proteins contain some of the hallmark eosinophil proteins such as Eosinophil peroxidase (EPX), which is stored in the secretory granules within eosinophils. EPX induces the oxidation of halide ions particularly bromide into HOBr that is very cytotoxic for microorganisms[99] by e.g. disruption of bacterial cell walls[100]. Proteoglycan 3 (PRG3) also known as pro eosinophil major basic protein 2, is similar in its cytotoxic and cytostimulatory activities to major basic protein (PRG2). It stimulates neutrophil superoxide production and the release of IL8, induces histamine and leukotriene C4 from basophils[101]. Eosinophil cationic protein (ECP) is cytotoxin and helminthotoxin, it exhibits antibacterial activity by inducing membrane depolarization[102, 103]. Arachidonate 15-lipoxygenase (ALOX15), catalyzes the stereo-specific peroxidation of free and esterified polyunsaturated fatty acids generating a spectrum of bioactive lipid mediators. The resulting metabolites including lipoxin, resolving and protectin have been suggested to inhibit, limit and resolve pathogen induced inflammatory responses[104, 105]. Indoleamine 2,3-dioxygenase 1 (IDO1) catalyzes the initial step of the catabolism of tryptophan, which limits pathogens growth by means of tryptophan deprivation[106]. Platelet-activating factor acetylhydrolase 2 (PAFAH2), is suggested to have a similar function compared to PAFAH which inactivates PAF[107], which is a mediator of leukocyte functions, platelet aggregation and degranulation and inflammation[108]. Synaptosomal-associated protein 29 (SNAP29) is a soluble N-ethylmaleimide-sensitve factor-attachment protein receptor (SNARE) this class of proteins is essential for fusion of cellular membranes[109], and this particular SNARE is involved in autophagy which needs to be triggered amongst others in case of intracellular pathogens[110-112]. TBC1 domain family member 5 (TBC1D5) is also involved with autophagy[113], putatively acting as a regulator in trafficking ATG9A-AP2 containing vesicles to autophagic membranes[114]. And thymocyte-expressed molecule involved in selection protein 2 (THEMIS2) has been suggested to constitute a control point in macrophage inflammatory responses by regulating TNF expression downstream of TLR4[115]. Eosinophil-derived neurotoxin (EDN/Rnase2) is missing as a specific protein, it has been identified in all three granuloctyes with an abundance in eosinohils of ~3 fold higher compared to neutrophils and basophils, in line with previously reported literature49. Similarly major

basic protein (MBP/PRG2) is found in all granulocytes, its abundance in eosinophils is ~3.5 times higher compared to the other granulocytes, also in line with previously reported ratios[49].

Proteome signature of basophils

The basophil specific proteins contain only few proteins with known function, most functions of which are inferred by similarity. Tryptase alpha/beta-1 (TPSAB), is known to be a major neutral protease in mast cells, its mouse homologue has been implicated to play a role in innate immunity[116]. N(G),N(G)-dimethylarginine dimethylaminohydrolase 2 (DDAH2) hydrolyzes N(G),N(G)-dimethyl-L-arginine (ADMA) and N(G)-monomethyl-L-arginine (MMA) which act as inhibitors of NOS and thus regulates nitric oxide generation[117]. Synaptotagmin-like protein 3 (SYTL3) may be involved in vesicle tracking by acting as a Rab effector protein (inferred by similarity). Dipeptidyl peptidase 4 (DPP4) is a cell surface glycoprotein receptor, involved in T-cell receptor mediated T-cell activation[118]. Apart from the receptor function DPP4 also acts as a serine exopeptidase regulating several physiological processes by cleaving peptides in the circulation e.g. chemokines and growth factors[119]. V-type proton ATPase 16 kDa proteolipid subunit is involved in forming proton-conducting pores and responsible for acidifying a variety of intracellular components in eukaryotic cells, such as the phagosome. This is required to activate acid-activated host enzymes to digest the internalized pathogen[120]. Hematopoietic prostaglandin D synthase catalyzes the conversion of PGH2 to prostaglandin D2 which is involved in the recruitment of Th2 cells and eosinophils, it is also a powerful inhibitor of platelet aggregation[121]. Plasminogen receptor (KT) is involved in the regulation of the inflammatory response, it regulates monocyte chemotactic migration and the activation of matrix metalloproteinases such as MMP2 and MMP9[122]. Interleukin-3 receptor subunit alpha is the receptor for IL-3. IL-3 has proven to be important for host defence against infection of in particular parasites by expanding the population of haematopoietic effector cells[123].

Concluding Remarks

Here we used high-end proteomics techniques to chart in-depth the proteomes of the neutrophil, eosinophil and basophil granulocytes, providing some of the most comprehensive datasets currently available. Several well-known key components were identified for each of the granulocyte types, and also a large overlap was observed in between their proteomes. Still clear proteome signatures could be extracted from which they can be easily distinguished. The protein signature for neutrophils is mostly comprised of proteins with a very well documented neutrophil related function, which comes as no surprise as it is the best characterized cell type of the three here analyzed. As a consequence most of the specific proteins found can directly be linked to an immunological function. In the proteome signatures of eosinophils and in particular basophils there are less proteins with well-documented roles in immune response. An explanation might be that these cell types are known to synthesize effectors *de novo* upon stimulation. Considering that this study has been performed with resting state

cells, these proteins may either be too low abundant or non-differential when compared to the other cell types. Besides this explanation, our analysis suffers from the fact that many of the eosinophil and basophil signature proteins have no clear annotated function or one inferred by similarity. These proteins may thus still be interesting but require further studies to reveal if that is the case.

The provided proteome resources will aid future studies, for instance wherein these cell-types are challenged with a stimulus, to be able to more clearly define the different roles these cell-types execute in immune response.

Acknowledgements
This work was supported by The Netherlands Organisation for Scientific Research (NWO) *via* the Roadmap Initiative Proteins@Work (project number 184.032.201).

References

1. Duffin R, Leitch AE, Fox S, Haslett C, Rossi AG. Targeting granulocyte apoptosis: Mechanisms, models, and therapies. ***Immunological reviews.*** 2010;236:28-40
2. Geering B, Stoeckle C, Conus S, Simon HU. Living and dying for inflammation: Neutrophils, eosinophils, basophils. ***Trends Immunol***. 2013;34:398-409
3. Farahi N, Singh NR, Heard S, Loutsios C, Summers C, Solanki CK, Solanki K, Balan KK, Ruparelia P, Peters AM, Condliffe AM, Chilvers ER. Use of 111-indium-labeled autologous eosinophils to establish the in vivo kinetics of human eosinophils in healthy subjects. ***Blood***. 2012;120:4068-4071
4. Park YM, Bochner BS. Eosinophil survival and apoptosis in health and disease. ***Allergy, asthma & immunology research***. 2010;2:87-101
5. Pillay J, den Braber I, Vrisekoop N, Kwast LM, de Boer RJ, Borghans JA, Tesselaar K, Koenderman L. In vivo labeling with 2h2o reveals a human neutrophil lifespan of 5.4 days. ***Blood***. 2010;116:625-627
6. Hallett JM, Leitch AE, Riley NA, Duffin R, Haslett C, Rossi AG. Novel pharmacological strategies for driving inflammatory cell apoptosis and enhancing the resolution of inflammation. ***Trends in pharmacological sciences***. 2008;29:250-257
7. Savill JS, Wyllie AH, Henson JE, Walport MJ, Henson PM, Haslett C. Macrophage phagocytosis of aging neutrophils in inflammation. Programmed cell death in the neutrophil leads to its recognition by macrophages. ***The Journal of clinical investigation***. 1989;83:865-875
8. Mayadas TN, Cullere X, Lowell CA. The multifaceted functions of neutrophils. ***Annu Rev Pathol***. 2014;9:181-218
9. Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, Weinrauch Y, Zychlinsky A. Neutrophil extracellular traps kill bacteria. ***Science***. 2004;303:1532-1535
10. McDonald B, Urrutia R, Yipp BG, Jenne CN, Kubes P. Intravascular neutrophil extracellular traps capture bacteria from the bloodstream during sepsis. ***Cell host & microbe***. 2012;12:324-333
11. Yipp BG, Petri B, Salina D, Jenne CN, Scott BN, Zbytnuik LD, Pittman K, Asaduzzaman M, Wu K, Meijndert HC, Malawista SE, de Boisfleury Chevance A, Zhang K, Conly J, Kubes P. Infection-induced netosis is a dynamic process involving neutrophil multitasking in vivo. ***Nature medicine***. 2012;18:1386-1393
12. Amulic B, Cazalet C, Hayes GL, Metzler KD, Zychlinsky A. Neutrophil function: From mechanisms to disease. ***Annual review of immunology***. 2012;30:459-489
13. Kolaczkowska E, Kubes P. Neutrophil recruitment and function in health and inflammation. ***Nat Rev Immunol***. 2013;13:159-175
14. Mantovani A, Cassatella MA, Costantini C, Jaillon S. Neutrophils in the activation and regulation of innate and adaptive immunity. ***Nat Rev Immunol***. 2011;11:519-531
15. Rosenberg HF, Dyer KD, Foster PS. Eosinophils: Changing perspectives in health and disease. ***Nat Rev Immunol***. 2013;13:9-22
16. Fabre V, Beiting DP, Bliss SK, Gebreselassie NG, Gagliardo LF, Lee NA, Lee JJ, Appleton JA. Eosinophil deficiency compromises parasite survival in chronic nematode infection. ***Journal of immunology***. 2009;182:1577-1583

17. Gebreselassie NG, Moorhead AR, Fabre V, Gagliardo LF, Lee NA, Lee JJ, Appleton JA. Eosinophils preserve parasitic nematode larvae by regulating local immunity. ***Journal of immunology***. 2012;188:417-425
18. Jacobsen EA, Ochkur SI, Lee NA, Lee JJ. Eosinophils and asthma. ***Current allergy and asthma reports***. 2007;7:18-26
19. Wegmann M. Targeting eosinophil biology in asthma therapy. ***Am J Respir Cell Mol Biol***. 2011;45:667-674
20. Ortega HG, Liu MC, Pavord ID, Brusselle GG, FitzGerald JM, Chetta A, Humbert M, Katz LE, Keene ON, Yancey SW, Chanez P, Investigators M. Mepolizumab treatment in patients with severe eosinophilic asthma. ***The New England journal of medicine***. 2014;371:1198-1207
21. Haldar P, Brightling CE, Hargadon B, Gupta S, Monteiro W, Sousa A, Marshall RP, Bradding P, Green RH, Wardlaw AJ, Pavord ID. Mepolizumab and exacerbations of refractory eosinophilic asthma. ***The New England journal of medicine***. 2009;360:973-984
22. Pavord ID, Korn S, Howarth P, Bleecker ER, Buhl R, Keene ON, Ortega H, Chanez P. Mepolizumab for severe eosinophilic asthma (dream): A multicentre, double-blind, placebo-controlled trial. ***Lancet***. 2012;380:651-659
23. Lee JJ, Dimina D, Macias MP, Ochkur SI, McGarry MP, O'Neill KR, Protheroe C, Pero R, Nguyen T, Cormier SA, Lenkiewicz E, Colbert D, Rinaldi L, Ackerman SJ, Irvin CG, Lee NA. Defining a link with asthma in mice congenitally deficient in eosinophils. ***Science***. 2004;305:1773-1776
24. Yu C, Cantor AB, Yang H, Browne C, Wells RA, Fujiwara Y, Orkin SH. Targeted deletion of a high-affinity gata-binding site in the gata-1 promoter leads to selective loss of the eosinophil lineage in vivo. ***The Journal of experimental medicine***. 2002;195:1387-1395
25. Neves JS, Perez SA, Spencer LA, Melo RC, Reynolds L, Ghiran I, Mahmudi-Azer S, Odemuyiwa SO, Dvorak AM, Moqbel R, Weller PF. Eosinophil granules function extracellularly as receptor-mediated secretory organelles. ***Proceedings of the National Academy of Sciences of the United States of America***. 2008;105:18478-18483
26. Spencer LA, Melo RC, Perez SA, Bafford SP, Dvorak AM, Weller PF. Cytokine receptor-mediated trafficking of preformed il-4 in eosinophils identifies an innate immune mechanism of cytokine secretion. ***Proceedings of the National Academy of Sciences of the United States of America***. 2006;103:3333-3338
27. Wang HB, Ghiran I, Matthaei K, Weller PF. Airway eosinophils: Allergic inflammation recruited professional antigen-presenting cells. ***Journal of immunology***. 2007;179:7585-7592
28. Spencer LA, Szela CT, Perez SA, Kirchhoffer CL, Neves JS, Radke AL, Weller PF. Human eosinophils constitutively express multiple th1, th2, and immunoregulatory cytokines that are secreted rapidly and differentially. ***Journal of leukocyte biology***. 2009;85:117-123
29. Wang HB, Weller PF. Pivotal advance: Eosinophils mediate early alum adjuvant-elicited b cell priming and igm production. ***Journal of leukocyte biology***. 2008;83:817-821
30. Yang D, Rosenberg HF, Chen Q, Dyer KD, Kurosaka K, Oppenheim JJ. Eosinophil-derived neurotoxin (edn), an antimicrobial protein with chemotactic activities for dendritic cells. ***Blood***. 2003;102:3396-3403

31. Yang D, Chen Q, Su SB, Zhang P, Kurosaka K, Caspi RR, Michalek SM, Rosenberg HF, Zhang N, Oppenheim JJ. Eosinophil-derived neurotoxin acts as an alarmin to activate the tlr2-myd88 signal pathway in dendritic cells and enhances th2 immune responses. ***The Journal of experimental medicine***. 2008;205:79-90
32. Elishmereni M, Alenius HT, Bradding P, Mizrahi S, Shikotra A, Minai-Fleminger Y, Mankuta D, Eliashar R, Zabucchi G, Levi-Schaffer F. Physical interactions between mast cells and eosinophils: A novel mechanism enhancing eosinophil survival in vitro. ***Allergy***. 2011;66:376-385
33. Durrani SR, Montville DJ, Pratt AS, Sahu S, DeVries MK, Rajamanickam V, Gangnon RE, Gill MA, Gern JE, Lemanske RF, Jr., Jackson DJ. Innate immune responses to rhinovirus are reduced by the high-affinity ige receptor in allergic asthmatic children. ***The Journal of allergy and clinical immunology***. 2012;130:489-495
34. Cromheecke JL, Nguyen KT, Huston DP. Emerging role of human basophil biology in health and disease. ***Curr Allergy Asthma Rep***. 2014;14:408
35. Rorvig S, Ostergaard O, Heegaard NH, Borregaard N. Proteome profiling of human neutrophil granule subsets, secretory vesicles, and cell membrane: Correlation with transcriptome profiling of neutrophil precursors. ***Journal of leukocyte biology***. 2013;94:711-721
36. Tomazella GG, daSilva I, Thome CH, Greene LJ, Koehler CJ, Thiede B, Wiker HG, de Souza GA. Analysis of detergent-insoluble and whole cell lysate fractions of resting neutrophils using high-resolution mass spectrometry. ***Journal of proteome research***. 2010;9:2030-2036
37. Zhou JY, Krovvidi RK, Gao Y, Gao H, Petritis BO, De AK, Miller-Graziano CL, Bankey PE, Petyuk VA, Nicora CD, Clauss TR, Moore RJ, Shi T, Brown JN, Kaushal A, Xiao W, Davis RW, Maier RV, Tompkins RG, Qian WJ, Camp DG, 2nd, Smith RD, Inflammation, the Host Response to Injury Large Scale Collaborative Research P. Trauma-associated human neutrophil alterations revealed by comparative proteomics profiling. ***Proteomics Clin Appl***. 2013;7:571-583
38. Wilkerson EM, Johansson MW, Hebert AS, Westphall MS, Mathur SK, Jarjour NN, Schwantes EA, Mosher DF, Coon JJ. The peripheral blood eosinophil proteome. ***Journal of proteome research***. 2016;15:1524-1533
39. Straub C, Pazdrak K, Young TW, Stafford SJ, Wu Z, Wiktorowicz JE, Haag AM, English RD, Soman KV, Kurosky A. Toward the proteome of the human peripheral blood eosinophil. ***Proteomics Clin Appl***. 2009;3:1151-1173
40. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. ***Nat Methods***. 2009;6:359-362
41. Kall L, Canterbury JD, Weston J, Noble WS, MacCoss MJ. Semi-supervised learning for peptide identification from shotgun proteomics datasets. ***Nat Methods***. 2007;4:923-925
42. Vizcaino JA, Cote R, Reisinger F, Barsnes H, Foster JM, Rameseder J, Hermjakob H, Martens L. The proteomics identifications database: 2010 update. ***Nucleic Acids Res***. 2010;38:D736-742
43. Scholten A, Mohammed S, Low TY, Zanivan S, van Veen TA, Delanghe B, Heck AJ. In-depth quantitative cardiac proteomics combining electron transfer dissociation and the metalloendopeptidase lys-n with the silac mouse. ***Mol Cell Proteomics***. 2011;10:O111 008474

44. Kumar L, M EF. Mfuzz: A software package for soft clustering of microarray data. ***Bioinformation***. 2007;2:5-7
45. Schwammle V, Jensen ON. A simple and fast method to determine the parameters for fuzzy c-means cluster analysis. ***Bioinformatics***. 2010;26:2841-2848
46. Risso A. Leukocyte antimicrobial peptides: Multifunctional effector molecules of innate immunity. ***Journal of leukocyte biology***. 2000;68:785-792
47. Atallah M, Krispin A, Trahtemberg U, Ben-Hamron S, Grau A, Verbovetski I, Mevorach D. Constitutive neutrophil apoptosis: Regulation by cell concentration via s100 a8/9 and the mek-erk pathway. ***PloS one***. 2012;7:e29333
48. Borregaard N, Sorensen OE, Theilgaard-Monch K. Neutrophil granules: A library of innate immunity proteins. ***Trends Immunol***. 2007;28:340-345
49. Abu-Ghazaleh RI, Dunnette SL, Loegering DA, Checkel JL, Kita H, Thomas LL, Gleich GJ. Eosinophil granule proteins in peripheral blood granulocytes. ***Journal of leukocyte biology***. 1992;52:611-618
50. Acharya KR, Ackerman SJ. Eosinophil granule proteins: Form and function. ***J Biol Chem***. 2014;289:17406-17415
51. Ueki S, Melo RC, Ghiran I, Spencer LA, Dvorak AM, Weller PF. Eosinophil extracellular DNA trap cell death mediates lytic release of free secretion-competent eosinophil granules in humans. ***Blood***. 2013;121:2074-2083
52. Morshed M, Hlushchuk R, Simon D, Walls AF, Obata-Ninomiya K, Karasuyama H, Djonov V, Eggel A, Kaufmann T, Simon HU, Yousefi S. Nadph oxidase-independent formation of extracellular DNA traps by basophils. ***Journal of immunology***. 2014;192:5314-5323
53. Urban CF, Ermert D, Schmid M, Abu-Abed U, Goosmann C, Nacken W, Brinkmann V, Jungblut PR, Zychlinsky A. Neutrophil extracellular traps contain calprotectin, a cytosolic protein complex involved in host defense against candida albicans. ***PLoS Pathog***. 2009;5:e1000639
54. Zhang LT, Yao YM, Lu JQ, Yan XJ, Yu Y, Sheng ZY. Recombinant bactericidal/permeability-increasing protein inhibits endotoxin-induced high-mobility group box 1 protein gene expression in sepsis. ***Shock***. 2008;29:278-284
55. Curran CS, Demick KP, Mansfield JM. Lactoferrin activates macrophages via tlr4-dependent and -independent signaling pathways. ***Cell Immunol***. 2006;242:23-30
56. Jaillon S, Peri G, Delneste Y, Fremaux I, Doni A, Moalli F, Garlanda C, Romani L, Gascan H, Bellocchio S, Bozza S, Cassatella MA, Jeannin P, Mantovani A. The humoral pattern recognition receptor ptx3 is stored in neutrophil granules and localizes in extracellular traps. ***The Journal of experimental medicine***. 2007;204:793-804
57. Hirose T, Hamaguchi S, Matsumoto N, Irisawa T, Seki M, Tasaki O, Hosotsubo H, Yamamoto N, Yamamoto K, Akeda Y, Oishi K, Tomono K, Shimazu T. Presence of neutrophil extracellular traps and citrullinated histone h3 in the bloodstream of critically ill patients. ***PloS one***. 2014;9:e111755
58. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. Gorilla: A tool for discovery and visualization of enriched go terms in ranked gene lists. ***BMC Bioinformatics***. 2009;10:48
59. Belaaouaj A, Kim KS, Shapiro SD. Degradation of outer membrane protein a in escherichia coli killing by neutrophil elastase. ***Science***. 2000;289:1185-1188

60. van Kessel KP, Bestebroer J, van Strijp JA. Neutrophil-mediated phagocytosis of staphylococcus aureus. ***Front Immunol***. 2014;5:467
61. Nizet V, Ohtake T, Lauth X, Trowbridge J, Rudisill J, Dorschner RA, Pestonjamasp V, Piraino J, Huttner K, Gallo RL. Innate antimicrobial peptide protects the skin from invasive bacterial infection. ***Nature***. 2001;414:454-457
62. Ganz T. Defensins: Antimicrobial peptides of innate immunity. ***Nat Rev Immunol***. 2003;3:710-720
63. Okamoto T, Tanida T, Wei B, Ueta E, Yamamoto T, Osaki T. Regulation of fungal infection by a combination of amphotericin b and peptide 2, a lactoferrin peptide that activates neutrophils. ***Clin Diagn Lab Immunol***. 2004;11:1111-1119
64. Tkalcevic J, Novelli M, Phylactides M, Iredale JP, Segal AW, Roes J. Impaired immunity and enhanced resistance to endotoxin in the absence of neutrophil elastase and cathepsin g. ***Immunity***. 2000;12:201-210
65. Greenlee MC, Sullivan SA, Bohlson SS. Cd93 and related family members: Their role in innate immunity. ***Curr Drug Targets***. 2008;9:130-138
66. Wu D, LaRosa GJ, Simon MI. G protein-coupled signal transduction pathways for interleukin-8. ***Science***. 1993;261:101-103
67. Gaudet P, Dessimoz C. Gene ontology: Pitfalls, biases, and remedies. ***Methods in molecular biology***. 2017;1446:189-205
68. Rankin SM, Conroy DM, Williams TJ. Eotaxin and eosinophil recruitment: Implications for human disease. ***Mol Med Today***. 2000;6:20-27
69. Uguccioni M, Mackay CR, Ochensberger B, Loetscher P, Rhis S, LaRosa GJ, Rao P, Ponath PD, Baggiolini M, Dahinden CA. High expression of the chemokine receptor ccr3 in human blood basophils. Role in activation by eotaxin, mcp-4, and other chemokines. ***The Journal of clinical investigation***. 1997;100:1137-1143
70. Forssmann U, Uguccioni M, Loetscher P, Dahinden CA, Langen H, Thelen M, Baggiolini M. Eotaxin-2, a novel cc chemokine that is selective for the chemokine receptor ccr3, and acts like eotaxin on human eosinophil and basophil leukocytes. ***The Journal of experimental medicine***. 1997;185:2171-2176
71. Klebanoff SJ, Kettle AJ, Rosen H, Winterbourn CC, Nauseef WM. Myeloperoxidase: A front-line defender against phagocytosed microorganisms. ***Journal of leukocyte biology***. 2013;93:185-198
72. Ericksen B, Wu Z, Lu W, Lehrer RI. Antibacterial activity and specificity of the six human {alpha}-defensins. ***Antimicrob Agents Chemother***. 2005;49:269-275
73. Zou G, de Leeuw E, Li C, Pazgier M, Li C, Zeng P, Lu WY, Lubkowski J, Lu W. Toward understanding the cationicity of defensins. Arg and lys versus their noncoded analogs. ***J Biol Chem***. 2007;282:19653-19665
74. Tschopp CM, Spiegl N, Didichenko S, Lutmann W, Julius P, Virchow JC, Hack CE, Dahinden CA. Granzyme b, a novel mediator of allergic inflammation: Its induction and release in blood basophils and human asthma. ***Blood***. 2006;108:2290-2299
75. Reitamo S, Klockars M, Adinolfi M, Osserman EF. Human lysozyme (origin and distribution in health and disease). ***Ric Clin Lab***. 1978;8:211-231

76. Voganatsi A, Panyutich A, Miyasaki KT, Murthy RK. Mechanism of extracellular release of human neutrophil calprotectin complex. ***Journal of leukocyte biology***. 2001;70:130-134
77. Ryckman C, Vandal K, Rouleau P, Talbot M, Tessier PA. Proinflammatory activities of s100: Proteins s100a8, s100a9, and s100a8/a9 induce neutrophil chemotaxis and adhesion. ***Journal of immunology***. 2003;170:3233-3242
78. Thomas RJ, McLeish G, McDonald IA. Electroejaculation of the paraplegic male followed by pregnancy. ***Med J Aust***. 1975;2:789-789
79. Bellamy W, Takase M, Yamauchi K, Wakabayashi H, Kawase K, Tomita M. Identification of the bactericidal domain of lactoferrin. ***Biochim Biophys Acta***. 1992;1121:130-136
80. Le Cabec V, Maridonneau-Parini I. Annexin 3 is associated with cytoplasmic granules in neutrophils and monocytes and translocates to the plasma membrane in activated cells. ***The Biochemical journal***. 1994;303 (Pt 2):481-487
81. Larsson M, Majeed M, Ernst JD, Magnusson KE, Stendahl O, Forsum U. Role of annexins in endocytosis of antigens in immature human dendritic cells. ***Immunology***. 1997;92:501-511
82. Murakami M, Lopez-Garcia B, Braff M, Dorschner RA, Gallo RL. Postsecretory processing generates multiple cathelicidins for enhanced topical antimicrobial defense. ***Journal of immunology***. 2004;172:3070-3077
83. Goetz DH, Holmes MA, Borregaard N, Bluhm ME, Raymond KN, Strong RK. The neutrophil lipocalin ngal is a bacteriostatic agent that interferes with siderophore-mediated iron acquisition. ***Mol Cell***. 2002;10:1033-1043
84. Flo TH, Smith KD, Sato S, Rodriguez DJ, Holmes MA, Strong RK, Akira S, Aderem A. Lipocalin 2 mediates an innate immune response to bacterial infection by sequestrating iron. ***Nature***. 2004;432:917-921
85. Tralau T, Meyer-Hoffert U, Schroder JM, Wiedow O. Human leukocyte elastase and cathepsin g are specific inhibitors of c5a-dependent neutrophil enzyme release and chemotaxis. ***Exp Dermatol***. 2004;13:316-325
86. Yang D, Chertov O, Oppenheim JJ. The role of mammalian antimicrobial peptides and proteins in awakening of innate host defenses and adaptive immunity. ***Cell Mol Life Sci***. 2001;58:978-989
87. Lachmann P. Genetics of the complement system. ***J Med Genet***. 1975;12:372-377
88. van den Berg CW, Tambourgi DV, Clark HW, Hoong SJ, Spiller OB, McGreal EP. Mechanism of neutrophil dysfunction: Neutrophil serine proteases cleave and inactivate the c5a receptor. ***Journal of immunology***. 2014;192:1787-1795
89. Capodici C, Berg RA. Neutrophil collagenase activation: The role of oxidants and cathepsin g. ***Agents Actions***. 1991;34:8-10
90. Tester AM, Cox JH, Connor AR, Starr AE, Dean RA, Puente XS, Lopez-Otin C, Overall CM. Lps responsiveness and neutrophil chemotaxis in vivo require pmn mmp-8 activity. ***PloS one***. 2007;2:e312
91. Balbin M, Fueyo A, Tester AM, Pendas AM, Pitiot AS, Astudillo A, Overall CM, Shapiro SD, Lopez-Otin C. Loss of collagenase-2 confers increased skin tumor susceptibility to male mice. ***Nature genetics***. 2003;35:252-257

92. Ohlsson K, Olsson I. The extracellular release of granulocyte collagenase and elastase during phagocytosis and inflammatory processes. ***Scand J Haematol***. 1977;19:145-152
93. Devarajan P, Johnston JJ, Ginsberg SS, Van Wart HE, Berliner N. Structure and expression of neutrophil gelatinase cdna. Identity with type iv collagenase from ht1080 cells. ***J Biol Chem***. 1992;267:25228-25232
94. Opdenakker G, Van den Steen PE, Dubois B, Nelissen I, Van Coillie E, Masure S, Proost P, Van Damme J. Gelatinase b functions as regulator and effector in leukocyte biology. ***Journal of leukocyte biology***. 2001;69:851-859
95. Tapper H, Karlsson A, Morgelin M, Flodgaard H, Herwald H. Secretion of heparin-binding protein from human neutrophils is determined by its localization in azurophilic granules and secretory vesicles. ***Blood***. 2002;99:1785-1793
96. McCabe D, Cukierman T, Gabay JE. Basic residues in azurocidin/hbp contribute to both heparin binding and antimicrobial activity. ***J Biol Chem***. 2002;277:27477-27488
97. Munder M, Mollinedo F, Calafat J, Canchado J, Gil-Lamaignere C, Fuentes JM, Luckner C, Doschko G, Soler G, Eichmann K, Muller FM, Ho AD, Goerner M, Modolell M. Arginase i is constitutively expressed in human granulocytes and participates in fungicidal activity. ***Blood***. 2005;105:2549-2556
98. Jacobsen LC, Theilgaard-Monch K, Christensen EI, Borregaard N. Arginase 1 is expressed in myelocytes/metamyelocytes and localized in gelatinase granules of human neutrophils. ***Blood***. 2007;109:3084-3087
99. van Dalen CJ, Kettle AJ. Substrates and products of eosinophil peroxidase. ***The Biochemical journal***. 2001;358:233-239
100. Borelli V, Vita F, Shankar S, Soranzo MR, Banfi E, Scialino G, Brochetta C, Zabucchi G. Human eosinophil peroxidase induces surface alteration, killing, and lysis of mycobacterium tuberculosis. ***Infect Immun***. 2003;71:605-613
101. Plager DA, Loegering DA, Weiler DA, Checkel JL, Wagner JM, Clarke NJ, Naylor S, Page SM, Thomas LL, Akerblom I, Cocks B, Stuart S, Gleich GJ. A novel and highly divergent homolog of human eosinophil granule major basic protein. ***J Biol Chem***. 1999;274:14464-14473
102. Torrent M, de la Torre BG, Nogues VM, Andreu D, Boix E. Bactericidal and membrane disruption activities of the eosinophil cationic protein are largely retained in an n-terminal fragment. ***The Biochemical journal***. 2009;421:425-434
103. Tollin M, Bergman P, Svenberg T, Jornvall H, Gudmundsson GH, Agerberth B. Antimicrobial peptides in the first line defence of human colon mucosa. ***Peptides***. 2003;24:523-530
104. Romano M, Cianci E, Simiele F, Recchiuti A. Lipoxins and aspirin-triggered lipoxins in resolution of inflammation. ***Eur J Pharmacol***. 2015;760:49-63
105. Russell CD, Schwarze J. The role of pro-resolution lipid mediators in infectious disease. ***Immunology***. 2014;141:166-173
106. van Baren N, Van den Eynde BJ. Tryptophan-degrading enzymes in tumoral immune resistance. ***Front Immunol***. 2015;6:34

107. Hattori K, Adachi H, Matsuzawa A, Yamamoto K, Tsujimoto M, Aoki J, Hattori M, Arai H, Inoue K. Cdna cloning and expression of intracellular platelet-activating factor (paf) acetylhydrolase ii. Its homology with plasma paf acetylhydrolase. *J Biol Chem*. 1996;271:33032-33038
108. Chao W, Olson MS. Platelet-activating factor: Receptors and signal transduction. *The Biochemical journal*. 1993;292 (Pt 3):617-629
109. Itakura E, Kishi-Itakura C, Mizushima N. The hairpin-type tail-anchored snare syntaxin 17 targets to autophagosomes for fusion with endosomes/lysosomes. *Cell*. 2012;151:1256-1269
110. Levine B, Deretic V. Unveiling the roles of autophagy in innate and adaptive immunity. *Nat Rev Immunol*. 2007;7:767-777
111. Schmid D, Munz C. Innate and adaptive immunity through autophagy. *Immunity*. 2007;27:11-21
112. Klionsky DJ. Autophagy: From phenomenology to molecular understanding in less than a decade. *Nat Rev Mol Cell Biol*. 2007;8:931-937
113. Popovic D, Akutsu M, Novak I, Harper JW, Behrends C, Dikic I. Rab gtpase-activating proteins in autophagy: Regulation of endocytic and autophagy pathways by direct binding to human atg8 modifiers. *Mol Cell Biol*. 2012;32:1733-1744
114. Popovic D, Dikic I. Tbc1d5 and the ap2 complex regulate atg9 trafficking and initiation of autophagy. *EMBO Rep*. 2014;15:392-401
115. Peirce MJ, Brook M, Morrice N, Snelgrove R, Begum S, Lanfrancotti A, Notley C, Hussell T, Cope AP, Wait R. Themis2/icb1 is a signaling scaffold that selectively regulates macrophage toll-like receptor signaling and cytokine production. *PloS one*. 2010;5:e11465
116. Jackson NE, Wang HW, Bryant KJ, McNeil HP, Husain A, Liu K, Tedla N, Thomas PS, King GC, Hettiaratchi A, Cairns J, Hunt JE. Alternate mrna splicing in multiple human tryptase genes is predicted to regulate tetramer formation. *J Biol Chem*. 2008;283:34178-34187
117. Leiper JM, Santa Maria J, Chubb A, MacAllister RJ, Charles IG, Whitley GS, Vallance P. Identification of two human dimethylarginine dimethylaminohydrolases with distinct tissue distributions and homology with microbial arginine deiminases. *The Biochemical journal*. 1999;343 Pt 1:209-214
118. Ikushima H, Munakata Y, Ishii T, Iwata S, Terashima M, Tanaka H, Schlossman SF, Morimoto C. Internalization of cd26 by mannose 6-phosphate/insulin-like growth factor ii receptor contributes to t cell activation. *Proceedings of the National Academy of Sciences of the United States of America*. 2000;97:8439-8444
119. Durinx C, Lambeir AM, Bosmans E, Falmagne JB, Berghmans R, Haemers A, Scharpe S, De Meester I. Molecular characterization of dipeptidyl peptidase activity in serum: Soluble cd26/dipeptidyl peptidase iv is responsible for the release of x-pro dipeptides. *Eur J Biochem*. 2000;267:5608-5613
120. Ip WK, Sokolovska A, Charriere GM, Boyer L, Dejardin S, Cappillino MP, Yantosca LM, Takahashi K, Moore KJ, Lacy-Hulbert A, Stuart LM. Phagocytosis and phagosome acidification are required for pathogen processing and myd88-dependent responses to staphylococcus aureus. *Journal of immunology*. 2010;184:7071-7081
121. Bundy GL, Morton DR, Peterson DC, Nishizawa EE, Miller WL. Synthesis and platelet aggregation inhibiting activity of prostaglandin d analogues. *J Med Chem*. 1983;26:790-799

122. Lighvani S, Baik N, Diggs JE, Khaldoyanidi S, Parmer RJ, Miles LA. Regulation of macrophage migration by a novel plasminogen receptor plg-r kt. ***Blood***. 2011;118:5622-5630

123. Lantz CS, Boesiger J, Song CH, Mach N, Kobayashi T, Mulligan RC, Nawa Y, Dranoff G, Galli SJ. Role for interleukin-3 in mast-cell and basophil development and in immunity to parasites. ***Nature***. 1998;392:90-93`

5.
Outlook

Future perspectives on immune cell analysis

Over the course of the last decades the "omics" technologies have matured substantially. Genomics and transcriptomics have come a long way with respect to the first micro-arrays, now followed up by next generation sequencing, which drastically has increased sensitivity, specificity, speed and ease of use to elucidate the genome and transcriptome. Also in the field of metabolomics major advancements have been made in terms of identification and quantification by either MS or NMR or a combination of both these analytical technologies. The omics approach which has been the focus of this thesis, proteomics, has too transformed from a laborious endeavor to identify and quantify a mere protein in a significant amount of time to elucidating the full proteome in a matter of hours. In order to fully understand and appreciate the intricacies of something as complex as a biological system, all omics strategies need to be integrated in a systems biology approach[1, 2] as applying just a single one of those strategies may not be enough[3]. I foresee that such multi-omics approaches may represent the standard for any type of research question concerning a biological system in the future.

But that is still beyond the scope of this thesis, that heavily focuses on proteomics. Recent advancements in terms of sample preparation[4], peptide separation[5, 6] and MS instrumentation[7, 8] have made MS-based proteomics the method of choice regarding the interrogation of biological systems at the protein level both qualitatively and quantitatively, including the analysis of post-translational modifications and protein-protein interactions[9, 10]. These advances have led to the point that a proteome up to a depth of up to 10 000 proteins in a mammalian system is attainable, which is considered to be comprehensible[11].

Currently a lot is known about circulating immune cells but there are also still large gaps in our knowledge that need to be filled in order to further understand the mechanisms involved in either subduing a pathogen directly or in sequestering other immune cells to aid in the process. At the genome level an effort is underway to chart the gene expression and its regulation across the immune system in the form of "The Immunological Genome project"[12]. Projects that map surface molecules including proteins are underway, but a similar project that focuses on the proteome level using aforementioned capabilities of the techniques available is so far missing. With the current technologies available in terms of isolating cells at a purity much higher than previously possible a comprehensive proteome for every type of circulating immune cell could be achieved. Several immune cells are able to secrete their granular content, which could be accurately determined as we demonstrated using reversed releasate profiling. Neutrophils are known to have different types of granules of which the release is triggered by different factors[13], by subjecting these cells to different stimuli it is possible to determine what specifically triggers the release of specific proteins from these granules. Besides a fundamental biological interest in the released content of circulating cells there are also clinical applications. Several platelet function disorders are known involving secretion, for instance abnormalities in the granule secretory mechanism where normal

granules fail to release their content upon activation. By employing a phospho-proteomics approach the molecular mechanisms leading to this defect can possibly be further studied mechanistically.

In the previous paragraph I briefly touched the subject of clinical applications, which also featured in the first chapter of this thesis. Biomarker research has transformed greatly in the -omics era and has come with great promise in terms of future clinical applications. However, it has yet to fully deliver on that very promise. For proteomics a vast amount of data generated is typically generated, making it difficult to convert this huge datasets to clinically valuable information, hampering the verification novel putative protein biomarkers. To start with the identification, the dynamic range of the cells of interest comes into play as it likely exceeds the dynamic range that can adequately be handled by current mass spectrometry equipment. To circumvent this problem a solution is already in place when it comes to handling serum as a potential source for biomarkers, where the most abundant proteins can be depleted from the sample[14, 15]. This will need to be tailored to the cell-type of interest. Another important aspect is the verification, a typical biomarker discovery experiment will likely yield many candidates, and especially now that the depth of information has increased. These putative biomarkers need to be verified using orthogonal techniques, such as ELISA, which is also quite laborious in nature. To reduce the sheer number of candidates for verification interfacing data from the biomarker experiment with potential contaminating proteomes is of utmost importance, as it will eliminate false-positives at a much earlier stage, as clearly demonstrated in this thesis.

Still, with all modules of a proteomics workflow becoming more sensitive, more specific and faster it is just a measure of time, before proteomics will deliver valuable new biomarkers, not only for disease but also for general wellness[16]. The era wherein a visitor to a hospital is asked for a drop of blood to rapidly measure the plasma proteome is on the horizon[17].

References

1. Kitano H. Systembiology: A brief overview. ***Science***. 2002;295:1662-1664
2. Munoz J, Low TY, Kok YJ, Chin A, Frese CK, Ding V, Choo A, Heck AJ. The quantitative proteomes of human-induced pluripotent stem cells and embryonic stem cells. ***Mol Syst Biol***. 2011;7:550
3. Gygi SP, Rochon Y, Franza BR, Aebersold R. Correlation between protein and mrna abundance in yeast. ***Mol Cell Biol***. 1999;19:1720-1730
4. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. ***Nature methods***. 2009;6:359-362
5. Di Palma S, Hennrich ML, Heck AJ, Mohammed S. Recent advances in peptide separation by multidimensional liquid chromatography for proteome analysis. ***J Proteomics***. 2012;75:3791-3813
6. Kocher T, Pichler P, Mazanek M, Swart R, Mechtler K. Altered mascot search results by changing the m/z range of ms/ms spectra: Analysis and potential applications. ***Anal Bioanal Chem***. 2011;400:2339-2347
7. Michalski A, Damoc E, Lange O, Denisov E, Nolting D, Muller M, Viner R, Schwartz J, Remes P, Belford M, Dunyach JJ, Cox J, Horning S, Mann M, Makarov A. Ultra high resolution linear ion trap orbitrap mass spectrometer (orbitrap elite) facilitates top down lc ms/ms and versatile peptide fragmentation modes. ***Mol Cell Proteomics***. 2012;11:O111 013698
8. Michalski A, Damoc E, Hauschild JP, Lange O, Wieghaus A, Makarov A, Nagaraj N, Cox J, Mann M, Horning S. Mass spectrometry-based proteomics using q exactive, a high-performance benchtop quadrupole orbitrap mass spectrometer. ***Mol Cell Proteomics***. 2011;10:M111 011015
9. Yates JR, 3rd, Mohammed S, Heck AJ. Phosphoproteomics. ***Anal Chem***. 2014;86:1313
10. Hebert AS, Richards AL, Bailey DJ, Ulbrich A, Coughlin EE, Westphall MS, Coon JJ. The one hour yeast proteome. ***Mol Cell Proteomics***. 2014;13:339-347
11. Mann M, Kulak NA, Nagaraj N, Cox J. The coming age of complete, accurate, and ubiquitous proteomes. ***Mol Cell***. 2013;49:583-590
12. Benoist C, Lanier L, Merad M, Mathis D, Immunological Genome P. Consortium biology in immunology: The perspective from the immunological genome project. ***Nat Rev Immunol***. 2012;12:734-740
13. Rorvig S, Ostergaard O, Heegaard NH, Borregaard N. Proteome profiling of human neutrophil granule subsets, secretory vesicles, and cell membrane: Correlation with transcriptome profiling of neutrophil precursors. ***J Leukoc Biol***. 2013;94:711-721
14. Patel BB, Barrero CA, Braverman A, Kim PD, Jones KA, Chen DE, Bowler RP, Merali S, Kelsen SG, Yeung AT. Assessment of two immunodepletion methods: Off-target effects and variations in immunodepletion efficiency may confound plasma proteomics. ***Journal of proteome research***. 2012;11:5947-5958
15. Ringrose JH, van Solinge WW, Mohammed S, O'Flaherty MC, van Wijk R, Heck AJ, Slijper M. Highly efficient depletion strategy for the two most abundant erythrocyte soluble proteins improves proteome coverage dramatically. ***Journal of proteome research***. 2008;7:3060-3063
16. Hood L, Friend SH. Predictive, personalized, preventive, participatory (p4) cancer medicine. ***Nature reviews. Clinical oncology***. 2011;8:184-187
17. Geyer PE, Kulak NA, Pichler G, Holdt LM, Teupser D, Mann M. Plasma proteome profiling to assess human health and disease. ***Cell systems***. 2016;2:185-195

6.

Summary
Samenvatting
Curriculum Vitae
List of publications
Acknowledgements

Summary

This thesis describes both the development and the implementation of novel strategies of mass spectrometry based proteomics used to characterize and investigate human primary circulating immune cells. The studies were performed to gain better insights into protein constituents of these cells and to explore the option of using these cells for obtaining clinically relevant biomarkers.

In **Chapter 1**, a general introduction is given, separated into 2 parts. The first part describes in brief the biology of several different circulating cells; including platelets, and neutrophil, eosinophil and basophil granulocytes. The second part describes the different experimental modules used in proteomics, i.e. sample preparation, reducing complexity by separation, mass spectrometry and quantitative proteomics.

In **Chapter 2**, our aim was to find potential biomarkers in circulating cells for coronary artery disease. Most protein based biomarker studies target easily accessible patient material such as plasma and urine, but these approaches are substantially hampered due to dynamic range issues in protein abundance. The circulating cells should in theory hold disease state specific biochemical information with a more favorable dynamic range. The protein content of isolated pooled platelets of small cohorts of carefully matched patients, suffering from coronary artery disease, and healthy controls were first digested, where after the resulting peptides were differentially labeled using stable isotope dimethyl labeling, separated by SCX and subsequently identified and quantified by LC-MS analysis. This yielded a total of 2440 quantified proteins of which 62 proteins initially were revealed as significantly different ($p<0.0001$). However upon closer inspection unavoidable and variable contamination of plasma, erythrocyte and leukocyte components became apparent, which could be traced back to the individual patient level. In order to deduce which protein expression change is related to disease, we present that correlating the quantitative patient/control study with reported quantitative proteomes of plasma, erythrocyte and leukocytes allows for filtering of false positive biomarkers into the verification phase. Screening for these contaminating proteins from different matrices aids in the use of circulating cells as a source for biomarkers. Overall our study reveals that uniform sample purification protocols are essential for clinical biomarker studies.

In **Chapter 3**, a novel method for identifying the platelet releasate is introduced. In contrast to previously reported methods where isolated platelets granules were isolated and analyzed or where cells are stimulated and the secreted content is extracted and analyzed, we introduce a reversed proteomics approach to unambiguously and quantitatively determine the platelets released from activated platelets against a resting platelet background. In short the proteomes of both mock and fully stimulated platelets are subjected to analysis after the released content is removed. There after the protein content of the samples are

digested, the resulting peptides differentially labeled using stable isotope dimethyl labeling, and separated using SCX followed by LC-MS analysis for identification and quantification. By using this strategy we were able to quantitatively discriminate the released proteins from uncontrolled lysis products. Monitoring the copy numbers of roughly 4500 platelet proteins it became apparent that following a full stimulation using thrombin and collagen only 124 proteins (<3% of the proteome) were significantly released ($p<0.05$). The released proteins span a concentration range of at least 5 orders of magnitude. Secretion of several of these proteins could be confirmed by ELISA analysis. The released proteins were highly enriched in exhibiting the known secretion tags. Known high abundant factors such as Thrombospondin, von Willebrand factor and Platelet factor 4 were abundant in the reaseate. Besides these well-known secreted proteins many novel low abundant proteins could be identified.

In **Chapter 4**, the proteomes of neutrophil, eosinophil and basophil granulocytes are quantitatively characterized and compared. At present no comprehensive proteomes are available for these 3 granulocytes that are to some extent similar but have diverging roles as well. For each of these cell types a comprehensive proteome was identified and quantified, exposing many novel specific proteins for these three granulocyte types. A first step was taken to find the similarities between neutrophils, eosinophils and basophils, as well as to expose the proteins that may be involved in their diverging roles.

Samenvatting

Wij als mens worden constant blootgesteld aan organismen die geïnhaleerd of ingeslikt worden maar ook organismen die op onze slijmvliezen of huid leven. Of deze organismen schadelijk zijn voor ons is afhankelijk van in hoeverre het organisme pathogeen is door de virulente factoren die deze tot zijn beschikking heeft maar ook van de integriteit van onze eigen afweer. Ons immuunsysteem is een uitgebreid netwerk van lymfoide organen, cellen en effectoren in de slijmvliezen en circulatie. In dit proefschrift worden de circulerende cellen nader onderzocht. Alhoewel er al ontzettend veel informatie beschikbaar is word er nog met regelmaat nieuwe informatie met betrekking tot de functie van deze cellen ontdekt. De afgelopen jaren decennia is dit in een stroomversnelling terecht gekomen omdat er op verschillende manieren op -omics niveau: genomics (DNA), transcriptomics (mRNA) en proteomics (eiwit) naar deze cellen gekeken kan worden. Elk van deze niveaus geeft een dwarsdoorsnede van wat er met een cel gebeurd op een bepaald moment van tijd, dit wil echt niet zeggen dat deze strategieën 1 op 1 met elkaar vergelijkbaar zijn. Je zou dit heel versimpeld kunnen vergelijken met het bouwen van een huis, waarbij de blauwdruk het DNA is, deze blauwdruk ligt veilig bij de architect (nucleus). Vervolgens moet de relevante informatie van de blauwdruk gekopieerd worden en naar de bouwplaats gebracht worden (transcriptie), deze kopie is het mRNA. Er zijn ook werkers nodig om bakstenen en andere materialen naar de bouwplaats te brengen (tRNA). Deze bakstenen zijn de aminozuren. Om de bakstenen aan elkaar te kunnen zetten is er cement nodig (energie om aminozuren te koppelen tot een eiwit). Om van de bakstenen muren en in het verlengde het huis te kunnen maken dient het bouwplan vervolgens gevolgd te worden (translatie van DNA naar mRNA). Logischerwijs kunnen bakstenen anders geordend worden om zo tot verschillende huizen te komen, het zelfde is van toepassing op aminozuren die op verschillende manieren geordend kunnen worden om verschillende eiwitten te creëren. Alhoewel de blauwdruk enorm veel informatie verschaft over hoe een huis er mogelijk uit gaat zien, bij elke stap van blauwdruk naar huis kunnen veranderingen optreden en dit is zeer zeker ook het geval met de stappen die er nodig zijn van DNA naar eiwit. Het DNA is dus informatief maar reflecteert niet noodzakelijkerwijs wat er zich op het eiwit niveau afspeelt. Om echt te kunnen kijken naar wat er zich in een cel afspeelt is het van belang dat er gekeken naar de moleculen die het werk verzette: de eiwitten en dat is dan ook de focus van dit proefschrift: proteomics.

Dit proefschrift beschrijft de ontwikkeling en implementatie van nieuwe op massa spectrometrie gebaseerde proteomics strategieën voor de karakterisering en het onderzoeken van humane primaire circulerende immuuncellen. Deze studies zijn uitgevoerd om een beter inzicht te krijgen in de eiwitten die zich in deze cellen bevinden en om te onderzoek of deze cellen gebruikt kunnen worden om klinisch relevante biomarkers te verkrijgen.

In **hoofdstuk 1** wordt een algemene inleiding gegeven. Het eerste gedeelte van de introductie beschrijft in het kort de biologie van de verschillende circulerende immuuncellen die voor de

studies in dit proefschrift zijn onderzocht, dit zijn: bloedplaatjes en neutrofiele, eosinofiele en basofiele granulocyten. Het tweede gedeelte van de inleiding beschrijft de experimentele stappen die essentieel zijn in een proteomics experiment, zoals: monster voorbewerking, chromatografie, massaspectrometrie en de mogelijkheden voor kwantitatieve proteomics.

In **hoofdstuk 2** was ons doel het vinden van potentiële biomarkers in circulerende cellen voor ischemische hartklachten. De meeste op eiwit gebaseerde biomarker studies richten zich op gemakkelijk te verkrijgen patiënten materiaal zoals plasma en urine. Het grote nadeel van dit materiaal is het grote dynamische bereik van de concentraties van de verschillende eiwitten. De circulerende cellen zouden in theorie dezelfde ziekte specifieke biochemische informatie moeten hebben maar met een veel gunstiger dynamisch bereik. Voor deze studie is een cohort van zorgvuldig geselecteerde patiënten met ischemische hartklachten en een cohort van gezonde donors onderzocht en met elkaar vergeleken. Voor beide cohorten werden de bloedplaatjes geïsoleerd, gedigesteerd, gelabeld met stabiele isotoop dimethyl labeling alvorens er scheiding plaats vond middels SCX waarna de identificatie en kwantificatie uitgevoerd werd met LC-MS analyse. In totaal werden er 2440 eiwitten gekwantificeerd waarvan er in eerste instantie 62 eiwitten als statistisch significant werden aangeduid. Toen deze 62 eiwitten beter bekeken werden bleek dat een deel hiervan onvermijdbare en variabele contaminatie aan plasma, rode en witte bloedcellen bevatte wat terug getraceerd kon worden tot individuele patiënten. Om te kunnen deduceren welk verschil in eiwit expressie gerelateerd is aan het een ziektebeeld stellen wij voor om de patiënt/controle studie te correleren aan de kwantitatieve proteoom van de mogelijke contaminanten om op deze manier het filteren vals positieve biomarkers mogelijk te maken voor de verificatie fase van het biomarker onderzoek. Tevens is een uniform sample opzuiverings protocol essentieel voor klinische biomarker studies.

In **hoofdstuk 3** wordt een nieuwe methode geïntroduceerd om de eiwitten in het plaatjes releasaat te identificeren. In eerder beschikbare methodes worden of de granules geïsoleerd uit bloedplaatjes en geanalyseerd of worden de cellen gestimuleerd en wordt de uitgescheiden factoren geëxtraheerd en geanalyseerd. Wij introduceren een omgekeerde proteomics methode om zonder twijfel en kwantitatief te kunnen determineren welke eiwitten door een geactiveerd bloedplaatje wordt uitgescheiden tegen de achtergrond van een bloedplaatje in rust. In het kort komt het er op neer dat het proteoom van bloedplaatjes in rust en het proteoom van volledig gestimuleerde bloedplaatjes worden geanalyseerd nadat de uitgescheiden factoren zijn verwijderd. Hierna worden de eiwitten die achterblijven gedigesteerd en differentieel gelabeld met stabiele isotoop dimethyl labeling, gevolgd door scheiding met SCX waarna LC-MS analyse volgt voor de identificatie en kwantificatie. Door deze strategie te gebruiken kunnen we een kwantitatief onderscheid maken tussen uitgescheiden eiwitten en ongecontroleerde lysis producten. De hoeveelheid moleculen van ruwweg 4500 bloedplaatjes eiwitten werd vergeleken wat leidde tot de conclusie dat slechts

124 eiwitten (<3% van het proteoom) werden uitgescheiden met voldoende significantie. De uitgescheiden eiwitten hebben een dynamisch bereik van ten minste 5 ordes van grootte en de secretie van verscheidene van deze eiwitten kon worden bevestigd door ELISA. Binnen de uitgescheiden eiwitten was een grote verrijking waarneembaar aan secretietags. Thrombospondin, von Willebrand factor en Platelet factor 4, enkele zeer bekende hoog abundante uitgescheiden eiwitten waren ook zeer abundant in het door ons gedefinieerde releasaat. Daarnaast deze bekende uitgescheiden eiwitten zijn ook veel nieuwe laag abundante eiwitten geïdentificeerd.

In **hoofdstuk 4** worden het proteoom van neutrofiele, eosinofiele en basofiele granulocyten geïdentificeerd, gekwantificeerd en vergeleken. Op het moment zijn er geen diepgaande proteomes beschikbaar voor deze 3 granulocyten, welke tot op zekere hoogte vergelijkbaar zijn maar ook zeker een divergerende rol hebben. Voor elk van deze 3 granulocyten is een diepgaand proteoom geïdentificeerd en gekwantificeerd. Deze studie is een eerste stap om de verschillen op eiwitniveau tussen deze granulocyten bloot te leggen wat resulteerde in eiwit vingerafdruk voor elk van deze granulocyt types.

Curriculum Vitae

Patrick Wijten was born at 29th of August 1984 in Eindhoven, the Netherlands. In 2003 he finished his Higher General Secondary Education after which his initial exploration of science was done by studying Biology and Applied Medical Laboratory Technology at Avans University (Breda, NL) with a specialization in Biotechnology. For this Bachelor degree he was required to do two internships, the first one at MSD (former Organon, Oss, NL). During this internship his interest in proteins and mass spectrometry was piqued as he studied the prediction of the bio-activity of heavily glycosylated gonadotrophins using analytical techniques of which mass spectrometry was one of the explored techniques, under the supervision of Rick Scheurs and Dr. Michel Eppink. Following this internship he did an internship at Philips Research where he first came into contact with proteomics as he studied the potential of the selective enrichment of sub-proteomes to be implemented for biomarker discovery under the supervision of Dr. Ir. Eef Dirksen and Dr. Edwin Romijn.

Following the completion of his Bachelor degree he continued with the Master program Molecular Life Sciences at Wageningen University where he specialized in Biological Chemistry. He did my thesis at the department of Phytopathology where he studied the differences between the secretomes of plant pathogen ***V. dahliae*** race 1 (resistance genes in tomato) and race 2 (no known resistance) strains using proteomics under the supervision of Dr. Ir. Peter van Esse and Dr. Ir. Bart Thomma. The final internship he did was at MSD (former Organon, Oss, NL), where he studied the elucidation of small molecule and steroid structures and distinguishing between epimers using Residual Dipolar Couplings derived from comparing anisotropic and isotropic spectra under the supervision of Ruud Aspers and Dr. Martin Jaeger.

Due to his continued interest in mass spectrometry and proteomics, in 2011 Patrick joined the Biomolecular Mass Spectrometry and Proteomics Group at the University of Utrecht as a Ph.D. student under the supervision of Dr. Arjen Scholten and Prof. Dr. Albert J.R. Heck. The research was focused around method development and (clinical) application on human circulating cells, the results of which are presented in this doctoral thesis.

List of publications

Wijten P, Bleijerveld OB, Cappadona S, McClellan EA, de Groot PG, Hoefer IE, Pasterkamp G, Heck AJ, Scholten A. Biomarker Discovery in Circulating Cells Benefits from Curating for Prospective Blood Matrix Contaminants. *Manuscript in preparation.*

Bleijerveld OB, **Wijten P**, Cappadona S, McClellan EA, Polat AN, Raijmakers R, Sels JW, Colle L, Grasso S, van den Toorn HW, van Breukelen B, Stubbs A, Pasterkamp G, Heck AJ, Hoefer IE, Scholten A. Deep proteome profiling of circulating granulocytes reveals bactericidal/permeability-increasing protein as a biomarker for severe atherosclerotic coronary stenosis. Journal of proteome research. 2012;11:5235-5244

Wijten P, van Holten T, Woo LL, Bleijerveld OB, Roest M, Heck AJ, Scholten A. High precision platelet releasate definition by quantitative reversed protein profiling--brief report. Arteriosclerosis, thrombosis, and vascular biology. 2013;33:1635-1638

van Holten TC, Bleijerveld OB, **Wijten P**, de Groot PG, Heck AJ, Barendrecht AD, Merkx TH, Scholten A, Roest M. Quantitative proteomics analysis reveals similar release profiles following specific par-1 or par-4 stimulation of platelets. Cardiovascular research. 2014;103:140-146

Tak T, **Wijten P**, Heeres M, Pickkers P, Scholten A, Heck AJ, Vrisekoop N, Leenen L, Borghans JA, Tesselaar K, Koenderman L. Human hypersegmented neutrophils identified as a separate subset by proteome profiling and in vivo pulse-chase labelling. Blood, *accepted.*

Acknowledgements

Albert, first and foremost I would like to thank you for giving me the opportunity to work in your lab. I felt privileged to be able to work in this lab amongst colleagues with such high level of expertise as well as be able to work with state-of-the-art equipment. Especially now that I have been gone from the lab for a while, I realize just how much I have learned during the time I spent there. From you in particular I learned to be pragmatic and give a positive spin to results that don't necessarily reflect that at a first glance. Besides this, you also gave me my much needed wake-up call after my health deteriorated and my progression became stagnant. The message at the time was a harsh one, but on retrospect this inspired me to follow through with the completion of my thesis and I don't think I could have done without it and your continued support.

The next person I have to thank is Arjen who was my daily supervisor until your departure to an industry position late in my PhD. I think we worked quite well together right from the start as you gave a large degree of freedom to work on my projects which I appreciated a lot. Reflected by how you once described your style of management to me as "liefdevolle verwaarlozing" which is Dutch for loving neglect. We would have talks initially about the project and which general direction we should go and beyond that it was largely up to me as to how I would interpret and execute it. I think it is a style that suited me well in the initial and middle phases of a project, during which I learned a lot by fending for myself but towards the ending stage I think we can all agree I needed a bit more of a hands-on approach to let go of the insecurity which made me reiterate steps that I already completed and move on towards completion. But that is also part of the learning curve and that has provided me with valuable insight into myself.

Which leads me to thank Onno, together we were in the part of the CTMM consortium with our projects and since you were in the group for quite some time already, you were a natural fit to teach me a lot of the techniques involved with working on circulating cells. You helped me out a lot, either by together performing the experiments and processing or by one your famous lengthy emails with all necessary information. I really enjoyed our discussions about both relevant and irrelevant matters. With the latter mostly being about the TV-shows we both liked, by the way I still haven't watched "the Sopranos", but I plan on remedying that soon!

A big thank you goes out to my various collaborators without whom my thesis would be much less interesting to read. Prof. Gerard Pasterkamp and Imo Hoefer from the department of Experimental Cardiology department of UMCU and Elizabeth McClellan of Erasmus MC for the work on biomarker research, Mark Roest and Thijs van Holten of the department of Clinical Chemistry and Haematology of UMCU for the work on the platelet releasate, Prof. Leo Koenderman and Tamar Tak of UMCU for the work on the granulocytes. To me these collaborations were very pleasant and the scientific discussions I had with each of you helped broaden my knowledge about the subject matter a lot.

Of course I need to thank everyone who has been in office Z602 during my time at the lab: Thomas, Serena, Natalie, Thin Thin, Thierry and Andrey. Each one of you made my stay more enjoyable by talking about work or non-work related matters and several of you have served as my cookie alert for whenever something could be found in the kitchen, for which I am of course eternally grateful! Kidding aside I really appreciated every one of you! P.S. Thierry, you once sent me a study called: The unsuccessful self-treatment of a case of "writer's block". I was so close to publishing a replication of those results but as you are reading this now, we can luckily conclude I have failed at that. ;-)

A special thanks goes out to the technicians of the lab: Mirjam, Harm, Soenita, Anja, Arjan, Jamilla and Luc who have helped me to varying degrees but all are instrumental to the smooth operation of the lab!

A special thanks also goes out to Corine for the occasional chat and helping me with various things related to my PhD which also extended beyond my contract as you helped me deal with my problems with IT that has tried to and succeeded to shut my account down on many occasions!

Also, I would to thank my students Liy Liy, Huibert-Jan and Henry that helped me on my projects.

Besides these people there have been many who I have shared my time in the lab with who I would like to thank, in no particular order: Esther, Teck, Fabio, Nicolas, Salvatore, Reinout, Henk, Erik, Hoijiang, Donna, Pepijn, Geert, Sara, Renske P, Christian Sr., Alba, Basak, Nadine, Nicolai, Jessica, Christian Jr., Hongtao, Gianluca, Lucrece, Vincent, Marco B, Benjamin, Marco H, Bas, Monique, Shabaz, Ana, Rebecca, Renske vG, Pierro, Javier, Mao, Christian, Maarten, Eleonora, Kees, Adja, Joost and Patrycja.

I saved the best for last, of course I need to thank my paranymphs. Ayse for starters, we shared a long and not always pleasant journey together at the lab. We were both responsible for the Q-TOF early on which lead to us spending a lot of time together, working at the various of problems the machine threw at us. We got to know each other quite well that way and I really appreciate our talks during the time we spend at the lab and extending beyond that. De andere paranimf die ik moet bedanken is Michiel, we kennen elkaar al zo ongeveer ons gehele leven en je hebt gedurende die tijd altijd voor de al dan niet nodige afleiding gezorgd. Dit heeft me zeker in de afgelopen jaren veel geholpen! En naast mijn paranimfen ben ik ook Wendy, mijn broer(tje) en ouders, in het bijzonder mijn moeder, veel dank verschuldigd om te komen tot waar ik nu ben.

I almost forgot… I also have to thank Antonín Dvořák for composing Symphony No. 9. Better known as New World Symphony, for some odd reason that was the music that best allowed me to write, which in case it wasn't clear already… wasn't always easy!