Ghent University

Faculty of Pharmaceutical Sciences

Laboratory for Pharmaceutical Biotechnology

Sciensano

Expertise, service provision and customer relations

Quality of vaccines and blood products

Prevention is better than cure:

Non-targeted and sensitive screening of impurities in plasma-derived immunoglobulin with mass spectrometry-based proteomics

Franck Limonier

Promoter: Prof. Dr. Apr. Dieter Deforce

Promoter: Dr. Myriam Sneyers

Promoter: Dr. Maarten Dhaenens

Thesis submitted in the fulfillment of the requirements for the degree of Doctor in Pharmaceutical Sciences.

Members of the examination committee

Prof. dr. apr. Sarah De Saeger
Department of Bio-analysis, Faculty of Pharmaceutical Sciences, Ghent University

Prof. dr. apr. Christophe Stove
Department of Bio-analysis, Faculty of Pharmaceutical Sciences, Ghent University

Prof. dr. Francis Impens
Center for Medical Biotechnology, Vlaams Instituut voor Biotechnologie - Ghent University

Prof. dr. David Communi
Institute of Interdisciplinary Research in human and molecular Biology, Medical Faculty - Université Libre de Bruxelles

Dr. Celine Vanhee
Medicines and Health Products, Department of Chemical and Physical Health Risks, Sciensano

Dr. Caroline Bellac
Official Medicines Control Laboratory, Swissmedic, Swiss Agency for Therapeutic Products

This work was funded by a PhD grant of Sciensano (formerly Scientific Institute of Public Health WIV-ISP)

Table of Contents

Dankwoord / Mot de remerciement

Mes remerciements vont en premier lieu à six personnes sans qui rien de tout cela n'aurait été possible.

Geneviève et **Koen**, vous resterez comme les premières personnes à m'avoir confié les clés d'un projet professionnel, qui plus est dans un secteur pour moi plein de sens : la santé publique. Votre modestie n'a d'égal que votre expertise et votre investissement dans votre travail quotidien pour la qualité des vaccins et produits sanguins. Ik herinner me nog mijn interview in 2013 in de Halter zaal. Tientallen meetings, mails en koffiepauzes later, hoop ik dat jullie tevreden zijn met dit allereerste doctoraat in de dienst Biologische Standaardisatie, of well 'QVBP' zoals we het nu moeten noemen. Koen, van harte bedankt dit doctoraatproefschrift te hebben herlezen.

Maarten en **Katleen**, de zo kostbare post-docs voor al de *Mass spec'ers* van de Lab FBT. De omstandigheden hebben ervoor gezorgd dat ik jullie twee successief als begeleider had. Wat een kans eigenlijk! Jullie zijn een beetje *het yin en het yang* van het onderzoek. De stille kracht en de terughoudendheid op één kant, de passie en het communicatieve enthousiasme op de andere. En tenslotte zijn jullie twee zeer inspirerende wetenschappers. Maarten, Ik ben ook super blij dat jij eindelijk mijn derde copromotor wordt. Al deze intense discussies over SWATH en DIA waren echt boeiend. Ik hou ook zeker contact met U als *the* Belgische expert van de Synapt!

Dieter en **Myriam**, mijn promotors, les garants bienveillants de cette thèse. Ik ben altijd onder de indruk van de zo grote hoeveelheid werk dat jullie kunnen beheren met overtuiging en aandacht voor iedereen. Professeur, chef de laboratoire de recherche et membre de l'EMA, ou directrice et pilote de la fusion de deux des plus grands instituts scientifiques fédéraux belges, *total respect* !

Verder zou ik ook graag van hart bedanken de leden van de examencommissie om mijn thesis grondig te lezen en een constructieve bijdrage te leveren.

Je remercie aussi **Stéphane**, **David**, **Koen** et **Christophe**, les partenaires industriels de ce projet, pour avoir fourni les échantillons, ainsi que les colonnes protéine A – protéine G. Merci particulièrement Stéphane pour ta disponibilité et les échanges téléphoniques instructifs au début du projet.

Daarna zou ik graag al de mensen bedanken die ik ontmoette in mijn twee verschillende werkplaatsen tijdens deze vijf jaar thesis: het Lab FBT van de UGent en de dienst Biologische Standaardisatie in het WIV/Sciensano.

Het LabFBT team, een leuke kweekvijver van gemotiveerde *young researchers* en *histones addicts*. Taalbarrière en mijn deeltijds aanwezigheid deden me soms eruitzien als een *alien* maar uiteindelijk was het echt leuk en interessant voor mij om bij jullie te komen werken!

In het bijzonder zou ik graag een paar teammaten bedanken. **Sander**, de bioinformaticus van de ploeg, bedankt voor je kostbare hulp in de codering van de Python script, en je scherpziende reviews van de SWATH publicatie. **Sophie** *aka the gold fingers*, bedankt voor je onschatbare technische hulp met "onze" TT! Bedankt **Elisabeth**, voor de aangename en constructieve uitwisselingen tijdens onze parallelle thesis avonturen vanaf 2013, **Paulien** voor je vriendelijkheid en pedagogie, **Ellen**, voor de West-Vlaamse vastberadenheid. **Laura**, **Simon** en **Bart**, de volgende *Mass spec'ers* generatie, succes met jullie thesis en als jullie het mist, kunnen jullie me toch af en toe een Teamviewer wachtwoord doorsturen! **Liesbeth**, **Yens** en **Veerle**, voor het goede humeur en de paar duwtjes in het oude labo tijdens de eerste jaren, **Yannick**, voor je communicatieve lach en de enkele ICT boosten, **Bart**, **Senne** en **Christoph** voor jullie toegankelijkheid en de een of andere leuke discussie, **Inge** voor de positieve energie en de administratieve follow-up.

Toute la grande équipe, présente et passée, du service Standardisation Biologique/Qualité des vaccins et produits sanguins à l'Institut/Sciensano. Entre deux release dans le LIMS, ring test et non-conformités résolues, vous étiez là, consciemment ou pas, pour me changer les idées et faire de la pause-café de 9h45 et la dégustation du plat du jour et autre *take-away* des rendez-vous incontournables. Et pour l'équipe actuelle, ce n'est pas encore prêt de s'arrêter !

Mes voisins du fameux C77, open-space parfois bruyant mais surtout convivial, et parfois lieu de canulars et recettes mémorables : **Olivier**, capitaine des produits sanguins et prophète des ABL et du magret de canard (« le gras, c'est la vie ! ») ; **Valerian**, aussi calé en médicine chinoise qu'en table dance sous la pluie (et merci pour tes coups de pouce avec Excel !) ; **Fabrice**, roi du bon mot malicieux et responsable en chef de la bonne ambiance ; **Julien**, avec qui discuter aussi bien cuisine (sans fromages !) que jeux vidéo ; **Alexandre**, boss du D&R et chasseur de rongeurs sauvages ; **Camille**-je-te-raconte-pas-ce-qu'-Oscar-a-encore-fait ; **Lieke**, ma voisine d'en face ; **Eléonore**, cavalière et pâtissière/boulangère de talent ; **Virginie S.**, aussi déterminée en analyse de NC qu'en trampoline ; **Margaux**, qui n'a plus vingt ans mais une thèse sur le feu (bonne chance !) ; **Aurélie** et **Valérie**, médaille des collègues sympas et relax.

Pascale, la fameuse Speedy Gonzalez du service, gardienne des formulaires, merci pour ta bonne gestion de toutes ces commandes et notes de frais. Tous les comparses du mess ensuite : **Lorenzo**, amateur autant de sport (parfois un peu extrême) que de bruschettas ; **Wim** en zijn perfecte balans tussen efficiëntie en humor, **Joëlle**, **Vinciane** et **Virginie M.**, les piliers de la pause-café et reines du saumon poché au Martini ; **Saloua**, molenbeekoise écolo avec conviction et humour ; **Laurence V.**, pourfendeuse des grèves de train ; **Dominique**, la force tranquille entre 2 boites eFarmz ; **Mathias** et **Lucas**, les râleurs attachants ; **Céline R.** et ses cours de cuisine ; **Céline**

F. et son rugby anti-gueule de bois ; **Céline D.** et ses merveilleux ; **Romain**, **Simon**, **Ellen** et **Boris**, la table séparatiste des djeuns cools où il était toujours sympa de faire un tour ; **Charles**, **Adrien**, **Quentin**, **Christophe**, les (déjà plus tant) nouveaux gars sympas ; **Julie** et ses recettes culinaires à l'accent chantant ; **Hanh** et son badminton ; **Hind** et **Laurence D.**, aussi mobiles et « Sciensano-transversal » que moi ; **Laure** et son hydromel 0% machisme ; **Florence** et ses réunions familiales au Mess ; **Sydney** et ses fêtes techno jusqu'à 8h du matin, et enfin les transferts sympas du CODA, **Pierre**, **Brigitte** et **Françoise**. Merci aussi **Fatiha**, **Sabah** et **Rachida** pour m'avoir permis de suivre vos tests au début de ma thèse. (Marc ?) **Antoine** enfin, beaucoup de succès dans ta thèse!

Ik zou ook graag **Joris** en **Tim** hartelijk willen bedanken voor hun vertrouwen met mijn nieuwe positie die ik in juni 2017 kon beginnen. Oriëntatie chemische (en biologische) risico's is een uitdagende en boeiende taak die ik blij ben om verder te ontwikkelen in Sciensano. Et ce avec des chouettes nouveaux collègues : **Koen D.**, **Tom**, **Julien**, **Laure J.** et **Philippe**, **David** …

Je vais évidemment terminer mes remerciements par mes proches, soutien ô combien précieux et irremplaçable.

Les amis de (très) longue date, mes immanquables et inénarrables carabistouilleurs **Hervé**, **Alex**, **Raph**, **Minh** et **Greg** mais aussi **Caro**, **Oli, Tim**, **John**, **Victoria**, **Martina**, **Vincent**, **Héloïse**, **Dorine Dédé**, **Lorenzo** et **Max**. Toute la grande clique de lascars Saint-Hubertois (vous vous reconnaitrez !) que je prends toujours un vrai plaisir à voir depuis toutes ces années, les camarades LS, OSEL, La Vague / Divagation, et de toutes ces bonnes soirées à Bruxelles.

Maman, **Papa**, et **Mamy**, votre soutien indéfectible et amour inconditionnel sont pour moi un pilier, et je serai toujours là pour vous le rendre. Beaucoup de tendresse et des remerciements sincères aussi pour **Tantine** et **Edouard**, **Mireille** et **Philippe.** Une pensée enfin pour la famille plus éloignée que le temps et la distance ne me font pas pour autant oublier.

En parlant de famille élargie, je pense aussi bien sûr à la grande et belle tribu de Mont-Sur-Marchienne et environs : **Eric** et **Danielle**, **Jules**, **Jean** et **Anne-Marie, Franz** et **Bernadette**, **Papy René** et **Danielle**, **Yvon** et **Andrée**, les cousins **Quentin** et **Marine**, **Guillaume** et **Caroline**, **Valérie** et **Guillaume**, les SNS **Marie-Alix**, **Emeline** et **Maëlle**, **Françoise**, **Joël** et leurs trois petits diables, **Philippe**, **Sigrid** et **Charles** et leurs deux petits anges. Eeeeet saantéééé à tous, amélioré(e) bien sûr.

Comme dans les bons articles scientifiques, le dernier nom occupe une place toute particulière. **Delphine**, je tairai ici tous les surnoms que je te donne depuis ces plus de 6 ans de relation et 2 ans de vie commune. Je sais que tu n'aimes pas les grands débordements romantiques, mais je dirai simplement que je ne souhaite que de passer encore de très nombreuses années avec toi !

Franck

List of abbreviations

ACN	Acetonitrile
ApoH	Apolipoprotein H
CID	Collision-induced dissociation
ColV	Column volume
CPLL	Combinatorial Peptide Ligand Library
CV	Coefficient of variation
Da	Dalton
DDA	Data-dependent acquisition
DIA	Data-independent acquisition
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDQM	European directorate for the quality of medicines & healthcare
ELISA	Enzyme-linked immunosorbent assay
EMA	European Medicines Agency
ESI	Electrospray ionization
F	Factor
FA	Formic acid
FDA	Food and Drug Administration
FDR	False discovery rate
FXI(a)	Coagulation factor XI (activated)
FXII	Coagulation factor XII
GMP	Good manufacturing practices
HBV	Hepatitis B virus
HCP	Host cell proteins
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HSA	Human serum albumin
HYE	Human-yeast-E.coli
Ig	Immunoglobulin
IT	Ion trap
ITP	Idiopathic thrombocytopenic purpura
IVIg	Intravenous immunoglobulin
kDa	kiloDalton
LAP	Low-abundant proteins
LC	Liquid chromatography
LOQ	Limit of quantification
mAb	Monoclonal antibody
MALDI	Matrix assisted laser desorption ionization
MGF	Mascot generic format
MRM	Multiple reaction monitoring
mRNA	Messenger RNA
MS	Mass spectrometry
m/z	Mass to charge ratio
NAPTT	Non-Activated partial thromboplastin time
OCABR	Official control authority batch release
OMCL	Official medicines control laboratory

PBS	Phosphate buffered saline
PEI	Paul-Ehrlich-Institut (German OMCL)
Ph. Eur.	European Pharmacopoeia
PKA	Prekallikrein activator
PMF	Plasma master file
Q	Quadrupole
QC	Quality control
QTOF	Quadrupole-time-of-flight
RBC	Red blood cell
RhD+/-	Rhesus D positive/negative
RNA	Ribonucleic acid
ScIg	Subcutaneous immunoglobulin
S/D	Solvent/detergent
SELDI	Surface-enhanced laser desorption ionization
SOP	Standard operating procedure
SRM	Selected reaction monitoring
SWATH	Sequential windows acquisition of all theoretical fragment ion
TEE	Thromboembolic events
TGA	Thrombin generation assay
TOF	Time-of-flight
TRFE	Serotransferrin
T	Technical replicate
TT	TripleTOF
vWF	Von Willebrand factor
WHO	World Health Organization

Chapter 1:

Biological introduction

1. Biological introduction

1.1 Plasma-derived therapeutics

Collected blood plasma is a unique source material for the manufacture of therapeutic products. Plasma is so much more than just the liquid part of blood that carries the red and white blood cells and the platelets. The clear, yellowish liquid which constitutes about 55% of the total blood volume serves critical functions: regulating the blood pressure and volume, transporting vital minerals, maintaining a proper pH balance in the body, and supplying crucial proteins for blood clotting and immunity. Plasma contains hundreds of proteins, ranging from abundant (> 10 mg/mL) and multifunctional albumin and immunoglobulin (Ig) to diverse less-abundant (down to pg/mL) proteins with specific physiologic activities including protease inhibitors (α1-antitrypsin, antithrombin …) and coagulation factors (factor VIII, fibrinogen, von Willebrand factor …). Plasma transfusion thus makes accessible a wide reservoir of components that demonstrate clear therapeutic interest.

At the dawn of the 20th century, von Behring was awarded the first Nobel Prize in Medicine for his work on serum therapy against diphtheria [1]. Curative or preventive blood serum was obtained from immunized animals or convalescent human. In the 1930's, better knowledge on the serum protein composition allowed to localize antibodies in the immunoglobulin (Ig) compartment [2]. Few years later, Cohn and Oncley developed separation techniques of plasma proteins into individual stable and concentrated fractions with improved therapeutic impact as compared to whole crude plasma [3,4]. At the same period, the Red Cross launched a national blood donations service to collect blood to treat injured soldiers during the World War II, mainly with albumin and dried plasma [5]. Before the late 1980's, viral transmission was the major concern when using medicinal products derived from human plasma. Over the years, manufacturing processes were progressively refined with dedicated and validated inactivation or removal steps for possible infectious agents, in compliance with good manufacturing practices and quality assurance. This resulted in a sophisticated and global plasma fractionation industry. Nowadays, around 25 million litres of human plasma are fractionated each year in the world, in over 50 factories [6].

As reported in Table 1.1, around 20 plasma proteins are known for exhibiting therapeutic effect in different life-threatening conditions such as bleeding trauma, immunological disorders, congenital deficiencies and/or tissue degenerating diseases, or infections. These proteins mostly belong to four main groups: albumin, coagulation factors, protease inhibitors & immunoglobulin [6,7]. When specific concentrate is not available, whole-blood plasma is still used for direct infusion [8]. Normal immunoglobulin, specific immunoglobulin (anti-D, anti-rabies & anti-tetanus) and coagulation factors VIII & IX are listed in the 18th WHO Model List of Essential Medicines [9].

Table 1.1 Plasma proteins with therapeutic role.

Proteins in italic present potential therapeutic indications but have no marketing authorization as therapeutic product yet. *** means that recombinant protein production as alternative to purification from plasma is licensed or under advanced development.** [10,11] **Total protein concentration in plasma is 60-75 g/L. Plasma proteins concentration dynamic range is also illustrated in Figure 9.1 in the Addendum.**

Plasma protein	Functional role in plasma	Plasma concen-tration (g/L)	Plasma-derived product	Therapeutic use
Albumin	Molecule transport & regulation of the oncotic pressure	35 - 45	Concentrate *	Blood volume replacement and oncotic pressure (Surgery, trauma, burns, sepsis, hypovolaemia)
Immunoglobulin	Circulating agents of the adaptive immune system	7 - 15	Normal polyvalent immunoglobulin G (Intravenous, subcutaneous or intramuscular)	Immune replacement therapies & immunomodulatory effects
			Hyperimmune specific immunoglobulin G *	Passive immunisation towards infectious agents, foetal Rhesus compatibility
Coagulation Factors	Coagulation cascade			
Factor VIII	Acceleration of Factor X activation	< 0.001	Concentrate *	Replacement (Haemophilia A, von Willebrand syndrome)
Factor IX	Factor X activation	0.005 - 0.01	Concentrate *	Replacement (Haemophilia B)
Prothrombin (Factor II)	Once activated, transform fibrinogen into fibrin	0.15 - 0.2	Prothrombin complex concentrate (factors II, VII, IX, X, proteins C, S & Z)	Correction of the coagulation system, complex liver disease treatment
Factor VII	Factor IX and X activation (extrinsic pathway)	< 0.001	Concentrate *	Replacement (Bleeding control)
von Willebrand factor	Platelet adhesion & Factor VIII binding	0.01 - 0.02	Concentrate (+ factor VIII) *	von Willebrand disease treatment
Factor XI	Factor IX activation (intrinsic pathway)	0.005 - 0.01	Concentrate	Replacement

Factor XIII	Transglutaminase, fibrin crosslink	0.03	Concentrate *	Replacement
Fibrinogen	Blood clotting, once converted into fibrin	2 - 4	Fibrin sealants (fibrinogen + thrombin)	Surgical haemostasis , adhesion or suture
			Concentrate *	(hypo)fibrinogenemia treatment
Protein C (Factor XIV)	Anticoagulation, inflammation, cell death	0.005 - 0.01	Concentrate	Replacement (Venous thrombosis and purpura fulminans)
Protease inhibitors	Regulation of the coagulation, platelet aggregation and complement system activation			
α1-antitrypsin	Inhibitor of neutrophil protease-induced host tissue damage	1.5	Concentrate *	Replacement (Pulmonary emphysema)
Antithrombin	Thrombin inhibitor, clotting reduction	0.3	Concentrate *	Surgery, thrombosis treatment, replacement
C1 esterase inhibitor	Inhibitor of the complement system to prevent spontaneous activation	0.2	Concentrate *	Replacement (Hereditary/acquired angioedema)
Activated protein C	*Anticoagulation, inflammation, cell death*	*0.005 - 0.01*	*- **	*Sepsis treatment*
Inter-α-trypsin inhibitor	*role in inflammatory process*	*0.15 - 0.5*	*Concentrate*	*Sepsis treatment*
Other proteins				
Transferrin	*Iron binding*	*1.5-3.5*	*Concentrate*	*Sepsis treatment*
Apolipoprotein A-I	*Lipase inhibitor*	*1*	*Concentrate*	*Treatment of hypocholesterolaemia and atherosclerosis*
Mannan-binding lectin	*Innate (aspecific) immune system component*	*< 0.001*	*Concentrate*	*Replacement (infantile infections)*
Plasmin(ogen)	*Fibrinolytic enzyme*	*< 0.001*	*Concentrate*	*Peripheral arterial occlusion treatment, replacement*

von Willebrand factor cleaving protease	*von Willebrand factor inactivation*	*< 0.001*	*-*	*Replacement (Thrombotic thrombocytopenic purpura)*
Butyrylcholinesterase	*serine hydrolase that cleaves esters of choline*	*< 0.001*	*Concentrate*	*Treatment of organophosphorous nerve agents poisoning*

From the 1980's, recombinant DNA technology applied to cell culture expression systems has emerged as an alternative production platform for plasma proteins. If pathogen safety has been the original justification, an additional driving force for new developments is currently the need for adequate supply of the plasma proteins at affordable cost worldwide, which has not been reached yet. Compared to plasma-derived products, the main advantage of recombinant biotechnology is the unlimited and relatively low-cost source material. On the other hand, caution should be paid during product development to avoid immunogenicity when non-human/non-physiological expression systems are used [10].

Contrary to most of the plasma proteins, recombinant antibodies cannot substitute for plasma-derived normal Ig [10,12]. Consequent to this particularity and because of an increasing demand due to various therapeutic uses, normal polyvalent Ig are today the leading product of the plasma-derived therapeutics' market: the global demand is not satisfied and drives the needs of plasma for fractionation [13].

1.2 Immunoglobulin G therapeutic products

1.2.1 Functions and structure of the immunoglobulin

Immunoglobulin is a group of proteins that function as antibodies, namely the circulating "weapons" of the adaptive humoral immune system against pathogens. Genetic combinatorial multiplication leads to a highly diverse repertoire of antibodies capable of specifically binding theoretically all kind of exogenous structures. Once the antigen-antibody interaction is established at the level of the lymphocytes B membranes, lymphocytes B are differentiated into secretory cells (called plasmocytes) producing numerous circulating antibodies. Monoclonal antibodies (mAb) are produced by secretory cells arising from a same given lymphocyte and thus all recognize a single epitope on the antigen. Polyclonal antibodies are produced by secretory cells arising from many distinct lymphocytes and thus target numerous epitopes. Even when the antigen is no longer present, memory B cells archive all the antibodies generated in response to immunogenic insults so that each organism has its own exposures record reflected in a unique antibodies repertoire in the serum, at any given moment.

Each immunoglobulin is composed of two heavy and two light chains linked by disulphide bonds. Each heavy and light chain is about 50 and 25 kDa respectively, resulting in a monomer molecular weight of approximately 150 kDa. One part of the heavy chains forms a constant region, named Fc domain, while the second part of the heavy chains and the light chains form a variable region, named Fab domain (Figure 1.1).

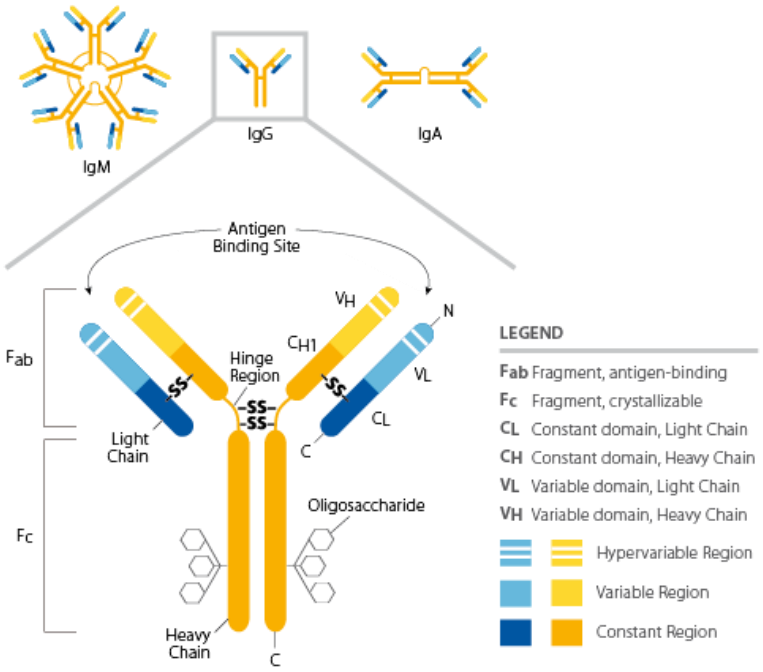


Figure 1.1 Immunoglobulin structure [14]

While there are only two types of light chains in mammals (kappa and lambda), the variability of the Fab domain comes from the so-called complementary determining regions, which result from gene rearrangements events and ensure specific binding to an antigen. Following aggregation of the Ig molecule on the surface of an antigen, the Fc domain mediates effector functions such as activating the complement system, inducing antibody-dependent cellular cytotoxicity via binding the Fc receptor on natural killer cells, macrophages or neutrophils, or inducing opsonisation by phagocytes [15,16].

Five Ig isotypes - IgG, IgA, IgM, IgD and IgE - are distinguished based on the Fc domain sequence, as result of genetic recombination in the heavy chain locus of mature B cells, known as “class switch recombination”. IgG isotype is subdivided into four subclasses of decreasing abundance – IgG1, IgG2, IgG3 and IgG4 – while IgA into two – IgA1 and IgA2 [17]. Each Ig isotype and subclass has different distribution, structure and specialized effector functions in the immune response, as summarized in Table 1.2. The IgG isotype is the main type in blood, representing about 75% of the circulating Ig, and has a much longer half-life than IgA and IgM (23 days for IgG versus 7 and 5 days for IgA and IgM respectively) [18].

The Fc domain also carries N-glycans that modulate the biological activity and vary depending on the age, the gender and the environment [19,20].

Table 1.2 Ig isotypes characteristics [17,21]

Ig iso-type	Predo-minant polyme-risation	Distribution	Propor-tion in blood	Ig Sub-type	Main functional activity
IgG	Monomer	Main type in blood and extracellular fluid, transport across placenta (mainly IgG1&3)	75%	**IgG1 (60%)**	Response to soluble and membrane protein antigens
				IgG2 (29%)	Response to bacterial carbohydrate antigens
				IgG3 (7%)	Induction of effector functions, potent pro-inflammatory, viral infection
				IgG4 (4%)	Repeated or long term antigen exposure in non-infectious setting
IgA	Dimer	Main Ig in mucosal secretions, eventually membrane-bound, transport across epithelium, lower level in blood. IgA1 predominates in blood, IgA2 in mucosal secretions.	15%	**IgA1**	Immune function of mucosal membrane and secretions (tears, sweat, milk, gastrointestinal tract, respiratory epithelium …).
				IgA2	
IgM	Pentamer	Cell membrane of immature lymphocyte B and secreted form	10%	-	First humoral response to initial antigen exposure, induction of effector functions
IgD	Monomer	Cell membrane of immature lymphocyte B, low secretion level in blood	0.25%	-	Antigen receptor of lymphocytes, mast cells activation for antimicrobial factors production
IgE	Monomer	Very low levels in blood or extracellular fluids, but avidly bound by mast cells receptors beneath the skin and in mucosa and connective tissue	0.05%	-	Response to parasites (such as helminths and protozoa) and allergens, potent pro-inflammatory

1.2.2 The development of concentrated immunoglobulin as therapeutic products

The unique and complex structural characteristics and biological functions of the immunoglobulin highlight both the interest and the difficulty to use them as therapeutic products. The Fab and Fc domains native structure should be preserved to ensure functional efficacy.

In the early days of plasma fractionation, the large scale plasma fractionation process developed by Cohn and Oncley already allowed to obtain 70-80% concentrate of IgG, with also substantial amounts of IgA and IgM [3,4]. When administered subcutaneously or intramuscularly, such concentrates demonstrated to be beneficial for both pathogen infections treatment and prevention (measles, hepatitis A, pertussis …) and reduction of infections in patient with immune deficiency (agammaglobulinemia or hypogammaglobulinemia). Bruton was the first to treat agammaglobulinemia with subcutaneous Ig in 1952 [22].

However, these Ig preparations were not suited for severe cases as they could not be administered intravenously due to their poor purity. When entering the blood circulation upon intravenous injection, immunoglobulin aggregates were causing non-specific spontaneous complement activation and life-

threatening anaphylactic reactions, while residual active proteases could cause side-effects such as hypotension [23]. Further process refinement had to be performed to allow alternative administration to the painful, irritating and slowly absorbing (4 to 6 days) repeated small doses of the intramuscular route [24]. Artificial aggregation was suspected to be caused by the exposition to hydrophobic surfaces or high temperatures during the manufacture. Enzymatic (pepsin, plasmin) or chemical (alkylation, acylation, sulfonation) modifications were considered to overcome aggregation but particular caution had to be taken to avoid IgG loss of integrity and subsequent shortened half-life and altered effector functions. By the end of the 1970s, the first Ig products suitable for intravenous administration, abbreviated as IVIg, resulted from improved fractionation processes that managed to significantly reduce aggregation and anticomplementary activity while increasing the purity and preserving the Ig intact. Modern IVIg preparations are more than 99% IgG, and more than 95% monomeric [2,16,25].

Another key point in the development of IVIg manufacture was the pathogen safety and the incorporation of various dedicated virus elimination procedures. The efforts invested in the refinement and the standardisation of the Ig production were even more strenuously motivated by the discovery of new therapeutic indications since the 1980s. Beside replacement therapies and passive immunisation towards infectious agents, Ig demonstrated therapeutic immunomodulation and anti-inflammatory efficacy in various autoimmune and inflammatory diseases: idiopathic thrombocytopenic purpura (ITP), Guillain-Barré syndrome, Kawasaki syndrome, dermatomyostosis, chronic inflammatory demyelinating polyneuropathy (CIDP) and multifocal neuropathy myasthenia gravis. Prevention of graft versus host disease in recipients of bone marrow transplants has also been reported. Moreover, numerous other off-label indications were also reported through the years (e.g multiple sclerosis, graft-versus-host disease in transplant patients or HIV) [26–31]. Taken together, this makes Ig products the driving force of what can be now called the plasma fractionation industry.

In the 1970s, Kohler and Milstein fused splenic cells from an immunized animal with an immortalized cell line derived from mouse plasmacytoma to produce a mAb secreting hybridoma [32]. In 1986, the first therapeutic mAb was approved and opened the path for a whole new range of specific antibody therapeutic products for targeted indications. Nowadays, plasma fractionation and biotechnology industry complement each other in supplying an extended range of therapeutic antibody products [10,33].

1.2.3 Plasma-derived hyperimmune specific immunoglobulin

Specific hyperimmune Ig products are derived from plasma collected from selected immunized or convalescing individuals. They aim to contain a great number of polyclonal antibodies against a specific pathogen or toxic substance, e.g. cytomegalovirus, diphtheria toxin, snake poison or hepatitis B virus, in order to provide passive immunisation and eventually curative treatment of receivers towards that specific agent. High-tittered Ig products can be administered either by subcutaneous or intramuscular infusion or intravascular injection and are indicated when immunisation/vaccination is ineffective, unavailable, contraindicated or has not been given before the exposure [7,33].

Anti-D (RhoD) immunoglobulin is a specific type of Ig products which are used to ensure immune compatibility between Rhesus D negative (RhD-) mothers and their Rhesus D positive (RhD+) foetus. Administered anti-D Ig is aiming to neutralize any RhD+ fetal red blood cell (RBC) penetrating the blood

stream of the mother, preventing any maternal immune reaction ("alloimmunisation") that could cause anaemia and possibly the death for a RhD+ baby in a subsequent pregnancy ("haemolytic disease of the newborn") [34].

Contrary to recombinant mAb, the development of resistance to hyperimmune Ig has never been reported. On the other hand, the efficacy of hyperimmune Ig could be suboptimal since only a fraction of the administered antibodies exerts the desired effect. Moreover, hyperimmune Ig production strongly relies on specific blood donors' availability, both in terms of quantity and suitability, resulting in considerable variation between batches. The emergence of oligoclonal recombinant antibodies is expected to provide a convincing alternative [35].

1.2.4 Plasma-derived normal polyvalent immunoglobulin

Normal polyvalent immunoglobulin solutions consist in a broad polyclonal immune repertoire, deriving from the pool of numerous (typically 10000 to 40000) plasma donations from standard healthy donors, who have not been intentionally immunised with particular antigens. They offer an unrivalled way to restore the levels of IgG in the case of primary or secondary immunodeficiency. Different genetic disorders can result in a *primary* immune deficiency: X-linked agammaglobulinaemia, class-switch recombination defects, hypogammaglobulinaemia and selective immunoglobulin deficiencies [36]. *Secondary* immune deficiencies are acquired and disease-related. Cancers with certain B cell malignancies, mainly myeloma and chronic lymphocytic leukaemia, neonatal HIV infection or an hematopoietic stem cell (bone marrow) transplant can lead to such secondary deficiency [18].

Using pooled plasma from different donors ensures that the patient immunity against the most common pathogens is restored, since the entire antibodies diversity of the population is statistically represented in the final product. In addition to resolving the lack of antibodies, normal polyvalent Ig stimulate the adaptive immune response and favour the maturation of dendritic cells and the co-stimulation of T lymphocytes [37].

Beside replacement therapies, several mechanisms are suggested and still discussed to explain observed immunomodulatory and anti-inflammatory beneficial properties of normal polyvalent Ig, e.g. in the treatment of ITP, Guillain-Barré or Kawasaki syndromes. More commonly cited mechanisms are listed hereafter (for more details, see [26,27,36–38]):

- Interactions with membranes molecules & receptors of the immune system cells (B and T lymphocytes, macrophages)
- Modulation of the complement cascade: attenuation of complement mediated injury
- Modulation of cytokine production: induction of anti-inflammatory cytokines
- Neutralisation of autoimmune antibodies
- Regulation of cell proliferation and apoptosis

Polyvalent immunoglobulin products are usually infused intravenously (peripheral veins or central catheters). Alternatively, immunoglobulin can be administered subcutaneously (under the skin, into the abdomen or thighs). Intramuscular infusion is no longer preferred as it is painful and suboptimal.

Subcutaneous administration allows treatment at home and for patients with difficult venous access. Still, differences in bioavailability and pharmacokinetics properties might require higher doses (1:1 to 1:1.5 IVIg:ScIg dose adjustment coefficients) [36,39,40]. Consequently, subcutaneous preparations are often more concentrated (10-20%) than the intravenous ones (5-10%). The average recommended dose for antibody replacement therapy is 0.4-0.8 g/kg body weight monthly. The exact replacement dose should be individualized for each patient, based on medical follow-ups [36,38,41]. For immunomodulatory/anti-inflammatory indications, administered dose can be increased to about 2g/kg body weight every 3 to 4 weeks, which is 3-5 times more than for replacement therapy [42,43].

1.2.5 Recombinant antibodies

Modern recombinant antibody engineering biotechnology is expected to provide around 70 approved monoclonal products on the market by 2020 and is already driving the overall sales of the biopharmaceutical market. mAb are indicated in the specific prophylaxis (treatment or prevention) of various diseases, ranging from orphan diseases to multiple sclerosis, asthma, rheumatoid arthritis and various cancers [10,44]. Recently, single-domain mAb fragments mimicking heavy chain only antibodies demonstrated interesting pharmacokinetics properties [45]. However, mAb are mostly administered as a single molecule to address a single disease. Escape variants to mAb are easily observed in lab settings and this hampers the efficacy for complex diseases. Nowadays, synergistic oligoclonal cocktails are increasingly produced to combine antibodies directed towards multiple epitopes of the target, in order to avoid these drawbacks. These products are expected to outmatch the efficacy of plasma-derived hyperimmune specific products, since here the whole antibody content is supposed to exert the desired effect [35]. On the other hand, they still fail to truly mimic the whole immune repertoire and are thus not expected to be substitutive for plasma-derived normal Ig. The general manufacture approach for recombinant antibodies – in vitro DNA modification followed by expression in controlled cells, animals or plants lines – moves a part of the plasma therapeutics products into a totally different field than the plasma fractionation. Nevertheless, while recombinant technology largely circumvents the pathogen transmission risks, its manufacture developments are driven by the two same key safety principles as the plasma fractionation: integrity and purity. Plasma proteins to be administered should be as close as possible to their native status in a healthy individual, and as free as possible from any undesired residual, be it from the plasma pool, the process or the production host cell.

1.3 Modern manufacture and safety of human plasma-derived immunoglobulin

1.3.1 The modern manufacture of immunoglobulin from plasma

1.3.1.1 Plasma collection

Plasma for fractionation is collected by licensed blood collection centres. Candidate donors undergo a medical interview to prove compliance for a donation. Eligibility criteria such as age, body weight, protein and ferrous levels ensure protection of donor's health [46]. By definition, anticoagulant is used during plasma collection, so that coagulation factors are prevented from clotting and are preserved, in contrast

to serum collection, where blood is allowed to clot. Plasma more accurately reflects the actual circulating blood.

Recovered plasma is obtained from whole blood donations (primarily done for red blood cell concentrates supply). Whole blood is centrifuged, so that plasma is separated from cellular components [16]. Good mixing with anticoagulant solution during the whole collection process and controlling blood temperature should preserve the integrity of the plasma. About 300mL plasma can be recovered from a single blood donation. Recovered plasma represents about 35% of the plasma fractionated in the world [6]. *Source* plasma is directly collected from the donors by plasmapheresis device where removed blood is anticoagulated (typically with sodium citrate solution), and immediately separated by physical means (centrifugation, filtration or combination of both) into components, some (e.g. red cells) being injected back to the donor [16]. A healthy adult can tolerate 600 to 800 mL plasma donation through plasmapheresis at a frequent base, with a maximum of 2L per month and 15L per year in Belgium [33,47]. Source plasma represents 65% of the plasma fractionated in the world [6].

Both recovered and source plasma is suitable for the production of all plasma-derived products. Recovered plasma contains less coagulation factors than source plasma possibly due to (i) longer processing time before freezing, (ii) higher relative dose of anticoagulant added and (iii) higher cellular contamination potentially releasing enzymes that could lyse coagulation factors. On the other hand, source plasma contains less IgG when collected from frequent donors. The use of filtration membrane (plasmapheresis) or leukoreduction on charged filters (recovered plasma) could lead to more activated complement component 3 and 5 [6].

In some countries, like the US, China, Russia or Germany, blood donations are remunerated to ensure greater supply. Nevertheless, WHO recommends voluntary non-remunerated donors [48]. Turning blood into a commodity may incite some donors to hide risk factors or diseases to earn the remuneration [49]. Source plasma from frequent remunerated US donors also show reduced total protein and total Ig levels and somewhat altered protein composition, as compared to unpaid donors [50].

Specific hyperimmune immunoglobulin are produced from high-tittered plasmapheresis plasma collected from selected successfully immunized donors [7,33].

1.3.1.2 Plasma processing

According to the European Pharmacopoeia, plasma should be frozen at a temperature of -25 °C or below within 24 hours of collection. Plasma can be kept frozen for several months or years but storage temperature should be as constant as possible, including for transportation to the fractionation plants [51].

Once delivered to the fractionation plants, plasma is quarantined for some quality control testing (see section 1.3.2 hereafter). Thousands of donations are pooled together to ensure homogeneity and representativeness of the normal polyclonal immune repertoire. The European Pharmacopoeia requires a minimum of thousand individual donations to be pooled [52]. For a same plasma pool volume, the number of pooled donations is greater for recovered plasma because the mean volume of a single donation is smaller.

1.3.1.3 Plasma fractionation

Nowadays, plasma-derived products are manufactured in fractionation facilities that are designed, built, and operated in compliance with high level Good Manufacturing Practices (GMP) and Standard Operating Procedures (SOP) (see also section 1.3.2 on safety and quality control). Plasma fractionation historically relies on two core technologies: cryoprecipitation and cold ethanol precipitation.

Slowly thawing frozen plasma leads to an insoluble precipitate called cryoprecipitate which is enriched in human coagulation factor VIII, fibrinogen and von Willebrand factor. In most processes, that cryoprecipitate is separated from the cryosupernatant by continuous refrigerated centrifugation and stored below -30 °C until further pooling and processing [6,7,16].

The thawed plasma (the cryosupernatant if cryoprecipitation was performed) is fractionated by cold ethanol precipitation as developed by Cohn, Oncley and colleagues, so that plasma proteins are separated according to their solubility [3,4]. Ethanol concentration, pH, temperature, ionic strength and protein concentration are important parameters. The separation method is thus stepped into successive precipitations at defined ethanol concentrations (typically from 0% to 40%), with pH shifts from neutral to 4.8 at temperature from 0°C to -8°C [6,7]. At each step, precipitated proteins are removed by centrifugation or depth filtration. Figure 1.2 gives a general overview of the common plasma fractionation flow that leads to the main plasma-derived therapeutics.

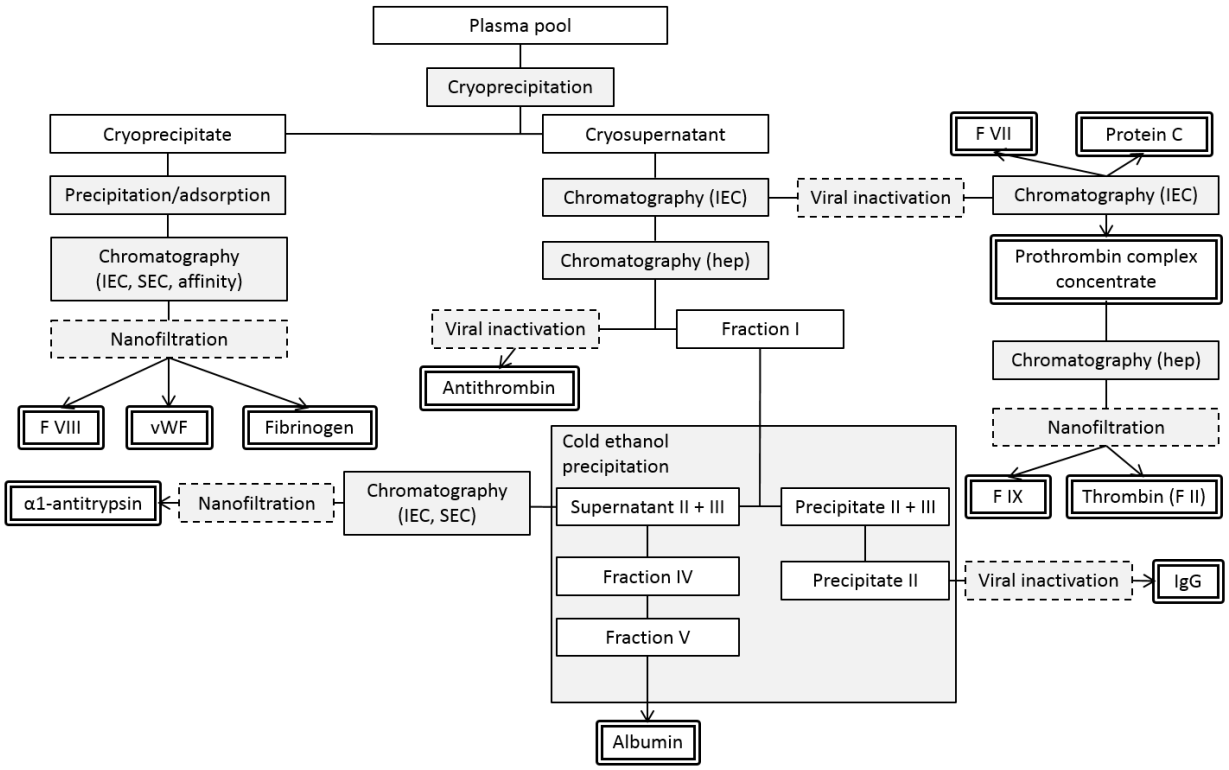


Figure 1.2 Typical plasma fractionation scheme leading to the main plasma-derived therapeutics (based on [6,7]**). Chromatography methods are: Ion-exchange (IEC), size-exclusion (SEC), heparin affinity (hep).**

In order to obtain products as pure and as stable as possible, additional isolation steps are implemented in the process, on the cryoprecipitate, the cryosupernatant (before ethanol precipitation) and the cold ethanol fractions. Chromatographic methods are mainly used, namely affinity chromatography, ion-exchange chromatography (IEC) and size-exclusion chromatography (SEC). Commonly used affinity ligands are immobilized heparin and monoclonal antibodies [6,7]. For examples, the so-called “fraction II”, namely the IgG precipitate obtained by cold ethanol fraction in the Cohn-Oncley process, can be subjected to mild acid (pH 4) treatment in the presence of traces of pepsin (about 1:10000) and purified by diethylaminoethyl ion-exchange chromatography [16,25].

Finally, different virus inactivation methods are implemented, mainly solvent/detergent (S/D) treatment, nanofiltration, caprylic (octanoic) acid treatment and pasteurization. Quality control and microbiological safety are further discussed in the section 1.3.2.2.

Taken together, all those process variants result in a large diversity of fine-tuned production methods that may result in a range of products from different manufacturers with slightly different composition and/or characteristics (see also section 1.4.2) [25].

1.3.1.4 Immunoglobulin products formulations

Highly concentrated liquid products at acid pH (around 4.5-5.5) show good stability and are well tolerated by patients [16]. Typical IgG concentration in commercial products is 50 or 100 mg/mL (corresponding to 5 or 10% respectively) which allows reasonable time of infusion. Purity of IgG in the protein fraction has to be over 95% (See section 1.3.2 for quality control guidelines).

Different possible stabilizers like polyols (sorbitol), sugars (maltose, glucose) or amino acid (glycine, proline, isoleucine) are usually added [16,53]. Sucrose was abandoned following observed negative effect on kidney function, including increased risks of fatal renal failure [16,53]. Variation in stabilizers and diluents contributes to relative uniqueness of each product, in terms of tolerability. For example, some people can be allergic to latex or intolerant to sorbitol, two possible additives [38].

Some immunoglobulin products are lyophilized in powders, requiring careful reconstitution with water before use. Aqueous diluent is often saline (up to 0.9%) and/or supplemented with sugars (at 1-5%) to minimize aggregate formation. Lyophilisation increases the product stability and thereby extends its shelf life; it also increases temperature tolerance (see following subsection 1.3.1.5). Too vigorous shaking for reconstitution may denature the protein. Caution should be also paid to the sodium and sugars content resulting in the osmolality of the reconstituted product, according to the diluent used and the Ig concentration desired (generally ranging from 3 to 12%) [42,53]. Liquid formulations are generally more convenient for patient and practitioner, because of the easier preparation and administration [54].

Specific IgA-free products are reserved for patients with selective IgA deficiency which leads to possible anaphylactic shock in the presence of IgA [55].

1.3.1.5 Immunoglobulin products stability

Contrary to packed red cells or platelets which are labile and perishable, plasma-derived proteins are stabile products with much longer shelf life. Liquid immunoglobulin product has a mean shelf life of 2-3

years when refrigerated (4-8 °C). Low pH favours product stability. Lyophilized product can be stored at room temperature (max. 25 °C) for up to five years [7,38].

Still, sugar excipients may induce undesired covalent attachment of sugars, the so-called “glycation”, during storage and affect protein structure and function [56,57].

1.3.2 Safety and quality control of human plasma-derived immunoglobulin

Nowadays, Ig products released on the market have reached high level of overall safety. Given that the product is properly infused and the patient is well hydrated, most adverse reactions are mild and transient, including headaches, fever, chills, diarrhoea, flushing, nausea, fatigue, blood pressure changes and tachycardia. Severe adverse reactions are rare and are mainly related to risk factors, product selection and/or incorrect administration [55,58–60].

This safety has to be associated with the implementation of a strict quality system within the modern plasma fractionation. Plasma collection centres are licensed and inspected by National Regulatory Authorities (NRA) and audited by plasma fractionators. Plasma fractionators, in turn, are also licensed and inspected by NRA. In Europe, information about collection procedure, testing, handling and transportation of the plasma for fractionation are assembled into a ‘Plasma Master File’ (PMF). Quality control (QC) tests to perform at the level of both the plasma pool and the final Ig products are defined and controlled by the Official Control Authority Batch Release (OCABR), in accordance with the guidelines and GMP referenced by international organizations, like the European Pharmacopoeia (Ph. Eur.) from the European Directorate for the Quality of Medicines & Healthcare (EDQM) of the Council of Europe or the World Health Organization. The applied norms and the techniques used to assess them are registered in a marketing authorization dossier for each product. Some of the QC tests are independently verified by an Official Medicines Control Laboratory (OMCL) that performs batch release and certification [16,51,52,61–64].

Table 9.1 in the Addendum (Chapter 9) is listing the Ph. Eur. guidelines as well as the tests verified by the OMCL. The quality control aspects can be divided into four main categories: the microbiological safety, the immunoglobulin content and its integrity, the measurement of specific antibodies and the purity and residual composition. Additional aspects include physico-chemical properties, formulation & labelling, tolerance, traceability and stability.

1.3.2.1 Highly concentrated preparations with intact Ig

As introduced in section 1.2.2, early manufacturer’s target was to obtain high tittered Ig products that still preserve structural and functional integrity. Nitrogen determination with sulphuric acid digestion allows controlling the total protein level, while immunoelectrophoresis and zone electrophoresis are recommended to verify respectively the identity and the proportion of IgG in this total protein level. During the process development, manufacturers should also demonstrate that 1 viral and 1 bacterial reference antibodies (with International Standard or Reference Preparation available) were enriched in the final Ig, as compared to the starting plasma pool. Immunoglobulin G should represent more than 90% of the protein fraction, 95% in the case of intravenous administration. Polymers and aggregates are controlled using size exclusion chromatography, and the related anticomplementary activity is titrated. Preserved Fc function is assayed by incubating the sample with antigen-coated RBC in the presence of

complement and measuring the haemolysis reaction [52,62–68] (see also Table 9.1). Finally, subclass distribution should be defined and ideally not far from normal physiological balance (IgG1 60%; IgG2 30%; IgG3 and IgG4 5% each), as some fractionation steps can lead to alteration: affinity chromatography with protein A preferentially binds IgM and IgG isotypes but not the IgG3 subclass or ion-exchange chromatography and ethanol-based purification can decrease IgG_4 level. Radial immune diffusion or immunoprecipitation techniques combined with infrared nephelometry and polyclonal monospecific antisera can be used to determine subclass distribution [16,52,62,63,69].

1.3.2.2 Microbiological safety

Since plasma can be contaminated by various blood-borne pathogens, ensuring microbiological safety of plasma-derived therapeutics has always been of first concern. Susceptibility is increased by the pooling of thousands different plasma donations. Some historical examples evoke the dramatic possible consequences, such as the hepatitis infection of 23.000 soldiers during World War II due to contaminated human serum used as stabilizer for vaccines [70]. The deaths of haemophiliacs in the US consequent to HIV transmission through donated blood achieved to make microbiological safety the first driving force of the plasma-related therapeutics QC. It reached successful achievement: there has been no transmission of HIV, HCV and HBV by IVIg administration since 1994 [7].

Donor selection (medical history scrutiny and absence of infection symptoms or potential risks) and screening procedure (genomic viral markers or antibodies to HIV, HCV and HBV) at the levels of both individual donations and plasma pools already greatly reduce the infectious risk. Nevertheless, the infection zero risk cannot be achieved in transfusion, especially with the increasing population mobility and in the less developed countries. Of note, bacterial contamination is not a serious threat, because of the plasma freezing and cold chain that preclude bacteria of replicating and surviving and the easy implementation of a bacteria retentive filtration. To address viral contamination risks, dedicated inactivation/removal steps were implemented in the manufacture process, consisting of pasteurization (heat treatment), S/D treatment, acid-pH treatment or virus filtration. Those viral reduction steps do not exempt from prior plasma donations and pools testing, since their effectiveness is still conditioned by the initial viral load and to a lesser extent some manufacture parameters (pH, incubation time, protein and solute content…). Moreover, equipment failure or human error at some manufacturing stage can never be totally excluded. Anyway, combination of two or three of those specific viral reduction steps in each IgG product manufacturing process, added to a last final product verification (sterility, bacterial endotoxins and antibody to surface antigen of HBV - HBSAg) ensure unprecedented margin of microbiological safety [7,16,49,51,52,62,63,71].

1.3.2.3 Ensuring appropriated product administration

Tests in animals and evaluation during clinical trials should ensure that the product is well tolerated when administered in the correct way. Reconstitution (if freeze-dried product) and administration instruction should be stated on the product label. Product stability within referred shelf-life is ensured by maintaining the cold chain through the whole process and respecting storage conditions, while the OMCL verifies the solubility (if freeze-dried product) and the appearance. The pH, the water content and the osmolality are other physico-chemical properties that are controlled by the manufacturers [52,62–64].

Some contraindications can be also linked to additives in the final formulation, namely antimicrobial preservatives (only accepted for multidose intramuscular Ig) and stabilizing agents. For example, acute renal failure can be linked to osmotic injury with sucrose-stabilized products. The excipients should be specified on the marketing authorization and should not exhibit any deleterious effect in the amount present. The anticoagulant added during the plasma collection is specified in the Plasma Master File [51,52,61–63].

Still in the perspective of selecting the appropriated product, IgA maximal content should be specified, so that IgA-deficient patients can opt for low IgA or IgA-free products and avoid any risk of anaphylactic reaction [52,62,63,72].

In the case of specific hyperimmune Ig products, each product should be demonstrated to contain a sufficient level of the particular antibody for which it is referred. Suitable immunoassay is used to determine the potency by comparing the antibody titre of the product to be controlled with that of a reference calibrated in International Units. The estimated potency should not be less than the stated potency, which is in turn higher than a fixed threshold [72–82].

Finally, any unexpected safety issue with given batches of Ig products could be rapidly contained thanks to an efficient traceability of both the plasma pools and the product released on the market. The OMCL collects and verifies information about production steps and controls respectively through the Plasma Master File and the protocol submission [61,64].

1.3.2.4 Recent safety issues

Over the last years, the focus for Ig products safety and QC has shifted from pathogen transmission risks, which are now greatly secured, to other problems, namely the increase in frequency of haemolytic and thrombotic events [25,83].

a) Haemolysis and anti-A, anti-B and anti-D antibodies

Severe haemolytic reactions were reported with high dose administration of IVIg in patients with blood group A, B and AB, suggesting an association with the anti-A and anti-B antibodies present in the purified IgG mixture. Anti-A and anti-B antibodies are of the IgM and IgG types. The IgM ones appear in the early days of life and are called *haemagglutinins* because they agglutinate the RBC. The IgG ones are secreted in the case of a humoral immune response and activate the complement, inducing haemolysis. They are called *haemolysins*. There are indications that recent optimisations in the purification process may have increased the anti-A titre up to two fold [84–86].

In the “direct method” recently prescribed by the Ph. Eur. for both IV and ScIg, test samples are incubated with papain-treated RBC with A or B antigen. The papain treatment allows the visual observation of the haemagglutination in case of positive result since it reduces the negative charges at the cell surface and thus diminishes the distance between RBC. A maximal titre of 64 is specified, meaning that the highest sample dilution that can give agglutination is 1:64 (initial concentration being set at 25mg/mL) [52,63,84,86,87].

Anti-D antibodies present in normal immunoglobulin products could be related to some haemolytic reactions at high dose injection in Rhesus positive (RhD+) receivers. These antibodies, which are normally absent in their phenotype, will react to the Rhesus factor proteins present on the surface of the RBC. Many manufacturers have thus set a limit for the level of anti-D antibodies. Ph.Eur. recommends a titration of anti-D antibodies based on incubation with papain-treated RhD+ or RhD- RBC, observing the limit sample dilution for agglutination with RhD+ RBC. The titre of the preparation must not be greater than fixed positive control (starting from the same concentration) [84,88].

For the particular case of specific anti-D Ig products, high level of anti-D antibodies is precisely expected and anti-D potency is assessed with an ELISA [73,89].

b) Thromboembolism and coagulation factors

Normally, thromboembolic complications are mainly of concern for patients of advanced age, with previous thrombosis history, hypertension, dyslipidemia, renal insufficiency or diabetes mellitus, and receiving high dose therapy at rapid infusion rate [55,83].

However, in 2010-2011, an unexpected increase of thromboembolic events (TEE) was reported with some Ig products administration, including stroke, myocardial infarction, pulmonary embolism and deep vein thrombosis. Detailed investigation in incriminated immunoglobulin batches identified increased concentrations of plasmatic procoagulant clotting factors as the main cause of increased TEE risks [90].

Previously, prekallikrein activator (PKA), being a proteolytic fragment of Coagulation Factor XII (FXII), had already demonstrated sufficient risk to justify a dedicated chromogenic assay in the case of IVIg. PKA activates prekallikrein to kallikrein which liberates the vasoactive peptide bradykinin from kininogen, possibly leading to hypotension when present in administered IVIg. In patients with compromised cardiac function, kallikrein could have adverse vasomotor effect [16,91].

This time, activated coagulation factor XI (FXIa) appeared as the predominant marker as there was clear correlation between trace level of FXIa, thrombin generation and TEE [90–92]. Apolipoprotein H (ApoH), also called beta-2-glycoprotein 1, from the clotting cascade appeared as a potential additional contributor to the procoagulant activity and TEE risks [93]. Finally, kallikrein could play a minor role as it is a common plasmatic residual contaminant in IVIg preparation. Of note, as kallikrein can cleave the FXII into PKA, it is also involved in hypotension events [91,94].

The TEE outbreak led to Ph.Eur. revision stating that IVIg and ScIg production method should yield to a product that “does not exhibit thrombogenic (procoagulant) activity”. No specific batch release testing was indicated but all Ig manufacturers now have to prove that their production process includes a step or steps “that have been shown to remove thrombosis-generating agents with emphasis to the identification of activated coagulation factors and their zymogens and process steps that may cause their activation” [52,63].

1.4 The importance of residuals characterization

Zone electrophoresis and immunoelectrophoresis as recommended by the Ph. Eur. only allow rapid verification and global view on the Ig products purity. The non-Ig plasma proteins are visualized as other-than-the-principal bands on the zone gel electrophoresis and the immunoelectrophoresis but they are not identified and even less quantified. Moreover, some non-IgG proteins could be hidden in the principal band, due to co-migration.

In all the cases, reaching an effective purity of more than 90-95 % in immunoglobulin, as recommended by the Ph.Eur., still means having potentially several percentages of non-immunoglobulin residuals. They are certainly not problematic *per se* but constitute a shadow zone that may in some cases impact the product's safety as observed in the recent events.

1.4.1 The 2010-2011 thromboembolic events outbreak

1.4.1.1 Description

The overall reported average incidence of IVIg-induced thrombosis ranges from 0.5 to 15%, with both venous and arterial ischemia events. Those adverse events are thus as severe as they are rare, considering that higher incidence is generally linked to patient-related risk factors. The way Ig administration induces TEE is still not completely elucidated but three main mechanisms have been evidenced: (i) an increased blood viscosity causing hypercoagulability, (ii) the passive transfer of anticardiolipin antibodies and (iii) the passive transfer of highly active factors [83,95].

In 2010, the European Medicines Agency (EMA) has temporarily suspended the marketing authorisation for the IVIg product Octagam (Octapharma GmbH) and requested a recall of Octagam on the market at that time in Europe. Simultaneously, the withdrawal from the US market was notified by the Food and Drug Administration (FDA). This recommendation followed notification by the German and Swedish OMCLs of an unexpected increase in reports of TEE consequent to administration of this product. The presence of procoagulant activity was evidenced in various batches [91,96,97].

In 2011, new TEEs were notified to the Paul-Ehrlich-Institut (PEI, the German OMCL), this time following treatment with a ScIg from another manufacturer, Vivaglobin (CSL Behring). The FDA subsequently emitted important safety information. This outbreak was even more unexpected since TEEs were not previously linked to subcutaneous administration of Ig due to the retarded absorption [90,98]. The same year, the IVIg product Omr-IgG-am (Omrix Biopharmaceuticals) was withdrawn from the Israeli market due to TEE reports again [99].

1.4.1.2 Retrospective analysis

Both American laboratories and European OMCL collected pharmacovigilance and laboratory data to retrospectively analyse this outbreak and better characterize its incidence. The FDA reported 209 TEEs linked to Ig treatment via its Adverse Events Reporting System during the period 2006-2010. The PEI analysed 228 reports of TEEs associated with six different Ig products over the period 2006-2011, 198 with IVIg and 30 with ScIg. A total of 100 cases with IVIg and 10 cases with ScIg had an onset within maximum 48 hours and were categorized by the PEI as drug related. Among these, 74 patients had been treated

with Octagam and 8 with Vivaglobin. This clearly evidences that the outbreak was specifically related to these Ig products. Moreover, the TEE-reporting rate for Octagam administration substantially increased in 2010, while it was not the case for the other IVIg. An increased pro-coagulant potential was demonstrated for the corresponding 2010 Octagam batches. More specifically, FXIa, which had been previously measured as potential IVIg products contaminant, was detected at an increased level that exceeded threshold concentration inducing thrombin formation. FXIa was most likely co-purified with Ig, consequent to process modifications intended to increase Ig recovery or improve viral safety [90,92,94,95].

Octapharma also immediately investigated the biochemical root cause of the incriminated Octagam batches, after having initiated a massive voluntary recall of these. Again, FXIa was unambiguously identified as the root cause of the increased thrombogenicity. At the same time, the other potential contributions, namely microvesicles/microparticles, other (activated) coagulation factors or cytokine and growth factors, were excluded. Kallikrein was detected, but at a level that is negligible for thrombin generation assay [100].

The situation appeared to be less clear for the TEE ScIg. First, the exact definition of drug related TEE onset is less straightforward, due to the time window of the slower absorption. Moreover, the PEI did not observe clear correlation between increased pro-coagulant activity and higher TEE-reporting rate, the FXIa resorption rate being also unknown. Nevertheless, manufacturing process revision as recommended by the Ph. Eur. ensured to have no more TEE reported as drug-related [90].

1.4.1.3 The thrombogenicity of FXIa

FXIa is a serine protease that triggers the middle phase of the intrinsic pathway of blood coagulation by activating factor IX. An overview of the blood coagulation cascade with its most important components is given in Figure 9.2 of the Addendum. FXI is activated by FXIIa, thrombin and FXIa itself. FXI level in healthy individuals is about 5 µg/mL. FXI deficiency can cause mild bleeding tendency; however, FXI replacement should be used sparingly due to potential prothrombotic effects [101,102].

The Wessler stasis model in rabbits is the state-of-the-art *in vivo* assay to measure the procoagulant potential of a sample. The evaluation is based on the excision of the jugular vein, after a 10 min of stasis following sample administration. The thrombus is scored using a scale from 0 to 4 (with 0 meaning no clot, and 4 a complete clot in a vein segment formed by a single thrombus) [94,100].

The Wessler model indicated that the FXIa threshold level inducing thrombin formation is between two and seven mU FXIa per mL for a 5% IgG solution (thus equivalent to 0.04 to 0.14 mU per mg of Ig product). Using FXIa spiking experiments, this threshold level is approximately 3 ng of FXIa per mg of Ig product [100,103].

Since FXIa can activate the zymogen FXI, this zymogen FXI could be inadvertently activated due to a process modification or progressively during storage (“autoactivation”) [92,94].

1.4.1.4 Resolution of the outbreak

Following Ph. Eur. revision (see 1.3.2.4 b)), all Ig manufacturers evaluated both intermediate and final products to validate the ability of their process to remove both procoagulant activity and FXIa contamination, so that routine testing of products would not be required. Different (pre-existing or to-be-implemented) production steps were identified to significantly reduce prothrombotic markers: immune-adsorption, intermediate filtration steps, pasteurization, ion exchange chromatography or octanoic acid precipitation [94,100,103–105].

Corrective and preventive measures to implement in the manufacture process were validated using both in-process and final product testing with sufficient sensitivity. The Wessler stasis model is useful for the validation and qualification of a manufacture process but is not suited for routine testing. Generic coagulation assays like Prothrombin Time (PT) and Activated Partial Thromboplastin Time (aPTT) did not show sufficient sensitivity. Adapted assays had to be designed or adjusted, like thrombin generation assay (TGA) and the Non-Activated Partial Thromboplastin Time (NAPTT) while targeted approaches including chromogenic assay and ELISA were used to quantify FXIa. Those tests demonstrate excellent sensitivity for FXIa : around 1 ng/mL for NAPTT and ELISA and around 0.1 ng/mL for TGA and chromogenic assay [94,100,103,106].

(i) Thrombin-generation assay (TGA)

TGA is a global assay that measures the overall tendency of a sample to generate thrombin after initiation of coagulation. Thrombin formation is monitored through the fluorescence of a peptide substrate for thrombin, after the sample has been mixed with citrated or FXI-deficient plasma, phospholipids, tissue factor, the fluorogenic peptide substrate and calcium as initiator. The read-out parameters (peak height, time to the peak, and lag time) reflect the three phases of coagulation (initiation, propagation and termination) and also the amount of thrombin formed. On the other hand, caution should be paid to reduce tissue factor and lipid concentrations, since they could mask triggering effects in the test sample and modify assay results. Consequently, absolute TGA values comparison requires detailed description of the assay impact factors and reagents, along with an individual validation [94,107].

TGA has been included in the routine batch release testing of Octagam for the absence of procoagulant activity [100].

(ii) Non-activated partial thromboplastin time (NAPTT)

After having mixed the test sample with citrated or FXI-deficient plasma, phospholipids and calcium, the time for clotting to occur is measured (in seconds). NAPTT of regular Ig products is >200s. NAPTT and TGA strongly correlate in distinguishing TEE-positive batches. Nevertheless, there is no compendial NAPTT method for Ig testing. No threshold value has been validated to ensure compliance with the Ph.Eur [94,106].

(iii) Chromogenic assay

Chromogenic substrates can be used for serine proteases determination since the latter cleave the substrate proportionally to their activity. Here, the product formation from FXa-specific chromogenic

substrate is measured by absorbance, since the activation of FX to FXa is directly related to the amount of FXIa [94].

(iv) ELISA

FXI can be measured by a sandwich-type ELISA with affinity-purified antibodies specific to FXIa as capture antibody and horseradish peroxidase anti-human FXI polyclonal antibody as detection antibody. Optimal antigen binding for accurate quantification should be optimised with the capture antibody concentration and the sample dilution [94].

Consequent to the process revisions, both the manufacturers and the PEI demonstrated a normalization of the thombogenic activity as well as a reduction of the FXIa level below its thrombogenic threshold. At the same time, the TEE reporting rate of both Octagam and Vivaglobin decreased to reach standard values of other Ig products, with no TEE evidenced as drug-related. This confirmed that both the evidenced pathomechanism and the correction measures to resolve the outbreak were relevant [90,100]. In May 2011, the EMA recommended the lifting of the suspension of Octagam [108]. In 2012, Omr-IgG-am was also cleared to return to the market [109].

Still, in 2013, the FDA mandated that the label on the IVIg products includes a prominent warning box about the risk of TEE [95]. In 2017, FDA recalled that “based on the available data, the TEE risk with Ig products is small but not zero. Accordingly, continued vigilance on the part of the regulatory authorities, the manufacturers, the healthcare providers and the patients is still warranted” [110].

1.4.1.5 The need for more?

The TEE outbreak thus revealed that the PKA assay, which only identifies a single component of the coagulation system, was not sufficient to properly characterize the thrombogenic activity in Ig products.

The outbreak also highlighted that even plasma proteins not belonging to top abundant ones can end up in the final product, due to co-purification issues during the manufacturing. Moreover, these impurities can have significant health effects at trace levels.

Among the OMCLs network, there was a debate whether testing for pro-coagulant impurities should be specified and detailed in the release specifications and in the Ph. Eur. Precise recommendations for batch testing rely on harmonized test methods and congruent results. TGA and NAPTT became the unofficially recommended tests but are still not explicitly mentioned in the Ph. Eur. today. Enquired process revision and verification demonstrated to be sufficient to avoid further drug-related TEE [90,94].

Still, in 2014, two studies identified respectively ApoH [93] and CD154-related proteins [111] as potential contributors to TEE risk in Ig products. While the first is directly implicated in the coagulation process, the latter could be involved in the mediation of monocyte-platelet aggregates and blood cell activation.

Considering the nature of its biological origin, it would be utopic to obtain totally pure plasma-derived Ig. Nevertheless, since impurities can have significant impact at trace levels, this episode raises the question whether corrective and targeted resolution of impurity issues could be substituted by a preventive and generic QC approach.

1.4.2 Diversity of commercial Ig products

Prior to the 2010-2011 TEE outbreak, few studies had already emphasized that residual proteins from the plasma source might lead to batch-to-batch variations [112]. Impurities might influence the anti-complementary activity (ACA) outcome [113], IgG fragmentation could be caused by proteases such as plasmin and kallikrein [114], and FXIa had been already detected as contaminant [92].

The TEE outbreak of 2010-2011 very clearly demonstrated that, even within an elaborated regulatory and QC framework, selecting one Ig batch instead of another can have dramatic consequences for the receiver. Ig products differ in source material origin and manufacturing process, and so the final composition can differ.

Pooling a large number of donations and applying strict microbiological safety measures have largely contributed to significantly reduce the variability coming from the plasma pool material. Still, the collection method (source plasma or recovered plasma), the donor remuneration status and the donation frequency all can affect the protein composition of the pool (see section 1.3.1.1).

Both the product demand and the QC requirements can continuously push manufactures into revising their processes and introducing modifications. Increased productivity, purity, absence of aggregation, viral safety, particular needs for particular receivers (e.g. IgA-deficient patients) or recently reduction of the procoagulant activity and the FXIa level: all these criteria have triggered further process and/or formulation modifications. Considering that each criterion can be reached through different means (see for example section 1.4.1.4), this inescapably results in a diversification of the Ig products on the market.

As a result, some differences in the biological and biochemical properties can be evidenced among Ig products on the market. Of note, the effective impact on the treatment efficacy and safety is not always easy to decipher. Better understanding of the mechanisms of action could help providing objective criteria to compare different products. Nevertheless, careful product selection is recommended. This product selection should integrate information about receiver's susceptibility and risk factors as well as the most detailed information possible that is available on the product. This information should be always reflected into innovative QC guidelines to ensure compliance with associated criteria [25,42,53,54,116,117].

Table 1.3 is summarizing the main product features that can contribute to the diversity of Ig products in the current regulatory and QC framework. While IgA univocally refer to a single characteristic (the presence of an alpha type heavy chain) that is easily assayed, formulation features are intentionally implemented and can be thus easily measured and referenced on the product label. On the other hand, IgG content and plasma residuals can both intrinsically contribute to (slight) differences in Ig products, for which current QC guidelines only propose mitigation.

Table 1.3 Product features potentially contributing to biodiversity within the current regulatory & QC framework. The potency concerns the hyperimmune specific Ig. (Partially based on [42,115]**)**

Product feature potentially contributing to diversity	**Patient risk factors**	**Risk control or *mitigation* (see also Table 9.1)**
IgA content	IgA deficiency (anti-IgA antibodies)	**Immunochemical method**
IgG content		
Potency	/	**Potency tests**
Anti-A, anti-B, anti-D antibodies	A, B, AB and RhD+ blood groups	**Direct agglutination methods**
Other specific IgG?	?	*Diversity and representativeness (at least 1000 donors per plasma pool)*
Plasma residuals		
Procoagulant proteins	thromboembolic history, diabetic, renal dysfunction, elderly	**Procoagulant activity and FXIa follow-up + specific removal steps**
PKA	cardiac impairment	**PKA chromogenic assay**
Other residuals?	?	*Zone electrophoresis and immunoelectrophoresis*
Formulation		
Sugar content	renal dysfunction, (pre)diabetic, elderly, intolerance	**(measured and) stated on the product label**
Sodium content	neonates, elderly, cardiac impairment, renal dysfunction, thromboembolic history	
Osmolality	neonates, elderly, cardiac impairment, renal dysfunction, thromboembolic history	
pH	neonates	
Concentration	neonates, elderly, cardiac impairment, renal dysfunction, thromboembolic history	
Excipients	allergies	

1.4.3 Towards a new QC paradigm?

From its early days in the 1930's until today, overall safety of plasma-derived products has undoubtedly made enormous progress to reach an unprecedented level. Still, unpredicted safety issues recently emerged from the relative products heterogeneity, at the levels of both IgG and non-IgG contents (see section 1.3.2.4). Current regulatory framework was unable to anticipate and prevent these adverse events prior to the introduction of the concerned batches on the market. So far, QC guidelines are mainly corrective and address already observed issues point-to-point.

Over the last decade however, new technology fields emerged with new promises: this is what can be called the "omics" era. "Omics" technologies aim to provide universal and comprehensive approaches to detect genes ("genomics"), mRNA ("transcriptomics"), peptides and proteins ("proteomics") and metabolites ("metabolomics") and by this adopting a holistic view of a biological sample. Trying to apply this field to QC means a shift of paradigm: moving from traditional hypothesis-driven, corrective and targeted approaches towards non-biased, preventive and non-targeted approach.

Genomics approaches already entered the world of plasma-derived therapeutics QC through the use of nucleic acid-based amplification tests (NAT) to detect viral genomes [51,61]. Massive next-generation sequencing (NGS) could provide further insights in pathogen organisms' detection and identification [118].

Proteomics approaches by using mass spectrometry (MS) could provide benefits for better characterization of IgG but also non-IgG components, thus lifting the veil on these reservoirs of heterogenity and by this preventing QC issues like the ones recently observed [112,119].

Chapter 2:

Technical background

2. Technical background

2.1 The era of MS-based proteomics

Over the last 20 years, research and developments in biological sciences have undergone prodigious evolution. Technological advances have opened the path to diverse large-scale and high-throughput applications, while preserving or even increasing accuracy. Mass spectrometry (MS) is definitely one of the pillars of these advances. In the sequel of the precursor work of J.J. Thomson at the dawn of the twentieth century, MS first provided insights for the analysis of volatile compounds in the 1960s. In the 1980s, the technique became available for non-volatile compounds, even with high (more than hundred thousands of Daltons) molecular masses, such as proteins. During the last two decades, progress in the field took off with entirely new instruments developed to meet the needs for proteomics [120].

Despite the current diversity of available technologies and devices, a mass spectrometer always contains the following elements [120] (Figure 2.1):

1. Sample inlet: the compound to analyse is introduced, directly, or via a separation device.
2. Ionization source: ions are produced from the sample.
3. Mass analyser(s): the ions produced (called precursor or MS1 ions) are separated according to their mass-to-charge ratio. When performing tandem MS or MS/MS, the selected ions are fragmented and the resulting fragments, also called product or MS2 ions, are analysed in a second analyser.
4. Detector: the ions emerging from the last analyser are detected and 'counted' in order to measure their abundance, by converting the ions into electrical signals.
5. Data processing system: the electrical signals from the detector are processed into a suitable mass spectrum that can be either interpreted manually or searched against an annotated database.

Of note, the sample inlet and the ionization source as well as the mass analyser and the detector are sometimes combined respectively. From the source to the detector, the mass spectrometer operates under vacuum.

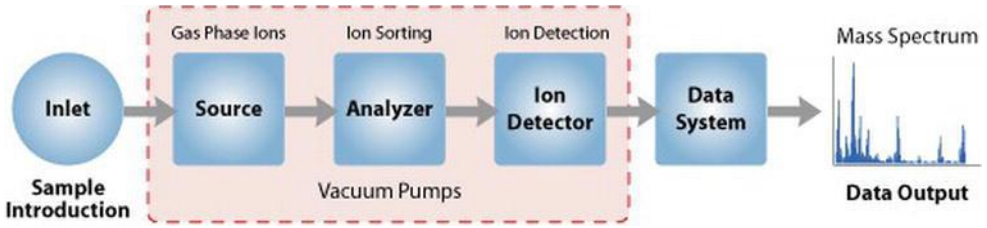


Figure 2.1 Components of a mass spectrometer [121]

Proteomics is the science of measuring proteomes, namely the complete protein profile of a biological sample, being a cell, an organism, blood, food, drugs or any other product or entity containing peptide and/or protein.

MS-based proteomics has emerged as a broadly effective resource for identification, characterization and quantification of proteins, providing a wide range of both biological fundamental insights and clinical applications: system biology analysis, biomarker discovery, epigenetics & post-translational modifications study, microbiological identification, drug discovery or food control, to name a few.

Still, successful application of MS-based proteomics asks to overcome challenges at different stages of the process, from sample preparation to data analysis. So far, MS strategies were mainly used for discovery-based applications or targeted methods.

One major bottleneck concerns the sensitivity. While typical dynamic range of MS detection is 4-6 orders of magnitude, the protein concentration in complex biological samples can span 12 orders of magnitude [122]. Unfortunately, unlike nucleic acids (with the polymerase chain reaction), there is no amplification technique for peptides and proteins. Targeted MS methods, the so-called selected/multiple reaction monitoring (S/MRM) acquisition, allow increasing sensitivity by focusing only on a defined subset of peptides. This however somewhat takes away from the “omics” perspective. As a consequence, sample fractionation prior to the MS injection is frequently mandatory to reach relevant sensitivity, especially with untargeted applications.

Qualification into validated quantitative methods is even more challenging. The unique ionization properties of each peptide prevent direct absolute quantification. In S/MRM approach, stable isotope-labelled peptides corresponding to the targets can be used for absolute quantification, but this again is not feasible in a truly comprehensive proteomics approach.

2.2 Meeting the challenge: proteomics in quality control of plasma-derived therapeutic proteins

As introduced in Chapter 1, plasma is a complex biological fluid and the range of plasma protein concentrations represents at least 10 orders of magnitude. Consequently, it is one of the most challenging types of sample to analyse with MS [123]. Yet the potential of plasma proteomics largely justifies the challenge: biomonitoring and biomarker discovery both take advantage of the high reflectivity of each individual’s general health state - allele differences, metabolic risk and inflammatory status - into its plasma proteome [124]. The combination of various protein separation techniques with MS analysis, either gel-based or gel-free with chromatography, provided an unprecedented view on the human plasma complexity. The Human Plasma Proteome Project (HPPP) from the Human Proteome Organization (HUPO) allowed to build a Human Plasma Peptide Atlas with a set of 1929 protein sequences at a false discovery rate of 1% [125].

MS-based proteomics could equally play a role in plasma fractionation and quality control of plasma-derived therapeutic proteins. As discussed in the previous chapter, plasma-derived therapeutics on the

market still contain impurities that are poorly characterized and can eventually cause QC issues. In addition, the diversification of manufacturing processes questions the exact equivalence between products. One could logically expect for MS-based proteomics to help in both process validation and final product quality control. This QC includes the detection of impurities as well as the determination of product-to-product or batch-to-batch variations but eventually also the characterization of the active component [112,119,126,127]. Table 2.1 is summarizing how MS-based proteomics has informed plasma fractionation so far.

Table 2.1 How MS-based proteomics has informed plasma fractionation and plasma-derived products quality control so far (partly based on [112,128]**)**

Plasma-derived product	MS-based proteomics workflow	Output	Year	Refe-rence
Intermediate fractions	Anion-exchange chromatography; in-gel & in solution digestion; LC-ESI-MS/MS; QTOF, DDA	Fractionation process optimization	2008	[129]
	Anion-exchange chromatography; in-gel & in solution digestion; LC-ESI-MS/MS; QTOF & IT, DDA		2009	[130]
Human Serum albumin	CPLL; in gel digestion; LC-ESI-MS/MS; IT, DDA	Impurities profiling : haptoglobin, hemopexin, alpha2-HS glycoprotein, ferroxidase, afamin precursor, beta-2-glycoprotein-I, peptidoglycan recognition protein, L precursor, C1 esterase inhibitor, alpha-1B-glycoprotein precursor, transthyretin, haemoglobin chain D, vitamin D-binding protein, immunoglobulin lambda chain	2006	[131]
	Size-exclusion chromatography; in-gel digestion; MALDI-MS/MS; TOF, DDA	Active component characterization : lipid ligands	2007	[132]
Antithrombin	Electrophoresis, in gel digestion; LC-ESI-MS/MS; QIT, DDA	Active component characterization : glycans pattern analysis	2004	[133]
	in solution digestion; ESI-MS; QIT DDA or MALDI-MS; TOF DDA		2005	[134]
	Electrophoresis, in gel digestion; LC-ESI-MS/MS; QIT, DDA		2005	[135]
Factor VIII/von Willebrand factor	Anion-exchange/size-exclusion chromatography; in gel or in solution digestion; LC-ESI-MS/MS; QTOF,	Impurities profiling for batch-to-batch variations: inter-alpha inhibitor proteins, fibrinogen, fibronectin, coagulation factor II (prothrombin), hyaluronan binding protein 2, kininogen 1, alpha-1-microglobulin	2009	[136]

	DDA + iTRAQ quantification			
	Electrophoresis; in-gel digestion; LC-ESI-MS/MS; IT, DDA	Impurities profiling: inter-alpha inhibitor proteins, alpha-1-microglobulin, fibrinogen, kininogen, galectin 3 binding protein, complement C3b, IgM	2010	[137]
	In-solution digest; 2D LC-ESI-MS/MS; IT, DDA	Impurities profiling: fibronectin, inter-alpha inhibitor proteins, fibrinogen, galectin 3 binding protein, clusterin, apolipoprotein B-100, platelet glycoprotein Ib, vitronectin, complement C3 & C4, albumin, vitamin D-binding protein, Ig fragments, serpin peptidase inhibitor, serine proteinase inhibitor, kininogen 1, lumican, angiotensinogen preproprotein, alpha-2-plasmin inhibitor, plexin-A4	2010	[138]
	Electrophoresis; in-gel digestion; LC-ESI-MS/MS; IT, DDA + Western Blot	Active compound integrity	2011	[139]
Factor IX	in-gel digestion; MALDI-MS; TOF	Impurities profiling: vitronectin	2001	[140]
	in-gel or in solution digestion; LC-ESI-MS/MS; IT, DDA + iTRAQ quantification	Impurities profiling for process validation : mannan binding lectin serine protease, hyaluronan binding protein 2, inter-alpha-inhibitor proteins, C1 inhibitor	2010	[141]
	1) electrophoresis; in gel digestion; MALDI-MS, TOF 2) in solution digest; LC-ESI-MS/MS; Orbitrap, DDA	Impurities profiling for process validation : complement C4, inter-alpha-trypsin inhibitors, kininogen-1, prothrombin, clusterin and vitronectin	2018	[142]
Prothrombin complex concentrate	2D electrophoresis; in-gel digestion; MALDI-MS/MS; TOF, DDA	Impurities profiling & products comparison: fibrinogen (alpha, beta, gamma chains), inhibitors (C1 inhibitor, inter-alpha-trypsin inhibitor, antithrombin-III, vitronectin, vitamin K-dependent protein C, kininogen, protein Z-dependent protease inhibitor), complement factors (C1s, C3, C4, C9, clusterin, factor H-related proteins, factor B), alpha-2-HS-glycoprotein, albumin, AMBP protein, apolipoproteins (A-I, A-IV, E), C4b-binding protein, carboxypeptidase N, CD5 antigen-like, ceruloplasmin, haptoglobin, serum paraoxonase, plasma retinol-binding protein, serum amyloid P, serotransferrin, transthyretin, alpha-1-antitrypsin, Ig kappa-chain C	2006	[143]
Alpha-1-antitrypsin	Isoelectric focusing, in gel digestion; LC-ESI-MS/MS; QTOF, DDA	Active component characterization : glycans pattern analysis	2006	[144]

Inter-alpha inhibitor proteins	electrophoresis; in gel digestion; MALDI-MS, TOF	Active component characterization for process validation	1997	[145]
	electrophoresis; in gel digestion; SELDI-MS, TOF or LC-ESI-MS/MS, QTOF, DDA	Active component characterization & impurities profiling for process validation: prothrombin (Factor II), fibrinogen, von Willebrand factor, vitronectin, Factors IX & X, protein S, plasminogen activator inhibitor	2006	[146]
Immunoglobulin	2D electrophoresis; in-gel digestion; MALDI-MS/MS; TOF, DDA + immunoblotting	Active component characterization : self-reactive repertoire + process optimization : impact of acidic pH	2009	[147]
	in solution digestion, isolation, modification; LC-MALDI-TOF-MS	Active component characterization : glycans pattern analysis	2006	[148]
	in solution digestion, isolation, modification; LC-ESI-MS/MS, Orbitrap, DDA		2013, 2014	[149,150]
	in solution digestion, LC-ESI-MS/MS, QTOF, DDA		2015	[151]
	in solution digestion; LC-ESI-MS/MS, QTOF, DDA		2014, 2016	[56,57]
	electrophoresis; in gel digestion; MALDI-MS/MS, TOF or LC-ESI-MS/MS, IT, DDA	Impurities profiling: beta-2-glycoprotein-I, transferrin, albumin	2014	[93]
C1 esterase inhibitor	electrophoresis; in gel digestion; MALDI-MS/MS, TOF, DDA	Impurities profiling & products comparison: α1-antichymotrypsin, ceruloplasmin	2014	[152]
Factor X	in solution digestion; LC-ESI-MS/MS, Orbitrap, DDA or ESI-MS, QTOF	Active component characterization : post-translational modifications	2015	[153]
Butyrylcholin-esterase	Electrophoresis, in gel digestion; LC-ESI-MS/MS; QTOF, DDA	Active component characterization : glycans pattern analysis	2008	[154]
	Electrophoresis, in gel digestion; LC-ESI-MS/MS; QTOF, DDA	Active component integrity & impurities profiling for process validation: no detected impurities	2011	[155]

The oldest application of MS to the plasma fractionation may be dated back to 1997, when a possible isolation of inter-alpha inhibitor protein from a by-product of the plasma fractionation was evaluated [145]. Still, MS-based proteomics only really started to inform the plasma fractionation and the transfusion medicine from around 2005. Application of modern proteomic technologies led to uncover unexpectedly diverse impurity profiles in human serum albumin (HSA) preparations, prothrombin complex concentrate or Factor VIII/von Willebrand factor concentrates [131,136,137,143]. As example of potential QC concern, some contaminants like serum proteases could alter the very sensitive FVIII molecule by creating neoepitopes that could stimulate undesired immunological complication by the

receiver [112]. This MS-based profiling not only allowed to perform product-to-product and/or batch-to-batch comparison, but also to help manufacturers to tune their purification processes and avoid potentially harmful contaminants, like hyaluronan binding protein 2 in the manufacture of concentrated Factor IX [129,130,141]. MS was also used for active component characterization, i.e. the investigation of the glycans pattern, for concentrated antithrombin-III and alpha-1-antitrypsin [134,135,144].

MS-based profiling of Ig glycosylation (enzyme-mediated post-translational modification) and/or glycation (often undesired random modification) was progressively investigated in circulating serum [156–160], in monoclonal antibodies [161–165], and finally also in plasma-derived Ig, but to a lesser extent, due to the increased analytical complexity caused by the inherent heterogeneity of polyclonal Ig [56,57,148–151]. This provided insights in sugar excipients effect on products comparison [151], Ig products stability [57] and inflammatory modulation mechanisms [148].

Strikingly, while Ig had become the leading product of the plasma fractionation, no proteomic impurities profiling and product-to-product or batch-to-batch comparison had been performed for commercial Ig in 2010 [112]. Previously, only the self-reactive IgG repertoire (directed towards auto-antigens) had been investigated with MS in Ig products on the market [147]. In the aftermath of the 2010-2011 TEE outbreak, a study using MS approach for impurities profiling in Ig products was finally published in 2014. Using two-dimensional gel electrophoresis and MS analysis of the spots of interest, trace amounts of beta-2-glycoprotein I (also called apolipoprotein H or ApoH), transferrin and albumin were identified as contaminants. The potential involvement of ApoH in thrombogenicity was further questioned with parallel TGA tests. The TGA assay values appeared to increase with increased spiking of pure ApoH [93].

The other mainstream production source for plasma therapeutic proteins, namely the recombinant engineering biotechnologies, had equally started to be informed by MS-based proteomics from the mid-90s [166,167]. From around 2014, even more publications about MS application for plasma therapeutics involved recombinant preparations. As already mentioned previously in section 1.2.5, the potential impurities here are bioprocess-related host cell proteins (HCP) arising from cell death or secretion from cells used for recombinant production. While traditionally measured using ELISA based on polyclonal antibodies raised against the host cell, more comprehensive characterization of HCP was obtained using MS [168–173]. MS methods also demonstrated to be useful for in-depth characterization of the active component, namely the monoclonal antibodies or recombinant therapeutic proteins, e.g. to study post-translational modifications or to evaluate biosimilar candidates [93,174–181]. Nowadays, proteomics methods appear to be more widely used for recombinant production than for plasma-derived production [128,182].

Nevertheless, in 2013, a review of proteomics application to transfusion plasma stated that it was only “the beginning of the story” [183]. And indeed, the past years, MS-based proteomics was e.g. used to compare C1 esterase inhibitor concentrates (angioedema treatment) [152], reveal novel post-translational modifications on plasma-derived factor X [153], or investigate the effect of nanofiltration on plasma-derived factor IX clinical lots [142].

Following all these insightful applications, MS-based proteomics is now slowly entering the official QC guidelines like the Ph. Eur. The oldest mention of MS for protein analysis dates back to 2008, for the composition determination of recombinant Human coagulation factor VIII [184]. In 2011, MS was referred for the glycan analysis of glycoproteins [185]. In 2014, it was mentioned as a tool for the active compound characterization of alpha-1-antitrypsin products [186]. In 2017, MS was cited in the recommended orthogonal analytical methods to characterize the various HCP in a recombinant product and by this supporting the development and selection of an ELISA assay [187]. As last addition to-date, in 2017-2018, MS was associated to genomics in the new molecular methods pipeline to substitute in vivo tests for the detection of viral agents [188]. Even if MS is increasingly cited, a reference method or protocol is not defined. The general description of MS as analytical tool dates back to 2008 but has not been updated since then [189]. It is only in 2016 that first basic guidelines for the "interpretation of screening results for unknown peptides and proteins by mass spectrometry based methods" are emitted by the EDQM [190].

The rise of MS-based proteomics in plasma-derived therapeutics QC, and more generally biological therapeutics QC, appears to be conditioned as follows: the technique should provide new insights as compared to QC tests already in place, while preserving the typical requirements for a QC method. Aforementioned applications clearly demonstrate the innovative potential of MS in the QC field: providing insightful and comprehensive characterization of the sample, being the active components or the low-abundant impurities.

2.3 MS-based proteomics strategies for untargeted profiling of low abundant proteins

Reaching sufficient sensitivity and ensuring appropriate repeatability and quantitative accuracy still asks for careful MS strategy development. This is especially true because MS technologies are still in their growing phase, resulting in a diverse and constantly evolving supply of promising but very expensive devices. Below, the main strategies for untargeted profiling of low-abundant proteins (LAP) in samples with high dynamic range are reviewed, and more particularly source material impurities in plasma derivatives.

2.3.1 Protein fractionation

With more than 90% of the sample being the active component(s), plasma-derived therapeutics typically ask for sample prefractionation before MS analysis can be conducted. Otherwise, the most abundant proteins exert a suppressive effect by monopolizing the instrument signal and making the detection of lower abundant proteins difficult, if not impossible. MS acquisition is always coupled to a peptides separation technique (see section 2.3.3). However, this generally turns out not to be enough for samples with concentration dynamic ranges of above 6 orders [122]. To overcome this challenge, two main approaches can be considered: either the selective depletion of the highly abundant active component(s), or the non-specific enrichment of the impurities.

2.3.1.1 Specific depletion of high abundant components

For years, high-abundance protein depletion has been the standard methodology for first line fractionation in plasma proteomics. Previously, albumin could be removed by using a high affinity dye but the method was lacking specificity. Nowadays, a wide range of commercial immunoaffinity ligands are available to perform antibody capture against targeted abundant plasma proteins. Current products enable single-stage separation that removes up to 20 high-abundance proteins. Further efforts have been invested in serial depletions or moderate abundance proteins depletion, e.g. for biomarker discovery [191–193].

When dealing with plasma-derived concentrates however, one must only focus to deplete the active component in the most selective way. Indeed, the other proteins are impurities that should be profiled and thus not depleted. With regard to Ig products, protein G and protein A are two recombinant proteins of microbiological origin with strong binding strength and high selectivity for mammalian IgG [194–196]. Affinity chromatography with protein G - protein A coating is one of the best-defined and most efficient processes to isolate antibodies [197–200].

The main advantages of a depletion strategy are that the low number of generated fractions – flow through and eluate – enables to perform a wide range of downstream analysis and automated systems with higher reproducibility and throughput. On the other hand, the main drawback is that co-depletion of non-targeted proteins can often occur due to non-specific binding interactions [201–203]. Caution should also be paid to possible binding capacity reduction over time [204]. Still, the overall benefit of depletion on the analytical depth has maintained its popularity in MS-based plasma proteomics [192].

2.3.1.2 Non-specific enrichment of low abundant components

Direct specific affinity-enrichment of low-abundant impurities cannot be considered since their exact nature is not known a priori. The use of a large, highly diverse combinatorial (hexa)peptide ligand library (CPLL) has been established as a strong alternative to immunoaffinity depletion to access LAP identifications.

The CPLL methodology utilizes solid-phase library of millions of different hexapeptides which are generated by a modified Merrifield approach and used for affinity-based binding to proteins, mostly via ionic interactions but also hydrophobic associations, hydrogen bonding, and a number of other weak interactions [205,206]. The library diversity is so high that it should contain a ligand to almost all possible proteins. The binding capacity for a given protein is limited by the number of ligands available but largely sufficient to bring protein traces at the detectability level. Highly abundant proteins (HAP) will saturate their ligands quickly and the unbound excess will be washed away. On the other hand, depending on the amount of loading, LAP will be concentrated on their ligands, resulting in a decreased concentration dynamic range (Figure 2.2) [205–208]. CPLL enrichment has proven to be a versatile and powerful tool to access the low abundance proteome of various biological samples [209–212], including plasma impurities in plasma-derived [131] as well as host cell protein impurities in recombinant therapeutics [213]. Numerous protocol variants have been described to optimize the CPLL enrichment, as summarized by Boschetti and Righetti [214,215].

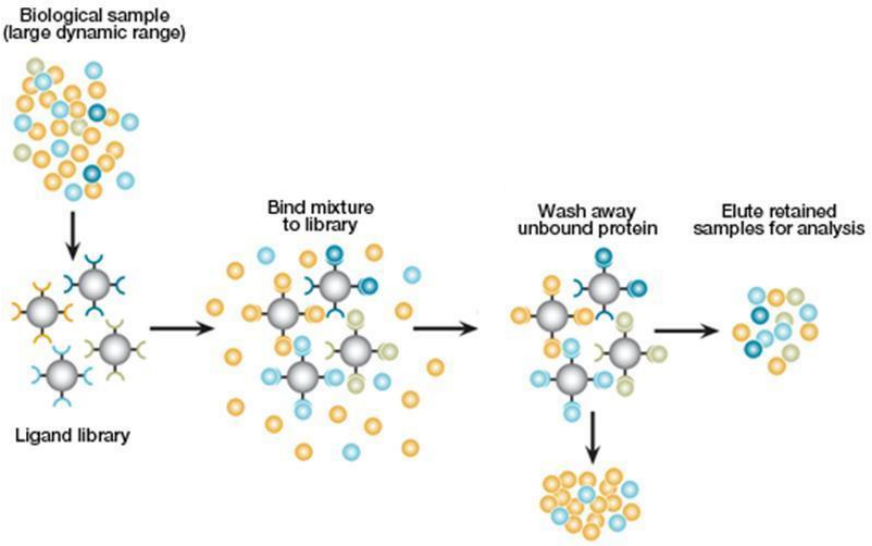


Figure 2.2 Schematic representation of the CPLL enrichment process [216]

2.3.1.3 Gel electrophoresis and other protein fractionation methods

Further protein prefractionation can be performed but many techniques only enrich a specific subset of plasma proteins: glycosylated proteins (glycoenrichment using lectins or hydrophilic interaction chromatography), proteins subjected to proteolysis (N-terminal enrichment using N-terminal-specific enzyme or terminal amine isotopic labelling of substrates) or carbonylated proteins (biotinylation and affinity chromatography) [191,192]. Consequently, these techniques are not useful for untargeted screening.

Protein electrophoretic separation according to physical characteristics such as charge, isoelectric point (pI) or size was one of the pioneer strategies to explore human plasma [217]. Two dimensional electrophoresis on polyacrylamide gel electrophoresis (2D PAGE) emerged as a well-established standard from the earliest proteomic investigations [218]. In the first dimension, proteins are separated according to their pI onto a pH gradient strip in an electric field (isoelectric focusing - IEF) or according to their size, once the native proteins have been negatively charged (blue native (BN) PAGE). In the second dimension, sodium dodecyl sulphate (SDS) surfactant addition leads to denatured and negatively charged proteins that migrate according to their molecular size. Migration profiles are then revealed using staining (coomassie blue, fluorophores ...) and/or antibody capture (Western Blot). Spots can be also excised to be further analysed by MS.

Although gel electrophoresis has been used in countless studies, one must remember that proteins with extremes of size, hydrophobicity (such as membrane proteins) or charge fail to enter the gel and are thus missed. For example, low abundance proteins such as leptin and ghrelin and peptides such as bradykinin cannot be detected using two-dimensional gel technology [219]. Moreover, high throughput and automation are difficult when working with gels [192,220].

Gel-free protein profiling strategies allow to circumvent these pitfalls by combining whole sample digest *in solution*, high-efficiency liquid chromatography separation and high resolution tandem MS. This is the so-called "shotgun" large-scale proteomics [220]. Liquid chromatography separation based on physico-chemical properties could be performed at the protein level [183], but digestion into peptides is usually performed first.

2.3.2 Enzymatic digest

Recent advances in MS instrumentation now allow for the so-called "top-down" analysis of intact proteins, preserving some information such as the connectivity of post-translational modifications. Still, further developments are required to improve both the information content and analysis in top-down approaches. The more commonly applied "bottom-up" approach relies on the digestion of proteins into peptides. These peptides are more efficiently separated using liquid chromatography and their lower molecular mass, an average size of ~500-3000 Da, and lower charge states enable better sensitivity to detect them with MS [221].

Even chemical cleavage is feasible and different enzymes are known to digest proteins with high cleavage efficiency and specificity, yet almost all proteomics studies are conducted using trypsin digestion. Trypsin cleaves peptide bonds C-terminal to the basic amino acid residues Lysine (Lys) and Arginine (Arg), except when followed by Proline (Pro). Since Lys and Arg are relatively abundant and well-distributed throughout a protein, tryptic peptides have an average length of ~14 amino acids and carry multiple positive charges, which makes them ideally suited for MS analysis [221].

Typical trypsin digest protocol includes a denaturation step using a chaotropic agent such as urea or guanidine, followed by reduction of disulphide bridges with dithiothreitol (DTT) and cysteines alkylation by iodoacetic acid or iodoacetamide. After buffer exchange, the trypsin digestion itself is conducted overnight at neutral pH in an ammonium bicarbonate buffer at 37°C [221]. Supplementing trypsin with Lys-C further enhances proteolysis by minimizing the risk of missed cleavages [222].

2.3.3 Peptide separation

Tryptic digestion results in a complex mixture of peptides arising from the different proteins. Current MS devices do not have sufficient resolving power to handle such complexity. Some peptide-based prefractionation may be performed to target specific subproteomes [191]. However, for comprehensive investigation, liquid chromatography (LC) coupled to the MS device has become the golden standard.

Chromatography as separation technique dates back to the early twentieth century, when it was first used to separate plant pigments (*chroma graphein* in Greek: "to write with colors"). Nowadays, chromatography refers to a large range of separation techniques, including thin layer chromatography, gas chromatography, high pressure liquid chromatography or capillary chromatography. In all cases, the sample is dissolved in a mobile phase which will be conducted through a stationary phase. Sample components will then separate according to their respective affinity for the mobile and the stationary phase: the lower the interaction with the stationary phase, the quicker the component flows out (Figure 2.3). The retention time is the time required for a given compound to elute from the column (from the injection in the mobile phase to the elution peak maximum). A broad range of stationary and mobile

phases are available, using different properties to separate the components: the size, the charge, the hydrophobicity or specific affinity.

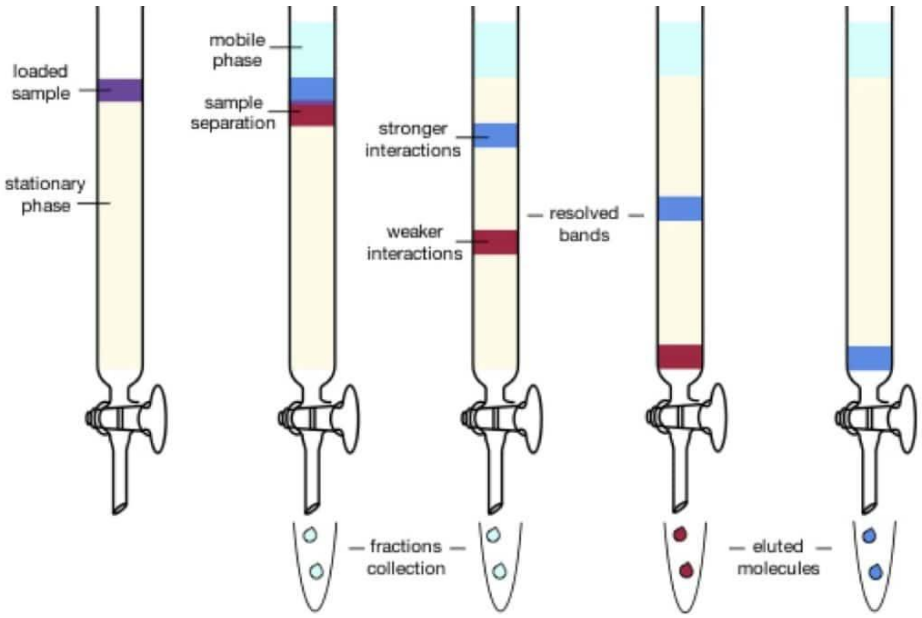


Figure 2.3 Sample separation by column chromatography [223]

The so-called "reversed-phase" LC (RP-LC), where peptides are separated according to their hydrophobic interactions with non-polar groups (typically C_{18} alkyl groups) bound on the stationary phase, is most commonly used on-line to MS [183]. The eluting peptides are directly sprayed into the MS ionization source.

A second orthogonal LC separation like "strong cation exchange" (SCX) can be performed upfront to increase the resolution power. On the other hand, having a larger number of fractions decreases the throughput [192].

Nowadays, LC flow rate ranges from high flow (about 1 mL/min) to nanoflow (about 300 nL/min). Lower flow rates favour better ionisation and higher sensitivity but increase the run time and the risks of instability. Microflow (few µL/min) provides convenient compromise between sensitivity, stability and throughput for routine and reproducible quantitative experiments [224–227].

2.3.4 MS acquisition

2.3.4.1 Ionization

In 2002, Fenn and Tanaka shared a Nobel prize for the introduction of electrospray ionization (ESI) [228] and matrix-assisted laser desorption ionization (MALDI) [229] respectively. ESI and MALDI are still the two main "soft" ionization sources available to ionize high-molecular-weight molecules such as peptides, without fragmenting them.

In MALDI, the sample is co-crystallized with an excess of matrix on a target and irradiated by a laser beam, which produces mainly monocharged ions. Combined to gel electrophoresis, MALDI has provided some

insights for impurities profiling in plasma derivatives (Table 2.1). Still, ESI is often preferred for complex peptides mixtures, since it can be easily coupled to LC, contrary to MALDI.

In ESI, the sample is introduced in liquid phase and forced through a needle held at high voltage (1500-3500 V), resulting in a fine spray of charged droplets that are directed into the vacuum chamber of the MS. The size of the droplets is reduced by solvent evaporation, helped by a drying gas, typically nitrogen. Finally, naked charged ions are generated in the gas phase due to Coulomb fission, consequent to surface tension and electrostatic repulsion (Figure 2.4).

From a sample preparation point of view, it is important to note that MALDI and even more ESI are intolerant to salts, which form sodium and/or potassium adducts and by this decrease ionization efficiency. A desalting step consisting in low cut-off dialysis, precipitation or hydrophobic solid phase extraction is thus usually performed [183].

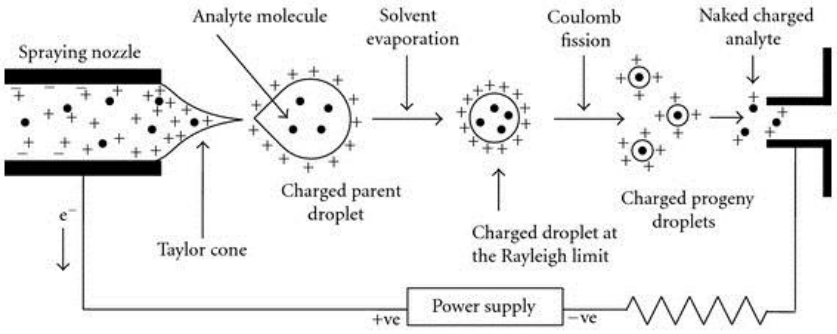


Figure 2.4 Schematic representation of the electrospray ionization process [230]

2.3.4.2 Mass analysers

Mass analysers are the central core of mass spectrometers, as they enable to discriminate the ions based on their mass-to-charge (*m/z*) ratio. They are directly linked to key parameters such as sensitivity, resolution, mass accuracy, scanning speed and dynamics [120,183].

Quadrupole (Q) analysers separate ions according to their *m/z* ratio by using the stability of the trajectories in oscillating electric fields, usually generated by four perfectly parallel cylindrical rods [120]. Quadrupoles are extensively used as triple quadrupole "QqQ" devices wherein the first isolates the precursor ions, the second acts as collision cell and the third separates the generated product ions. These devices offer high mass selectivity, very fast scan speed and 4 to 5 orders of linear dynamic range, which make them mainstay instruments for targeted quantification in S/MRM mode.

For high-resolution non-targeted profiling and global expression studies, modern high performance tandem mass spectrometers rely on high resolving power and high mass accuracy analysers.

TOF analysers are a golden standard for high-molecular-weight proteins, thus fitting the purpose of plasma proteomics [183]. Briefly, the ions are accelerated to known kinetics energies over a known fixed distance, so their time-of-flight can be measured. This time is a function of the mass, so the *m/z* can be

calculated. A reflectron (or ion mirror) retards and then reverses ion velocities so that ions having the same mass but slightly different kinetic energy spend the same flight time [120,183].

Other high resolution mass analysers can be also used for non-targeted proteomics. Ion trap analysers (IT) capture ions in a radio frequency field delimited by hyperbolic shaped, flat rings or linear electrodes. The possibility to perform multiple stage fractionation makes them particularly powerful for structure elucidation. Fourier Transform ion cyclotron resonance (FT-ICR) measures mass by detecting the image current produced by ions cyclotroning in the presence of a magnetic field. Since the frequency of an ion's cycling is determined by its mass-to-charge ratio, this can be deconvoluted by performing a Fourier transform on the signal. FT-ICR has the advantage of high sensitivity (since each ion is "counted" more than once) and very high mass accuracy and thus precision. In Orbitrap analyser, ions are electrostatically trapped in an orbit around a central, spindle shaped electrode. The oscillation generates an image current in the detector plates which is recorded by the instrument. The frequencies of these image currents depend on the mass-to-charge ratios of the ions. Mass spectra are obtained by Fourier transformation of the recorded image currents. Orbitrap have high mass accuracy, good dynamic range and very high resolution [120,183].

Importantly, most contemporary high resolution mass spectrometers are hybrid MS/MS instruments combining different mass analysers, such as quadrupole time-of-flights (Q-TOF) and linear ion trap/quadrupole Orbitraps (LTQ-Orbitrap/Q-exactive) [231]. Despite lower selectivity and scan speed compared to the triple quadrupoles devices, high resolution and high accuracy MS2 scan allows for non-targeted relative quantification (the full precursor *m/z* range is scanned) but also targeted absolute quantification with isotope labelling (predefined list of precursor *m/z* is selected) [231].

In the present work, the ESI-Q-TOF device TripleTOF 5600 (Sciex) will be used (Figure 2.5). This system combines high sensitivity and high resolution, fast acquisition speeds, and stable mass accuracy [232].

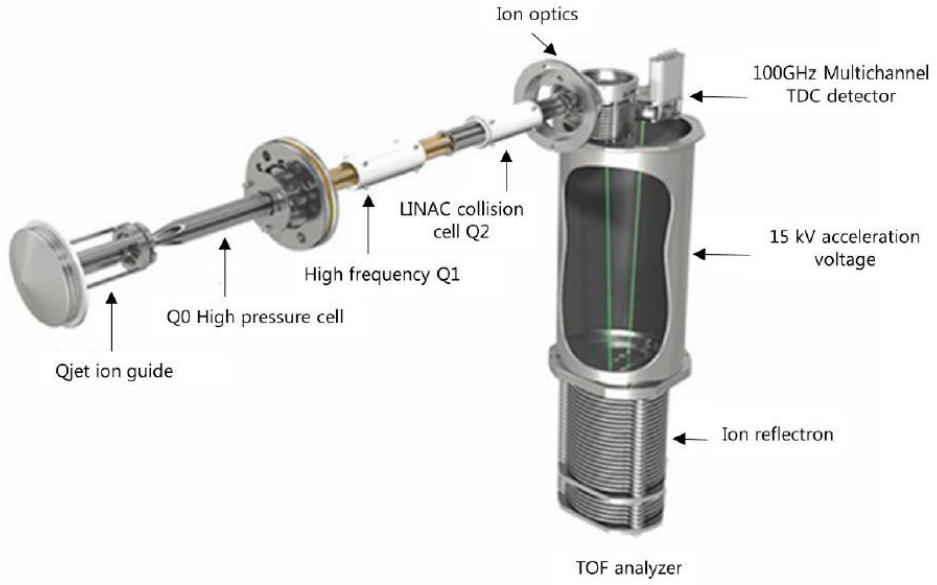


Figure 2.5 Representation of the TripleTOF 5600 QTOF device [105, adapted by Paulien Meert]

2.3.4.3 Fragmentation

Once they have been analysed, peptide-ions (precursor or MS1 ions) are generally further fragmented into product ions (or MS2 ions). Fragmentation patterns obtained by analysing these MS2 ions provide crucial information for peptide to spectrum matches by reaching a higher level of specificity. In most cases, fragmentation occurs by collision-induced dissociation (CID): MS1 ions are fragmented by colliding with neutral molecules of inert gas like Argon that is streamed into the collision cell. Kinetic energy from the collision is converted into internal energy which results in bond cleavage and fragmentation into smaller ions. Bond cleavage mainly occurs through the lowest energy pathways, namely the amide bonds. CID thus generates b-ions when the charge is retained by the amino-terminal fragment and y-ions when it is retained by the carboxy-terminal fragment [233]. Other ions (a, c, x, z) can also be generated when other bonds break in the peptide chain, but this mainly occurs with other fragmentation techniques that are mainly useful for post-translational modifications research and top-down proteomics (Electron Capture Dissociation-ECD, Electron Transfer Dissociation-ETD...) (Figure 2.6).

The so-called MS/MS or tandem MS can be conceived and sequenced either in space or in time. Like QqQ, QTOF are space instruments that combine two mass analysers and one collision cell. During the full scan or MS1 scan, the Q1 is open for all the ions in the MS1 *m/z* range, the ions are not fragmented and the TOF detects all of them. During the product ion scans, the Q1 selects a defined MS1 *m/z* range; these ions are fragmented in the collision cell and the resulting MS2 ions are analysed in the TOF.

The full MS1 scan and the consecutive MS2 scans result in a sequence that is repeated as long as peptides elute from the LC column. The cycle time, or the duration of one iteration of the scan sequence, is set to ensure that enough data points, typically 8 to 10, are acquired during the elution of one peak (LC peak width).

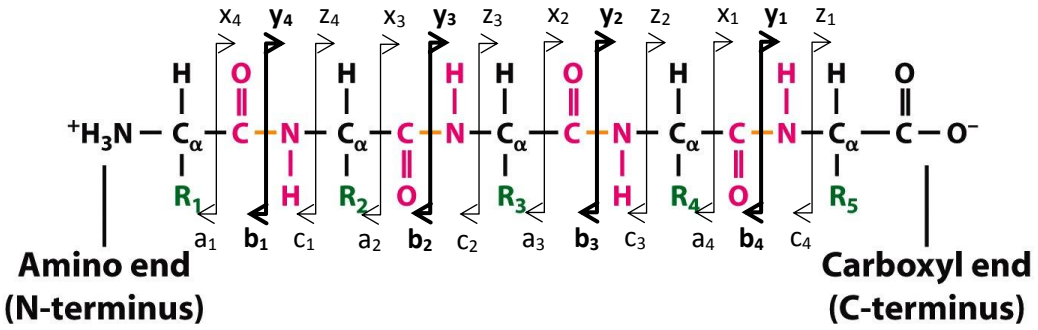


Figure 2.6 MS/MS fragmentation : The Biemann nomenclature for MS2 ions [Adapted from 107,108]

2.3.4.4 Detection

The detector converts both MS1 and MS2 ions detected in the last mass analyser to provide an interpretable spectrum. Several types of detectors exist. Still, they all use the charge, the mass or the velocity of the ions to produce a signal related to the amount of ions detected. The detector in the TT5600 is a multichannel time-to-digital converter (TDC) detector with 4-anode/channel detection that relies on kinetic energy transfer [232]. Briefly, incident ions collide with a surface that in turn generates secondary electrons, which are further amplified to give an electronic current [120].

Each detector possesses an intrinsic dynamic range of linearity where the signal is correlated to the amount of ions, resulting in usable quantitative data. This linear dynamic range is delimited by upper and lower limits of quantification that correspond respectively to detector saturation and detector unresponsiveness (Figure 2.7). TT5600 is expected to provide at least four orders of dynamic range, in both MS1 and MS2 scanning modes [232].

Importantly, each ion has unique ionization properties, so the limit of quantification (LOQ) cannot be inferred from one peptide to another *per se*.

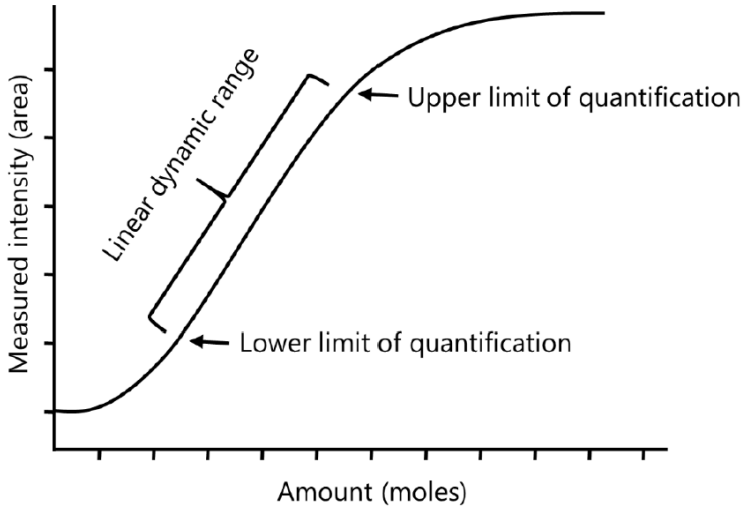


Figure 2.7 Linear dynamic range and limits of quantification

2.3.4.5 Untargeted acquisition strategies

For the sake of sensitivity and repeatability, clinical and quality control (QC) applications of liquid chromatography-tandem mass spectrometry (LC-MS/MS) are often operated by selecting only a defined subset of MS1 and MS2 ions to be analysed ("Selected/Multiple Reactions Monitoring" or S/MRM). However, this requires a priori information on the targets of interest and restricts the gained information. For non-targeted and comprehensive screenings, shotgun data-dependent acquisition (DDA) has long been the golden standard approach in MS. In recent years however, data-independent acquisition (DIA) emerged as a strong alternative [236]. Main acquisition strategies on QTOF device are summarized in Figure 2.8.

a) Data-dependent acquisition (DDA)

In traditional data-dependent acquisition (DDA), the MS2 acquisition relies on the data coming from the MS1 scan of the ionized peptides. Only a limited number of ions surpassing a certain signal threshold are selected for fragmentation. This produces tandem MS data where precursor and product ion signals can be associated and quite easily annotated (see 2.3.4.6). Although extremely powerful, DDA presents inherent technical variability [237]. Small differences in sample loading and/or slight shifts in LC retention times alter the instant composition of the peptides mix eluting from the column. This results in a certain variation of the MS1 signal intensity between repeated runs. Therefore, each DDA run can sample slightly different peptides for MS/MS (i.e. stochastic sampling), leading to some variation in gained spectral data.

It is generally accepted that 5 to 10 repeated injections of one sample are needed to uncover almost all possibly identifiable peptides of a complex sample [238–240].

b) Data-independent acquisition (DIA)

The development of faster and higher resolution MS instruments opened the path to new acquisition strategies that could overcome the main DDA limitations. Data-independent acquisition theoretically combines the advantages of DDA and S/MRM by offering non-targeted reproducible screening. In DIA, all peptides are subjected to fragmentation, irrespective of their MS1 signal. A single DIA run thus theoretically contains all the MS/MS spectral information that a sample can deliver at a given limit of detection.

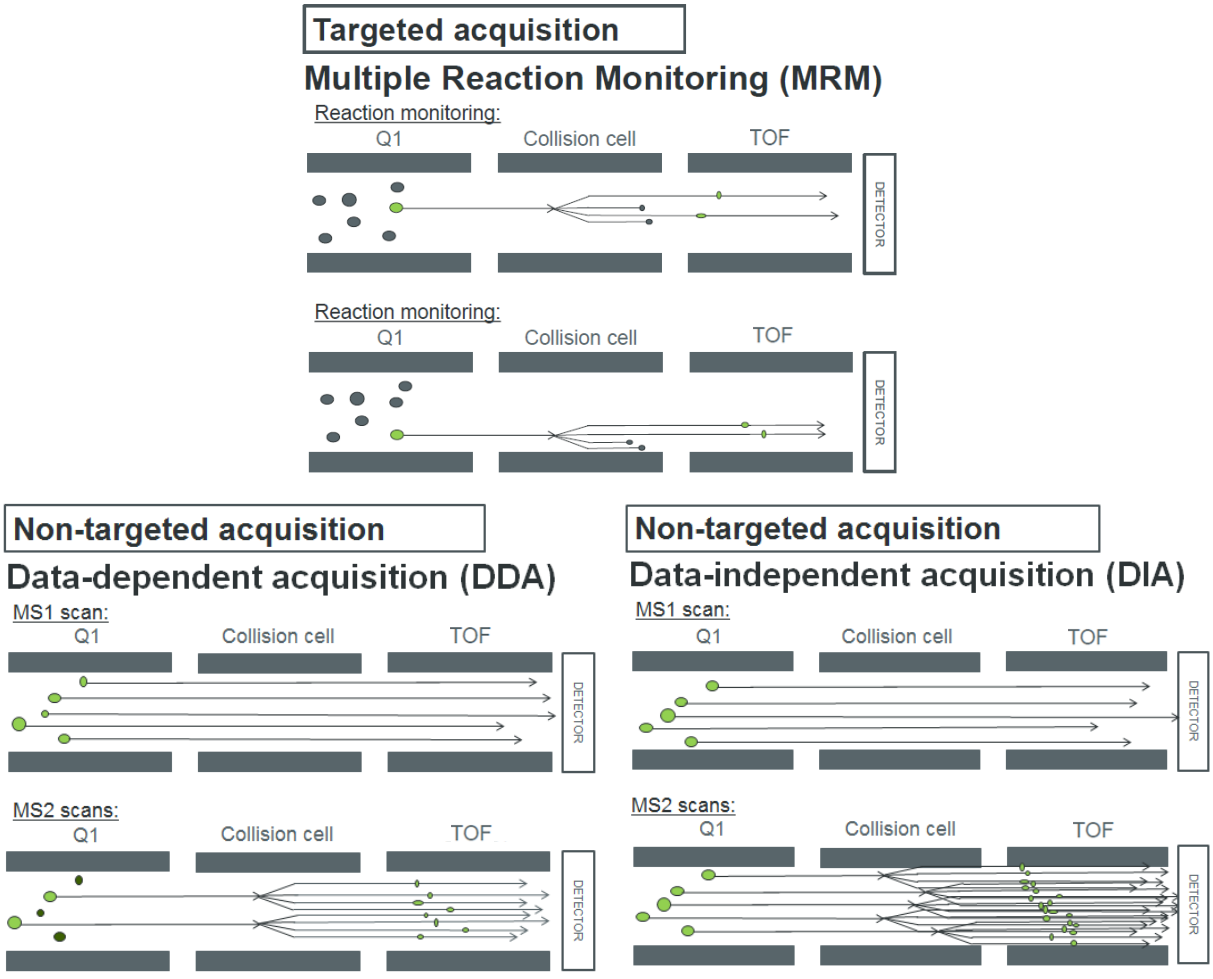


Figure 2.8 Main acquisition strategies on QTOF MS device

While the first DIA approaches were published about 15 years ago [241,242], effective DIA implementations have only been developed over the last five years [236]. Two main approaches are currently promoted by the QTOF instrument vendors, namely MS^E and SWATH.

In MS^E, the CID cell rapidly alternates between low energy and high energy, to acquire respectively the MS1 and the MS2 ions [242]. Nowadays, MS^E is most often implemented as $HDMS^E$, where ultra-high pressure LC and ion mobility (mobility of the ions in a carrier buffer gas) are used as orthogonal separations of the ions [243].

In SWATH (Sequential Windows Acquisition of All Theoretical Fragment Ion Mass Spectra), the MS1 *m/z* range is divided into adjacent *m/z* isolation windows (or "swaths") which are independently and consecutively selected, fragmented and analysed [244]. Consequently, MS2 data is related to a given range of possible precursor *m/z* that was selected by the quadrupole, which provides increased specificity as compared to other DIA like MS^E [244]. The acquisition cycle time results from the product of the number of windows multiplied by their accumulation time (with the addition of the initial MS1 scan) (Figure 2.9). For a fixed cycle time (set to ensure enough data points per LC peak), an increased number of shorter accumulation windows will increase the specificity, while reduced number of longer accumulation windows will favour the sensitivity. Variable window widths can now be scaled across the *m/z* range, depending on the number of precursors found: higher number of shorter windows will be allocated to the regions with higher density of measured precursors (sciex.com/support/knowledge-base-articles/how-to-use-the-swath-variable-calculator-excel-sheet).

The main challenge with DIA is the complexity of the obtained convoluted data that requires elaborate data analysis software tools.

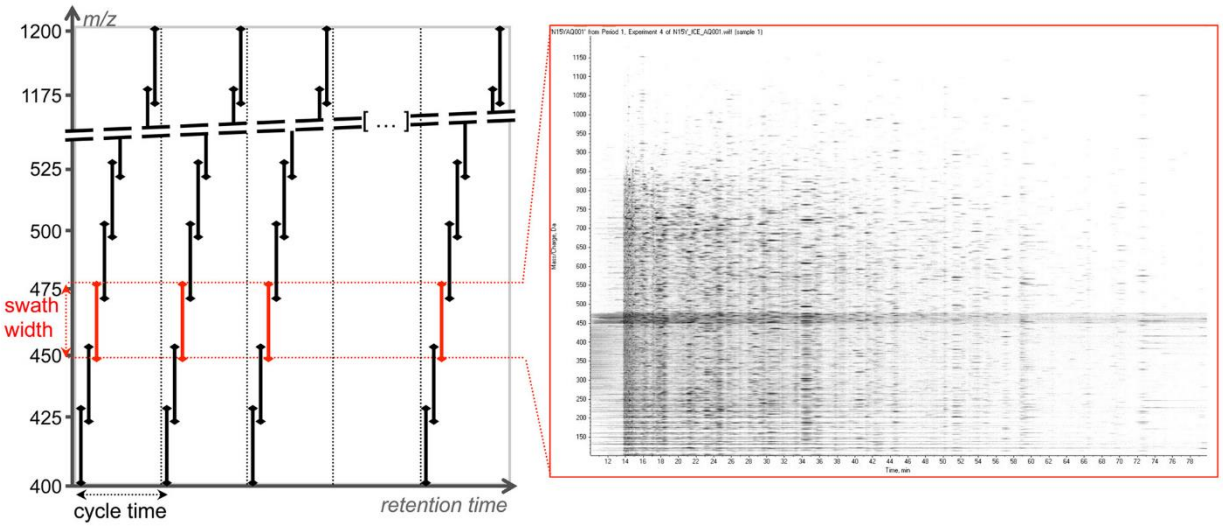


Figure 2.9 SWATH MS data-independent acquisition [244]

2.3.4.6 Data analysis

MS acquisition results in a set of measured MS1 precursor ions *m/z* and MS2 fragment ions *m/z*, with associated intensities as produced by the detector: the MS/MS spectrum. Correct interpretation and annotation of this spectrum allows peptide identification, and even quantitative information.

a) Identification

Different identification strategies can be considered to annotate spectral data with corresponding peptides (Figure 2.10).

Manual annotation of an MS/MS spectrum can be performed by considering the *m/z* distance between consecutive MS2 peaks ("de novo sequencing"). Still, this approach becomes quickly hard to handle when the complexity and the number of spectra increase.

In most cases, raw data files are converted into peak list file format that can be compared to publicly available protein databases. Non-redundant databases like UniProtKB/Swiss-Prot contain only previously identified proteins (in contrast to redundant databases that also contain predicted proteins from sequenced genomes) [245]. Search engine software like Mascot (Matrix Science) match the experimental peak lists to *in silico* generated ions of the database, using both peptide mass fingerprint (very accurate molecular mass of the enzymatic peptides) and MS/MS ions search (fragment ion mass and relative intensity values) [246]. Following parameters are defined: the type of enzyme, the expected peptide charge, fixed or variable modifications, the number of tolerated missed cleavages, the MS instrument type, the MS1 and MS2 mass tolerance.

Other search engines like Protein Pilot (Sciex) use so-called "sequence tag" approach, where small sequences of few amino acids ("tags") are generated from the spectrum and the protein database is searched for entries where many tags match. The main advantage of this approach is that the MS1 precursor mass is not used as a precursor filter and possible modifications do not have to be specified [247].

A false discovery rate (FDR) is most often calculated to validate the annotation by matching with the protein database. FDR estimates the rate (%) of matches that are actually false positives by dividing the number of matches in a decoy database by the number of matches in the real database. A decoy database contains randomized or reverse sequences, so that any match to this database is expected to be a false positive.

A minimum of two significant unique peptide matches is generally required to consider confident protein identification. A peptide match is unique if it contains sequences that unambiguously belong to only one protein in the database. In Mascot, the significance is defined by a probability-based score that is calculated to reflect the quality of the data [190].

A fourth identification strategy consists in matching the obtained spectrum directly with previously acquired spectrum on same (or very similar) device that was already annotated. This is the main approach adopted to overcome DIA data complexity in SWATH-MS.

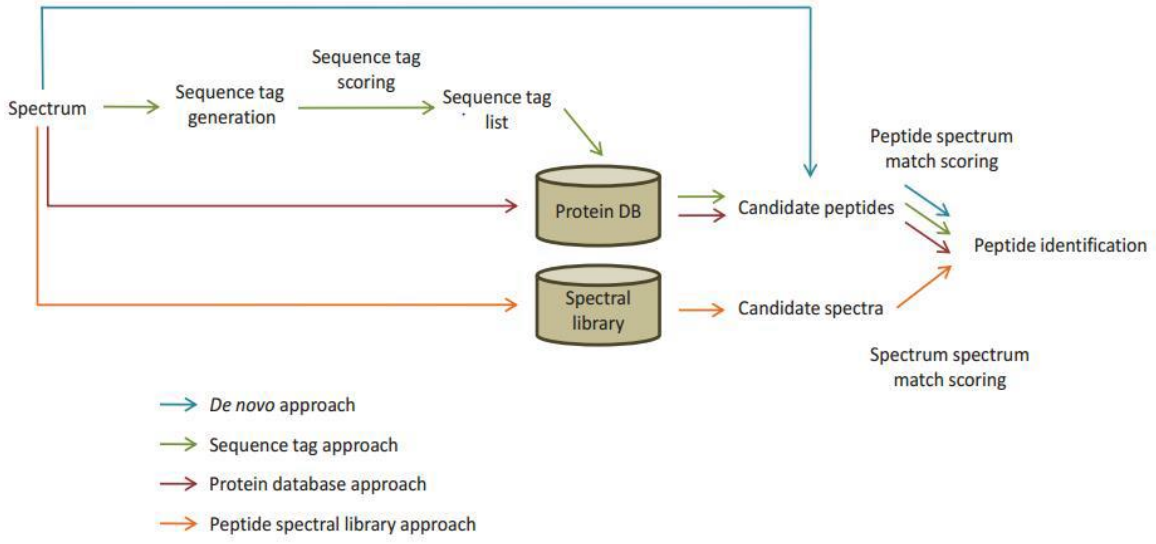


Figure 2.10 Peptide identification strategies using MS-based proteomics [248]

In SWATH-MS, spectral assay libraries generated on a high resolution MS instrument with DDA (either in-house or by a reference lab which made them publicly available) are generally used to perform data extraction by matching SWATH spectral data with annotated spectra from the library. Both chromatographic and spectral components of the peptides indexed in the spectral library, including the retention time, the peak width and area, the mono-isotopic mass accuracy or the presence of contiguous ion series, are used to mine the MS2 maps for signals that correlate with these coordinates. The overall quality of this correlation ("match") is reflected into a score that is compared to the score obtained by matching a library of false decoy sequences, leading to a second FDR calculation (the first being the FDR of the annotation of the DDA library runs with a database search). Adopting a stringent control of this second FDR has been an important pillar of recent SWATH data analysis optimization [249–252]. Retention time alignment to overcome potential LC variation has also become widespread practice. Retention times are indexed as dimensionless values that refer to the elution of spiked reference peptides spaced over the LC gradient [249,253]. Recently, a multicentre benchmark study of the main SWATH data analysis pipelines available demonstrated a very high consistency in MS2 peak picking and resulting peptide identification across all software tools, exceeding the overlap between different DDA search engines. Recommended default settings for each of these tools were also published [254].

Importantly, extended spectral libraries can be built to query the SWATH data over a large amount of features, which largely exceeds what was identified in the sample with DDA [250,251]. For the sake of impurities screening, the library can be built directly on the (fractionated) source material to index the largest possible amount of potential residuals, that could be then screened in SWATH runs of final products, over and over.

Of note, SWATH data analysis without spectral library has been also developed. The principle is to assemble precursor and fragment co-eluting features into pseudo-tandem MS data that can be identified with conventional database searching ('data-centric approach') [255]. While presenting overall good

performance, this approach results in somewhat different reported peptides, as compared to the library-based tools ('peptide-centric approach') [254].

b) Quantification

For some harmful and/or undesired contaminants, confident detection already provides important information. Still, in most cases, quantitative information will be required to properly consider the impact of a contamination. An overview of possible quantification strategies with MS-based proteomics is given in Figure 2.11.

The detector measures peak intensities, but MS does not allow for direct absolute quantification, since individual peptides differ in ionization susceptibilities and in isolation and detection efficiencies, being also dependent on the MS equipment. S/MRM, most often with synthetic isotope-labelled peptides or proteins spiked as internal standards, is still the only convincing way to perform absolute quantification with MS, at least for LAP [183,256–258]. Nevertheless, this approach asks for a priori information and dedicated development: it is thus not suitable for non-targeted profiling.

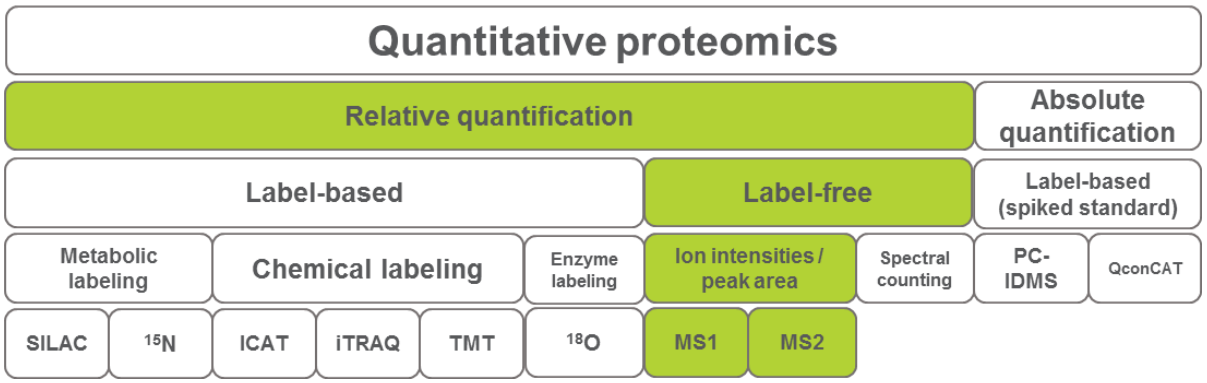


Figure 2.11 Main peptide quantification strategies using MS-based proteomics (adapted from Paulien Meert). Green boxes highlight the strategies considered in the present work.

Most large-scale non-targeted proteomics studies thus perform relative quantification. Stable isotope labels were often considered to perform multiplex runs where samples to compare are pooled together to account for technical variability during the analytical process, while the unique label of each sample allows accurate comparisons. Isotope labels can be incorporated (i) during cell culture ("stable isotope labelling of amino acids in cell culture - SILAC" using $^{13}C/^{15}N$ arginine and lysine residues or ^{15}N culture media), (ii) using enzymatic labelling (digest in ^{18}O water) or (iii) with chemical covalent labeling of exogenous reagents ("isotope-coded affinity tags - ICAT" using derivative reduced cysteine residues, "tandem mass tags – TMT" and "isobaric tag for relative and absolute quantification – iTRAQ" using N-termini and lysine side chains labels which are isobaric and chromatographically indistinguishable but generates different reporter ions after fragmentation). However, these standards are sometimes not easy to obtain, often expensive, increase the sample complexity and always require time-consuming additional steps. Moreover, multiplexing tends to decrease the dynamic range of measured protein abundances [231]. Opposed to absolute quantification, isotope-labelled standards can be avoided. Label-free methods are simpler, applicable to any sample and any protein/peptide, higher throughput and more economical approaches.

Importantly, a label-free approach requires LC retention times alignment to account for run-to-run and or device-to-device chromatographic variability. Using adequate computational resources, ion abundances can be evaluated through spectral counting (the total number of spectra identified for a protein) or peak measurements (intensity and peak area, also called area under the curve-AUC). The latter is generally preferred since the first might be biased towards more abundant proteins [259,260]. MS signal stability should be also carefully monitored and run-to-run signal normalization might be asked.

Using DDA, label-free quantitation can only occur at the MS1 level, since the MS2 acquisition is inherently stochastic. On the other hand, DIA benefits from full MS/MS datasets and thus opens the path for MS2 quantitation with greater specificity. DIA is expected to provide comprehensive large scale profiling with quantitative characteristics which were hitherto S/MRM privilege [261].

In SWATH, the ion chromatograms associated to the peptides passing the FDR of the spectral library matching are extracted at the MS2 level. This results in MS2 peak area values that are simply summed up to give peptide- and protein level relative abundances. Recent inter-laboratory study evidenced that proteins can be reliably quantified across 4 orders of dynamic range using SWATH on Triple TOF5600 [261].

Chapter 3:
Outline and aims of this thesis

3. Outline and aims of this thesis

Plasma-derived therapeutics constitute a unique class of medicinal products due to the nature of their source material. In particular, normal polyvalent Ig succeeds to provide administrable immune cocktails that can reconstitute deficient immune repertoire as well as demonstrate beneficial immunomodulatory effects in ever-growing indications (ITP, Guillain-Barré syndrome, Kawasaki syndrome, ...). Being irreplaceable by recombinant products, normal Ig drives the plasma fractionation market.

From the early cold ethanol precipitation of Cohn and Oncley in the 1940s, plasma fractionation progressively evolved in a refined and structured large-scale industry. With the help of a strict regulatory framework, manufacturers achieved to provide high integrity and high purity products with unprecedented microbiological safety. Still, the complexity of the plasma pool source material makes absolute purity challenging, if not impossible. Variation in the source material origin as well as the diversification of production processes raises the question of Ig products interchangeability. In 2010-2011, a TEE outbreak with administered Ig was associated with trace levels (ng/mg) of plasma residual FXI in the incriminated batches. A posteriori corrective and preventive actions including process revision and additional testing solved the issue but remained specific and dedicated.

Since co-purification during the manufacture can theoretically lead to the residual presence of any of the plasma proteins from the plasma pool, non-targeted generic screening could detect such safety issues due to source material impurities. This thesis aims to consider the use of MS-based proteomics to provide relevant profiling of plasma impurities in polyvalent Ig.

While a more detailed biological background was introduced in **Chapter 1**, **Chapter 2** presented the main MS strategies that can be developed to achieve such non-targeted screening of LAP. This work will be centred on the use of a high resolution ESI-Q(q)TOF, the TripleTOF5600 (Sciex), preceded by reversed-phase LC separation of the tryptic peptides obtained by *in solution* enzymatic digest.

In **Chapter 4**, the two main prefractionation strategies – namely highly abundant component specific depletion vs. non-targeted enrichment of LAP – will be compared. The classical shotgun DDA approach will also be discussed and label-free MS1 relative quantification will be performed. Due to inherent issues in repeatability and coverage with the DDA, SWATH-MS will be further considered in **Chapter 5**. As DIA method, SWATH-MS enables label-free MS2 relative quantification. An innovative use of the variability to signal dependence will be proposed to decipher the confidence of the complex data generated. Both in Chapters 4 and 5, the QC relevance will be challenged by the detection of FXI spiked at the ng/mg as well as the comparison of five different Ig products from three different manufacturers, including two batches of a same product and one TEE-positive batch.

Overall conclusion and final back-to-back comparison of the two acquisition methods will be wrapped up in **Chapter 6**. Finally, **Chapter 7** will question the broader relevance of MS-based proteomics in the QC of biological therapeutics. The future perspectives that this work suggests will be considered.

Chapter 4:

Mass spectrometry-based sensitive screening of plasma residuals in human plasma-derived immunoglobulin *(part I)*:

sample fractionation and data-dependent acquisition

Adapted from "**An application of mass spectrometry for quality control of biologicals: Highly sensitive profiling of plasma residuals in human plasma-derived immunoglobulin**"

Franck Limonier, Katleen Van Steendam, Geneviève Waeterloos, Koen Brusselmans, Myriam Sneyers and Dieter Deforce

Journal of Proteomics (2017) 152, 312-320
DOI: 10.1016/j.jprot.2016.11.007

4. Mass spectrometry-based sensitive screening of plasma residuals in human plasma-derived immunoglobulin *(part I)*: sample fractionation and data-dependent acquisition

4.1 Abstract

Thromboembolic events (TEE) associated to trace amounts of plasmatic activated coagulation factor XI (FXIa) in administrated immunoglobulin (Ig) have recently raised concerns and hence there is a need for highly sensitive profiling of residual plasma source proteins. This study aims to consider LC-ESI-QTOF data-dependent acquisition in combination with sample fractionation for this purpose. Sample fractionation proved mandatory to enable identification of plasma residuals. Two approaches were compared: Ig depletion with protein G - protein A affinity chromatography and low-abundant protein enrichment with a combinatorial peptide ligand library (ProteoMiner™, Bio-Rad). The latter allowed a higher number of identifications. Highly sensitive detection of prothrombotic FXIa was assessed with confident identification of a 1 ng/mg spike. Moreover, different residuals compositions were profiled for various commercial Ig products. Using a MS1-level quantitative label free analysis, a TEE-positive Ig batch was distinguished from other regular Ig products, with increased levels of FXIa but also other unique proteins. This could have prevented the recently observed TEE problems with Ig. The method is a convenient tool to better characterize Ig products after any plasma pool or manufacture process change, gaining insights in the product quality profile without any prior information required. Still, the inherent stochastic nature of data-dependent acquisition asks for repeated injections and extended acquisition time.

4.2 Introduction

As explained in Chapter 1, the 2010-2011 TEE outbreak highlighted the importance of the residual co-purified fraction in immunoglobulin (Ig) products, even at trace levels. The *in vivo* Wessler thrombosis rabbit model indicated that the minimal dose of FXIa spike in intravenously administrated products causing a prothrombotic effect is as low as 300 ng FXIa/mL product. As it was a 100 mg Ig/mL product, it corresponds to 3 ng FXIa/mg Ig [103]. With a sensitivity as high as 0.2 ng FXIa/mL product, the thrombin generation assay (TGA) is increasingly recommended by several papers and promoted by the US Food and Drug Administration [94]. However, this targeted test is dedicated to FXI level and procoagulant activity measurements. Given that relevant sensitivity could be achieved, non-targeted screening of plasma residuals in final Ig product after each process or plasma pool source modification may warn against any

composition change. In this Chapter, an application of bottom-up mass spectrometry (MS) is developed to provide such convenient screening.

As introduced in Chapter 2, MS-based techniques have already been applied to characterize residuals in plasma-derived products, such as HSA preparations, clotting factor VIII or IX and prothrombin complex concentrates, allowing the identification of numerous impurities (Table 2.1). In 2013, Lackner *et al.* used gel-based MS to identify beta-2-glycoprotein 1 or apolipoprotein-H (ApoH), serotransferrin (TRFE) and HSA as contaminants of Ig preparations [93]. Nevertheless, no consensus has emerged so far concerning the adopted method and neither sensitivity nor reproducibility was discussed. Despite continuous developments in the MS technologies, sample pre-fractionation techniques remain paramount to access the low abundant contaminant fraction [192].

Depletion of high abundant proteins – here the Ig – is a common strategy to access identification of lower abundant proteins (LAP) but can be somewhat hindered by issues of co-depletion due to non-specific binding interactions [201–203,262]. The use of a large, highly diverse combinatorial (hexa)peptide ligand library (CPLL) has been established as a strong alternative to reduce the samples protein concentration dynamic range and access LAP identifications [131,206,209,210,213,263–265]. Here, both specific depletion of Ig – by means of protein G - protein A affinity chromatography – and CPLL enrichment strategies were evaluated. Impurities coverage was also optimised considering inherent variability of the classical shotgun MS acquisition mode, the so-called data-dependent acquisition (DDA) mode [237].

QC relevance was put to the test by the identification of spiked FXIa at 1 ng/mg and comparison of different Ig products, including a TEE-positive Ig batch, and two batches from a same product. MS precursor normalized abundance levels were also reported to obtain relative quantification of impurities.

4.3 Material & methods

4.3.1 Samples

Intravenous Ig (IVIg) products were provided by two manufacturers. Products from the first manufacturer will be referred as A and B while product from the second as C. Samples from different batches of product A will be referred as A1 and A2. A 2011 TEE-positive subcutaneous Ig (ScIg) product batch from a third manufacturer (referred as D) was also obtained. Protein concentrations are 50 mg/mL (A and B), 100 mg/mL (C) and 160 mg/mL (D) respectively.

Native coagulation factor XIa from human plasma (Calbiochem, Merck Millipore, Darmstadt, Germany) was used to spike the samples at 1 and 0.2 ng/mg total protein.

4.3.2 Sample fractionation

4.3.2.1 Ig depletion by protein G - protein A affinity columns

Sepharose 4 Fast Flow gel filtration medium was coated with either protein A or G and served to pack two XK 26 columns (GE Healthcare, Chalfont St. Giles, UK). Gel medium height was 10 cm in each column. The two columns were put in series on a ÄKTA purifier 10 system (GE Healthcare), giving a total column volume (ColV) of approximately 100 mL. All the buffers were filtered with 0.22 µm vacuum filters and the flow was set at 5 mL/min. The columns were washed with 3 ColV of milliQ water and then equilibrated with 3 ColV of starting buffer (20 mM sodium phosphate solution, pH 7). Eight different sample loads of Ig product were tested: 0.2-0.4-0.8-2-5-10-15-20 mL. Non-Ig impurities were measured at 280 nm and collected.

Bound Ig fraction was eluted with 0.1M glycine buffer (adjusted to pH 2.7 with HCl_{cc}). Prior to a new sample injection, columns were washed with 3 ColV of milliQ water and reequilibrated with 3 ColV of starting buffer. If stored before next use, columns were equilibrated with 3 ColV of ethanol 20% in milli-Q water and stored at 4°C. Collected flowthroughs were desalted and concentrated with Centricon Plus-70 10 kDa (Merck Millipore, Darmstadt, Germany) and dried under vacuum.

4.3.2.2 Low-abundant protein (LAP) enrichment with combinatorial peptide ligand library (CPLL)

The bead-based CPLL was obtained from ProteoMiner™ protein enrichment kits (Bio-Rad, Hercules, CA, USA). Instructions from the Bio-Rad kits were first followed [266]. In short, proteins were captured in physiological conditions with phosphate buffered saline (Na_2HPO_4 10mM pH 7.4, NaCl 150 mM). “Small-capacity” (20 µL settled bead volume) and “large capacity” (100 µL settled bead volume) kits, requiring respectively 200 µL and 1 mL of ≥50mg/mL sample load, were compared. Bound proteins were desorbed using either single-elution (urea 8M, CHAPS 2% and acetic acid 5%) or sequential elution protocol (reagent 1 is NaCl 1M, HEPES 20mM (pH 7.4), reagent 2 is Glycine 200 mM (pH 2.4), reagent 3 is ethylene glycol 60% in water and reagent 4 is 2-propanol 33.3%, ACN 16.7%, TFA 0.1%).

Eluted fractions were desalted using either protein precipitation (ReadyPrep 2-D Cleanup kit, Bio-Rad, Hercules, CA, USA) or filtration with 10 or 3 kDa cut-off (Amicon Ultra 0.5 mL 10 or 3 kDa, Merck Millipore, Darmstadt, Germany).

Secondly, protein capturing in a low ionic strength environment (Phosphate buffer 25mM pH7, NaCl 50 mM), followed by sequential recapturing of the excess at pH 9.3 (Tris-HCl 25 mM) and then pH 4 (acetate buffer 25 mM) and coupled to on-bead digestion was performed according to the optimized protocols published by Boschetti and Righetti [214,215].

To study the technical repeatability of the fractionation, five replicate enrichments were performed for one given sample.

4.3.2.3 Total protein assay

The sample recovery after each fractionation (except CPLL with on-bead digestion) was assayed using the Micro BCA Protein Assay kit (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's instructions.

4.3.3 Tryptic digest

Depleted or CPLL-eluted samples were digested *in solution* as following. Dried protein pellets were suspended in triethylammonium bicarbonate (TEABC) 0.5 M. DTT was added at 1mM as reducing agent and samples were incubated 1 hour at 60°C. S-methyl methanethiosulfonate (MMTS) in isopropanol was added at 10 mM and incubated for 10 minutes at room temperature to alkylate the proteins. Finally, Trypsin/lysC (Promega Corp., Madison, WI, USA) in TEABC 0.5M was added at a mass ratio enzyme/protein of 1:100 and complemented with calcium dichloride ($CaCl_2$) 1mM and acetonitrile (ACN) 5%. Complete tryptic digestion was achieved by overnight incubation at 37°C. Samples treated with CPPL performed according to optimized methods from Boschetti and Righetti were digested "on-bead", following published protocol [214,215]. Digested samples were vacuum dried and stored at – 20°C until injection.

4.3.4 LC-MS/MS analysis

4.3.4.1 Sample injection and LC separation

Dried peptides were dissolved to 1µg/µL in 0.1% formic acid (FA) in water. An extra desalting step was also tested by means of 10 µL pipette tips with a bed of C18 reversed-phase chromatography media fixed at its end (ZipTip C18 Pipette Tips, Merck Millipore, Darmstadt, Germany). Peptides were separated on an Eksigent nanoLC 425 (Eksigent, Sciex, Framingham, MA, USA). Two micrograms of each sample was first trapped on a pre-column (YMC-Triart C18, pore size 12nm, particle size 3µm, 0.5 mm internal diameter by 5 mm; YMC, Dinslaken, Germany), and separated on reversed-phase (C18) microflow (5µL/min) LC column (YMC-Triart C18 15 cm, particle size 3 µm, 0.3 mm internal diameter by 150 mm; YMC, Dinslaken, Germany) using a linear gradient of 95:5 buffer A (95% H_2O; 5% DMSO; 0.1% FA) /buffer B (95% ACN; 5% DMSO; 0.1% FA) to 20:80 buffer A/buffer B at 5 µL/min over 1 hour. With the on-bead digested sample, dried peptides were dissolved in 500 µL FA 0.1%, assuming a total binding capacity of 1000 µg per large capacity spin column. Total ion count (TIC) was checked to be between 5E6 and 5E7 in intensity; injecting 2µL per run appeared to be adequate.

4.3.4.2 MS data acquisition

A TripleTOF 5600 (Sciex, Framingham, MA, USA) was used in a data dependent mode to analyse peptides. Dynamic accumulation allowed to switch automatically from MS (400-1250 *m/z*; accumulation time 0.25 s) to MS/MS (65-2000 *m/z*; accumulation time 0.025 s) when a threshold of 50cps was exceeded. Surpassing ions were excluded from further analysis for 30 seconds. For quantification purposes, the sample list was scrambled.

4.3.4.3 Data analysis

a) Protein identification

For method optimisation, raw spectra data (Wiff) files were converted to Mascot Generic Format (MGF) with Protein Pilot Software 4.5 (Sciex, Framingham, MA, USA) and data were searched against the Swissprot Mammalia database (66,431 sequences) using Mascot Daemon version 2.5 (Matrix Science, Boston, MA, USA) with methylthio on cysteine as fixed modification, and oxidation on methionine and deamidation on glutamine and asparagine as variable modifications. Peptide mass tolerance was set at 10 ppm and fragment mass tolerance at 0.1 Da, and peptide charge was set at +2, +3 and +4. Two missed cleavages were allowed and identity threshold for peptides was set at p-value <0.01. False discovery rates for peptide identification of all searches were less than 0.2% with automatic decoy database search. Minimum 2 unique peptides-sequences per protein and a Mascot score above 50 or only 1 unique peptide-sequence and a Mascot score above 80 were requested as protein identity threshold. Expression in blood was verified using NexProt repository (available at http://www.nextprot.org/db). Fragments of G class Ig (IgG), the main type of antibodies constituting the Ig products, and epidermal and hair proteins (keratin variants, dermcidin and hornerin), if identified, were never reported.

To overcome DDA variability, up to 16 LC-MS injections were performed for one given sample. 8 repeated LC-MS injections were performed for each tested condition/technical replicate to calculate the mean (± standard deviation (SD)) number of detected proteins that satisfy the identity threshold. Normal distribution assumption was tested with the Shapiro-Wilk test. As data were not normally distributed, the non-parametric Mann-Whitney U test for two independent samples was used. All statistics were performed in Microsoft Excel, using Analysis ToolPak and Real Statistics add-ins.

b) Protein relative quantification

For Ig products comparison, raw spectra profile data were analysed with Progenesis QI for Proteomics (Nonlinear Dynamics, Waters, Durham, NC, USA) to obtain relative quantification based on MS precursors intensity. After alignment and normalization, a multivariate statistical analysis was performed on all MS precursors, without any prior peptide identification. The MS/MS spectra of the differential MS precursors were exported as a MGF peak list for database searching against the Swissprot Mammalia database using Mascot, as described above. Results were filtered for confident hits using a 0.01 expectancy cut-off. Protein-level abundances are weighted sums of normalized abundances from non-conflicting peptides matched to a given protein. Such abundances were reported for every (non-IgG) plasma residual protein with a confidence score above 50 and at least 1 unique peptide. Considering all the reported protein abundances, a median coefficient of variation (CV) between all the MS injections of one sample preparation and another between the different technical replicates (considering all the repeated injections) were calculated and reported as percentages. Fold changes in average protein abundances compared to the median abundance in every tested Ig samples were reported. Finally, a heatmap with hierarchical clustering and data normalization (mean=0, variance=1) was generated in Qlucore Omics Explorer (Qlucore AB, Lund, Sweden).

4.4 Results

4.4.1 Method development

4.4.1.1 Sample fractionation

Injection of commercial IVIg product (product A1) spiked with FXIa 1 ng/mg without any sample fractionation did not return any confident identification except IgG fragments, trypsin and sample handling-associated epidermal proteins (keratin, dermcidin and hornerin), which are never reported hereafter. Therefore, sample fractionation was mandatory to identify residual (non-IgG) plasma proteins (at least 2 identified peptides, Mascot score ≥ 50 or only 1 peptide, Mascot score ≥ 80). Two strategies were compared:

(1) Ig depletion by protein G - protein A affinity chromatography
(2) LAP enrichment with CPLL

Correct sample load was assessed for both strategies: at least 2 mL had to be loaded on the protein G - protein A columns, while for CPLL, the "large capacity" kit (1mL load) gave better results than the "small capacity" (200 µL load) (Figure 9.3 in the Addendum).

Optional "sequential elution" from ProteoMiner™ kits, which uses 4 elution reagents instead of 1 ("normal elution"), allowed additional identifications but at the same time few proteins identified with normal elution were also lost, resulting in a borderline significant increase of identifications (Figure 9.4 and Table 9.2 in the Addendum).

Desalting method refinement was conducted, in order to minimize possible protein loss at this level. Different methods were compared, namely 10 or 3 kDa cut-off filtration, protein precipitation (ReadyPrep 2-D Cleanup kit, Bio-Rad), with or without additional C18 ZipTip peptides treatment, but no clearly significant impact on the results was outlined (data not shown).

With sequential elution and precipitation desalting, CPLL fractionation allowed identifying up to 31 plasma proteins in product A1, which is the double of identifications made with Ig depletion (Figure 4.1 and Table 9.2).

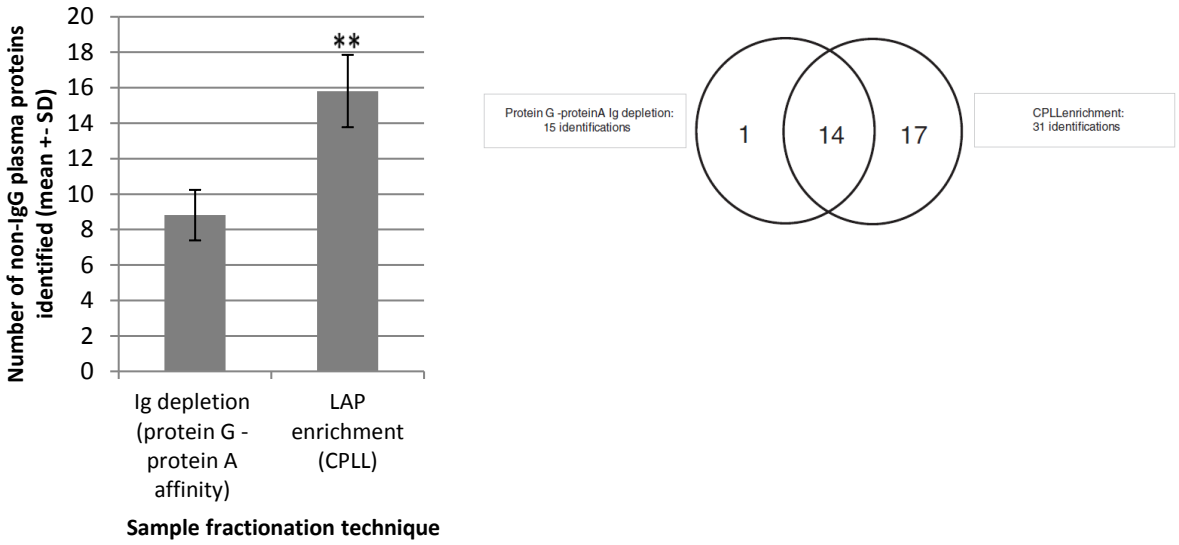


Figure 4.1 Non-IgG plasma protein identifications found in product A1, either depleted with protein G – protein A affinity chromatography or (optimised) CPLL enrichment. Optimised CPLL enrichment protocol consisted in using large capacity – sequential elution kit of ProteoMiner™, followed by precipitation desalting. Each fractionation was tested twice and each time injected 8 times on LC-MS/MS. This gives a total of 16 runs per condition. (A) Mean number ± SD (n = 16) of identifications. ** indicates Mann-Whitney U test p-value <0.01. (B) Venn diagram of identifications overlap (with all 16 runs for each technique merged for database search).

Alternative CPLL protocol, namely protein capturing in a low ionic strength environment, coupled to on-bead digestion, led to decreased number of identifications. Further addition of subsequent recapturing of the flow through excess, first at pH 9.3 and finally at pH 4 did not improve the residuals coverage (Figure 9.4 and Table 9.2).

Combination of both fractionation techniques (protein G/A depletion followed by CPLL enrichment) allowed deeper coverage of the residuals compared to single step fractionation (Figure 9.4 and Table 9.2). Nevertheless, sample preparation becomes burdensome (Ig depletion process takes one day per sample).

Considering increased coverage of plasma residuals within convenient execution time, CPLL enrichment – ProteoMiner™ large capacity with protein capture in physiological conditions and sequential elution (combined to protein precipitation desalting) – was used as the preferred sample fractionation method.

4.4.1.2 Variability and sample coverage

Two sources of variability may affect the samples' residuals coverage, namely the list of reported proteins which satisfy identification threshold:

(1) Inherent variability coming from MS run in DDA mode
(2) Technical variability linked to quite extensive sample preparation (fractionation/desalting)

Up to 16 repeated MS injections were performed to gain the highest possible number of peptide identifications in a given sample. Maximal protein sequence coverage was generally reached after 8 repeated runs (Figure 4.2).

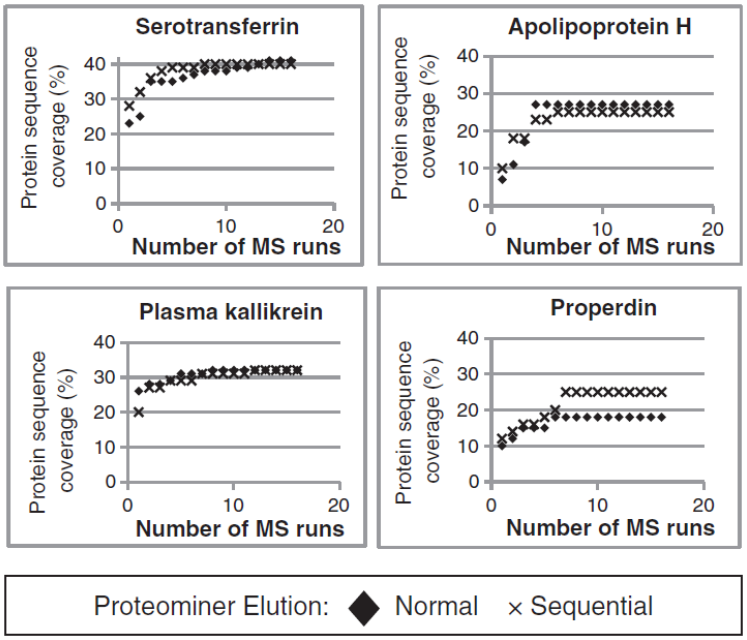


Figure 4.2 Sequence coverage (%) for four identified proteins in product A1, depending on the number of replicated injections merged for data analysis. Sample was enriched with CPLL – ProteoMiner™ large capacity - normal or sequential elution.

Median CV in normalized abundances of (non-IgG) plasma proteins identified was 10.7% between the 8 injections for each of five replicates. 88.6% of data had a CV below 30%.

Technical repeatability of the sample treatment was controlled by performing five CPLL fractionations in parallel ("technical replicates (T)") for one given sample. Every T was injected 8 times. 75% of total plasma LAP identifications (15 to 20) were reported in every T. 94% (17 to 18) with at least two peptides were made in at least four on five replicates, meaning that all the identifications reported in less than three to five replicates were single-peptide identifications. Concerning quantitative repeatability, median CV in normalized abundances of (non-IgG) plasma proteins identified was 20.2% between the five replicates. 81% of data had a CV below 30%. For given plasma protein identified, the highest observed fold change was 4 between two technical replicates and 2 when compared to the median of abundances in all five replicates.

In order to find a compromise between residuals coverage and time spent on MS acquisition, two technical replicates were performed, each injected 8 times, giving a total of 16 DDA runs per sample.

4.4.2 Method application and QC relevance

The developed method with sample CPLL enrichment (ProteoMiner™ large capacity, sequential elution and protein precipitation desalting) was put to the test to identify a sensitivity marker (spiked FXIa) and to profile plasma residual proteins of different Ig products, including a TEE-positive batch. For each application, sample preparation was performed twice and each T was injected 8 times on MS, giving a total of 16 DDA runs per sample.

4.4.2.1 Identification of sensitivity marker: spiked FXIa

Spiking of FXIa was used as a sensitivity marker. Cross-testing with manufacturer's TGA confirmed that spiking error was below 0.05 ng/mg (confidential data not shown). FXIa spiked at 1 ng/mg Ig in product A1 was confidently identified. Single CPLL fractionation (sequential elution and precipitation desalting) surpassed combined depletion and CPLL, with up to 20 identified peptides versus 12 peptides. Depletion alone led to only 2 peptides identification.

Using the same sample preparation, number of replicates and data analysis, FXIa spiked at 0.2 ng/mg Ig was only identified in one of two spiking experiments. This indicates that the method limit of detection for FXIa is situated between 0.2 and 1 ng/mg Ig.

4.4.2.2 Residuals characterization in different Ig products

The developed method led to the identification of about 70 different plasma residual proteins among the five different tested Ig products. Data analysis with Progenesis software (Nonlinear Dynamics, Waters) allowed relative quantification of these proteins by calculating normalized abundances of each of them in every tested sample (Table 9.3 in the Addendum). A satisfying automatic alignment (>80%) was obtained among the different CPLL-enriched Ig samples. A heatmap with hierarchical clustering of identified proteins in each tested sample according to their normalized abundances was generated to easily reveal residuals profiles (Figure 4.3).

Additionally, only fold changes exceeding a factor of 10 were further considered because this variation is more than 2 times the maximum fold change observed between two technical replicates for a given protein. Comparison of two batches from given IVIg product (A1 and A2) revealed only 3 proteins with more than 10-fold variation in abundance, further with CV above 30% (data not shown). On the other hand, more than twenty proteins were reported when comparing two different products, even from the same manufacturer (A and B). Table 4.1 highlights proteins which have at least 10-fold greater abundance in given product (with CV<30% among the 16 injections) compared to the median of their abundances in every tested product. This gives clear indication that each Ig product has its own plasma residual profile. In the TEE-positive batch, FXIa was found to be 66-fold more abundant. TEE batch was also distinguished from other regular products by several other proteins with higher abundance, such as fibrinogen, angiotensinogen, antithrombin-III, complement factor I, complement component C8, ficolin-3 or ectonucleotide pyrophosphatase/ phosphodiesterase family member 2 (See Table 4.1, Figure 4.3 and also Table 9.3 in the Addendum).

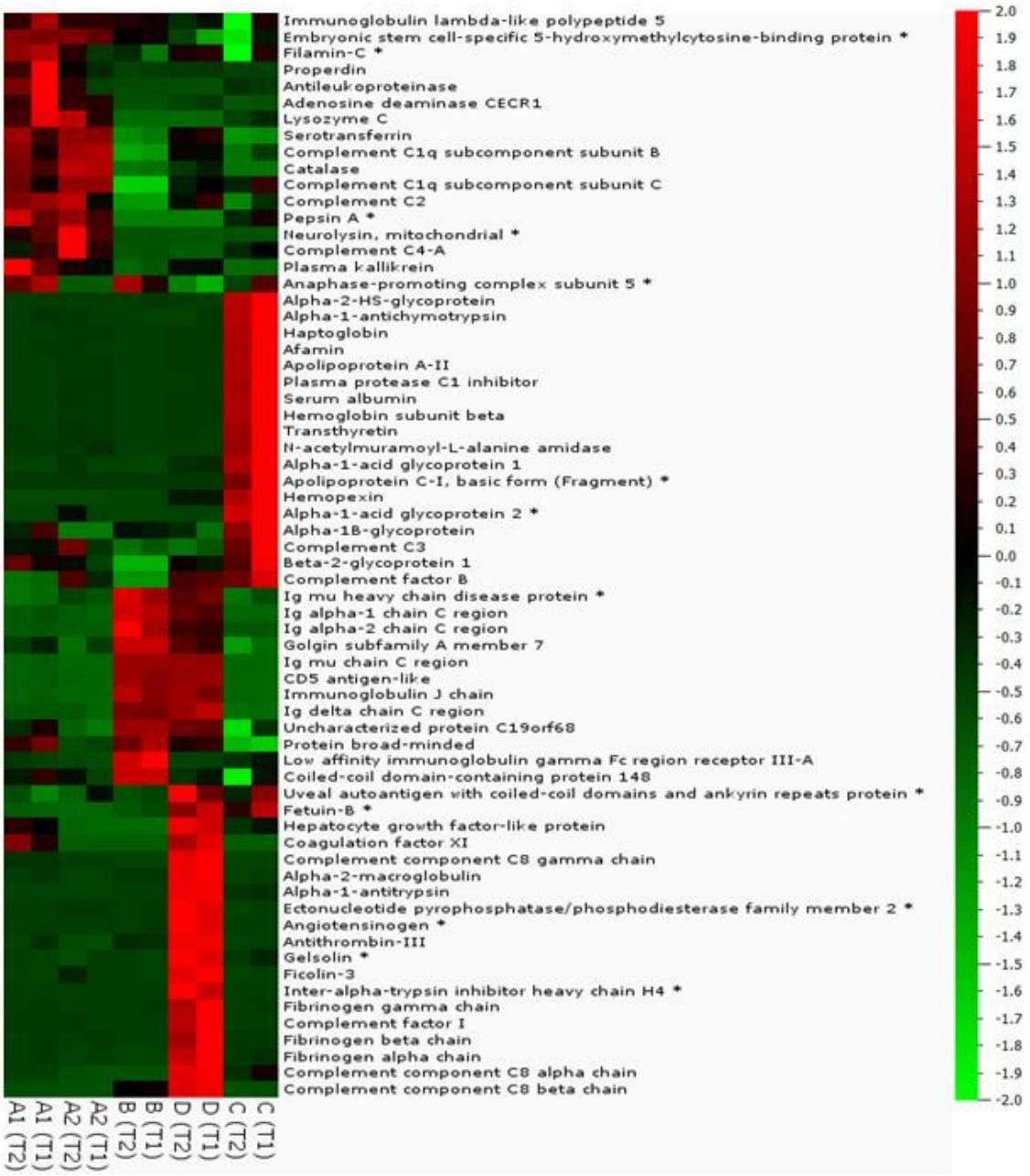


Figure 4.3 Heatmap of normalized abundances of (non-IgG) plasma proteins in two batches from IVIg product A (A1 and A2), two other IVIg products, one from the same manufacturer (B), one from another (C) and a TEE-positive ScIg batch (D). A1 and A2 are spiked with FXIa 0.2 ng/mg Ig. Each sample was enriched with ProteoMiner™ large capacity – sequential elution and desalted with protein precipitation (two technical replicates, T1 and T2). Each technical replicate was injected 8 times, giving a total of 16 runs per sample. Progenesis was used to calculate MS precursors normalized abundances and MS/MS spectra of those precursors were exported as MGF file to be searched in Mascot. Confidence score above 50 and at least 1 unique peptide were requested as identity threshold. * denotes single unique peptide matches. Heatmap is generated in Qlucore Omics Explorer, with data normalization for each protein across the five tested samples (mean= 0, variance= 1) and hierarchical clustering option.

Table 4.1 Comparison of plasma residual protein profile in five different Ig products. Those include: two IVIg products from a first manufacturer, including two different batches from one of the two products (A1, A2 and B), one IVIg product from another manufacturer (C) and one TEE-positive batch from ScIg product (D). Each sample was enriched with ProteoMiner™ large capacity – sequential elution and desalted with protein precipitation (two technical replicates, 8 injections of each). This table reports identified (non-IgG) plasma proteins with confidence score above 50 and at least 1 unique peptide and with normalized abundance exceeding more than 10-fold the median of their abundances in every tested sample. CVs in normalized abundance (among the 16 injections) are below 30% (except for ectonucleotide pyrophosphatase/phosphodiesterase family member 2: 33%). The proteins are listed in alphabetical order. * denotes single unique peptide matches.

Regular commercial IVIg				TEE-positive ScIg (2011)
Manufacturer 1			Manufacturer 2	Manufacturer 3
A1	A2	B	C	D
Neurolysin, mitochondrial *		CD5 antigen-like	Afamin	Alpha-1-antitrypsin
		Ig alpha-1 chain C region	Alpha-1-acid glycoprotein 1	Alpha-2-macroglobulin
		Ig alpha-2 chain C region	Alpha-1-antichymotrypsin	Angiotensinogen
		Ig mu chain C region	Alpha-2-HS-glycoprotein	Antithrombin-III
		Immunoglobulin J chain	Apolipoprotein A-II	CD5 antigen-like
			Apolipoprotein C-I, basic form (Fragment)*	Coagulation factor XI
			Haptoglobin	Complement component C8 alpha chain
			Hemoglobin subunit beta	Complement component C8 beta chain
			Hemopexin	Complement component C8 gamma chain
			N-acetylmuramoyl-L-alanine amidase	Complement factor I
			Plasma protease C1 inhibitor	Ectonucleotide pyrophosphatase/ phosphodiesterase family member 2*
			Serum albumin	Fibrinogen alpha chain
			Transthyretin	Fibrinogen beta chain
				Fibrinogen gamma chain
				Ficolin-3
				Gelsolin
				Ig alpha-1 chain C region
				Ig alpha-2 chain C region
				Ig mu chain C region
				Immunoglobulin J chain

4.5 Discussion

TEE outcome with Ig linked to co-purified plasma FXIa pointed out the need for additional control of trace amounts of plasma residuals. Each Ig product is expected to have its own profile of trace-level residual plasma proteins since both the origin and composition of the plasma pool source and the production process differ among manufacturers, products and even batches. MS-based shotgun proteomics is the ideal approach for the screening of complex biological samples and is gradually penetrating the world of plasma-derived therapeutics QC [112,119,267]. A key advantage of this approach over common QC assays (ELISA, Western blot, activity assays …), is that it does not require any prior information or dedicated development for each target. Despite continuous developments in MS technologies and the use of a high performance QTOF MS devices (TripleTOF 5600, Sciex), the extreme difference (about 6 orders of magnitude) between protein abundances of source material residuals and purified component in biologicals remains highly challenging: abundant peptides, here the Ig fragments, monopolize the total ion current in the MS [268]. The dynamic range of protein concentration should be reduced by means of protein-level fractionation, in addition to peptide-level LC separation [269]. In this work, two commonly utilized sample fractionation strategies were compared, either abundant component (Ig) selective depletion – with protein G - protein A affinity chromatography – or LAP non-targeted enrichment – with CPLL technology.

Protein G - protein A affinity chromatography is a well-known and effective process to specifically isolate antibodies, minimizing, by this, the risks of co-depletion [198–200,270]. Still, for our purpose, a major drawback of this depletion procedure is that the LAP of interest are found in the highly diluted flow through, which requires a concentration step, possibly causing protein trace losses [271]. Direct specific affinity-enrichment of the LAP cannot be considered here since their exact nature is not known a priori. One might think to use anion exchange chromatography to selectively elute the Ig fragments in the flow through, based on their pI, but LAP with similar pI could be lost [272]. Actually, anion exchange chromatography is already used by various Ig manufacturers during the purification process. Since impurities are still found in the purified Ig end-product, this demonstrates that, although efficient, the technique has insufficient specificity for the purpose here.

A way to circumvent this dilemma is to consider non-specific enrichment of LAP, with concomitant dilution of abundant Ig fragments. CPLL enrichment can be easily performed by using the ready-to-use ProteoMiner™ kits from Bio-Rad. “Large capacity” kit (100 µL settled bead volume) gave better results than the “small capacity” (20 µL settled bead volume): a larger volume of beads enlarges the ligand diversity and by this, the chance to fish more proteins. In all the cases, at least 5 to 50 mg of proteins per 100 µL of beads have to be loaded, in order to reach saturation of the high abundant proteins [215,266].

Generally, protein capture is performed under physiological conditions, using phosphate buffered saline (PBS) [214,266,273]. Since biological therapeutics are made with physiological buffers, this is the simplest option as no sample conditioning is required. Reducing the ionic strength of PBS – just by decreasing the NaCl concentration in PBS from 150 mM to 50 or 25 mM – enhances the CPLL beads binding capacity, and by this the potential harvest of rare proteins [207,214,274]. Nevertheless, this was found to be detrimental in our application: the increased binding capacity results in a lower discrimination for protein

capture and a less effective reduction of the Ig fragments concentration [214]. Subsequent recapture of the flow through at different pHs (to play on the influence of ionic interactions on binding capacity) was also not fruitful here. This comes most probably from the restricted diversity of residual proteins in Ig, compared to the more complex samples which were used by almost all publications (namely whole plasma, cell lysate, urine or plant extracts) [209,274–278]. To our knowledge, such application of CPLL to the characterization of extremely low-abundant residuals dates back to 2006-2007 [131,213]. In the two studies, CPLL was conducted with protein capture in PBS and elution reagent really close to the one of the Bio-Rad kit (9M urea-citric acid, pH 3.3-3.5, 2% CHAPS *versus* 8M urea-acetic acid 5%, 2% CHAPS) [131,213,266]. The more stringent "sequential elution" was found to be slightly beneficial. On-bead digestion is an interesting way to circumvent any risk of partial elution of captured proteins from the beads. Nevertheless, the absence of protein assay implies approximation in the enzyme to protein ratio during the digest. Moreover, applying a drastic harvest from the beads also means collecting more IgG peptides, which will be monopolizing a part of the MS ion current.

Both depletion and CPLL revealed (non-IgG) plasma proteins but CPLL enrichment allowed a higher number of identifications over depletion, even though one identification was lost. Complementarity of the two approaches was described previously using whole plasma samples [273,279,280]. Combining CPLL enrichment to prior Ig depletion chromatography process allowed an increase in the number of identifications but required about six more hours, plus an additional concentration/desalting step. Each fractionation step added means increasing the resolution of trace components but also the costs, the time spent on sample preparation and the risks of unspecific sample loss or artefacts. Single CPLL enrichment with capture under physiological conditions and sequential elution was further applied as sample fractionation method because it showed the best LAP coverage with convenient execution.

Downstream shotgun MS analysis was conducted in DDA, which remains the most conventional non-targeted acquisition mode, available on all MS devices. Nevertheless, DDA has an inherent variability [237]. In this study, considering a given trypsinized sample, maximum number of identified peptides (and thus proteins sequence coverage) was mostly reached after 8 merged runs, even if further improvement with more injections was also reported for a few proteins (including spiked FXIa). "Data-independent acquisition" (DIA) modes, where all the MS ions are further fragmented in MS/MS during one run (and not only the most abundant ones like in DDA), are expected to solve this repeatability issue. However, recent studies emphasize that, even with DIA, sample fractionation is still required to access LAP identifications [273,280]. The use of such a DIA mode in the context of sensitive residuals profiling will be discussed in the next chapter.

Since no information about given residuals concentrations in Ig is available, pure plasma-native FXIa (Calbiochem) was spiked in as a sensitivity marker, at QC-relevant ratios of 1 and 0.2 ng per mg of total protein. This spiking approach is limited by inherent variability (because of micropipetting small volumes and/or highly diluting the spike) but cross-testing with manufacturer's TGA (confidential data) revealed that the spiking error was below 0.05 ng/mg. The 1 ng/mg spike was identified with up to 20 peptides, which reflects the high method sensitivity. The 0.2 ng/mg spike was only identified in one of the two spiking experiments, indicating that the sensitivity limit was achieved. Due to unique MS ionization properties of each peptide, this sensitivity for FXIa cannot be extrapolated to other proteins. For relevant

QC proteins, sensitivity should always be checked by spiking experiments and parallel testing with a quantitative method.

Intensity-based label-free relative quantification on CPLL-enriched Ig samples allowed calculating fold changes in plasma LAP normalized abundances between different samples. The median CV in normalized abundances was 20% between five technical replicates and maximal fold change observed between two replicates was 4. The technical repeatability of CPLL technology with ProteoMiner™ spin columns was previously discussed elsewhere [273,279–283]. Consequent to this method variability, at least 10-fold changes in protein abundance were considered as significant biological variation. While a very similar pattern was found between two batches from a given IVIg product, each different Ig product revealed its own plasma residual proteins profile. About 15 to 40 residual proteins were identified in a regular IVIg product. This is a significant larger number of proteins than identified by Lackner *et al.* with a gel-based approach (ApoH, TRFE and HSA, the three proteins they identified, were also found)[93]. Nevertheless, this remains a quite low number of identifications compared to other proteomics applications, especially with high-end MS instrument and repeated runs. This corroborates the extremely low initial levels of residuals and the major contribution of IgG peptides to the total ion current in the MS, even after fractionation.

A TEE-positive batch was distinguished from regular Ig products with unique higher abundance identifications, including (non-spiked) FXIa but also various blood stream-regulator proteins (fibrinogen, angiotensinogen, antithrombin-III, complement component C8…). Interestingly, those more abundant proteins did not include ApoH and plasma kallikrein, which were previously reported as possible pro-TEE factors [91,93]. This profiling was possible without any initial information on residual fraction composition of the tested samples and without any target-specific development, which is a huge advantage. The TEE outbreak reported in 2010-2011 would most likely have been prevented if such a method had been applied. MS-based profiling could be applied every time the purification process (or the plasma pool source) is modified; to be sure no unwanted plasma residual is concentrated. This should not be seen as a substitute but rather as a complement to the routine specific quantitative testing for proteins of QC interest (like FXIa). For such specific quantitative purpose, MS also offers to build dedicated multiple reaction monitoring assays with (synthetic labelled) peptides spiking and standard curves.

4.6 Conclusion

In summary, this study highlights the use of MS non-targeted screening of source material residuals in biologicals. Sample pre-fractionation remains the mandatory first step to reach relevant sensitivity and bring residual plasma traces to the detectability level. Every technique will have slightly different LAP coverage and each added step should be discussed regarding the balance between relevant new identifications and the time spent on sample preparation (with associated risks of sample loss and/or artefacts). In a second step, conventional shotgun LC-MS/MS workflow provides spectral data which can be used for robust non-targeted identification, given the use of a database and selected quality criteria (search engine score, unique peptide-sequences). DIA mode is expected to solve the issue of conventional DDA inherent variability which asks for repeated injections and extended acquisition time. Finally, for the proteins with known critical impact on product safety and/or quality, sensitivity should be validated with spiking and ideally parallel testing with other specific assays.

4.7 Acknowledgements

Protein G - protein A affinity columns were provided by Plasma-Industries Belgium (formerly CAF-DCF). The authors are grateful to Sofie Vande Casteele for her excellent technical assistance.

Chapter 5:

Mass spectrometry-based sensitive screening of plasma residuals in human plasma-derived immunoglobulin *(part II)*:

data-independent acquisition

Adapted from "**Estimating the reliability of low-abundant signals and limited replicate measurements through MS2 peak area in SWATH.**"

Franck Limonier, Sander Willems, Geneviève Waeterloos, Myriam Sneyers, Maarten Dhaenens* and Dieter Deforce*

*Equal contribution

Proteomics (2018) 18 (24), e1800186
DOI: 10.1002/pmic.201800186

5. Mass spectrometry-based sensitive screening of plasma residuals in human plasma-derived immunoglobulin *(part II)*: data-independent acquisition

5.1 Abstract

SWATH-MS provides large-scale protein quantification with high accuracy and selectivity. As data-independent acquisition (DIA), SWATH is comprehensive in nature, as a single DIA run theoretically contains all the spectral information that a sample can deliver at a given limit of detection. Spectral assay libraries generated on a high resolution MS instrument are generally used to perform data extraction by matching SWATH spectral data with annotated spectra from the library. For sake of Ig profiling, an extended spectral library could be built directly from the plasma source material. Hereby, MS2 relative quantification data could be extracted over a larger number of potential contaminants during a single MS run. Nevertheless, reliable quantification of low-abundant signals in complex samples remains challenging, as recently illustrated in a multicentre benchmark study of different label free software tools. Here, the SWATH Replicates Analysis 2.0 template from Sciex is used to highlight that the relationship between the MS2 peak area and the variability can be described by a function. This functional relationship appears to be largely insensitive to variation in samples or acquisition conditions, suggesting a device-intrinsic property. By using a power regression, we show that the MS2 peak area can be used to predict the quantification repeatability without relying on replicate injections, thus contributing to high-throughput confident quantification of low-abundant signals with SWATH-MS. As proof of applicability for LAP quantification, in-house Python tool was built to flag the proteins that integrate a minimal number of transitions in the peak area range of desired repeatability. By doing so, a set of highly robust quantifications of plasma residual proteins in fractionated Ig samples is secured from large spectrum SWATH screening using extended spectral library.

5.2 Introduction

In search of a comprehensive and non-targeted characterization of complex proteomes by mass spectrometry, data-independent acquisition (DIA) has emerged as a very promising strategy [236]. Where traditional data-dependent acquisition (DDA) produces only partial tandem spectral information by fragmenting only a certain number of the most abundant precursor ions (MS1 ions), DIA systematically fragments all the MS1 ions, providing a full tandem spectral record of the sample. This minimizes the risk of stochastic under-sampling, especially for low-abundant proteins (LAP). At the same time, contrary to selected/multiple reaction monitoring (S/MRM), DIA does not require any target-specific information prior to analysis and is suitable for non-targeted applications. 'Sequential Windows Acquisition of All Theoretical Fragment Ion Mass Spectra' (SWATH) performs DIA by dividing the MS1 *m/z* range into adjacent *m/z* isolation windows which are independently and consecutively analysed [244]. Spectral assay libraries generated on a high resolution MS instrument (either in-house or by a reference lab which made

them publicly available) are generally used to perform data extraction by matching SWATH spectral data with annotated spectra from the library [249]. Since only transitions that are indexed in the library can be extracted, spectral assay libraries are often extended by fractionation and repeated DDA injections. Notably, prudence is called for to control false discovery rate (FDR), retention time alignment, and handling of modified and shared peptides [249–251,284]. With these quality measures in place, numerous successful discovery-based applications using SWATH-MS have been described so far, demonstrating SWATH's capability to provide high throughput relative quantification of up to thousands of proteins in complex samples [285–292].

SWATH-MS has an outstanding potential to go beyond discovery-based applications and could be used in clinical and quality control (QC) applications as well [261]. Indeed, DIA is comprehensive in nature, as a single DIA run theoretically contains all the spectral information that a sample can deliver at a given limit of detection. As any of the hundreds of different proteins present in the plasma pool source material could theoretically be unintentionally co-purified during the Ig manufacture process, rendering S/MRM less practical, an extended spectral library could be built directly from the plasma source material. Hereby, MS2 relative quantification data could be extracted over a larger number of potential contaminants during a single MS run. In the previous chapter, it was evidenced that up to 8 repeated injections per sample are recommended when performing DDA based MS1 relative quantification. A spectral library was built by using samples of plasma pool source material obtained from the Ig manufacturers. The SWATH 2.0 in Peakview data analysis pipeline was then applied as validated in a recent multicentre study benchmarking software tools for label-free SWATH quantification [254]. Important repeatability issues were observed.

Indeed, reliable SWATH quantification of lower abundant signals in complex samples remains challenging. Numerous studies, including a recent inter-laboratory study, report that proteins can be reliably quantified across four orders of dynamic range in complex samples using SWATH-MS on a TripleTOF 5600 system [244,261,291,293].

In common practice, multiple injections are performed and peptide or protein level CVs are reported to assess SWATH quantification repeatability. Importantly, these CVs are inferred by summing the MS2 ion intensities, the root level of SWATH quantification. While summing peak areas is indeed the most robust metric available today, proteins with CV below 20% can actually rely on a majority of ion quantifications with very poor repeatability. The MS2 ions level is the root level of SWATH quantification, so the repeatability criterion should thus be assessed there. An intrinsic property of analytical instruments is that their variability in measurement within a certain signal range can be predefined. In other domains, such a specification is depicted in the instrument specifications, as it is the case for e.g. a balance. In mass spectrometry however, this property is less well defined. In their multicentre study benchmarking software tools for label-free SWATH quantification, Navarro and colleagues evidenced marked differences in the lower abundance range between the software tools and found that selection of the most intense peptides generally resulted in lower variance [254]. "Intensity score" is incorporated in the mProphet data processing algorithm (now pyProphet) that is used by all these software tools for separating targets from decoys, i.e. "identification" by peptide-centric scoring [252,294]. Its weight in the confidence measure is calculated by semi-supervised learning. Quantification of peptides on the other hand is always done by summing the intensities of the transitions, wherein their weight is given by their peak areas.

For SWATH data, the SWATH Replicates Analysis 2.0 Excel template (Sciex, 2016) allows to plot average peak area against corresponding CV for up to forty thousand extracted MS2 ions drawn uniformly across

the data. Remarkably, this evidenced a functional relationship between variability (as estimated through the CV), and the root measurement signal (here the MS2 peak area) that appeared to be greatly conserved across all tested samples, conditions, and devices. To explore the generalization of this relationship, publicly available reference data from Navarro and colleagues was also analysed: equivalent relationship between CV and MS2 peak area was observed. Using a simple power regression to model what appears to be a device-intrinsic property, we were able to perform accurate predictions on quantification repeatability based only on the MS2 measurement signal and without relying on replicates and CV calculation.

Consequently, the applicability of the functional relationship in the context of LAP profiling in Ig samples was explored. We built a Python tool that enables to automatically flag transitions falling in a user-defined peak area range corresponding to desired repeatability (as calculated through the power regression), while the peptides and proteins that integrate at least a certain number of such transitions are evidenced. This enabled to secure highly repeatable quantifications of plasma residual proteins in the Ig samples. As these samples were the same as the ones profiled in the previous chapter, DDA-based MS1-level and SWATH-based MS2-level relative quantifications could be finally compared.

5.3 Material & methods

5.3.1 Samples

Fractionated Ig samples were the same as the ones used for DDA MS1 relative quantification in the previous chapter. Aliquots of plasma pools used as source material for products A, B and C respectively were also obtained from manufacturers to build an in-house plasma library. MS compatible human protein extract digest (Promega, Madison, USA) was used to confirm MS2 area to CV relationship with another sample type. Plasma pools were fractionated using optimized incubation at 3 different pH (pH7, pH9.3 & pH4) in parallel and on-bead digestion, as recommended by Boschetti, Righetti and colleagues [214,215]. Samples were tryptic digested *in solution* or on-bead as referred in [295].

5.3.2 Liquid chromatography

Dried peptides were dissolved in 0.1% formic acid (FA) in water. Every sample to be injected was spiked with iRT peptides (iRT Kit, Biognosys, Zurich, CH) at 0.1 µL of 10x iRT Standard (see manufacturer user guide) per injection for retention time (RT) calibration. Peptides were separated on an Eksigent nanoLC 425 (Eksigent, Sciex, Framingham, MA, USA). Two, four or eight micrograms of sample was first trapped on a pre-column (YMC-Triart C18, pore size 12nm, particle size 3µm, 0.5 mm internal diameter by 5 mm; YMC, Dinslaken, Germany), and separated on reversed-phase (C18) microflow (5µL/min) LC column (YMC-Triart C18 15 cm, particle size 3 µm, 0.3 mm internal diameter by 150 mm; YMC, Dinslaken, Germany) using a linear gradient of 95:5 buffer A (95% H_2O; 5% DMSO; 0.1% FA) /buffer B (95% ACN; 5% DMSO; 0.1% FA) to 20:80 buffer A/buffer B at 5 µL/min over 1 or 2 hour(s). LC peak width was around 30s.

5.3.3 Mass spectrometry

MS analysis was performed using a TripleTOF 5600 mass spectrometer (Sciex, Framingham, MA, USA) operated in positive ionization mode with a Duospray Ion Source. All MS parameters were controlled by AnalystTF software 1.7 (Sciex). The source conditions were as follows: temperature: 80°C; curtain gas (CUR): 30 psi; ion source gas (GS) 1 and 2: 13 and 0 psi respectively; and ion-spray voltage floating (IVSF): 5.5 kV. MS1 and MS2 declustering potential was set at 100V. MS2 Ion release delay and ion release width were 67 and 25 respectively.

5.3.3.1 DDA runs for library generation

Fractionated Ig sample library was run (8 times) with the same DDA method as used in previous DDA work [295]. Extended spectral library in-house generation by using fractionated plasma samples (8 injections for each pH of fractionation) and the pure spikes (FXI, FXII & PK, 3 injections each) was performed in accordance to recent published guidelines [249,250,284]. Briefly, each cycle consisted of one full MS1 scan (*m/z* 400-1250) of 250 ms followed by MS2 data-dependent trigger events (m/z 65-2000, high sensitivity mode). A maximum of 20 candidate ions (charge state +2 to +4) exceeding 150 cps were monitored per cycle, with an accumulation time of 150 ms and using a rolling collision energy (CE) with a spread of 15V. Cycle time was 3.3 s, in order to have 8 to 10 data points per LC peak. Dynamic accumulation and dynamic background subtraction were switched on. Surpassing ions were excluded from further analysis for 20s.

5.3.3.2 SWATH runs

SWATH runs consisted in successive cycles of one full MS1 scan followed by consecutive Q1 isolation windows for MS2 fragmentation, covering a precursor mass range of *m/z* 400-1250. Optimized isolation window widths across the *m/z* range were calculated using the Variable Window Calculator Excel tool from Sciex (download from http://www.absciex.com/software-downloads). Cycle time was fixed at 2.8 s, in order to have 8 to 10 data points per LC peak. The MS1 full scan had an accumulation time of 250 ms, with collision energy (CE) of 10 eV. Each isolation window had an accumulation time of 25, 50 or 100 ms, depending on the number (respectively 100, 50 or 25) of windows (with 1 *m/z* overlap on the lower side of the window). Rolling CE (charge state: +2) with a CE spread of 15 eV was used, and MS2 was conducted either in high sensitivity (HS) or high resolution (HR) mode.

5.3.3.3 Spectral libraries generation

Raw spectra data (wiff) files from library DDA runs were converted into peak list format and searched against human reviewed protein database (downloaded from UniProtKB /Swiss-Prot in June 2016 and complemented with internal standards, iRT and crap peptides [296]) using Protein Pilot (V 4.5, Sciex) with the Paragon algorithm. The search parameters were set as follows: sample type: identification; Cys alkylation: MMTS; digestion: trypsin; special factors: none; ID focus: allow biological modifications; search effort: thorough ID; results quality: detected protein threshold > 0.05; false discovery rate option ticked. The .group files resulting from the database search were directly imported into Peakview (V2.2 with SWATH Acquisition MicroApp 2.0, Sciex) with exclusion of shared peptides. Protein FDR threshold was set at 1% by limiting the maximal number of proteins to import, according to the ProteinPilot FDR analysis. Libraries were saved as text files. Full lists of indexed proteins (FDR <1%, shared peptides excluded) in the different libraries can be found in Table 9.4 of the Addendum. Of note, individual DDA fraction RT alignment as suggested by Zi et al. was found to have almost no impact (linear regression $y = a.x + b$ with 0.98 < a < 1.02 and $|b|$ < 0.25, R^2 > 0.99) [284]. An external fractionated plasma library was also provided by Sciex as .txt file. Library generation details as referenced by Sciex are to be found in Note 9.1 of the Addendum.

5.3.3.4 SWATH-MS data processing

Peakview V2.2 with SWATH Acquisition MicroApp 2.0 (Sciex) was used to perform peak extraction with maximum 6 transitions per peptide, XIC extraction window of 10 min and XIC width (mass tolerance) of 50 ppm (Table 5.1). Peptide filter settings included a peptide confidence threshold being the ProteinPilot reported confidence corresponding to 1% FDR during the database search of the library files. Shared and modified peptides were both excluded. The SWATH 2.0 integrated FDR threshold was set at 1%. RT calibration was performed using reference peptides from the iRT-kit (Biognosys) spiked into each sample, so that every extracted peptide was aligned to this LC device-independent RT unit system. After the peak extraction, the software automatically sums up extracted MS2 ion areas to give peptide areas which in turn are summed up to give protein areas. The ion, peptide and protein level peak areas were exported in Excel format.

Table 5.1 SWATH data processing parameters. * Protein level FDR analysis in Protein Pilot: Number of proteins detected at critical FDR of 1%. ** ProteinPilot reported confidence (%) corresponding to critical FDR (1%) at peptide level during the database search for the library building. For external libraries and merged libraries, default value of 99% is applied.

Data analysis step		Parameter	Value
Import library (Peakview)		Maximum number of proteins to import	See Protein Pilot protein FDR*
		Do not import shared peptides	ticked
		Select sample type	unlabelled
		Convert sample type to	unticked
Processing settings (Peakview)	Peptide filter	Number of peptides per protein	2000
		Number of transitions per peptide	6
		Peptide confidence threshold % (0-99)	See Protein Pilot peptide FDR**
		False discovery rate threshold % (0-100)	1.0
		Exclude modified peptides	ticked
		Exclude shared peptides	ticked
		Fix Rank	unticked
	XIC options	XIC extraction window (min)	10.0
		XIC width (ppm)	50
MS2 repeatability flagging (Python script)		CV threshold % *(if CV based criterion)*	20
		Lower bound area range (l) *(if area based criterion)*	20000
		Upper bound area range (u) *(if area based criterion)*	1000000
		Minimal number of replicates passing the criterion (r)	1
		Analysis depth (d)	3 (to protein level)
		Minimal number of flagged ions required per flagged peptide (n)	1 to 6
		Minimal number of flagged peptides required per flagged protein (m)	1 to 6

Graphs of the relationship between MS2 area and CV were generated with the SWATH Replicates Analysis 2.0 Excel sheet from Sciex (download from http://www.absciex.com/software-downloads). This relationship was also evidenced in publicly available reference data (ProteomeXchange identifier PXD002952) [254].

A Python script was developed to automatically flag the extracted ions that have their peak area in a defined range and/or their replicate injections CV below a certain threshold. Peptides having at least a minimal number of flagged ions were flagged accordingly. Finally, proteins had to have at least a minimal number of flagged peptides to be flagged. The Python script is available on Dropbox [1] and its parameters are detailed in the Table 5.1.

5.4 Results & discussion

5.4.1 Automatic results and repeatability issues

Automatic results were generated using Peakview SWATH 2.0 with processing settings following recommendations of the benchmark study and summarized in Table 5.1 [254]. As expected, extended libraries built from plasma allow extracting more ions and thus quantify more proteins as compared to the homologous library built from the Ig sample itself. At the same time, less than 50% of the quantified proteins using such extended library had a CV below 20% (Table 5.2). Normalization to total extracted areas in each run did not improve repeatability (Table 5.2). Two main reasons for this technical variability are that extended libraries are expected to (i) extract lower abundant ion signals with intrinsically higher CV that were missed during the DDA analysis of the sample, and (ii) index proteins that are in fact not present in the sample at all, increasing the number of false signals disproportionally to the amount of added true signals.

Table 5.2 Automatic results with Peakview SWATH 2.0 processing. Sample is fractionated Ig (A1) spiked with FXIa 1ng/mg. Three different libraries were tested: (i) homologous fractionated Ig (ii) in-house fractionated plasma source material and (iii) external fractionated plasma. Total counts of quantifications and percentage with CV < 20% at MS2 ion, peptide and protein levels.

Library	SWATH quantifications in fractionated Ig (A1) spiked with FXIa 1ng/mg				
		MS2 ions level		Peptide level	Protein level
			Total area normalized		
Fractionated Ig library	Total count	1802	1802	327	91
	Percentage with CV < 20%	68.48	65.32	84.40	85.71
Fractionated Plasma library (in-house)	Total count	1925	1925	348	124
	Percentage with CV < 20%	32.88	32.10	49.43	47.58
Fractionated Plasma library (external)	Total count	2597	2597	461	167
	Percentage with CV < 20%	29.73	29.77	44.69	46.71

[1] https://www.dropbox.com/sh/yvmdgjxe9ck5jrv/AACgSqm00rfmUnF5tDRFEH1ca?dl=0

Moreover, some proteins with CV below 20% demonstrated to mainly integrate ion quantifications with very poor repeatability (see examples in Table 9.5 in the Addendum). Two main phenomena contribute to this inference bias: (i) A “weight effect”: the elements with the highest peak area skew the error of the combined output because of their larger relative weight on the final summed peak area. (ii) The reduction in CV from ion to peptide and peptide to protein in part results from the fact that by summing the individual ion or peptide areas, upwards and downwards deviations in the measurements can average out and reduce CV of the higher level. Together, while summing peak areas is indeed the most robust metric available today, it is at least misleading to only report the summed CV. The repeatability was thus next investigated at the root level of SWATH quantification, namely the MS2 ions level.

5.4.2 MS2 peak area to CV functional relationship

5.4.2.1 A device intrinsic property

As already introduced, the variability in measurement and the measurement signal on an analytical instrument are expected to be linked. To our knowledge, this relationship is less well-defined in mass spectrometry. We used the SWATH Replicates Analysis 2.0 Excel template (Sciex, 2016) to investigate both in-house and reference public data generated using SWATH 2.0 in Peakview.

Publicly available data consisted in the high-complexity benchmarking datasets of Navarro and colleagues. Briefly, they measured two human-yeast-E.coli (HYE) hybrid proteome sample sets, providing several thousand proteins present at defined relative ratios, enabling in-depth statistical evaluation of quantification across a dynamic range of several orders of magnitude. A first sample set (“HYE124”) was tested on two TripleTOF systems, respectively 5600+ and 6600, with either 32 fixed or 64 variable windows, and a second (“HYE110”) exclusively on TripleTOF 6600, with 32 or 64, fixed or variable, windows [254]. The Peakview SWATH 2.0 results from this study (samples A of each sample set - ProteomeXchange identifier PXD002952) were imported into the Replicates Analysis tool. Figure 5.1 shows how the relationship between variability (as estimated through the CV), and the root measurement signal (here the MS2 peak area) can be visualized by using the template.

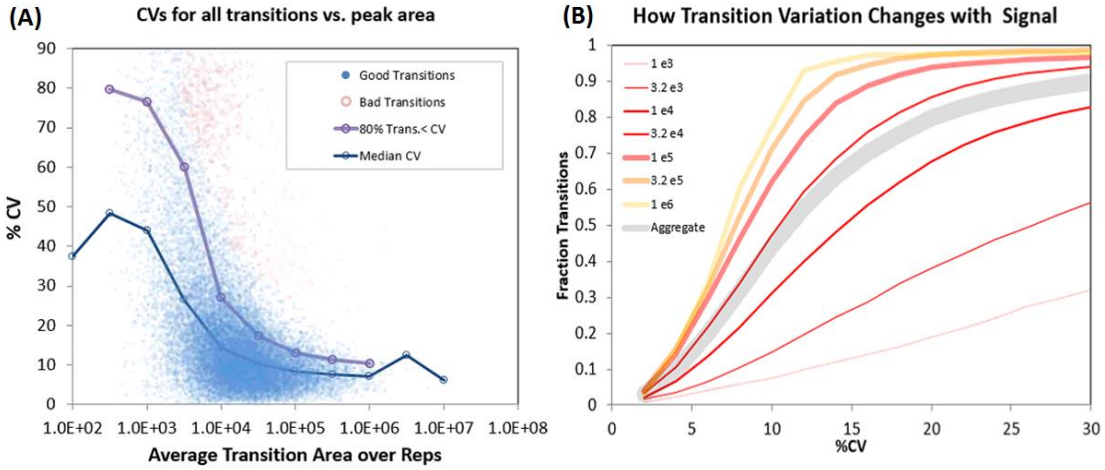
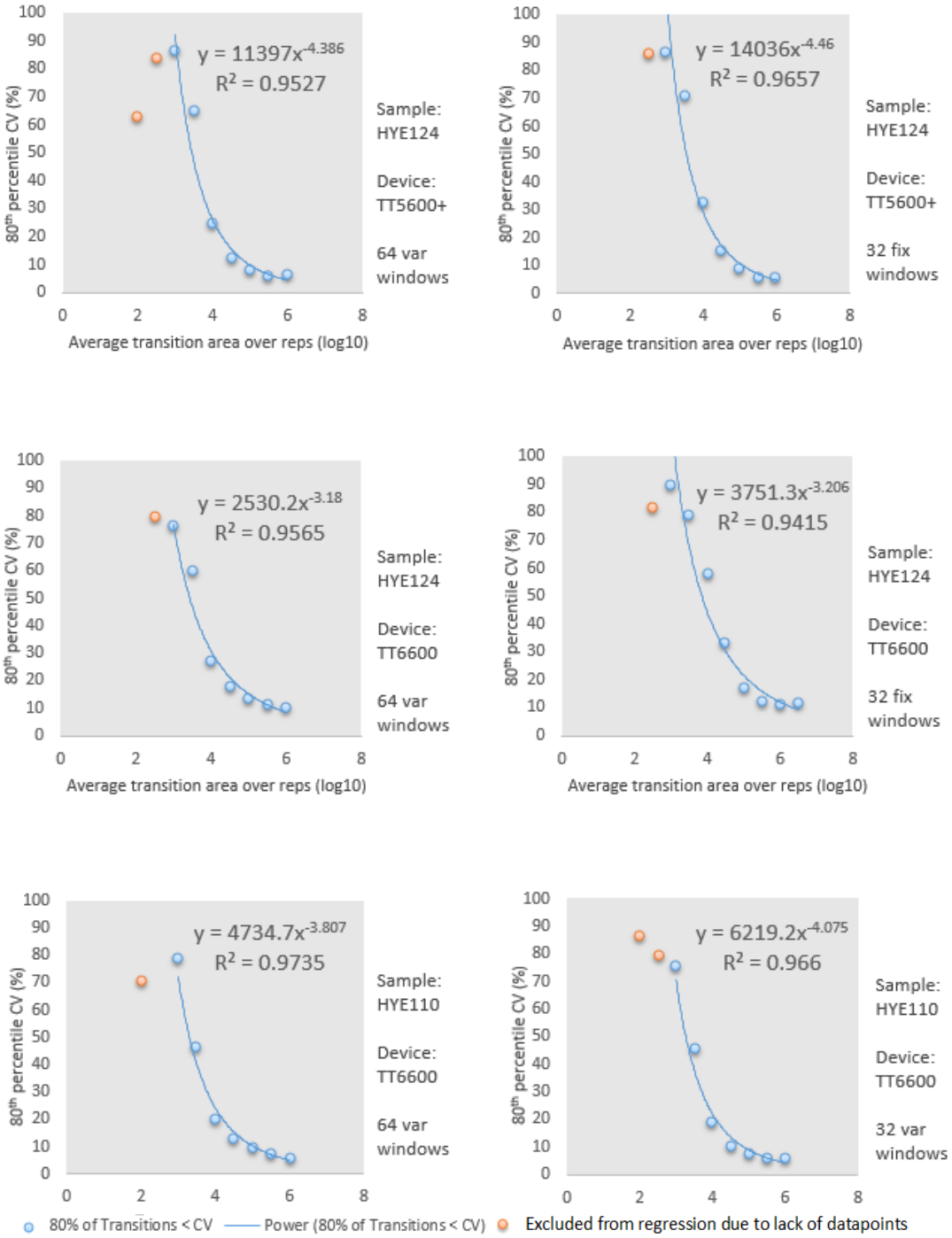


Figure 5.1 Transition peak area to CV relationship generated with the SWATH Replicates Analysis 2.0 Excel sheet (HYE124 sample A on TT6600, 64 variable windows).
(A) Dot plot representing the %CV measurements over triplicate injections. Transitions are defined by the software as "good" or "bad" based on whether the %CV is more or less than 3 times the median CV for area. Nevertheless, all the ions are included for the calculation of the median and 80th percentile lines. The percentile value is user-defined.
(B) How transition variation changes with signal (transitions are clustered into signal intensity groups)

Equivalent relationship between CV and MS2 peak area was observed in all the investigated samples. It has long been understood that CV increases with decreasing signal. Here, it was further evidenced that a functional dependence between the 80th percentile CV and the log-transformed peak area can be modelled by a power regression ($y = a \cdot x^b$) ($R^2 > 0.9$, see Figure 5.2). Additionally, the overall power regression is greatly conserved (correlation coefficient (r) > 0.99) across all tested samples (HYE124 and HYE110 from Navarro or in-house Ig and human protein extract samples), conditions (sample load, LC gradient length, number of windows, MS2 acquisition mode and spectral library) and devices (TripleTOF 5600, 5600+ or 6600) (Figure 5.3). This supports the hypothesis that the functional relation between repeatability and signal intensity is an intrinsic property related to the instrument, such as detector capabilities.

This in turn implies that power can be borrowed from this function. Indeed, the depicted functional relationship integrates around 40K data points, where usual CV approach only considers the number of replicates for each point individually. Note that it was not possible to model the instrument performance in the lowest peak area range ($\log_{10}$ peak area < 3), and the sometimes observed re-decrease in variability was neglected as well (orange dots in Figure 5.2), because there is a lack of data points in this region (as seen in Figure 5.1). On the other extremity of the peak area scale, the sometimes observed re-increase in variability might be linked to detector saturation, but there are too few data points in the region to confirm this hypothesis (Figures 5.1 & 5.2). When considering the Ig samples, the peak area range above 100 000 (log10 peak area > 5, orange dots in Figure 5.2) was neglected because this lack of data points hampers the regression.

(A)


y = 11397x^-4.386
R² = 0.9527
Sample: HYE124
Device: TT5600+
64 var windows
y = 14036x^-4.46
R² = 0.9657
Sample: HYE124
Device: TT5600+
32 fix windows
y = 2530.2x^-3.18
R² = 0.9565
Sample: HYE124
Device: TT6600
64 var windows
y = 3751.3x^-3.206
R² = 0.9415
Sample: HYE124
Device: TT6600
32 fix windows
y = 4734.7x^-3.807
R² = 0.9735
Sample: HYE110
Device: TT6600
64 var windows
y = 6219.2x^-4.075
R² = 0.966
Sample: HYE110
Device: TT6600
32 var windows
80th percentile CV (%)
Average transition area over reps (log10)
80% of Transitions < CV
Power (80% of Transitions < CV)
Excluded from regression due to lack of datapoints

(B)

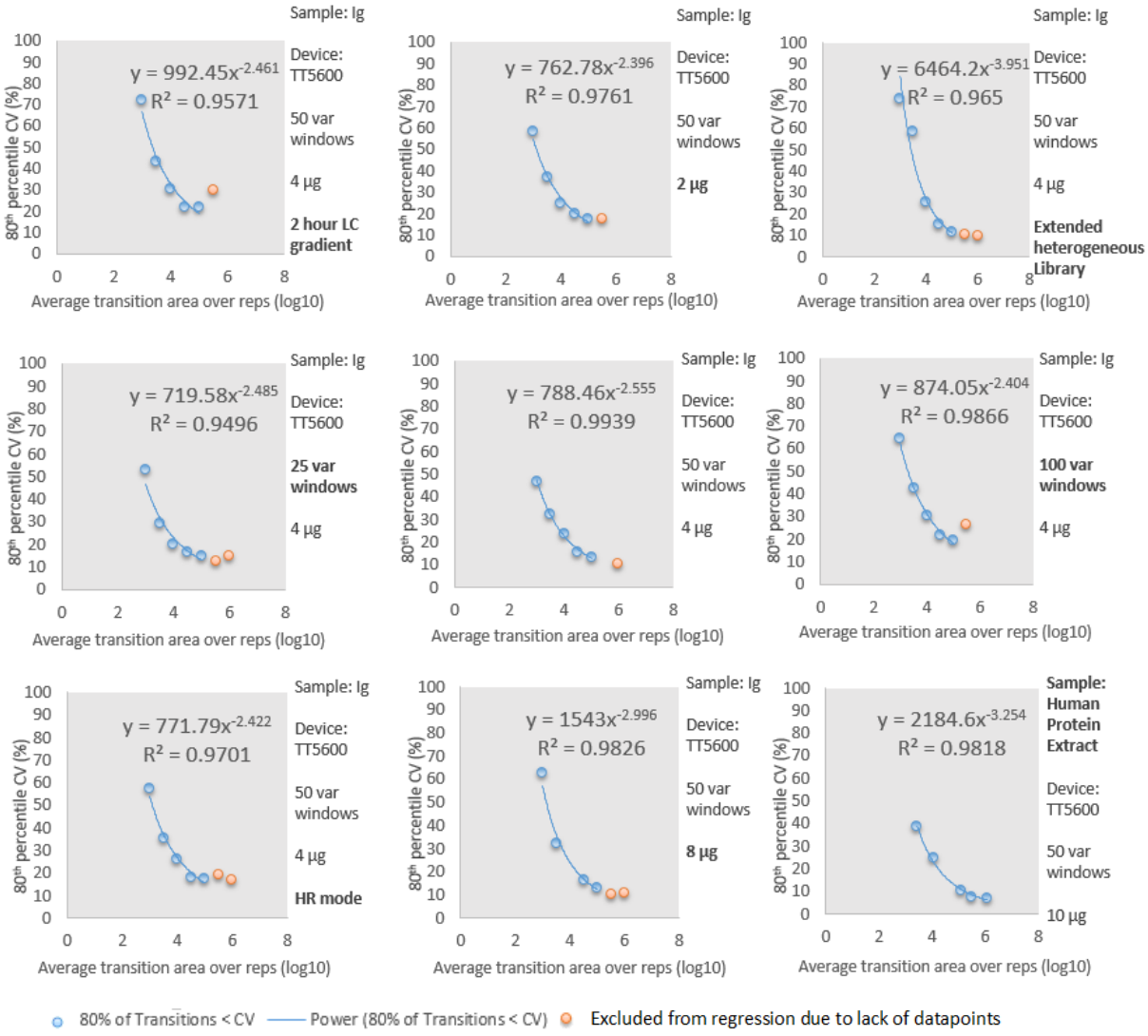


Figure 5.2 MS2 peak area (log10) to 80th percentile CV functional relationship modelling with power regression - Power regression from Replicates Analysis 2.0 graphs for different samples and conditions:

(A) Publicly available data [254]**: HYE124 and HYE110 samples using different number of windows (32 or 64, fixed or variable) and devices (TripleTOF 5600+ or 6600).**

(B) In-House generated data: Ig and human protein extract samples on TripleTOF5600, using different sample loads (in µg), LC gradient length (1 or 2 hours), MS2 acquisition mode (High Sensitivity-HS by default or High Resolution-HR), number of windows (25, 50 or 100 variable windows) and spectral library (homologous or extended heterogeneous)

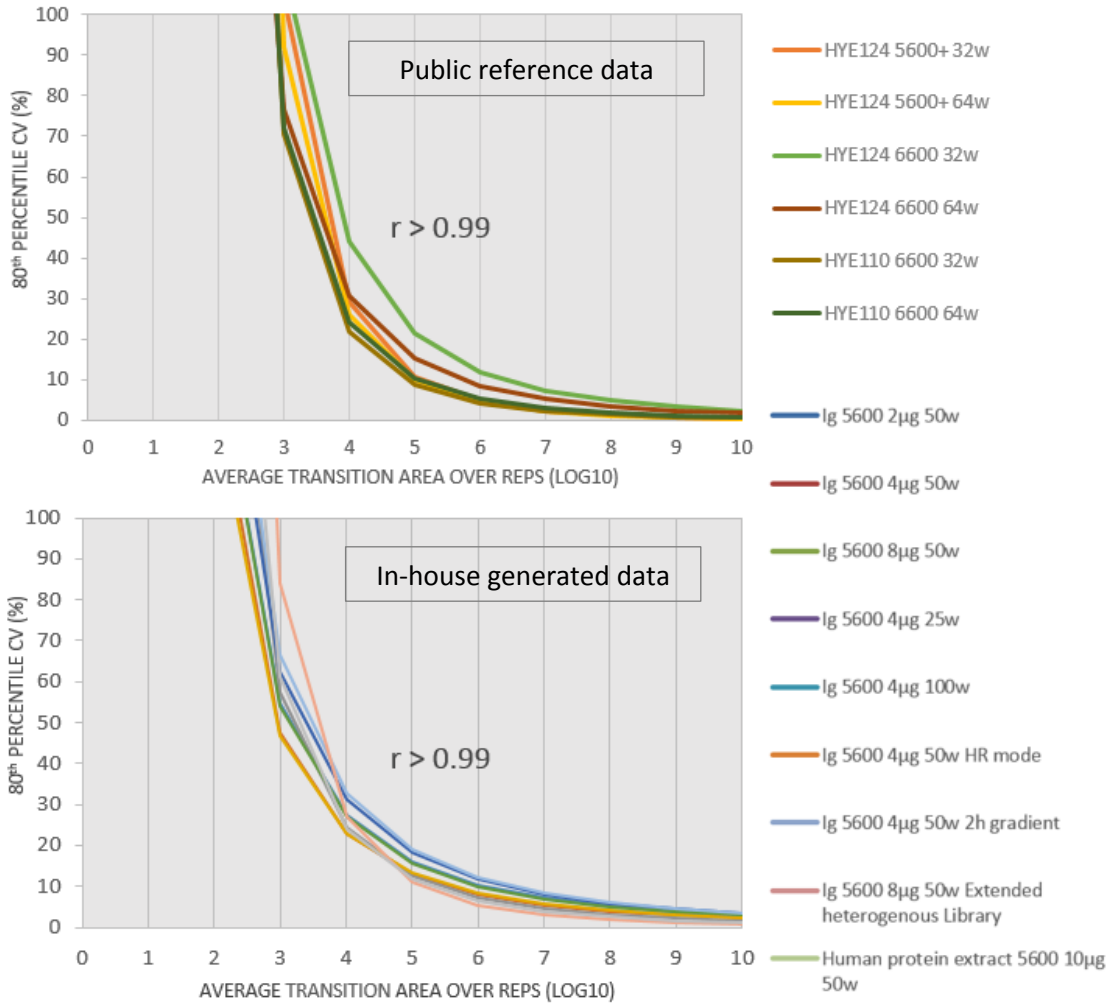


Figure 5.3 MS2 peak area (log10) to 80th percentile CV (%) functional relationship modelling with power regression - Correlation of the power regression across the different conditions tested: different samples (HYE 124 or HYE110, Ig or human protein extract), sample load (in µg), LC gradient length, MS2 acquisition mode (High Sensitivity by default or High Resolution-HR), number of windows (w), spectral library (homologous by default or extended heterogeneous) and devices (TripleTOF 5600, 5600+ or 6600). Correlation coefficient (r) is above 0.99.

5.4.2.2 An actionable point for repeatability assessment

Once the user has outlined the functional dependence of the employed device in his/her own lab, he/she can now estimate the probability that a given measurement will meet given repeatability requirements, e.g. 80% chance that the data has a CV below 20%. This can be illustrated by a concrete example. In the data of Navarro and colleagues, the estimate function between the 80th percentile CV (y) and MS2 peak area (x in $\log_{10}$ scale) for the TripleTOF6600 is $y = 2530.2 \times x^{-3,18}$ using the triplicate injection of the sample HYE124. This means that an ion with a peak area of 1000 has 80% chance that the CV is below $y = 2530.2 \times 3^{-3,18}$ = 76.90%. On the other hand, an ion with a peak area of 100 000 has 80% chance that the CV is below $y = 2530.2 \times 5^{-3,18}$ = 15.15%. In other words, one would consider an ion with a peak area of 100 000 as more reliable for quantification than an ion with a peak area of 1000.

Obviously, the function also enables to estimate the peak area value from which a certain level of repeatability is reached. Here, it can be predicted that the MS2 ions with a peak area equal to or greater than $x = \sqrt[b]{y/a} = \sqrt[-3.18]{20/2530.2}$ = 4.58 i.e. around 38 000, have at least 80% chance to have a CV<20%. This gives the user a better idea on how low-abundant signals can be quantified without compromising the repeatability and quantification confidence. To verify these predictions, the power regression function was applied to the other sample HYE110 that was independently acquired on the same device. For this, we picked the transitions with an area between 950 and 1050 as well as the transitions with a peak area between 95000 and 105 000 and compared their CV on triplicate injections (Table 9.6). 79% of the first group had a CV below 77%, while 95% of the second group had a CV below 15%. Then, all the transitions with a peak area above 38000 were selected. The increased reliability of this subset of transitions was clear-cut: 96% of these transitions had a CV<20%, while only 55% of all the extracted transitions passed the same CV (Table 9.7).

To further explore the potential benefits of the functional relationship in the context of LAP profiling, we built a Python tool that enables to automatically flag transitions falling in a user-defined peak area range corresponding to desired repeatability (as calculated through the power regression), while the peptides and proteins that integrate at least a certain number of such transitions are evidenced. Using the power regression of the 80th percentile CV to MS2 peak area relationship of our in-house TT5600, it can be calculated that at least 80% of the data with a peak area above 20000 has a CV below 20%. (With x = MS2 peak area in $\log_{10}$ scale and y = 80th percentile CV, and given the power regression $y = a \cdot x^b$, then $x = \sqrt[b]{20/a}$.) Given the lack of data points and the eventual detector saturation, an upper limit of 1.0E+06 was also set. Consequently, the Python tool was used to flag MS2 ions in the 2.0E+04 – 1.0E+06 peak area range. Distinction of different subgroups among the obtained quantifications based on the MS2 peak area through this Python script should never be seen as a hard cut-off criterion. Still, it will be demonstrated below that the use of this script gives an extra evidence that the highlighted relationship can serve as an actionable point for increased confidence in LAP quantification with SWATH, without relying on the number of replicates.

5.4.3 Optimizing confident quantification of LAP in an abundant background matrix

In general, acquisition parameters are adjusted in discovery proteomics in a way that yields more protein quantifications. However, the present work requires surfacing low abundant contaminants, namely the non G class Ig (IgG) plasma proteins, from an abundant background matrix, namely the IgG fragments. Therefore, the impact of different acquisition parameters was compared regarding not only the total number of transitions, but also the number of transitions in the peak area range of high repeatability (here 2.0E+04 – 1.0E+06), as flagged with the Python tool (Figure 5.4). The sample coverage (total number of quantifications) was increased by increasing specificity with a higher number of windows (with shorter accumulation times) and an increased LC gradient length, while the sample load and the mass resolution mode have only minor impact. On the other hand, more repeatable quantifications were evidenced with the proposed peak area-based flagging approach when sensitivity is preferred: increased sample load, fewer windows with longer accumulation time, shorter LC gradients and high sensitivity mass resolution mode (Figure 5.4). Thus, flagged quantifications appeared to be remarkably robust and largely insensitive to interferences.

Here, the method application for the Ig products profiling was performed by injecting 8µg, using a 1-hour gradient and performing the SWATH acquisition with 50 windows operated in HS mode.

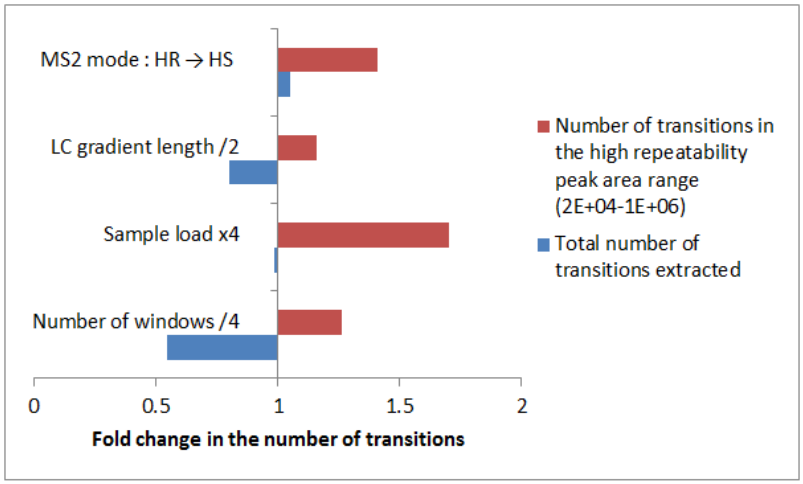


Figure 5.4 Impact of different conditions on the MS2 peak area distribution: the MS sample load (2,4 or 8µg), the LC gradient length (1 or 2h), the MS2 acquisition mode (High Resolution-HR or High Sensitivity-HS) and the number of variable windows (25, 50 or 100). Cycle time is 2.8s. Sample is in-house fractionated Ig, default parameters are: 4µg injection, 1h LC gradient, MS2 in HS and 50 variable windows.

5.4.4 Method application and relevance for Ig profiling

5.4.4.1 Choice of the spectral library

Extended heterogeneous spectral libraries were built with fractionated plasma coming from the Ig manufacture source material, in order to index a large number of possible residuals in the Ig samples under investigation. Full lists of indexed proteins (FDR <1%, shared peptides excluded) in the different libraries can be found in Table 9.4 (in the Addendum). External library downloaded as .txt file as provided by Sciex was found to index the greatest number of proteins, even if each library exhibited its own profile (Figure 5.5 A). When using three different plasma pools (each related to the manufacture of product A, B and C, respectively) to build corresponding libraries in-house, around 80% of their content at the protein level was found to be common.

Of note, a single library had to be chosen to compare quantification values across different samples. In first instance, an externally available plasma library appeared as the most convenient option. Still, we have shown that automatic results should be taken cautiously (Tables 5.2 & 9.5). With the Python tool to flag transitions in the area range of high repeatability at hands, no increase in the total number of flagged quantifications was reported using one of the plasma libraries compared to the Ig library (Figure 5.5 B). Obviously, some Ig fragments were better quantified with the fractionated Ig library. At the same time, few non-IgG proteins were flagged with the Ig library and not the plasma libraries. This can be explained by one of the three following cases: (i) the protein is not indexed in one of the libraries, (ii) the protein is always indexed and extracted but not always flagged or (iii) the protein is always indexed but not always extracted. The two first cases could be solved by respectively complementing the library or tolerating a less strict flagging criterion.

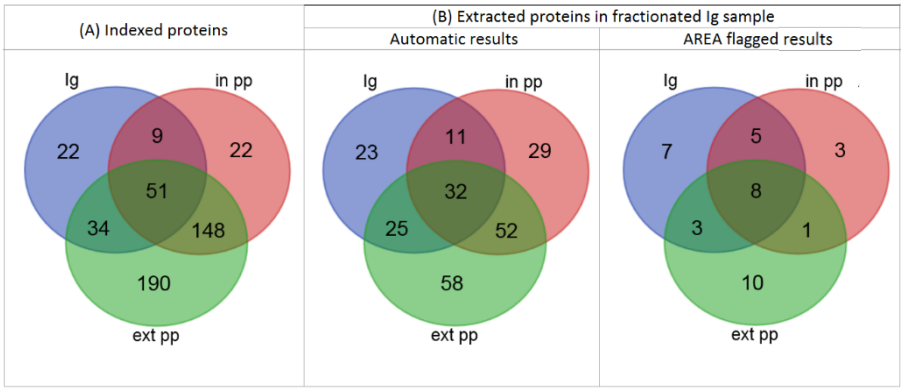


Figure 5.5 Spectral libraries comparison. Considering a given Ig sample (A1 spiked with FXIa 1 ng/mg), the three compared libraries are (i) in-house homologous library built with DDA runs of the fractionated Ig sample itself ("Ig") (ii) in-house heterogeneous plasma library built with DDA runs of fractionated plasma pool sample of the corresponding source material for manufacture ("in pp") and (iii) external heterogeneous plasma library downloaded as txt file ("ext pp").

(A) Indexed proteins overlap in the three libraries.

(B) Protein quantifications overlap in the considered (fractionated) Ig sample, depending on the chosen library. SWATH 2.0 automatic results and flagged results using peak area criterion (at least 2 peptides which have at least two ions passing the criterion) are successively considered.

Concerning the indexed proteins which were not extracted, this scenario was only observed when comparing the results with an in-house library to those with the external library. By investigating this issue, it was found that the external library sometimes indexes different peptides for a given protein, as compared to the ones reported in the in-house runs. This emphasizes that increased quality control becomes increasingly important as the library becomes more distant from the actual sample, being in-house fractionated or from an external deep-mining origin. To find a compromise between library completeness and indexed spectral information homology, the in-house plasma library was preferred. Moreover, this is the only one which can capture plasma pool source material variability, as still 20% of the indexed spectral information was found to differ at the protein level. In order to compare products manufactured from different plasma pools, all the three generated in-house plasma libraries (coming from the manufacture of products A, B and C respectively) were here merged together.

5.4.4.2 Sensitivity markers

Three different plasma proteins, namely FXI, coagulation factor XII (FXII) and prekallikrein (PK), were spiked at different mass ratios in a non-fractionated Ig sample to outline the link between the proposed flagging approach and the sensitivity and dynamic range. To ensure library adequacy, DDA runs of these pure spikes were included into the spectral assay library. Mean extracted area for these three spikes in the different experiments are reported in Figure 5.6. The spike was considered as quantified above the LOQ if both the preceding and following dilution steps gave significantly different areas (Mann-Whitney test, p-value < 0.05). It is to be noted that 8µg injections were performed to maximize sensitivity (cf section 5.4.3) but quantification of the spikes at 1:100 ratio apparently suffered from detector saturation at that sample load, as higher than expected CV were observed. Each spike was quantified within three to four orders of magnitude in the concentration dynamic range. Note that with each peptide having its own ionization properties, the sensitivity of a label-free MS assay cannot be extrapolated as is to other proteins. Still, other publications report comparable sensitivity using SWATH on a TripleTOF5600 [244,261,291,293].

The Python script allowed to count for each spike the number of transitions falling in the high repeatability peak area range (here 2.0E+04 – 1.0E+06). Singularly, all the proteins with at least 4 flagged transitions were above the LOQ while the number of flagged transitions tends to zero when the spike is below the estimated LOQ (Figure 5.6). The Pearson's correlation coefficient *r* between the mean extracted area (log scale) and the number of flagged transitions is always above 0.9. This is not the case when CV criterion is used: there, *r* fluctuates between 0.49 and 0.95 (Figure 5.6). Even this is not a strict sensitivity assessment *per se*, this is again an indication that the CV approach is less robust than the peak area approach.

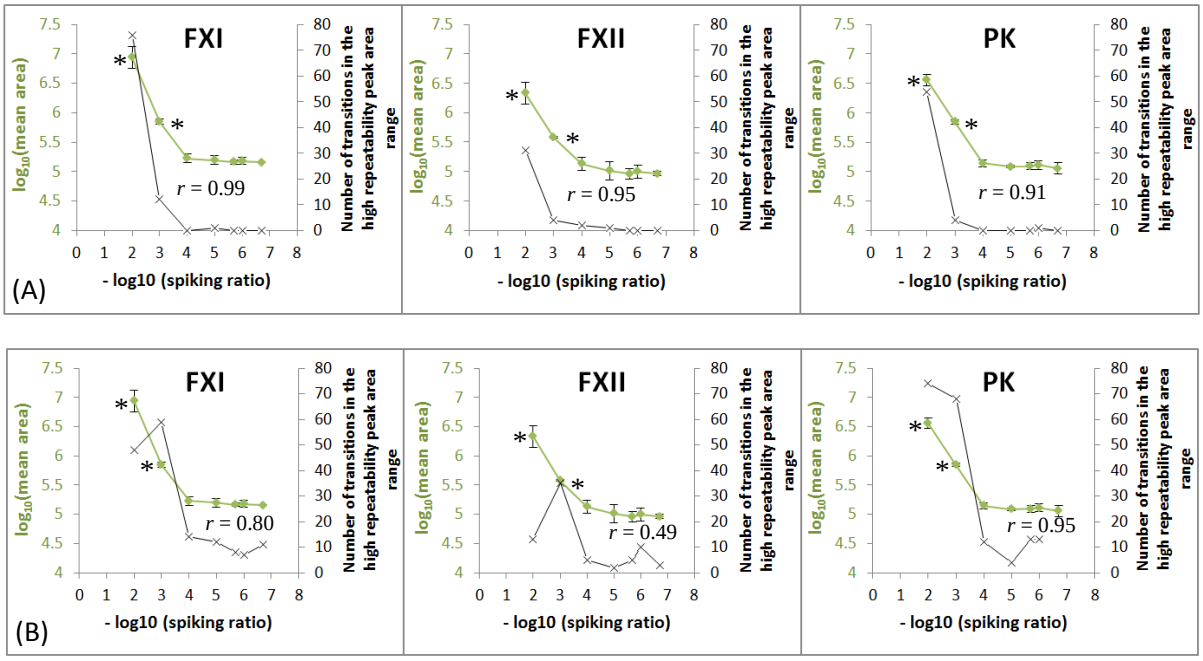


Figure 5.6 Sensitivity markers. Three plasma proteins, namely coagulation factor XI (FXI), coagulation factor XII (FXII) and prekallikrein (PK), were spiked at different mass ratios in non-fractionated Ig sample (A1). Each sample was injected in triplicates and mean area was reported in log scale. The spike is considered as quantified above the LOQ if both the preceding and following dilution steps gave significantly different areas with Mann-Whitney test (* denotes p-value < 0.05). In-house Python tool allowed to count for each spike the number of transitions: (A) falling in the high repeatability peak area range (2.0E+04 – 1.0E+06) (B) having a CV below 20%. The Pearson's correlation coefficient r between the mean peak area and the number of flagged transitions was also calculated.

Those results also emphasize that in our application, sample prefractionation is still required to reach relevant sensitivity. By means of CPLL enrichment, FXI spiked at 1ng/mg, the level from which thrombogenic effects can be observed *in vivo*, was confidently quantified with three flagged peptides (at least two ions with a peak area in the range 2.0E+04 – 1.0E+06) among the 18 peptides indexed in the library.

5.4.4.3 Residuals characterization in different Ig products

As final proof of applicability, we used our MS2 repeatability flagging approach for the impurity profiling in our batch of five different Ig products, including a TEE-positive one. Each product was fractionated twice to take fractionation variability into account (CPLL technical variability for this application was previously discussed in chapter 4). Each fractionated sample was injected in triplicate for back-to-back comparison with our MS2 peak area criterion. Spectral library resulted from merging together all the three plasma

libraries generated in-house. We also complemented the library with the DDA runs of three proteins of interest, FXI, FXII and PK.

Table 5.3 Comparison of plasma residual protein profile in five different Ig products. Those include: two IVIg products from a first manufacturer, including two different batches from one of the two products (A1, A2 and B), one IVIg product from another manufacturer (C) and one TEE-positive batch from ScIg product (D). Each sample was enriched using CPLL strategy (two technical replicates, three injections of each). Table reports every non-IgG protein for which peptide(s) that infer a minimum of two flagged ions (peak area between 2.0E+04 and 1.0E+06) were reported and for which the abundance in the considered sample equals or exceeds ten times the median value of its reported quantifications in all the tested samples. The proteins are listed in alphabetical order.

Regular commercial IVIg				TEE-positive SCIg (2011)
Manufacturer 1			Manufacturer 2	Manufacturer 3
A1	A2	B	C	D
		Ig alpha-1 chain C region	Afamin	Alpha-1-antitrypsin
		Ig alpha-2 chain C region	Alpha-1-acid glycoprotein 2	Alpha-2-macroglobulin
		Ig mu chain C region	Alpha-1-antichymotrypsin	Angiotensinogen
		Immunoglobulin J chain	Alpha-1B-glycoprotein	Antithrombin-III
			Alpha-2-HS-glycoprotein	Coagulation factor XI
			Apolipoprotein A-II	Complement component C8 beta chain
			Cartilage acidic protein 1	Complement component C8 gamma chain
			Haptoglobin	Complement factor I
			Hemoglobin subunit beta	Fibrinogen alpha chain
			Hemopexin	Fibrinogen beta chain
			Plasma protease C1 inhibitor	Fibrinogen gamma chain
			Serum albumin	Ficolin-3
			Transthyretin	Gelsolin
				Hemopexin
				Ig alpha-1 chain C region
				Ig mu chain C region
				Ig delta heavy constant region
				Immunoglobulin J chain
				Pregnancy zone protein

Among the 217 non-IgG proteins for which spectral information was indexed in the library, data was extracted for 165 of them (Figure 9.5 in the Addendum). By using the Python tool, 45 different non-IgG plasma proteins were flagged for having peptide(s) with at least two ions in the high repeatability peak area range, in at least one sample (Table 9.8 in the Addendum). As expected, these flagged proteins exhibited a very high repeatability with triplicates measurements. 84% of all flagged non-IgG plasma

proteins quantifications had a CV below 20%, while this proportion even rose up to 94% for those having at least two flagged peptides (Table 9.8). Across the different Ig samples tested, 82 to 100% of the proteins having at least two peptides with at least two ions in the area range where 80% of data with CV<20% (here the 2.0E+04 – 1.0E+06 area range) demonstrated a CV below 20%, while at the same time, only 32 to 49% of the other proteins passed the same CV threshold. Table 5.3 further reports every flagged protein for which the abundance in the considered sample equals or exceeds ten times the median value of its reported quantifications in all the tested samples. The differences between samples tends to increase when comparing a given product (A1) respectively to another batch of the same product (A2), another product of the same manufacturer (B), another product from another manufacturer (C) and finally a TEE-positive batch from 2010 (D). The last one was clearly distinguished from all the four regular products by increased levels of coagulation factor XIa but also, alpha-1-antitrypsin, angiotensinogen, antithrombin-III, fibrinogen, ficolin-3, gelsolin or complement component C8.

Importantly, those results are exceedingly consistent with previously obtained results using DDA approach on the same samples (see Table 4.1 in chapter 4 and Table 6.2 in Chapter 6). Nevertheless, where DDA provided stochastic MS1 level quantifications that require repeated injections, SWATH now provides comprehensive MS2 level quantifications that can be trusted with single injections and can be interrogated in the future with an extended DDA library.

5.5 Conclusion

In conclusion, we here refer to a well-known concept in analytical chemistry to bring it to a new level of applicability. More specifically, it is well established that the repeatability is linked to the measurement intensity. It was evidenced by Navarro and colleagues that this equally holds for SWATH-MS. Transition intensity by means of the MS2 peak area is already included in data processing pipelines by selecting the top n (usually top 3 to top 6) most abundant transitions. As observed in the Navarro manuscript however, this is not sufficient to guarantee a reliable quantification of low-abundant signals. To our knowledge, this is the first time that the functional relationship between CV and MS2 peak area is described (by using the relatively under-exploited Replicates Analysis 2.0 template from Sciex). With only a simple power regression to model the relation, it was demonstrated that the MS2 peak area can confidently contribute to predict the reliability of MS2 quantification. Importantly, this metric does not rely on the number of replicates and can even be considered for single injections, contributing to increased high-throughput. The robustness of the observed relationship across all the investigated conditions suggests that it is intrinsic to the instrument. Assuming that its stability would be regularly verified, this functional relationship could thus further refine the scoring algorithm of SWATH-MS and other DIA software tools. As proof of applicability for LAP quantification, in-house Python tool was built to flag the proteins that integrate a minimal number of transitions in the peak area range of desired repeatability. By doing so, a set of highly robust quantifications of plasma residual proteins in fractionated Ig samples was secured from large spectrum SWATH screening using extended spectral library.

Chapter 6:

Final discussion and general conclusion

6. Final discussion and general conclusion

This thesis aimed at considering MS to profile undesired residual proteins in Ig, the leading product of the plasma fractionation market. By offering non-targeted and comprehensive screening, MS method development can be expected to prevent issues related to undesired low-abundant residuals in Ig, like the 2010-2011 TEE outbreak related to co-purified plasma coagulation factor XI.

A reference *in solution* digest protocol using trypsin and reverse phase C18 microflow LC - YMC-Triart C18 column connected to an Eksigent nanoLC 425 - coupled to high resolution QTOF MS - TripleTOF 5600, Sciex - operating in non-targeted acquisition (Q1 open for all the peptide-ions in the 400-1250 m/z range, resulting in a full MS1 scan) were used as the fixed framework for our investigations.

In the first part, golden standard bottom-up MS was investigated through data-dependent acquisition and MS1 relative quantification. To reach relevant sensitivity, sample fractionation strategy was also optimized. In the second part, data-independent acquisition SWATH-MS with MS2 quantification was challenged to provide an even more exhaustive and repeatable profiling.

Thanks to a collaboration with two different Ig manufacturers, we obtained a set of five different samples: two products, referred as A and B, including two different batches of A (referred as A1 and A2) from the first manufacturer and one product, referred as C, from the second manufacturer. To compare these regular batches released on the market with a batch presenting a quality issue, the first manufacturer also provided us with an aliquot of TEE-positive ScIg from 2011, further referred as D. With this set of samples at hand, it was possible to challenge an MS method for relevant QC application, namely the distinction of a TEE-positive batch and the biodiversity assessment of different Ig products in terms of plasma residuals.

6.1 Two challenges named sensitivity and repeatability

Only the levels of PKA, FXI and IgA in regular products are ensured to be below 0.1-1 ng/mL, 35 IU/mL (at 3% protein content) and the stated label respectively, thanks to dedicated testing (chromogenic assay for PKA; NAPTT, ELISA, TGA or chromogenic assay for FXI; immunochemical method for IgA) (Table 9.1). For this reason, highly pure native FXI, PK and FXII (Calbiochem) were spiked at different mass ratios ranging from 10 µg/mg (1:100) to 0.2 ng/mg (1:5000000) to estimate if relevant sensitivity, namely the range of the ng/mg, could be achieved. Of note, a global sensitivity cannot be specified as such in a label-free MS assay, since each peptide will have its own ionization properties and the link between signal intensity and the concentration in the sample will vary accordingly. However, each MS device will exhibit an intrinsic capability to distinguish and quantify different signals across a certain linear dynamic range, related to the detection system. For a TT5600 device preceded by a single LC column, peptides are expected to be quantified through a dynamic range of around 4 orders of magnitude when quantified at the MS2 level [244,261,291,293]. Early results thus confirmed what could be expected: direct injection of digested Ig

sample on the TT5600 preceded by C18 LC separation did not return any confident identification of plasma residuals and FXI spiked at 1 ng/mg. The total ion current is monopolized by the Ig fragments. Mining the samples with additional fractionation appeared as mandatory to obtain relevant profiling. The application of an adequate sample treatment strategy was thus the first milestone of this thesis. Next to this challenge of dynamic range, a second challenge quickly appeared. When trying to quantify LAP in the lower signal range that the instrument can handle, particular attention should be paid to the reliability of the quantification. Especially when the method development aims at orientating manufacturers (and possibly regulators) in qualifying the status of a batch: conform or not. Concretely, this called for cautious and elaborated data acquisition and interpretation.

6.1.1 The need for sample pre-fractionation

Thanks to years of studies mainly for biomarkers discovery, sample fractionation in plasma proteomics has been already thoroughly investigated [269,280,297,298]. Because non-biased comprehensive profiling was desired, fractionation techniques that isolate proteins with specific properties (e.g. glycoproteins) were not suitable. Biophysical fractionation methods based on generic physico-chemical properties (pI, size, ...) were also not favoured because of the large number of fractions generated. Moreover, LAP with properties similar to the main component might be lost; anion exchange chromatography is e.g. already used by various manufacturers and does not prevent impurities ending up in the final product. Two main strategies thus remained to be considered for uncovering the largest possible amount of LAP: the 'as specific as possible' depletion of the high abundant components and the 'as generic as possible' enrichment of low abundant components. In each case, downstream desalting and protein assay were conducted to prevent ionization-inhibitory salts and calculate the exact MS sample load, respectively. LAP enrichment trough CPLL technology uncovered more plasma proteins in the samples than the specific Ig depletion through the use of protein G-protein A affinity column (Figure 4.1). Moreover, FXIa 1ng/mg spike was identified with up to 20 peptides with single CPLL, while only 2 were reported after depletion. While sometimes controversial due to the diversity of protocols, superiority of CPLL over depletion for plasma proteomics had already been observed previously [274,298,299]. Despite being efficient, depletion strategies require specific ligand/stationary phase and particular caution should always be paid to the risk of undesired co-depletion [201–203,262]. Here, protein G-protein A affinity chromatography apparently suffered from the high dilution of the residuals in the flow through, since co-depletion was known to be minimal with these ligands [198–200,270]. The associated extra concentration step increases the risk of trace losses [271].

On the other hand, CPLL technology is readily available through commercial kits and does not require application-specific ligands. Except when the sample amount is limiting, a "large capacity" kit should be preferred to a "small capacity" kit: a larger volume of beads enlarges the ligand library diversity and by this, the chance to fish more proteins. Then, the benefits of the extra "sequential elution" should be weighed against the extra manipulation cost. Surprisingly, the kits' vendor did not include the most recent optimization work, considering fine-tuning of the binding buffer (ionic strength), further recapture of the flow through at different pHs and on-bead digest [214]. Still, these refinements did not demonstrate to be beneficial for our application (Figure 9.4 and Table 9.2). Contrary to the more complex samples used in the optimization studies (whole plasma, cell lysate, plant extracts ...), samples with only very few LAP

among a vast majority of given active compound(s) behave differently: drastic harvest from the beads also means collecting more peptides from the abundant compound(s) and thus reducing their concentration to a lesser extent. In other studies dealing with the same type of sample, CPLL was indeed conducted with a protocol really close to the commercial kits [131,213,266]. On the other hand, when we built a spectral library for SWATH starting from plasma pool samples, we used the optimized protocol with incubation at 3 pH (pH7, pH9.3 & pH4) in parallel and on-bead digestion, as its benefits had been demonstrated with this matrix [214,215,300].

Despite overall better performance of CPLL enrichment over Ig depletion, one identification (plasma serine protease inhibitor) was only made with protein G-protein A, recalling that different fractionation techniques can yield somewhat different results and can be thus complementary [280,297,298]. Coupling CPLL enrichment to prior Ig depletion with protein A-protein G chromatography led to an increase in the number of identifications (Figure 9.4 and Table 9.2). Still, this extended sample treatment required about six more hours per sample, plus an additional concentration/desalting step. On the other hand, only a minor benefit of two extra identifications and no increase in the number of reported peptides for the FXI 1 ng/mg spike were reported as compared to the optimized single CPLL. This is inducing a rule of thumb: the increased coverage in LAP gained with each extra fractionation step must be worth the time and cost invested. Moreover, each extra step increases the risks of unspecific sample loss or artefacts and higher technical variability. Therefore, each fractionation protocol should be assessed in terms of technical repeatability.

Previously, CPLL enrichment demonstrated suitable repeatability for both qualitative and quantitative profiling [273,279,280,283]. Despite the modification of the concentration dynamic range in the sample, the CPLL technology can be used for differential studies as long as the proteins under investigation do not reach saturation on the beads, i.e. the proteins of low to medium abundance in the original sample [301]. Here, five technical replicates of the sample treatment were performed. 94% (17 to 18) of plasma LAP identifications with at least two peptides were made in at least four out of five replicates. The median CV of MS1 quantification was 20.2% and 81% of the data had a CV below 30%. Nevertheless, a maximum of 4-fold change in normalized protein abundance was reported (Chapter 4, section 4.4.1.2). Obviously, variation in protein abundance can be only considered as biologically relevant if this variation is unequivocally above the technical variability of the protocol. Here, it was decided to consider at least 10-fold changes in protein normalized abundance as significant biological variation (more than 2 times the maximum fold change observed between two technical replicates for any protein). Of note, the choice of the minimal margin between the technical variability and the biological significance is user-defined, depending on the susceptibility of the investigated samples.

In conclusion, the sample fractionation approach should provide the best compromise between relevant LAP mining and gain in sensitivity on one hand and adapted throughput and secured technical repeatability on the other hand. In this thesis, single CPLL enrichment with capture under physiological conditions and sequential elution (performed with two technical replicates) was further applied as sample fractionation method because it showed the best LAP coverage and FXI spike identification within convenient execution.

6.1.2 The importance of the data: from acquisition to interpretation

Non-targeted acquisition on TripleTOF devices can be performed in two different ways, as summarized in Table 6.1: the golden standard DDA or the more recent DIA-SWATH strategy.

Table 6.1 Overview of the two acquisition workflows considered with the TripleTOF5600

	DDA acquisition	SWATH acquisition
MS run		
MS1 scan (400-1250 m/z)	Full scan	
MS2 scans (65-2000 m/z)	MS1-triggered, stochastic	Full scan (adjacent m/z windows), comprehensive
Output	Linked tandem data	Complex convoluted data
Identification		
Software	Protein Pilot (Sciex), Mascot (Matrix Science)	SWATH 2.0 in Peakview (Sciex)
Database	General protein sequence database (Swissprot db)	Dedicated spectral library
Process	Peak picking, annotation with db searching engine, scoring and FDR calculation	Library filtering, RT alignment, peak extraction, spectral matching with the library
Relative quantification		
Level of quantification	MS1-level	MS2-level
Software	Progenesis QI (Waters)	SWATH 2.0 in Peakview (Sciex)
Process	Runs alignment and abundance normalization, differential statistical analysis, identification, scoring, reporting	Peak area of the extracted MS2 ions are reported and summed to give corresponding peptide and protein abundances

6.1.2.1 Operating in DDA

Thanks to years of practice, DDA or "shotgun" MS runs are easy to implement. MS2 ions result from selected MS1 ions that surpassed a defined signal threshold and are thus linked to their precursor ("tandem data"). A well-established search engine, like the probability-based Mascot, is used for identification by matching public reference database through peptide mass fingerprint, sequence query and MS2 ions search [302]. Of note, as for every probability-based approach, a carefully selected threshold of significance has to be defined. Mascot includes several quality control measures. First, the quality of the peptide match is reflected into a score (the higher the better) linked to a p-value (the lower the better) for which a cut-off can be defined (usually 0.05 or 0.01). Then, it can be ensured that (i) sequence matching is the best hit in the database (rank=1) (ii) a peptide univocally corresponds to a single protein in the database and (iii) a maximal number (usually 1 or 2) of enzymatic miscleavages is tolerated. False discovery rate calculation using automatic decoy database search is also a key indicator and should always be less than 1%.

Once peptide-level criteria are defined, the peptide to protein is *per se* a matter of inference, as tryptic peptides are the real analytes for bottom-up MS. Usually, it is recommended to have at least two unique peptide-sequences identified to confirm the presence of a protein [190]. The definition of an identification cut-off for the protein score given by the software algorithm is more empirical and its value is software-

dependent. However, biological context can help. Here, as plasma-derived samples were profiled, protein expression in blood was verified through the NextProt repository (except for reported adjuvants in the formulation) and a score threshold below which a large majority of the reported proteins were not expressed in blood was set.

For our purpose of LAP screening, we favoured stringent peptide criteria to ensure quality data and relatively loose protein criteria to avoid missing some contaminants. At peptide level, p-value was below 0.01, FDR below 0.2% and best ranking was required. At protein level, the number of unique peptides were reported and the minimal Mascot score was set to 50 when at least two unique peptides were reported and 80 when only one unique peptide was reported.

One major drawback of DDA is that the acquisition is inherently stochastic, inducing repeatability issues [237]. It is well-known that repeated injections have to be performed to reach maximal coverage [238–240]. Indeed, it was observed that about 8 replicate runs per sample were needed to reach maximal protein sequence coverage (Figure 4.2). Together with the two technical replicates of the CPLL fractionation, a set of 16 runs were thus acquired for each Ig sample to profile. When considering MS1 relative quantification, run alignment and signal normalization are also an essential concern. Fortunately, software pipelines are devoted to such differential analysis. We used Progenesis QI for Proteomics (for details, see section 4.3.4.3 b). Once the quality of the automatic alignment has been controlled (here >80%), quantification CVs are reported to reflect the level of repeatability, between replicate injections and between replicate fractionation. As presented in Chapter 4 and summed up in the next section 6.4, insightful LAP profiling based on MS1-level quantification was obtained. Relevant sensitivity was also achieved, as the FXIa spike at 1ng/mg was confidently identified. Nevertheless, it was emphasized that this comes with an extended acquisition time (more than 16h per sample, as LC gradient for each run is lasting 1h) and careful data interpretation.

6.1.2.2 Operating in DIA-SWATH

When introducing SWATH, Gillet and colleagues evidenced that fragment ion signal extraction from MS2 scans using a spectral library is more reliable than precursor ion detection/picking in the MS1 scan in DDA [244]. Indeed, MS2 signals are even more component-specific and the dynamic range could be modified, as the signal of each MS1 ion is split into multiple MS2 ions. Finally, SWATH is DIA and thus comprehensive by nature: all the detected MS1 ions are fragmented, producing an exhaustive MS/MS map at a given limit of detection during a single run. Altogether, using SWATH for our profiling was thus definitely worth an investigation, especially since it is readily available on the TripleTOF devices.

Compared to a DDA run, a SWATH run requires more complex elaboration. First, a spectral library has to be available. For the purpose of extracting data about the largest possible amount of potential plasma residuals, in-house spectral libraries were here built with DDA runs of plasma pools collected from manufacturers and fractionated using optimized CPLL (see sections 5.3.1 & 5.3.3.1). When importing the spectral library, shared and modified peptides were excluded and only the proteins corresponding to <1% FDR identification during the database search of the DDA runs library were retained. Then, the sequence of consecutive *m/z* windows has to be specified. The number of windows has to be specified within the fixed cycle time (to preserve enough data points per LC peak): a higher number of windows favours

specificity, while a smaller number favours sensitivity. The Variable Window Calculator Excel tool from Sciex allows defining windows of variable m/z width by considering the signal distribution of the sample under investigation: more (shorter) windows will be allocated to the m/z regions of more intense signal, which improves the balance between specificity and sensitivity [303]. Finally, exogenous reference peptides have to be spiked in fixed amount in each injection vial for further RT calibration.

The output of a SWATH run is intricately convoluted data that has to be deciphered by matching to a spectral library. To do so, different software pipelines exist and demonstrate comparable performance, as assessed by a recent multicentre study [254]. Still, this study also highlighted that low abundant signals quantification remains challenging irrespective of the software tested. While applying one of these software pipelines - SWATH 2.0 in Peakview - with processing settings as recommended in the multicentre study (see 5.3.3.4 and Table 5.1), this pitfall was quickly verified in our own data: only about 30% of the extracted transitions had a CV below 20%. Even more problematic, proteins with good quantitative CV could in fact rely on a vast majority of poorly quantified ions (as illustrated in Table 9.5 of the Addendum). Protein and peptide-level quantifications are simply inferred by summing the peak area of corresponding extracted MS2 ions that passed the library matching processing settings. An additional MS2-level data quality criterion was thus desired.

CV calculation relies on a certain number of replicates and the benefit of single run injections is thus lost. The “SWATH Replicates Analysis 2.0” Excel sheet allowed to plot average peak area against corresponding CV for up to forty thousand extracted MS2 ions drawn uniformly across the data. By doing so, it was outlined that the well-known signal-intensity-to-repeatability relationship also holds for label-free quantitative MS. Moreover, it can be modelled by a power regression that is greatly conserved across all samples and acquisition parameters, both in in-house data and in reference data from the multicentre study (Figure 5.2). Once the inherent MS2 peak area to CV relationship of an MS device has been modelled, and assuming that its stability is verified, the user can make repeatability predictions from a single run, only based on the MS2 extracted signals (see section 5.4.2.2) Moreover, the functional relationship can serve as an actionable point for optimizing acquisition parameters (section 5.4.3) and ensuring more reliable LAP profiling (section 5.4.4). To illustrate this, we built a Python script that enables to flag automatically the MS2 ions falling in a user-defined peak area range corresponding to the desired level of repeatability. This allows defining the peptides and the proteins that integrate a minimal number of such ions, i.e. quantified with higher confidence. As presented in Chapter 5 and summed up in the next section, this approach led to secure highly confident LAP profiling in the Ig products. Still, distinction of different subgroups among the obtained quantifications based on the MS2 peak area through this script should never be seen as a hard cut-off criterion. Consequently, we do not present it as a ready-made tool for users to implement in their workflow. This is because there are currently five main software tools that are being used for SWATH data analysis, each one with its own specifications. The developers of these tools already are in close contact and inspire one another to implement new functionalities. The findings of the current work are actually an open invitation to the existing software developers to incorporate this intrinsic instrumental feature into their pipelines in any way they see fit.

6.2 The reward: a better characterization of important biological therapeutics

Throughout this thesis, sample fractionation and two MS data generation and interpretation pipelines were thus carefully optimized to provide relevant residual plasma LAP profiling in Ig products, being both sensitive and repeatable. Without fractionation, spiking experiments did tend to confirm that TripleTOF5600 allows quantifying components across up to four orders of dynamic range, when using MS2-level quantification (Figure 5.6). Optimized CPLL fractionation allowed to confidently reveal FXI spiked at the ng/mg with both DDA (20 identified peptides) and SWATH (three peptides flagged as repeatable with the Python tool), thus confirming that LAP could be surfaced at relevant sensitivity with the two pipelines.

With 16 DDA runs per sample at hands (8 injections for each of the two replicated CPLL fractionations), about 70 different plasma residual proteins were confidently identified among our set of five Ig products. The Progenesis pipeline allowed attributing MS1-level normalized abundances for these plasma LAP. With only 1 SWATH run per CPLL fractionation and a spectral library built from fractionated plasma pools samples (spectral index of 217 proteins), MS2-level quantification was obtained for 165 proteins. In comparison, the only previous study profiling impurities in Ig products with MS identified three proteins [93]. Both approaches clearly revealed that each Ig product exhibits its own impurity profile, as summarized in the form of two heatmaps with data normalization and hierarchical clustering (Figures 4.3 & 9.3).

Still, strongly reliable and significant variations had to be outlined, in order to evidence potential QC impact. The MS2 peak area to repeatability relationship allowed developing strict repeatability guarantee for single injection SWATH data. Flagging proteins with peptide(s) having at least two MS2 ions in the peak area range of desired repeatability (here 80% data with CV < 20%) secured a set of 45 non-IgG plasma proteins, with 84% of them having a CV below 20% (Table 9.8).

Subsequently, a minimum of 10-fold change in the protein normalized abundance was required to consider this change as a significant biological variation. Indeed, this is more than 2 times the maximum fold change observed between two technical replicates for any protein, i.e. the highest observed noise due to extensive sample preparation and which has no biological significance. As a batch of five Ig products was considered, fold changes were referred to the median of the abundances in every tested sample. This led to extract a short list of residual plasma proteins that significantly vary among the tested products. While initially presented in Tables 4.1 & 5.3, the Table 6.2 hereunder is summarizing these findings.

While being obtained from two totally different data acquisition and processing pipelines, the plasma residuals profiles with both SWATH MS2-based and DDA MS1-based quantifications were found to be very consistent, which confirms the reliability of these results.

Table 6.2 Final comparison of plasma residual protein profile in five different Ig products. Those include: two IVIg products from a first manufacturer (A and B), including two different batches from one of the two products (A1, A2), one IVIg product from another manufacturer (C) and one TEE-positive batch from 2011 ScIg product (D). Each sample was enriched using CPLL strategy (two technical replicates). Listed proteins (alphabetical order) demonstrate an abundance in the considered sample which equals or exceeds ten times the median value of its reported quantifications in all the tested samples. Considered abundances are normalized MS1 abundances using DDA strategy and extracted MS2-based abundances using SWATH-MS strategy. Proteins in bold are reported with the two types of abundance/strategy, while proteins in italic are reported with only one strategy (specified in the superscript).

Regular commercial IVIg				TEE-positive ScIg (2011)
Manufacturer 1			Manufacturer 2	Manufacturer 3
A1	A2	B	C	D
Neurolysin, mitochondrial^{DDA}		*CD5 antigen-like*^{DDA}	**Afamin**	**Alpha-1-antitrypsin**
		Ig alpha-1 chain C region	*Alpha-1-acid glycoprotein 1*^{DDA}	**Alpha-2-macroglobulin**
		Ig alpha-2 chain C region	*Alpha-1-acid glycoprotein 2*^{SWATH}	**Angiotensinogen**
		Ig mu chain C region	**Alpha-1-antichymotrypsin**	**Antithrombin-III**
		Immunoglobulin J chain	*Alpha-1B-glycoprotein*^{SWATH}	*CD5 antigen-like*^{DDA}
			Alpha-2-HS-glycoprotein	**Coagulation factor XI**
			Apolipoprotein A-II	*Complement component C8 alpha chain*^{DDA}
			Apolipoprotein C-I, basic form (Fragment)^{DDA}	**Complement component C8 beta chain**
			Cartilage acidic protein 1^{SWATH}	**Complement component C8 gamma chain**
			Haptoglobin	**Complement factor I**
			Hemoglobin subunit beta	*Ectonucleotide pyrophosphatase/ phosphodiesterase family member 2*^{DDA}
			Hemopexin	**Fibrinogen alpha chain**
			N-acetylmuramoyl-L-alanine amidase^{DDA}	**Fibrinogen beta chain**
			Plasma protease C1 inhibitor	**Fibrinogen gamma chain**
			Serum albumin	**Ficolin-3**
			Transthyretin	**Gelsolin**
				Hemopexin^{SWATH}

	Ig alpha-1 chain C region
	Ig alpha-2 chain C region^{DDA}
	Ig delta heavy constant region^{SWATH}
	Ig mu chain C region
	Immunoglobulin J chain
	Pregnancy zone protein^{SWATH}

The few observed divergences in part resulted from homologous peptides (Ig alpha-1 and alpha-2 chains, alpha-1-acid glycoproteins and the different chains of Complement component C8). Concerning the remaining proteins quantified with DDA and not SWATH, Neurolysin mitochondrial, Apolipoprotein C-I basic form (Fragment), CD5 antigen-like and Ectonucleotide pyrophosphatase/phosphodiesterase family member 2 were not indexed in the in-house plasma spectral library. N-acetylmuramoyl-L-alanine amidase was just outside the flagging criterion (only 1 transition in the peak area range of desired repeatability). This again recalls the crucial importance of both library completeness and data interpretation when using SWATH-MS. Finally, while it was expected to gain additional quantifications with SWATH in comparison to DDA, it was observed that this gain is very modest when strict data quality criteria have been applied.

These final results succeed in achieving one of the main goals of this thesis: confirming the diversity of Ig products in terms of the low-abundant residual fraction. While the two batches of the same product were found to be very similar, two products from a same manufacturer can exhibit different properties, thanks to modulated processing or plasma selection. For example, the lower reported level of IgA in the product A as compared to the product B confirms that product A will most likely be preferred for IgA-deficient patients, since very low level of IgA will minimize the risk of anaphylactic shock. Next, the heterogeneity in terms of both plasma pool composition and fractionation process among different manufacturers is strongly reflected in the distinct impurities profile of product C (second manufacturer) as compared to products A and B (first manufacturer). Finally, the TEE-positive batch (product D) is clearly distinguished from all the other regular batches. The role of FXI in the thrombogenicity is confirmed since its level is more than 10 times higher in this product as compared to the others. Other proteins also appear to be at markedly higher level only in the TEE-positive batch. Some of these proteins, like fibrinogen, angiotensinogen or antithrombin-III, are involved in blood-stream regulation and might be pro-TEE candidates. Nevertheless, this hypothesis should be confirmed by activity assays and/or a larger batch of products. Notably, previously hypothesized implication of plasma kallikrein [91], CD154 protein [94] and apolipoprotein H [93] in TEE events cannot be confirmed here.

In conclusion, this work is demonstrating that adequate sample fractionation followed by label-free quantitative MS strategy with careful data interpretation does provide a relevant opportunity for large-scale and non-targeted profiling of impurities that could advantageously complement dedicated testing and prevent any unexpected issue at clinical level associated with undesired residuals.

Chapter 7:

Broader international context, relevance and future perspectives

7. Broader international context, relevance and future perspectives

“With the plasma market, we’ve traditionally seen some type of supply disruption every three to five years. Due to the small number of manufacturers and the long lead time between plasma collection and getting product on the market, any manufacturing issue or recall hits the marketplace hard.” said in 2016 a general manager of specialty distributor in a bio/pharma business magazine [304]. Nowadays, the adequate supply of blood-sourced therapeutics derived from the process of blood and plasma collected from voluntary donors still remains challenging.

Despite the development of engineering techniques enabling to synthetize proteins from genetic material, the so-called “recombinant” technology, the demand for plasma-derived protein therapeutics remains high. In particular, normal polyvalent Ig products, namely the isolated antibodies repertoire from the donors, are almost impossible to substitute with recombinant equivalents due to the extensive diversity of proteins contained in such products. Moreover, the demand for these Ig products is maintained high through their use in replacement therapies for immune deficiencies (both native or disease-related) but also increasing indications for immunomodulation in various autoimmune and inflammatory diseases. Therefore, Ig products are the leading product of the current plasma market. Nowadays, this market is representing more than 14 billion dollars of annual sales, and several projections foresee a 10% annual growth. Nevertheless, the annual cost of a plasma-derived therapy can still exceed 200 000 dollars [304]. A first reason is that the global blood supply can be stressed by endemic or epidemic viruses (e.g. Ebola, dengue or Zika). Then, both collection and manufacturing process of the plasma are highly elaborated and regulated. On the other hand, plasma-derived products are mainly intended to treat relatively rare disorders. Patient populations range from hundreds to hundreds of thousands and do not allow high economies of scale in the production facilities. Consequent to both the increasing specific demand and the overall high cost of supply, preventing any quality issue and/or market withdrawal is of prime importance in the plasma-derived therapeutics market.

Initiated at the dawn of the 20th century with the launch of the Red Cross to treat injured soldiers during the WWII and the early development of plasma proteins separation techniques, the blood and plasma therapeutics supply evolved into an elaborate large-scale process. Now starting from thousands of donations in licensed collection centres, it ends with complex manufacturing in industrial facilities under GMP and SOP. On top of this, a strict quality system has been implemented, including licensing, inspection and quality control by independent regulatory authorities. Thanks to this evolution, unprecedented safety has been achieved for the first main risk associated with blood-derived products: the transmission of microbiological and viral agents. Not any transmission of HIV, HCV and HBV with Ig administration has been reported since 1994 [7]. Moreover, most eventual adverse reactions are mild and transient, given that the product is correctly infused [55,58–60].

Still, over the last years, safety concerns other than pathogen transmission emerged. These were related to the fine-tune composition of the products and demonstrated to potentially cause severe issues with both public health and economic impact. An increase in frequency of haemolytic and thrombotic events (TEE) was associated with the presence of anti-blood group determinants antibodies (anti-A, anti-B and anti-D antibodies) and unintended residues of plasma coagulation factor XI respectively [84–86] [90–92]. With high dose indications increasing their susceptibility, receivers suffered from both severe haemolysis and TEE including stroke, myocardial infarction, pulmonary embolism and deep vein thrombosis.

In 2010-2011, the outbreak of TEE following Ig administration led to a temporary market withdrawal of three Ig products by the regulatory authorities [96,97,99]. In addition to these events, an increasing demand for products' bio-diversity assessment and personalized product selection emerged. For example, patients with a given blood group or deficiency in one of the classes of antibodies (most often the IgA class) may require Ig products with certified low amounts/absence of the reacting type of antibodies. Altogether, this clearly called for an even more detailed characterization of the products, namely the composition of the antibodies (IgG) repertoire on one hand and the low-abundant undesired residual (non-IgG) plasma proteins on the other hand.

Current regulatory requirements include overall purity assessment (less than 10 % of non G-class Ig in the final protein content, 5 % in the case of intravenous administration) and maximal authorized levels for some particular compounds known to cause potential adverse effects (Table 9.1). While IVIg preparations on the market effectively contain at least 99% IgG, reaching 100% purity is impossible when starting from such complex biological material as a pool of human plasma [2,16,25]. A Wessler *in vivo* model revealed that only a few ng of coagulation factor XI per mg of Ig product can induce thrombin formation [100,103]. Importantly, FXI does not belong to the top abundant plasma proteins, but manufacturing process modifications led to unintentional co-purification. This means potentially every of the hundreds of proteins from the plasma source material can end up as traces in the final product and strongly affect its quality.

Dedicated testing (TGA, NAPTT or ELISA) to control FXI level and procoagulant activity enabled *a posteriori* resolution of the TEE outbreak. At the same time, only non-targeted comprehensive profiling might have prevented it. Since 2010-2011, manufacturers most probably performed this kind of assessment on their own but did not necessarily publish the results. Altogether, targeted methods can only bring dedicated solutions to particular problems and comprehensive fine-tuned composition is not provided to both the practitioner and the patient.

From 2005 onwards, mass spectrometry-based proteomics demonstrated its relevance for plasma-derived therapeutics characterization, being impurities profiling, products comparison or active component characterization (Table 2.1). Surprisingly, we had to wait until 2014 to see a first publication on impurities profiling in Ig using MS [93]. Actually, MS was quite extensively used for plasma-derived products characterization between 2006 and 2013, consequent to both improved technological maturity and also above-mentioned increased concern for better characterization of these therapeutics. From 2014, the point of attention shifted towards the recombinant engineered proteins, including the monoclonal antibodies, thanks to the new promises of this alternative production approach (faster,

cheaper and more controlled). There too, MS provided interesting insights and MS instruments vendors even developed dedicated software packages and analysis workflows. Characterization of the Ig component, including glycan analysis, is even more critical for recombinant proteins, especially if they have been expressed in non-human cell lines [162,165,181,305,306]. Impurities, here the residual host cell proteins (HCP), should be best screened too [168–173,307]. Anyhow, this extension of MS applicability to the recombinant production should not stop the investigations in plasma-derived therapeutics. As explained, this is particularly true for normal polyvalent Ig, which are not likely to be substituted by recombinant equivalents.

Throughout this thesis, we dedicated our work to provide MS-based comprehensive and quantitative profiling of impurities in Ig products. While targeted S/MRM approaches require a priori knowledge on the compounds to watch for, an entire pan of MS-based proteomics is dedicated to non-targeted acquisition providing large-scale screening capabilities. Moreover, label-free relative quantification can be obtained without requiring time-consuming and/or expensive specific development with labelled standards. Given adequate sample pre-fractionation and cautious data acquisition and interpretation were conducted, insightful results were obtained. Relevant sensitivity was achieved: FXI spiked at the ng/mg, the prothrombotic (thrombin formation) onset level in reference *in vivo* model, was confidently detected. Repeatable results (CV<20%) were ensured. Distinct impurities profile, including label-free relative quantification, was revealed in each of the five tested Ig products, with increasing distinction between batches from a given product, products from the same manufacturer and products from different manufacturers. A TEE-positive batch exhibited a unique profile with higher level impurities, as compared to all the regular products. Finally, very consistent results were obtained between the two data acquisition strategies. Proposed methods thus do fill the gap of having generic and preventive screening possibilities for impurities, each time a modification in the source material composition or fractionation process occurs. Having confirmed its relevance and valorisation potential, MS-based non-targeted screening is still suffering from several pitfalls that, in part, call for further developments.

First, label-free non-targeted MS is only providing relative quantification and accurate sensitivity cannot be extrapolated as such from one protein to another. Peak intensities measured by the detector rely on ionization susceptibilities and isolation/detection efficiencies, being dependent on both the peptide and the MS device. For this reason, blanco/reference material for comparison is needed to perform the analysis and a standard has to be available to further validate the sensitivity for a given compound. Consequently, label-free MS screening of impurities is complementary and not substitutive to targeted testing with accurate absolute quantification for components of well-known adverse effect. One might consider building targeted multiple reaction monitoring (MRM) assays using transitions detected during the screening and spiking labelled standards to obtain absolute quantification. Still, these assays would have to compete with the extreme sensitivity of the targeted tests in place. For example, dedicated testing of FXI level implemented after the TEE outbreak (NAPTT, ELISA, TGA or chromogenic assay) is reaching the 10 to 1 ppb (ng/g) sensitivity [94,100,103,106]. Of note, strict sensitivity comparison is often challenging, since MS quantification results always refer to mass units, while activity units are often preferred for biologically relevant threshold level values. International reference standards are thus needed to establish an official equivalence between mass and activity units. Such biological reference material should be

released from widely recognized institutes like the National Institute for Biological Standards and Control of the UK (NIBSC - http://www.nibsc.org/standardisation/international_standards.aspx).

Second, MS technologies are still in their growing phase, resulting in a diverse technical offer to consider for each potential application. Each user should define which MS device should be selected, and which sample preparation and (chromatographic) separation should be performed upfront.

What the sample preparation and fractionation concerns, it should provide sufficient isolation of the components of interest so that remaining interfering compounds can be resolved by the MS instrument used downstream. Obviously, active compound characterization, e.g. post-translational modifications or oxidation, will call for a very different strategy than low-abundant impurities screening. The first requires the specific enrichment of selected proteins/peptides, e.g. glycosylated proteins (glycoenrichment using lectins or hydrophilic interaction chromatography), proteins subjected to proteolysis (N-terminal enrichment using N-terminal-specific enzyme or terminal amine isotopic labelling of substrates) or carbonylated proteins (biotinylation and affinity chromatography) [191,192]. The second, on the other hand, is asking for non-biased enrichment of all the impurities. This can be performed with either gel electrophoresis, selective depletion of the abundant component or low-abundant component non-specific enrichment. Gel-based techniques suffer from several drawbacks. First, some proteins are missed because of their physico-chemical properties (size or charge) that prevent them from entering the gel, due to low solubility or tendency to aggregate [143,219]. Secondly, gel-based techniques do not easily suit for high throughput and automation [192,220]. At the same time, the development of high-efficiency liquid chromatography separation and high resolution tandem MS allowed to consider direct injection after whole sample digest *in solution*, which is circumventing these drawbacks. Having compared selective depletion of Ig through protein A – protein G affinity chromatography and generic enrichment of residuals through the so-called combinatorial peptide ligand library (CPLL), this work confirmed the benefits of the latter approach. Coupled with downstream standard tryptic digest and C18 chromatography, it revealed relevant impurity profiles. Notably, CPLL demonstrated to be a highly versatile tool to access low-abundant proteins (LAP), so the exposed approach might be worth the transfer to other applications [209–213,308]. Nevertheless, more elaborate sample fractionation could be performed, like the combination of both depletion and enrichment. Two-dimension chromatography or multi-proteases digestion might also be considered to increase sample separation and sequence coverage respectively [192,309]. As we emphasized, the user should define his needs in sensitivity / LAP mining depending on the desired application, while keeping satisfactory throughput and repeatability. For these two last aspects, the possibility to (semi-)automate the sample preparation is obviously of great added value. Except for the coupled chromatographic separation of the peptides, this is still not common practice in mass spectrometry, and only few initiatives have been published [310,311].

Concerning the MS devices themselves, a very wide range of instruments, acquisition strategies and data analysis softwares are coexisting and can be thus considered for a single given application. Currently, different MS devices relying on different mass analysers are generally used for non-targeted screening and targeted MRM. Triple quadrupole and quadrupole-ion trap are mostly used for MRM targeted quantitation and provide over 10^5 wide dynamic range of protein quantification [312]. High resolution mass analysers for non-targeted proteomics on the other hand include TOF analyser, linear ion trap

analyser, Fourier Transform ion cyclotron resonance (FT-ICR) and Orbitrap analyser [120,183]. Of note, targeted experiments on these devices are increasingly proposed, even it remains challenging to maintain good sensitivity when the list of compounds to watch for increases (sensitivity inversely proportional to the degree of multiplexing) [313,314]. Still, other (new) MS devices might always demonstrate increased resolution and sensitivity, due to alternative mass analyser, ion separation or detector technology, e.g. Orbitrap or ion mobility [122,243]. At the same time, high price of MS devices (and as such lower accessibility for laboratories) may hinder or inhibit fast implementation of new MS technology. On top of this, each MS device can often be operated in different acquisition modes and multiple data processing software pipelines.

So-called data-independent acquisition (DIA) strategies are very promising to the field. Contrary to the classical data-dependent acquisition, DIA provides full MS/MS fragmentation map at a given limit of detection. Instead of the repeated injections needed to overcome the inherent stochastic nature of DDA, single injection per sample could here provide high-throughput characterization of a large amount of samples, given that appropriate interpretation of the complex data output is conducted. DIA data mining is a very active field with high impact emphasizing its growth potential [255,315–319]. In this thesis, DIA application for LAP profiling was investigated through the SWATH-MS mode available on the high resolution QTOF TripleTOF series [244]. Recent multicentre studies provided an important step forward in the harmonization and maturity of this technique [254,261]. Still, it was here confirmed that reliable quantification of low abundant signals remains challenging. The use of the MS2 peak area was evidenced as an actionable point for data quality assessment applicable for single injections. The findings of the current work are an open invitation to the existing software developers to incorporate this intrinsic instrumental feature into their pipelines in any way they see fit, just like other studies promoted more stringent control of the false discovery rate (FDR) [249–252]. In the future, the limiting requirement of having all the quantifiable components indexed in a spectral library might also disappear, as library-free SWATH data analysis pipeline is developing [255,318,320]. Notably, DIA can be also performed with other strategies, like SWATH on Orbitrap, or (HD)MS^E on QTOF from other vendor (often including ion mobility) [243,321–323]. This thus calls for critical and independent comparison of these different approaches.

Altogether, harmonization in data presentation and annotation is badly needed to enable correct data dissemination and interpretation across different labs. For this, initiatives like the ProteomeXchange Consortium for the submission and dissemination of globally coordinated standard data (www.proteomexchange.org/) are helpful. Moreover, collaborative and multicentre studies to compare the performance of different set-ups on a same set of samples are also of prime importance to see MS screening becoming an official part of the biological therapeutics quality control. In doing so, consensus protocols should emerge from the technical diversity and progressively integrate the official QC guidelines, like recently started in the Ph.Eur [190]. Similarly, the large consortium European Horizon2020 project VAC2VAC, which aims at providing detailed consistency approach for vaccine batch release with non-animal tests, has included MS-based proteomics in its physicochemical methods work package (http://www.vac2vac.eu/).

Importantly, the rise of such MS-based screening methods contributes to bring the quality control of therapeutics into a new era, the “omics era” [324]. Quality testing is not only hypothesis-driven, targeted and dedicated to well-known issues anymore. It is also comprehensive, non-targeted and anticipative. While proteomics provides the detailed protein composition, genomics can reveal the gene sequence of any microorganism present and metabolomics detects small molecules and metabolites. In this era, it is not the procurement of data which is limiting, but the correct interpretation of the abundantly generated information. Manufacturers and regulators can now access an unprecedented view on the inherent complexity of the biological products, be it the active product characterization or the presence of impurities. Collateral undesired issues due to process modification might be rapidly evidenced, so production losses and market withdrawal risks can be minimized. However, one should avoid to over-interpret non-significant variation and to apply too stringent criteria that would lead to unnecessary batch rejection. As we emphasized in this thesis, the repeatability of the whole sample treatment procedure should be evaluated and the cut-off of biological relevance should clearly exceed the technical variability. Then, selected consensus data quality indicators, being a signal intensity, an algorithm matching score or a number of congruent signals, should be reported to further define its reliability. Finally, an experience-based auto-refining approach should incorporate all available clinical data. Given that the manufacturing process and plasma sourcing remains largely unchanged at a manufacturer’s level, routine QC testing by MS may not be necessary for Ig products. However, MS should be considered at least for the initial characterization of the Ig product, and is also strongly recommended as characterization assay in case of important changes, e.g. changes in the manufacturing process or in the plasma donor population (to demonstrate that the safety profile of the Ig product is still within acceptable limits). Over time, recording the comprehensive composition along with the corresponding manufacturing process and the observed clinical tolerance and safety for each released product/batch might constitute an unparalleled data resource for imaging bio-therapeutics diversity, contributing to fine-tuned risk assessment and personalized treatments.

Chapter 8:

Summary / samenvatting

8. Summary / samenvatting

With increasing indications in both immune deficiencies and autoimmune/inflammatory diseases, human polyvalent immunoglobulin (Ig) is the current leading product of the plasma fractionation market. Resulting from the refined processing of thousands of plasma donations, Ig offer a unique substitutive antibodies repertoire that is not expected to be recreated with recombinant engineering. Thanks to Good Manufacturing Practices and Standard Operating Procedures in the manufacturing facilities as well as licensing, inspection and quality control testing by independent regulatory authorities, unprecedented overall safety has been achieved with Ig treatments. Microbiological safety is highly secured and most reported reactions are mild and transient. Still, over the last years, new safety concerns related to the fine-tune composition of the products emerged. In particular, an outbreak of severe thromboembolism events (TEE) with Ig in 2010-2011 was found to be caused by trace amounts – about the ng/mg of Ig – of unintentionally co-purified plasma residual. Importantly, this residue, the coagulation factor XI, does not belong to top abundant plasma proteins, meaning that every of the hundreds to thousands of different proteins in the plasma pool source material could potentially end up in the Ig products following the fractionation process. While dedicated implementation of specific testing for both thrombogenicity and FXI level resolved the outbreak, this episode raised the interest in a preventive and non-targeted screening of plasma residuals in Ig products. In doing so, high-cost market withdrawal of products might be prevented and practitioners might gain a better view on the exact biodiversity in the available therapeutics. This thesis dedicates to considering mass spectrometry (MS)-based proteomics to enable such non-targeted screening of plasma protein impurities in Ig products.

Chapter 1 provides a detailed view on the modern manufacturing of plasma-derived Ig, including historic perspective, current formulations, safety guidelines and quality control testing. The 2010-2011 TEE outbreak is described to emphasize the importance of non-targeted and sensitive residuals characterization. **Chapter 2** is summarizing the current state-of-art in what MS-based proteomics has to offer for both quality control of plasma-derived therapeutics and non-targeted quantitative screening. The main protein fractionation techniques to increase sensitivity are presented and untargeted MS data acquisition strategies with label-free relative quantification are introduced. In this thesis, *in solution* trypsin digest and reverse phase C18 microflow liquid chromatography (LC) coupled to high resolution QTOF MS – TripleTOF 5600, Sciex – operating in non-targeted acquisition, form the fixed framework of the investigations. To test the applicability of the developed methods, we obtained a set of five different Ig batches, including two products (A, B) from a first manufacturer, with two different batches (A1, A2) from one of the two, one product (C) from a second manufacturer, and finally a TEE-positive batch from 2010-2011 (D).

Chapter 3 is summarizing the objectives of this thesis. In the first part of **Chapter 4**, upstream sample fractionation proves to be mandatory to uncover plasma residuals in the Ig products. Consequently, non-biased enrichment of the plasma impurities is conducted through either highly selective depletion of the Ig component with protein A – protein G affinity chromatography or the non-specific enrichment of low-abundant residues with combinatorial peptide ligand library (CPLL). The latter demonstrates better results, uncovering more impurities and a more reliable identification of FXI spiked at 1 ng/mg. The depletion technique might suffer from the high dilution of the residuals in the flow through, with increased risk of losses during the extra concentration step. Still, the two fractionation methods do not

yield to exactly the same coverage of the impurities, and more extensive fractionation, e.g. depletion followed by CPLL, can be considered. This leads to the following rule of thumb: for each extra fractionation step, the gain in resolution and sensitivity should be worth the extra costs and time spent. Moreover, the repeatability of the overall sample treatment has to be assessed. Threshold of biologically relevant variation has to be above the maximal technical variability observed. Here, optimized CPLL enrichment demonstrates satisfactory performance for our application. 10-fold change variation was chosen as a robust threshold of biological significance, as it is more than two times the highest variation observed between two technical replicates.

In the second part of **Chapter 4**, the golden standard data-dependent acquisition (DDA) was challenged to provide MS1-level relative quantification. In this mode, only a selected number of MS1 ions that surpass a given intensity threshold are further fragmented in MS2 ions. Since the MS1 intensity can fluctuate due to e.g. small differences in sample loading and/or slight shifts in LC retention times, each DDA run can sample somewhat different peptides for MS2. Therefore, repeated injections – here about eight – have to be performed to overcome this inherent technical variability and gain maximal coverage. With a set of 16 runs per sample (8 repeated injections for two CPLL technical replicates), we used data alignment and normalization software Progenesis (Waters) to perform differential analysis, resulting in a heatmap of about 70 plasma residuals identified in the tested Ig products. By evidencing 10-fold variations in the reported abundances, distinct profiles were obtained, with increasing difference when comparing two batches from a given product, two products from a given manufacturer and two products from two different manufacturers. Finally, a TEE-positive batch is distinguished from regular batches by a unique set of proteins with higher abundance, including FXI but also other blood stream-regulator proteins (fibrinogen, angiotensinogen, antithrombin-III, complement component C8…).

In **Chapter 5**, the same Ig samples are reanalysed on the QTOF using a data-independent acquisition (DIA) mode called SWATH-MS. In this mode, the MS1 m/z range is divided into adjacent m/z isolation windows which are independently and consecutively fragmented and analysed, producing an exhaustive MS2 ions map at a given limit of detection, during a single run. In contrast to DDA, quantification with DIA is performed with MS2 signals, which are expected to be even more component-specific and provide a reliable quantification across a modified dynamic range. Signal extraction from the complex output data is usually performed by matching against a spectral library generated with high resolution DDA. Here, an in-house library is built using plasma source material fractionated with optimized CPLL, in order to index the largest possible amount of potential residuals. One of the reference data analysis software pipelines – SWATH 2.0 in Peakview (Sciex) – is used with processing settings as recommended in a recent multicentre benchmark study [254]. Just as observed in this study, reliable quantification of low-abundant signals in complex samples proves challenging. In particular, proteins with good quantitative CV can in fact rely on a vast majority of poorly quantified ions. Here, the SWATH Replicates Analysis 2.0 template from Sciex was used to highlight that the relationship between the MS2 peak area and the variability can be described by a function. This functional relationship appears to be largely insensitive to variation in samples or acquisition conditions, suggesting a device-intrinsic property. By using a power regression, it was shown that the MS2 peak area can be used to predict the quantification repeatability without relying on replicate injections, thus contributing to high-throughput confident quantification of low-abundant signals with SWATH-MS. As proof of applicability for low-abundant proteins quantification, a Python tool is further built to evidence transitions with peak area falling in the area range of desired repeatability, and then flag peptides and proteins that integrate a certain number of such transitions. By doing so, a set of

highly robust quantifications with more than 10-fold variations in abundance was secured and revealed distinct impurity profiles in the different Ig products.

In conclusion, developing relevant non-targeted quantitative MS screening with both relevant sensitivity, namely 6 orders of concentration dynamic range, and satisfactory repeatability (CV<20%) requires both adapted upstream sample fractionation and careful data acquisition and interpretation. Importantly, the two developed data acquisition and processing pipelines provide very consistent profiling of the plasma impurities (**Chapter 6**). MS thus succeeds in filling in the gap of generic and preventive quantitative screening of impurities, which could be performed each time a modification in the source material composition or fractionation process occurs. By applying MS, unexpected clinical safety and market recall issues linked to residuals like the 2010-2011 TEE outbreak can be thus prevented in the future.

As outlined in **Chapter 7** however, some pitfalls still remain. Label-free non-targeted MS only provides relative quantification and accurate sensitivity cannot be extrapolated as such from one protein to another. Consequently, downstream testing with absolute quantification remains necessary for proteins known for causing quality issues and for which a defined maximal level is allowed. MS provides the opportunity to perform such absolute quantification by building targeted selected reactions monitoring (SRM/MRM) lists and spiking labelled standards. Still, these MRM are currently conducted on MS devices with other specifications. Moreover, official reference standards should be available to define an unequivocal correspondence between activity and mass units. In the future, the release of such references associated to the ever-improving offer of MS devices in terms of both sensitivity and resolution might achieve making MS an essential and flexible tool for fine-tune characterization of biological therapeutics, given that consensus on harmonized protocols and data sharing and interpretation can emerge from collective and multi-centre efforts.

Door het toenemende aantal patiënten met immuniteitsaandoeningen of met auto-immuunziekten vormen polyvalente immunoglobulines (Ig) momenteel het belangrijkste product van de plasma-fractionatie-industrie. De Ig-productie gebeurt door afzondering en opzuivering uit een 'pool' van meerdere duizenden plasma-donaties. Hierdoor bevatten Ig een uniek repertoire aan antilichamen tegen een zeer groot aantal verschillende antigenen, hetgeen bijna onmogelijk kan bekomen worden via recombinante moleculaire biologie-technieken. Doordat farmaceutische firma's verplicht zijn om 'Good Manufacturing Practices' (GMP, goede praktijken van produceren) en standaardprocedures in acht te nemen, en doordat er tevens een strikte controle gebeurd van geneesmiddelen door de bevoegde autoriteiten onder de vorm van evaluatie van het registratiedossier, inspecties en kwaliteitscontroletesten, hebben Ig net zoals andere geneesmiddelen een zeer goed veiligheidsprofiel. Het Ig-productieproces is strikt gecontroleerd en de eindproducten zijn steriel. Nevenwerkingen komen niet veel voor en zijn eerder licht of mild en van voorbijgaande aard. Niettemin zijn vrij recent (enkele jaren geleden) toch meerdere problemen opgedoken met Ig. In het bijzonder waren er in 2010-2011 meerdere gevallen van trombo-embolie (TEE) bij patiënten die met Ig behandeld werden. Onderzoek heeft uitgewezen dat deze TEEs veroorzaakt werden door restanten van bepaalde plasma-eiwitten die tijdens de Ig-productie niet of onvoldoende verwijderd waren. De Ig loten die TEEs veroorzaakten, bevatten een lage hoeveelheid van de coagulatiefactor FXI die niet voldoende verwijderd was tijdens de opzuivering. FXI is in vergelijking met andere plasma-eiwitten in zeer lage hoeveelheden aanwezig in humaan plasma. Aangezien plasma meerdere duizenden plasma-eiwitten bevat, betekent dit dat elk van deze eiwitten theoretisch gezien als restant kan aanwezig zijn in plasma-afgeleide geneesmiddelen zoals Ig. Sinds de toename van het aantal gevallen van TEE in 2010-2011, zijn producenten van Ig verplicht om bijkomende specifieke testen uit te voeren op Ig voor thrombogeniciteit en voor de hoeveelheid residuele FXI. Sinds deze voorvallen wordt ook steeds meer geopperd om preventieve, niet-specifieke 'screenings' uit te voeren voor de mogelijke aanwezigheid van plasma-restanten in plasma-afgeleide geneesmiddelen zoals Ig. Deze niet-specifieke 'screening' zou bijkomende veiligheid garanderen en voorkomen dat mogelijke 'probleemloten' (na waarneming van ernstige neveneffecten) van de markt moeten teruggeroepen worden (hetgeen steeds een moeilijk en dure onderneming is). Bovendien kunnen de gegevens van zulke testen mogelijks bijkomende informatie geven over eventuele 'diversiteit' van Ig-producten die op de markt beschikbaar zijn. Deze thesis bestudeert de mogelijkheid om 'proteomics' door middel van massaspectrometrie (MS) te gebruiken voor de niet-specifieke 'screening' van plasma-restanten in Ig.

Hoofdstuk 1 geeft een overzicht van de hedendaagse productiemethodes van plasma-afgeleide Ig, alsook van de geldende richtlijnen en kwaliteitscontroletesten voor Ig. Dit deel beschrijft ook de TEE-problematiek van 2010-2011 en benadrukt het belang van een voldoende gevoelige, niet-specifieke screening voor eventuele restanten of onzuiverheden in plasma-producten.

Hoofdstuk 2 beschrijft meer in detail de huidige 'state-of-the-art' methodes voor MS-gebaseerde proteomics voor controle van plasma-afgeleide geneesmiddelen en voor niet-specifieke (kwantitatieve) screening. Er wordt uitgelegd welke de belangrijkste eiwit-fractionatie-technieken zijn die nodig zijn om de gevoeligheid van MS-methodes te verhogen. Tevens wordt een introductie gegeven over 'untargeted' MS-strategieën met 'label-free' relatieve kwantificatie. Voor het onderzoek van deze thesis werd gebruik gemaakt van 'in solutie' digestie met trypsine en 'omgekeerde fase C18 microflow liquid chromatography' (LC) gekoppeld met hoge resolutie QTOF-MS (TripleTOF 5600, Sciex) in 'non-targeted' data acquisitie modus. Om de toepasbaarheid na te gaan van de ontwikkelde methoden, werden stalen van 5

verschillende Ig loten getest: twee Ig-producten (A, B) van een eerste producent, waarvan twee verschillende loten (A1, A2) van product A en één lot (B) van product B; een Ig-product (C) van een tweede producent; en een Ig lot (D) waarvoor in 2010-2011 meerdere TEE-gevallen werden vastgesteld.

Hoofdstuk 3 vat de doelstellingen van deze scriptie samen. Deel 1 van **hoofdstuk 4** toont aan dat een voorbehandeling door middel van fractionatie noodzakelijk is om resterende plasma-eiwitten te kunnen detecteren en meten in Ig-producten. Aanrijking van plasma-onzuiverheden werd uitgevoerd ofwel door verwijdering van immunoglobulines (Ig) met behulp van proteïne A-proteïne G affiniteitschromatografie, ofwel door niet-selectieve aanrijking van lage-concentratie-restanten met behulp van een 'combinatorial peptide ligand library' (CPLL). Deze laatste methode gaf de beste resultaten: er werden meer onzuiverheden waargenomen en tevens kon extra toegevoegd FXI (aan 1 ng/mg) op een meer betrouwbare manier gedetecteerd worden. De eerste methode (Ig-depletie) heeft het nadeel dat de plasma-restanten veel verdund worden in de elutievloeistof van de chromatografie-stap, met als gevolg een hoger risico voor verlies bij de concentratiestap. Er werd echter waargenomen dat de twee verschillende fractionatiemethoden niet tot exact dezelfde onzuiverheden leidden. Bijkomende fractionatie, bijvoorbeeld Ig-depletie gevolgd door CPLL, kan daarom ook overwogen worden. Hierbij moet wel rekening gehouden worden met het volgende: de winst in resolutie en gevoeligheid bekomen door een bijkomende fractionatiestap moet in verhouding staan tot de extra kosten en tijd. Bovendien moet ook de herhaalbaarheid van een fractionatie-voorbehandeling goed geëvalueerd worden. De drempelwaarde van de biologisch relevante variatie moet hoger zijn dan de maximale technische variatie. Geoptimaliseerde CPLL-aanrijking gaf in deze studie goede resultaten voor het beoogde doel. Een 10-voudige verandering werd gekozen als drempelwaarde voor de biologisch relevante variatie, aangezien dit meer dan 2 keer hoger was dan de hoogste variatie waargenomen voor 2 technische replica's.

In deel 2 van **hoofdstuk 4** werd MS met data-afhankelijke acquisitiemethode ('data-dependent acquisition', DDA) toegepast voor MS1 relatieve kwantificatie. In deze modus worden enkel die MS1 ionen verder gefragmenteerd in MS2 ionen die een bepaalde minimale intensiteit hebben. Aangezien de MS1 intensiteit kan schommelen ten gevolge van bijvoorbeeld kleine verschillen in hoeveelheid geladen staal of kleine veranderingen van de LC-retentietijd, kan elke DDA-analyse resulteren in een lichtjes verschillende set van peptiden voor MS2. Daarom is het noodzakelijk om meerdere injecties (hier: ongeveer 8) uit te voeren om de impact van deze inherente technische variabiliteit te minimaliseren en een maximale 'coverage' te verkrijgen. Een set van 16 analyses per staal (8 herhaalde injecties voor 2 CPLL technische replica's) werd geëvalueerd met behulp van de 'data alignment' en normalisatie-software "Progenesis" (Waters), hetgeen resulteerde in een 'heat map' van ongeveer 70 residuele plasma-eiwitten die werden geïdentificeerd in de geteste Ig-producten. Analyse van de verschillende profielen van deze plasma-eiwitten toonde een toenemende graad van verscheidenheid wanneer 2 loten van hetzelfde product werden vergeleken, wanneer 2 verschillende producten van dezelfde producent werden vergeleken, en wanneer 2 producten van verschillende producenten werden vergeleken. Bovendien toonde het 'TEE-positieve' lot een duidelijk verschillend patroon in vergelijking met de andere, 'normale' loten. Dit lot vertoonde een specifiek set van bepaalde eiwitten met hogere concentratie, waaronder FXI maar ook enkele bloeddruk-controlerende eiwitten (fibrinogeen, angiotensinogeen, antithrombine-III, complement component C8…).

Hoofdstuk 5 beschrijft de analyse van dezelfde Ig-stalen maar nu geanalyseerd in een data-onafhankelijke acquisitiemethode ('data-independent acquisition', DIA), genaamd 'SWATH-MS'. In deze modus wordt de MS1 *m/z* output verdeeld in opeenvolgende zogenaamde '*m/z* isolation windows', die onafhankelijk en opeenvolgend gefragmenteerd en geanalyseerd worden. Dit geeft per analyse een uitgebreide set van MS2 ionen voor een bepaalde detectielimiet. In tegenstelling tot DDA wordt de kwantificatie bij DIA uitgevoerd op MS2 signalen, die normaal gezien nog meer component-specifiek zijn en die een betrouwbare kwantificatie toelaten over een aangepast dynamisch bereik. Signaal-extractie uit de complexe 'output data' gebeurt door vergelijking met een MS-spectra 'bibliotheek' die gegenereerd werd met hoge resolutie DDA. Voor deze studie werd een 'in-house' bibliotheek ontwikkeld op basis van plasma gefractioneerd met geoptimaliseerde CPLL, met als doel een zo groot mogelijke set van eventuele residuele plasma-eiwitten te bekomen. "SWATH 2.0" in "Peakview" software (Sciex) werd gebruikt voor analyse van referentiedata met verwerkingsinstellingen zoals aanbevolen in een recente multicenter studie [285]. Zoals ook in deze studie werd waargenomen, blijkt een betrouwbare kwantificatie van lage signalen in complexe stalen moeilijk te zijn. Eiwitten die in het eindrapport werden gekwantificeerd met een goede kwantitatieve variatiecoëfficiënt (CV) zijn soms gemeten aan de hand van een grote meerderheid aan slecht gekwantificeerde ionen. In deze studie werd de "SWATH Replicates Analysis 2.0 template" van Sciex gebruikt om aan te tonen dat er een verband bestaat tussen de MS2 piek-oppervlakte en de variabiliteit. Dit verband kan in een functie beschreven worden, die relatief ongevoelig is voor variatie in stalen of verwerkings-condities, hetgeen suggereert dat dit een instrument-intrinsieke eigenschap is. Met behulp van regressieanalyse werd aangetoond dat de MS2 piek-oppervlakte gebruikt kan worden om de herhaalbaarheid van de kwantificatie te voorspellen zonder dat er meerdere injecties nodig zijn. Dit draagt bij aan de betrouwbare 'high throughput' kwantificatie van laagfrequente signalen met behulp van SWATH-MS. Als bewijs van toepasbaarheid voor kwantificatie van eiwitten die in lage concentratie aanwezig zijn, werd een "Python"-programma ontwikkeld om transities te analyseren met een piek-oppervlakte binnen het gebied van de gewenste herhaalbaarheid, en om vervolgens peptiden en proteïnen te identificeren die een bepaald aantal dergelijke transities integreren. Door dit uit te voeren, werd een set van zeer robuuste kwantificaties vastgelegd met meer dan 10-voudige variaties in hoeveelheid, en werden verschillende onzuiverheidsprofielen geïdentificeerd in de verschillende Ig-producten.

Er kan besloten worden dat voor de ontwikkeling van 'non-targeted' MS-screeningsmethodes met voldoende hoge gevoeligheid en herhaalbaarheid, zowel een aangepaste fractionatie-voorbehandeling van de stalen vereist is als een zorgvuldige data-acquisitie en -interpretatie. Belangrijk om te vermelden is dat de twee verschillende manieren van data-acquisitie (DDA en DIA) redelijk gelijklopende en consistente profielen van residuele plasma-eiwitten geven (**hoofdstuk 6**). Dit toont aan dat MS-analyse nuttig kan zijn voor preventieve screening van residuele plasma-eiwitten, bijvoorbeeld wanneer er veranderingen zijn in de oorsprong van het plasma of wanneer er aanpassingen aangebracht worden in het productieproces van de Ig-producten. Door toepassing van MS kunnen eventuele onverwachte klinische problemen en terugroeping van vrijgegeven Ig-producten, zoals bij de TEE-problematiek in 2010-2011, mogelijks voorkomen worden in de toekomst.

Zoals beschreven in **hoofdstuk 7**, blijven er echter nog een aantal problemen en tekortkomingen. 'Label-free untargeted' MS geeft enkel een relatieve kwantificatie. De gevoeligheid bekomen voor een bepaald eiwit kan dus ook niet zomaar geëxtrapoleerd worden naar andere eiwitten. Bijgevolg blijft het noodzakelijk om na de hoger beschreven MS-analyse bijkomende absolute kwantificatie uit te voeren voor residuele eiwitten die mogelijks een veiligheidsrisico inhouden voor patiënten en/of waarvoor per dosis een maximaal toegelaten limiet bestaat. MS biedt de mogelijkheid om zulke absolute kwantificatie uit te voeren met behulp van datasets van doelgerichte geselecteerde reactie-monitoring (SRM/MRM) en extra toevoeging van gemerkte standaarden. Deze MRM-analyses worden momenteel echter meestal uitgevoerd met andere MS-instrumenten die andere specificaties hebben. Bovendien moeten er ook officiële referentiestandaarden beschikbaar zijn om een eenduidige relatie te kunnen definiëren tussen activiteit en massa. In de toekomst kan de beschikbaarheid van zulke referentiestandaarden en van steeds betere en performantere MS-instrumenten op vlak van gevoeligheid en resolutie, in combinatie met geharmoniseerde protocollen voor data-verwerking en -interpretatie, verder bijdragen tot het gebruik van MS als een essentieel en flexibel hupmiddel voor een betere karakterisering van biologische geneesmiddelen, en in het bijzonder voor plasma-afgeleide geneesmiddelen.

Chapter 9:

Addendum

9. Addendum

9.1 Supporting Information Chapter 1

Table 9.1 Ph. Eur. guidelines and tests verified by the OMCL for the batch release of Ig products. ***Specific guidelines for hyperimmune Ig products are in italic.*** **Tests performed by the OMCL (OCABR guidelines) are in bold.**

Quality Criterion	Process step	Regulatory status	Performed by	Reference method(s)	Expected value or result
Microbiological safety					
Donors selection	Plasma donation	Ph. Eur. (0853)	Blood establishment	Epidemiological surveillance, donor identification, questionnaire, medical examination, laboratory blood test and study of medical history	Free from detectable agents of infection transmissible by plasma-derived products
Viral markers: anti-HIV-1 and 2, anti-HCV antibodies, HBsAg	Plasma donation	Ph. Eur. (0853)	Blood establishment	Antibodies or antigen test methods	No repeat-reactive result
Viral markers: **anti-HIV-1 and 2 antibodies, HBsAg, HCV RNA**	**Plasma pool**	**Ph. Eur. (0853), OCABR**	**Plasma fractionator, OMCL**	**Antibodies or antigen test methods, NAT testing for HCV (Ph.Eur. 2.6.21)**	**Negative/non-reactive results**
Viral marker: ***B19 virus DNA***	***Plasma pool (for human anti-D)***	***Ph.Eur. (0557,1527), OCABR***	***Plasma fractionator, OMCL***	***NAT testing***	***Negative/non-reactive results*** ***(<10 IU/µL)***

Infectious agents specific removal or inactivation	Fractionation method development	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	heat, S/D or acid treatment, nanofiltration	Implementation of at least 1 dedicated and validated process step for this purpose
Bacteria filtration	Fractionation method development	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	Bacteria retentive sterile filter	/
Sterility	Final product	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	*In vitro* safety test (Ph. Eur. 2.6.1)	Sterile
Antibody to HBSAg	Final product	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	Immunochemical method (Ph. Eur. 2.7.1)	≥ 0.5 IU/g immunoglobulin
Pyrogens or bacterial endotoxins	Final product	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	*In vivo* test for pyrogens (Ph. Eur. 2.6.8) or preferably validated *in vitro* endotoxin test (Ph. Eur. 2.6.14)	Bacterial endotoxin test: IV: < 0.5 IU/mL (if protein content ≤ 5%); < 1.0 IU/mL (if 5% < protein content ≤ 10%) IM, SC: < 5 IU/mL
Integrity: Immunoglobulin content and function					
Total protein	Individual plasma units	Ph. Eur. (0853)	Blood establishment	Determination of nitrogen by sulphuric acid digestion (Ph. Eur. 2.5.9)	≥ 50 g/L
Total protein content	**Final product (except hyperimmune Ig)**	**Ph. Eur. (0338, 0918, 2788), OCABR**	**Plasma fractionator, OMCL**	**Determination of nitrogen by sulphuric acid digestion (Ph. Eur. 2.5.9)**	**Between 90 and 110% of the label-stated quantity (protein content in the range : ≥ 3% (IV), 10-22% (SC) , 10-18 % (IM))**
Diversity and representativeness	Fractionation method development	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	/	At least 1000 donors

	(except for hyperimmune Ig)				
Enrichment of 1 viral and 1 bacterial reference antibodies	Fractionation method development	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	/	≥ 3x enrichment compared to starting plasma pool (at 5% protein content)
Molecular size distribution	**Final product**	**Ph. Eur. (0338, 0918, 2788), OCABR**	**Plasma fractionator, OMCL**	**Size exclusion chromatography (Ph. Eur. 2.2.30)**	**Monomer and dimer peak areas ≥ 85% (SC,IM) or 90% (IV) ; Polymer and aggregates peak areas ≤ 10% (SC,IM) or 3% (IV)**
Anticomplementary activity	Final product (IV)	Ph. Eur. (0918)	Plasma fractionator	Complement consumption titration (Ph. Eur. 2.6.17)	≤ 50% (1 CH_{50} / mg Ig)
IgG subclasses distribution	Fractionation method development	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	Radial immune diffusion or immunoprecipitation techniques	Defined and comparable to physiological balance: IgG_1 = 60 %; IgG_2 = 29.4 %; IgG_3 = 6.6 % ; IgG_4 = 4.2 %
Fc function	Fractionation method development (IV, SC)	Ph. Eur. (0918, 2788)	Plasma fractionator	Haemolysis reaction (Ph. Eur. 2.7.9)	Not less than for reference preparation
Specific antibodies level					
Immunoglobulin A content	Final product	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	Immunochemical method (Ph. Eur. 2.7.1)	Not greater than stated on the label
Anti-A and anti-B hemagglutinins	**Final product (IV, SC)** (except human anti-D Ig IV	**Ph. Eur. (0918, 2788), OCABR**	**Plasma fractionator, OMCL**	**Direct agglutination method** **(Ph. Eur. 2.6.20, method B)**	**Titer < 64 (at 2.5% protein content)**

	with content < 25 g/L)				
Anti-D antibodies	**Final product (IV, SC)** (except anti-D Ig preparations)	**Ph. Eur. (0918, 2788), OCABR**	**Plasma fractionator, OMCL (only IV)**	**Direct agglutination method (Ph. Eur. 2.6.26)**	**ratio to reference preparation (nominal titre of 8) ≤ 1**
Antibody to HAV	***Final product (IM : if intended for prophylaxis of hepatitis A)***	**Ph. Eur. (0338), OCABR**	**Plasma fractionator, OMCL**	***Immunoassay (Ph. Eur. 2.7.1)***	***≥ stated potency ≥ 100 IU/mL***
Potency	***Final product***	**Ph. Eur. (0557, 1527,0769, 1016, 0722, 0397,0723, 0617, 0398, 0724), OCABR**	**Plasma fractionator, OMCL**	***Immunoassay (Ph. Eur. 2.7.1), haemagglutination(-inhibition) test (Ph. Eur. 2.7.13), and/or* in vivo/in vitro *neutralising capacity***	***Anti-D: ≥ 0.9 x stated potency*** ***Measles, hepatitis B (IV): ≥ stated potency≥ 50 IU/mL*** ***Varicella, tetanus, hepatitis B (IM, SC): ≥ stated potency ≥ 100 IU/mL*** ***Hepatitis A: ≥ stated potency ≥ 600 IU/mL*** ***Rabies: 1-2 x stated potency (≥ 150 IU/mL)*** ***Rubella: ≥ 4500 IU/mL***
Purity: residual composition					
Plasma cells and cells debris removal	Individual plasma units	Ph. Eur. (0853)	Plasma units	Sterile separation method (without adding antimicrobial agent)	/
Protein composition: IgG proportion	**Final product**	**Ph. Eur. (0338, 0918, 2788), OCABR**	**Plasma fractionator, OMCL**	**Zone electrophoresis (Ph. Eur. 2.2.31)**	**≤ 10% (SC, IM) or 5% (IV) of protein with mobility different from that of the principal band**

Protein composition: IgG identification	Final product	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	Immunoelectrophoresis with antiserum to normal human serum	/
Virus inactivation substances (if used) residues	Final product	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	/	To be specified on the marketing authorization (No adverse effects on the patients)
Prekallikrein activator	Final product (IV) (except for anti-D Ig)	Ph. Eur. (0918)	Plasma fractionator	Chromogenic assay (Ph. Eur. 2.6.15)	≤ 35 IU/mL (at 3% protein content)
Thrombosis-generating agents removal	Fractionation method development (IV, SC)	Ph. Eur. (0918, 2788)	Plasma fractionator	/	Implementation of at least 1 dedicated and validated process step for this purpose
Coagulation factors and other procoagulant agents follow-up	Fractionation method development (IV, SC)	Ph. Eur. (0918, 2788)	Plasma fractionator	/	/
Thrombogenic (procoagulant) activity	Final product (IV, SC)	Ph. Eur. (0918, 2788)	Plasma fractionator	/	No activity
Physico-chemical properties					
Solubility	**Final product (if freeze-dried)**	Ph. Eur. (0338, 0918, 2788), OCABR	Plasma fractionator, OMCL	**Reconstitution (Ph. Eur. 0918)**	**Complete dissolution within 20 (IM, SC) to 30 (IV) min at 20-25°C (with stated diluent volume)**
pH	Final product	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	pH meter (Ph. Eur. 2.2.3)	4.0 - 7.4 (IV)

					4.6 - 7.2 (SC) 5.0 - 7.2 (IM)
Water content	Final product	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	Semi-micro determination (Ph. Eur. 2.5.12), loss on drying (Ph. Eur. 2.2.32) or near-infrared spectroscopy (Ph. Eur. 2.2.40)	Within stated limits
Osmolality	Final product (IV) (except for hyperimmune Ig)	Ph. Eur. (0918)	Plasma fractionator	Osmometer (Ph. Eur. 2.2.35)	≥ 240 mosmol/kg
Formulation & labelling					
Anticoagulant	Plasma donation	Ph. Eur. (0853)	Blood establishment	/	To be specified on the PMF (appropriated mixing)
Stabilizing agent	Fractionation method development	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	/	To be specified on the marketing authorization (no deleterious effect)
Antimicrobial preservative	Final product (multi-dose IM)	Ph. Eur. (0338)	Plasma fractionator	/	To be specified on the marketing authorization (no deleterious effect)
Product Labelling	Final product	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	/	protein content, amount of Ig, administration route, IgG subclasses distribution, the max IgA content. If applicable: reconstitution instructions, amount of albumin as stabilizer, name and amount of antimicrobial preservative, anti-HAV activity if intended for hepatitis A prophylaxis, number of IU per container if specific Ig

Tolerance					
In vivo tolerance	Fractionation method development	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	Tests in animals and clinical trials evaluation	Tolerance with the labelled administration route
Traceability					
Records of donors and donations traceability	Plasma donation	Ph. Eur. (0853)	Blood establishment & plasma fractionator	Labelling	The origin and acceptance procedure of each donation in a plasma pool is known
Plasma pool traceability	**Plasma pool**	**OCABR**	**OMCL**	**Plasma master file & protocol submission**	**Information about collection procedure, testing, handling and transportation of the plasma for fractionation**
Ig product traceability	**Final product**	**OCABR**	**OMCL**	**Protocol submission**	**Information about production steps and controls as outlined in the Marketing Authorisation and the relevant Ph. Eur. Monographs**
Stability					
Cold chain	Plasma donation	Ph. Eur. (0853)	Blood establishment	Freezing apparatus	For the recovery of labile compounds: <-25°C at the core of each plasma within 12h. For the recovery of stable compounds: <-20°C within 24 (source plasma) or 72 (recovered plasma) hours
Labile proteins conservation	Plasma donation	Ph. Eur. (0853)	Blood establishment	Coagulation factor VIII Activity assay (Ph. Eur. 2.7.4)	≥ 0.7 IU/mL

Cold chain	**Plasma storage and transport**	**Ph. Eur. (0853), OCABR**	**Blood establishment, OMCL**	**Freezing apparatus**	**<-20°C**
Plasma appearance	Plasma pool	Ph. Eur. (0853)	Blood establishment	Visual inspection	Clear or slightly turbid liquid, from light yellow to green
Appearance	**Final product**	**Ph. Eur. (0338, 0918, 2788), OCABR**	**Plasma fractionator, OMCL**	**Visual inspection**	**liquid: clear and colourless; freeze-dried: white hygroscopic powder**
Final product stability	Fractionation method development	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator	/	Stable within the specified shelf life
Product storage	Final product	Ph. Eur. (0338, 0918, 2788)	Plasma fractionator, practitioner	/	Liquid : colourless glass container, protected from light, at labelled temperature Freeze-dried : airtight colourless glass container, protected from light, at temperature ≤ 25°C

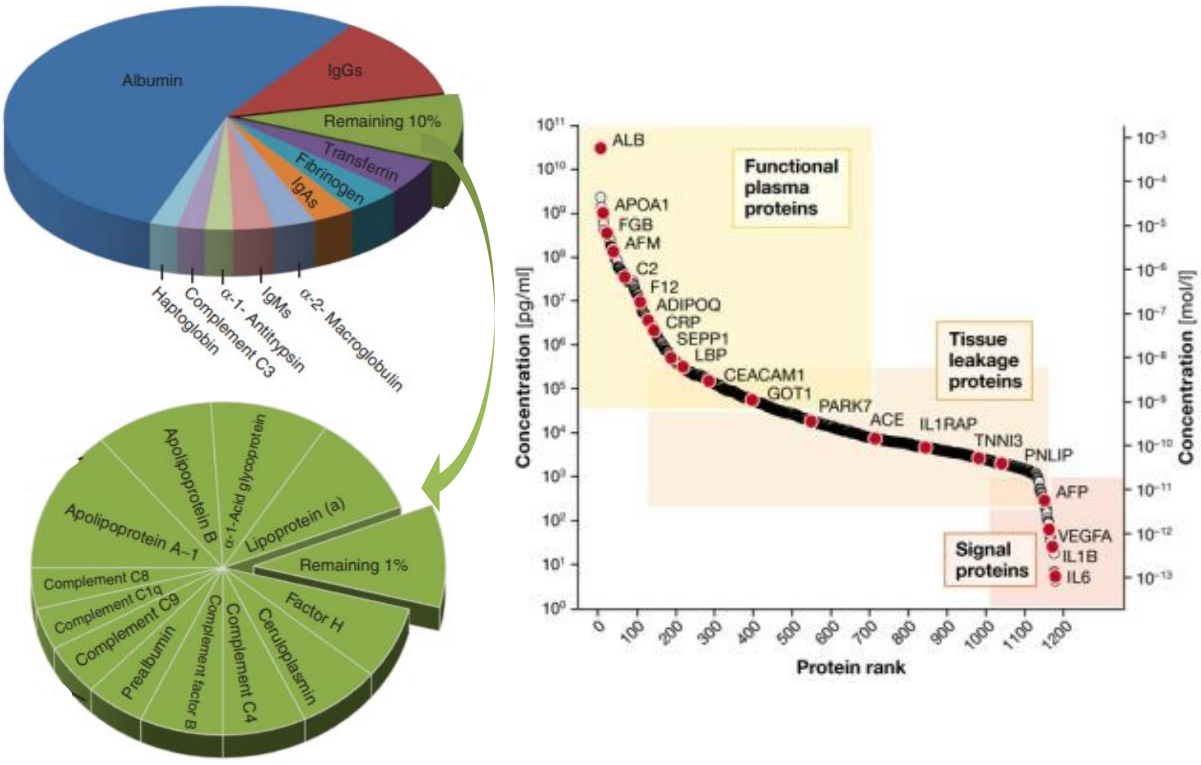


Figure 9.1 Protein concentration dynamic range and percentage in plasma.
22 most abundant proteins represent about 99% of the total protein content, while remaining 1% represent thousands of proteins (pie charts are reproduced from [325] **and graph from** [326]**)**

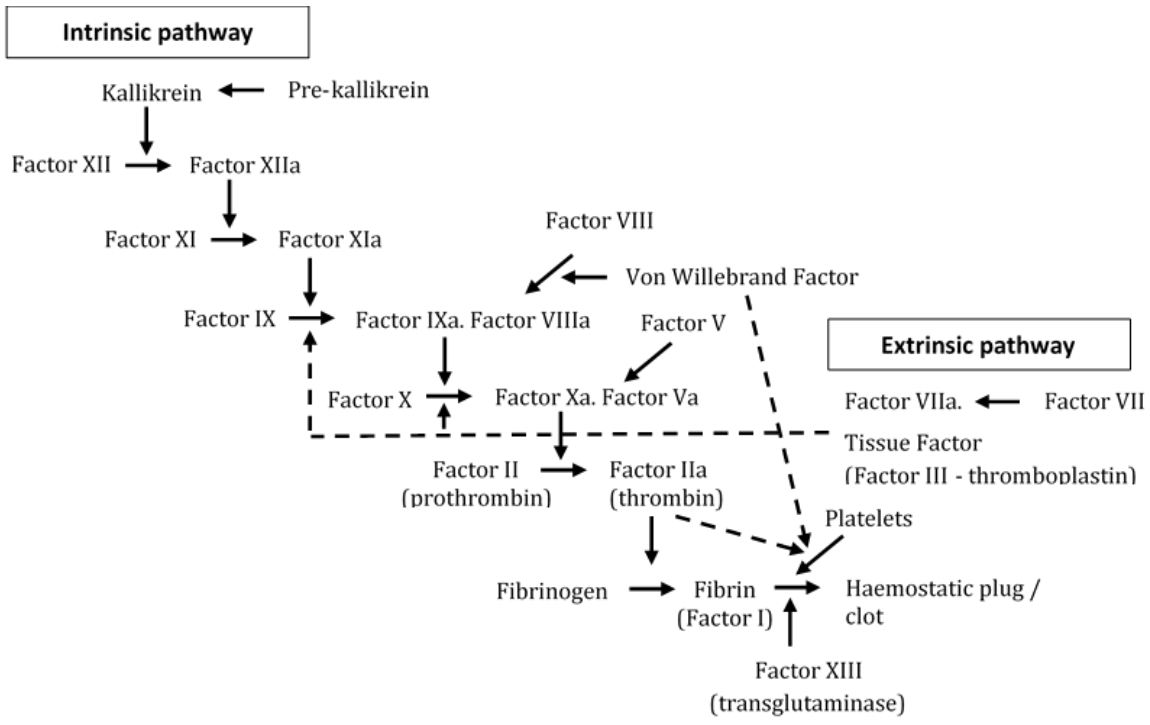


Figure 9.2 Overview of the blood coagulation cascade and its main components (adapted from [327]**)**

9.2 Supporting Information Chapter 4

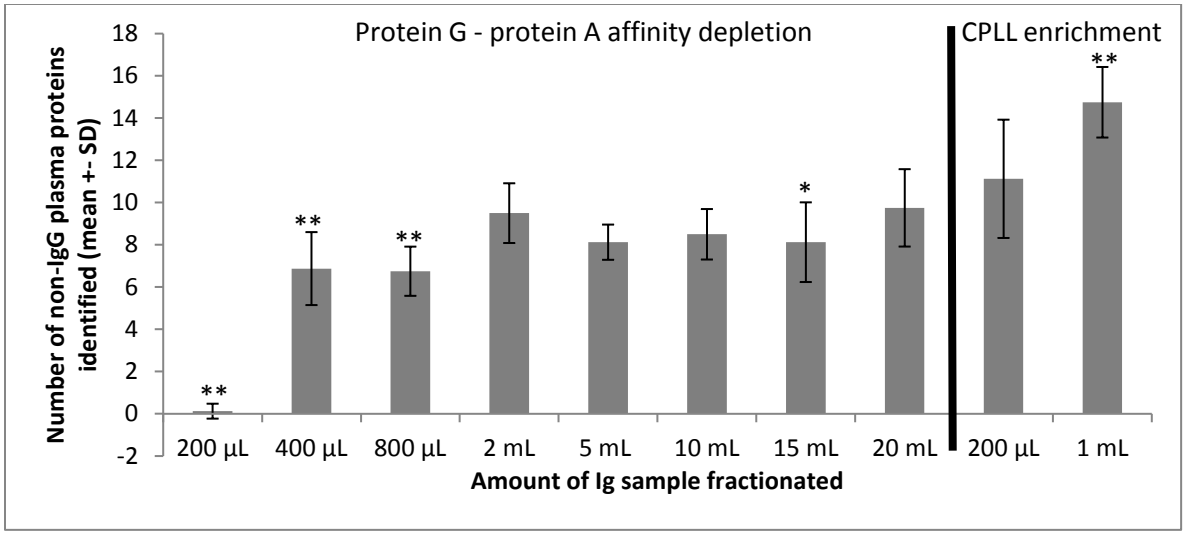


Figure 9.3 Number (mean ± SD, n=8) of (non-IgG) plasma proteins identified in Ig sample (product A1) depending on sample amount fractionated, either with protein G – protein A affinity depletion or CPLL enrichment. (Each sample was injected eight times on LC-MS/MS.) Mann-Whitney U test was used to compare the different loads for each technique (for depletion, the load of 2mL is taken as reference). ** indicates p-value < 0.01, * indicates p-value between 0.01 and 0.05.

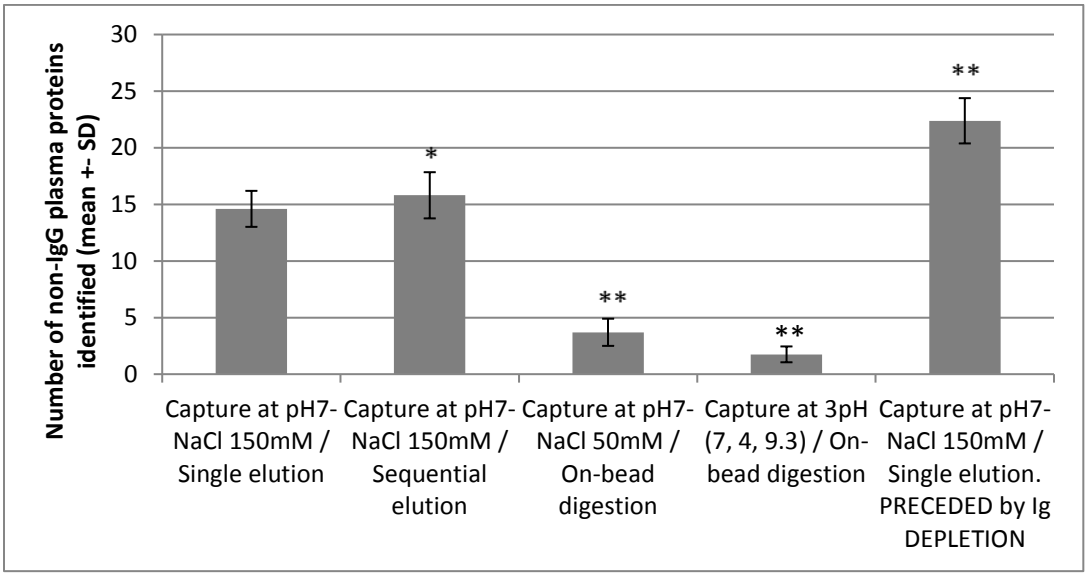


Figure 9.4 Optimization of CPLL enrichment, considering mean number ± SD (n=16) of non-IgG plasma protein identifications found for each protocol variant in Ig product (A1) spiked with FXIa 1ng/mg. Each protocol variant was tested twice and each time injected 8 times on LC-MS/MS, giving a total of 16 runs per condition. ** indicates Mann-Whitney U test p-value < 0.01, * indicates p-value between 0.01 and 0.05. (The capture at pH7 – NaCl 150 mM / normal elution is taken as reference.)

Remarks:
- Large capacity (1mL load) ProteoMiner™ spin columns were always used, except when CPLL was preceded by depletion. In this case, small capacity (200µL load) spin columns were used, in order to maintain required concentration (≥ 50 mg/mL)
- In case of multistep ("sequential") elution, eluted fractions can be injected either apart or pooled (maintaining always 2 µg per injection). Those two ways of injecting did not exhibit significantly different results (p-value>0.05, m=4, where m is a series of 4 injections, either 1 injection of each fraction, either 4 injections of pooled fractions). Here, fractions obtained with sequential elution were always pooled before injection.

Table 9.2 List of identified (non-IgG) plasma proteins in product A1, depending on the sample fractionation technique. Each fractionation was tested twice and each time injected 8 times on LC-MS/MS, giving a total of 16 runs per condition. X denotes identification with at least 2 unique peptides and Mascot score ≥ 50 while (X) denotes identification with only 1 unique peptide but with Mascot score ≥ 80.

Protein	Ig depletion	CPLL enrichment				
	Protein G – Protein A affinity chromatography	Capture at pH7 – NaCl 150 mM / single elution	Capture at pH7 – NaCl 150 mM / sequential elution	Capture at pH7 – NaCl 50 mM / On-bead digestion	Capture at 3 pH (7,4,9.3) –/ On-bead digestion	Capture at pH7 – NaCl 150 mM / Single elution. PRECEDED by Ig DEPLETION
Adenosine deaminase	X	X	X	(X)		X
Albumin	X	X	X	X		X
Alpha-2-HS-glycoprotein				(X)		
Antileukoproteinase			(X)			
Antithrombin-III		X				X
Apolipoprotein H	X	X	X			X
ATP-dependent RNA helicase DHX29		(X)				
Catalase	X	X	X	X		X
CD5 antigen-like						X
Complement C1q subunit A		X	X			(X)
Complement C1q subunit B		X	X			X
Complement C1q subunit C		X	X			X
Complement C2	X	X	X			(X)
Complement C3		X	X			
Complement C4-A(B)			X	(X)	(X)	X
Complement component C6	X	X	X			X
Complement component C7		(X)	(X)			(X)
Complement component C8 gamma chain		(X)	(X)			X
Complement factor B	X	X	X			X
Complement factor I	X	(X)	(X)			X
Fibrinogen alpha-chain			X			X
Gelsolin	X	X	X	X		X
Hemopexin						X
Hepatocyte growth factor-like protein		X	X	(X)		X
Ig alpha-1 chain C region	X	X	X	X	X	X
Ig alpha-2 chain C region						X
Ig delta chain C region		X	X			X

Ig mu chain C region	(X)	X	X			X
Lysozyme C		X	X	X	X	X
Methionine adenosyltransferase 2 subunit beta					(X)	
Macrophage receptor MARCO						X
Pepsin A		(X)	(X)			
Peroxiredoxin-1						X
Plasma kallikrein	X	X	X	(X)	(X)	X
Plasma serine protease inhibitor	(X)					X
Properdin	X	X	X			X
Proteoglycan 4		X	X	X		
Ribonuclease pancreatic			(X)			
Serotransferrin	X	X	X	X	X	X
SH3 domain-containing protein 21				(X)		
Synaptogamin-like protein 4		(X)	(X)	(X)		X
Thrombospondin-4		(X)	(X)			X
Total	**15**	**29**	**31**	**14**	**6**	**33**

Table 9.3 MS precursors normalized abundances calculated with Progenesis of (non-IgG) plasma proteins in two batches from IVIg product A [A1 and A2], two other IVIg products, one from the same manufacturer [B], one from another [C] and a TEE-positive ScIg batch [D]. A1 and A2 are spiked with FXIa 0.2 ng/mg Ig. Each sample was enriched with Proteominer large capacity – sequential elution and desalted with protein precipitation (two technical replicates, T1 and T2). Each technical replicate was injected 8 times, giving a total of 16 runs per sample. Progenesis was used to calculate MS precursors normalized abundances. MS/MS spectra of those precursors were exported as MGF file to be searched in Mascot. Confidence score above 50 and at least 1 unique peptide were requested as identity threshold. * denotes single unique peptide matches.

Sample		Product A - batch 1 - FA11 0.2ng/mg spiked		Product A - batch 2 - FA11 0.2ng/mg spiked		Product B		Product C		Product D – TEE-positive batch	
Uniprot entry	Protein name	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
CECR1	Adenosine deaminase CECR1	1099	385	338	370	35	29	28	23	102	80
AFAM	Afamin	666	502	510	653	587	501	59812	38382	667	692
A1AG1	Alpha-1-acid glycoprotein 1	184	163	160	277	145	173	2640	1878	356	321
A1AG2	Alpha-1-acid glycoprotein 2 *	0	1	0	53	0	0	300	250	0	2
AACT	Alpha-1-antichymotrypsin	54	21	24	38	10	20	971	685	29	12
A1AT	Alpha-1-antitrypsin	19	9	20	27	42	46	139	93	1961	1940
A1BG	Alpha-1B-glycoprotein	8420	4979	2138	2047	5266	5618	15348	10673	2410	4741
FETUA	Alpha-2-HS-glycoprotein	73	56	141	159	374	149	39851	28053	120	110
A2MG	Alpha-2-macroglobulin	2186	1267	946	865	1539	1846	1208	803	32140	30400
APC5	Anaphase-promoting complex subunit 5 *	98	86	55	55	75	97	83	59	41	49
ANGT	Angiotensinogen *	0	0	1	1	2	0	0	0	136	152
SLPI	Antileukoproteinase	190	100	10	64	6	7	16	19	4	4
ANT3	Antithrombin-III	755	328	166	271	731	803	462	334	6151	6199
APOA2	Apolipoprotein A-II	52	20	118	73	1	1	15482	10208	94	109
APO1B	Apolipoprotein C-I, basic form (Fragment) *	1	1	2	0	0	1	536	254	0	1
APOH	Beta-2-glycoprotein 1	5050	7023	2931	4316	112	153	10689	6538	3958	4881
CATA	Catalase	1427	1749	2219	2301	78	26	55	35	666	505
CD5L	CD5 antigen-like	1148	576	435	503	27803	30276	341	298	29847	29680
FA11	Coagulation factor XI	821	1482	17	37	21	15	30	19	2152	1742
CC148	Coiled-coil domain-containing protein 148	1673	1520	1428	1547	2009	2013	1531	1059	1531	1463
C1QB	Complement C1q subcomponent subunit B	736	1095	1254	1205	135	70	442	242	619	687

C1QC	Complement C1q subcomponent subunit C	661	903	1010	1006	78	67	769	525	611	542
CO2	Complement C2	550	515	334	599	144	150	289	180	416	333
CO3	Complement C3	4029	3812	2852	7756	1871	907	12632	7454	1879	967
CO4A	Complement C4-A	4151	1728	3643	11507	319	202	2629	1519	158	143
CO8A	Complement component C8 alpha chain	167	156	48	68	46	51	646	252	1965	1874
CO8B	Complement component C8 beta chain	87	101	72	80	699	630	87	64	2274	2180
CO8G	Complement component C8 gamma chain	563	555	179	181	88	135	123	111	9611	9214
CFAB	Complement factor B	895	711	1779	3305	303	298	5250	3751	3356	3144
CFAI	Complement factor I	87	75	30	53	43	31	7	10	4015	2791
ENPP2	Ectonucleotide pyrophosphatase/phosphodiesterase family member 2 *	1	2	6	5	5	4	2	0	172	184
HMCES	Embryonic stem cell-specific 5-hydroxymethylcytosine-binding protein *	116	118	112	117	99	96	92	58	68	85
FETUB	Fetuin-B *	3	4	2	1	7	4	48	25	51	41
FIBA	Fibrinogen alpha chain	406	236	207	85	86	167	596	378	8641	6755
FIBB	Fibrinogen beta chain	129	85	153	78	123	121	178	147	6498	4078
FIBG	Fibrinogen gamma chain	326	298	195	131	409	308	405	308	14330	9800
FCN3	Ficolin-3	823	671	1124	2652	820	1033	1091	588	20368	18749
FLNC	Filamin-C *	214	205	156	179	141	160	178	106	179	176
GELS	Gelsolin *	14	18	21	5	15	21	48	28	327	362
GOGA7	Golgin subfamily A member 7	608	535	539	404	1185	1183	419	300	809	898
HPT	Haptoglobin	3	2	103	70	3	8	33225	21491	70	63
HBB	Hemoglobin subunit beta	12	3	70	83	1	1	5024	3247	94	91
HEMO	Hemopexin	1308	1329	2718	3165	4918	4271	285233	191641	40356	39255
HGFL	Hepatocyte growth factor-like protein	966	1317	91	184	109	112	820	579	2695	2969
IGHA1	Ig alpha-1 chain C region	27290	18681	9723	13094	367159	397029	4203	957	232117	256744
IGHA2	Ig alpha-2 chain C region	516	101	67	73	22090	30458	343	181	12173	13508
IGHD	Ig delta chain C region	181	176	88	185	1280	1299	57	37	1651	1468
IGHM	Ig mu chain C region	23510	10122	3602	2857	525279	528370	4957	3556	507065	516447
MUCB	Ig mu heavy chain disease protein *	2344	644	5176	7719	37030	45610	5984	551	31432	27676

IGJ	Immunoglobulin J chain	2363	796	371	466	43372	52536	275	198	41598	47292
IGLL5	Immunoglobulin lambda-like polypeptide 5	386344	344550	345564	345333	336746	339915	340526	217619	324876	313261
ITIH4	Inter-alpha-trypsin inhibitor heavy chain H4 *	0	0	0	0	0	0	0	0	228	296
FCG3A	Low affinity immunoglobulin gamma Fc region receptor III-A	434	499	287	384	3748	3055	840	580	389	400
LYSC	Lysozyme C	50995	25888	21126	43848	157	196	10126	5815	261	199
PGRP2	N-acetylmuramoyl-L-alanine amidase	22	33	83	28	2	3	1862	1064	6	12
NEUL	Neurolysin, mitochondrial *	171	121	156	381	0	3	3	5	0	0
PEPA_PIG	Pepsin A *	264	389	228	318	2	1	181	104	5	5
KLKB1	Plasma kallikrein	29060	57379	14738	20369	4315	2556	2568	1619	14863	14449
IC1	Plasma protease C1 inhibitor	1415	1535	2936	4577	1708	1697	209990	133559	2257	1696
PROP	Properdin	9951	3103	635	1880	162	74	73	46	222	248
BROMI	Protein broad-minded	1408	1186	754	717	1640	1403	228	301	1191	1105
TRFE	Serotransferrin	8843	12891	11996	12441	1894	1156	1293	884	8189	7144
ALBU	Serum albumin	73674	75462	155190	251367	57731	33141	7563012	4966755	185631	236394
TTHY	Transthyretin	603	784	1329	2392	461	361	76552	44583	375	791
CS068	Uncharacterized protein C19orf68	4030	3532	2983	3168	5001	4961	3477	2261	4512	4695
UACA	Uveal autoantigen with coiled-coil domains and ankyrin repeats protein *	185	235	287	218	244	232	420	272	390	538

9.3 Supporting Information Chapter 5

Figure 9.5 Heatmap of mean SWATH abundances of (non-IgG) plasma proteins in five different Ig products, as reported in Peakview automatic results. Tested products include two batches from IVIg product A (A1 and A2) and two other IVIg products, one from the same manufacturer (B), one from another (C) and finally a TEE-positive ScIg batch (D). A1 and A2 are spiked with FXIa 0.2 ng/mg Ig. Each sample was enriched using CPLL strategy (two technical replicates, T1 and T2) and each fractionated sample was injected three times. Heatmap was generated in Qlucore Omics Explorer, with data normalization for each protein across the five tested samples (mean=0, variance=1) and hierarchical clustering option.

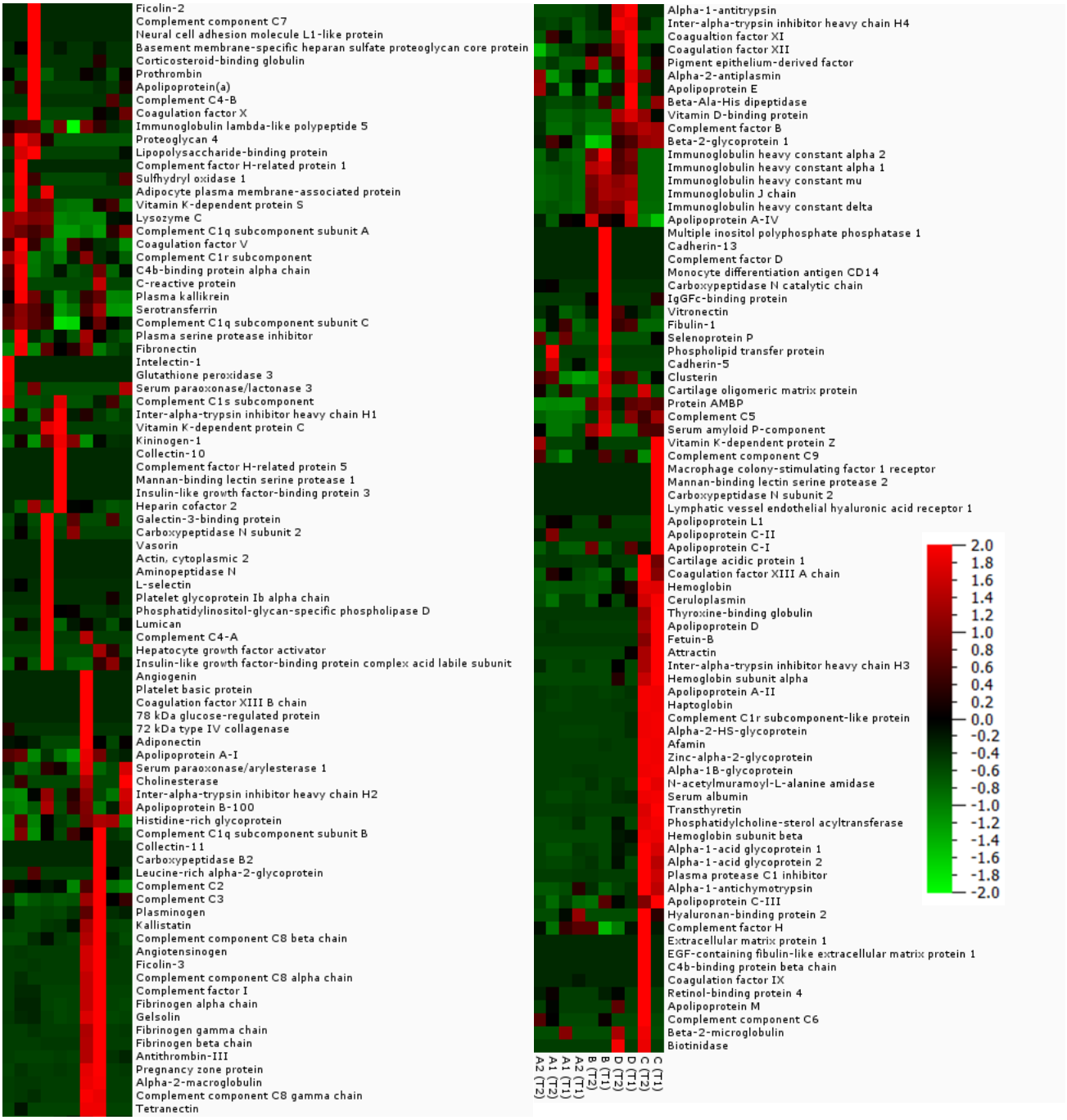


Table 9.4 List of the proteins (alphabetical order) for which spectral information is indexed in each of the used libraries (FDR <1%, shared peptides excluded). Refer to the Material & method section of Chapter 5 for libraries generation details.

The Excel file is downloadable using the following Dropbox link:

https://www.dropbox.com/sh/yvmdgjxe9ck5jrv/AACgSqm00rfmUnF5tDRFEH1ca?dl=0

Table 9.5 Automatic results with Peakview SWATH 2.0 processing. Sample is fractionated Ig (A1) spiked with FXIa 1ng/mg. Three different libraries were tested: (i) homologous fractionated Ig (ii) in-house fractionated plasma source material and (iii) external fractionated plasma. Two examples of proteins quantified with CV<20% in spite of poorly reproducible peptides and ions integration.

Two examples of proteins quantified with CV<20% in spite of poorly reproducible peptides and ions integration								
Protein			**Peptide**			**Ions**		
Name	**CV**	**Peak Area**	**Name**	**CV**	**Peak Area**	**Name**	**CV**	**Peak Area**
FIBG	0.09	5620	IHLISTQSAIPY	0.80	2018	b10	0.78	535
						b9	0.84	1209
						b7	1.20	163
						b8	1.73	75
						b4	1.73	18
						b6	1.73	17
			LTIGEGQQHHLGGAK	0.39	1036	y12	0.91	46
						y10	1.73	167
						y8	0.00	110
						b10	1.33	483
						b11	1.00	0
						y11	0.72	231
			HAGHLNGVYYQGGTYSK	0.10	484	y9	1.00	55
						b8	0.43	179
						y8	0.35	92
						y7	1,00	0
						b7	0.49	126
						b9	0.88	33
			TVQIHDITGK	1.00	2082	y8	0.87	458
						y7	1.27	97
						y6	1.14	412
						b5	0.97	57
						b8	0.73	254
						b6	1.55	805
ITIH4	0.12	44194	ANTVQEATFQMELPK	0.29	3838	y10	1.05	290
						y9	0.87	597
						y8	0.62	600
						y7	0.86	218

						y6	0.81	1536
						y5	0.68	597
			AGFSWIEVTFK	1.03	2308	y6	0.97	1047
						y8	1.62	382
						y7	1.61	604
						b9	1.73	45
						y9	0.44	73
						y10	1.56	157
			ETLFSVMPGLK	0.13	16479	y8	0.21	4418
						y7	0.03	4149
						y9	0.24	1122
						y6	0.36	1747
						y10	0.20	1047
						y4	0.22	3996
			TGLLLLSDPDK	0.76	2751	y7	0.84	697
						y6	0.98	633
						y8	0.80	575
						y9	0.99	206
						y10	0.93	119
						y5	0.69	521
			ITFELVYEELLKR	0.63	1444	y7	0.87	37
						y8	1.00	55
						y6	0.83	962
						y11	0.65	128
						y9	0.87	110
						y5	0.63	153
			GIDIYSLTVDSR	1.18	3550	y8	0.87	373
						y7	1.42	2245
						y10	1.73	4
						y6	0.74	260
						y9	1.03	366
						y5	1.32	302
			TSMVVTKPDDQEQSQVAEKPME GESR	0.78	1372	y10	0.90	281
						y19	1.17	333
						y9	1.73	151
						y11	1.00	0
						y21	1.00	0
						y22	0.93	608
			FKPTLSQQQK	0.22	570	y8	0.43	73
						b8	1.03	58
						y7	0.92	46
						y6	0.47	225
						y5	1.15	149

						b9	1.73	18
			YFAPEGLTTMPK	0.82	4640	y9	1.63	2411
						y10	0.87	358
						y7	1.71	1117
						y8	1.73	333
						y6	1.18	173
						y11	0.69	249
			DQFNLIVFSTEATQWR	1.55	4403	y9	1.46	902
						y10	1.35	479
						b5	1.61	808
						y4	1.51	761
						b4	1.56	266
						b8	1.70	1187
			SIQNNVR	0.12	2836	y5	0.16	2544
						y4	0.87	37
						y6	0.87	37
						b6	0.91	31
						b5	1.73	37
						b4	0.61	151

Table 9.6 Example of CV predictability using the MS2 peak area: predicted 80th percentile CV

The Excel file is downloadable using the following Dropbox link:

https://www.dropbox.com/sh/yvmdgjxe9ck5jrv/AACgSqm00rfmUnF5tDRFEH1ca?dl=0

Table 9.7 Example of CV predictability using the MS2 peak area: Predicted peak area from which at least 80% of the data has a CV<20%

The Excel file is downloadable using the following Dropbox link:

https://www.dropbox.com/sh/yvmdgjxe9ck5jrv/AACgSqm00rfmUnF5tDRFEH1ca?dl=0

Table 9.8. Flagged (non-IgG) protein quantifications (number of flagged peptides, mean, CV) in the five tested Ig products. Those include: two IVIg products from a first manufacturer, including two different batches from one of the two products (A1, A2 and B), one IVIg product from another manufacturer (C) and one TEE-positive batch from SclG product (D). Flagged peptides include at least two ions with a peak area in the high reproducibility range (here 2e4-1e6).

Product	Flagged (non-IgG) protein	Number of flagged peptides	Mean abundance	CV
A1 (T1)	sp|P01871|IGHM_HUMAN	7	1.65E+06	0.07
	sp|P03952|KLKB1_HUMAN	4	1.36E+06	0.07
	sp|P01876|IGHA1_HUMAN	3	1.25E+06	0.09
	sp|P61626|LYSC_HUMAN	2	4.88E+05	0.07
	sp|P02787|TRFE_HUMAN	1	3.95E+05	0.06
	sp|P02749|APOH_HUMAN	1	2.10E+05	0.09

	sp|P02768|ALBU_HUMAN	1	8.31E+05	0.21
A1 (T2)	sp|P02768|ALBU_HUMAN	31	7.60E+05	0.16
	sp|P03952|KLKB1_HUMAN	18	2.12E+06	0.03
	sp|P01871|IGHM_HUMAN	11	8.39E+05	0.08
	sp|P01876|IGHA1_HUMAN	3	9.32E+05	0.13
	sp|P61626|LYSC_HUMAN	2	5.69E+05	0.05
	sp|P02647|APOA1_HUMAN	1	8.79E+04	0.52
	sp|P02787|TRFE_HUMAN	1	4.73E+05	0.1
	sp|P01877|IGHA2_HUMAN	1	1.32E+05	1.47
	sp|P02749|APOH_HUMAN	1	2.76E+05	0.05
A2 (T1)	sp|P02768|ALBU_HUMAN	40	3.36E+06	0.02
	sp|P03952|KLKB1_HUMAN	16	4.96E+05	0.01
	sp|P01876|IGHA1_HUMAN	3	2.82E+05	0.34
	sp|P61626|LYSC_HUMAN	2	5.03E+05	0.06
	sp|P04114|APOB_HUMAN	1	1.12E+05	0.57
	sp|P02749|APOH_HUMAN	1	1.11E+05	0.04
	sp|P01011|AACT_HUMAN	1	7.01E+04	1.15
	sp|B9A064|IGLL5_HUMAN	1	3.54E+06	0.02
A2 (T2)	sp|P02768|ALBU_HUMAN	15	3.43E+06	0.06
	sp|P03952|KLKB1_HUMAN	3	8.12E+05	0.11
	sp|P01876|IGHA1_HUMAN	3	4.70E+05	0.21
	sp|P61626|LYSC_HUMAN	2	5.18E+05	0.11
	sp|P02749|APOH_HUMAN	1	1.65E+05	0.09
	sp|P01871|IGHM_HUMAN	1	1.02E+05	0.4
B (T1)	sp|P02768|ALBU_HUMAN	35	1.83E+06	0.40
	sp|P01871|IGHM_HUMAN	16	3.09E+07	0.04
	sp|P01876|IGHA1_HUMAN	6	1.50E+07	0.05
	sp|P02790|HEMO_HUMAN	1	1.24E+05	0.04
	sp|P08571|CD14_HUMAN	1	2.42E+04	1.03
	sp|P02760|AMBP_HUMAN	1	9.15E+04	0.02
	sp|P01591|IGJ_HUMAN	1	1.19E+06	0.03
	sp|P01877|IGHA2_HUMAN	1	7.91E+05	0.05
	sp|P01009|A1AT_HUMAN	1	1.12E+05	0.23
	sp|B9A064|IGLL5_HUMAN	1	1.90E+06	0.88
B (T2)	sp|P01871|IGHM_HUMAN	11	2.47E+07	0.07
	sp|P02768|ALBU_HUMAN	9	1.81E+06	0.45
	sp|P01876|IGHA1_HUMAN	3	1.38E+07	0.06
	sp|P02760|AMBP_HUMAN	1	7.05E+04	0.81
	sp|P01591|IGJ_HUMAN	1	9.23E+05	0.04
	sp|P01877|IGHA2_HUMAN	1	4.99E+05	0.05
C (T1)	sp|P02768|ALBU_HUMAN	43	8.71E+07	0.04
	sp|P02790|HEMO_HUMAN	13	5.90E+06	0.06
	sp|P00738|HPT_HUMAN	10	7.72E+06	0.06

	sp|P05155|IC1_HUMAN	8	4.04E+06	0.08
	sp|P04217|A1BG_HUMAN	5	6.79E+05	0.1
	sp|P02766|TTHY_HUMAN	4	1.62E+06	0.07
	sp|P43652|AFAM_HUMAN	4	9.51E+05	0.08
	sp|P02652|APOA2_HUMAN	2	5.08E+05	0.06
	sp|P02765|FETUA_HUMAN	2	4.74E+05	0.06
	sp|P68871|HBB_HUMAN	2	2.95E+05	0.07
	sp|P19652|A1AG2_HUMAN	1	9.62E+04	0.08
	sp|P00751|CFAB_HUMAN	1	4.18E+05	0.08
	sp|P01876|IGHA1_HUMAN	1	8.93E+04	1.21
	sp|P04114|APOB_HUMAN	1	1.18E+05	0.5
	sp|P02749|APOH_HUMAN	1	3.52E+05	0.04
	sp|Q9NQ79|CRAC1_HUMAN	1	7.70E+04	0.73
	sp|B9A064|IGLL5_HUMAN	1	3.95E+06	0.04
	sp|P61626|LYSC_HUMAN	1	2.35E+05	0.1
	sp|P01024|CO3_HUMAN	1	2.17E+05	0.09
C (T2)	sp|P02768|ALBU_HUMAN	43	8.53E+07	0.03
	sp|P00738|HPT_HUMAN	11	7.94E+06	0.04
	sp|P02790|HEMO_HUMAN	11	5.97E+06	0.1
	sp|P05155|IC1_HUMAN	9	5.16E+06	0.03
	sp|P04217|A1BG_HUMAN	5	6.76E+05	0.07
	sp|P02766|TTHY_HUMAN	4	1.42E+06	0.02
	sp|P43652|AFAM_HUMAN	4	9.70E+05	0.04
	sp|P02652|APOA2_HUMAN	2	5.18E+05	0.06
	sp|P02765|FETUA_HUMAN	2	4.62E+05	0.05
	sp|P68871|HBB_HUMAN	2	3.05E+05	0.05
	sp|P19652|A1AG2_HUMAN	1	1.34E+05	0.07
	sp|P00751|CFAB_HUMAN	1	4.43E+05	0.09
	sp|P02760|AMBP_HUMAN	1	7.22E+04	0.04
	sp|P01876|IGHA1_HUMAN	1	7.95E+04	1.28
	sp|P02749|APOH_HUMAN	1	3.47E+05	0.07
	sp|Q9NQ79|CRAC1_HUMAN	1	1.77E+05	0.84
	sp|B9A064|IGLL5_HUMAN	1	4.10E+06	0.05
	sp|P61626|LYSC_HUMAN	1	2.20E+05	0.05
	sp|Q12805|FBLN3_HUMAN	1	5.57E+04	0.77
	sp|P01024|CO3_HUMAN	1	1.56E+05	0.12
D (T1)	sp|P01023|A2MG_HUMAN	28	1.13E+07	0.03
	sp|P02768|ALBU_HUMAN	21	8.96E+06	0.01
	sp|P02790|HEMO_HUMAN	12	2.32E+06	0.02
	sp|P01871|IGHM_HUMAN	11	3.05E+07	0.03
	sp|P03952|KLKB1_HUMAN	5	1.28E+06	0.06
	sp|P01008|ANT3_HUMAN	4	4.50E+05	0.03
	sp|O75636|FCN3_HUMAN	3	8.19E+05	0.02

	sp|P01876|IGHA1_HUMAN	3	1.00E+07	0.04
	sp|P02679|FIBG_HUMAN	3	7.31E+05	0.03
	sp|P01009|A1AT_HUMAN	2	6.67E+05	0.03
	sp|P01019|ANGT_HUMAN	2	3.36E+05	0.09
	sp|P03951|FA11_HUMAN	2	5.14E+05	0.07
	sp|P06396|GELS_HUMAN	2	5.77E+05	0.07
	sp|P07360|CO8G_HUMAN	2	4.64E+05	0.1
	sp|P07358|CO8B_HUMAN	1	2.58E+05	0.03
	sp|P02675|FIBB_HUMAN	1	7.06E+05	0.04
	sp|Q14624|ITIH4_HUMAN	1	1.20E+05	0.01
	sp|P05156|CFAI_HUMAN	1	2.25E+05	0.03
	sp|P02671|FIBA_HUMAN	1	4.10E+05	0.17
	sp|P00751|CFAB_HUMAN	1	3.88E+05	0.07
	sp|P20742|PZP_HUMAN	1	2.16E+05	0.03
	sp|P01591|IGJ_HUMAN	1	1.15E+06	0.07
	sp|P01880|IGHD_HUMAN	1	8.30E+04	0.03
	sp|P02749|APOH_HUMAN	1	2.74E+05	0.05
	sp|P01877|IGHA2_HUMAN	1	5.22E+05	0.05
	sp|B9A064|IGLL5_HUMAN	1	4.93E+06	0.1
D (T2)	sp|P01023|A2MG_HUMAN	28	1.09E+07	0.02
	sp|P02768|ALBU_HUMAN	21	8.16E+06	0.03
	sp|P02790|HEMO_HUMAN	12	1.85E+06	0.11
	sp|P01871|IGHM_HUMAN	10	2.72E+07	0.03
	sp|P03952|KLKB1_HUMAN	5	1.04E+06	0.03
	sp|P01008|ANT3_HUMAN	3	3.12E+05	0.15
	sp|P01876|IGHA1_HUMAN	3	9.01E+06	0.02
	sp|P07360|CO8G_HUMAN	3	4.85E+05	0.02
	sp|O75636|FCN3_HUMAN	2	6.71E+05	0.03
	sp|P01009|A1AT_HUMAN	2	5.65E+05	0.06
	sp|P03951|FA11_HUMAN	2	4.19E+05	0.08
	sp|P01591|IGJ_HUMAN	1	1.12E+06	0.01
	sp|P01877|IGHA2_HUMAN	1	4.09E+05	0.03
	sp|P02671|FIBA_HUMAN	1	3.15E+05	0.11
	sp|P02749|APOH_HUMAN	1	2.44E+05	0.04
	sp|P06396|GELS_HUMAN	1	5.07E+05	0.05
	sp|P20742|PZP_HUMAN	1	2.01E+05	0.03

Note 9.1 External fractionated plasma library generation details.

Sample. 1mg human serum (estimate 80mg/ml; Seralabs Human serum) was denatured with 6.4M Urea/10mM DTT/50mM Tris (pH 8) at 56°C for 45min and then alkylated with 25mM IAA for 30min in the dark at RT. Sample was diluted to 1ml with 20mM Tris (pH 8) and digested at 1:20 with trypsin (Promega sequencing grade) overnight at 30°C.

2D LC Fractionation. 100µg of total digest was fractionated using high pH reverse phase chromatography using a Shimadzu Nexera system (UV detection at 214nm). A Durashell RP column (250 x 4.6mm, 5µm, Agela Technologies) operating at 1 mL/min was used, running a gradient of 2 – 90% acetonitrile in 2mM ammonium hydroxide. 15 fractions were collected every 2 min and dried down.

1D LC-MS/MS Chromatography. Prior to the LC-MS injection, fractions were reconstituted into 40 µl 5% acetonitrile/0.1% FA with HRM peptides (Biognosys). 8 µl of each fraction (5% of digested fraction, ~3-5 µg on column) was then separated using low pH reverse phase gradient on the NanoLC™ 425 system operating in microflow mode. A Triart C18 150 x 0.3mm column (YMC) was used at 5 µL/min flow rate with a 45 min gradient from 2-40% acetonitrile in 0.1% formic acid for a total run time of 1 hour per fraction. The eluent was analysed using the TripleTOF® 6600 system equipped with a DuoSpray™ Source and a 25 µm ID hybrid electrode. Data dependent acquisition was performed with 30 MS/MS per cycle, each with 40 msec accumulation.

Data Processing. All DDA data was processed using ProteinPilot™ Software 5.0, all in one large search. Identification results were visualized from the full ProteinPilot reports that were automatically created.

Library generation. Library was generated with PeakView® 2.2 Software /SWATH® acquisition micro app 2.0 by importing the previously created ProteinPilot group file. All peptide and protein information at 1% global FDR level were imported and saved as text file.

Chapter 10:

References

10. References

[1] E. Von Behring, S. Kitasato, Uber das zustandekommen der diphtherie-immuni- tat und der tetanus-immunitat bei tieren, Dtsch Med Wochenschr. 16 (1890).

[2] M.M. Eibl, History of Immunoglobulin Replacement, Immunol. Allergy Clin. North Am. 28 (2008) 737–764. doi:10.1016/j.iac.2008.06.004.

[3] E.J. Cohn, L.E. Strong, W.L. Hughes, D.J. Mulford, J.N. Ashworth, M. Melin, H.L. Taylor, Preparation and Properties of Serum and Plasma Proteins. IV. A System for the Separation into Fractions of the Protein and Lipoprotein Components of Biological Tissues and Fluids, J. Am. Chem. Soc. 68 (1946) 459–475. doi:10.1021/ja01207a034.

[4] J.L. Oncley, M. Melin, D.A. Richert, J.W. Cameron, P.M. Gross, The Separation of the Antibodies, Isoagglutinins, Prothrombin, Plasminogen and β1 -Lipoprotein into Subfractions of Human Plasma, J. Am. Chem. Soc. 71 (1949) 541–550. doi:10.1021/ja01170a048.

[5] American Red Cross, History of Blood Transfusions, (n.d.). https://www.redcrossblood.org/learn-about-blood/blood-transfusions/history-blood-transfusions (accessed March 9, 2018).

[6] T. Burnouf, Modern Plasma Fractionation, Transfus. Med. Rev. 21 (2007) 101–117. doi:10.1016/j.tmrv.2006.11.001.

[7] E. Mathys, Plasma derivatives: from blood to medicinal products, (2009).

[8] G. Liumbruno, F. Bennardello, A. Lattanzio, P. Piccoli, G. Rossetti, as I.S. of T.M. and I. (SIMTI) W. Italian Society of Transfusion Medicine and Immunohaematology (SIMTI) Work Group, Recommendations for the transfusion of plasma and platelets., Blood Transfus. 7 (2009) 132–50. doi:10.2450/2009.0005-09.

[9] WHO, WHO | WHO Model Lists of Essential Medicines, World Heal. Organ. 19 (2015) 55. doi:10.1016/S1473-3099(14)70780-7.

[10] T. Burnouf, Recombinant plasma proteins, Vox Sang. 100 (2011) 68–83. doi:10.1111/j.1423-0410.2010.01384.x.

[11] H.A.D. Lagassé, A. Alexaki, V.L. Simhadri, N.H. Katagiri, W. Jankowski, Z.E. Sauna, C. Kimchi-Sarfaty, Recent advances in (therapeutic protein) drug development., F1000Research. 6 (2017) 113. doi:10.12688/f1000research.9970.1.

[12] S.J. Lee, J. Chinen, A. Kavanaugh, Immunomodulator therapy: Monoclonal antibodies, fusion proteins, cytokines, and immunoglobulins, J. Allergy Clin. Immunol. 125 (2010) S314–S323. doi:10.1016/j.jaci.2009.08.018.

[13] Creativ-Ceutical / EU Commission, An EU-wide overview of the market of blood, blood components and plasma derivatives focusing on their availability for patients, (2015). https://ec.europa.eu/health//sites/health/files/blood_tissues_organs/docs/20150408_cc_report_en.pdf.

[14] ThermoFischer Scientific, Immunoglobulin Structure and Classes, (n.d.). https://www.thermofisher.com/be/en/home/life-science/antibodies/antibodies-learning-

center/antibodies-resource-library/antibody-methods/immunoglobulin-structure-classes.html (accessed March 9, 2018).

[15] P.J. Späth, Structure and function of immunoglobulins, Sepsis. 3 (1999) 197–218. doi:10.1023/A:1009899803032.

[16] M. Radosevich, T. Burnouf, Intravenous immunoglobulin G: Trends in production methods, quality control and quality assurance, Vox Sang. 98 (2010) 12–28. doi:10.1111/j.1423-0410.2009.01226.x.

[17] G. Vidarsson, G. Dekkers, T. Rispens, IgG subclasses and allotypes: From structure to effector functions, Front. Immunol. 5 (2014) 520. doi:10.3389/fimmu.2014.00520.

[18] R.J. Looney, J. Huggins, Use of intravenous immunoglobulin G (IVIG)., Best Pract. Res. Clin. Haematol. 19 (2006) 3–25. doi:10.1016/j.beha.2005.01.032.

[19] M. Pučić, A. Knežević, J. Vidič, B. Adamczyk, M. Novokmet, O. Polašek, O. Gornik, S. Šupraha-Goreta, M.R. Wormald, I. Redžić, H. Campbell, A. Wright, N.D. Hastie, J.F. Wilson, I. Rudan, M. Wuhrer, P.M. Rudd, D. Josić, G. Lauc, High Throughput Isolation and Glycosylation Analysis of IgG–Variability and Heritability of the IgG Glycome in Three Isolated Human Populations, Mol. Cell. Proteomics. 10 (2011) M111.010090. doi:10.1074/mcp.M111.010090.

[20] A. Kobata, The N-linked sugar chains of human immunoglobulin G: their unique pattern, and their functional roles., Biochim. Biophys. Acta. 1780 (2008) 472–8. doi:10.1016/j.bbagen.2007.06.012.

[21] C. Janeway, P. Travers, W. M, Immunobiology: The Immune System in Health and Disease. 5th edition., Garland Sc, New York, 2001.

[22] O.C. Bruton, Agammaglobulinemia., Pediatrics. 9 (1952) 722–8. http://www.ncbi.nlm.nih.gov/pubmed/14929630 (accessed November 22, 2017).

[23] S. Barandun, P. Kistler, F. Jeunet, H. Isliker, Intravenous Administration of Human γ-Globulin, Vox Sang. 7 (1962) 157–174. doi:10.1111/j.1423-0410.1962.tb03240.x.

[24] G.N. Smith, B. Griffiths, D. Mollison, P.L. Mollison, Uptake of IgG after intramuscular and subcutaneous injection., Lancet (London, England). 1 (1972) 1208–12. http://www.ncbi.nlm.nih.gov/pubmed/4113191 (accessed November 22, 2017).

[25] A. Farrugia, I. Quinti, Manufacture of immunoglobulin products for patients with primary antibody deficiencies - the effect of processing conditions on product safety and efficacy, Front. Immunol. 5 (2014) 665. doi:10.3389/fimmu.2014.00665.

[26] L. Emmi, F. Chiarini, The role of intravenous immunoglobulin therapy in autoimmune and inflammatory disorders., Neurol. Sci. 23 Suppl 1 (2002) S1–8. doi:10.1007/s100720200010.

[27] A. Ephrem, N. Misra, G. Hassan, S. Dasgupta, S. Delignat, J.-P.D. Van Huyen, S. Chamat, F. Prost, S. Lacroix-Desmazes, S. V. Kavery, M.D. Kazatchkine, Immunomodulation of autoimmune and inflammatory diseases with intravenous immunoglobulin, Clin. Exp. Med. 5 (2005) 135–140. doi:10.1007/s10238-005-0079-y.

[28] I. Knezevic-Maramica, M.S. Kruskall, Intravenous immune globulins: An update for clinicians,

Transfusion. 43 (2003) 1460–1480. doi:10.1046/j.1537-2995.2003.00519.x.

[29] Conseil Supérieur de la Santé, Recommandations en matière d'indications pour l'administration des immunoglobulines., Publ. Du Cons. Supérieur La Santé N°8366. (2010). https://www.health.belgium.be/sites/default/files/uploads/fields/fpshealth_theme_file/19073133/Recommandations en matière d'indications pour l'administration des immunoglobulines (juin 2010) (CSS 8366).pdf.

[30] U. Katz, Y. Shoenfeld, G. Zandman-Goddard, Update on intravenous immunoglobulins (IVIg) mechanisms of action and off- label use in autoimmune diseases., Curr. Pharm. Des. 17 (2011) 3166–75. http://www.ncbi.nlm.nih.gov/pubmed/21864262 (accessed November 22, 2017).

[31] H. Leong, J. Stachnik, M.E. Bonk, K.A. Matuszewski, Unlabeled uses of intravenous immune globulin, Am. J. Heal. Pharm. 65 (2008) 1815–1824. doi:10.2146/ajhp070582.

[32] G. Köhler, C. Milstein, Continuous cultures of fused cells secreting antibody of predefined specificity, Nature. 256 (1975) 495–497. doi:10.1038/256495a0.

[33] E.R. Stiehm, M. a. Keller, G.N. Vyas, Preparation and use of therapeutic antibodies primarily of human origin, Biologicals. 36 (2008) 363–374. doi:10.1016/j.biologicals.2008.07.002.

[34] R.D. McBain, C.A. Crowther, P. Middleton, Anti-D administration in pregnancy for preventing Rhesus alloimmunisation, in: R.D. McBain (Ed.), Cochrane Database Syst. Rev., John Wiley & Sons, Ltd, Chichester, UK, 2015: p. CD000020. doi:10.1002/14651858.CD000020.pub3.

[35] X.-Z. Wang, V.W. Coljee, J.A. Maynard, Back to the future: recombinant polyclonal antibody therapeutics., Curr. Opin. Chem. Eng. 2 (2013) 405–415. doi:10.1016/j.coche.2013.08.005.

[36] Y. Camcioglu, Immunoglobulin Treatment of Immunodeficient Patients, in: K. Metodiev (Ed.), Immunodeficiency, InTech, 2012. doi:10.5772/51660.

[37] B. Chaigne, L. Mouthon, Mechanisms of action of intravenous immunoglobulin, Transfus. Apher. Sci. 56 (2017) 45–49. doi:10.1016/j.transci.2016.12.017.

[38] M.C.Z. Novaretti, C.L. Dinardo, Immunoglobulin: production, mechanisms of action and formulations., Rev. Bras. Hematol. Hemoter. 33 (2011) 377–82. doi:10.5581/1516-8484.20110102.

[39] M. Berger, M. Rojavin, P. Kiessling, O. Zenker, Pharmacokinetics of subcutaneous immunoglobulin and their use in dosing of replacement therapy in patients with primary immunodeficiencies, Clin. Immunol. 139 (2011) 133–141. doi:10.1016/j.clim.2011.01.006.

[40] M. Fadeyi, T. Tran, Calculating the dose of subcutaneous immunoglobulin for primary immunodeficiency disease in patients switched from intravenous to subcutaneous immunoglobulin without the use of a dose-adjustment coefficient., P T. 38 (2013) 768–70. http://www.ncbi.nlm.nih.gov/pubmed/24391400 (accessed April 11, 2018).

[41] L. Kobrynski, Subcutaneous immunoglobulin therapy: a new option for patients with primary immunodeficiency diseases., Biologics. 6 (2012) 277–87. doi:10.2147/BTT.S25188.

[42] E.W. Gelfand, Differences between IGIV products: Impact on clinical outcome, Int.

Immunopharmacol. 6 (2006) 592–599. doi:10.1016/j.intimp.2005.11.003.

[43] L. Gilardin, J. Bayry, S. V Kaveri, Intravenous immunoglobulin as clinical immune-modulating therapy., CMAJ. 187 (2015) 257–64. doi:10.1503/cmaj.130375.

[44] D.M. Ecker, S.D. Jones, H.L. Levine, The therapeutic monoclonal antibody market., MAbs. 7 (2015) 9–14. doi:10.4161/19420862.2015.989042.

[45] Ablynx, Understanding Nanobodies, (2018). http://www.ablynx.com/technology-innovation/understanding-nanobodies/ (accessed December 7, 2018).

[46] World Health Organization, Blood donor selection: Guidelines on assessing donor suitability for blood donation, World Health Organization, Geneva, 2012.

[47] Loi relative au sang et aux dérivés du sang d'origine humaine., Santé Publique et environnement, Belgium, 1994.

[48] World Health Organization, Blood transfusion safety - Voluntary non-remunerated blood donation, WHO. (2017). http://www.who.int/bloodsafety/voluntary_donation/en/ (accessed October 17, 2017).

[49] V. Salunkhe, P.F. van der Meer, D. de Korte, J. Seghatchian, L. Gutiérrez, Development of blood transfusion product pathogen reduction treatments: A review of methods, current applications and demands, Transfus. Apher. Sci. 52 (2015) 19–34. doi:10.1016/j.transci.2014.12.016.

[50] R. Laub, S. Baurin, D. Timmerman, T. Branckaert, P. Strengers, Specific protein content of pools of plasma for fractionation from different sources: Impact of frequency of donations, Vox Sang. 99 (2010) 220–231. doi:10.1111/j.1423-0410.2010.01345.x.

[51] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Monograph on human plasma for fractionation (0853), Eur. Pharmacopoeia Online. 9.2 (2014) 0853.

[52] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Monograph on human normal immunoglobulin for intravenous administration (0918), Eur. Pharmacopoeia Online. 9.2 (2012) 0918.

[53] P. Chérin, J. Cabane, Relevant criteria for selecting an intravenous immunoglobulin preparation for clinical use, BioDrugs. 24 (2010) 211–223. doi:10.2165/11537660-000000000-00000.

[54] M.R. Stein, The new generation of liquid intravenous immunoglobulin formulations in patient care: A comparison of intravenous immunoglobulins, Postgrad. Med. 122 (2010) 176–184. doi:10.3810/pgm.2010.09.2214.

[55] U. Katz, A. Achiron, Y. Sherer, Y. Shoenfeld, Safety of intravenous immunoglobulin (IVIG) therapy., Autoimmun. Rev. 6 (2007) 257–9. doi:10.1016/j.autrev.2006.08.011.

[56] Y. Leblanc, M. Romanin, N. Bihoreau, G. Chevreux, LC–MS analysis of polyclonal IgGs using IdeS enzymatic proteolysis for oxidation monitoring, J. Chromatogr. B. 961 (2014) 1–4. doi:10.1016/j.jchromb.2014.04.053.

[57] Y. Leblanc, N. Bihoreau, M. Jube, M.-H. Andre, Z. Tellier, G. Chevreux, Glycation of polyclonal

IgGs: Effect of sugar excipients during stability studies, Eur. J. Pharm. Biopharm. 102 (2016) 185–190. doi:10.1016/j.ejpb.2016.03.016.

[58] J. Carbone, Adverse reactions and pathogen safety of intravenous immunoglobulin., Curr. Drug Saf. 2 (2007) 9–18. http://www.ncbi.nlm.nih.gov/pubmed/18690945 (accessed April 24, 2015).

[59] J.B. Caress, B.L. Kennedy, K.D. Eickman, Safety of intravenous immunoglobulin treatment, Expert Opin. Drug Saf. 9 (2010) 971–979. doi:10.1517/14740338.2010.484419.

[60] N. Rezaei, H. Abolhassani, A. Aghamohammadi, H.D. Ochs, Indications and safety of intravenous and subcutaneous immunoglobulin therapy, Expert Rev. Clin. Immunol. 7 (2011) 301–316. doi:10.1586/eci.10.104.

[61] EDQM - The European Directorate for the Quality of Medicines & Health Care, Protocol for approval of plasma pools, EU Off. Control Auth. Batch Release Hum. Blood Deriv. Med. Prod. (2015).

[62] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Monograph on human normal immunoglobulin for intramuscular administration (0338), Eur. Pharmacopoeia Online. 9.2 (2015) 0338.

[63] European Pharmacopoeia Commission - Council of Europe (European Directorate for the Quality of Medicines), Monograph on human normal immunoglobulin for subcutaneous administration (2788), Eur. Pharmacopoeia Online. 9.2 (2015) 2788.

[64] EDQM - The European Directorate for the Quality of Medicines & Health Care, Guideline for human immunoglobulin., EU Off. Control Auth. Batch Release Hum. Blood Deriv. Med. Prod. (2015).

[65] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Determination of nitrogen by sulfuric acid digestion, Eur. Pharmacopoeia Online. 9.3 (2008) 2.5.9.

[66] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Test for anticomplementary activity of immunoglobulin, Eur. Pharmacopoeia Online. 9.3 (2018) 2.6.17.

[67] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Test for Fc function of immunoglobulin, Eur. Pharmacopoeia Online. 9.3 (2018) 2.7.9.

[68] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Size-Exclusion chromatography, Eur. Pharmacopoeia Online. 9.3 (2008) 2.2.30.

[69] E.S. Bergmann-Leitner, R.M. Mease, E.H. Duncan, F. Khan, J. Waitumbi, E. Angov, Evaluation of immunoglobulin purification methods and their impact on quality and yield of antigen-specific antibodies., Malar. J. 7 (2008) 129. doi:10.1186/1475-2875-7-129.

[70] C. Kempf, M. Stucki, N. Boschetti, Pathogen inactivation and removal procedures used in the production of intravenous immunoglobulins, Biologicals. 35 (2007) 35–42. doi:10.1016/j.biologicals.2006.01.002.

[71] T. Burnouf, A. Padilla, Current strategies to prevent transmission of prions by human plasma derivatives, Transfus. Clin. Biol. 13 (2006) 320–328. doi:10.1016/j.tracli.2006.11.001.

[72] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Immunochemical methods, Eur. Pharmacopoeia Online. 9.2 (2008) 2.7.1.

[73] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Monograph on human anti-D immunoglobulin, Eur. Pharmacopoeia Online. 9.3 (2015) 0557.

[74] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Monograph on human anti-D immunoglobulin for intravenous administration, Eur. Pharmacopoeia Online. 9.3 (2008) 1527.

[75] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Monograph on human hepatitis A immunoglobulin, Eur. Pharmacopoeia Online. 9.3 (2015) 0769.

[76] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Monograph on human hepatitis B immunoglobulin for intravenous administration, Eur. Pharmacopoeia Online. 9.3 (2015) 1016.

[77] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Monograph on human hepatitis B immunoglobulin, Eur. Pharmacopoeia Online. 9.3 (2015) 0722.

[78] European Pharmacopoeia Commission - Council of Europe (European Directorate for the Quality of Medicines), Monograph on human measles immunoglobulin, Eur. Pharmacopoeia Online. 9.3 (2015) 0397.

[79] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Monograph on human rabies immunoglobulin, Eur. Pharmacopoeia Online. 9.3 (2015) 0723.

[80] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Monograph on human rubella immunoglobulin, Eur. Pharmacopoeia Online. 9.3 (2015) 0617.

[81] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Monograph on human tetanus immunoglobulin, Eur. Pharmacopoeia Online. 9.3 (2015) 0398.

[82] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Monograph on human varicella immunoglobulin, Eur. Pharmacopoeia Online. 9.3 (2015) 0724.

[83] P.J. Späth, G. Granata, F. La Marra, T.W. Kuijpers, I. Quinti, On the dark side of therapies with immunoglobulin concentrates: The adverse events, Front. Immunol. 6 (2015) 11. doi:10.3389/fimmu.2015.00011.

[84] J. Kahwaji, E. Barker, S. Pepkowitz, E. Klapper, R. Villicana, A. Peng, R. Chang, S.C. Jordan, A.A. Vo,

Acute hemolysis after high-dose intravenous immunoglobulin therapy in highly HLA sensitized patients, Clin. J. Am. Soc. Nephrol. 4 (2009) 1993–1997. doi:10.2215/CJN.04540709.

[85] R.F. Padmore, Hemolysis upon intravenous immunoglobulin transfusion, Transfus. Apher. Sci. 46 (2012) 93–96. doi:10.1016/j.transci.2011.11.004.

[86] C.L. Bellac, D. Polatti, T. Hottiger, P. Girard, M. Sänger, M. Gilgen, Anti-A and anti-B haemagglutinin levels in intravenous immunoglobulins: Are they on the rise? A comparison of four different analysis methods and six products, Biologicals. 42 (2014) 57–64. doi:10.1016/j.biologicals.2013.10.004.

[87] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Anti-A and anti-B haemagglutinins, Eur. Pharmacopoeia Online. 9.3 (2011) 2.6.20.

[88] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Test for anti-D antibodies in human immunoglobulin, Eur. Pharmacopoeia Online. 9.3 (2011) 2.6.26.

[89] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Assay of human anti-D immunoglobulin, Eur. Pharmacopoeia Online. 9.3 (2008) 2.7.13.

[90] M.B. Funk, N. Gross, S. Gross, A. Hunfeld, A. Lohmann, S. Guenay, K.M. Hanschmann, B. Keller-Stanislawski, Thromboembolic events associated with immunoglobulin treatment, Vox Sang. 105 (2013) 54–64. doi:10.1111/vox.12025.

[91] M. Etscheid, S. Breitner-Ruddock, S. Gross, A. Hunfeld, R. Seitz, J. Dodt, Identification of kallikrein and FXIa as impurities in therapeutic immunoglobulins: implications for the safety and control of intravenous blood products., Vox Sang. 102 (2012) 40–46. doi:10.1111/j.1423-0410.2011.01502.x.

[92] A.S. Wolberg, R.H. Kon, D.M. Monroe, M. Hoffman, Coagulation factor XI is a contaminant in intravenous immunoglobulin preparations, Am. J. Hematol. 65 (2000) 30–34.

[93] F. Lackner, G. Beck, S. Eichmeir, M. Gemeiner, K. Hummel, D. Pullirsch, E. Razzazi-Fazeli, A. Seifner, I. Miller, Contamination of therapeutic human immunoglobulin preparations with apolipoprotein H (β2-glycoprotein I), Electrophoresis. 35 (2014) 515–521. doi:10.1002/elps.201300319.

[94] W.A. Germishuizen, D.C. Gyure, D. Stubbings, T. Burnouf, Quantifying the thrombogenic potential of human plasma-derived immunoglobulin products, Biologicals. 42 (2014) 260–270. doi:10.1016/j.biologicals.2014.04.002.

[95] E.M. Ammann, M.P. Jones, B.K. Link, R.M. Carnahan, S.K. Winiecki, J.C. Torner, B.D. McDowell, B.H. Fireman, E.A. Chrischilles, Intravenous immune globulin and thromboembolic adverse events in patients with hematologic malignancy, Blood. 127 (2016) 200–207. doi:10.1182/blood-2015-05-647552.

[96] EMA - European Medicines Agency, European Medicines Agency recommends suspension of Octagam in all EU Member States, 2010. http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/news/2010/09/news_

detail_001123.jsp&mid=WC0b01ac058004d5c1.

[97] FDA - U.S. Food and Drug Administration, Biologics Recalls. Voluntary Market Withdrawal - September 23, 2010 Octagam [Immune Globulin Intravenous (Human)] 5% Liquid Preparation, Biologics (2010) ucm227133. http://www.fda.gov/BiologicsBloodVaccines/SafetyAvailability/Recalls/ucm227133.htm.

[98] FDA - U.S. Food and Drug Administration, Biologics Recalls. Important Safety Information: Risk of Thrombotic Adverse Events with Subcutaneous or Inappropriate Intravenous Use of Vivaglobin (Immune Globulin Subcutaneous), Biologics (2011) ucm246863. http://www.fda.gov/BiologicsBloodVaccines/SafetyAvailability/ucm246863.htm.

[99] G.W. Daniel, M. Menis, G. Sridhar, D. Scott, A.E. Wallace, M. V. Ovanesov, B. Golding, S.A. Anderson, J. Epstein, D. Martin, R. Ball, H.S. Izurieta, Immune globulins and thrombotic adverse events as recorded in a large administrative database in 2008 through 2010, Transfusion. 52 (2012) 2113–2121. doi:10.1111/j.1537-2995.2012.03589.x.

[100] J.R. Roemisch, W. Kaar, A. Zoechling, C. Kannicht, Identification of Activated FXI as the Major Biochemical Root Cause in IVIG Batches Associated with Thromboembolic Events . Analytical and Experimental Approaches Resulting in Corrective and Preventive Measures Implemented into the Octagam ® Manufacturing, WebmedCentral Immunother. 2 (2011) WMC002002.

[101] P.H. Bolton-Maggs, Factor XI deficiency and its management., Haemophilia. 6 Suppl 1 (2000) 100–9. http://www.ncbi.nlm.nih.gov/pubmed/10982275 (accessed January 23, 2018).

[102] F. Bauduer, E. de Raucourt, C. Boyer-Neumann, M. Trossaert, P. Beurrier, A. Faradji, J. Peynet, J.-Y. Borg, P. Chamouni, C. Chatelanaz, C. Henriet, F. Bridey, J. Goudemand, French Postmarketing Study Group, Factor XI replacement for inherited factor XI deficiency in routine clinical practice: results of the HEMOLEVEN prospective 3-year postmarketing study., Haemophilia. 21 (2015) 481–9. doi:10.1111/hae.12655.

[103] M. Komenda, D. Stadler, T. Malinas, M. Moses, I. Pragst, E. Herzog, P. Schmutz, K. Minnig, I. El Menyawi, Assessment of the ability of the Privigen® purification process to deplete thrombogenic factor XIa from plasma, Vox Sang. 107 (2014) 26–36. doi:10.1111/vox.12119.

[104] M. José, N. Marzo, B. Pons, A. Herrerias, L. López, M. Faro, M. López, J.I. Jorquera, Flebogamma® DIF (intravenous immunoglobulin) purification process effectively eliminates procoagulant activities, Biologicals. 41 (2013) 393–399. doi:10.1016/j.biologicals.2013.08.002.

[105] R. Voges-Haas, V. Braun, R. Schmeidl, G. Kuchta, S. Kistner, Manufacturing process of Intratect efficaciously eliminates thrombogenic potential., WebmedCentral Immunother. 5 (2014) WMC004514.

[106] J. Roemisch, C. Zapfl, A. Zoechling, K. Pock, Comparison of Thrombin Generation Assay (TGA) and Non-Activated Partial Thromboplastin Time (NAPTT) for the Assessment of Enhanced Procoagulant Activity in Immunoglobulin Solutions, WebmedCentral Immunother. 3 (2012) 1–23. doi:10.9754/journal.wmc.2012.003247.

[107] E. Castoldi, J. Rosing, Thrombin generation tests, Thromb. Res. 127 (2011) S21–S25. doi:10.1016/S0049-3848(11)70007-X.

[108] EMA - European Medicines Agency, European Medicines Agency recommends lifting of suspension of Octagam, 2011.

[109] Omrix Biopharmaceuticals, Omr-IgG-am 5% Cleared to Return to Market, 2012. http://www.pharmaline.co.il/images/newsletterregistration/omrix/2012/27072012/omr_igg_aml_letter_e.pdf.

[110] M. V. Ovanesov, M.D. Menis, D.E. Scott, R. Forshee, S. Anderson, W. Bryan, B. Golding, Association of immune globulin intravenous and thromboembolic adverse events, Am. J. Hematol. 92 (2017) E44–E45. doi:10.1002/ajh.24644.

[111] U. Salge-Bartels, M. Heiden, N. Groß, R. Seitz, Significance of platelet and monocyte activation for therapeutic immunoglobulin-induced thromboembolism, Thromb. Res. 133 (2014) 244–253. doi:10.1016/j.thromres.2013.11.026.

[112] D. Gaso-Sokac, D. Josic, The role of proteomics in plasma fractionation and quality control of plasma-derived therapeutic proteins., Blood Transfus. 8 Suppl 3 (2010) s86–91. doi:10.2450/2010.014S.

[113] A. Buchacher, P. Schluga, J. Müllner, M. Schreiner, C. Kannicht, J. Weinberger, Anticomplementary activity of IVIG concentrates - Important assay parameters and impact of IgG polymers, Vox Sang. 98 (2010) e209–18. doi:10.1111/j.1423-0410.2009.01271.x.

[114] M. Page, C. Ling, P. Dilger, M. Bentley, T. Forsey, C. Longstaff, R. Thorpe, Fragmentation of therapeutic human immunoglobulin preparations., Vox Sang. 69 (1995) 183–94. http://www.ncbi.nlm.nih.gov/pubmed/8578729 (accessed January 30, 2018).

[115] G. Lemm, Composition and properties of IVIg preparations that affect tolerability and therapeutic efficacy., Neurology. 59 (2002) S28–32. http://www.ncbi.nlm.nih.gov/pubmed/12499468 (accessed January 24, 2018).

[116] F. Dhainaut, P.-O. Guillaumat, H. Dib, G. Perret, A. Sauger, C. de Coupade, M. Beaudet, M. Elzaabi, L. Mouthon, In vitro and in vivo properties differ among liquid intravenous immunoglobulin preparations., Vox Sang. 104 (2013) 115–26. doi:10.1111/j.1423-0410.2012.01648.x.

[117] S. Elluru, J.-P.D. Van Huyen, J. Bayry, S. Delignat, F. Prost, A. Ephrem, S. Siberil, N. Misra, S. Lacroix-Desmzes, M.D. Kazatchkine, S. V. Kaveri, Comparative study of the anti-inflammatory effect of two intravenous immunoglobulin preparations manufactured by different processes, Immunol. Lett. 107 (2006) 58–62. doi:10.1016/j.imlet.2006.07.009.

[118] K.G. Frey, K.A. Bishop-Lilly, Next-Generation Sequencing for Pathogen Detection and Identification, Methods Microbiol. 42 (2015) 525–554. doi:10.1016/BS.MIM.2015.06.004.

[119] T. Thiele, L. Steil, U. Völker, A. Greinacher, Proteomics of blood-based therapeutics: A promising tool for quality assurance in transfusion medicine, BioDrugs. 21 (2007) 179–193. doi:10.2165/00063030-200721030-00005.

[120] E. de Hoffmann, V. Stroobant, Mass Spectrometry: Principles and Applications, Third Edit, John Wiley & Sons, Ltd, Chichester, UK, 2007.

[121] U.S. Baghel, A. Singh, D. Singh, M. Sinha, Application of Mass Spectroscopy in Pharmaceutical and Biomedical Analysis, in: Spectrosc. Anal. - Dev. Appl., InTech, 2017. doi:10.5772/intechopen.70655.

[122] T.E. Angel, U.K. Aryal, S.M. Hengel, E.S. Baker, R.T. Kelly, E.W. Robinson, R.D. Smith, Mass spectrometry-based proteomics: existing capabilities and future directions, Chem. Soc. Rev. 41 (2012) 3912. doi:10.1039/c2cs15331a.

[123] G.L. Hortin, D. Sviridov, The dynamic range problem in the analysis of the plasma proteome, J. Proteomics. 73 (2010) 629–636. doi:10.1016/j.jprot.2009.07.001.

[124] P.E. Geyer, N.A. Kulak, G. Pichler, L.M. Holdt, D. Teupser, M. Mann, Plasma Proteome Profiling to Assess Human Health and Disease, Cell Syst. 2 (2016) 185–195. doi:10.1016/j.cels.2016.02.015.

[125] T. Farrah, E.W. Deutsch, G.S. Omenn, D.S. Campbell, Z. Sun, J.A. Bletz, P. Mallick, J.E. Katz, J. Malmström, R. Ossola, J.D. Watts, B. Lin, H. Zhang, R.L. Moritz, R. Aebersold, A High-Confidence Human Plasma Proteome Reference Set with Estimated Concentrations in PeptideAtlas, Mol. Cell. Proteomics. 10 (2011) M110.006353. doi:10.1074/mcp.M110.006353.

[126] N. Lion, J.-D. Tissot, Application of proteomics to hematology: the revolution is starting, Expert Rev. Proteomics. 5 (2008) 375–379. doi:10.1586/14789450.5.3.375.

[127] G. Liumbruno, G.M. D'Amici, G. Grazzini, L. Zolla, Transfusion medicine in the era of proteomics, J. Proteomics. 71 (2008) 34–45. doi:10.1016/j.jprot.2008.02.004.

[128] D. Gaso-Sokac, S. Kovac, J. Clifton, D. Josic, Therapeutic plasma proteins - application of proteomics in process optimization, validation, and analysis of the final product, Electrophoresis. 32 (2011) 1104–1117. doi:10.1002/elps.201000641.

[129] S. Kovac, X. Yang, F. Huang, D. Hixson, D. Josic, Proteomics as a tool for optimization of human plasma protein separation, J. Chromatogr. A. 1194 (2008) 38–47. doi:10.1016/j.chroma.2008.04.026.

[130] X. Yang, J. Clifton, F. Huang, S. Kovac, D.C. Hixson, D. Josic, Proteomic analysis for process development and control of therapeutic protein separation from human plasma, Electrophoresis. 30 (2009) 1185–1193. doi:10.1002/elps.200800501.

[131] F. Fortis, L. Guerrier, L.B. Areces, P. Antonioli, T. Hayes, K. Carrick, D. Hammond, E. Boschetti, P.G. Righetti, A new approach for the detection and identification of protein impurities using combinatorial solid phase ligand libraries., J. Proteome Res. 5 (2006) 2577–85. doi:10.1021/pr060090s.

[132] O. Belgacem, G. Stübiger, G. Allmaier, A. Buchacher, K. Pock, Isolation of esterified fatty acids bound to serum albumin purified from human plasma and characterised by MALDI mass spectrometry, Biologicals. 35 (2007) 43–49. doi:10.1016/j.biologicals.2006.01.004.

[133] U.M. Demelbauer, A. Plematl, L. Kremser, G. Allmaier, D. Josic, A. Rizzi, Characterization of glyco isoforms in plasmaderived human antithrombin by on-line capillary zone electrophoresis-electrospray ionization-quadrupole ion trap-mass spectrometry of the intact glycoproteins, Electrophoresis. 25 (2004) 2026–2032. doi:10.1002/elps.200305936.

[134] U.M. Demelbauer, A. Plematl, D. Josic, G. Allmaier, A. Rizzi, On the variation of glycosylation in human plasma derived antithrombin., J. Chromatogr. A. 1080 (2005) 15–21. http://www.ncbi.nlm.nih.gov/pubmed/16013610 (accessed January 31, 2018).

[135] A. Plematl, U.M. Demelbauer, D. Josic, A. Rizzi, Determination of the site-specific and isoform-specific glycosylation in human plasma-derived antithrombin by IEF and capillary HPLC-ESI-MS/MS, Proteomics. 5 (2005) 4025–4033. doi:10.1002/pmic.200401238.

[136] J.G. Clifton, F. Huang, S. Kovac, X. Yang, D.C. Hixson, D. Josic, Proteomic characterization of plasma-derived clotting factor VIII-von Willebrand factor concentrates, Electrophoresis. 30 (2009) 3636–3646. doi:10.1002/elps.200900270.

[137] A.M. Timperio, F. Gevi, G. Grazzini, S. Vaglio, L. Zolla, Comparison among plasma-derived clotting factor VIII by using monodimensional gel electrophoresis and mass spectrometry., Blood Transfus. 8 Suppl 3 (2010) s98–104. doi:10.2450/2010.016S.

[138] F. Basilico, I. Nardini, F. Mori, E. Brambilla, L. Benazzi, A. De Palma, E. Rosti, C. Farina, P. Mauri, Characterization of factor VIII pharmaceutical preparations by means of MudPIT proteomic approach., J. Pharm. Biomed. Anal. 53 (2010) 50–7. doi:10.1016/j.jpba.2010.03.027.

[139] G.M. D'Amici, B. Blasi, A. D'Alessandro, S. Vaglio, L. Zolla, Plasma-derived clotting factor VIII: Heterogeneity evaluation in the quest for potential inhibitory-antibody stimulating factors, Electrophoresis. 32 (2011) 2941–2950. doi:10.1002/elps.201100246.

[140] D. Josic, C. Kannicht, K. Löster, K. Pock, G. Iberer, A. Buchacher, Vitronectin in clotting factor IX concentrates., Haemophilia. 7 (2001) 250–7. http://www.ncbi.nlm.nih.gov/pubmed/11380628 (accessed January 31, 2018).

[141] J. Clifton, F. Huang, D. Gaso-Sokac, K. Brilliant, D. Hixson, D. Josic, Use of proteomics for validation of the isolation process of clotting factor IX from human plasma, J. Proteomics. 73 (2010) 678–688. doi:10.1016/j.jprot.2009.09.020.

[142] R. Russo, G. Focà, C. Rega, A. Sandomenico, N. Doti, F. Mori, M. Maddaluno, C. Farina, M. Ruvo, A. Chambery, A multianalytical approach to investigate the effect of nanofiltration on plasma-derived factor IX clinical lots, Anal. Biochem. 542 (2018) 1–10. doi:10.1016/j.ab.2017.11.006.

[143] M. Brigulla, T. Thiele, C. Scharf, S. Breitner-Ruddock, S. Venz, U. Völker, A. Greinacher, Proteomics as a tool for assessment of therapeutics in transfusion medicine: Evaluation of prothrombin complex concentrates, Transfusion. 46 (2006) 377–385. doi:10.1111/j.1537-2995.2006.00732.x.

[144] D. Kolarich, P.L. Turecek, A. Weber, A. Mitterer, M. Graninger, P. Matthiessen, G.A.F. Nicolaes, F. Altmann, H.P. Schwarz, Biochemical, molecular characterization, and glycoproteomic analyses of α1-proteinase inhibitor products used for replacement therapy, in: Transfusion, 2006: pp. 1959–1977. doi:10.1111/j.1537-2995.2006.01004.x.

[145] C. Mizon, A. Héron, C. Capon, P. Sautiere, C. Michalski, R. Sesboüé, J. Mizon, Human pre-alpha-inhibitor: isolation from a by-product of industrial scale plasma fractionation and structural analysis of its H3 heavy chain., J. Chromatogr. B. Biomed. Sci. Appl. 692 (1997) 281–91. http://www.ncbi.nlm.nih.gov/pubmed/9188816 (accessed January 31, 2018).

[146] D. Josic, M.K. Brown, F. Huang, Y.-P. Lim, M. Rucevic, J.G. Clifton, D.C. Hixson, Proteomic

characterization of inter-alpha inhibitor proteins from human plasma, Proteomics. 6 (2006) 2874–2885. doi:10.1002/pmic.200500563.

[147] G. Bussone, H. Dib, J.D. Dimitrov, L. Camoin, C. Broussard, N. Tamas, L. Guillevin, S. V. Kaveri, L. Mouthon, Identification of target antigens of self-reactive IgG in intravenous immunoglobulin preparations, Proteomics. 9 (2009) 2253–2262. doi:10.1002/pmic.200800819.

[148] Y. Kaneko, F. Nimmerjahn, J. V Ravetch, Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation., Science. 313 (2006) 670–3. doi:10.1126/science.1129594.

[149] I. Ritamo, J. Räbinä, S. Natunen, L. Valmu, Nanoscale reversed-phase liquid chromatography-mass spectrometry of permethylated N-glycans., Anal. Bioanal. Chem. 405 (2013) 2469–80. doi:10.1007/s00216-012-6680-5.

[150] I. Ritamo, M. Cloutier, L. Valmu, S. Néron, J. Räbinä, Comparison of the glycosylation of in vitro generated polyclonal human IgG and therapeutic immunoglobulins, Mol. Immunol. 57 (2014) 255–262. doi:10.1016/j.molimm.2013.10.005.

[151] W.J.R. Fokkink, D. Falck, T.C.M. Santbergen, R. Huizinga, M. Wuhrer, B.C. Jacobs, Comparison of Fc N-Glycosylation of Pharmaceutical Products of Intravenous Immunoglobulin G., PLoS One. 10 (2015) e0139828. doi:10.1371/journal.pone.0139828.

[152] A. Feussner, U. Kalina, P. Hofmann, T. MacHnig, G. Henkel, Biochemical comparison of four commercially available C1 esterase inhibitor concentrates for treatment of hereditary angioedema, Transfusion. 54 (2014) 2566–2573. doi:10.1111/trf.12678.

[153] G. Chevreux, N. Tilly, V. Faid, N. Bihoreau, Mass spectrometry based analysis of human plasma-derived factor X revealed novel post-translational modifications, Protein Sci. 24 (2015) 1640–1648. doi:10.1002/pro.2756.

[154] D. Kolarich, A. Weber, M. Pabst, J. Stadlmann, W. Teschner, H. Ehrlich, H.-P. Schwarz, F. Altmann, Glycoproteomic characterization of butyrylcholinesterase from human plasma., Proteomics. 8 (2008) 254–63. doi:10.1002/pmic.200700720.

[155] A. Weber, H. Butterweck, U. Mais-Paul, W. Teschner, L. Lei, E.-M. Muchitsch, D. Kolarich, F. Altmann, H.J. Ehrlich, H.P. Schwarz, Biochemical, molecular and preclinical characterization of a double-virus-reduced human butyrylcholinesterase preparation designed for clinical use., Vox Sang. 100 (2011) 285–97. doi:10.1111/j.1423-0410.2010.01415.x.

[156] M. Wuhrer, J.C. Stam, F.E. van de Geijn, C.A.M. Koeleman, C.T. Verrips, R.J.E.M. Dolhain, C.H. Hokke, A.M. Deelder, Glycosylation profiling of immunoglobulin G (IgG) subclasses from human serum., Proteomics. 7 (2007) 4070–81. doi:10.1002/pmic.200700289.

[157] C. Huhn, M.H.J. Selman, L.R. Ruhaak, A.M. Deelder, M. Wuhrer, IgG glycosylation analysis., Proteomics. 9 (2009) 882–913. doi:10.1002/pmic.200800715.

[158] A.M. Goetze, Z. Zhang, L. Liu, F.W. Jacobsen, G.C. Flynn, Rapid LC–MS screening for IgG Fc modifications and allelic variants in blood, Mol. Immunol. 49 (2011) 338–352. doi:10.1016/j.molimm.2011.09.002.

[159] J.E. Huffman, M. Pučić-Baković, L. Klarić, R. Hennig, M.H.J. Selman, F. Vučković, M. Novokmet, J.

Krištić, M. Borowiak, T. Muth, O. Polašek, G. Razdorov, O. Gornik, R. Plomp, E. Theodoratou, A.F. Wright, I. Rudan, C. Hayward, H. Campbell, A.M. Deelder, U. Reichl, Y.S. Aulchenko, E. Rapp, M. Wuhrer, G. Lauc, Comparative performance of four methods for high-throughput glycosylation analysis of immunoglobulin G in genetic and epidemiological research., Mol. Cell. Proteomics. 13 (2014) 1598–610. doi:10.1074/mcp.M113.037465.

[160] G. Dekkers, L. Treffers, R. Plomp, A.E.H. Bentlage, M. de Boer, C.A.M. Koeleman, S.N. Lissenberg-Thunnissen, R. Visser, M. Brouwer, J.Y. Mok, H. Matlung, T.K. van den Berg, W.J.E. van Esch, T.W. Kuijpers, D. Wouters, T. Rispens, M. Wuhrer, G. Vidarsson, Decoding the Human Immunoglobulin G-Glycan Repertoire Reveals a Spectrum of Fc-Receptor- and Complement-Mediated-Effector Activities., Front. Immunol. 8 (2017) 877. doi:10.3389/fimmu.2017.00877.

[161] D. Reusch, M. Haberger, T. Kailich, A.-K. Heidenreich, M. Kampe, P. Bulau, M. Wuhrer, High-throughput glycosylation analysis of therapeutic immunoglobulin G by capillary gel electrophoresis using a DNA analyzer., MAbs. 6 (2014) 185–96. doi:10.4161/mabs.26712.

[162] L. Zhang, S. Luo, B. Zhang, Glycan analysis of therapeutic glycoproteins., MAbs. 8 (2016) 205–15. doi:10.1080/19420862.2015.1117719.

[163] S. Fischer, J. Hoernschemeyer, H.C. Mahler, Glycation during storage and administration of monoclonal antibody formulations, Eur. J. Pharm. Biopharm. 70 (2008) 42–50. doi:10.1016/j.ejpb.2008.04.021.

[164] C. Quan, E. Alcala, I. Petkovska, D. Matthews, E. Canova-Davis, R. Taticek, S. Ma, A study in glycation of a therapeutic recombinant humanized monoclonal antibody: Where it is, how it got there, and how it affects charge-based behavior, Anal. Biochem. 373 (2008) 179–191. doi:10.1016/j.ab.2007.09.027.

[165] G. Dekkers, R. Plomp, C.A.M. Koeleman, R. Visser, H.H. Von Horsten, V. Sandig, T. Rispens, M. Wuhrer, G. Vidarsson, Multi-level glyco-engineering techniques to generate IgG with defined Fc-glycans, Sci. Rep. 6 (2016) 36964. doi:10.1038/srep36964.

[166] N. Bihoreau, J.F. Veillon, C. Ramon, J.M. Scohyers, J.M. Schmitter, Characterization of a recombinant antihaemophilia-A factor (factor VIII-delta II) by matrix-assisted laser desorption/ionization mass spectrometry., Rapid Commun. Mass Spectrom. 9 (1995) 1584–8. doi:10.1002/rcm.1290091524.

[167] K.F. Medzihradszky, M.J. Besman, A.L. Burlingame, Structural characterization of site-specific N-glycosylation of recombinant human factor VIII by reversed-phase high-performance liquid chromatography-electrospray ionization mass spectrometry., Anal. Chem. 69 (1997) 3986–94. http://www.ncbi.nlm.nih.gov/pubmed/9322435 (accessed February 21, 2018).

[168] C.E. Doneanu, W. Chen, Analysis of host-cell proteins in biotherapeutic proteins by LC/MS approaches, Methods Mol. Biol. 1129 (2014) 341–350. doi:10.1007/978-1-62703-977-2_25.

[169] J.A. Madsen, V. Farutin, T. Carbeau, S. Wudyka, Y. Yin, S. Smith, J. Anderson, I. Capila, Toward the complete characterization of host cell proteins in biotherapeutics via affinity depletions, LC-MS/MS, and multivariate analysis, MAbs. 7 (2015) 1128–1137. doi:10.1080/19420862.2015.1082017.

[170] A. Farrell, S. Mittermayr, B. Morrissey, N. Mc Loughlin, N. Navas Iglesias, I.W. Marison, J. Bones,

Quantitative Host Cell Protein Analysis Using Two Dimensional Data Independent LC–MS E, Anal. Chem. 87 (2015) 9186–9193. doi:10.1021/acs.analchem.5b01377.

[171] D.G. Bracewell, R. Francis, C.M. Smales, The future of host cell protein (HCP) identification during process development and manufacturing linked to a risk-based management for their control, Biotechnol. Bioeng. 112 (2015) 1727–1737. doi:10.1002/bit.25628.

[172] N.E. Levy, K.N. Valente, L.H. Choe, K.H. Lee, A.M. Lenhoff, Identification and characterization of host cell protein product-associated impurities in monoclonal antibody bioprocessing, Biotechnol. Bioeng. 111 (2014) 904–912. doi:10.1002/bit.25158.

[173] G. Zhu, L. Sun, J. Heidbrink-Thompson, S. Kuntumalla, H.Y. Lin, C.J. Larkin, J.B. Mcgivney, N.J. Dovichi, Capillary zone electrophoresis tandem mass spectrometry detects low concentration host cell impurities in monoclonal antibodies, Electrophoresis. 37 (2016) 616–622. doi:10.1002/elps.201500301.

[174] G.M. D'Amici, A.M. Timperio, F. Gevi, G. Grazzini, L. Zolla, Recombinant clotting factor VIII concentrates: Heterogeneity and high-purity evaluation, Electrophoresis. 31 (2010) 2730–2739. doi:10.1002/elps.201000216.

[175] J. Kemptner, M. Marchetti-Deschmann, R. Müller, A. Ivens, P. Turecek, H.P. Schwarz, G. Allmaier, A comparison of nano-electrospray gas-phase electrophoretic mobility macromolecular analysis and matrix-assisted laser desorption/ionization linear time-of-flight mass spectrometry for the characterization of the recombinant coagulation glycoprotein von Willebrand factor., Rapid Commun. Mass Spectrom. 24 (2010) 761–7. doi:10.1002/rcm.4440.

[176] K. Sandra, I. Vandenheede, P. Sandra, Modern chromatographic and mass spectrometric techniques for protein biopharmaceutical characterization, J. Chromatogr. A. 1335 (2014) 81–103. doi:10.1016/j.chroma.2013.11.057.

[177] L. Chen, L. Wang, H. Shion, C. Yu, Y.Q. Yu, L. Zhu, M. Li, W. Chen, K. Gao, In-depth structural characterization of Kadcyla® (ado-trastuzumab emtansine) and its biosimilar candidate, MAbs. 8 (2016) 1210–1223. doi:10.1080/19420862.2016.1204502.

[178] M. Haberger, M. Leiss, A.-K. Heidenreich, O. Pester, G. Hafenmair, M. Hook, L. Bonnington, H. Wegele, M. Haindl, D. Reusch, P. Bulau, Rapid characterization of biotherapeutic proteins by size-exclusion chromatography coupled to native mass spectrometry, MAbs. 8 (2016) 331–339. doi:10.1080/19420862.2015.1122150.

[179] Y. Wang, X. Li, Y.-H. Liu, D. Richardson, H. Li, M. Shameem, X. Yang, Simultaneous monitoring of oxidation, deamidation, isomerization, and glycosylation of monoclonal antibodies by liquid chromatography-mass spectrometry method with ultrafast tryptic digestion, MAbs. 8 (2016) 1477–1486. doi:10.1080/19420862.2016.1226715.

[180] R. O'Flaherty, I. Trbojević-Akmačić, G. Greville, P.M. Rudd, G. Lauc, The sweet spot for biologics: recent advances in characterization of biotherapeutic glycoproteins., Expert Rev. Proteomics. 15 (2018) 13–29. doi:10.1080/14789450.2018.1404907.

[181] O. Montacir, H. Montacir, M. Eravci, A. Springer, S. Hinderlich, F. Mahboudi, A. Saadati, M. Parr, Bioengineering of rFVIIa Biopharmaceutical and Structure Characterization for Biosimilarity Assessment, Bioengineering. 5 (2018) 7. doi:10.3390/bioengineering5010007.

[182] N. Iwamoto, T. Shimada, Recent advances in mass spectrometry-based approaches for proteomics and biologics: Great contribution for developing therapeutic antibodies, Pharmacol. Ther. (2017). doi:10.1016/J.PHARMTHERA.2017.12.007.

[183] A. Ortiz, L. Richa, C. Defer, D. Dernis, J.J. Huart, C. Tokarski, C. Rolando, Proteomics applied to transfusion plasma: The beginning of the story, Vox Sang. 104 (2013) 275–291. doi:10.1111/j.1423-0410.2012.01663.x.

[184] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Human coagulation factor VIII (rDNA), Eur. Pharmacopoeia Online. 9.5 (2008) 1643.

[185] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Glycan analysis of glycoproteins, Eur. Pharmacopoeia Online. 9.5 (2011) 2.2.59.

[186] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Human α-1-proteinase inhibitor, Eur. Pharmacopoeia Online. 9.5 (2014) 2387.

[187] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Host-cell protein assays, Eur. Pharmacopoeia Online. 9.5 (2017) 2.6.34.

[188] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Substitution of in vivo method(s) by in vitro method(s) for the quality control of vaccines, Eur. Pharmacopoeia Online. 9.5 (2018) 5.2.14.

[189] European Pharmacopoeia Commission – Council of Europe (European Department for the Quality of Medicines), Mass Spectrometry, Eur. Pharmacopoeia Online. 9.5 (2008) 2.2.43.

[190] EDQM - The European Directorate for the Quality of Medicines & Health Care, Interpretation of screening results for unknown peptides and proteins by mass spectrometry based methods, Gen. Eur. OMCL Netw. Qual. Manag. Doc. (2016) PA/PH/OMCL (15) 04 2R.

[191] M. Pernemalm, R. Lewensohn, J. Lehtiö, Affinity prefractionation for MS-based plasma proteomics., Proteomics. 9 (2009) 1420–7. doi:10.1002/pmic.200800377.

[192] M. Pernemalm, J. Lehtiö, Mass spectrometry-based plasma proteomics: state of the art and future outlook, Expert Rev. Proteomics. 11 (2014) 431–448. doi:10.1586/14789450.2014.901157.

[193] M. Kullolli, J. Warren, M. Arampatzidou, S.J. Pitteri, Performance evaluation of affinity ligands for depletion of abundant plasma proteins., J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci. 939 (2013) 10–6. doi:10.1016/j.jchromb.2013.09.008.

[194] L. Björck, G. Kronvall, Purification and some properties of streptococcal protein G, a novel IgG-binding reagent., J. Immunol. 133 (1984) 969–74. http://www.jimmunol.org/content/133/2/969.abstract (accessed September 1, 2016).

[195] M. Eliasson, A. Olsson, E. Palmcrantz, K. Wiberg, M. Inganas, B. Guss, M. Lindberg, M. Uhlen, Chimeric IgG-binding receptors engineered from staphylococcal protein A and streptococcal protein G, J. Biol. Chem. 263 (1988) 4323–4327. http://www.jbc.org/content/263/9/4323.short (accessed September 1, 2016).

[196] B. Nilsson, T. Moks, B. Jansson, L. Abrahmsén, A. Elmblad, E. Holmgren, C. Henrichson, T.A. Jones, M. Uhlén, A synthetic IgG-binding domain based on staphylococcal protein A, "Protein Eng. Des. Sel. 1 (1987) 107–113. doi:10.1093/protein/1.2.107.

[197] O.P. Dancette, J.L. Taboureau, E. Tournier, C. Charcosset, P. Blond, Purification of immunoglobulins G by protein A/G affinity membrane chromatography., J. Chromatogr. B. Biomed. Sci. Appl. 723 (1999) 61–8. http://www.ncbi.nlm.nih.gov/pubmed/10080633 (accessed September 1, 2016).

[198] B.V. Ayyar, S. Arora, C. Murphy, R. O'Kennedy, Affinity chromatography as a tool for antibody purification., Methods. 56 (2012) 116–29. doi:10.1016/j.ymeth.2011.10.007.

[199] S. Sheng, F. Kong, Separation of antigens and antibodies by immunoaffinity chromatography., Pharm. Biol. 50 (2012) 1038–44. doi:10.3109/13880209.2011.653493.

[200] R.M. Hnasko, J.A. McGarvey, Affinity Purification of Antibodies., Methods Mol. Biol. 1318 (2015) 29–41. doi:10.1007/978-1-4939-2742-5_3.

[201] N. Zolotarjova, J. Martosella, G. Nicol, J. Bailey, B.E. Boyes, W.C. Barrett, Differences among techniques for high-abundant protein depletion., Proteomics. 5 (2005) 3304–13. doi:10.1002/pmic.200402021.

[202] L.A. Echan, H.-Y. Tang, N. Ali-Khan, K. Lee, D.W. Speicher, Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma., Proteomics. 5 (2005) 3292–303. doi:10.1002/pmic.200401228.

[203] C. Tu, P.A. Rudnick, M.Y. Martinez, K.L. Cheek, S.E. Stein, R.J.C. Slebos, D.C. Liebler, Depletion of abundant plasma proteins and limitations of plasma proteomics., J. Proteome Res. 9 (2010) 4982–91. doi:10.1021/pr100646w.

[204] T. Ichibangase, K. Moriya, K. Koike, K. Imai, Limitation of immunoaffinity column for the removal of abundant proteins from plasma in quantitative plasma proteomics, Biomed. Chromatogr. 23 (2009) 480–487. doi:10.1002/bmc.1139.

[205] V. Thulasiraman, S. Lin, L. Gheorghiu, J. Lathrop, L. Lomas, D. Hammond, E. Boschetti, Reduction of the concentration difference of proteins in biological liquids using a library of combinatorial ligands., Electrophoresis. 26 (2005) 3561–71. doi:10.1002/elps.200500147.

[206] P.G. Righetti, E. Boschetti, L. Lomas, A. Citterio, Protein Equalizer Technology : the quest for a "democratic proteome"., Proteomics. 6 (2006) 3980–92. doi:10.1002/pmic.200500904.

[207] L. Guerrier, V. Thulasiraman, A. Castagna, F. Fortis, S. Lin, L. Lomas, P.G. Righetti, E. Boschetti, Reducing protein concentration range of biological samples using solid-phase ligand libraries., J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci. 833 (2006) 33–40. doi:10.1016/j.jchromb.2005.12.048.

[208] E. Boschetti, L. Lomas, A. Citterio, P.G. Righetti, Romancing the "hidden proteome", Anno Domini two zero zero seven., J. Chromatogr. A. 1153 (2007) 277–90. doi:10.1016/j.chroma.2007.01.136.

[209] A. Castagna, D. Cecconi, L. Sennels, J. Rappsilber, L. Guerrier, F. Fortis, E. Boschetti, L. Lomas, P.G. Righetti, Exploring the hidden human urinary proteome via ligand library beads., J. Proteome Res.

4 (2005) 1917–30. doi:10.1021/pr050153r.

[210] L. Sennels, M. Salek, L. Lomas, E. Boschetti, P.G. Righetti, J. Rappsilber, Proteomic analysis of human blood serum using peptide library beads., J. Proteome Res. 6 (2007) 4055–62. doi:10.1021/pr070339l.

[211] A. Altomare, E. Fasoli, M. Colzani, X.M.P. Parra, M. Ferrari, F. Cilurzo, C. Rumio, L. Cannizzaro, M. Carini, P.G. Righetti, G. Aldini, An in depth proteomic analysis based on ProteoMiner, affinity chromatography and nano-HPLC-MS/MS to explain the potential health benefits of bovine colostrum, J. Pharm. Biomed. Anal. (2016). doi:10.1016/j.jpba.2016.01.013.

[212] J. Matysiak, J. Hajduk, F. Mayer, R. Hebeler, Z.J. Kokot, Hyphenated LC–MALDI–ToF/ToF and LC–ESI–QToF approach in proteomic characterization of honeybee venom, J. Pharm. Biomed. Anal. 121 (2016) 69–76. doi:10.1016/j.jpba.2016.01.009.

[213] P. Antonioli, F. Fortis, L. Guerrier, S. Rinalducci, L. Zolla, P.G. Righetti, E. Boschetti, Capturing and amplifying impurities from purified recombinant monoclonal antibodies via peptide library beads: a proteomic study., Proteomics. 7 (2007) 1624–33. doi:10.1002/pmic.200600778.

[214] E. Boschetti, P.G. Righetti, Low-abundance protein discovery: State of the Art and protocols, Elsevier, Massachusetts, 2013.

[215] E. Boschetti, P.G. Righetti, Optimized sample treatment protocol by solid-phase peptide libraries to enrich for protein traces, Amino Acids. 45 (2013) 1431–1442. doi:10.1007/s00726-013-1596-8.

[216] Bio-Rad, ProteoMiner Protein Enrichment Kits | Life Science Research | Bio-Rad, (n.d.). http://www.bio-rad.com/en-be/product/proteominer-protein-enrichment-kits?ID=1dd94f06-7658-4ab4-b844-e29a1342a214 (accessed March 9, 2018).

[217] L. Anderson, N.G. Anderson, High resolution two-dimensional electrophoresis of human plasma proteins., Proc. Natl. Acad. Sci. U. S. A. 74 (1977) 5421–5. doi:VL - 74.

[218] T. Rabilloud, A.R. Vaezzadeh, N. Potier, C. Lelong, E. Leize-Wagner, M. Chevallet, Power and limitations of electrophoretic separations in proteomics strategies, Mass Spectrom. Rev. 28 (2009) 816–843. doi:10.1002/mas.20204.

[219] K. Rose, L. Bougueleret, T. Baussant, G. Böhm, P. Botti, J. Colinge, I. Cusin, H. Gaertner, A. Gleizes, M. Heller, S. Jimenez, A. Johnson, M. Kussmann, L. Menin, C. Menzel, F. Ranno, P. Rodriguez-Tomé, J. Rogers, C. Saudrais, M. Villain, D. Wetmore, A. Bairoch, D. Hochstrasser, Industrial-scale proteomics: from liters of plasma to chemically synthesized proteins., Proteomics. 4 (2004) 2125–50. doi:10.1002/pmic.200300718.

[220] A. D'Alessandro, L. Zolla, Proteomics for quality-control processes in transfusion medicine, Anal. Bioanal. Chem. 398 (2010) 111–124. doi:10.1007/s00216-010-3799-0.

[221] L. Switzar, M. Giera, W.M.A. Niessen, Protein Digestion: An Overview of the Available Techniques and Recent Developments, J. Proteome Res. 12 (2013) 1067–1077. doi:10.1021/pr301201x.

[222] S. Saveliev, M. Bratz, R. Zubarev, M. Szapacs, H. Budamgunta, M. Urh, Trypsin/Lys-C protease mix for enhanced protein mass spectrometry analysis, Nat. Methods. 10 (2013) i–ii. doi:10.1038/nmeth.f.371.

[223] J. Torres, The Basics of Running a Chromatography Column - Bitesize Bio, (n.d.). https://bitesizebio.com/29947/basics-chromatography-column/ (accessed March 9, 2018).

[224] M. Olkowicz, I. Rybakowska, S. Chlopicki, R.T. Smolenski, Development and analytical comparison of microflow and nanoflow liquid chromatography/mass spectrometry procedures for quantification of cardiac troponin T in mouse hearts., Talanta. 131 (2015) 510–20. doi:10.1016/j.talanta.2014.08.029.

[225] A. Uclés Moreno, S. Herrera López, B. Reichert, A. Lozano Fernández, M.D. Hernando Guil, A.R. Fernández-Alba, Microflow liquid chromatography coupled to mass spectrometry--an approach to significantly increase sensitivity, decrease matrix effects, and reduce organic solvent usage in pesticide residue analysis., Anal. Chem. 87 (2015) 1018–25. doi:10.1021/ac5035852.

[226] C. Hunter, X. Zhu, R. Van Soest, Exploring the Sensitivity Differences for Targeted Peptide Quantification in the Low Flow Rate Regime, SCIEX Tech. Note. RUO-MKT-02 (2013) 2013.

[227] C.L. Hunter, N. Morrice, Microflow SWATH ® Acquisition for Industrialized Quantitative Proteomics, SCIEX Tech. Note. RUO-MKT-02 (2016) 1–5.

[228] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules., Science. 246 (1989) 64–71. http://www.ncbi.nlm.nih.gov/pubmed/2675315 (accessed March 2, 2018).

[229] K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, T. Matsuo, Protein and polymer analyses up tom/z 100 000 by laser ionization time-of-flight mass spectrometry, Rapid Commun. Mass Spectrom. 2 (1988) 151–153. doi:10.1002/rcm.1290020802.

[230] S. Banerjee, S. Mazumdar, Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte, Int. J. Anal. Chem. 2012 (2012) 1–40. doi:10.1155/2012/282574.

[231] A.M. Hawkridge, Practical considerations and current limitations in quantitative mass-spectrometry-based proteomics, in: C. Eyers, S.J. Gaskell (Eds.), Quant. Proteomics, The Royal Society of Chemistry, 2014: pp. 3–25.

[232] G.L. Andrews, B.L. Simons, J.B. Young, A.M. Hawkridge, D.C. Muddiman, Performance characteristics of a new hybrid quadrupole time-of-flight tandem mass spectrometer (TripleTOF 5600), Anal. Chem. 83 (2011) 5442–5446. doi:10.1021/ac200812d.

[233] H. Steen, M. Mann, The abc's (and xyz's) of peptide sequencing, Nat. Rev. Mol. Cell Biol. 5 (2004) 699–711. doi:10.1038/nrm1468.

[234] H. Lodisch, A. Berk, L. Zipursky, P. Matsudaira, D. Baltimore, J. Darnell, Molecular Cell Biology., 6th Editio, W.H. Freeman, New York, 2008. doi:10.1016/S1470-8175(01)00023-6.

[235] K. Biemann, Contributions of mass spectrometry to peptide and protein structure, Biol. Mass Spectrom. 16 (1988) 99–111. doi:10.1002/bms.1200160119.

[236] A. Doerr, DIA mass spectrometry, Nat. Methods. 12 (2014) 35–35. doi:10.1038/nmeth.3234.

[237] D.L. Tabb, L. Vega-Montoto, P.A. Rudnick, A.M. Variyath, A.-J.L. Ham, D.M. Bunk, L.E. Kilpatrick,

D.D. Billheimer, R.K. Blackman, H.L. Cardasis, S.A. Carr, K.R. Clauser, J.D. Jaffe, K.A. Kowalski, T.A. Neubert, F.E. Regnier, B. Schilling, T.J. Tegeler, M. Wang, P. Wang, J.R. Whiteaker, L.J. Zimmerman, S.J. Fisher, B.W. Gibson, C.R. Kinsinger, M. Mesri, H. Rodriguez, S.E. Stein, P. Tempst, A.G. Paulovich, D.C. Liebler, C. Spiegelman, Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry., J. Proteome Res. 9 (2010) 761–76. doi:10.1021/pr9006365.

[238] H. Liu, R.G. Sadygov, J.R. Yates, A model for random sampling and estimation of relative protein abundance in shotgun proteomics., Anal. Chem. 76 (2004) 4193–201. doi:10.1021/ac0498563.

[239] T. Kislinger, A.O. Gramolini, D.H. MacLennan, A. Emili, Multidimensional protein identification technology (MudPIT): technical overview of a profiling method optimized for the comprehensive proteomic investigation of normal and diseased heart tissue., J. Am. Soc. Mass Spectrom. 16 (2005) 1207–20. doi:10.1016/j.jasms.2005.02.015.

[240] R.J.C. Slebos, J.W.C. Brock, N.F. Winters, S.R. Stuart, M.A. Martinez, M. Li, M.C. Chambers, L.J. Zimmerman, A.J. Ham, D.L. Tabb, D.C. Liebler, Evaluation of strong cation exchange versus isoelectric focusing of peptides for multidimensional liquid chromatography-tandem mass spectrometry., J. Proteome Res. 7 (2008) 5286–94. doi:10.1021/pr8004666.

[241] S. Purvine, J.T. Eppel, E.C. Yi, D.R. Goodlett, Shotgun collision-induced dissociation of peptides using a time of flight mass analyzer, in: Proteomics, WILEY-VCH Verlag, 2003: pp. 847–850. doi:10.1002/pmic.200300362.

[242] R.S. Plumb, K.A. Johnson, P. Rainville, B.W. Smith, I.D. Wilson, J.M. Castro-Perez, J.K. Nicholson, UPLC/MSE; a new approach for generating molecular fragment information for biomarker structure elucidation, Rapid Commun. Mass Spectrom. 20 (2006) 1989–1994. doi:10.1002/rcm.2550.

[243] D. Helm, J.P.C. Vissers, C.J. Hughes, H. Hahne, B. Ruprecht, F. Pachl, A. Grzyb, K. Richardson, J. Wildgoose, S.K. Maier, H. Marx, M. Wilhelm, I. Becher, S. Lemeer, M. Bantscheff, J.I. Langridge, B. Kuster, Ion Mobility Tandem Mass Spectrometry Enhances Performance of Bottom-up Proteomics, Mol. Cell. Proteomics. 13 (2014) 3709–3715. doi:10.1074/mcp.M114.041038.

[244] L.C. Gillet, P. Navarro, S. Tate, H. Röst, N. Selevsek, L. Reiter, R. Bonner, R. Aebersold, Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis., Mol. Cell. Proteomics. 11 (2012) O111.016717. doi:10.1074/mcp.O111.016717.

[245] Uniprot, UniProtKB Reviewed - Swiss-Prot, (n.d.). http://www.uniprot.org/uniprot/?query=reviewed:yes (accessed March 9, 2018).

[246] Matrix Science, Mascot Server | Protein identification software for mass spec data, (n.d.). http://www.matrixscience.com/server.html (accessed March 9, 2018).

[247] M. Mann, M. Wilm, Error-tolerant identification of peptides in sequence databases by peptide sequence tags., Anal. Chem. 66 (1994) 4390–9. http://www.ncbi.nlm.nih.gov/pubmed/7847635 (accessed December 21, 2018).

[248] N. Colaert, Digging deeper in the data, novel approaches and analysis methods in quantitative peptide-centric proteomics, (2011). https://biblio.ugent.be/publication/1898634 (accessed

December 21, 2018).

[249] O.T. Schubert, L.C. Gillet, B.C. Collins, P. Navarro, G. Rosenberger, W.E. Wolski, H. Lam, D. Amodei, P. Mallick, B. MacLean, R. Aebersold, Building high-quality assay libraries for targeted analysis of SWATH MS data., Nat. Protoc. 10 (2015) 426–41. doi:10.1038/nprot.2015.015.

[250] J.X. Wu, X. Song, D. Pascovici, T. Zaw, N. Care, C. Krisp, M.P. Molloy, SWATH Mass Spectrometry Performance Using Extended Peptide MS/MS Assay Libraries., Mol. Cell. Proteomics. 15 (2016) 2501–14. doi:10.1074/mcp.M115.055558.

[251] J.X. Wu, D. Pascovici, V. Ignjatovic, X. Song, C. Krisp, M.P. Molloy, Improving Protein Detection Confidence Using SWATH Mass Spectrometry with Large Peptide Reference Libraries, Proteomics. 17 (2017) 1700174. doi:10.1002/pmic.201700174.

[252] G. Rosenberger, I. Bludau, U. Schmitt, M. Heusel, C.L. Hunter, Y. Liu, M.J. MacCoss, B.X. MacLean, A.I. Nesvizhskii, P.G.A. Pedrioli, L. Reiter, H.L. Röst, S. Tate, Y.S. Ting, B.C. Collins, R. Aebersold, Statistical control of peptide and protein error rates in large-scale targeted data-independent acquisition analyses, Nat. Methods. (2017). doi:10.1038/nmeth.4398.

[253] S.J. Parker, H. Rost, G. Rosenberger, B.C. Collins, L. Malmström, D. Amodei, V. Venkatraman, K. Raedschelders, J.E. Van Eyk, R. Aebersold, Identification of a Set of Conserved Eukaryotic Internal Retention Time Standards for Data-independent Acquisition Mass Spectrometry, Mol. Cell. Proteomics. 14 (2015) 2800–2813. doi:10.1074/mcp.O114.042267.

[254] P. Navarro, J. Kuharev, L.C. Gillet, O.M. Bernhardt, B. MacLean, H.L. Röst, S.A. Tate, C.-C. Tsou, L. Reiter, U. Distler, G. Rosenberger, Y. Perez-Riverol, A.I. Nesvizhskii, R. Aebersold, S. Tenzer, A multicenter study benchmarks software tools for label-free proteome quantification., Nat. Biotechnol. 34 (2016) 1130–1136. doi:10.1038/nbt.3685.

[255] C.-C. Tsou, D. Avtonomov, B. Larsen, M. Tucholska, H. Choi, A.-C. Gingras, A.I. Nesvizhskii, DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics., Nat. Methods. 12 (2015) 258–264. doi:10.1038/nmeth.3255.

[256] S.A. Gerber, J. Rush, O. Stemman, M.W. Kirschner, S.P. Gygi, Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS, Proc. Natl. Acad. Sci. 100 (2003) 6940–6945. doi:10.1073/pnas.0832254100.

[257] K. Kito, T. Ito, Mass Spectrometry-Based Approaches Toward Absolute Quantitative Proteomics, Curr. Genomics. 9 (2008) 263–274. doi:10.2174/138920208784533647.

[258] F. Calderón-Celis, J.R. Encinar, A. Sanz-Medel, Standardization approaches in absolute quantitative proteomics with mass spectrometry, Mass Spectrom. Rev. (2017). doi:10.1002/mas.21542.

[259] D.H. Lundgren, S.-I. Hwang, L. Wu, D.K. Han, Role of spectral counting in quantitative proteomics, Expert Rev. Proteomics. 7 (2010) 39–53. doi:10.1586/epr.09.69.

[260] E.J. Soderblom, J.W. Thompson, M.A. Moseley, Overview and implementation of mass spectrometry-based label-free quantitative proteomics, in: C. Eyers, S.J. Gaskell (Eds.), Quant. Proteomics, The Royal Society of Chemistry, 2014: pp. 131–153.

[261] B.C. Collins, C.L. Hunter, Y. Liu, B. Schilling, G. Rosenberger, S.L. Bader, D.W. Chan, B.W. Gibson, A.-C. Gingras, J.M. Held, M. Hirayama-Kurogi, G. Hou, C. Krisp, B. Larsen, L. Lin, S. Liu, M.P. Molloy, R.L. Moritz, S. Ohtsuki, R. Schlapbach, N. Selevsek, S.N. Thomas, S.-C. Tzeng, H. Zhang, R. Aebersold, Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry, Nat. Commun. 8 (2017) 1–12. doi:10.1038/s41467-017-00249-5.

[262] Y. Shen, J. Kim, E.F. Strittmatter, J.M. Jacobs, D.G. Camp, R. Fang, N. Tolié, R.J. Moore, R.D. Smith, Characterization of the human blood plasma proteome., Proteomics. 5 (2005) 4034–45. doi:10.1002/pmic.200401246.

[263] E. Boschetti, P.G. Righetti, The ProteoMiner in the proteomic arena: a non-depleting tool for discovering low-abundance species., J. Proteomics. 71 (2008) 255–64. doi:10.1016/j.jprot.2008.05.002.

[264] L. Guerrier, P.G. Righetti, E. Boschetti, Reduction of dynamic protein concentration range of biological extracts for the discovery of low-abundance proteins by means of hexapeptide ligand library, Nat. Protoc. 3 (2008) 883–890. doi:10.1038/nprot.2008.59.

[265] P.G. Righetti, G. Candiano, A. Citterio, E. Boschetti, Combinatorial Peptide Ligand Libraries as a " Trojan Horse " in Deep Discovery Proteomics, Anal. Chem. 87 (2015) 293–305.

[266] Bio-Rad, ProteoMiner Protein enrichment Kits. Instruction Manual., n.d.

[267] J.M. Burkhart, M. Vaudel, S. Gambaryan, S. Radau, U. Walter, L. Martens, J. Geiger, A. Sickmann, R.P. Zahedi, The first comprehensive and quantitative analysis of human platelet protein composition allows the comparative analysis of structural and functional pathways, Blood. 120 (2012). doi:10.1182/blood-2012-04-416594.

[268] A. Michalski, J. Cox, M. Mann, More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS., J. Proteome Res. 10 (2011) 1785–93. doi:10.1021/pr101060v.

[269] S. Camerini, P. Mauri, The role of protein and peptide separation before mass spectrometry analysis in clinical proteomics, J. Chromatogr. A. 1381 (2015) 1–12. doi:10.1016/j.chroma.2014.12.035.

[270] T.C. Petric, P. Brne, B. Gabor, L. Govednik, M. Barut, A. Strancar, L.Z. Kralj, Anion-exchange chromatography using short monolithic columns as a complementary technique for human serum albumin depletion prior to human plasma proteome analysis., J. Pharm. Biomed. Anal. 43 (2007) 243–9. doi:10.1016/j.jpba.2006.06.019.

[271] P.G. Righetti, E. Boschetti, Combinatorial peptide libraries to overcome the classical affinity-enrichment methods in proteomics., Amino Acids. 45 (2013) 219–29. doi:10.1007/s00726-013-1505-1.

[272] A.C. Grodzki, E. Berenstein, Antibody purification: ion-exchange chromatography., Methods Mol. Biol. 588 (2010) 27–32. doi:10.1007/978-1-59745-324-0_4.

[273] A. Hakimi, J. Auluck, G.D.D. Jones, L.L. Ng, D.J.L. Jones, Assessment of reproducibility in depletion and enrichment workflows for plasma proteomics using label-free quantitative data-independent

LC-MS., Proteomics. 14 (2014) 4–13. doi:10.1002/pmic.201200563.

[274] F. Di Girolamo, E. Boschetti, M.C.M. Chung, F. Guadagni, P.G. Righetti, “Proteomineering” or not? The debate on biomarker discovery in sera continues., J. Proteomics. 74 (2011) 589–94. doi:10.1016/j.jprot.2011.01.013.

[275] A. D’Amato, A. Bachi, E. Fasoli, E. Boschetti, G. Peltre, H. Sénéchal, J.P. Sutra, A. Citterio, P.G. Righetti, In-depth exploration of Hevea brasiliensis latex proteome and “hidden allergens” via combinatorial peptide ligand libraries., J. Proteomics. 73 (2010) 1368–80. doi:10.1016/j.jprot.2010.03.002.

[276] E. Fasoli, A. Farinazzo, C.J. Sun, A. V Kravchuk, L. Guerrier, F. Fortis, E. Boschetti, P.G. Righetti, Interaction among proteins and peptide libraries in proteome analysis: pH involvement for a larger capture of species., J. Proteomics. 73 (2010) 733–42. doi:10.1016/j.jprot.2009.10.008.

[277] B.R. Fonslow, P.C. Carvalho, K. Academia, S. Freeby, T. Xu, A. Nakorchevsky, A. Paulus, J.R. Yates, Improvements in proteomic metrics of low abundance proteins through proteome equalization using ProteoMiner prior to MudPIT., J. Proteome Res. 10 (2011) 3690–700. doi:10.1021/pr200304u.

[278] E. Fasoli, A. D’Amato, A. V Kravchuk, E. Boschetti, A. Bachi, P.G. Righetti, Popeye strikes again: The deep proteome of spinach leaves., J. Proteomics. 74 (2011) 127–36. doi:10.1016/j.jprot.2010.10.009.

[279] R.C. Dwivedi, O. V Krokhin, J.P. Cortens, J.A. Wilkins, Assessment of the reproducibility of random hexapeptide peptide library-based protein normalization., J. Proteome Res. 9 (2010) 1144–9. doi:10.1021/pr900608z.

[280] D. Baiwir, G. Mazzucchelli, N. Smargiasso, F. Quesada-Calvo, E. De Pauw, M. Malaise, E. Louis, M.-A. Meuwis, Comparison of serum fractionation methods by data independent label-free proteomics, EuPA Open Proteomics. 9 (2015) 14–22. doi:10.1016/j.euprot.2015.07.009.

[281] E. Mouton-Barbosa, F. Roux-Dalvai, D. Bouyssié, F. Berger, E. Schmidt, P.G. Righetti, L. Guerrier, E. Boschetti, O. Burlet-Schiltz, B. Monsarrat, A. Gonzalez de Peredo, In-depth exploration of cerebrospinal fluid by combining peptide ligand library treatment and label-free protein quantification., Mol. Cell. Proteomics. 9 (2010) 1006–21. doi:10.1074/mcp.M900513-MCP200.

[282] M. Fertin, O. Beseme, S. Duban, P. Amouyel, C. Bauters, F. Pinet, Deep plasma proteomic analysis of patients with left ventricular remodeling after a first myocardial infarction., Proteomics. Clin. Appl. 4 (2010) 654–73. doi:10.1002/prca.200900178.

[283] F. Di Girolamo, P.G. Righetti, M. Soste, Y. Feng, P. Picotti, Reproducibility of combinatorial peptide ligand libraries for proteome capture evaluated by selected reaction monitoring., J. Proteomics. 89 (2013) 215–26. doi:10.1016/j.jprot.2013.05.037.

[284] J. Zi, S. Zhang, R. Zhou, B. Zhou, S. Xu, G. Hou, F. Tan, B. Wen, Q. Wang, L. Lin, S. Liu, Expansion of the Ion Library for Mining SWATH-MS Data through Fractionation Proteomics, Anal. Chem. 86 (2014) 7242–7246. doi:10.1021/ac501828a.

[285] B.C. Collins, L.C. Gillet, G. Rosenberger, H.L. Röst, A. Vichalkovski, M. Gstaiger, R. Aebersold, Quantifying protein interaction dynamics by SWATH mass spectrometry: application to the 14-3-

3 system, Nat. Methods. 10 (2013) 1246–1253. doi:10.1038/nmeth.2703.

[286] J.-P. Lambert, G. Ivosev, A.L. Couzens, B. Larsen, M. Taipale, Z.-Y. Lin, Q. Zhong, S. Lindquist, M. Vidal, R. Aebersold, T. Pawson, R. Bonner, S. Tate, A.-C. Gingras, Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition, Nat. Methods. 10 (2013) 1239–1245. doi:10.1038/nmeth.2702.

[287] Y. Liu, A. Buil, B.C. Collins, L.C.J. Gillet, L.C. Blum, L.-Y. Cheng, O. Vitek, J. Mouritsen, G. Lachance, T.D. Spector, E.T. Dermitzakis, R. Aebersold, Quantitative variability of 342 plasma proteins in a human twin population., Mol. Syst. Biol. 11 (2015) 786. doi:10.15252/msb.20145728.

[288] T. Guo, P. Kouvonen, C.C. Koh, L.C. Gillet, W.E. Wolski, H.L. Röst, G. Rosenberger, B.C. Collins, L.C. Blum, S. Gillessen, M. Joerger, W. Jochum, R. Aebersold, Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps, Nat. Med. 21 (2015) 407–413. doi:10.1038/nm.3807.

[289] R. Bruderer, O.M. Bernhardt, T. Gandhi, S.M. Miladinović, L.-Y. Cheng, S. Messner, T. Ehrenberger, V. Zanotelli, Y. Butscheid, C. Escher, O. Vitek, O. Rinner, L. Reiter, Extending the Limits of Quantitative Proteome Profiling with Data-Independent Acquisition and Application to Acetaminophen-Treated Three-Dimensional Liver Microtissues, Mol. Cell. Proteomics. 14 (2015) 1400–1410. doi:10.1074/mcp.M114.044305.

[290] O.T. Schubert, C. Ludwig, M. Kogadeeva, M. Zimmermann, G. Rosenberger, M. Gengenbacher, L.C. Gillet, B.C. Collins, H.L. Rost, S.H.E. Kaufmann, U. Sauer, R. Aebersold, Absolute proteome composition and dynamics during dormancy and resuscitation of mycobacterium tuberculosis, Cell Host Microbe. 18 (2015) 96–108. doi:10.1016/j.chom.2015.06.001.

[291] N. Selevsek, C.-Y. Chang, L.C. Gillet, P. Navarro, O.M. Bernhardt, L. Reiter, L.-Y. Cheng, O. Vitek, R. Aebersold, Reproducible and Consistent Quantification of the Saccharomyces cerevisiae Proteome by SWATH-mass spectrometry, Mol. Cell. Proteomics. 14 (2015) 739–749. doi:10.1074/mcp.M113.035550.

[292] E.G. Williams, Y. Wu, P. Jha, S. Dubuis, P. Blattmann, C.A. Argmann, S.M. Houten, T. Amariuta, W. Wolski, N. Zamboni, R. Aebersold, J. Auwerx, Systems proteomics of liver mitochondria function, Science (80-.). 352 (2016) aad0189–aad0189. doi:10.1126/science.aad0189.

[293] J.M.B. Simbürger, K. Dettmer, P.J. Oefner, J. Reinders, Optimizing the SWATH-MS-workflow for label-free proteomics, J. Proteomics. 145 (2016) 137–140. doi:10.1016/j.jprot.2016.04.021.

[294] L. Reiter, O. Rinner, P. Picotti, R. Hüttenhain, M. Beck, M.-Y. Brusniak, M.O. Hengartner, R. Aebersold, mProphet: automated data processing and statistical validation for large-scale SRM experiments, Nat. Methods. 8 (2011) 430–435. doi:10.1038/nmeth.1584.

[295] F. Limonier, K. Van Steendam, G. Waeterloos, K. Brusselmans, M. Sneyers, D. Deforce, An application of mass spectrometry for quality control of biologicals: Highly sensitive profiling of plasma residuals in human plasma-derived immunoglobulin, J. Proteomics. 152 (2017) 312–320. doi:10.1016/j.jprot.2016.11.007.

[296] R. Craig, J.P. Cortens, R.C. Beavis, Open source system for analyzing, validating, and storing protein identification data, J. Proteome Res. 3 (2004) 1234–1242. doi:10.1021/pr049882h.

[297] M. Pernemalm, L.M. Orre, J. Lengqvist, P. Wikström, R. Lewensohn, J. Lehtiö, Evaluation of Three Principally Different Intact Protein Prefractionation Methods for Plasma Biomarker Discovery, J. Proteome Res. 7 (2008) 2712–2722. doi:10.1021/pr700821k.

[298] M. De Bock, D. de Seny, M.-A. Meuwis, A.-C. Servais, T.Q. Minh, J. Closset, J.-P. Chapelle, E. Louis, M. Malaise, M.-P. Merville, M. Fillet, Comparison of three methods for fractionation and enrichment of low molecular weight proteins for SELDI-TOF-MS differential analysis, Talanta. 82 (2010) 245–254. doi:10.1016/j.talanta.2010.04.029.

[299] O. Beseme, M. Fertin, H. Drobecq, P. Amouyel, F. Pinet, Combinatorial peptide ligand library plasma treatment: Advantages for accessing low-abundance proteins., Electrophoresis. 31 (2010) 2697–704. doi:10.1002/elps.201000188.

[300] R. Meng, M. Gormley, V.B. Bhat, A. Rosenberg, A.A. Quong, Low abundance protein enrichment for discovery of candidate plasma protein biomarkers for early detection of breast cancer., J. Proteomics. 75 (2011) 366–74. doi:10.1016/j.jprot.2011.07.030.

[301] P.G. Righetti, E. Boschetti, The ProteoMiner and the FortyNiners: Searching for gold nuggets in the proteomic arena, Mass Spectrom. Rev. 27 (2008) 596–608. doi:10.1002/mas.20178.

[302] D.N. Perkins, D.J.C. Pappin, D.M. Creasy, J.S. Cottrell, Probability-based protein identification by searching sequence databases using mass spectrometry data, Electrophoresis. 20 (1999) 3551–3567. doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2.

[303] Sciex, Improved Data Quality Using Variable Q1 Window Widths in SWATH® Acquisition, 2017. https://sciex.com/Documents/tech notes/Variable_Window_SWATH_Acquisition_TripleTOF_RUO-MKT-02-2879-B.pdf.

[304] S. Shelley, Immunoglobulin (IG) drives the blood-plasma therapeutics market, (2016). http://pharmaceuticalcommerce.com/brand-marketing-communications/immunoglobulin-ig-drives-the-blood-plasma-therapeutics-market/.

[305] A. Beck, E. Wagner-Rousset, D. Ayoub, A. Van Dorsselaer, S. Sanglier-Cianférani, Characterization of therapeutic antibodies and related products, Anal. Chem. 85 (2013) 715–736. doi:10.1021/ac3032355.

[306] BiopharmaLynx : Waters, (n.d.). http://www.waters.com/waters/en_US/BiopharmaLynx-/nav.htm?cid=513793&locale=en_US (accessed September 25, 2018).

[307] C. Doneanu, S. Lennon, M. Anderson, I. Reah, M. Ross, S. Anderson, I. Morns, Y. Qin Yu, A. Chakraborty, L. Denbigh, W. Chen, A Data Independent Acquisition Mode for Identification of Host Cell Proteins in Biopharmaceuticals, 2018. http://www.waters.com/webassets/cms/library/docs/720006262en.pdf (accessed September 25, 2018).

[308] F. Fortis, L. Guerrier, P.G. Righetti, P. Antonioli, E. Boschetti, A new approach for the removal of protein impurities from purified biologicals using combinatorial solid-phase ligand libraries., Electrophoresis. 27 (2006) 3018–27. doi:10.1002/elps.200500847.

[309] L. Tsiatsiani, A.J.R. Heck, Proteomics beyond trypsin, FEBS J. 282 (2015) 2612–2626. doi:10.1111/febs.13287.

[310] V. Guryča, D. Roeder, P. Piraino, J. Lamerz, A. Ducret, H. Langen, P. Cutler, Automated sample preparation platform for mass spectrometry-based plasma proteomics and biomarker discovery., Biology (Basel). 3 (2014) 205–19. doi:10.3390/biology3010205.

[311] Q. Fu, M.P. Kowalski, M. Mastali, S.J. Parker, K. Sobhani, I. van den Broek, C.L. Hunter, J.E. Van Eyk, Highly Reproducible Automated Proteomics Sample Preparation Workflow for Quantitative Mass Spectrometry, J. Proteome Res. 17 (2018) 420–428. doi:10.1021/acs.jproteome.7b00623.

[312] D.C. Liebler, L.J. Zimmerman, Targeted Quantitation of Proteins by Mass Spectrometry, Biochemistry. 52 (2013) 3797–3806. doi:10.1021/bi400110b.

[313] N. Rauniyar, Parallel reaction monitoring: A targeted experiment performed using high resolution and high mass accuracy mass spectrometry, Int. J. Mol. Sci. 16 (2015) 28566–28581. doi:10.3390/ijms161226120.

[314] T. Shi, E. Song, S. Nie, K.D. Rodland, T. Liu, W.-J. Qian, R.D. Smith, Advances in targeted proteomics and applications to biomedical research, Proteomics. 16 (2016) 2160–2182. doi:10.1002/pmic.201500449.

[315] Y.S. Ting, J.D. Egertson, S.H. Payne, S. Kim, B. MacLean, L. Käll, R. Aebersold, R.D. Smith, W.S. Noble, M.J. MacCoss, Peptide-Centric Proteome Analysis: An Alternative Strategy for the Analysis of Tandem Mass Spectrometry Data, Mol. Cell. Proteomics. 14 (2015) 2301–2307. doi:10.1074/mcp.O114.047035.

[316] Y. Li, C.-Q. Zhong, X. Xu, S. Cai, X. Wu, Y. Zhang, J. Chen, J. Shi, S. Lin, J. Han, Group-DIA: analyzing multiple data-independent acquisition mass spectrometry data files, Nat. Methods. 12 (2015) 1105–1106. doi:10.1038/nmeth.3593.

[317] J. Wang, M. Tucholska, J.D.R. Knight, J.-P. Lambert, S. Tate, B. Larsen, A.-C. Gingras, N. Bandeira, MSPLIT-DIA: sensitive peptide identification for data-independent acquisition, Nat. Methods. 12 (2015) 1106–1108. doi:10.1038/nmeth.3655.

[318] Y.S. Ting, J.D. Egertson, J.G. Bollinger, B.C. Searle, S.H. Payne, W.S. Noble, M.J. MacCoss, PECAN: library-free peptide detection for data-independent acquisition tandem mass spectrometry data, Nat. Methods. 14 (2017) 903–908. doi:10.1038/nmeth.4390.

[319] R. Peckner, S.A. Myers, A.S.V. Jacome, J.D. Egertson, J.G. Abelin, M.J. MacCoss, S.A. Carr, J.D. Jaffe, Specter: linear deconvolution for targeted analysis of data-independent acquisition mass spectrometry proteomics, Nat. Methods. 15 (2018) 371–378. doi:10.1038/nmeth.4643.

[320] C.-C. Tsou, C.-F. Tsai, G.C. Teo, Y.-J. Chen, A.I. Nesvizhskii, Untargeted, spectral library-free analysis of data-independent acquisition proteomics data generated using Orbitrap mass spectrometers., Proteomics. 16 (2016) 2257–71. doi:10.1002/pmic.201500526.

[321] R.S. Plumb, K.A. Johnson, P. Rainville, B.W. Smith, I.D. Wilson, J.M. Castro-Perez, J.K. Nicholson, UPLC/MSE; a new approach for generating molecular fragment information for biomarker structure elucidation, Rapid Commun. Mass Spectrom. 20 (2006) 1989–1994. doi:10.1002/rcm.2550.

[322] A. Bilbao, E. Varesio, J. Luban, C. Strambio-De-Castillia, G. Hopfgartner, M. Müller, F. Lisacek, Processing strategies and software solutions for data-independent acquisition in mass

spectrometry, Proteomics. 15 (2015) 964–980. doi:10.1002/pmic.201400323.

[323] R. Bruderer, O.M. Bernhardt, T. Gandhi, Y. Xuan, J. Sondermann, M. Schmidt, D. Gomez-Varela, L. Reiter, Optimization of Experimental Parameters in Data-Independent Mass Spectrometry Significantly Increases Depth and Reproducibility of Results, Mol. Cell. Proteomics. 16 (2017) mcp.RA117.000314. doi:10.1074/mcp.RA117.000314.

[324] F. Girolamo, I. Lante, M. Muraca, L. Putignani, The Role of Mass Spectrometry in the “Omics” Era, Curr. Org. Chem. 17 (2013) 2891–2905. doi:10.2174/13852728178881311181627 25.

[325] E. Baker, T. Liu, V.A. Petyuk, K.E. Burnum-Johnson, Y.M. Ibrahim, G.A. Anderson, R.D. Smith, Mass spectrometry for translational proteomics: progress and clinical implications, Genome Med. 4 (2012) 63. doi:10.1186/gm364.

[326] P.E. Geyer, L.M. Holdt, D. Teupser, M. Mann, Revisiting biomarker discovery by plasma proteomics, Mol. Syst. Biol. 13 (2017) 942. doi:10.15252/msb.20156297.

[327] D.J. Bowen, Haemophilia A and haemophilia B: molecular insights., Mol. Pathol. 55 (2002) 127–44. http://www.ncbi.nlm.nih.gov/pubmed/11950963 (accessed September 29, 2017).

Limonier Franck

Scientist at *Sciensano*
Bioengineer in chemistry and bio-industries

Avenue Molièrelaan, 9
1190 Vorst-Forest, Belgium

franck.limonier@gmail.com

+32 476 462 758

Born on March 31th 1990, Belgium

Experience

June 2017-Present — Scientist, *Mass Spectrometry Platform*, Sciensano
Chemical risk orientation and MS screening methods

October 2013-January 2019 — Scientific collaborator and researcher, *Quality of Vaccines and Blood Products*, Sciensano (ex-Scientific Institute of Public Health WIV-ISP)
PhD student in Pharmaceutical Sciences, *Laboratory of pharmaceutical biotechnologies* (Prof. D. Deforce), Ghent University
"Prevention is better than cure: Non-targeted and sensitive screening of impurities in plasma-derived immunoglobulin with mass spectrometry-based proteomics."

August 2012-June 2013 — Master thesis in cellular and nutritional biochemistry (Pr. Y-J Schneider)
"Effect of apple polyphenolic extracts on obesity-related inflammation of the adipose tissue."

August 2012 — Student Job, Croix-Rouge de Belgique,
Blood donation service of Saint Pierre University Hospital

July-August 2010 — Trainee, Ceres S.A. (Cereals industry), quality control laboratory

Education

2013 - 2019 — PhD student in Pharmaceutical Sciences
Universiteit Gent

2008 - 2013 — Master degree in Bioengineer in chemistry and bio-industries
With specialization in food sciences, technologies and quality
Université Catholique de Louvain

2002 - 2008 — Secondary education
Collège Saint-Hubert, Brussels

Publications in peer-reviewed journals

Estimating the Reliability of Low-Abundant Signals and Limited Replicate Measurements through MS2 Peak Area in SWATH.
Limonier, F., Willems, S., Waeterloos, G., Sneyers, M., Dhaenens, M. & Deforce, D. (2018)
Proteomics 18 (24): e1800186
DOI: 10.1002/pmic.201800186

An application of mass spectrometry for quality control of biologicals: Highly sensitive profiling of plasma residuals in human plasma-derived immunoglobulin.
Limonier, F., Van Steendam, K., Waeterloos, G., Brusselmans, K., Sneyers, M. & Deforce, D. (2017)
Journal of Proteomics 152:312-320.
DOI: 10.1016/j.jprot.2016.11.007

National and international conferences

Belgian Proteomics Association (BePA) Biannual conference

Prevention is better than cure: Non-targeted and sensitive screening of impurities in plasma-derived immunoglobulin with MS-based proteomics
2018 - oral presentation, Brussels, Belgium

Towards application of non-targeted MS for quality control of biologicals: highly sensitive profiling of plasma residuals in human immunoglobulin.
2016 - poster, Gent, Belgium

Mass spectrometry-based proteomics for the quality control of intravenous immunoglobulin: Detection of possible thrombogenic impurities
2014 - poster, Gent, Belgium

Human Proteome Organization (HUPO) 13th World congress
Mass spectrometry-based proteomics for the quality control
of intravenous immunoglobulin: Detection of possible thrombogenic impurities
2014 - poster, Madrid, Spain

European Directorate for the Quality of Medicines & Healthcare (EDQM) 50 years
Mass spectrometry-based proteomics for the quality control
of intravenous immunoglobulin: Detection of possible thrombogenic impurities
2014 - poster, Strasbourg, France

Activities

Writer for independent music webzine *La Vague Parallèle*
Organiser of live music events
Violinist playing classical and tango music, formerly musician and member of the organization committee at the University Students Orchestra OSEL.


UNIVERSITEIT
GENT


sciensano