Analytical Characterization of Protein-based Biopharmaceuticals

Ana Martins Rodrigues

Thesis to obtain the Master of Science Degree in

Pharmaceutical Engineering

Supervisor: Doctor Marco André Miranda Galésio

Professor Ana Margarida Nunes da Mata Pires de Azevedo

Examination Committee

Chairperson: Professor José Monteiro Cardoso de Menezes

Supervisor: Doctor Marco André Miranda Galésio

Members of Committee: Doctor Joana Margarida Lopes da Silva Cristóvão

November 2019

I declare that this document is an original work of my own authorship and that it fulfils all the requirements of the Code of Conduct and Good Practices of the *Universidade de Lisboa*.

Preface

The work presented in this thesis was performed at the company Hovione (Lumiar, Lisbon, Portugal), during the period of November 2018 to July 2019, under the supervision of doctor Marco Galésio and doctor Susana Ramos. The thesis was co-supervised at Instituto Superior Técnico by Professor Ana Azevedo.

Acknowledgements

A conclusão desta dissertação não teria sido possível sem a valiosa contribuição de algumas pessoas, às quais gostaria de expressar o meu sincero agradecimento.

Primeiramente gostaria de expressar a minha gratidão à Constança Cacela, por tão bem me ter acolhido e dado a oportunidade de me integrar no grupo de desenvolvimento e por me ter proporcionado a excelente experiência na Hovione.

A todos os orientadores envolvidos na minha tese e os melhores que podia pedir, Marco Galésio, Susana Ramos e Ana Azevedo, por todo o suporte e apoio ao longo deste grande projeto, com o intuito de o trabalho ser sempre melhor, com novas metas e conquistas.

A todas as pessoas do grupo da analítica que tive o prazer de partilhar o laboratório e que se predispuseram desde o início do processo a ajudar-me em tudo aquilo que era necessário.

À Rute Nunes, colega e amiga para a vida, porque me fez crescer científica e pessoalmente, porque todos os dias eram para ser conquistados o melhor possível, por todos os conselhos construtivos, por todas as ajudas no laboratório e fora dele, porque todos os dias recebia e partilhava sorrisos, por todas as conversas de ânimo a tomar o pequeno almoço. Levo-te comigo para a vida. *"Bom dia minha perninhas de rã".*

À Liliana Silva, colega e amiga para a vida, porque também me fez crescer muito, a ser ponderada, a saber enfrentar as adversidades da melhor maneira possível, por todas as nossas conversas, partilhas, rotina diária de entreajuda, até por ser tão chata com as amostras de LC-MS…em especial, por teres partilhado comigo momentos da tua vida e obrigada por essa deposição de confiança. Levo-te comigo para a vida!

À Lúcia Sousa, por ser a melhor pessoa que conheci neste mundo e no outro, porque me ensinou diariamente a perceber que sempre que há dedicação e esforço, tudo é alcançado, e quando não é, é porque já não teve de ser. Porque as gargalhadas são características, porque o tempo que despendeste comigo para todo o trabalho foi precioso e levo cada bocadinho de ensinamento, cada conselho, cada partilha, cada vitória e cada derrota, tudo muito graças a ti. Levo-te comigo sempre, para a vida inteira!

À Joana Cristóvão, doutorada com distinção e louvor, obrigada por tudo, por toda a paciência, porque deste um excelente apoio ao meu projeto, porque contigo aprendi a ser determinada, ponderada, por toda a ajuda possível e imaginária, porque foste incansável e caminhei com o maior esforço sempre contigo neste projeto.

Ao bus crew, ao John, Minh, Grainne e James, por me terem enriquecido com a sua cultura, por terem melhorado o meu inglês pela prática contínua durante os almoços divertidos e no dia-a-dia. Por todos os passeios ao sol no Lumiar, por todas as conversas e partilhas.

À Sofia Rodrigues e Sónia Ferreira, porque as meninas da inovação estiverem sempre presentes, obrigada pelas pausas para gargalhadas, por todas as reflexões, pela arte de todos os conselhos que

desempenharam na perfeição e, porque quer com tripeiras quer com canadenses, as vibrações foram sempre as melhores.

Aos meus pais e irmãos um profundo obrigada, com o meu coração inteiro, por terem suportado os meus estudos e porque no fim tudo vale a pena, por serem exemplo de coragem e determinação, por estar sempre presente o amor incondicional, e por me ensinarem que tudo nesta vida é possível, porque o cansaço se transforma em louvores e o esforço e dedicação constantes são a maior arma da vida.

Ao João, a metade que me completa, os sorrisos que se complementam e um amor indescritível. Obrigada pela compreensão de todas as ausências, por todo o tempo em que a atenção não foi a melhor, obrigada por todo o apoio, pela estabilidade emocional que me proporcionaste e, que foi preponderante para este projeto. Obrigada simplesmente por teres sempre caminhado comigo, sem nunca desistirmos e, desta forma, por tão bem me completares!

Abstract

Biopharmaceuticals are produced through biotechnological processes, using molecular biology methods, encompassing several classes, such as nucleic acids and proteins. The main goal of this project was to develop, understand and implement fit for purpose analytical methodologies to characterize biological large molecules, such as proteins.

This project includes the characterization of model proteins (Bovine Serum Albumin, Ovalbumin and Lysozyme) and a monoclonal antibody, ARD_2019. It includes the development of a Karl Fischer Titration (KFT) to determine the process water content and the development of Size Exclusion Chromatography (SEC) together with Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE) methodologies to determine protein purity, size variants and stability of proteins. An Ion Exchange Chromatography (IEX) was also developed and implemented to assess the presence of different charged variants.

By KFT, it was possible to determine the process water content of BSA [6.87% (w/w)] and ARD_2019 [0.55 - 3.02% (w/w)]. The SEC method was first optimized for the model proteins with known molecular weights, and later applied to ARD_2019, which allowed the estimation of its molecular weight, and the High and Low Molecular Weight species. An SDS-PAGE methodology was also developed. The results obtained for the isolated antibody are in accordance with SEC results. By SDS-PAGE was possible to observe both antibody heavy (50 kDa) and light (25 kDa) chains. The determination of charge heterogeneity analysed by Cation Exchange Chromatography was performed with both pH and salts gradients. For the ARD_2019 mAb it was possible to determine the presence of 8 different charged variants.

Keywords

Biopharmaceuticals • Biocharacterization • Bovine Serum Albumin • Lysozyme • ARD_2019 protein •Analytical Methodologies

Resumo

Os biofarmacêuticos são produzidos através de processos biotecnológicos, recorrendo a métodos de biologia molecular. Abrangem diversas classes, tais como ácidos nucleicos e proteínas. O objetivo deste projeto foi desenvolver, entender e implementar metodologias analíticas adequadas para a caracterização de macromoléculas biológicas, nomeadamente proteínas.

Este projeto engloba a caracterização de proteínas modelo (Albumina de Soro Bovino, Ovalbumina e Lisozima) e um anticorpo monoclonal, ARD_2019, através do desenvolvimento de técnicas analíticas, como titulação de Karl Fischer (KFT), Cromatografia de Exclusão de Molecular (SEC), Electroforese em Gel de Poliacrilamida (SDS-PAGE) e Cromatografia de Troca Iónica (IEX).

A partir do método de KFT, determinou-se o conteúdo de água do processo de BSA [6.87% (w/w)] e do ARD_2019 [0.55 - 3.02% (w/w)]. O método de SEC, otimizado primeiramente para proteínas modelo com pesos moleculares conhecidos e aplicado ao anticorpo ARD_2019, permitiu calcular o seu peso molecular nativo, bem como de espécies de alto e baixo peso molecular. A técnica SDS-PAGE foi também desenvolvida. Os resultados obtidos para o anticorpo isolado, estão de acordo com os resultados obtidos pela técnica de SEC. A avaliação da pureza do anticorpo, foi realizada por SDS-PAGE, concluindo-se que o anticorpo se encontrava puro, observando-se a separação das cadeias pesada (50 kDa) e leve (25 kDa) do anticorpo. A determinação da heterogeneidade de carga por IEX foi realizada em gradiente de pH e de força iónica. Para o anticorpo monoclonal, ARD_2019, foi possível determinar a presença de 8 variantes de carga diferentes.

Palavras-chave

Biofarmacêuticos • Biocaracterização • Albumina de Soro Bovino • Lisozima • Proteína ARD_2019 •Metodologias Analíticas

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ADCs − Antibody-Drug Conjugates

BEH − Ethylene-Bridged Hybrid inorganic–organic

BSA − Bovine Serum Albumin

CATH − Class Architecture Topology/fold Homologous Superfamily

CHO − Chinese Hamster Ovary

DLS − Dynamic Light Scattering

DNA −Deoxyribonucleic Acid

ELISA − Enzyme-Linked Immunosorbent Assay

ESI - QTOF − Electrospray Ionization Quadrupole Time of Flight

GC − Gas Chromatography

HC − Heavy Chain

HMW − High Molecular Weight

HPLC − High Performance Liquid Chromatography

ICH − International Conference Harmonization

IEX − Ion Exchange Chromatography

IgG − Immunoglobulin G

KFT− Karl Fischer Titration

LC − Light Chain

LC − Liquid Chromatography

LogMW – Logarithm of Molecular Weight

LMW − Low Molecular Weight

mAb − monoclonal antibody

MALDI-TOF − Matrix-Assisted Laser Desorption Ionization Time of Flight

MS − Mass Spectroscopy

PABA − Paraminobenzoic Acid

pI − Isoelectric Point

pKa − acid ionization constant

PTM − Post Translational Modification

RI −Refractive Index

RSD −Relative Standard Deviation

RT − Retention Time

SCOP − Structural Classification of Proteins

SDS-PAGE − Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis

SEC − Size Exclusion Chromatography

TFA −Trifluoroacetic Acid

UHPLC − Ultra High Performance Liquid Chromatography

Chapter 1

Introduction

1.1 Biopharmaceuticals

1.1.1 Overview

In the 20th century, the pharmaceutical industry has deeply evolved due to a developing appreciation of the principles of organic chemistry together with an increased knowledge in molecular biology, in disease-associated pathways and disease mechanisms [1]. More recently, the combination of increased knowledge with technological advances in areas such as genomics, proteomics and high-throughput screening are significantly changing the field of new drug development.

The term “biopharmaceutical” was first used in the 1980s to describe classes of cytokines, enzymes, hormones, clotting factors, vaccines, cell therapies, antisense drugs and therapeutic proteins, produced by modern biotechnological techniques, specifically via genetic [2]. Proteins with identical amino acid sequence could be obtained by different approaches, such as recombinant DNA or extraction from animal tissue. However, the characteristics of the two identical proteins may vary if extracted from different sources. This idea becomes more clear when approaching the different terms used to described the type of biopharmaceutical between *reference*, *biosimilars* and *biobetters*. The *reference* medicine is the first biopharmaceutical version of the same protein. *Biosimilars* also referred to as a “follow-on biologic”, share highly similar, but not identical, safety and efficacy profiles to an already approved reference medicine. Their main differences arise from post-translational modifications or from the use of different manufacturing process. The production and marketing of biosimilars require meticulous comparison with the reference biologic, which involves a variety of methods, including comparability exercises [3] [4] [5]. The *biobetters* are therapeutic macromolecules modified either by chemical and/or engineered methods, which increases significantly their pharmacologic properties, such as higher activity, enhanced stability, fewer side effects and lower immunogenicity.

The United States and/or European Union witnessed over the period from January 2014 until July 2018, the approval of 155 biopharmaceutical products, being 81 (52%) of the these genuinely new to the market, with the remaining products representing biosimilars, me-too products (a product introduced by a company after it has seen that other companies are successful with the same type of product) and products previously approved elsewhere [6]. Biopharma manufacturing has become a strategic driver in a pharmaceutical industry, with the ability to create and maintain market access through scalable and flexible operations, controlled costs and high quality products. The implementation of these strategies is supported by the impact that this industry is currently reflecting with increased therapeutic competition, greater prevalence of large molecules, expansion in the number of personalized or target products, and a rise of treatments for many orphan diseases [7] [4].

It was notorious that advances in biomedical research have identified various biomolecules synthesized naturally by the body, whose therapeutic potential was obvious. Biopharmaceutical proteins are emerging as an important class of drugs for the treatment of various diseases including cancer, inflammatory diseases and autoimmune disorders, among others. Several examples include insulin, interferons, interleukins and other cytokines, which regulate aspects of immunity, inflammation and other

processes of central importance to maintaining a healthy state [8]. In this class, growth factors can also be included, such as erythropoietin and neurotrophic factors, which regulate the development and maintenance of neural tissue. While the pharmaceutical potential of these regulatory molecules was generally appreciated, its widespread medical application was, in most cases, impracticable due to the small quantities in which they were produced naturally. Unlike blood products and various hormones that are naturally produced in enough quantities, which facilitates their extraction from biological source materials, most cytokines are produced in exceedingly low concentrations in the body. Consequently, this made their isolation difficult and routine large-scale production impossible. The boosting and discovery of recombinant DNA (genetic engineering) and hybridoma technology overcomes many of these difficulties, and marks a new era in pharmaceutical sciences [2]. There are a wide variety of biopharmaceutical proteins formats, enhanced by new nucleic acid modalities and cellular therapies that are being propelling on the market. However, monoclonal antibodies (mAbs) and antibody-drug conjugates (ADCs), recombinant proteins, produced by engineered Chinese hamster ovary (CHO) cells and, therefore, extracted from mammalian cells, have been emerged as the largest group of biopharmaceuticals, dominating the biopharmaceutical market, attributed to its obvious benefits in terms of safety and efficacy [6]. Mammalian cell derived recombinant therapeutic proteins, is based on the principle that the transformation of these cells using a genetic vector, proliferate *in vitro* on a top of a complex glycoprotein production [9].

The greatest advantages of biopharmaceutical drugs over conventional drugs are their high specificity and potency. Biopharmaceuticals drugs also present fewer side effects. The most important attributes in biomolecules characteristics arise from their macromolecular nature, which provides the structural complexity that is often required for specificity. This structural complexity makes the characterization biomolecules critical quality attribute more challenging, since the structural integrity of the protein needs to be assured. Many biologics such as monoclonal antibodies and other recombinant therapeutic proteins are larger and complex proteins. Therefore, the analytical methodologies for the characterization of these molecules must be able to evaluate their primary structure and the high ordered secondary, tertiary and quaternary structures, that define the protein three-dimensional structure shape and function. The characterization of post translational modifications (PTMs), glycosylation, oxidation, phosphorylation, sulphation, lipidation, disulphide bond formation, and deamidation are crucial aspects of the analysis, since these phenomenon can impact in the three-dimensional structure of the protein and influence the immunogenicity of biologics [10].

1.1.2 Classes of Biopharmaceuticals

Biopharmaceuticals are characterized by their complexity, being totally different from the conventional small molecules, fundamentally in the average size – for example, while the molecular weight of an aspirin, small molecule, is 180 Da, the biopharmaceuticals might be over 100 times larger, which is the case of the interferon beta with a molecular weight of 19 000 Da. Biopharmaceuticals may also be distinguished from small molecules in terms of their higher potential heterogeneity, their safety and

efficacy profile, and even their manufacturing process. Process engineering has been applied to alter immunological or pharmacokinetic profile of a protein, or equally to generate novel fusion proteins. The main areas of biopharmaceuticals action are cancer, viral infections, diabetes, hepatitis, neurodegenerative and cardiac diseases, and multiple sclerosis. The areas of action of proteins can still be grouped and divided by cytokines, enzymes, hormones, clotting factors, vaccines, cell therapies and antisense drugs. Cytokines, such as interferon and interleukins, are mediator molecules, controlling reactions between cells, relatively to the immune system, activating lymphocytes and macrophages. These are characterized by reacting to the virus and preventing the uncontrolled cell proliferation, also acting as intermediates in the various steps of the immune response. Control of a proinflammatory reactions in response to injured tissues, is due to a class of a interleukins, IL-1, which is praised for being associated with inflammatory disorders, such as systemic juvenile idiopathic arthritis, adult-onset Still disease and rheumatoid arthritis. Enzymes, such as alteplase and dornase alfa, that dissolves blood clots and digests DNA in the mucous secretions in lungs, are characterized by triggering chemical changes into other substances, without changes in the chemistry or conformation in themselves. Vaccines, as an example, hepatitis B virus [Baraclude (Entecavir), Adefovir dipivoxil (Hepsera)], are microorganisms that, like cytokines, are associated with immune response system related phenomenon, activating the immune response and stimulating human resistance to specific diseases. Currently, approximately one third of all new treatments, are referenced using monoclonal antibodies, produced from immortal cells with an antibody producing spleen cells and encompassing a wide range of treatments at the level of breast cancers, leukemia, asthma, rheumatoid arthritis, psoriasis, chronic gastrointestinal inflammatory disease and transplant rejection. In order to minimizing immunogenic reactions, Humira® has emerged as a first fully human monoclonal antibody in 2003 [11]. Despite the wide range of applications involving mAbs, trends in benchmarks for various types of these proteins, have shown greater impetus, regard to anticancer mAbs [12].

The most recognized stem cell therapy is the bone marrow transplant, involved in processes of introducing new cells into tissues with the aim to treat a disease, with perspectives in regenerative medicine using stem-cell research, allied to tissue engineering and gene therapy, where the main goal is repair, replacement and regeneration of cells, tissues or organs. Liquid phase and solid phase technologies offered the possibility to produce an amount of therapeutics peptides, that currently, only selected antimicrobial peptides have been licensed or in preclinical or clinical stages contain non-natural amino acids, to make them more active and stable.

1.1.3 Biopharmaceutical Manufacturing and Technology Innovation

Biopharmaceutical drugs are large and complex molecules derived from living cells and therefore acquire a status of greater complexity, as previously mentioned. Currently, the great majority of commercially available biopharmaceuticals contain recombinant proteins as their active pharmaceutical ingredient. The inherent specificity of each recombinant protein, with respect to its properties, requires a wide range of expression systems, particularly in prokaryotic systems, mammalian cells, or insect cell

lines. It is notorious the number of recombinant protein products produced in mammalian cells, approved for the use of drugs in humans, being the elected expression system, with an increase of approximately 60%, over 2010–2014. Chinese hamster ovary cell line, a mammalian cell line, is the primary choice for recombinant protein production, due to its suitability to produce conveniently glycosylated proteins, a specific post-translational modification (PTM), only with specificity of occurrence in mammalian expression systems, where the same is not verified in bacterial expression system. However, there are some disadvantages regarding this expression system, since in addition to high complex nutritional requirements, slow cell growth and high fragility, there are known challenges related to contaminants by animal viruses that may arise from such systems. In contrast to what happens in mammalian systems, where the largest part of the proteins produced are monoclonal antibodies, in bacteria, the predominant group are insulins. *Escherichia Coli (E. coli)* bacteria is totally suitable for this type of system, since it is associated with rapid growth, high product yield, easy process scale-up and short timeframes. Bacterial expression systems may be affected by its inability to produce correct disulfide bonds, protein solubility issues and the presence of endotoxins. *Saccharomyces. cerevisiae and Pichia. pastoris*, yeast expression systems, can produce recombinant proteins, due to their rapid growth in protein free media and ability to secrete the product extracellularly. The same happens with insect cell-based, an expression system also known for its growth to higher densities, allowing for post translational modifications, but inversely to what happens with mammalian system, not preserve the original glycosylation pattern [13]. Despite the strength of mammalian cell lines as factories, microbial cells and specially *E. coli*, are still potent protein factories essentially supported by their versatility and cost-effective cultivation. An update performed from recombinant pharmaceuticals from microbial cells demonstrated that approximately 24% of the protein biopharmaceuticals marketed are antitumor drugs. From this 24%, only a small part is considered as being true antitumoral drugs, being the other drugs mainly used for supportive purposes intended to minimize the side effects of cancer treatments. From the true antitumoral drugs, 69% are expressed using *E. coli* expression system [14].

In order to achieve optimal conditions for cell growth and biopharmaceutical production, these expression systems are predominant in the upstream process during biopharmaceuticals production. The manufacturing process of biopharmaceutical products are referred to as upstream and downstream processing, as it may be seen in Figure 1. Upstream process involves factors associated with the culture medium and maintenance of cells [3]. Through upstream process optimization is critical to control process parameters, such as temperature, pH, oxygen and sterilization. Additionally, the production process can occur in batch, fed-batch or continuous. The control and optimization of all these conditions is demanding and may affect the correct transformation of substrates into the desired metabolic products. The downstream process are related to the chemical and physical separations and involves three major stages: (1) initial recovery, which includes extraction and isolation processes, (2) purification, with the aim of removing contaminants, and finally (3) purge specific contaminants that accumulated during the isolation, namely the polishing step.

Categorized as a robust, reliable and easily scaled up process, the purification step is preponderant during the production process. Its purpose is to remove host cell related impurities, as well as process and product related impurities in order to capture the target biomolecule for a final purified product. Initial

recovery its affected by the production conditions of the target molecule, whether extracellularly or intracellularly. Thus, the various steps of these process, centrifugation, filtration, sedimentation and flotation, are adjusted according to the specificity of production of the target molecule [3,11].

The highest outlay in biopharmaceutical manufacturing is attributed to downstream processing, which is responsible for a massive 40% of the total costs, adding lower rates of 30% and 20% of the costs of

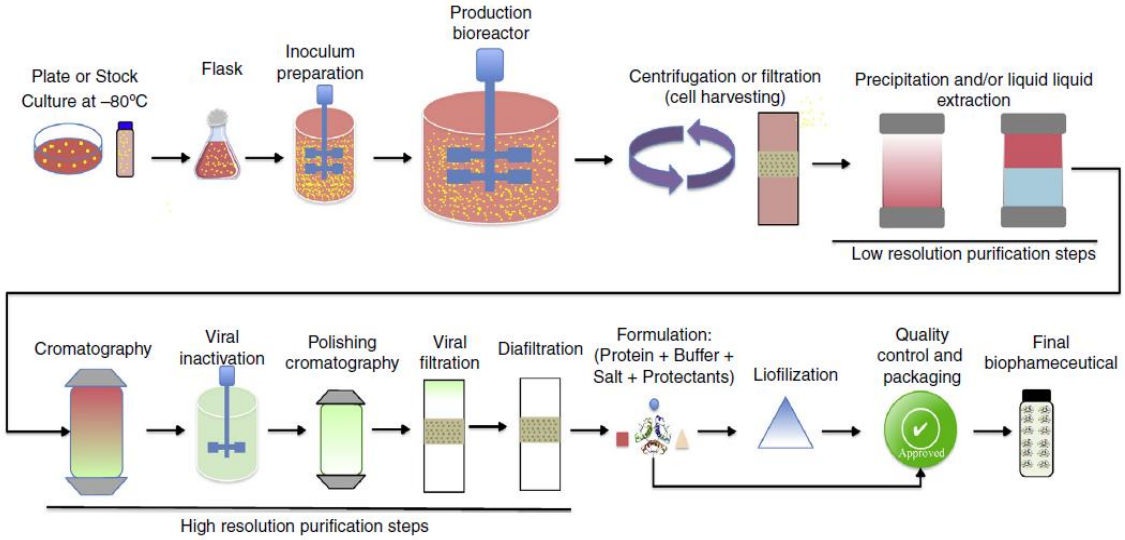

Figure 1.The biopharmaceutical manufacturing technology flowchart exemplifying the upstream and the downstream bioprocess. Adapted from A.Faustino *et al.* [3]

process development accounts and for upstream processing, respectively [11].

Biopharmaceutical manufacturing is characterized by the use of advanced technologies, harnessing of new scientific advances. Manufacturing technology innovation spans primary and secondary manufacturing, the two general steps in the drug production process, contributing for increasingly sophisticated enhancements to chemical and biological processes. Although biopharmaceutical manufacturers are innovating throughout the entire process from raw material to finished drugs product, the main goal is related with large molecule (biologic) production. Advances of great investment in this type of innovation technology, as shown in Figure 2, are based on continuous manufacturing, process analytical technology, single-use systems and other new areas of technology innovation such as, alternative techniques in downstream processing methods of vaccine and therapy production and green chemistry, which enables manufacturing flexibility and scalability while improving quality, throughput and controlling costs [7].

Figure 2. Key areas of manufacturing technology innovation. Adapted from Life Sciences [7]

1.1.4 Quality Control and Analytical Techniques for Biopharmaceuticals Analysis

The International Conference for Harmonization (ICH) released the guideline Q6B, “Test Procedures and Acceptance Criteria for Biotechnological/Biological Products”, with the aim to establish the relevant specifications and acceptance criteria for the complete characterization of a biotechnological or biological product, which includes the determination of physicochemical properties, biological activity, immunochemical properties, purity and impurities.

The determination of the composition, physical properties and primary structure of the desired product are essential to a complete physicochemical characterization, as it also has an active role in the assessment of the consistency of the manufacturing process [15,16]. Structural heterogeneity is an inherent characteristic of recombinant proteins, that can occur during the biosynthetic processes used by living organisms in their production or over manufacture and/or storage of the drug substance or drug product. Consequently, the therapeutic product might be a mixture of different protein variants that may or may not affect the safety and efficacy of the product. If these protein variants demonstrate equivalent activity, efficacy and safety to those of the desired product, are considered product related substances [15].

Protein-based biopharmaceuticals present a secondary, tertiary and, in some cases, quaternary structures. Comprising a high order structure, biopharmaceuticals require not only an extensive characterization of their physicochemical properties, such as molecular weight or size, isoform pattern, extinction coefficient, electrophoretic patterns and chromatographic patterns, but also biological assays, which is an assay where a biological response is induced in a biological assay system. The primary structure, or amino acid sequence, that defines the identity of a biological, refers to one of the first physical parameters to be analysed in the intact molecule, namely the molecular weight. Other relevant techniques to characterize the physicochemical properties of proteins are size exclusion chromatography, sodium dodecyl sulphate polyacrylamide gel electrophoresis (under reducing and/or non-reducing conditions) and mass spectrometry. Mass spectrometry is a key technology for sequence determination, post-translational modifications (PTMs) and higher order structure (HOS) [15,17].

The primary structure usually analysed by the “bottom up” approach, provides evidence of the amino acid composition, disulfide linkages and PTMs of biotherapeutics. The “bottom up” methodology is a strategy commonly used with mass spectrometry, in which a protein is digested with a specific enzyme to generate smaller protein fragments, peptides. Peptide mapping is a critical workflow in biotherapeutic protein characterization, but due to its complexity and inherent variability, is generally performed in a comparative manner, using both mass spectrometry and ultraviolet detection. The results of peptide mapping allow not only to confirm the Amino Acid sequence, but also the N- or C-terminal clipping, that can modify the molecular weight. Post translational modifications, such as glycosylation, encompass chemical modifications of a protein during or after its translation, affecting protein activity and pharmacokinetics. The ICH Q6B requests the analysis for glycoprotein based biotherapeutics, that includes the determination of carbohydrate content, neutral sugars, amino sugars, sialic sugars, the

structure of the oligosaccharide pattern and glycopeptides. This type of analysis usually resorts to gas chromatography coupled to mass spectroscopy (GC-MS), among others: the oligosaccharide pattern, may be analysed by LC−MS and MALDI-TOF-MS [16,17].

In order to measure the biological activity, a valid biological assay, which comprises *in vitro* and/or, *in vivo*, biochemical and physicochemical assays, such as animal-based, cell culture-based and enzymatic reaction rates or biological responses induced by immunological interactions, should be provided. The quantitative measure of biological activity and potency is an important property that is used to assess the ability or capacity of a product to achieve a defined biological effect.

The immunological properties should also be fully characterized, using immunochemical procedures such as Enzyme-Linked Immunosorbent Assay (ELISA) or western-blot, when an antibody is the desired product, which allows determining the monoclonal antibody affinity, avidity and immunoreactivity, using a defined antigen.

The desired product and multiple product-related substances of the drug substance or product may be constituents when assessing for evaluation of their purity. This evaluation is a starting point to assess the presence and identity of impurities, such as process and product-related impurities. Process-related impurities in the drug substance may include contaminants from synthesis and purification processes, leachables from columns and host cell proteins, whereas, product-related impurities (molecular variants with different properties of the desired product) may include (1) truncated forms, such as hydrolytic enzymes or chemicals, detected by HPLC or SDS-PAGE, (2) modified forms, such as deamidated, isomerized, oxidized, possible detected by HPLC, capillary electrophoresis, mass spectroscopy and circular dichroism and (3) aggregates, quantitated by SEC and capillary electrophoresis [15,17].

Analytical techniques, such as Size Exclusion Chromatography (SEC), Karl Fischer Titration (KFT), SDS-PAGE, and ion exchange chromatography (IEX), are methodologies that will be further discriminated throughout this study, as they were developed and implemented for the analytical characterization of both model proteins (Bovine Serum Albumin, Lysozyme and Ovalbumin) and a monoclonal antibody, ARD_2019. Relatively to other analytical techniques discussed herein, used in the quality control of biopharmaceuticals, they are listed in an annex table (Chapter 6), presenting the testing procedures and acceptance criteria that should be considered for biotechnological /biological products.

1.2 Protein-based Biopharmaceuticals

1.2.1 Protein Structure and Function

As previously described in section 1.1.2, proteins, such as monoclonal antibodies and therapeutic peptides, belong to a wide range of biopharmaceuticals classes. This requires a general-to-particular approach to the questions: How the relationship between protein structure and function can impact and provide information about biological and molecular mechanisms? How evolution of sequence leads to diversity in protein structure and function?

In fact, proteins are the embodiment of the transition from the one-dimensional world of sequences to the three-dimensional world of molecules capable of diverse activities [18]. Proteins are linear polymers built of monomer units called amino acids, constituting the primary structure of the protein. The secondary structure is related to the folding of a polypeptide, which must fold into a specific three-dimensional structure before it can perform its biological function and termed as *protein*. The association of several polypeptides leads to, the quaternary structure of a protein. Therefore, proteins admit various sizes and shapes, presenting as fibrous proteins and globular proteins, performing functions as enzymes, transport proteins, cytoskeleton anchor, recognition, and membrane receptor.

In the protein synthesis process, the polypeptide chain is formed by a wide range of amino acids, which have an ever-equal central structure, consisting of a carboxyl group, amine group, hydrogen atom and a side chain, which is variable. Despite the particularity of properties of each amino acid, it is possible to observe some extent of similarity of the side chains between different amino acids, converging on similar properties, which can be quite useful when comparing two polypeptide chains or for decode the function of the protein by replacing one of these amino acids, which is valuable for helping deduce general function. Thus, it considers four major groups, based on the side chain structure : side chains with (1) basic groups, which include arginine, lysine and histidine, (2) acidic groups, such as aspartic acid and glutamic acid (3) polar groups but uncharged, namely asparagine , glutamine, serine, threonine, tyrosine and cystine and (4) nonpolar groups: glycine, alanine, valine, leucine, isoleucine, phenylalanine, tryptophan, proline and methionine. The binding of these amino acids, upon loss of a water molecule, and wherein the alpha-carbon binds to a nitrogen group, is called a peptide bond. This peptide association, together with other peptide interactions, such as charge-charge, hydrophobic and disulphide bonds, are the basis for all the protein structures described above, primary, secondary, tertiary and eventually quaternary [19].

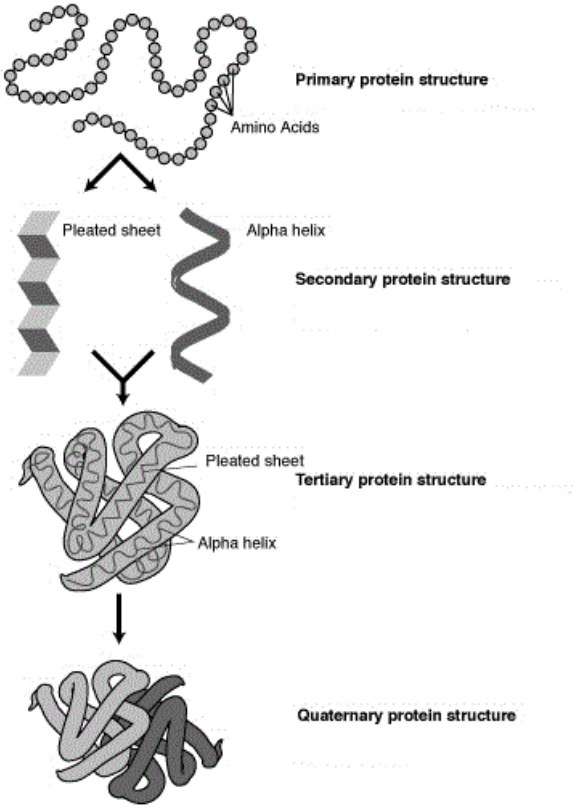

Figure 3. Dimensional Protein Structures. Adapted from N.H.G.R Institute [20]

From these structures can arise structural features, such as motifs, domains and folds. These characteristics are dependent on the rearrangement of the secondary structure of the protein in the case of motifs and how they rearrange relative to each other in three dimensions, namely the protein fold. A binding of about 100 amino acids results a sequence with specific functions, such as protein-protein interaction and DNA binding, presenting as domains. The proteins encoded in a genome, and how these genes are expressed, determine the material basis of an organism's anatomy and physiology. Diversity in protein sequence and structure can come from several factors associated with gradual changes in the gene sequence, with the accumulation of mutations and recombination events, which are consequently related to genetic mechanisms. These can generate new functionalities that arise from the emergence of multiple isoforms, during gene duplication, divergence of gene sequence through mutations and association of different gene domains. Although most mutations are deleterious and eliminated by natural selection, phenomenon associated with loss of function, no changes or new function can occur.

At the molecular level, three-dimensional structures facilitate the detection of distant evolutionary relationships, which can reveal how function is modified during evolution by sequence and structural changes and may be involved in safety assessment processes. Represented by two large databases, which classify protein structure, Structural Classification of Proteins (SCOP), and Class Architecture Topology/fold Homologous Superfamily (CATH), which uses automated computer comparisons to

classify, but also identify three-dimensional folding. Bioinformatic analysis is used to determine the structural lineage of a protein, their relationships, to provide insights on their evolutionary history, structural information, and can provide context by focusing on sequence changes in regions known to influence specificity or mode of action [21].

1.2.1.1 *Specific case of monoclonal antibodies: from structure to effector functions*

Antibody molecules, Y-shaped molecules, are constructed in the same way from paired heavy and light polypeptide chains, and the generic term immunoglobulin is used for all such proteins. Compared to glycoproteins, they only differ structurally and have different effector functions. In human serum, one of the most abundant proteins is Immunoglobulin G, accounting for about 10 - 20% of plasma protein content. In addition, immunoglobulins represent a wide range and can be grouped into classes, that can be distinguished by their constant regions, IgM, IgD, IgG, IgA, and IgE, and subclasses, in order of decreasing abundance relative to IgG, such as IgG1, IgG2, IgG3, and IgG4. Despite the similarity in the amino acid level, there are several factors that provide unique specificity to these subclasses regarding antigen binding, immune complex formation, complement activation, triggering of effector cells, half-life, and placental transport [22].

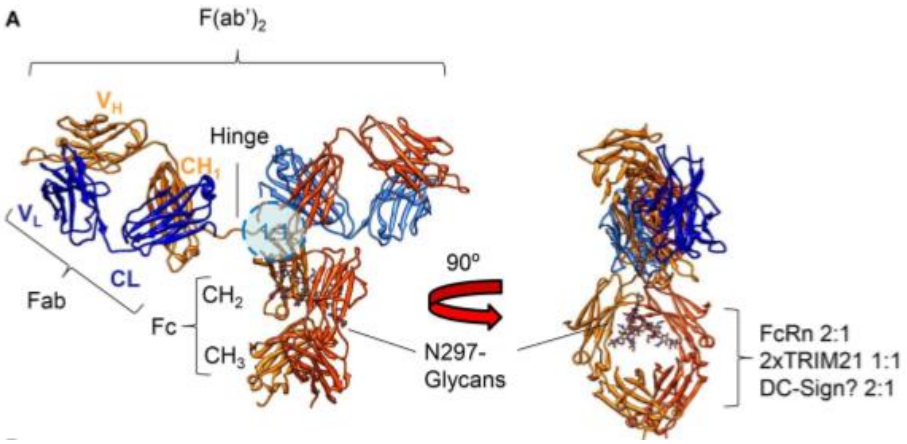

Figure 4. Crystal structure of an human IgG1 molecule. Adapted from G.Vidarsson *et al.* [22]

IgG antibodies are large molecules, having a molecular weight of approximately 150 kDa, composed of two different kinds of polypeptide chain, as shown in Figure 4. The two identical heavy chains of approximately 50 kDa, and two identical light chains of 25 kDa are binding by inter chain disulfide bonds. The similarity between the chains allows the antibody molecule two identical antigen binding sites, and thus the ability to bind simultaneously to two identical structures. The two arms of the Y end in regions that vary between different antibody molecules, the V regions, involved in antigen binding, whereas the stem of the Y, or the C region, is far less variable and is the part that interacts with effector cells and molecules. Each heavy chain consists of an N-terminal variable domain (VH) and three constant domains (CH1, CH2, CH3), that are numbered from the amino-terminal end to the carboxy terminus,

with an additional “hinge region” between CH1 and CH2. The light chains consist of an N-terminal variable domain (VL) and a constant domain (CL). The light chain associates with the VH and CH1 domains to form a Fab arm (fragment antigen binding). The part of the antibody formed by the lower hinge region and the CH2/CH3 domains is called Fc (fragment crystalline). The Fc fragment and hinge regions differ in antibodies of different isotypes, thus determining their functional properties [22,23].

Antibodies correlate the mechanisms between the adaptive immune system and the innate immune system by combining antigen binding sites and many innate receptors and adapter molecules. These action mechanisms are directly related to the immunoglobulin subclasses, being differentiated for each one of them. Despite being able to neutralize virus particles and toxins, IgG2 and IgG4 induce less aggressive responses, while IgG1 and IgG3 are potent triggers of effector mechanisms.

Antibodies, a secreted form of B-cell receptor, acquire similar conformation of the receptor, except for a small portion of the C-terminus of the heavy chain constant region. One of the parameters that differentiates antibodies from the B-cell receptors, it’s their hydrophobicity, in the case of antibodies, the C-terminus is a hydrophilic sequence and, in the C-terminus of B cell receptor, features a hydrophobic sequence. This hydrophilic sequence, provides high solubility and allows their secretion in large quantities , making them easily obtainable and easily studied [22,23].

Throughout this thesis, a monoclonal antibody, herein designated by ARD_2019, was used as case study. All the methodologies developed during this work were applied in the analytical characterization of this antibody.

Chapter 2

Materials and Methods

This chapter describes all the methodologies performed and materials used throughout method development and analytical characterization of the biomolecules under study.

2.1 Biomolecules

The development of the techniques described in the present work was performed with model proteins, such as Bovine Serum Albumin (BSA), Ovalbumin and Lysozyme, and a monoclonal antibody, ARD_2019.

Bovine serum albumin (BSA) is a globular protein with 66 kDa and an isoelectric point of 4.7, that is used in numerous biochemical applications due to its stability and lack of interference within biological reactions [24]. Categorized as an albumin, Ovalbumin it is a heat stable form of egg albumin formed when eggs are stored. Chicken egg albumin is a phosphorylated-glycoprotein and the major protein constituent of egg whites. It has 44.3 kDa and an isoelectric point of 4.5, adopting a serpin-like structure. Lysozyme is thermally stable, with a melting point reaching up to 72 ^{o}C at pH 5.0. Its isoelectric point is 11.4. Lysozyme can survive in a large range of pH (6-9) [25]. The model proteins were obtained from Sigma Aldrich, with a part numbers of 9048-46-8, 9006-59-1 and 12650-88-3, corresponding respectively to the proteins BSA, Ovalbumin and Lysozyme.

All these proteins were in the lyophilized state, whereby the BSA protein used in the development of the Size Exclusion Chromatography (SEC) method was dissolved in water Milli-Q, before use, being used at a final concentration of 2 mg/mL. In the degradation studies, this protein was also dissolved in water Milli-Q, at a final concentration of 1 mg/mL.

ARD_2019 protein is a monoclonal antibody. The isoelectric point of ARD_2019 is 8.2 and its molecular weight is around 122 kDa. The monoclonal antibody product was handled as a liquid solution at a concentration of 15mg/mL and as a dried powder containing about 20% (w/w).

2.2 Development of a Karl Fischer Titration Method

2.2.1 Introduction

Karl Fischer titration (KFT) is the most widely used technique in the pharmaceutical industry for water determination, due to its high sensitivity and selectivity. In this type of analysis, a stoichiometric reaction (first step in the sulfur-iodine cycle), occurs between water (H_2O) and iodine (I_2), in the presence of sulphate (SO_2), and the products obtained, sulfuric acid (H_2SO_4) floating on top and a mixture of hydrogen iodide (HI) and unreacted iodine on the bottom, represent the modified Bunsen reaction (Equation I).

$$2H_2O + SO_2 + I_2 \rightarrow H_2SO_4 + 2HI \qquad \text{(I)}$$

Water determination in the characterization of proteins is a critical parameter, since it can trigger several reactions, being able to induce degradation of drug compound, cause undesired crystalline phase transitions, as well as lower glass transition temperatures for amorphous lyophilized products. Additionally, water promotes microbiological activity, playing a critical role in physical and chemical stability, and, consequently, the safety of the drug product, could be compromised.

The volumetric and coulometric titrations are representative of the two classes that define the KFT methods, differing in the way iodine is supplied. While in coulometric KFT iodine is formulated *in situ* at the anode, electrolytically generated, while in volumetric KFT the iodine is formed in the titration reagent.

In these two types of titrations, samples can be introduced directly as solids, liquids or pre-dissolved solutions. Recurring problems associated with sample preparation are the low solubility in alcohols, particularly in the case of proteins, and any side reactions with KF reagents to release water or consume iodine, falsifying the results.

Alternatively, a heating method can be used (KF Oven), wherein water is thermally released, and the resulting water vapor is transferred by inert gas stream to the titration cell, mitigating the low solubility, unwanted side reactions and matrix effects, since it is only the water that is in contact with the titration cell and the sample itself does not come into contact with the KF reagent [26,27].

To determine the process residual water content of the proteins under study, a KF oven method was developed.

2.2.2 Materials and Methods

In the present work, a KF oven (874 Oven Sample Processor and 801 Stirrer), purchased from Metrohm, method was developed to analyse process residual water, using the model proteins (BSA, Ovalbumin

and Lysozyme). The quantification of process residual water of these proteins was carried out by KF oven, rather than by other KFT processes, as a result of the low solubility of proteins in most common KF solvents.

To perform this procedure, 20 mg of each protein were weighted in a glove box, with a relative humidity between 7 and 8%, since these proteins are hygroscopic. The samples were weighed out directly into vials that are then hermetically sealed, with PTFE-coated septa (VWR International), and placed in the oven apparatus, preventing the water content from being falsified by atmospheric moisture. With the aim to select the optimum oven temperature to determine the content of process water present in proteins, a oven temperature ramp between 50 ºC and 220 ºC was define. The results obtained are presented in section 3.1.1 of Results and Discussion chapter.

To verify if the equipment is properly working, three (3) samples of a Lactose Standard at 5% (w/w) H_2O (STD) (Merck) were weighted and analysed at 150 ºC.

After confirming the equipment is properly working, control samples (empty vials) were also analysed, prior to each sample measurement, at the same temperature defined after temperature ramp assignment (120 ºC).

In order to demonstrate the repeatability of the method, six samples of BSA were then weighted (10 mg), in a glove box at a relative humidity between 7 and 8%, and the vials were then placed in the oven apparatus.

To verify that the equipment is not saturated with water after the analysis of all samples, another lactose standard was weighed and analysed at 150 ºC. For all the Lactose Standards analyzed it was evaluated if the water content was within, 5.10% (w/w) ± 0.20% (w/w). Table 1 summarizes the Karl Fischer Oven analysis sequence during method's development.

Table 1. Method conditions and procedure of Karl Fischer Oven technique

Karl Fischer Oven Analysis Sequence	Method Procedure
Equipment Performance Evaluation	Lactose Standard at 5%(w/w) H_2O, analysed at 150 ºC
Step 1	Analysis of the model proteins performed with a temperature ramp between 50 ºC and 220 ºC
Step 2	Repeatability assessment by analysing six (6) different Bovine Serum Albumin samples preparations using a temperature of 120 ºC
Equipment Performance Evaluation	Lactose Standard at 5%(w/w) H_2O, analysed at 150 ºC

After establishing the conditions that should be used to determine the process residual water content of proteins, as shown in Table 1, the same conditions were used for the determination of the water content of ARD_2019 mAb. It was weighed 10 mg of the ARD_2019 sample, in a glove box at ≤ 5% Relative Humidity (RH).

For both model and ARD_2019 proteins, the water content results were analysed by the software indicated, in order to quantify process residual water, with results in sections 3.1.1 and 3.2.1 of Results and Discussion chapter.

2.3 Size Exclusion Chromatography Methodologies

2.3.1 Introduction

The complex structure and large size of proteins hamper the characterization of protein-based products, necessary to guarantee their efficacy and safety. [28]

The change in protein higher structure relatively with physical modifications of protein-based biopharmaceuticals, as being one of the most common issues, can induce a permanent (or transient) partial unfolding and aggregation. Smaller molecules and aggregates, a multimeric high molecular weight forms in proteins, considered as a process or product related impurities by the regulatory guidelines [15,29], can affect protein activity and solubility and may trigger undesired immune reactions [28,30].

Size Exclusion Chromatography (SEC) is a reference technique for characterizing proteins based in their molecular weight. It is a widely used methodology in the qualitative and quantitative evaluation of protein mixtures. It separates biomolecules according to their hydrodynamic diameter, by interacting with a porous structure. SEC can also be used for the routine monitoring of aggregates in biopharmaceuticals [28]. The elution order of the molecules, obtained as a chromatographic profile, typically follows an elution where large molecules, excluded from the pores, elute first and smaller molecules, that permeate a larger accessible volume within the column, since these are capable to access pores within the resin particles, elute later. The same is expected to occur with protein aggregates and other high molecular weight proteins, larger when compared with the size of the native protein, eluting earlier. [31]

In order to obtain a higher chromatographic performance, it is crucial to evaluate some parameters based on stationary phase and mobile phase principles. The stationary phase consists of spherical porous particles, controlled pore size and pore size distribution and to avoid physico-chemical interactions with the protein, requires the use of an inert material, in order to minimize the adsorption of

proteins and protein aggregates. Most common stationary phases are hydrophilic and consist of silica chemically bonded to 1,2-propanediol functional groups, in which the acidic silanol groups are derivatized. This type of stationary phase is commonly used for the separation of biopolymers and synthetic water-soluble polymers. Conversely, this stationary phase is not advise for separation of basic polymers, since it will ion exchange onto residual acidic silanol groups or even diol-modified silica, being able to block or react with most of these acidic groups and neutralize the surface. Adsorption of the proteins to the stationary phase can also be impacted by the mobile phase constitution.

Silica material presents higher versatility and it is advantageous for the analysis of non-ionic polymers with organic mobile phases, but due to its solubility in aqueous buffers it is not recommended. In order to mitigate this situation, a new material, ethylene-bridged hybrid inorganic–organic (BEH), has emerged, which reduces the reactivity and acidity of silanol and improves chemical stability.

Recently, small particle size packing materials, sub-2 mm SEC columns, have been introduced in the market. Liu *et al*. demonstrated that sub-2 mm columns provide improved chromatographic resolution and higher efficiencies when compared with larger particle sizes columns that commonly lead to inaccurate quantitative results [30, 32]. Additionally, these columns are used in low dispersion UHPLC (Ultra High Performance Liquid Chromatography) systems that allow higher pressures and reduced analysis time.

SEC analysis can be characterized as a two-stage process in which protein samples are separated and then revealed using different detectors, such as UV, the most used approach for quantitative determinations, Refractive Index (RI), Mass Spectroscopy (MS), Viscometer and Static or Dynamic Light Scattering (SLS and DLS). Although the combination of mass spectrometry coupled with SEC separation could be a great advantage for this technique, as it could provide simultaneous identification and quantification of size variants, the type of mobile phases commonly employed (non-volatile buffers and salts) make it less suitable for MS detection. In order to mitigate this challenge, in the last years, alternative elution solvents, such as denaturing mobile phases containing organic solvents, ion-pairing reagents (such as Trifluoroacetic Acid TFA) or charge modifiers have been used. To perform a complete characterization of protein samples, their structural variants, fragments or aggregates, and to get information of the protein purity and impurities, a combination of SEC analysis allied with UV and MS detection are the most recurring tools [28,30].

2.3.2 Materials and Methods

A SEC method was developed using a gel filtration standard, purchased from BioRad, containing the proteins Thyroglobulin (670 kDa), IgG (150 kDa), Ovalbumin (44 kDa), Myoglobin (17 kDa), and Vitamin B12 (14 kDa). This was reconstituted with 500µL of water (Milipak Gold 0.22 µm, purchased from Merck), diluted to 5 mL with water and then 100 µL were transferred to HPLC vials with an insert. It was also weighed 10 mg of 0.1 mg/mL 4-aminobenzoic acid (PABA) (0.14 kDa), purchased from Sigma Aldrich, and transferred to a 15 mL conical centrifuge tube, adding 10 mL of water and vortex to mix. After a 1:10 dilution with water Milli-Q, the PABA solution was also introduced into an HPLC vial, making

up to 3 quarters of the volume. Bovine Serum Albumin (BSA), purchased from Sigma Aldrich, was also used at a concentration of 2 mg/mL in water, in order to condition the column.

An Alliance HPLC was used with a TSKgel G3000SWxl (TosohBioscience) column, 30 cm×7.8 mm, 5 µm, 250 Å, from Tosoh Biosciences. The mobile phase was 92.5% of 25 mM sodium phosphate monobasic (NaH_2PO_4; Sigma Aldrich), 350 mM sodium chloride (NaCl; Sigma Aldrich) and 7.5% of 2-propanol (Merck) (v/v), at a pH of 6.8 (adjusted with sodium hydroxide, NaOH; Honeywell, Fluka). After preparation, the mobile phase was filtered (Nylon membrane filter 0.2 µm, 47 mm, Pall Corporation). The flow-rate was set to 0.8 mL/min, according to the diameter and the length of the column, which allow the time to pass through the storage solution and conditioning of the column with mobile phase, relative to a stipulated volume with respect to the storage 10 times the volume of the column. These volumes are determined by the Equation II:

$$V = \pi r^2 l \qquad \text{(II)}$$

where,

V is the column volume;
r^2 corresponds to the column radius (expressed in cm); and
l is the column length (expressed in cm).

The gel filtration standard, BSA and PABA, already introduced into HPLC vials, with the autosampler temperature set to 2-8 °C, were injected (10 µL) into the column (maintained at ambient temperature) and separated by running mobile phase for 40 minutes. The elution of the proteins and the organic compound from the column was monitored at a wavelength of 220 nm.

Based on chromatographic profile obtained for the proteins presented in the gel filtration standard, it was obtained a calibration curve of the logarithm of the molecular weight of the proteins (in kDa) as a function of their practical retention times (in minutes).

Regarding to ARD_2019 characterization by SEC, all the conditions of the method were maintained. The ARD_2019 sample preparation (18 mg) was performed at a relative humidity ≤ 5%, as it is a hygroscopic product, and in which theoretically 10 mg of the sample contained 120 µL of water, having been added 222 µL of the mobile phase (92.5% of 25 mM sodium phosphate monobasic (NaH_2PO_4) and 350 mM sodium chloride (NaCl) and 7.5% of 2-propanol (v:v)). From this diluted sample, an additional 1:3 dilution in water (ARD_2019 final concentration = 0.12 mg/mL) was performed and the sample solution was then transferred to an HPLC vial with an insert, counting on an injection volume of 10 µL. Elution of the protein was also monitored at a wavelength of 220 nm.

Based on the data obtained from the ARD_2019 chromatographic profile, it was possible to determine the molecular weight, from the equation obtained for the standard proteins, as well as analyzing high molecular weight proteins and impurities, as presented in section 3.2.2 of results and Discussion chapter.

In order to evaluate the sensitivity of the size exclusion chromatography method to detect small variations related with impurities, aggregates or proteolysis, a degradation study was also performed. Two different stress conditions, pH and temperature, were tested. For this purpose, the protein selected for the study was BSA (1 mg/mL). The study was performed at 2 different conditions: (1) 0.3 mg/mL of BSA subjected to 80 °C for 30 minutes, and (2) 0.3 mg/mL of BSA with the addition of 0.5% of formic acid (CH_2O_2), purchased from Merck. A control sample was also prepared with 0.3 mg/mL of BSA subjected to any stress condition (in its native form). After degradation, fresh BSA protein samples at the same concentration (0.3 mg/mL) were spiked with degraded protein already centrifuged for 5 minutes at 13.4 rpm, at the levels of 0.5% (w/w), 1% (w/w) and 2% (w/w). Additionally, three control solutions were prepared at 2% (w/w) of BSA for each stress degradation condition (final volume was 300 µL). Samples were then centrifuged prior to transfer to HPLC vials with inserts of 150 µL, and analysed by size exclusion chromatography at the same conditions as previously described in this section.

2.4 Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis

2.4.1 Introduction

Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE) is widely used to analyse proteins in complex extracts. This technique is used to separate proteins according to their molecular weights.

In gel electrophoresis, an electric field is used to move charged molecules through a matrix of a polymerized substance such as polyacrylamide, which is fueled by a battery, causing the proteins to migrate through the gel, from the negative pole to the positive pole. The rates of migration at which individual molecule moves through the gel matrix is dependent on the properties of the separation system and the structure and the charge of the proteins. The mobility of the molecules is related to the amount of resistance that the matrix presents, and in the diameter of the pore, size and geometry of themselves. The buffer system and the strength of the electrophoretic field used for the separation are the other two parameters that can affect the mobility of molecules. In view of the migration of molecules, smaller molecules, more highly charged, migrate more rapidly through gels than larger or less charged molecules.

Polyacrylamide gels are used to separate proteins, since it is an inert material and consequently does not trigger any interaction with the proteins. It is made from the chemical polymerization of acrylamide and a cross-linking monomer reagent, *N,N'* - Methylenebisacrylamide. Two other intervening agents are added to the process, allowing the polymerization of the polyacrylamide, upon reaction of the first catalyst, ammonium persulfate (APS), with the second catalyst, *N,N,N',N'*-Tetramethylethylenediamine

(TEMED), generating oxygen radicals. These free oxygen radicals react with the vinyl groups in acrylamide and bisacrylamide, completing the polymerization process, that forms a mesh-like matrix suitable for the separation of proteins.

In denaturating SDS-PAGE, proteins are separated solely based on their polypeptide chain length. The denaturing conditions of the sample preparation are achieved by adding SDS (Sodium Dodecyl Sulphate), anionic detergent with a strong protein-denaturing effect, with the combination of heat, since the proteins are denatured by boiling them with this detergent and the reducing agent (2-mercaptoethanol), breaking the many noncovalent bonds that stabilize protein folding, and cleaving the disulfide bonds in the proteins, respectively. In the presence of this solution, proteins unfold (in which the tertiary and secondary structure are converted into its primary structure) into linear chains with negative charge proportional to the polypeptide chain length.

In order to obtain a high resolving power of SDS-PAGE gels, the Laemmli (1970) SDS-PAGE system has two additional components, such as stacking gel and running gel. These have different pores sizes, characteristic that can be manipulated by the acrylamide concentration, ionic strengths and pHs. The electrophoresis buffer, referred to as the third component, is constituted mainly by glycine. In aqueous solution, glycine can adopt 3 forms, depending on whether the pH is acidic (protonated form, pKa~2.3), neutral (negatively charged carboxyl group and a positively charged amino group, pI = 5.95) and basic (negative charged, pka~ 9.6). The pH of the running buffer (alkaline solution) is closer to the pKa of the glycine, so a significant fraction of the glycine molecules assumes a negative charge. When proteins cross the stacking gel (pH = 6.8), these conditions are altered by the acidic pH, causing them to migrate more slowly than the chloride ions and the SDS-protein complex, being positioned at the interface between glycine and chloride ions. Regarding this, the pores in the stacking gel are larger than those of the running gel, so that the migration of all proteins occurs at the same time. When the glycine reaches the running gel (pH = 8.8), acquires a much more negative charge, given the alkalinity of the solution; in this way, migration to the positive pole occurs by the following order: chloride ions, glycine and, finally, the SDS-protein complex. The smaller pores present frictional resistance to the migration of proteins, being able to separate the proteins according to their molecular weight. [33] [34]

2.4.2 Materials and Methods

To confirm the results obtained by SEC (as per section 3.1.3.1), the degradation samples were also analysed by Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis. The BSA protein was again chosen for further comparative analysis of these two techniques and was prepared as described in section 2.3.2 and subjected to the same stress conditions, pH and temperature.

Additionally, the purity and the characterization of the various fragments of the monoclonal antibody ARD_2019 was also performed by SDS-PAGE, in reduced conditions.

For both studies, electrophoresis was performed as previously described (Laemmli, 1970) [35], on a 12% (v/v) polyacrylamide gel in 2.5 M Tris-HCl, pH 8.8, containing 10% SDS (m/v) (Mini Protean TGX Precast Gels, purchased from BioRad). The stacking gel had a polyacrylamide concentration of 5%

(m/v) in 0.5 M Tris-HCl, pH 6.8.

Firstly, 3 dried powder samples of ARD_2019 (referred in section 3.2.3) were prepared in a glove box, with a relative humidity under 5%. After weighting and then diluting with water Milli-Q, purchased from Merck, the samples were at a concentration of 83 mg/mL. To match the concentration of the dried powder samples with the fresh monoclonal antibody, which was at a concentration of 15 mg/mL, a 1:3 dilution with H_2O was made, obtaining a final concentration of 27.7 mg/mL (15 mg/mL of antibody and the remainder being the excipients that support the formulation). Given the high concentration of the antibody, a dilution was performed to a final concentration of 2 mg/mL.

To all ARD_2019 samples previously prepared it was necessary to add the denaturation solution, sample buffer, prepared with a total volume of 20 mL (5 mL of stacking gel, 8 mL 10% SDS, 1 mL β-mercaptoethanol, purchased from Sigma Aldrich, 2 mL Gycerol, purchased from Acros, and 4 mg of Bromophenol blue, purchased from BioRad), and was added in the same ratio of each sample, total volume of 10 µL, in a proportion of 1:1 to protein mixture. The standard protein mixture with a range of known molecular weights between 10 kDa and 250 kDa, used in the SDS-PAGE analysis gel performed for the degradation studies, and a standard protein mixture with known molecular weights of 17 kDa, 44 kDa, 158 kDa and 670 kDa, used to ARD_2019 gel electrophoresis, both purchased from BioRad, were prepared in the same way.

After the addition of the sample buffer, and in order to speed up the denaturation process, the samples were introduced in a dry bath for 10 minutes at 95 °C and then were subjected to a short spin in the centrifuge. Samples and the standard proteins mixture were loaded (10 µL each) with a pipette using gel loading tips, with special attention to avoid cross contamination between different samples in the different gel wells. This way, the electrophoresis started at constant 120 V, 400 mA, for 70 minutes. After electrophoresis was completed, the running buffer was discarded and was removed the gel cassette and washed with water Milli-Q for the subsequent gel colouring, with the staining solution Coomassie Blue R-250 (purchased from Sigma Aldrich), for 30 minutes in the stirrer plate. Finally, the staining solution was removed and added the detaining solution (7.5% (v/v) glacial acetic acid and 45% (v/v) methanol) in the gel for 30 minutes in the stirrer plate. The gel was kept in this solution until the analysis.

The results obtained for the electrophoresis gel for the degradation studies were intended to evaluate and to confirm the sensitivity of the method of size exclusion chromatography, described in section 2.3.2, to detect small variations in the level of impurities, aggregates or proteolysis. The electrophoresis gel performed for ARD_2019 allowed to evaluate the purity of the antibody, confirming its correct structure with respect to heavy chain and light chain, comparatively to the theoretical structure of monoclonal antibodies, as described in section 1.2.1.1 of Introduction chapter.

2.5 Ion Exchange Chromatography

2.5.1 Introduction

Ion Exchange Chromatography (IEX) is one of the most versatile and broadly applicability technique for the detailed characterization of therapeutic proteins and allows for the separation of ionizable molecules by evaluating the charge heterogeneity. [36,37]

Biomolecules, such as proteins, vary in their charge properties, displaying different types of interaction with charged chromatographic media in relation to their overall charge, charge density and surface charge distribution. Proteins are multivalent anions or cations and built up on their surface of charged amino acids, containing weak acidic and basic groups, and there may be changes in their net surface charges, since possess different *pKa* values (acid ionization constant). The specificity of each protein relatively to the relationship between net charge surface and pH of the medium can be demonstrated as a titration curve as shown in Figure 5. [38]

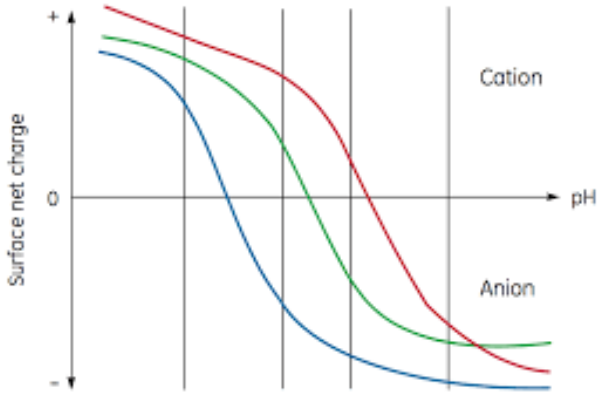

Figure 5. Net surface charge vs pH of theoretical proteins titration curves. Adapted from G.E. Healthcare [38]

A protein that has no net charge at a pH equivalent to its isoelectric point (pI) will not interact with a charged resin. However, at a pH above its pI, a protein will bind to a positively charged ion exchange chromatography or anion exchanger. Oppositely, at a pH below its pI, a protein will bind to a negatively charged medium or cation exchanger.

The anion exchange chromatography appeal to the positively charged functional groups, Diethylaminoethyl (DEAE) and quaternary ammonium (Q), while negatively charged carboxymethyl (CM), sulfomethyl (S), and sulfopropyl (SP) groups are typically cation exchangers. These two formats

of Ion Exchange Chromatography arise the state of equilibrium of counterions between the mobile and stationary phases.

IEX comprises both mobile and stationary phases and ion exchange medium constituted for an inert organic matrix of spherical particles, carry charge ionic groups, negatively or positively charged, providing a high internal surface area. The matrix is packed into a column to form a packed bed and then equilibrated with aqueous buffer, and when equilibrium is reached, all stationary phase charged groups are bound with exchangeable counterions, such as chloride or sodium. This first step assures that proteins of interest, when sample is loaded, bind to the medium, being guaranteed by the selection of pH and ionic strength of the start buffer. Proteins are then eluted by increasing the ionic strength (salt concentration) of the buffer or by a pH gradient. Thus, the order of elution of proteins are directly related to a lower or higher net charge, proteins with the highest charge at a certain pH will be most strongly retained and will be eluted later. The higher the net charge of the protein, the higher the ionic strength that is needed for elution. Upon realising protein and exchanger counterions into the mobile phase, and in order to avoid variation in matrix and protein net charge, the constant maintenance of the mobile phase pH is a critical parameter with regards to pH fluctuations.

The selection of an appropriate ion exchange matrix, with the desired ion exchanger charge and strength, the suitable exchanger functional groups and the selection of a suitable mobile phase (pH and ionic strength) should be adjusted to achieve the optimum target protein adsorption and elution [37,39].

IEX was used to determine the acidic and basic species as well as the different charged variants of the antibody ARD_2019, using a pH gradient and salt gradient respectively.

2.5.2 Materials and Methods

The work performed for the evaluation of charge heterogeneity of the ARD_2019 was supported by Cation Exchange Chromatography using either pH and salt gradients. For the development of this Ion Exchange Chromatography, Protein Pak Hi Res CM, 7µm, 4.6 ˣ 100 mm (Waters) column was used. The first step was the preparation of the equipment and the mobile phases, since the samples were diluted in mobile phase A. Relatively to the pH gradient, the optimized method encompassed a mobile phase A (10 mM Tris buffer, pH 7.5) and mobile phase B (10 mM Tris buffer, pH 10.5) and regarding to the salt gradient it was used the mobile phase A, and adding to mobile phase B, 300 mM NaCl, where for both the pH values (7.5 and 10.5), it was adjusted with HCl, purchased from PanReac, and NaOH, purchased from Honeywell, Fluka, respectively. Tris buffer was purchased from BioRad. Throughout the method development process, all mobile phases were filtered to avoid any contamination and to reduce the amount of salts present, in the case of a salt gradient. Regarding the column that was used, the storage solution and several water passages are essential to prevent any kind of microbiological growth and to avoid salt precipitation in the column, compromising the performance of the column and the analytical results. The mobile phases were filtered using a Nylon membrane filter 0.2 µm, 47 mm, from Pall Corporation, before being placed in the equipment. Additionally, the storage solution was prepared: 0.05% (w/w), sodium azide in high purity water purchased from Sigma Aldrich, which inhibit undesired microbial growth.

The preparation of the sample was performed in a glove box with relative humidity under 5%, where it was weighted 20 mg of the sample. As previously mentioned, 10 mg of ARD_2019 sample presents 120 µL of water, and was diluted with 238 µL of mobile phase (10 mM Tris buffer, pH 7.5). Finally, from this diluted sample, a new 1:3 dilution in mobile phase was performed: 150 µL of this solution were then transferred to an HPLC vial with an insert, counting on an injection volume of 10 µL, and the elution of the protein was monitored at a wavelength of 220 nm.

For each elution, the column was pre equilibrated with, at least, three column volumes of mobile phase A; prior to sample injection, water washes were carried out through the column as salts may precipitate and the column clogging. After the injection of the monoclonal antibody sample onto the column, a linear increasing in the percentage of mobile phase B was delivered. The linear gradient at a flow rate of 0.8 was started with 100% of mobile phase A, increasingly linearly to 100% of mobile phase B, in 25 minutes, and kept for 5 minutes; the percentage of mobile phase A was then increased to 100% in the preparation for subsequent analysis, summing up a 40 min of run, as shown in Table 2.

Table 2. Gradient Conditions for salt and pH gradient in Cation Exchange Chromatography

Condition	Mobile Phase	Gradient
Salt gradient	A: 10 mM Tris Buffer, pH 7.5 B: 10 mM Tris Buffer + 300 mM NaCl, pH 7.5	100% A: 0 min 65% A and 35% B: 25 min 65% A and 35% B: 30 min 100% A: 40 min
pH gradient	A: 10 mM Tris Buffer, pH 7.5 B:10 mM Tris Buffer, pH 10.5	100% A: 0 min 100% B: 25 min 100% B: 30 min 100% A: 40 min

Chapter 3

Results and Discussion

In this chapter are presented the most relevant results obtained by the analytical characterization of the proteins under study, model proteins and ARD_2019.

3.1 Model Proteins Characterization

3.1.1 Karl Fischer Oven

The development of a Karl Fischer Oven method, as described in section 2.2.2 of Materials and Methods chapter, had as purpose the quantification of the process residual water present in the proteins under study. This was the method chosen over other titration processes, such as volumetric KFT, as a result of the low solubility of proteins in of the common KF solvents.

The implementation of a temperature ramp was performed between 50 ºC and 220 ºC, as shown in Figure 6. Based on the profiles obtained, it was possible to conclude that the same temperature range returned similar profiles for Bovine Serum Albumin, Ovalbumin and Lysozyme.

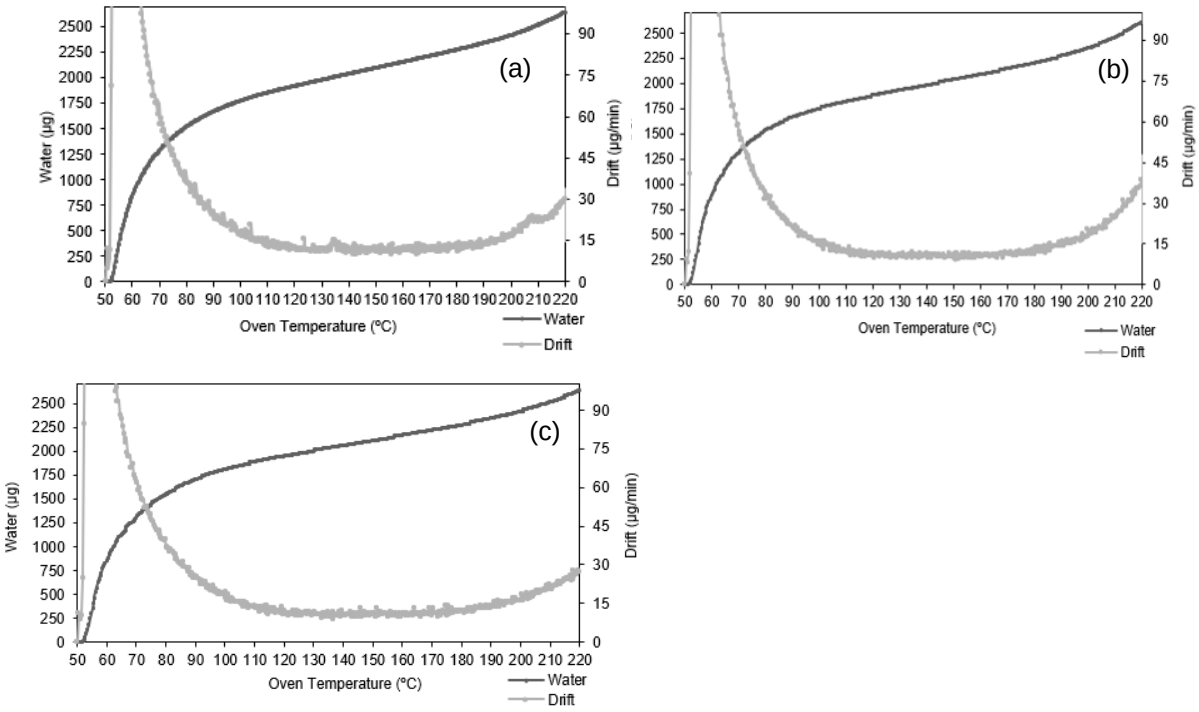

Figure 6. Karl Fischer Oven (874 Oven Sample Processor and 801 Stirrer) method development. Temperature ramps of model proteins. The standard used for temperature ramp analysis (50 ºC and 220 ºC) was Lactose Standard at 5% (w/w) H_2O, analysed at 150 ºC. (a) Lysozyme, (b) Bovine serum albumin, (c) Ovalbumin

The temperature ramps performed for the model proteins were intended to infer the optimal temperature for determining process residual water. Similarity of profiles for each of the temperature ramps was verified for all model proteins, which led to a temperature set of 120 ºC with a lactose system at 150 ºC, based on the drift that each protein presented when subjected to the temperature range. Optimum

temperature selection was made due to the absence of oscillations at this point of temperature, since it is indicative, by the results obtained, that the release of structural water of the proteins occurs from temperature higher than 180 ºC, at which the drift increases. In addition, 120 ºC is close to the water release temperature (boiling point at 100 ºC), preponderant to the release of process residual water.

The repeatability assessment was performed by analysing six different BSA sample preparations. (Table 3).

Table 3. Repeatability method by Karl Fischer Oven of six Bovine Serum Albumin samples. The standard used for water content analysis (120 ºC) was Lactose Standard at 5% (w/w) H_2O, analysed at 150 ºC

Sample	Water Content (µg)	Sample Size (mg)	Water Content [(% (w/w))]
BSA 1	839.83	11.12	7.01
BSA 2	840.94	11.05	7.07
BSA 3	896.38	11.89	7.03
BSA 4	865.65	11.71	6.88
BSA 5	802.43	11.11	6.68
BSA 6	809.92	11.43	6.56

From the individual values obtained for each of the BSA samples preparation, their average was calculated, obtained a process water content of 6.87% (w/w). The results and average obtained also allowed the calculation of the %RSD (Relative Standard Deviation), with a value of 3. According to pharmacopeia guidelines, the acceptance criterion for the %RSD is ≤ 10%, which proves the method is repeatable.

To ensure reliable results in the determination of BSA water content, it was necessary to analyse the results for the lactose standards, in order to verify their compliance with the established acceptance criterion, 5.10% (w/w) ± 0.20% (w/w), as shown in Table 4. The values obtained clearly demonstrate that the equipment is returning precise results, and therefore the results obtained for the determination of the water content in the BSA protein can be considered exact values.

Table 4. Results obtained from the Lactose Standard at 5% (w/w) H_2O system (STD1, STD2, STD3 and STD4), analysed at 150 ºC by Karl Fischer Oven. Acceptance criterion: 5.10% (w/w) ± 0.20% (w/w).

Standard Injection	Water Content (µg)	Water Weight (mg)	Water Content [(%(w/w))]	Water Content [(%(w/w))] Average	% RSD	% RSD Acceptance criteria
STD 1	2645.02	50.79	5.0914			
STD 2	2584.27	48.99	5.1545	5.1136	0.7	≤ 3%
STD 3	2590.57	49.37	5.1276			
STD 4	2595.58	49.92	5.0811			

The nature of the samples is a key parameter as it may influence the reliability of the results given their level of hygroscopicity. The titration process performed by KF Oven is an advantage, especially as regards to the amount of sample used. A study for the determination of the total water content in inulin using the volumetric KF titration, corroborates the advantage of the Oven technique [40]. The aim of this study was to determine the water content of inulin, using the conventional method of introducing the powder directly in the methanol base-working medium, or a medium modified by formamide (solvent used for solubilizing samples in working medium). In the development of the volumetric KF method, a 250 mg of inulin was used. This value, substantially higher when compared to the amount of BSA sample used in the present development, which was about 12 mg. The results obtained for the determination of water content by the conventional method was 1.3% (w/w), while using the formamide solvent, this percentage increased to 2.1% (w/w), which can be explained by the amount of water that is released from the formamide.

3.1.2 Size Exclusion Chromatography

The size exclusion chromatographic method, as described in section 2.3.2 of Materials and Methods chapter, was optimized by adjusting factors such as flow rate, column length, mass load, and volume load, which can impact in the resolution level, analysis time and sensitivity. In order to analyse the chromatographic profile of the proteins present in the standard filtration gel, Thyroglobulin (670 kDa), IgG (150 kDa), Ovalbumin (44 kDa), Myoglobin (17 kDa), and Vitamin B12 (14 kDa), a TSKgel G3000SWxl column (30 cm×7.8 mm, 5 µm, 250 Å) was used, being the chromatographic profile obtained, as shown in Figure 7. The column chosen has a particle size of 5 µm and a pore size of 25 nm, important factors for better peak shape, peak sensitivity, and resolution. Efficiency is derived from the theoretical plate model of chromatography. A plate refers to one complete equilibrated transfer (or partition) of a solute between the mobile and stationary phases. Efficiency is a qualitative term used to measure the number of theoretical plates in each column, or the degree to which an analyte partitions between the mobile and stationary phases. In relation to particle size, efficiency is inversely proportional, and with a lower particle size, efficiency increases, and more resolution is achieved. Aiming to mitigate secondary interactions (adsorption between the analyte and packing material) and electrostatic interactions, the mobile phase of 25 mM sodium phosphate monobasic and 350 mM sodium chloride

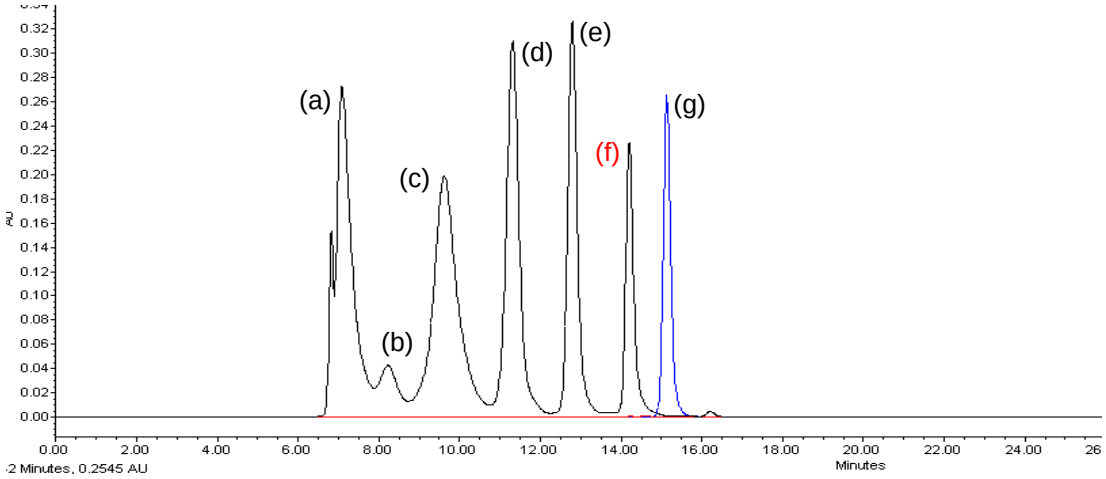

Figure 7. Size Exclusion Chromatographic profile of standard gel filtration with method conditions optimized: 25 mM sodium phosphate monobasic and 350 mM sodium chloride, pH 6.8. The injection volume was 10 µL. The flow rate was 0.8 mL/min and monitored at a wavelength of 220 nm. The column used was TSKgel G3000SWxl at a temperature of 25 ºC; Autosampler Tray Temperature: 2-8 ºC. (a) Thyroglobulin (670 kDa), (b) IgG aggregate, (c) IgG (150 kDa), (d) Ovalbumin (44 kDa), (e) Myoglobin (17 kDa), (f) Vitamin B12 (14 kDa), (g) PABA (4-aminobenzoic acid) (0.14 kDa).

was used at a pH 6.8, considering that the pH stability of the column is between 2.5 and 7.5 [41]. It is important to highlight that these values were optimized through method development, in which different concentrations of salt and organic modifier were tested. The pH and composition of the mobile phase can cause conformational changes, dissociation, or association of protein complexes, resulting in changes in non-ideal interactions with the stationary phase, which can be predicted based on the

relationship between mobile phase pH and the isoelectric point of the proteins [42].

As already mentioned, each of the proteins eluted at a retention time as a function of their previously known theoretical molecular weight. For most molecules, maximum resolution is obtained with a long column and a low flow rate. Based on the theoretical molecular weights of the model proteins, a calibration curve of the logarithm of the molecular weight was made as a function of the practical retention times, not being included in this calibration curve Vitamin B12, since Vitamin B12 does not have the same structure as a globular protein and therefore was considered as an outlier, and PABA, since it only served to evaluate the performance of the column (Figure 8). Table 4 shows the retention times of each protein obtained from the chromatographic profile as well as their molecular weights.

Table 5. Retention times and molecular weights of model proteins (Thyroglobulin, IgG, Ovalbumin, Myoglobin), obtained from Size Exclusion Chromatographic profile with method conditions optimized: 25 mM sodium phosphate monobasic and 350 mM sodium chloride, pH 6.8. The injection volume was 10 µL. The flow rate was 0.8 mL/min and monitored at a wavelength of 220 nm. The column used was TSKgel G3000SWxl at a temperature of 25 ºC; Autosampler Tray Temperature: 2-8 ºC.

Model Protein	Retention Time (min)	Molecular Weight (kDa)
Thyroglobulin	7.084	670
IgG	9.621	150
Ovalbumin	11.315	44
Myoglobin	12.791	17
Vitamin B12	14.204	−
PABA	15.154	−

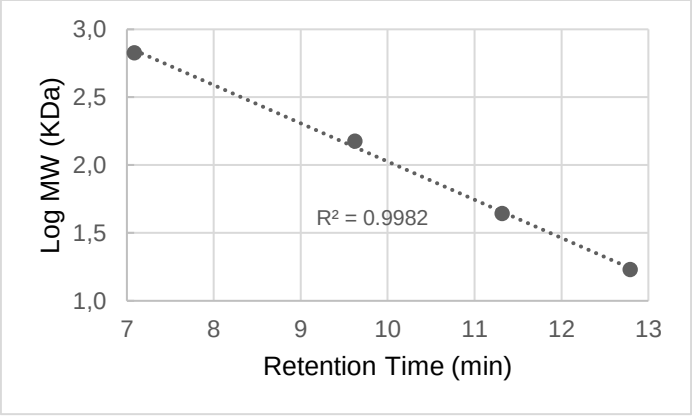

Figure 8. Model proteins (Thyroglobulin, IgG, Ovalbumin, Myoglobin) calibration curve, obtained from the molecular weights and retention times of Size Exclusion chromatographic results.

From the result obtained, it was possible to establish a calibration curve fitted by the equation $LogMW = -0.2815RT + 4.841$. This equation can be used to later determine the molecular weights of other proteins, particularly, the molecular weight of the protein under study, ARD_2019.

For most of the molecules, maximum resolution is obtained with a long column and a low flow rate. The larger the sample molecule, the more slowly it diffuses into and out of the pores in the packing. This slow diffusion results in peaks that broaden quickly with increased flow rate. As previously described by Ricker and Sandoval [43], a developed size exclusion method was applied to biological macromolecules, in which a 4-component protein mixture (BSA, Lysozyme, Ovalbumin) and sodium azide was separated on a Zorbax GF250 column (250 × 9.4 mm), using a mobile phase of 200 mM sodium phosphate, at a pH of 7.0, in which the effects of injection volume and flow rate at the level of chromatographic resolution were compared [43]. Analysis by size exclusion chromatography for the large molecules (BSA, Lysozyme, Ovalbumin), regarding the effects of flow rate on resolution, shows the increased resolution obtained by reducing flow-rate (0.25 - 0.5 mL/min), except for the small molecule, sodium azide, that increases at higher levels of flow- rate (2 mL/min).

Compared to the standard gel filtration chromatographic results, in which the flow rate used was 0.8 mL/min, it is noted that these results demonstrated a good resolution of the protein peaks, and contrary to what was described by Ricker and Sandoval [43], to the mobile phase was added ionic strength, which may be a justification factor in the elution of proteins, influencing the time of analysis, related to the ionic strength of each protein. The results obtained from the comparison of the injection volumes, demonstrating the increase in resolution that can be obtained by keeping sample volumes below a certain threshold, as with injection volumes between 10 µL and 20 µL, in accordance with the current developed method for the protein mixture (standard gel filtration).

3.1.3 Degradation Studies

3.1.3.1 *Size Exclusion Chromatography*

Degradation studies are a powerful tool in pharmaceutical development in order to assign stability and to understand the degradation pathways of the drug substances and drug products and degradation impurities. These experiments generally expose the substance to an external stress to assess the stability of the constituents or formulations.

In order to evaluate the sensitivity of the Size Exclusion Chromatographic method to detect small variations in the level of impurities, aggregates or proteolysis, a degradation study was implemented. In this degradation study, the BSA protein was the chosen model protein. The chromatographic profile of the BSA native protein, is shown in Figure 9.

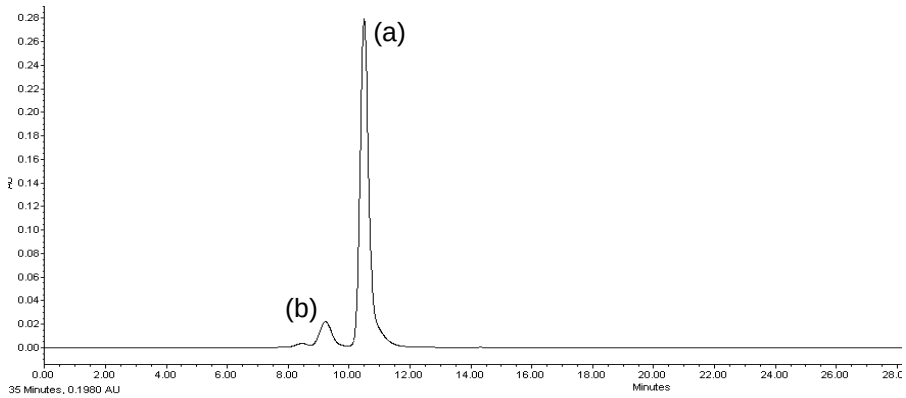

Figure 9. Size Exclusion Chromatogram profile of native protein, Bovine Serum Albumin (BSA) with method conditions optimized: 25 mM sodium phosphate monobasic and 350 mM sodium chloride, pH 6.8. The injection volume was 10 µL. The flow rate was 0.8 mL/min and monitored at a wavelength of 220 nm. The column used was TSKgel G3000SWxl at a temperature of 25 ºC; Autosampler Tray Temperature: 2-8 ºC. (a) BSA main peak, (b) High molecular weight (HMW) species.

Table 6. Bovine Serum Albumin size exclusion chromatographic results of the different spikes, 0.5% (w/w), 1% (w/w) and 2% (w/w), of fresh BSA at temperature stress condition (80 ºC) with method conditions optimized: 25 mM sodium phosphate monobasic and 350 mM sodium chloride, pH 6.8. The injection volume was 10 µL. The flow rate was 0.8 mL/min and monitored at a wavelength of 220 nm. The column used was TSKgel G3000SWxl at a temperature of 25 ºC; Autosampler Tray Temperature: 2-8 ºC.

Stress Condition	Main Peak Retention Time (min)	Main Peak % Area	HMW species Retention Time	HMW species % Area
80 ºC spike 0.5% of fresh BSA	10.493	88.17	6.697	0.01
80 ºC spike 1% of fresh BSA	10.488	87.83	6.674	0.14
80 ºC spike 2% of fresh BSA	10.495	86.75	6.677	0.38

In the chromatographic profile obtained for the native protein, the main peak has a retention time of 10.491 minutes and a percentage area of 88.13% (area). When then spiked with protein subjected to 80 ºC, it is apparent that retention times and percentage area do not vary relative to the main peak by the slight increase in the percentage of the spikes, as shown in Table 6. In the chromatographic profile of the protein, when subjected to a temperature of 80 °C, a peak corresponding to High Molecular Weight (HMW) species was observed, as shown in Figure 10. The purpose of these experiments was to evaluate the method's sensitivity to detect small variations that could occur according to the different spiking percentages of the sample subjected to stress conditions, since there are variations in HMW species peak, as shown in Figure 10, and the results demonstrated in Table 6.

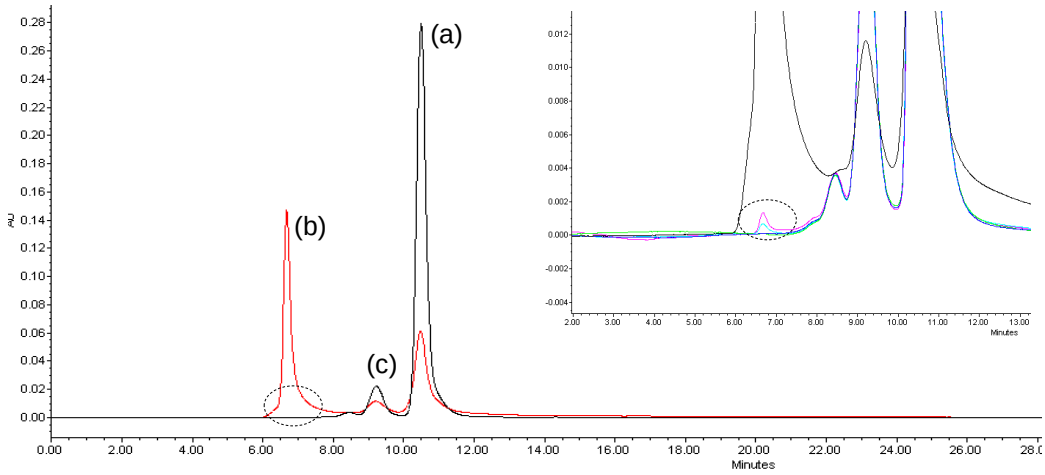

Figure 10. Size exclusion chromatographic profile of fresh BSA at temperature stress condition (80 ºC). Zoom of chromatographic profile of high molecular weight species at 80 ° C of spikes of 0.5% (w/w), 1% (w/w) and 2% (w/w) with method conditions optimized: 25 mM sodium phosphate monobasic and 350 mM sodium chloride, pH 6.8. The injection volume was 10 µL. The flow rate was 0.8 mL/min and monitored at a wavelength of 220 nm. The column used was TSKgel G3000SWxl at a temperature of 25 ºC; Autosampler Tray Temperature: 2-8 ºC. (a) BSA main peak, (b) High molecular weight (HMW) species, (c) HMW species.

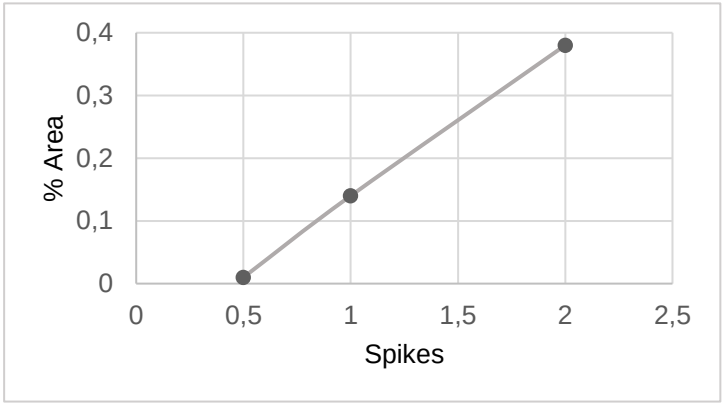

Figure 11. Linear increase at different spikes, 0.5% (w/w), 1% (w/w) and 2% (w/w) relative to the percentage of area of HMW species, obtained from size exclusion chromatographic optimized method of fresh BSA at temperature stress condition (80 ºC).

The fact that there was a decrease in the main peak area when BSA protein was subjected to stress conditions, which is explained since in the native chromatographic profile there was a shift towards high molecular weight species, could be indicative of aggregation. The linear increase in HMW species relative to a higher percentage of spikes is demonstrated in Figure 11.

In contrast to temperature stress condition, the addition of formic acid to protein, with impact on pH, reveals the appearance of Low Molecular Weight (LMW) species in the chromatographic profiles in comparison to the chromatographic profile of the native protein as shown in Figure 12, where the results are demonstrated in Table 7.

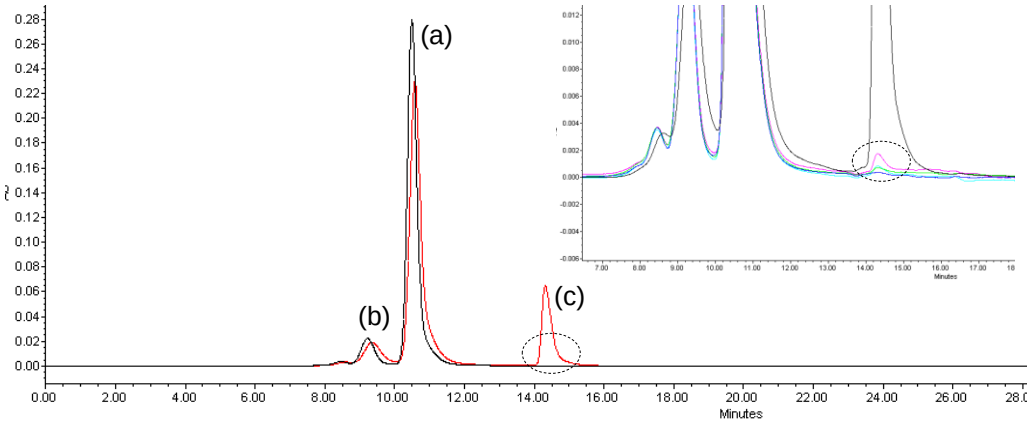

Figure 12. Size Exclusion Chromatographic profile of fresh BSA at temperature stress condition (0.5% formic acid). Zoom of chromatographic profile of low molecular weight species at 80 ° C of spikes of 0.5% (w/w), 1% (w/w) and 2% (w/w) with formic acid with method conditions optimized: 25 mM sodium phosphate monobasic and 350 mM sodium chloride, pH 6.8. The injection volume was 10 µL. The flow rate was 0.8 mL/min and monitored at a wavelength of 220nm. The column used was TSKgel G3000SWxl at a temperature of 25 ºC; Autosampler Tray Temperature: 2-8 ºC. (a) BSA main peak, (b) HMW species, (c) LMW species.

Table 7. Bovine Serum Albumin size exclusion chromatographic results of the different spikes, 0.5% (w/w), 1% (w/w) and 2% (w/w) of BSA with 0.5% of formic acid at pH stress condition (0.5% formic acid) with method conditions optimized: 25 mM sodium phosphate monobasic and 350 mM sodium chloride, pH 6.8. The injection volume was 10 µL. The flow rate was 0.8 mL/min and monitored at a wavelength of 220 nm. The column used was TSKgel G3000SWxl at a temperature of 25 ºC; Autosampler Tray Temperature: 2-8 ºC.

Stress Condition	Main Peak Retention Time (min)	Main Peak % Area	LMW species Retention Time	LMW species %Area
BSA spiked with 0.5% Formic acid	10.491	87.24	14.31	0.24
BSA spiked with 1% Formic acid	10.497	86.87	14.309	0.84
BSA spiked with 2% Formic acid	10.497	86.45	14.317	0.88

In the case of the addition of formic acid, compared to the temperature stress condition, the variation of the values of the different spikes is higher, as shown in Table 7. These small variations are presented in the peak of low molecular weight species, being indicative of proteolysis or impurities.

Thus, it is possible to infer that the size exclusion chromatography method for the two stress conditions, temperature and pH, can detect small variations, proving the sensitivity of the method.

3.1.3.2 *Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis*

In order to confirm the results obtained by size exclusion chromatography for the proteins subjected to stress conditions, a SDS-PAGE analysis was performed as described in section 2.4.2 of Materials and Methods chapter (Figure 13).

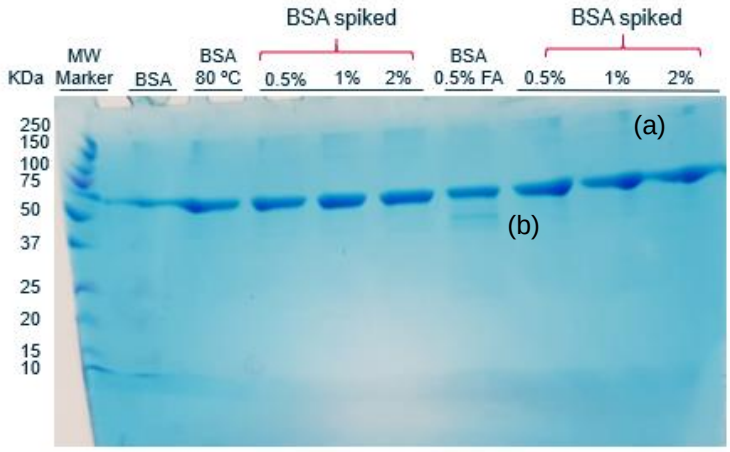

Figure 13. SDS-Polyacrylamide Gel electrophoresis: 12% (v/v) polyacrylamide gel in 2.5 M Tris-HCl, pH 8.8, containing 10% SDS (m/v). The stacking gel had a polyacrylamide concentration of 5% (m/v) in 0.5 M Tris-HCl, pH 6.8. Electrophoresis conditions: 120 V, 400 mA for 70 minutes. The samples were prepared under non reduced conditions, with the addition of sample buffer (5 mL of stacking gel, 8 mL 10% SDS, 2 mL Gycerol, and 4 mg of Bromophenol blue). The proteins were visualized by Coomassie Blue R-250. The samples used were obtained from different spikes, 0.5% (w/w), 1% (w/w) and 2% (w/w), of BSA with 0.5% of formic acid at pH stress condition (0.5% formic acid) and of fresh BSA at temperature stress condition (80 ºC). (a) High molecular weight species, (b) Low molecular weight species.

The samples were the same with respect to the two stress conditions under study. With the information obtained from the SDS-PAGE, which besides evaluating the purity of the components, provides their molecular weights, it was possible to confirm the results obtained by SEC, since it was observed bands corresponding to HMW (a) and LMW (b) species (Figure 13).

In the electrophoresis gel it is possible to observe more intense bands common to all samples corresponding to the molecular weight of the protein, 67 kDa. Bands with higher and lower molecular weights, relative to those of reference, although attenuated, are visible, could be indicative of aggregation and proteolysis phenomenon compared to chromatograms shown in Figure 10 (b) and Figure 12 (c), respectively.

3.2 Case Study: ARD_2019 Characterization

3.2.1 Karl Fischer Oven

Regarding the method for the determination of water content by Karl Fischer Oven, as described in section 2.2.2 of Materials and Methods chapter, 3 different ARD_2019 dried powder, were analysed. The results obtained for the water content determination were concordant between the different samples.

The results obtained for the water content was 0.55 - 3.02% (w/w), which are acceptable values for this material.

3.2.2 Size Exclusion Chromatography

In the characterization by size exclusion chromatography, the biomolecules are separated based on their hydrodynamic radius, allowing the evaluation of protein's molecular weight and purity. Based on this, the ARD_2019 mAb was analysed by Size Exclusion Chromatography. The chromatographic profile obtained, as shown in Figure 14, was performed at the same wavelength used to analyse the model proteins, 220 nm. At this wavelength, the peptide bond and some buffers absorbs, as opposed to a wavelength of 280 nm, characteristic of specific amino acids absorption, which makes the latter more sensitive and specific.

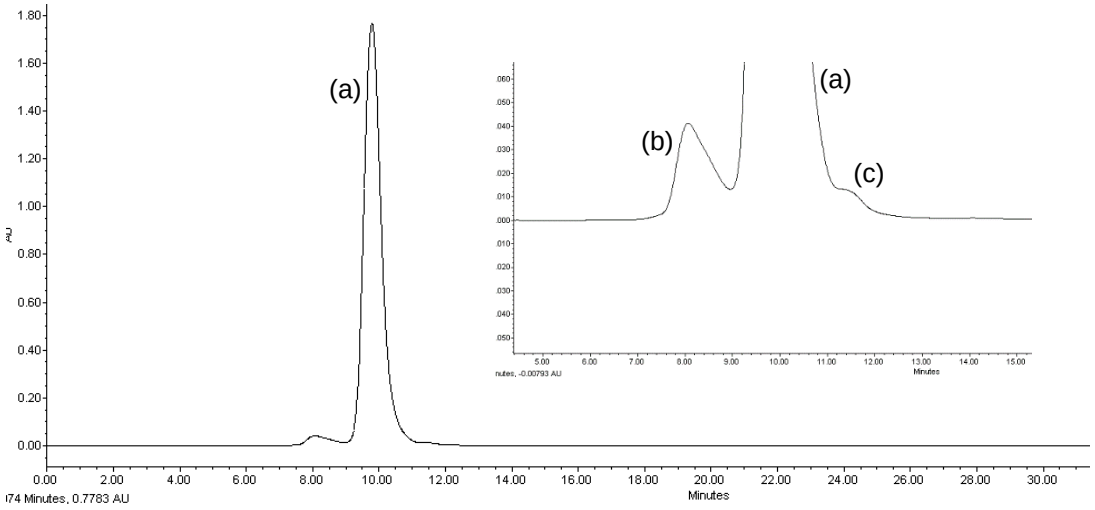

Figure 14. Size exclusion chromatographic profile of ARD_2019 with method conditions optimized: 25 mM sodium phosphate monobasic and 350 mM sodium chloride, pH 6.8. The injection volume was 10 µL. The flow rate was 0.8 mL/min and monitored at a wavelength of 220nm. The column used was TSKgel G3000SWxl at a temperature of 25 ºC; Autosampler Tray Temperature: 2-8 ºC. (a) ARD_2019 main peak, (b) HMW species, (c) LMW species.

The ARD_2019 chromatographic profile shows the elution of the native protein (main peak), High Molecular Weight and Low Molecular Weight species, in which retention times and percentage of areas are shown in the Table 8.

Table 8. ARD_2019 Size Exclusion Chromatographic results with method conditions optimized: 25 mM sodium phosphate monobasic and 350 mM sodium chloride, pH 6.8. The injection volume was 10 µL. The flow rate was 0.8 mL/min and monitored at a wavelength of 220 nm. The column used was TSKgel G3000SWxl at a temperature of 25 ºC; Autosampler Tray Temperature: 2-8 ºC.

Sample	Main Peak Retention Time (min)	Main Peak % Area	HMW species Retention Time (min)	LMW species Retention Time (min)	HMW species % Area	LMW species % Area
ARD_2019	9.79	96.6	8.06	11.3	3.11	0.46

The equation from the calibration curve obtained from the model proteins, by SEC, $LogMW = -0.2815RT + 4.841$, already described in section 3.1.2, was used to calculate the molecular weight of the ARD_2019 native protein, using the retention time obtained (9.79 minutes). Thus, the value of 122 kDa was obtained for the total molecular weight of the protein, approximate value of the theoretical value of the total molecular weight of an antibody, 150 kDa.

Analysis of the high molecular weight species was performed as these could be other high molecular weight proteins or aggregates. Using the retention time of the HMW species (8.06 minutes), it was obtained a value of 372 kDa, which makes approximately three times the molecular weight of the native protein, suggesting that aggregation of the protein itself may have occurred. The presence of Low Molecular Weight (LMW) species is also visible, with a lower percentage of area [(0.46% (area))], wherein the calculated molecular weight was 45.71 kDa taking into account its retention time (11.3 minutes), indicative of proteolysis phenomenon, influenced by antibody purification related impurities or protein instability processes caused by the low pH or temperature stress conditions.

In the case of the presence of HMW species (b), this phenomenon is justified from self-association of the native monomer (a) (unaltered conformationally and active) that occurs when conformationally

stable monoclonal antibody molecules form complexes through non-covalent interactions. This phenomenon of self-association may lead to reversible or irreversible aggregation, depending on the nature of the native protein.

A study performed by L.Chen *et al*. of the structural characterization of Kadcyla® (ado-trastuzumab emtansine) and its biosimilar candidate was also performed by size exclusion chromatography in order to evaluate the size variants [44]. Antibody-drug conjugates (ADCs) or immunoconjugates are a sub-class of biotherapeutics designed to facilitate the targeted delivery of potent cytotoxic drugs to cancer cells. The FDA approved ado-trastuzumab emtansine (Kadcyla®, Genentech), an antibody-drug conjugate, used for the treatment of human epidermal growth factor receptor (HER)2-positive breast cancer. The size exclusion chromatographic profile of the ADCs samples is shown in Figure 15.

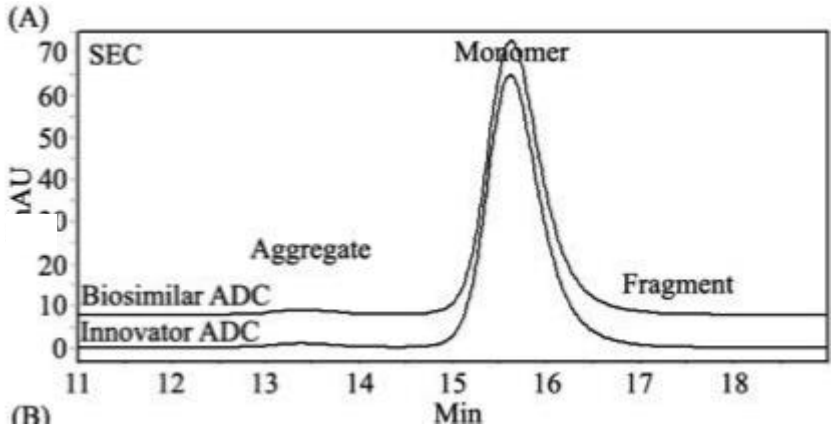

Figure 15. Size Exclusion Chromatographic profile of Kadcyla® and biosimilar ADC with the following method conditions: 0.2 M tripotassium phosphate and 15% isopropyl alcohol, at a flow rate of 0.5 mL/min. The UV absorbance was measured at a wavelength of 280 nm. Adapted from L.Chen et al. [44]

The results obtained from this size exclusion chromatographic study performed using the same column, showed that Kadcyla® sample contains 98.34% of monomer, 1.64% of aggregates and 0.02% of fragments. As expected, these results are slightly better than the results obtained for ARD_2019 mAb that was subjected to various freeze-thaw cycles throughout these studies. The results obtained, show a content of HMW species of about 3.11% area and LMW species of about 0.46% area. The monoclonal antibody ARD_2019 contains 96.6% of monomer.

The Sec methodology developed during this thesis proved to be essential for the qualitative and quantitative evaluation of protein aggregates. The MW range (17 – 670 kDa) of the protein markers used as calibration standards, confirmed to be suitable for the analysis of monoclonal antibodies, 150 kDa. Moreover, with this methodology it is possible to assess protein structural stability, as it allows the detection of both Low and High MW species.

3.2.3 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis

Analytical techniques, such as size exclusion chromatography and SDS-PAGE, can be considered as complementary techniques. The purpose of the characterization of ARD_2019 by SDS-PAGE that

separates different size proteins according to their electrophoretic mobility, was (1) to compare the results obtained with the ones obtained by SEC, and (2) to evaluate the purity of the antibody.

In the electrophoresis gel performed, as shown in Figure 16, four samples were run, fresh antibody and three dried powder samples, under reducing conditions, showing the separation of the chains. Considering the molecular marker as a reference for calculating the molecular weight of the antibody, the bands corresponding to the antibody light chains, 25 kDa, and the antibody heavy chains, 50 kDa are visible.

The electrophoresis gel shows consistent results regarding the size exclusion chromatographic technique, since it evidences the purity of the dried powder samples, given the absence of bands of molecular weight superior to the heavy chain, confirming the hypothesis of possible aggregation.

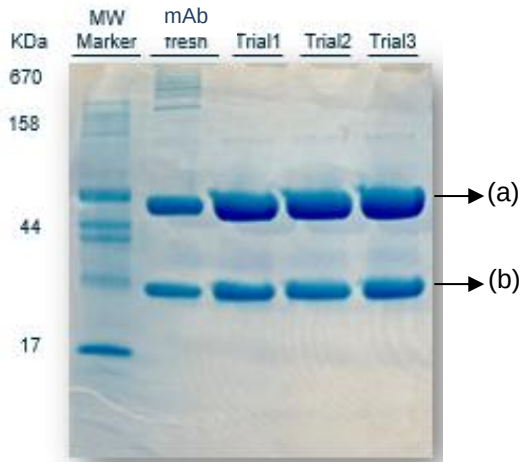

Figure 16. SDS-Polyacrylamide Gel electrophoresis: 12% (v/v) polyacrylamide gel in 2.5 M Tris-HCl, pH 8.8, containing 10% SDS (m/v). The stacking gel had a polyacrylamide concentration of 5% (m/v) in 0.5 M Tris-HCl, pH 6.8. Electrophoresis conditions: 120 V, 400 mA for 70 minutes. The fresh monoclonal antibody and the three trials at a concentration of 2 mg/mL, were prepared under non reduced conditions, with the addition of sample buffer (5 mL of stacking gel, 8 mL 10% SDS, 1 mL β-mercaptoethanol, 2 mL Gycerol, and 4 mg of Bromophenol blue). The proteins were visualized by Coomassie Blue R-250. (a) Antibody Heavy Chain (HC), (b) Antibody Light Chain (LC).

In the characterization of monoclonal antibodies, where fragments are revealed by SDS-PAGE under reducing conditions, and in the case of pure ARD_2019 samples, the heavy (50 kDa) and light (25 kDa) chains of the antibody are expected to be observed, as shown in Figure 16. Referring to an analysis performed by Z.C. Zhu *et al.,* under non-reducing conditions, in the specific case of the IgG 4 antibody, its heavy and light chain fragments are also preserved and revealed in the electrophoresis gel [45].

3.2.4 Ion Exchange Chromatography

In order to evaluate the charge heterogeneity of the monoclonal antibody, a Cation Exchange Chromatography was performed. Parameters such as the buffer and pH range were optimized, obtaining an optimization profile. Initially, the first condition tested was pH gradient, using the Tris buffer (10 mM), with a pH range between 7.5 and 10.5, considering the isoelectric point (pI) of the antibody, 8.2 (Figure 17).

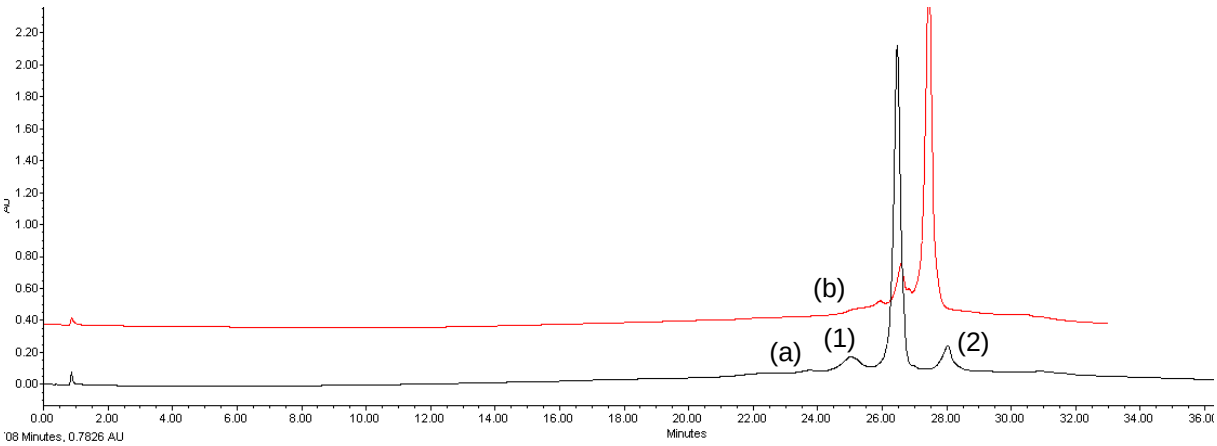

Figure 17. Cation exchange chromatographic profile of ARD_2019 by pH gradient with method conditions optimized: mobile phase A :10 mM Tris buffer, pH 7.5 and mobile phase B: 10 mM Tris buffer, pH 10.5. The injection volume was 10 µL. The flow rate was 0.8 mL/min and monitored at a wavelength of 220 nm. The column used was Protein Pak Hi Res CM at a temperature of 25 ºC; Autosampler Tray Temperature: 2-8 ºC. (a) pH 7.5 - 10.5, optimized profile (b) pH 7 - 9.5, (1) acidic species, (2) basic species.

In Figure 17 are presented the comparison of two chromatographic profiles obtained with different pH variations. Both assumed a linear gradient between the mobile phase A (10 mM Tris buffer, pH 7.5) and the mobile phase B (10 mM Tris buffer, pH 10.5). Initially the percentage of mobile phase A was higher than the mobile phase B, allowing conditioning of the column. The positively charged mobile phase will bind to the negatively charged carboxymethyl group of the stationary phase, and the antibody will have to compete to bind, and therefore protein desorption occurs, and it is eluted. Thus, it is possible to observe in the chromatographic profile the acidic (1) and basic (2) antibody species, relative to the main peak. In the case of a pH range between 7 and 9.5 the basic species are not visible, and it was necessary to extend the pH range of the buffer (7.5 and 10.5).

A salt gradient with impact on the electrostatic surface of the protein was also tested, giving the information on the different charged variants of the antibody. In this case, the mobile phase A was 10 mM Tris Buffer, pH 7.5, and the mobile phase B was 10 mM Tris Buffer + 300mM NaCl, pH 7.5, having obtained the optimized chromatographic profile as shown in Figure 18.

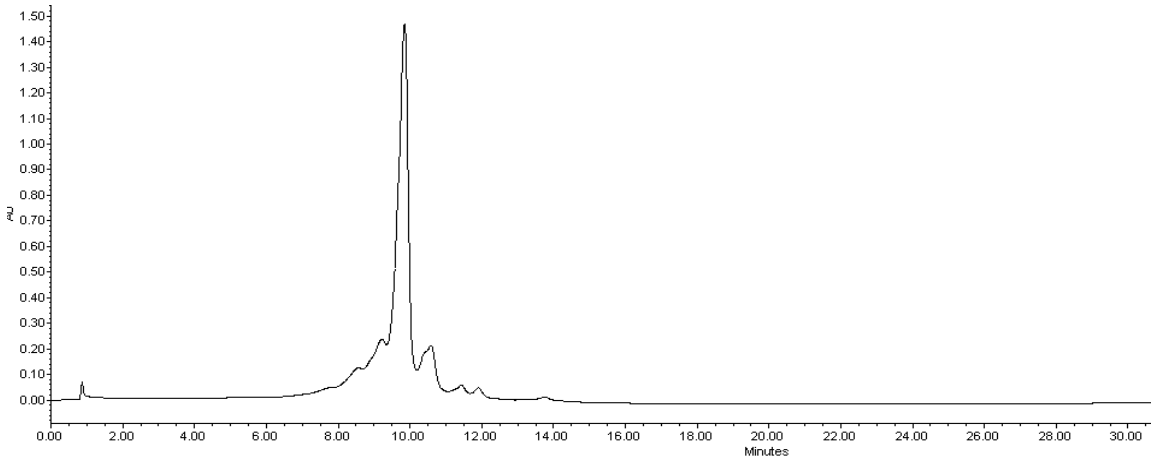

Figure 18. Cation ion exchange chromatographic profile obtained for ARD_2019 by salt gradient with method conditions optimized: mobile phase A :10 mM Tris buffer, pH 7.5 and mobile phase B: 10 mM Tris Buffer + 300 mM NaCl, pH 7.5. The injection volume was 10 µL. The flow rate was 0.8 mL/min and monitored at a wavelength of 220nm. The column used was Protein Pak Hi Res CM at a temperature of 25 ºC; Autosampler Tray Temperature: 2-8 ºC.

In this situation, the additional factor is the salt, and the higher the net charge of the protein, the higher the ionic strength that is needed for elution, in which protein must compete to be able to bind to the negatively charged carboxymethyl group. Contrary to what occurs in the pH gradient, in the salt gradient, initially with 100% of mobile phase A, up to 30 min of run there is a gradual incremental change of mobile phase B, which contains 300 mM NaCl, with 65% mobile phase A and 35% mobile phase B.

The cation exchange chromatographic methodologies developed during this thesis, which include pH and salt gradient elution, allowed the evaluation of the charged variant profile of the antibody under study. With these methods it was possible to identify 8 different charge variants for ARD_2019 mAb.

Chapter 4

Conclusions

The analytical characterization of the model proteins (BSA, Ovalbumin and Lyzozyme) was made through different analytical techniques. The process residual water was carried out by a Karl Fischer Oven method. Initially, a temperature ramp between 50 ºC and 220 ºC was performed, to infer the optimal temperature that should be considered. The optimal temperature determined was 120 ºC. All the model proteins presented a similar temperature ramp profile. Since this method was repeatable for BSA protein, was also used to determine the water content of the ARD_2019 monoclonal antibody, using the same optimal working temperature, 120 ºC. During the sample preparation of ARD_2019 mAb, hygroscopic protein, and in order to minimize any kind of interference from the environment with respect to the percentage of water, it was necessary to restrict the relative humidity values, with values less than 5% RH. Values between [0.55 and 3.02% (w/w)] for the water content were obtained. Thus, it can be concluded that the developed KF oven method, is suitable to determine the water content of process residual water in model protein BSA and ARD_2019 antibody.

The evaluation and characterization of the physicochemical parameters, size variants and process related impurities was performed by the development of a Size Exclusion Chromatographic method. A standard gel filtration was used [(Thyroglobulin (670 kDa), IgG (150 kDa), Ovalbumin (44 kDa), Myoglobin (17 kDa), and Vitamin B12 (14 kDa)], and the best results were obtained when applying a flow rate of 0.8 mL/min, 25 mM sodium phosphate monobasic and 350 mM sodium chloride, 7.5% IPA buffer, an injection volume of 10 µL and a pH value of 6.8, monitored at a wavelength of 220 nm. This chromatographic profile revealed better peak shape, peak sensitivity and resolution. Based on theoretical molecular weights of the known proteins present in the standard gel filtration and their retention times, a calibration curve was made. The optimized SEC method was also applied to the ARD_2019 antibody, to determine its molecular weight and to characterize the antibody size variants. From the SEC profile was observed, besides the main peak (122 kDa), High Molecular Weight species (HMW) (372 kDa) and Low Molecular Weight species (LMW) (45.71 kDa), with the molecular weights obtained from the calibration curve equation. The molecular weight values obtained allowed to infer that they could be indicative of (1) aggregation, in the case of HMW species, and (2) proteolysis, in the case of LMW species, which was also corroborated by the analysis of the electrophoresis gel obtained for ARD_2019 antibody.

After optimizing the method using the model proteins and applying it to the ARD_2019 antibody, a degradation study was carried out to evaluate the sensitivity of the SEC method to detect small variations in the level of impurities, aggregates or proteolysis. For this study, the BSA protein was subjected to stress conditions, temperature (80 ºC) and pH (0.5% formic acid). Regarding the chromatographic profile obtained for the temperature stress condition, it showed that when increasing the spiking levels, it was observed an increasing in the area corresponding to HMW species, which may be associated with the aggregation phenomenon. The same was verified with the addition of formic acid to the BSA protein, being observed the increase in the percentage area of the LMW species with the increase in the percentage of spikes, which can be associated with proteolysis phenomenon. Thus, it can be concluded that the stress conditions analysed, pH and temperature, can be detected by SEC, which confirms the sensitivity of the method. The same samples, subjected to the same stress

conditions, were also analysed by SDS-PAGE, in order to confirm the results obtained by Size Exclusion Chromatography, by evaluating the purity of the electrophoresis gel of the BSA protein. The results obtained corroborate the results achieved by the SEC method, since bands with higher and lower molecular weights, relative to those of reference, were visible, corresponding to aggregation and proteolysis phenomenon.

The electrophoretic pattern of ARD_2019 protein was also analysed by SDS-PAGE, under reducing conditions. The gel showed consistent results with the Size Exclusion Chromatography method, in terms of purity, in which only intense bands, corresponding to the heavy chain (50 kDa) and light chain (25 kDa) were visualized.

Cation Exchange Chromatography was also performed in order to evaluate the isoform pattern. The charge heterogeneity of the ARD_2019 antibody was evaluated by either pH gradient and salt gradient. In the pH gradient, the pH and buffer were optimized, where 10 mM Tris buffer with a pH range of 7.5 and 10.5 was used. These conditions may be considered optimal for the characterization of ARD_2019 antibody, since the chromatographic profile demonstrated well resolved acidic and basic antibody species. After optimization of the pH cation exchange methodology, the chromatographic conditions of the method were extrapolated for salt gradient analysis, where ionic strength was added by the addition of 300 mM sodium chloride at pH 7.5. It is possible to conclude that the Cation Exchange Chromatographic method, was reliable and consistent for charge heterogeneity analysis, since it was possible to analyse acidic and basic species, and charge variants of ARD_2019 protein.

References

[1] G. Walsh, Pharmaceutical Biotechnology: Concepts and Applications, 2007.

[2] G. Walsh, Biopharmaceuticals: Biochemistry and Biotechnology, 2003.

[3] A. Faustino, D. Costa, L. Lacalendola, V. De Araújo, C. Alexandre, S. Leite, P. Gava, L. De Oliveira-nascimento, C.D.O. Rangel-yagui, P. De Oliveira, M. Antonio, D. Oliveira, A. Pessoa, Biotechnology and Industry Microbiology Biopharmaceuticals from microorganisms : from production to purification, Brazilian J. Microbiol. 47 (2016) 51–63. doi:10.1016/j.bjm.2016.10.007.

[4] E. M.G., A. S., R.-P. M., E. A., Biosimilars: Potential implications for clinicians, Clin. Cosmet. Investig. Dermatol. 9 (2016) 135–142. doi:10.2147/CCID.S91691 LK - http://limo.libis.be/resolver?&sid=EMBASE&issn=11787015&id=doi:10.2147%2FCCID.S91691&atitle=Biosimilars%3A+Potential+implications+for+clinicians&stitle=Clin.+Cosmet.+Invest.+Dermatol.&title=Clinical%2C+Cosmetic+and+Investigational+Dermatology&volume=9&issue=&spage=135&epage=142&aulast=Eleryan&aufirst=Misty+G.&auinit=M.G.&aufull=Eleryan+M.G.&coden=&isbn=&pages=135-142&date=2016&auinit1=M&auinitm=G.

[5] S. Mitragotri, P.A. Burke, R. Langer, Overcoming the challenges in administering biopharmaceuticals: Formulation and delivery strategies, Nat. Rev. Drug Discov. 13 (2014) 655–672. doi:10.1038/nrd4363.

[6] G. Walsh, Biopharmaceutical benchmarks 2018, Nat. Publ. Gr. 36 (2018) 1136–1145. doi:10.1038/nbt.4305.

[7] L. Sciences, Advanced Biopharmaceutical Manufacturing : An Evolution Underway.

[8] G. Walsh, B. Murphy, Biopharmaceuticals, an Industrial Perspective, 1999.

[9] A.J. Brown, D. Kalsi, A. Fernandez-martell, N.O.W. Barber, Y.D. Patel, R. Turner, L. Claire, Y.B. Johari, D.C. James, Expression Systems for Recombinant Biopharmaceutical Production by Mammalian Cells in Culture, (2017) 423–467.

[10] S.A. Berkowitz, J.R. Engen, J.R. Mazzeo, G.B. Jones, Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars, Nat. Rev. Drug Discov. 11 (2012) 527–540. doi:10.1038/nrd3746.

[11] B.S. Sekhon, Biopharmaceuticals : an overview Abstract :, 34 (2010) 1–19.

[12] J. Reichert, Monoclonal Antibodies as Innovative Therapeutics, Curr. Pharm. Biotechnol. 9 (2008) 423–430. doi:10.2174/138920108786786358.

[13] M. Kesik-brodacka, Progress in biopharmaceutical development, (2017) 306–322.

doi:10.1002/bab.1617.

[14] L. Sanchez-Garcia, L. Martín, R. Mangues, N. Ferrer-Miralles, E. Vázquez, A. Villaverde, Recombinant pharmaceuticals from microbial cells: A 2015 update, Microb. Cell Fact. 15 (2016) 1–7. doi:10.1186/s12934-016-0437-3.

[15] International Council for Harmonisation, ICH Harmonised Tripartite Guideline. Specifications: Test Procedures and acceptance critera for biotechnological/ biological products Q6B, ICH Harmon. Tripart. Guidel. (1999) 1–20. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q6B/Step4/Q6B_Guideline.pdf.

[16] C.J. Robinson, C. Jones, Quality control and analytical techniques for biopharmaceuticals, Bioanalysis. 3 (2011) 81–95. doi:10.4155/bio.10.161.

[17] M. Kristina, O. Montacir, H. Montacir, Journal of Pharmaceutical and Biomedical Analysis Physicochemical characterization of biopharmaceuticals, 130 (2016) 366–389.

[18] R.I. Gumport, F.H. Deis, N.C. Gerber, Biochemistry , Fifth Edition.

[19] B.E. Tropp, 77-Protein Structure and Function, Mol. Biol. 4th Ed. (2012) 27–74. doi:10.1021/ja209464f.

[20] N.H.G.R. Institute, Dimensional Protein Structures, (n.d.). https://www.ncbi.nlm.nih.gov/Class/MLACourse/Original8Hour/Genetics/structure.html.

[21] J.M. Jez, Revisiting protein structure, function, and evolution in the genomic era, J. Invertebr. Pathol. 142 (2017) 11–15. doi:10.1016/j.jip.2016.07.013.

[22] G. Vidarsson, G. Dekkers, T. Rispens, IgG subclasses and allotypes : from structure to effector functions, 5 (2014) 1–17. doi:10.3389/fimmu.2014.00520.

[23] In Reply: BEHAVIOUR THERAPY, Br. J. Psychiatry. 112 (1966) 211–212. doi:10.1192/bjp.112.483.211-a.

[24] D.C. Carter, J.X. Ho, Structure of serum albumin, Adv. Protein Chem. 45 (1994) 153–176. doi:10.1016/S0065-3233(08)60640-3.

[25] Sigma-Aldrich, -glucosamine residues in peptidoglycan and between N-acetyl-, 0–1.

[26] P. Bruttel, R. Schlink, Water Determination by Karl Fischer Titration Water Determination by Karl Fischer Titration, (n.d.).

[27] P.D. Leon Zhou, J.M. Socha, P.D. Frederick G. Vogt, P.D. Sarah Chen, P.D. Alireza S. Kord, A Systematic Method Development Strategy for Water Determinations in Drug Substance Using Karl Fischer Titrations, (n.d.). https://www.americanpharmaceuticalreview.com/Featured-Articles/117884-A-Systematic-Method-Development-Strategy-for-Water-Determinations-in-Drug-Substance-Using-Karl-Fischer-Titrations/.

[28] G. Brusotti, E. Calleri, R. Colombo, G. Massolini, F. Rinaldi, C. Temporini, Advances on Size

Exclusion Chromatography and Applications on the Analysis of Protein Biopharmaceuticals and Protein Aggregates : A Mini Review Advances on Size Exclusion Chromatography and Applications on the Analysis of Protein Biopharmaceuticals and Pr, Chromatographia. (2017). doi:10.1007/s10337-017-3380-5.

[29] EMA, Guideline on Development, Production, Characterization and Specifications for Monoclonal Antibodies and Related Products, EMEA/CHMP/BWP/157653/2007, 44 (2016). www.ema.europa.eu/contact.

[30] S. Fekete, A. Goyon, Size Exclusion Chromatography of Protein Biopharmaceuticals : Past , Present and Future Conventional Size Exclusion Chromatography (HP-SEC) – The Past Modern Size Exclusion Chromatography (UHP-SEC) – The Present, i (2018) 1–6.

[31] G.E. Healthcare, Size Exclusion Chromatography Handbooks from GE Healthcare Life Sciences For more information refer to gelifesciences . com / handbooks, (n.d.).

[32] H. Liu, G. Gaza-Bulseco, C. Chumsae, Analysis of Reduced Monoclonal Antibodies Using Size Exclusion Chromatography Coupled with Mass Spectrometry, J. Am. Soc. Mass Spectrom. 20 (2009) 2258–2264. doi:10.1016/j.jasms.2009.08.015.

[33] T.L. Kirley, K.D. Greis, A.B. Norman, Domain unfolding of monoclonal antibody fragments revealed by non- reducing SDS-PAGE, Biochem. Biophys. Reports. 16 (2018) 138–144. doi:10.1016/j.bbrep.2018.10.004.

[34] S.- Page, S.- Page, Chapter 14.

[35] M. Chamberlin, J. McGrath, L. Waskell, © 1970 Nature Publishing Group, Nature. 228 (1970) 227–231.

[36] S. Fekete, A. Beck, J. Veuthey, D. Guillarme, Journal of Pharmaceutical and Biomedical Analysis Ion-exchange chromatography for the characterization of biopharmaceuticals, J. Pharm. Biomed. Anal. 113 (2015) 43–55. doi:10.1016/j.jpba.2015.02.037.

[37] F. Di Virgilio, J.M. Walker, Methods in Molecular Biology 1149, 1485 (2014) 209–223. doi:10.1007/978-1-4939-6412-3.

[38] G.E. Healthcare, Ion Exchange Chromatography Handbooks from GE Healthcare Life Sciences For more information refer to www.gelifesciences.com/handbooks, (2008).

[39] A. Grönberg, Ion Exchange Chromatography, Biopharm. Process. Dev. Des. Implement. Manuf. Process. (2018) 379–399. doi:10.1016/B978-0-08-100623-8.00018-9.

[40] S.N. Ronkart, M. Paquot, C. Fougnies, C. Deroanne, J.C. Van Herck, C. Blecker, Determination of total water content in inulin using the volumetric Karl Fischer titration, Talanta. 70 (2006) 1006–1010. doi:10.1016/j.talanta.2006.02.024.

[41] Tosoh Bioscience, Operating Conditions and Specifications - TSKgel Protein A-5PW, (2019) 5–6.

[42] P. Hong, S. Koza, E.S.P. Bouvier, W. Corporation, A review size-exclusion chromatography for

the analysis of protein biotherapeutics and their aggregates, (2012) 2923–2950. doi:10.1080/10826076.2012.743724.

[43] R.D. Ricker, L.A. Sandoval, Fast, reproducible size-exclusion chromatography of biological macromolecules, J. Chromatogr. A. 743 (1996) 43–50. doi:10.1016/0021-9673(96)00283-X.

[44] 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.

[45] Z.C. Zhu, Y. Chen, M.S. Ackerman, B. Wang, W. Wu, B. Li, L. Obenauer-Kutner, R. Zhao, L. Tao, P.M. Ihnat, J. Liu, R.B. Gandhi, B. Qiu, Investigation of monoclonal antibody fragmentation artifacts in non-reducing SDS-PAGE, J. Pharm. Biomed. Anal. 83 (2013) 89–95. doi:10.1016/j.jpba.2013.04.030.

Annex 1

Table 9. ICH Q6B: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products

	Type of Analysis	Analytical Techniques
Structural characterization (primary structure)	Amino acid sequence	Proteolytic digestion, HPLC, Capillary electrophoresis, LC-MS/HRMS, GC-MS (biopolymers/ carbohydrates)
	Amino acid composition	
	Terminal amino acid sequence	
	Peptide mapping	
	Sulfhydryl group(s) and disulfide bridges	
	Carbohydrate structure	
	Post translational modification	
Structural characterization (higher-order structure)	Spectroscopy	Differential Scanning Calorimetry (DSC) Differential Scanning Fluorimetry (DSF) Circular Dichroism (CD) Fourier Transformed Infra-Red Spectroscopy (FTIR) Nuclear Magnetic Resonance Spectroscopy (NMR)
Physicochemical properties	Molecular weight or size	LC-MS/MS, HPLC methods: IEX, SEC, RP, HIC, HILIC, NP, HPAEC, IC, Affinity Chromatography SDS-PAGE Electrophoresis, Isoelectric focusing, Western-blot
	Isoform pattern	
	Extinction coefficient	
	Electrophoretic patterns	
	LC patterns	
Product related impurities	Degradation	LC-MS/MS, immunoassays, Karl Fisher, HPLC methods (SEC, IEX, etc), Capillary Electrophoresis, SDS-PAGE Electrophoresis, Isoelectric focusing, Circular Dichroism (CD), Analytical ultracentrifugation, Zeta Potential
	Aggregation	
	Oxidation, deamination	
	Disulfide scrambling	
	Buffer / matrix composition	
	Residual water	
Process related impurities	Residual sovelnts	GC-FID, HPLC methods, Capillary Electrophoresis, SDS-PAGE Electrophoresis.
	Leachables	
	Cell media components	
Quantity and Biological activity	Quantity	LC-MS/MS, protein assay, immunoassays (ELISA, western blotting, etc...), enzymatic activity
	Activity / binding assays	