KTH ROYAL INSTITUTE OF TECHNOLOGY

Degree Project in Medical Biotechnology

Second cycle 30 credits

Identification of Monoclonal Antibodies

Peptide Mass Fingerprinting (PMF) with Matrix Assisted Laser Desorption Ionization (MALDI), Time of Flight (ToF), Mass Spectrometry (MS) and Protein Peptide Mapping (PPM) with Capillary Electrophoresis (CE)

SOFIA BENGTSSON

MAIN SUPERVISOR **JOHN LÖFBLOM, KTH**

EXTERNAL SUPERVISORS **AHMAD AMINI & JOHAN CARLSSON,**

THE SWEDISH MEDICAL PRODUCTS AGENCY

[1]

Stockholm, Sweden 2023

Abstract

The number of monoclonal antibodies used in pharmaceuticals is increasing sharply. These medicines are expensive, and the risk of counterfeiting is high. The need to develop a method for rapid and precise identification of monoclonal antibodies is therefore urgent. For identification, analyses were performed with Matrix Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry (MALDI-ToF-MS), Capillary Gel Electrophoresis (CGE) and Capillary Zone Electrophoresis (CZE) on nine monoclonal antibodies. The focus was to investigate whether significant physiochemical features and unique amino acid sequences were present and could be distinguished. Various analyses with MALDI-ToF-MS were used to both separate the monoclonal antibodies based on their physicochemical properties and annotate amino acid sequences containing key fragments. With the methods based on capillary electrophoresis, separation was also achieved. CZE is preferred over CGE as the amount of data obtained from CZE is greater and sample preparation is simpler. In summary, an identification process protocol was designed and is initiated with MALDI-ToF-MS analyses of reduced-form monoclonal antibodies against known references. A hypothesis is then formulated based on which antibodies look the most similar. Finally, these are analysed by CZE to determine the identity of the monoclonal antibody.

Keywords

Monoclonal Antibodies (mAbs), Biosimilars, Deglycosylation, Matrix-Assisted Laser Desorption/Ionization Time-of Flight Mass Spectrometry (MALDI-ToF-MS), Peptide Mass Fingerprinting, In Source-Decay (ISD), Capillary Electrophoresis (CE), Capillary Gel Electrophoresis (CGE), Capillary Zone Electrophoresis (CZE).

Sammanfattning

Antalet monoklonala antikroppar som används i läkemedel ökar kraftigt. Dessa läkemedel är dyra och risken för förfalskning är stor. Behovet att utveckla en metod för snabb och precis identifiering av monoklonala antikroppar är därför brådskande. För identifiering utfördes analyser med Matrix Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry (MALDI-ToF-MS), Capillary Gel Electrophoresis (CGE) and Capillary Zone Electrophoresis (CZE) på nio monoklonala antikroppar. Fokuset var att undersöka huruvida signifikanta fysiokemiska egenskaper och unika aminosyrasekvenser var närvarande och kunde urskiljas. Olika analyser med MALDI-ToF-MS användes till att både separera de monoklonala antikropparna baserat på dess fysiokemiska egenskaper, och annotera aminosyrasekvenser innehållande nyckelfragment. Med metoderna baserade på kapillärelektrofores uppnåddes också separation. CZE föredras framför CGE då mängden data som erhålls från CZE är större och provberedningen är enklare. Sammanfattningsvis utformades ett protokoll för identifieringsprocessen, vilket inleds med MALDI-ToF-MS-analyser av monoklonala antikroppar på reducerad form mot kända referenser. Därefter är en hypotes formulerad utifrån vilka antikroppar som ser mest lika ut. Slutligen analyseras dessa med CZE för fastställning av den monoklonala antikroppens identitet.

Table of Contents

1. Introduction

The Swedish Medical Products Agency is a state administrative authority, responsible for approval and supervision of pharmaceuticals, natural remedies and drugs classified as narcotics. Responsibility for supervision of medical technology products and cosmetics is also included. The Swedish Medical Products Agency belongs to the Ministry of Social Affairs. The main mission is to ensure that the medicinal products available for the individual patient, the health and medicinal care and the animal health care are both safe and effective. It is also of great importance that these products are used in an appropriate and cost-effective way.[2] Approval from the Swedish Medicinal Products Agency of pharmaceuticals for manufacturing and marketing authorisations is most often done through the European Medicines Agency (EMA).[3]

The use of monoclonal antibodies as pharmaceuticals are continuously increasing. Therapeutic monoclonal antibodies have a wide area of use in medication. Different antibodies can act through several different mechanisms, including inducement of apoptosis, modulation of signalling pathways and blocking of functions of targeted molecules. Monoclonal antibodies are mainly used to treat autoimmune diseases and cancer. These expensive pharmaceuticals can be falsified and therefore the need of a rapid and accurate procedure for identification is urgent.[4] Analyses based on MALDI-ToF-MS and capillary electrophoresis (CE) will identify, for each monoclonal antibody, amino acid key fragments and significant properties.

The aim of this project is to develop a rapid and specific MALDI-ToF-MS- & CE-based method for identification of monoclonal antibodies by peptide mapping.

1.1 Monoclonal Antibodies

Monoclonal antibodies (mAbs) are produced from B-cell clones, which entails that all produced antibodies are identical. This property makes it beneficial for use as medication. The structure of monoclonal antibodies (*figure 1*) consists of four polypeptide chains linked together with 16 disulfide bonds, covalent bonds between sulfur atoms on cysteine amino acids. Of these chains are two identical heavy chains (HC) and two identical light chains (LC). The molecular mass of the LC and HC is around 25 kDa and 50 kDa respectively, which results in an approximate total molecular mass of the whole mAb of 150 kDa. The top part of both chains has a variable domain with an antigen-binding site and the rest of the chains are constant regions.[5] In, addition, carbohydrates attached to the heavy chains contribute to the molecular mass of mAbs. Glycosylation is a post-translational modification generating several isoforms. These are created when different sizes and structures of oligosaccharides are N-linked by covalent addition to the asparagine residues in the sequences

Asn-X-Thr or Asn-X-Ser. The variations created by glycosylation affect the structure, stability and activity of the mAb.[6]

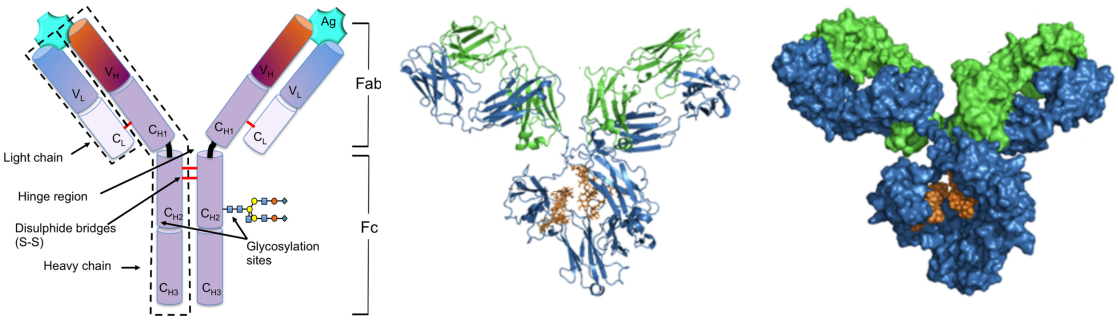

Figure 1: Structures of a mAb. The structure to the left is schematic with all parts of the mAb marked. The structure in the middle and to the right is two different kinds of the quaternary protein structure of a mAb. The light chains are green, and the heavy chains are blue. The glycosylation is orange.[7][8]

1.1.1 Biosimilars

A biological pharmaceutical, such as mAbs, that is produced and is highly similar to an already approved and existing biological medicine on the market is called a biosimilar. In terms of quality, safety and efficacy, there must therefore be no clinically significant difference between a biosimilar and its reference medicine. Biosimilars are not the same as generics since generics are identical chemical drugs. Regarding biological pharmaceuticals, the same level of similarity cannot be achieved due to its complexity and the natural variability. Biosimilars can be approved after the period of data exclusivity for the reference medicine has expired, which usually occurs after 8 years. The main benefit of biosimilars is that they improve the access for the patient.[9][10]

1.2 Analysis of mAbs

When analysing mAbs, it is often necessary to treat them to be able to collect the desired data. Reduction and alkylation are performed to separate the HC and LC from each other. Deglycosylation is performed to remove the glycans from the mAbs.

1.2.1 Reduction of mAbs

When analysing mAbs in reduced form, the polypeptide chains are separated from each other by heat denaturation and reduction. The disulfide bonds are broken when 2-Mercaptoethanol (2-ME) or dithiothreitol (DTT) is added (*figure 2*). Hydrogen atoms from the reducing agent bind to the sulfur atoms on the protein and two 2-ME/DTT molecules bind to each other by forming disulfide bonds. This makes it possible to break up the tertiary structure of a protein and allows analysis of heavy and light chains separated.[11]

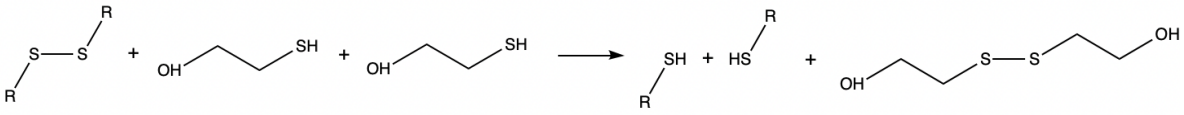

Figure 2: Chemical structure of 2-ME reacting with a substance containing a disulfide bridge.

1.2.2 Alkylation of mAbs with Iodoacetamide

When mAbs are reduced, the thiol (SH-) groups on cysteine that form disulfide bonds become free sulfur atoms. To avoid the disulfide bonds being recreated, iodoacetamide (IAM) is added. IAM reacts by condensation reaction with a free sulfur atom at cysteine and thereby decreases the number of reactive groups with free sulfur atoms (*figure 3*). The alkylated state is more stable, and the cleavage of the chains can thereby be assessed as irreversible.[11]

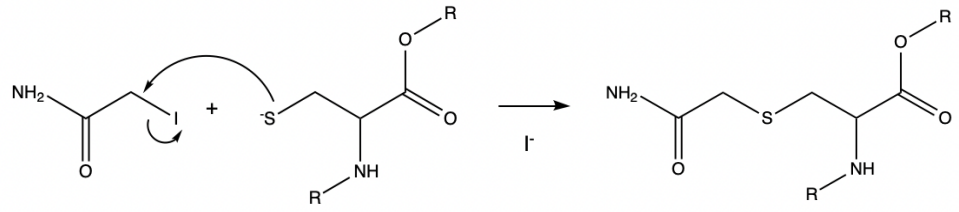

Figure 3: Chemical structure of the condensation reaction of IAM reacting with a free sulfur atom at the cysteine amino acid.

1.2.3 Deglycosylation of mAbs

Removing the glycans from the mAbs is called deglycosylation. After deglycosylation, the difference in mass between the intact glycoprotein and the deglycosylated glycopeptide can be calculated. Information about the composition and heterogeneity of the glycans can thereby be obtained. The process of deglycosylation involves proteolytic digestion of the glycoprotein into glycopeptides and cleavage of glycans from a glycoprotein. For the cleavage of the glycan, enzymes such as PNGase F can be used.[11][12]

1.3 MALDI-ToF-MS

Matrix-Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry (MALDI-ToF-MS) is a technique within the area of mass spectrometry. Mass spectrometry analysis implies detection of signals of separation based on the mass and the charge. The samples are exposed to a laser, which ionizes and mobilizes the matrix with the protein sample (*figure 4*). When the sample hits the detector a mass to charge ratio (m/z) can be determined since the

flight time is proportional to m/z. The collection of data is followed by analysis of the generated fragments.[13]

Trifluoroacetic acid (TFA) is often added as a counter ion source, to generate the ions [M+H] The protein samples are then mixed with a matrix (normally in 1:1 ratio) and loaded into marked spots on a stainless steel MALDI plate. The matrix is crystallized and laser energy absorbing. Sinapinic acid (3,5-Dimethoxy-4-hydroxycinnamic acid), DHB (2,5-dihydroxybenzoic acid) and α-CHCA (α-cyano-4-hydroxycinnamic acid) are the most used matrices. The matrices are mixed with acetonitrile, water and trifluoracetic acid (TFA) (0.1%), or with ethanol, water and TFA (0.1%). According to the choice of matrix, there are no general rules, mostly decided based on recommendations from the manufacturer regarding properties including reducing action and from trial and errors. Characteristics to consider are solubility, light absorbance, ionization capacity and volatility. In addition, obviously the matrix must not react with the analyte.[14]

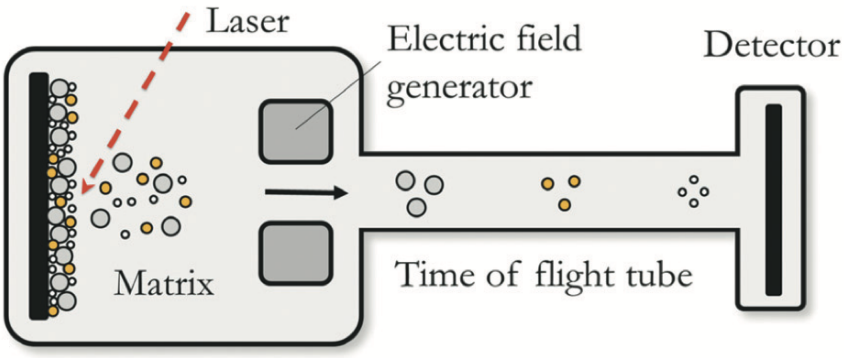

Figure 4: Simplified schematic figure of the process of MALDI-ToF-MS analysis including laser ionizes the protein sample on the plate, which is then passing through a electric field until it reaches the detector.[15]

1.3.1 Peptide Mass Fingerprinting (PMF)

Peptide mass fingerprinting (PMF) is an analytical tool used when identifying proteins. The protein is cleaved into smaller peptide fragments and then analysed with mass spectrometry, for example MALDI-ToF-MS, to determine the intensity and molecular weight of the fragments. The data obtained is then compared to a peptide database in order to identify the parent protein. By analysing the results, the best match can be found.[16]

1.3.2 LIFT

PMF is often followed by using LIFT (Laser Ionization Fragmentation Technology) in MALDI-ToF-MS, which enables MS analysis (amino acid sequencing). LIFT improves the

resolution by reaccelerating the specific ions from the sample and lowers the disrupting background signals. For accurate identification, the amino acid sequence must be annotated. LIFT technology is used for annotation of signature amino acids of peaks from the PMF spectra. An advantage of LIFT is the possibility to obtain low molecular weight fragments that are of low abundance.[17]

1.3.3 In Source-Decay (ISD)

In Source-Decay (ISD) is a rapid N-terminus fragmentation of selected proteins that occurs before the ion extraction in MALDI and can be used to annotate the partial amino acid sequence of the protein analysed. In comparison to conventional methods for mass spectrometric degradation, ISD allows determination of the amino-acid sequence without any pre-digestion.[18]

A flowchart of the identification process of an unknown protein, using above mentioned MALDI methods, is presented in *figure 5*.

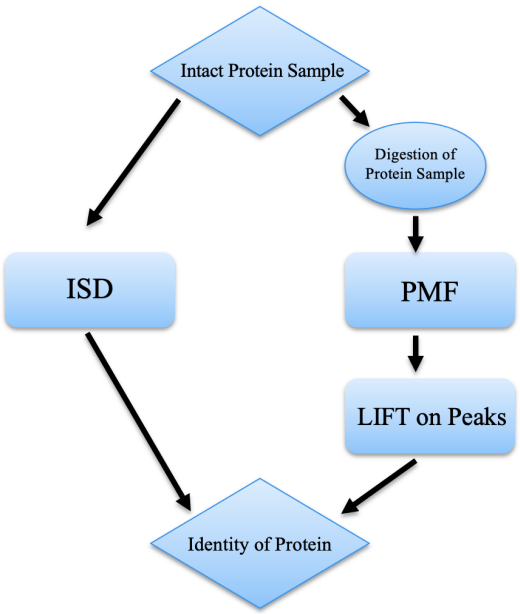

Figure 5: Flowchart of the process of identification of a protein. ISD is used on the protein in intact form, while for reduced form, PMF followed by LIFT of the peaks are performed.

1.4 Capillary Electrophoresis (CE)

Alternative analysis methods for separation and identification of mAbs are capillary electrophoresis (CE) methods, such as capillary zone electrophoresis (CZE) and capillary gel electrophoresis (CGE). These analysis techniques can also be combined with MALDI-ToF-

MS in an analysis method for identification of mAbs. The sample is hydrodynamically or electrokinetically injected into a capillary with silica coated inner walls. In CE, electrophoretic mobility separates molecules, which allows identification.

An electric field is created across the capillary as voltage is applied between the anode and cathode. This makes the analytes move at different mobilities towards the anode or cathode, depending on their hydrodynamic volume and charge. By applying voltage over the capillary, an electrophoretic flow is created which moves all molecules through the capillary. Electroosmotic flow arises from potential built up on the capillary surface. It is defined as the movement of a liquid with potential applied across the capillary (*figure 7*). The molecules are detected by a UV detector through a capillary window at the end of the capillary (*figure 6*).[19]

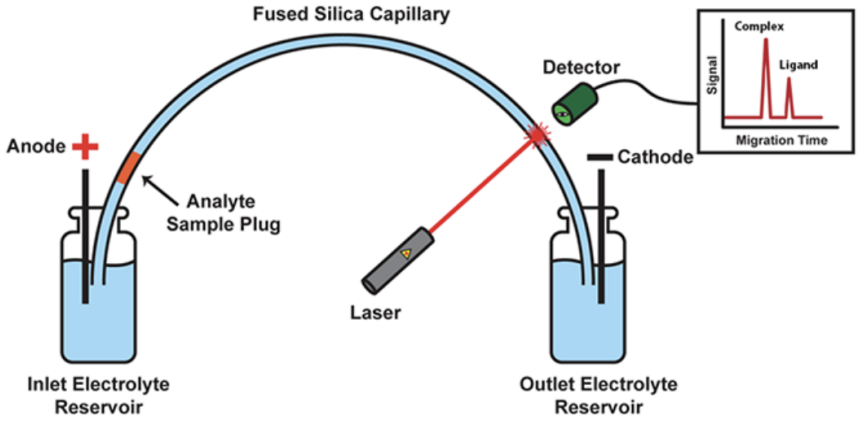

Figure 6: Schematic figure of CE analysis with capillary, anode, cathode, and detection.[20]

When performing CE analysis, a background electrolyte (BGE) keeps the pH constant. This is necessary because both electrophoretic mobility and electroosmosis are depending on pH. It also maintains the ion concentration necessary for the passage of the flow through the capillary. Other important properties of a BGE are good buffer capacity, low UV absorbance, low temperature coefficient and suitable conductivity for the passage of the flow.[21]

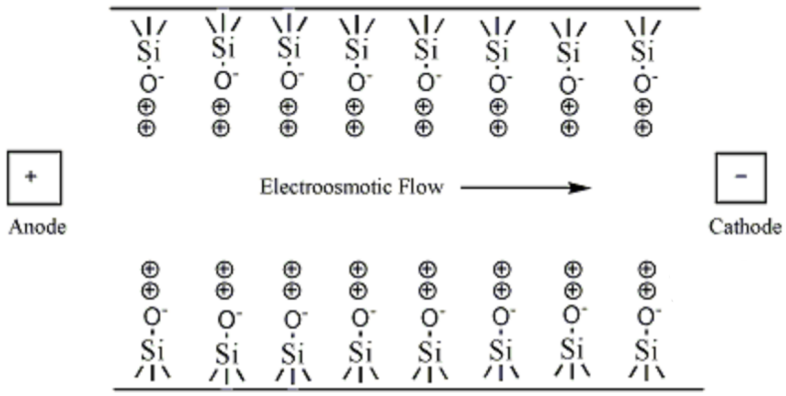

Figure 7: A buffer solution in a fused silica capillary.[22]

In both CGE and CZE, the capillary is rinsed with hydrochloric acid (HCl) to neutralize the silanol groups on the inner surface of the capillary. Sodium hydroxide is then used to rinse the capillary, remove the HCl and make the environment basic. The silanol groups are thereby deprotonated and the inner wall of the capillary is negatively charged.[23]

1.4.1 Capillary Gel Electrophoresis (CGE)

The BGE is viscous SDS-gel which separates the proteins, based on their hydrodynamic size. The sample preparation includes mixing the sample with SDS-MW Sample Buffer before denaturation at high temperature. The SDS will bind to the protein fragments and generate the same mass to charge ratios for all proteins. This is important for size-based separation by CE. This method can be applied with the purpose to validate the purity of the antibodies, by investigating if there are any significant differences between the HC and LC of the mAbs. The mAbs can be analysed in both non-reduced and reduced form.[24]

1.4.2 Capillary Zone Electrophoresis (CZE)

In capillary zone electrophoresis (CZE), the BGE is a buffer that keeps the pH stable and allows separation based on charge-to-size ratio. CZE can be applied to visualize isoform patterns of the mAbs as electrophoretic peaks with corresponding migration times. Triple injection capillary zone electrophoresis (TICZE) can be used to sandwich the protein sample between two known references, being sequentially injected into the capillary. TICZE generates higher accuracy because the reference and analyte are analysed simultaneously within a single run instead of multiple runs. The run-to-run variation is thereby avoided.[25]

1.5 Previous Studies

Previous studies involving analysis of mAbs have been done on a small scale, where a few mAbs were analysed with one type of analysis methods such as CE, CZE and CGE. In a study

by Dadouch *et. al.*, mAbs were analysed with both CGE and CZE. Based on the methods and results for the respective, a comparison could be made to ascertain which was most beneficial for the aim. The conclusion was that CZE, with its flexibility and simplicity, is more versatile, and generates comprehensive characterization of the physiochemical properties of the mAb.[11]

There is no general protocol for identification of mAbs or any analysis presented of mAbs consisting of a combinations of analysis methods. The possibilities for identification are known but not carried out on a large scale. Analysis of mAbs based on its immunological properties, such as antigen specificity is much more common. The reason for that is because the area of use in medicine for therapy this is generally of greater interest.

Identification of mAbs is done based on its structure and physiochemical characteristics. CE analyses provides, as already known, useful data of that. MALDI-ToF-MS is a simple and rapid analysis method, which suggests that it should be used for this aim if possible. Possibly, the mentioned analysis methods can complement each other and frame a protocol for rapid and accurate identification.

2. Materials

The project included 9 mAbs that are used for medicinal purposes. These are anonymized due to confidential reasons and therefore named the letters A-I. mAb B only consists of one chain instead of a heavy and a light chain, while mAb E and I are biosimilars.

2.1 Chemicals

The matrices used were purchased Bruker Daltonics. All other chemicals are purchased from MERCK, former Sigma Aldrich.

2.2 Equipment

Analysis instrument MALDI-ToF-MS autoflex maX-instrument from Bruker Daltonics with flexControl and flexAnalysis software (version 3.4) with appurtenant stainless-steel plate and CESI 8000 plus-instrument from Beckman Coulter with 32 Karat software (version 10.1) för MALDI and CE analyses respectively.

Other equipment used are heat block Thermomixer comfort (Eppendorf) for incubations. Scales XPR205DR, Mettler Toledo and Entris 623I-1S Sartorius and pH-meter SevenCompact, Mettler Toledo.

3. Methods

Several analyses of the mAbs were performed with MALDI-ToF-MS. The mAbs were also analysed with CE, non-reduced form with CZE and both reduced and non-reduced with CGE. Before analysis, the antibodies were treated as described under the following sections. For MALDI-ToF-MS, the matrix used with the dried-droplet method in each analysis respectively are also indicated.

3.1 MALDI-ToF-MS

Methods for all MALDI-ToF-MS-based analyses.

3.1.1 MALDI analysis of reduced mAbs

All mAbs included in the project were reduced by mixing the samples with 5 µl 2-ME followed by incubation in 80 °C for 5-10 min. The matrix was prepared by mixing a 1:1 ratio sinapinic acid in 50% ACN, 0,1% TFA and alpha-cyano-4-hydroxycinnamic acid (α-CHCA) in ethanol, ACN (1:1) and 0.1 % TFA. Samples of each mAb were placed on nine spots on the MALDI plate with reference series including all mAbs on each side of it (*figure 8*).

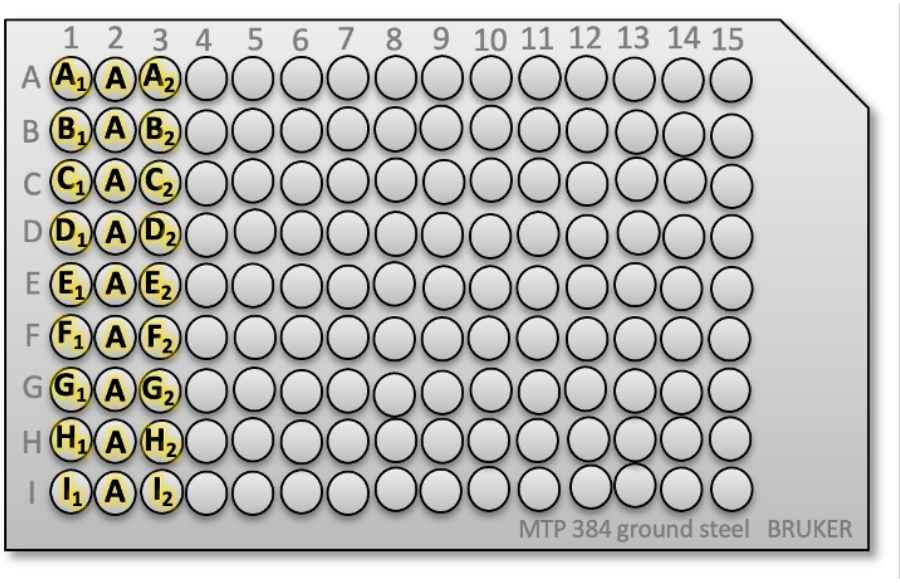

Figure 8: Sample A placed on nine spots in the middle with references A-I on both sides.

3.1.2 MALDI-In Source Decay (ISD)

Sample preparation included mAb protein, 1 µg/µl, in 50 mM ammonium bicarbonate buffer being reduced into light and heavy chains by 20 mM dithiothreitol (DTT) at 80°C for 15 min. The reduced sample was then buffer exchanged to 50% ACN/0,1% formic acid (FA), which is

compatible with MALDI. Firstly, an arbitrary mAb, C, was analysed by preparing two concentrations of the sample, 1 µg/µl and 17 µg/µl and centrifuge them (10 min, 14000 rpm) with varying centrifugal filters, 10 kDa and 30 kDa molecular weight cut off centrifugal filters. The filters select out every component smaller than 10 or 30 kDa. With filter 10 kDa, salts that can disturb MALDI analyses can be selected out. With filter 30 kDa, the LC and HC can be separated from each other. Combinations of the filters, 10 kDa & 10 kDa and 10 kDa & 30 kDa, were also included. After MALDI-analysis and based on the signals obtained, the most optimal concentration and filter for ISD-analysis could be determined, 17 µg/µl and 10 kDa, and used for further analysis of the remaining mAbs. The matrix used was sinapinic acid the same as in *3.1.1*.

The amino acid sequence of the light chain for each mAb was compared in order to identify unique signature amino acids for each mAb within the analysis range 1500-6000 m/z. These were then focused on when annotating the collected data.

3.1.3 Deglycosylation

For analysis of deglycosylated protein, ribonuclease was first used as a model protein followed by analysis of a mAb. The protein (2 mg) was dissolved in 0.1 % 500 µl RapiGest in 50 mM ammonium bicarbonate. The protein sample was reduced with 5 µl 2-ME and incubated in 40 ºC for 15 min. IAM 30 µl is added followed by incubation for 30 min away from light. Pronase is then added followed by incubation (37 ºC, 30 min). The sample was analysed with the SA matrix.

3.2 Capillary Electrophoresis (CE)

Methods for all CE-based analyses of reduced and non-reduced samples.

3.2.1 Capillary Gel Electrophoresis (CGE)

Methods for preparation and analysis of reduced and non-reduced samples. The samples were combined in 27 mixtures according to *table 1*. The mAb A was excluded due to its low initial concentration and the mAb B was excluded due its differing structure with only one chain. The analysis is performed in a fused silica capillary with the dimensions 30 cm × 20 cm × ID 50 µm (total length × effective length × inner diameter).

Table 1: All combinations of mAbs for analysis with CGE.

Combinations of the samples					
C+D	D+E	E+F	F+G	G+H	H+I
C+E	D+F	E+G	F+H	G+I	
C+F	D+G	E+H	F+I		
C+G	D+H	E+I			
C+H	D+I				
C+I					

Before separation, a single injection conditioning program (*table 2*) was performed to prepare the capillary for the separation. The voltage is negative because the injection is performed on the cathodic side instead of the anodic side and the EOF moves towards the anode. This is called reversed polarity mode.

Table 2: Conditioning Program performed before the separation, including time, event, pressure, voltage, duration, and additional comments. Cleaning of capillary ends is performed to avoid contamination to the next vial.

Conditioning Program					
Time [min]	**Event**	**Pressure [psi]**	**Voltage [kV]**	**Duration [min]**	**Comment**
	Rinse	30		3.0	NaOH
	Rinse	30		1.0	HCl
	Rinse	30		1.0	H_2O
	Rinse	30		10	Capillary is filled with the SDS-MW gel
	Waiting	-	-		Cleaning of the capillary ends.
	Injection		-5	10 s	Electrokinetic injection.
	Waiting	-	-		Cleaning of capillary ends.
0.00	Separation		-15	30	Reversed polarity mode.
5.00	Autozero	-	-		Resetting of baseline

The CGE separation was performed according to the separation program in *table 3*.

Table 3: Separation Program performed, including time, event, pressure, voltage, duration, and additional comments. Cleaning of capillary ends is performed to avoid contamination to the next vial.

Separation Program					
Time [min]	**Event**	**Pressure [psi]**	**Voltage [kV]**	**Duration [min]**	**Comment**
	Rinse	30		3.0	NaOH
	Rinse	30		1.0	BGE
	Injection		-5	1.0	Protein Sample
	Waiting			10	Cleaning of the capillary ends.
	Injection		-5	10 s	
	Waiting				Cleaning of the capillary ends.
0.00	Separation		-15	30	
5.00	Autozero				

3.2.1.1 Reduced Samples

The samples were prepared with a protein concentration of 1 mg/ml in MilliQ-H_2O in a total volume of 700 µl. SDS-MW Sample Buffer (350 µl) and 2-ME (10 µl) were added followed by incubation at 65°C for 10 min and then cooled down in RT. The samples were then centrifuged (3 min 1400 rpm). The samples were mixed in different combinations with the ratio 1:1 in a total volume of 200 µl in 0.2 ml micro tubes without lid placed in universal CE vials with caps. The samples were first combined according to *table 1*. After analysis of non-reduced samples (*3.2.1.2*), the samples that showed separation were analysed one by one in reduced form (C, D, E, F, H & I) to determine which peak belongs to which mAb.

In the case where the biosimilars are combined with sample C, the results were inconsistent although they should be the same according to the biosimilarity. With the aim to ensure this

result, new samples, C+E & C+I, were prepared and analysed. The new samples were prepared by first mixing the two mAbs to the concentration of 1 µl/ml in MilliQ-H_2O to a total volume of 100 µl followed by addition of SDS-MW Sample Buffer (100 µl) and 2-ME (10 µl).

Lastly, all samples in *table 1* were analysed in reduced form again. The sample preparation differed from the previous analysis of all combinations. Instead of treating each mAb sample separately before combining them, the mAbs in MilliQ-H_2O (1 mg/ml) were mixed 1:1 first, followed by treatment of SDS-MW Sample Buffer and 2-ME.

3.2.1.2 Non-Reduced Samples

The samples were prepared with a protein concentration of 1 mg/ml in MilliQ-H_2O in a total volume of 700 µl. SDS-MW Sample Buffer (350 µl) and IAM (10 µl) were added followed by incubation at 65°C for 10 min and then cooled down in RT. The samples were then centrifuged (3 min 1400 rpm). The samples were mixed in different combinations with the ratio 1:1 in a total volume of 200 µl in 0.2 ml micro tubes without lid placed in universal vials with caps. The samples analysed in non-reduced form are those combinations from *table 1* that showed separation at any peak, for LC or HC, in the analysis of reduced samples, (C+G, D+G, E+H & F+H) (*3.2.1.1*).

3.2.2 Capillary Zone Electrophoresis (CZE)

A fused silica capillary with dimensions 30 cm × 20 cm × ID 50 µm (total length × effective length × inner diameter) was used. To prevent electrostatic interactions between the positively charged mAb samples and the capillary wall, the capillary is rinsed with the cationic polymer polybrene (PB). PB is absorbed to the negatively charged inner surface of the capillary to make it positively charged [26]. The capillary is then washed with the buffer before injection of sample.

The mAbs are analysed with TICZE, one sample mAb and two references mAbs, in intact form. The BGE contains ε-aminocaproic acid (eACA), hydroxypropyl methylcellulose (HPMC) and lithium acetate (LiOAc). eACA has a zwitterionic form, which means that it consists of an equal number of functional groups that are positively and negatively charged. When using eACA at higher ionic strengths, the separation resolution is enhanced, and the electroosmotic flow is lowered. Its “neutral” charge make it not affecting the ionic strength of the BGE [27]. HPMC is a polymer that with water forms a colloid with gel consistency, which coats the inner surface of the capillary to avoid interaction with the mAb samples and enhances the viscosity of the BGE [28]. LiOAc neutralizes the silanol groups on the silica by positively charged lithium ions binding to them [29]. This facilitates the mAbs to pass easily through the capillary by decreasing the electrostatic interaction.[11]

The samples were prepared by diluting each mAb A-I in MilliQ-H_2O to a concentration of 1 mg/ml. The BGE was prepared by mixing three solutions, eACA (8 ml), HPMC (2 ml) and LiOAc (204 µl) to final concentrations of 39 mM eACA, 0.2 % HPMC and 25 mM LiOAc and a pH of 4.5. eACA (0.66 g) was dissolved in 95 ml MilliQ-H_2O. Glacial acetic acid was used to adjust the pH to 4.5 followed by adding MilliQ-H_2O to a total volume of 100 ml and final concentration of 50 mM eACA. The solution was then filtered through a 0.22 µm filter before use. HPMC (1 g) was dissolved in 80 ml heated water while stirring. The volume was then adjusted to 100 ml before cooling down the solution. LiOAc (3.3 g) was dissolved in 50 ml MilliQ-H_2O to the final concentration of 1000 mM. Polybrene (PB) was prepared by dissolving 25 mg in 5 ml H_2O to the concentration 5 mg/ml. TICZE with samples (200 µl) of each mAb A-I together with two references (200 µl) of the same mAb are analysed for each combination.

Before sample analysis, a program (*table 4*) was performed to prepare the capillary.

Table 4: Conditioning Program performed, including time, event, pressure, voltage, duration, direction, and additional comments.

Conditioning Program						
Time [min]	**Event**	**Pressure [psi]**	**Voltage [kV]**	**Duration [min]**	**Direction**	**Comment**
	Rinse	20		5.0	forward	HCl
	Rinse	20		5.0	forward	NaOH
	Rinse	30		5.0	forward	PB
	Rinse	30		5.0	forward	BGE
	Injection	1.0		10.0 s	forward	Sample injected.
0.00	Separation		30	16.0		Capillary ends dipped in BGE.
5.00	Autozero					
16.01	Stop data					
16.02	Rinse	30		2.5		Water
18.03	Wait			0.04		
18.04	End					

The CZE separation was performed according to the separation program in *table 5*.

Table 5: Separation Program performed, including time, event, pressure, voltage, duration, direction, and additional comments.

Separation Program						
Time [min]	**Event**	**Pressure [psi]**	**Voltage [kV]**	**Duration [min]**	**Direction**	**Comment**
	Rinse	30		3.0	forward	PB
	Rinse	30		5.0	forward	BGE
	Injection	1.0		10.0 s	forward	Override
	Separation		30	2.0		0.30 min ramp, normal polarity
	Injection	1.0		10.0 s	forward	Override
	Separation		30	2.0		0.30 min ramp, normal polarity
	Injection	1.0		10.0 s	forward	Override
0.00	Separation		30	15		0.30 min ramp, normal polarity
5.00	Autozero					

4. Results

The results for each analysis respectively are presented in the following sections.

4.1 MALDI-ToF-MS

The results from all MALDI-ToF-MS analyses.

4.1.1 Analysis of Reduced mAbs

The experimental data from the analysis of all mAbs in reduced form includes molecular weight of the light and heavy chains. These values are compared to the references to identify significant differences unique for each mAb respectively. The differences between the analysed samples and the references are obtained by calculating a value of the relative molecular mass (RMM) of both LC and HC respectively (*equation 1*).

$$RMM = \frac{S}{\left(\frac{Ref_1 + Ref_2}{2}\right)} \quad (1)$$

S, Ref_1 and Ref_2 are the molecular mass of the sample and references 1 and 2 respectively. The references are numbered according to the location on the MALDI plate, Ref_1 to the left of the sample and Ref_2 to the right (*figure 9*).

Theoretically, for a sample with references from the same mAb, the RMM value received should be closest to 1. In two cases, sample E with reference E and sample F with reference F, the samples could not be distinguished from others due to very similar results. The results are presented in *table 6* and *figure* 9 shows an example of how a spectrum can look like (sample E with reference A_2). Here peaks for LC and HC in mAb E are visible, similar spectrum were obtained for both references, 1 and 2 respectively.

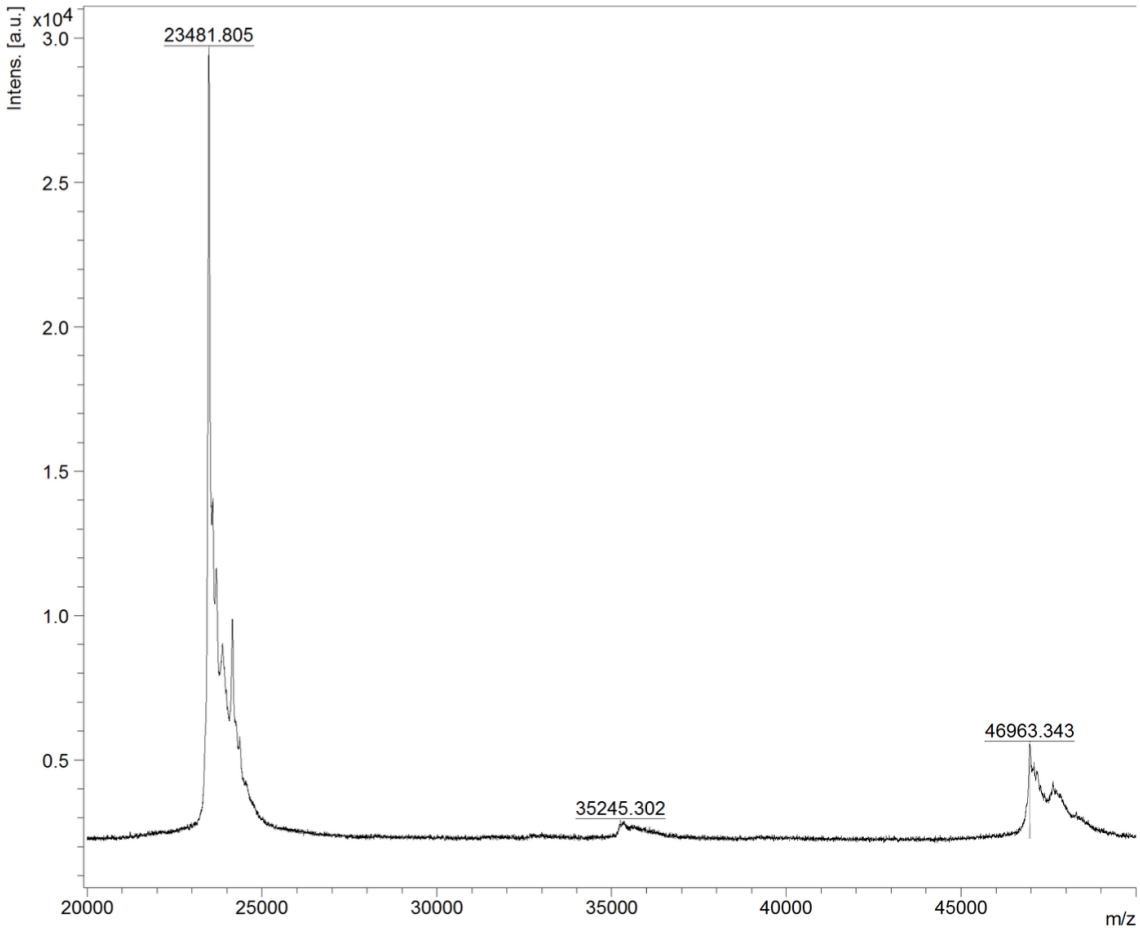

Figure 9: An example of MALDI-spectrum of, mAb E (with the reference E_2). For mAb E, the peak for LC is 23481.805 m/z and for HC 46963.343 m/z.

Table 6: The values that hypothetically should be closest to one are in grey cells. Those values that are closest to one are in extra bold type. The RMM values closest to 1 coincide in almost all cases except sample E with reference E and sample F with reference F (bold type underlined). The false positive results (1.0000 ± 0.0004, p = 95%) are in italics and underlined.

Sample		**Reference**								
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
A	**LC**	**1.00030**	0.94876	1.00205	1.00102	1.00172	0.99787	1.00535	1.00503	1.00162
	HC	**1.00047**	1.27250	1.00191	1.00121	1.00216	0.99804	1.00585	0.96175	1.01563
B	**LC**	1.05442	**0.99993**	1.05619	1.05574	1.03780	1.05187	1.05958	1.03605	1.03835
	HC	1.04467	**0.99951**	1.04898	1.04619	1.04647	1.04279	1.05030	1.00540	1.04774
C	**LC**	0.99808	0.96356	**1.00004**	0.99953	*0.99977*	0.99613	1.00345	0.95982	*0.99965*
	HC	0.99809	0.95402	**0.99992**	0.99955	*0.99988*	0.99618	1.00361	0.95986	*0.99984*
D	**LC**	0.99896	0.96421	1.00056	**0.99998**	*1.00032*	0.99667	1.00408	0.96041	*1.00038*
	HC	0.99858	0.95608	1.00051	**1.00017**	*1.00017*	0.99680	1.00389	0.96029	1.00041
E	**LC**	0.99837	0.96372	*1.00036*	*0.99961*	0.99987	0.99645	1.00376	0.96023	**1.00006**
	HC	0.99824	0.95518	*1.00021*	*0.99960*	0.99971	0.99651	1.00378	0.96032	**1.00017**
F	**LC**	1.00195	0.96687	1.00045	**1.00005**	1.00353	0.99990	1.00717	0.96369	1.00361
	HC	0.99972	0.95874	*1.00030*	**0.99996**	1.00352	0.99972	1.00740	0.96321	1.00344
G	**LC**	0.66108	0.96012	0.99661	0.99606	0.99644	0.99262	**1.00006**	*0.99971*	0.99620
	HC	0.99265	0.95191	0.99535	0.99605	0.99626	0.99282	**0.99984**	0.95641	0.99631
H	**LC**	1.04040	1.00364	0.99688	1.04127	1.04154	1.04220	1.04963	**1.00002**	1.04587
	HC	1.03148	0.99500	1.04169	1.04115	1.04141	1.03786	1.04541	**0.99998**	1.04169

I	**LC**	0.99863	0.96343	1.00044	*0.99960*	*1.00019*	0.99645	1.00368	1.00329	**0.99992**
	HC	0.99821	0.95411	*1.00017*	0.99932	*1.00005*	0.99647	1.00369	0.95995	**1.00000**

In almost all cases, the calculated RMM values for sample and reference of the same mAb are closest to 1. In two cases, sample E with reference E and sample F with reference F, the samples could not be distinguished from others due to very similar results. Instead, the RMM values closest to 1 are for sample E with reference I and sample F with reference D. This indicates that there is a risk of obtaining false positive results due to the small differences between the mAbs. All values that are within the range 1.0000 ± 0.0004 (p = 95%) are classified as false positives and marked in *table 6*.

According to the sample E with reference I, the result can be explained by the fact that they are biosimilars and therefore have a very similar molecular weight. For sample I with reference E, the RMM values are also very close to 1 but not as close as for sample I with reference I, compare reference E LC 1.00019 (1.00000 ± 0.00019) to reference I LC 0.99992 (1.00000 ± 0.00008). This similarity is expected due to the biosimilarity. According to sample F with reference D, CGE did not show any major similarity (*4.2.1.1*) and the result from CZE is partial separation of HC and no separation for LC (*4.2.2*). This indicates that the result is false positive.

Regarding all false positive results, sample C with references E and I, sample D with reference E, sample E with references C and D and sample I with reference E have false positive values for both LC and HC. According to the molecular mass of the chains, E and I have the same for both LC and HC and so do C and D. For the mAbs that have false positive results for only chain, the similarity is not as great, and they are still distinguishable.

4.1.2 MALDI-In Source Decay

Data was collected for each mAb, and the amino acid sequences were annotated with focus on the identified signature amino acids. There were varying qualities of collected data for the mAbs, but for the majority, the annotation was relatively good and for a few, longer continuous amino acid sequences could be identified. Estimation of the results for each mAb are presented in *table 7*. An example of an annotated amino acid sequence can be seen in *figure 11* (mAb G) and compared to the highlighted part of the sequence of the mAb in *figure 10*.

>DIQMTQSPSSLSASVGD**RVTITCQASQDIS**N**YLNWYQQ**KP**G**KAPKLLIYDASNLETGVPSRF
SGSGSGTDFTFTISSLQPEDIATYFCQHFDHLPLAFGGGTKVEIKRTVAAPSVFIFPPSDEQL
KSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
HKVYACEVTHQGLSSPVTKSFNRGEC

Figure 10: The amino acid sequence of the mAb G. The sequence within the range in the spectra (1500-6000 m/z) is underlined.

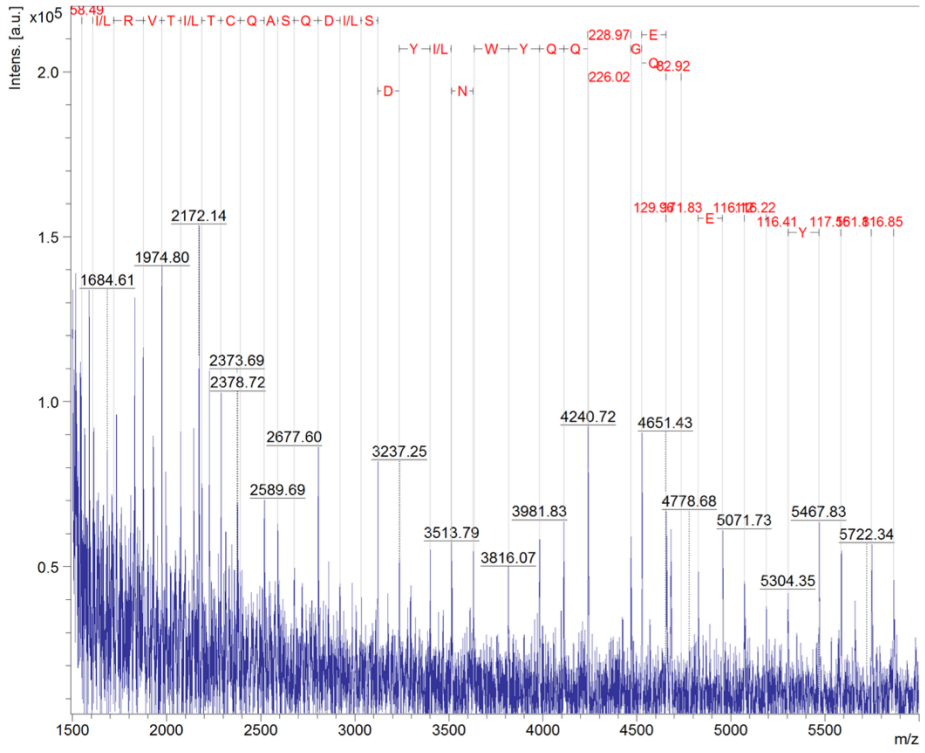

Figure 11: Annotated amino acid sequence of mAb G. The amino acids are presented with the 1-letter abbreviations. In the places where there are numbers instead of an amino acid, the distance corresponds to one or two of the amino acids in the sequence of the mAb. Annotation of these was not possible due to poor signals for the exact distances.

Comparing the annotation (*figure 11*) with the visible amino acid sequence (<6000 kDa) (*figure 10*), 51% of the amino acid sequence could be annotated. Among these, there are several signature amino acids annotated, thereby the result can be classified as good. The percentage annotated amino acid for each mAb is presented in *table 7*.

Table 7: Estimation of results of amino acid annotation of each mAb A-I.

mAb	Amino acids annotated [%]	Estimation of Amino Acid Annotation
A	3	Poor signals, single amino acid annotated but no sequence longer than one amino acid.
B	33	Relatively good signals for the most part of it, most of the amino acids, including unique signature amino acids, could be annotated.
C	43	Mostly good signals, where a longer amino acid sequence, including several unique ones, could be annotated.
D	10	Partially good signals, overall difficult to annotate, a few amino acids can be annotated.
E	7	Poor signals, a few amino acids could be annotated, but none are unique.
F	15	Partially good signals, unique amino acids could be annotated.
G	51	Good signals, substantially the entire sequence could be annotated, including unique amino acids.
H	12	Weak signals, some single amino acids could be annotated.
I	0	Poor signals, no amino acids were annotated.

For the majority of the mAbs, the amino acid annotation obtained with ISD was relatively good. For some of the mAbs longer continuous amino acid sequences could be identified. According to the comparison between the amino acid sequences of the LCs of respective mAb, the signature amino acids could be annotated of the signals for four of the nine mAbs.

4.1.3 Deglycosylation

The result from analysis of deglycosylated ribonuclease is presented in *figure 12*. The degree of deglycosylation can be obtained as each peak in the spectrum corresponds to the protein with a specific number of mannose units attached to it. For comparison, *figure 13* shows another MALDI analysis of deglycosylation of ribonuclease. The carbohydrates for each peak are drawn with the respective number of mannose units.

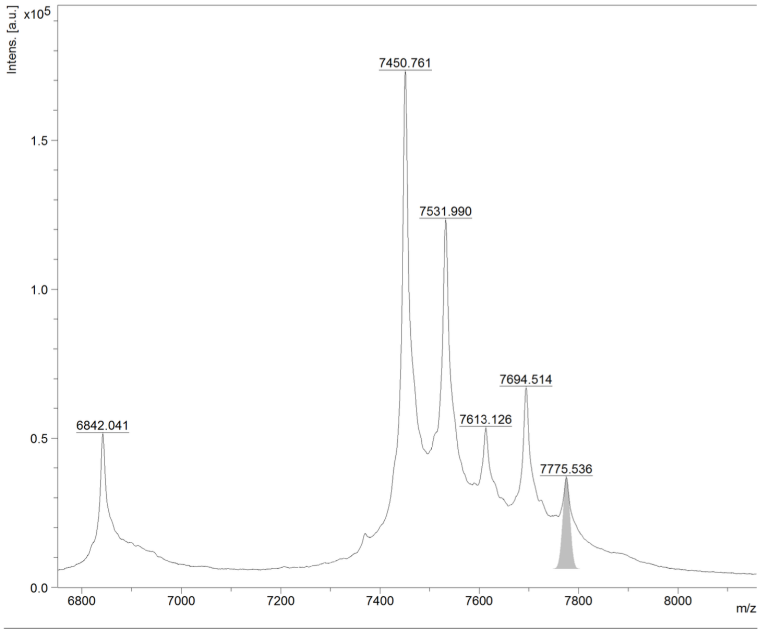

Figure 12: MALDI analysis of deglycosylation of ribonuclease.

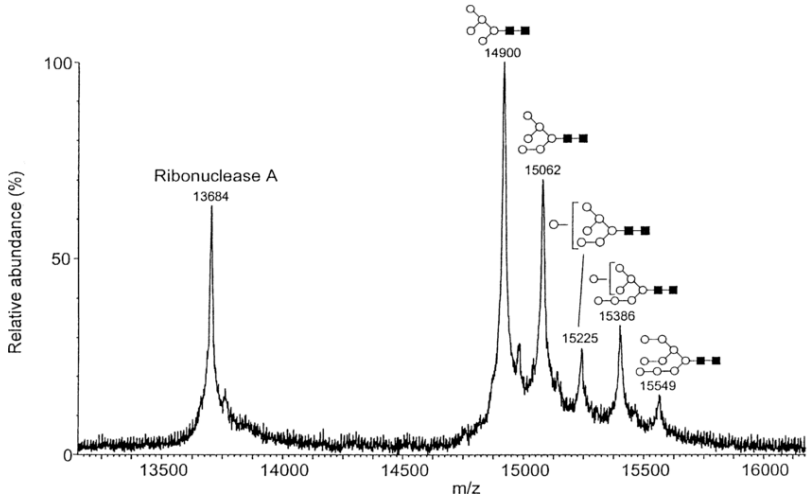

Figure 13: Example of deglycosylation of ribonuclease with carbohydrates specified for each peak.[30]

Clear similarities can be seen between the spectrum obtained (*figure 12*) and the example spectra (*figure 13*).

The deglycosylation of the mAb did not give any result. Possible reasons for that are discussed in *5.1.3*.

4.2 Capillary Electrophoresis (CE)

The results from all CE analyses, CGE and CZE.

4.2.1 Capillary Gel Electrophoresis (CGE)

Results of the capillary gel electrophoresis for reduced and non-reduced samples.

4.2.1.1 Reduced Samples

The results of analysis of reduced samples (*table 1*) are presented in *table 8*. The separation of the peaks for LC and HC between the two mAbs in the electropherograms are estimated and categorized as baseline separation (BS) (*figure 14*), partial separation (PS) (*figure 14*), tendency to separation (TS) (*figure 15*) and non-tendency to separation (NS) (*figure 16*).

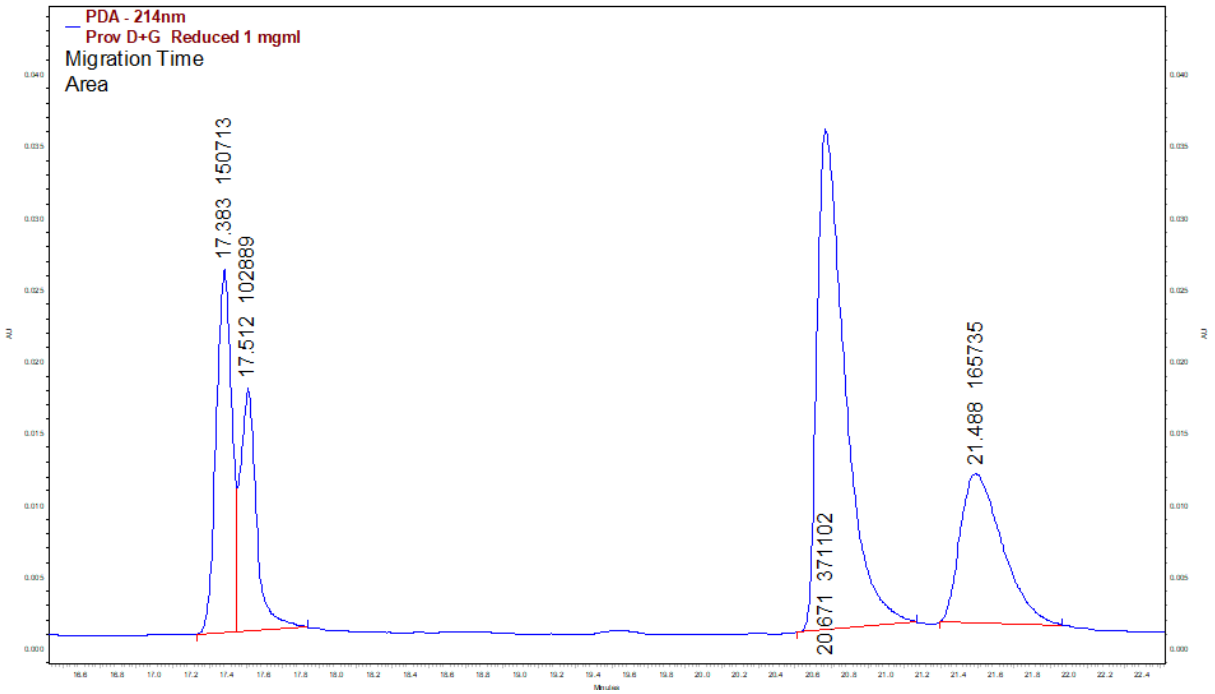

Figure 14: Sample D+G, partial separation (PS) for LC (peaks to the left) and baseline separation (BS) for HC (peaks to the right).

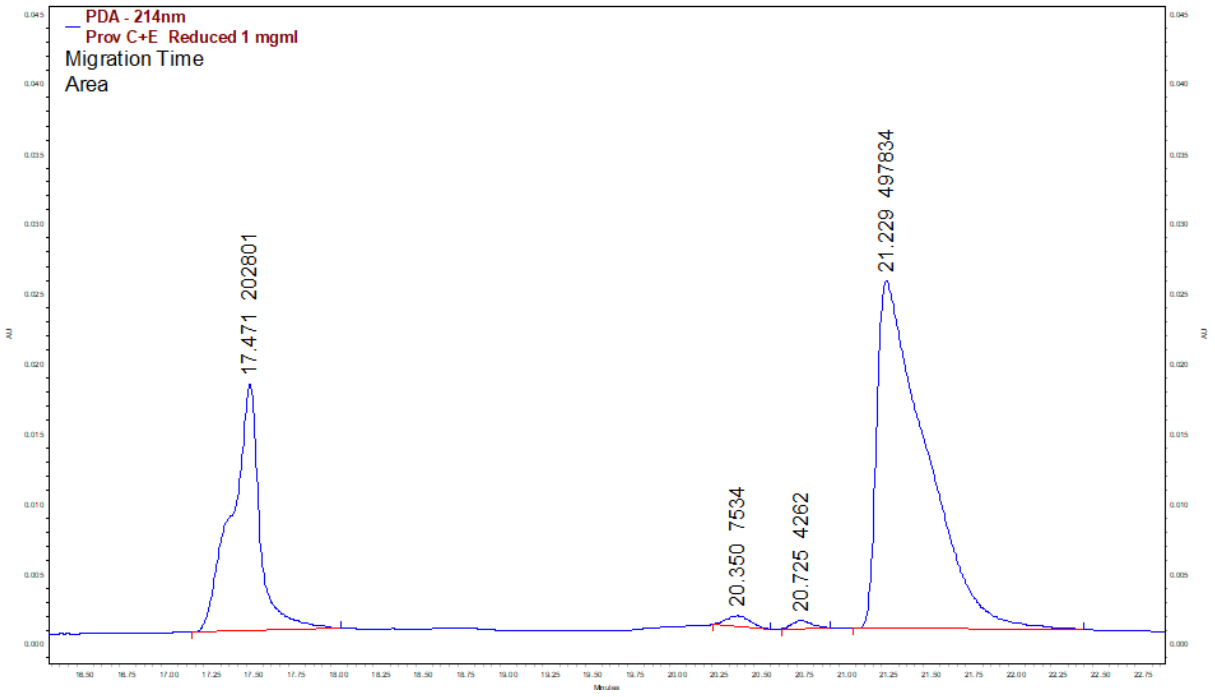

Figure 15: Sample C+E, tendency to separation (TS) for both LC and HC.

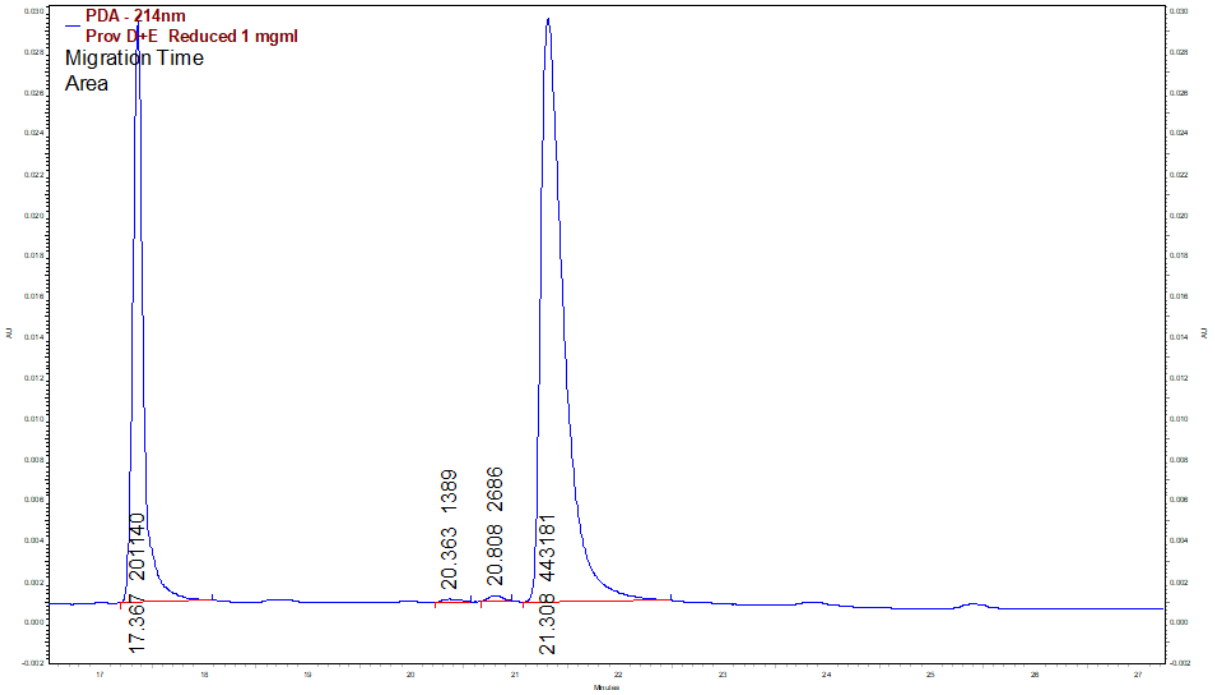

Figure 16: Sample D+E, non-tendency separation (NS) for LC and HC.

Table 8: Estimation of CGE Results. The separation between LC and HC for each sample is categorized as baseline separation (BS), partial separation (PS), tendency to separation (TS) and non-tendency separation (NS).

Separation of LC and HC for combinations of mAb C-I							
mAb		**D**	**E**	**F**	**G**	**H**	**I**
C	**LC**	**PS**	**TS**	**PS**	**PS**	**TS**	NS
	HC	NS	**TS**	NS	NS	NS	NS
D	**LC**		NS	NS	**PS**	NS	NS
	HC		NS	NS	**BS**	NS	NS
E	**LC**			NS	**TS**	NS	NS
	HC			NS	**TS**	**BS**	NS
F	**LC**				NS	NS	NS
	HC				**TS**	**BS**	NS
G	**LC**					NS	**TS**
	HC					NS	NS
H	**LC**						NS
	HC						**BS**

Reanalysis of all combinations except with A and B (*table 1*) with changed sample preparation (*table 9*). B is removed due to the results in *table 8*.

In *table 8*, both LC and HC achieved different separation for E+C (TS for both chains) and I+C (NS for LC and BS for HC). For E+G and I+G, the HC got TS and NS respectively. Because TS is more of a matter of interpretation when visually determining the results and closer to NS than clear PS and BS, this result should not be given so much importance. For HC, TS compared to BS differ a lot from each other and the risk for misinterpretation is minimal. This sample was therefore reanalysed and the new results for E+C and I+C were almost consistent and agree with the results in *table 9*.

Table 9: Estimation of results from CGE reanalysis of all combinations, except with A and B, and new sample preparation. The separation between LC and HC for each sample is categorized as baseline separation (BS), partial separation (PS), tendency to separation (TS) and non-tendency to separation (NS).

Separation of LC and HC for combinations of mAbs C-I in reanalysis							
mAb		**D**	**E**	**F**	**G**	**H**	**I**
C	**LC**	**PS**	**PS**	**PS**	**PS**	**PS**	**PS**
	HC	NS	**TS**	**TS**	NS	**BS**	NS
D	**LC**		NS	NS	NS	NS	NS
	HC		NS	**PS**	NS	**BS**	NS
E	**LC**			NS	NS	NS	NS
	HC			NS	**TS**	**BS**	NS
F	**LC**				NS	NS	NS
	HC				**TS**	**BS**	NS
G	**LC**					NS	NS
	HC					**BS**	NS
H	**LC**						NS
	HC						**BS**

In *table 9*, the separation of the HC peaks for E+C and I+C was TS and NS respectively and for HC for E+G and I+G, the separation also was TS and NS respectively. The same reasoning regarding the diffuse boundary between TS and NS can be adopted here as well. The separation of the LC peaks is consistent with each other between the biosimilars, PS for E+C and I+C and NS for E+G and I+G. These results were expected. The reason for separation depends on two properties of the mAbs, the molecular mass and the number of amino acids. All mAbs have the same number of amino acids in the LC, while the number of amino acids in the HC differs between them. The biggest difference is between H and the

others, 240 and 221-231 amino acids respectively. According to the molecular mass of the HC, H stands out here as well, 25646 Da. The molecular masses of the HC of the other mAbs are within 23654-24869 Da. This explains why BS is only achieved between H and the other mAbs.

The results of the analysis of the samples C, D, E, F, H & I on reduced form was done after analysis of non-reduced samples with the aim to determine which peaks belong to which mAb. This data (*figure 17a-f*) is used to confirm if separation has occurred in *4.2.1.2*.

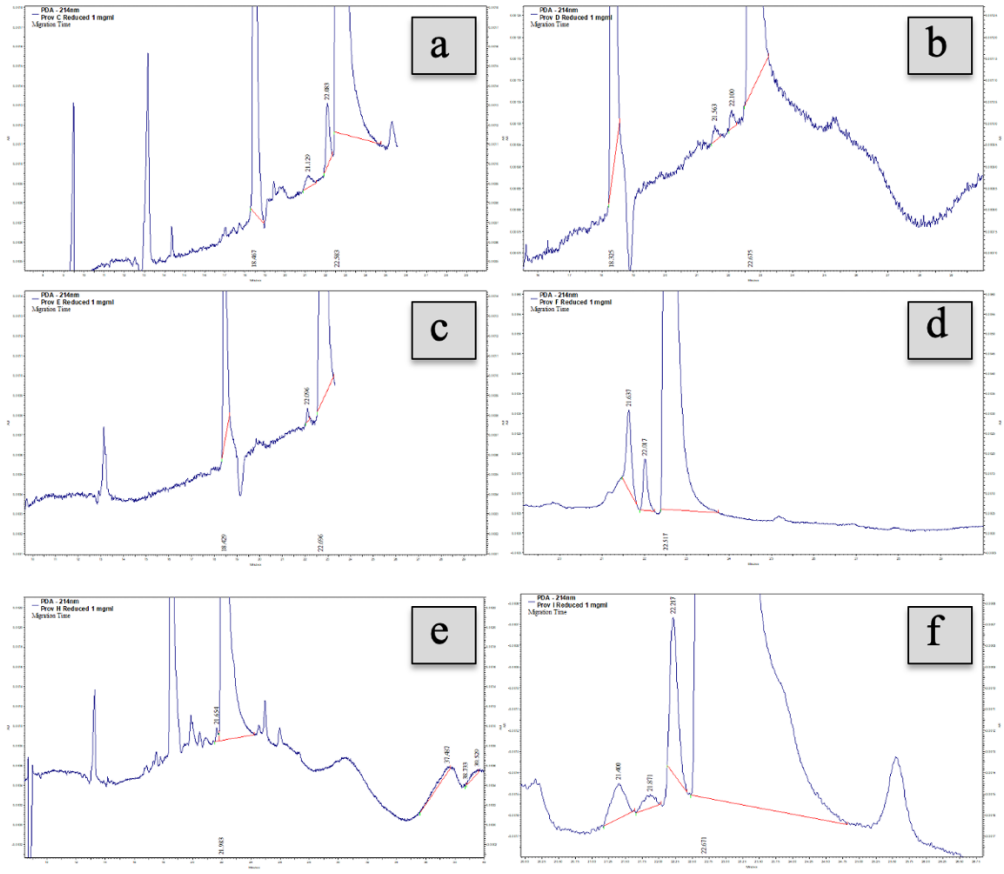

Figure 17a-f: mAb C, D, E, F, H, and I analysed separately.

The results of the reanalysis of the biosimilars E and I combined with C, C+E and C+I, were PS for LC in both combinations and for HC, TS, and NS respectively.

4.2.1.2 Non-Reduced Samples

The results of analysis of the samples that showed separation in table 8 (C+D, C+E, C+F, C+G, C+H, C+I, D+G, E+G, E+H, F+G, F+H G+I and H+I) in non-reduced form shows if separation of the peaks from the respective antibody is visible or if the separation can only be

seen between the two chains, LC and HC, in reduced form. The samples C+G, D+G, E+H, F+H, and H+I are shown in *figure 18a-e*. Partial separations between the mAbs can be seen for samples E+H, F+H, and H+I, while NS/TS prevails between the mAbs in C+G and D+G.

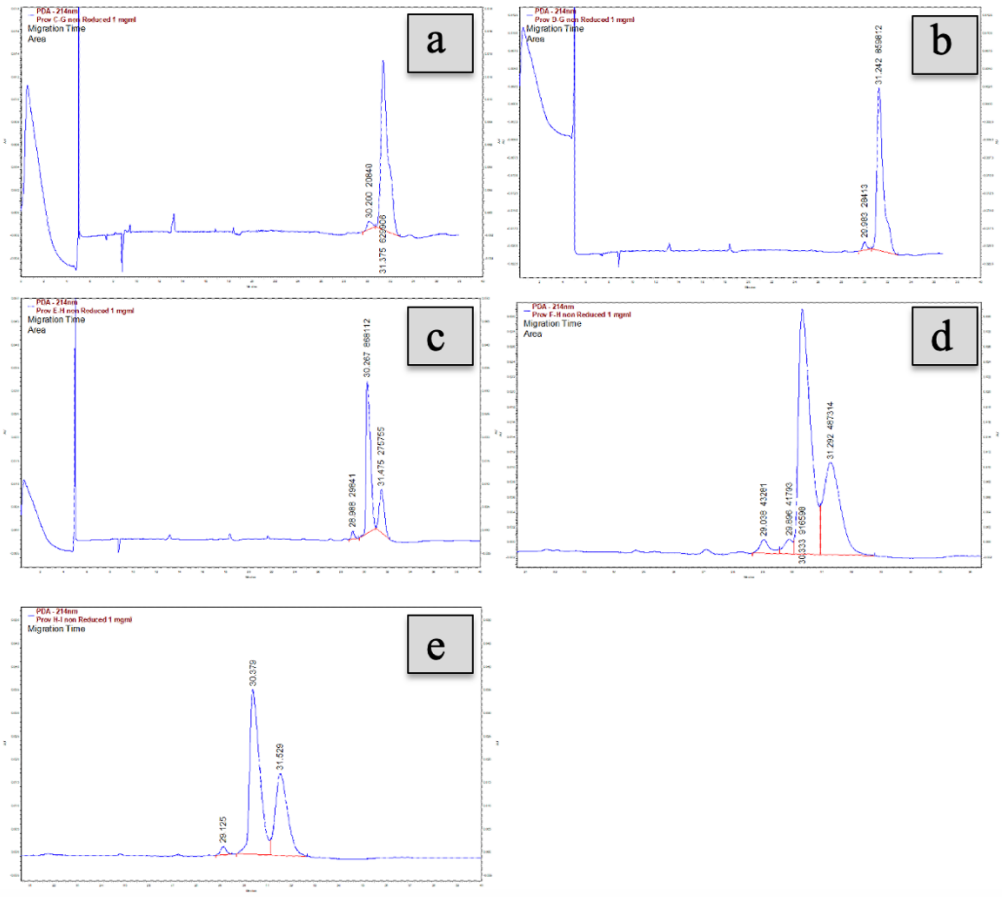

Figure 18 a-e: mAb combinations C+G, D+G, E+H, F+H and H+I in non-reduced form. NS and TS for samples C+G and D+G. Partial separation between the mAbs in samples E+H, F+H, and H+I.

4.2.2 Capillary Zone Electrophoresis (CZE)

The results of CZE analysis of every combination of A-I are presented in *table 10*. *Figure 19* and *20* show examples of one electropherograms of a protein sample with the same reference (sample C with reference C) and one electropherogram with another reference (sample C with reference E) respectively. The relative migration time (RMT) is calculated for each combination of sample and reference (*equation 2*) and presented in *table 10*.

$$RMT = \frac{\left(\frac{Ref_2 - Ref_1}{2} + S\right)}{Ref_2} \qquad (2)$$

S, Ref_1 and Ref_2 are the migration time of the sample and references 1 and 2 respectively. The references numbered according to the order they have been injected.

In some cases where the separation patterns of sample and reference from two different mAbs are very similar to each other, the area of the peaks can be used to determine to which mAb the peak belongs. By calculating and comparing the ratios of the areas of two peaks within the same separation pattern, the separation pattern from the references can be distinguished from the separation patterns from the sample.

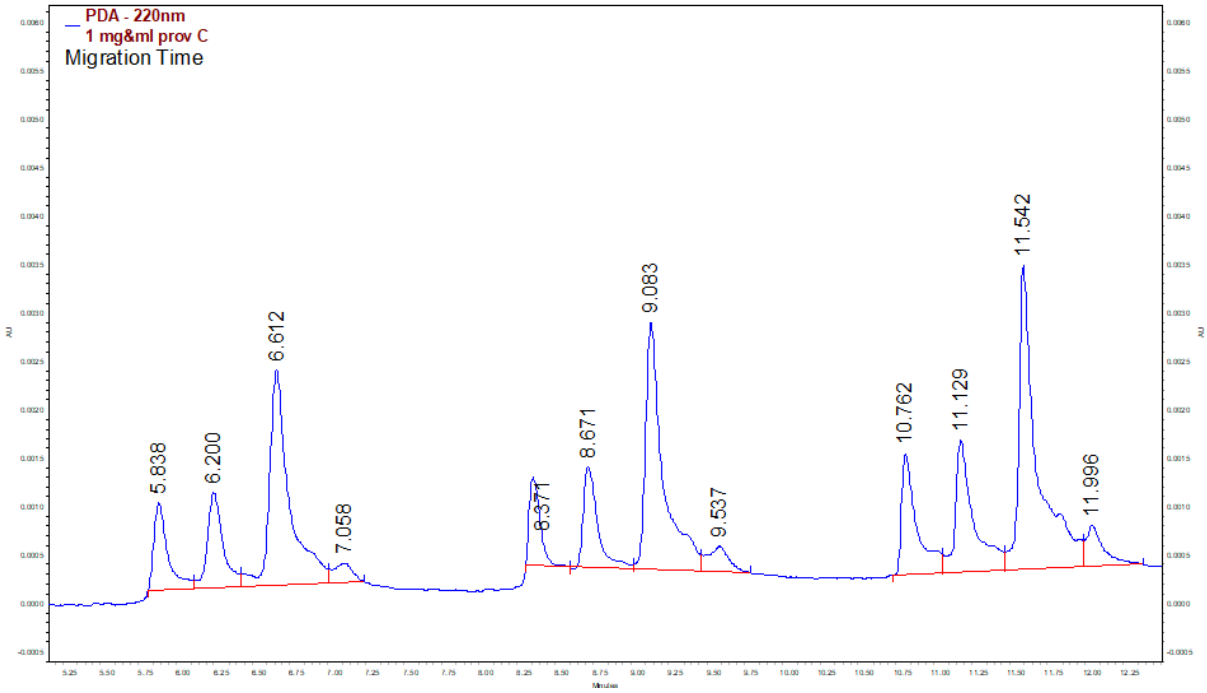

Figure 19: CZE electropherogram of sample C in the middle (4 peaks) with reference C on both sides of it (4 peaks respectively).

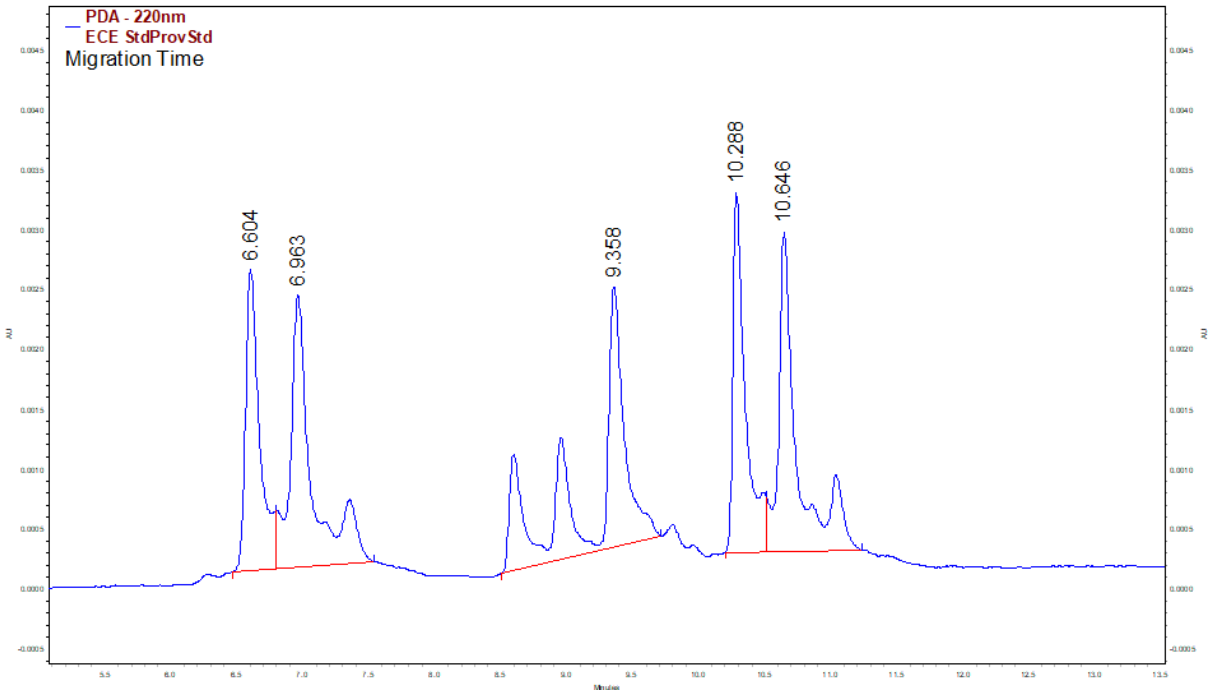

Figure 20: CZE electropherogram of sample C in the middle (4 peaks) with reference E on both sides of it (3 peaks respectively).

Table 10: CZE results including RMT values and separation patterns for each combination of sample and reference respectively. No peaks shown for sample I with reference B. In every combination, the separation pattern for each mAb is defined by the numbers of peaks. In cases where the peaks overlap and the number thus becomes fewer, are marked with an asterisk. The RMT values closest to 1 coincide in almost all cases except sample I with reference I (bold type underlined). The false positive results (1.000 ± 0.005, p = 95%) are in italics and underlined.

Sample		**Reference**								
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
A	**RMT**	**1.0000**	1.0423	1.1469	1.1187	1.2083	1.2750	0.7945	1.1196	1.0850
	Separation Pattern	A:3 A:3	A:3 B:1	A:1* C:4	A:2* D:3	A:2* E:3*	A:2* F:2	A:3 G:2*	A:2* H:2*	A:3 I:3*
B	**RMT**	0.9596	**1.0005**	1.0958	1.0911	0.9549	1.0068	0.7635	1.0736	0.9927
	Separation Pattern	B:1 A:2*	B:1 B:1	B:1 C:4	B:1 D:3	B:1 E:3*	B:1 F:2	B:1 G:3	B:1 H:2*	B:1 I:4
C	**RMT**	0.8714	0.9091	**1.0005**	0.9753	1.0886	1.1087	0.6924	0.9708	1.0817
	Separation Pattern	C:3* A:1*	C:4 B:1	C:4 C:4	C:4 D:3	C:4 E:3*	C:3* F:2	C:4 G:2*	C:4 H:2*	C:4 I:4
D	**RMT**	0.8920	0.9213	1.0381	**1.0020**	1.0785	1.2094	0.7143	*1.0049*	1.1162
	Separation Pattern	D:3 A:2*	D:3 B:1	D:3 C:4	D:3 D:3	D:3 E:3*	D:2* F:2	D:3 G:3	D:3 H:2*	D:3 I:4
E	**RMT**	0.8276	0.9648	0,9303	0.9000	**1.0026**	1.0212	0.6403	0.9340	*1.0032*
	Separation Pattern	E:2* A:2*	E:3* B:1	E:4 C:4	E:3* D:3	E:4 E:4	E:4 F:2	E:3* G:3	E:3* H:2*	E:4 I:4
F	**RMT**	0.7862	0.8160	0.9120	0.8837	0.9840	**1.0027**	0.6292	0.8852	0.9841

	Separation Pattern	F:2 A:2*	F:2 B:1	F:1* C:4	F:1* D:3	F:2 E:4	F:2 F:2	F:2 G:3	F:1* H:3	F:2 I:4
G	**RMT**	1.2544	1.2982	1.4507	1.4073	1.5681	1.5978	**1.0011**	1.4150	1.5764
	Separation Pattern	G:2* A:2*	G:3 B:1	G:3 C:4	G:3 D:3	G:3 E:4	G:3 F:2	G:3 G:3	G:3 H:3	G:3 I:4
H	**RMT**	0.8924	0.9269	1.0323	*1.0016*	1.1157	1.2082	0.7106	**1.0015**	1.1142
	Separation Pattern	H:2* A:2*	H:3 B:1	H:3 C:4	H:3 D:3	H:3 E:4	H:2* F:2	H:3 G:3	H:3 H:3	H:3 I:4
I	**RMT**	1.0128	-	0.9320	0.9009	**1.0030**	1.0218	0.6439	0.9057	1.0032
	Separation Pattern	I:4 A:3	I:- B:-	I:4 C:4	I:3* D:3	I:4 E:4	I:4 F:2	I:4 G:2*	I:3* H:3	I:4 I:4

Theoretically, for a sample with references from the same mAb, the RMT value received should be closest to 1. This was the case for all except for the sample and references from mAb I, 1.0032 (1.0000 ± 0.0032). Instead, the RMT value closest to 1 was for sample I with references from E, 1.0030 (1.0000 ± 0.0030). All values that are within the range 1.000 ± 0.005 (p = 95%) are classified as false positives and marked in *table 10*.

According to the number of peaks, the fewer peaks in each separation pattern, the more overlap has occurred in the spectrum. In *table* 10, the number of peaks in the separation pattern for each mAb A, B, C, D, E, F, G, H and I is 3, 1, 4, 3, 4, 2, 3, 3 and 4 respectively. For mAb B, only one peak is visible in the electropherogram most likely due to its different structure. For the biosimilars E and I, there was no overlap as expected. The sample that had the smallest number of overlaps with its references was G. This indicates that G has most similarities to the other mAbs. A, E and I had most overlaps, which indicates that these should be easier to identify.

5. Discussion

Discussion of the results from all analyses and conclusion about the most optimal method for identification of mAbs.

5.1 MALDI-ToF-MS

Discussion of the results from MALDI analyses of reduced mAbs, ISD and deglycosylation. General advantages with MALDI are that both sample preparation and execution of the analysis are easy to implement. The sample preparation also requires fewer reagents than other analytical methods, which can be found to be environmentally friendly. A disadvantage is that the ionization of some samples is bad, which results in difficulties when interpreting the data obtained.

5.1.1 MALDI analysis of reduced mAbs

The results indicated good possibilities to use the method for identification of the mAb, but in a wider perspective it must work for all mAbs on the market. In those cases, where many more mAbs are included, the probability that the molecular masses of the chains are much more similar between the mAbs and the risk for false positive results increases. The number of biosimilars is also increasing, which implies that the differences between the mAbs are reduced, and the possibility of identification is further reduced.

Deficient measurement reliability and calibration may have affected the results. The method could still possibly be used if the amount of data collected is larger, both for each sample but also from several identical samples. Larger data collection would reduce the errors occurring when marking the peaks areas in the spectrum. However, this would require more resources and time, which is desired to be kept as low as possible. Other alternative analyses should therefore be reviewed, but this method could perhaps be a complement, due to its advantageous properties that the data obtained is clear and easy to handle.

An alternative to MALDI analyses on reduced form can be to analyse the intact form of the mAbs. For this purpose, however, it would not generate any better results, partly because larger molecules (higher m/z) are gives lower precision and specificity in analyses. Another reason is that separation between the mAbs can occur for only one of the chains, which is not visible here.

5.1.2 MALDI-In Source Decay (ISD)

There are many improvement measures that can possibly generate even better results. In this case, only two different concentrations and only two different centrifuge filters were tested for one of the nine mAbs when preparing the protocol for the analyses. Both more concentrations and filters can be tested to generate a protocol. Also, different matrices can be used to determine the most appropriate one. Possibly, also the digestion of the mAb and the buffer

can be changed to obtain results of higher average quality. By adapting various combinations of these factors to the different mAbs, better results could possibly be obtained.

This method requires much time as it involves manual annotation of almost each signal and each amino acid. The possible improvement measures for the development of an even better adapted method would also require a lot of time. These are disadvantages of the method, due to the overarching goal, fast and accurate identification of mAbs.

5.1.3 Deglycosylation

The analysis of the model protein used, ribonuclease, provided good results. Unfortunately, the method did not work for the mAbs. The reason for that is most likely due to sample preparation. With another sample preparation that fulfills deglycosylation of the mAb, the result would look like the spectra for ribonuclease. After a successful deglycosylation, the mass difference can be calculated and the composition and heterogeneity of the glycans can be determined. That would further the identification process.

5.2 Capillary Electrophoresis (CE)

Discussion of the results from CGE analyses of mAbs in reduced and non-reduced form and CZE analyses of intact mAbs.

5.2.1 Capillary Gel Electrophoresis (CGE)

For the mAbs analysed in reduced form, separation occurred for most of the combinations, at least for one of the chains.

Analysis of the combination of two different mAbs was done two times with different results (*table 8* and *table 9*) The only difference between these analyses, except that B was excluded, was the method for sample preparation. Comparing the results of the two analyses (*table 8* and *table 9*), the separation is different for one or two chains for ten of the combinations. In total, for 13 chains of in total 42, LC or HC had differing separation. For the biosimilars E and I analysed together, no separation occurred in either of the cases in *table 8* and *9*, which was expected. When comparing the biosimilars in samples with the other mAbs, the results should be the same due to the similarity between E and I. This was consistent in most of the cases but for E and I combined with C and G, the results differed between the biosimilars in both *table 8* and *table 9*.

When comparing the results with E and I between the two *tables 8* and *9*, the samples with C in *table 9* are confirmed with an additional analysis and therefore can be seen as more

reliable. The samples with the biosimilars and G are more alike between the two results, TS and NS separates the results of LC in both cases. This similarity between TS and NS cannot strengthen which result is more reliable. Apart from the biosimilars, BS occurred for HC in H+C, H+D, H+G only in *table 9* but for D+G only in *table 8*. In all these cases the corresponding results were NS, which makes the difference between the results huge. The most likely reason for this is the different sample preparation, as that is the biggest known difference between the two analyses. The conclusion can be drawn that the sample preparation is of great importance for which result is obtained. Based on the results of analyses with the biosimilars and C, the sample preparation when the mAbs are mixed before reduction with 2-ME and addition of SDS-MW Sample Buffer are generating the most reliable results.

The analysis of the combinations of the mAbs in intact form was interpreted based on the results from the analysis of the mAbs one and one in reduced form. This is because the affiliation of the peaks could then be more easily determined. The results showed PS for E+H, F+H, and H+I and NS/TS for C+G and D+G. This shows that the difference between some of the antibodies that showed separation in reduced form is also visible and can be separated in intact form. This is demonstrably true for far from all mAbs, so therefore the mAbs should be analysed in reduced form for the most reliable result in CGE analysis.

5.2.2 Capillary Zone Electrophoresis (CZE)

The results from the CZE are presented with calculated RMT values for each combination of sample and reference and the number of peaks in each separation pattern. In those cases when the affiliation of the peaks was difficult to determine, the ratios between the areas of two of the peaks within the same separation patterns could be calculated and compared. This increases the reliability of the interpretation of the results.

Theoretically, the RMT value for a sample and references from the same mAb should be closest to 1. This was the case for all except for one case, the sample, and references from mAb I, Instead, the RMT value closest to 1 was for sample I with references from mAb E. Since these mAbs are biosimilars and the difference between the RMT values is very small, the result is still expected. Further, the number of peaks in each separation pattern and the look of it can indicate the similarity between the mAbs. The fewer peaks, the more overlap has occurred in the spectrum. Overlap occurs when the respective migration time of sample and reference are not the same and they are shifted unequally relative to each other in the spectrum. In other words, a lower number of peaks in a separation pattern than usual indicates a difference between sample and reference mAbs. In a spectrum where the sample and reference are from the same mAb, the migration time is equal in relation to each other. There the separation patterns for the respective analyte appear at regular intervals and without overlap in the spectrum.

CZE compared to CGE as analysis methods, generates more data, both graphic and numerical, for each sample to easily compare. In other words, CZE is preferable over CGE for this aim.

5.2.3 Conclusion

Analysis of **mAbs in reduced form with MALDI** generates mostly reliable and clear results. The method is simple and fast since many samples can be analysed at the same time. Due to the risk for false positive results, the method is to recommend for definition of a first hypothesis when identifying unknown mAbs.

MALDI-ISD can be concluded to be too time and resource demanding due the needs to both compare sequences before analysis and find signature amino acids. Because the annotation of the amino acids also must be done manually, the method can be considered to work well for additional confirmation after the final result of mAb identification is produced.

CGE works well if the protocol regarding the sample preparation is suitable, but visual interpretation of the results is not reliable enough in cases where the separation is difficult to classify. **CZE** involves a simpler and less risky sample preparation than CGE and generates a lot of data, both numerical and graphical. These properties make CZE preferable over CGE for the aim.

Lastly, a protocol for identification of unknown mAbs can be designed (*figure 21*). The most optimal identification process begins with MALDI analysis of the unknown mAb in reduced form with references from several known mAbs. Based on the results obtained, a hypothesis can be formulated about which mAbs are most similar to the unknown mAb (false positive results, (RMM 1.0000 ± 0.0004, p = 95%). A final analysis of the unknown mAb compared to the selected candidates similar to the mAb as references is done with TICZE to eliminate the false positive results (RMT 1.000 ± 0.005, p = 95%). The identity of the unknown mAb must then be ascertained. If false positive results remain, MALDI-ISD is performed as a final selection step. The result can be confirmed with PMF, but that is not included as part of the method due to time-consuming.

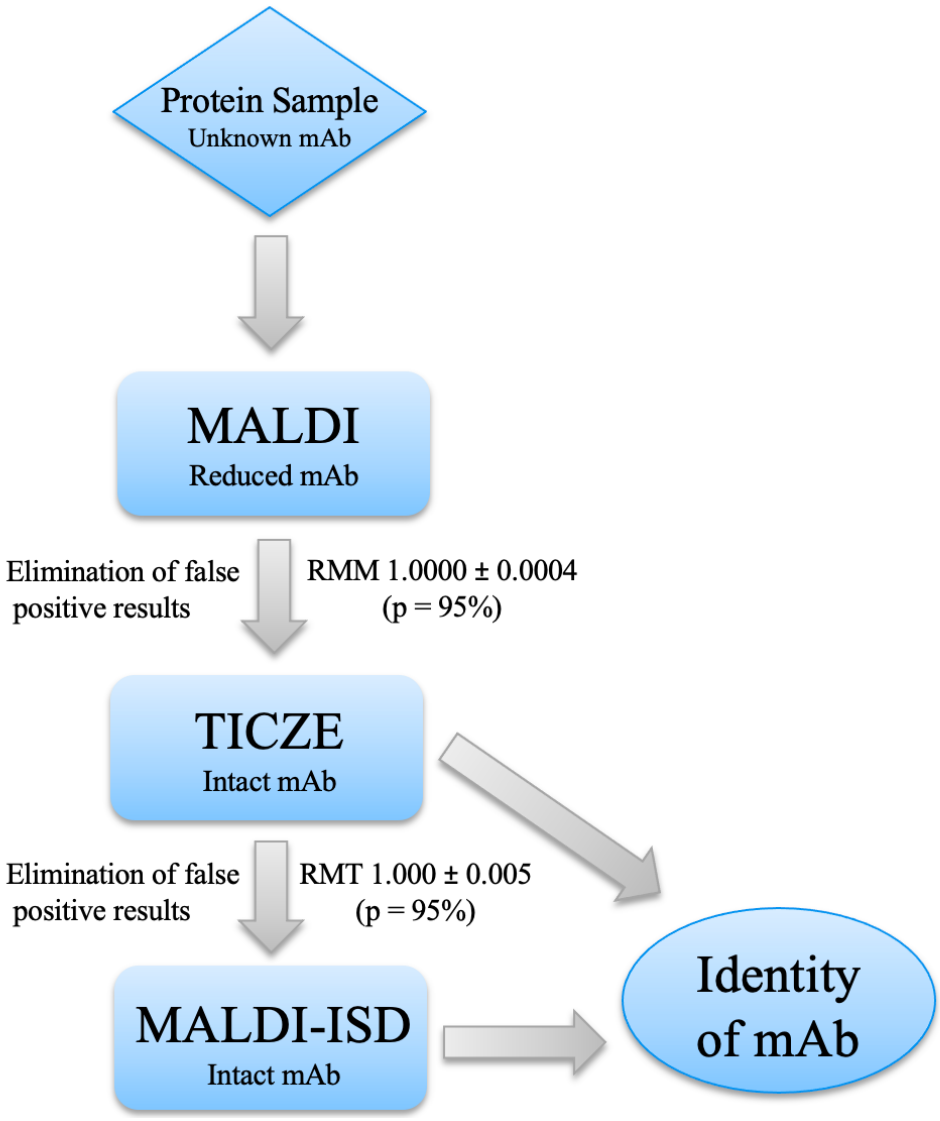

Figure 21: Flowchart describing the process of the final protocol for identification of mAbs.

6. Future Perspectives

The rapidly increasing number of mABs in medicine argues for that analysis for accurate and rapid identification of these is important to avoid counterfeiting. According to the MALDI analysis of reduced mAbs, a more extensive data collection should be invested in for faster identification. A future goal might be to create an easy-to-use database containing physiochemical data for all known mAbs. ISD is a good analytical tool but could be improved by developing a protocol for the most appropriate sample preparation. The same applies for deglycosylation. CZE is preferable over CGE, due to the results possible form each. On one hand, perhaps multiple injection CZE (MICZE) can generate more interesting and useful data but on the other hand, TICZE performs well and includes consumption of a little less resources.

7. Acknowledgements

I would like to express my deepest gratitude to my supervisors Ahmad Amini and Johan Carlsson at The Swedish Medical Products Agency, for the support through constant availability and guidance. With their long experience, specialist skills and extensive technical knowledge, they have guided me throughout my thesis project and enabled this work. I am extremely thankful for the opportunity to experience to carry out my project at The Swedish Medical Products Agency. I must also thank Maria Löfgren for great amount of assistance, caring support, and fun collaboration in the laboratory. Many thanks to all on the laboratory unit, for a warm welcome, generous introductory program and encouragement.

I would also like to extend my sincere thanks to my main supervisor John Löfblom at KTH, for the commitment and feedback throughout my project.

References

[1] Wageningen, University & Research. (Monoclonal) Antibodies. Retrieved 8 February 2023.

[Available at https://www.wur.nl/en/product/monoclonal-antibodies.htm]

[2] Läkemedelsverket. Om Läkemedelsverket. Retrieved 8 February 2023.

[Available at https://www.lakemedelsverket.se/sv/om-lakemedelsverket]

[3] Läkemedelsverket. Så godkänns ett läkemedel. Retrieved 15 February 2023.

[Available at https://www.lakemedelsverket.se/sv/tillstand-godkannande-och-kontroll/forsaljningstillstand/sa-godkanns-ett-lakemedel#hmainbody1]

[4] World Health Organization. Substandard and falsified medical products. Retrieved 15 February 2023.

[Available at https://www.who.int/health-topics/substandard-and-falsified-medical-products#tab=tab_1]

[5] Science Direct. Monoclonal Antibody. Retrieved 8 February 2023.

[Available at https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/monoclonal-antibody]

[6] Zheng K, Bantog C, Bayer R. The impact of glycosylation on monoclonal antibody conformation and stability. MAbs. 2011 Nov-Dec;3(6):568-76. doi: 10.4161/mabs.3.6.17922. Epub 2011 Nov 1. PMID: 22123061; PMCID: PMC3242843. Retrieved 21 February 2023.

[Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3242843/]

[7] O'Kennedy, Richard & Murphy, Caroline & Lysakova-Devine, Tatyana. (2016). Technology advancements in antibody purification. Antibody Technology Journal. Volume 6. 17-32. 10.2147/ANTI.S64762. Retrieved 25 April 2023.

[Available at https://www.researchgate.net/publication/307180889_Technology_advancements_in_antibody_purification]

[8] Absolute Antibody. Antibody structure. Retrieved 25 April 2023.

[Available at https://absoluteantibody.com/antibody-resources/antibody-overview/antibody-structure/]

[9] Läkemedelsverket. Biosimilarer. Retrieved 8 February 2023.

[Available at https://www.lakemedelsverket.se/sv/tillstand-godkannande-och-kontroll/tillverkningstillstand/biologiska-lakemedel/biosimilarer]

[10] European Medicines Agency. Biosimilar medicines: Overview. Retrieved 9 February 2023.

[Available at https://www.ema.europa.eu/en/human-regulatory/overview/biosimilar-medicines-overview]

[11] Dadouch M, Ladner Y, Perrin C. Analysis of Monoclonal Antibodies by Capillary Electrophoresis: Sample Preparation, Separation, and Detection. *Separations*. 2021; 8(1):4. https://doi.org/10.3390/separations8010004. Retrieved 30 Mars 2023.

[Available at https://www.mdpi.com/2297-8739/8/1/4]

[12] Geyer H, Geyer R. Strategies for analysis of glycoprotein glycosylation, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, Volume 1764, Issue 12, 2006, Pages 1853-1869, ISSN 1570-9639, https://doi.org/10.1016/j.bbapap.2006.10.007. Retrieved 27 Mars 2023.

[Available at https://www.sciencedirect.com/science/article/pii/S1570963906003487]

[13] Shimadzu. MALDI Basics. Retrieved 27 February 2023.

[Available at https://www.ssi.shimadzu.com/service-support/faq/maldi/index.html]

[14] Iowa State Univeristy. Common Matrices in MALDI-ToF. Retrieved 27 February 2023.

[Available at https://www.protein.iastate.edu/MALDI_matrix_information.html]

[15] André S. Landin H. Nilsson-Öst M. Mastitdiagnostik – när, var, hur? Retrieved 10 May 25, 2023.

[Available at http://www.juverportalen.se/media/1176/mastitdiagnostik-naer-var-hur.pdf]

[16] Creative Proteomics. The Principle, Characteristics and Application of Peptide Mass Fingerprinting. Retrieved 12 April 2023.

[Available at https://www.creative-proteomics.com/resource/the-principle-characteristics-and-application-of-peptide-mass-fingerprinting.htm]

[17] Suckau, D., Resemann, A., Schuerenberg, M. *et al.* A novel MALDI LIFT-TOF/TOF

mass spectrometer for proteomics. *Anal Bioanal Chem* **376**, 952–965 (2003). https://doi.org/10.1007/s00216-003-2057-0. Retrieved 27 Mars 2023.

[Available at https://www.sciencedirect.com/science/article/pii/S1044030503008353]

[18] Mitsuo T. In-source decay characteristics of peptides in matrix-assisted laser

desorption/ionization time-of-flight mass spectrometry. Journal of the American Society for Mass Spectrometry,Volume 12, Issue 4, 2001, Pages 420-427, ISSN 1044-0305, https://doi.org/10.1016/S1044-0305(01)00218-5. Retrieved 27 Mars 2023.

[Available at https://www.sciencedirect.com/science/article/pii/S1044030501002185]

[19] LibreTexts Chemistry. Capillary Electrophoresis. Retrieved 10 May 2023.

[Available at https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Supplemental_Modules_(Analytical_Chemistry)/Instrumentation_and_Analysis/Capillary_Electrophoresis]

[20] Bioptic. Introduction to CE. Retrieved 10 May 2023.

[Available at https://www.bioptic.com.tw/support/introduction-to-ce]

[21] Kok W. The Backgound Electrolyte. Retrieved 19 May 2023.

[Available at https://link.springer.com/chapter/10.1007/978-3-322-83133-0_7]

[22] Wikipedia. Capillary Electrophoresis. Retrieved 10 May 2023.

[Available at https://en.wikipedia.org/wiki/Capillary_electrophoresis#/media/File:Capillarywall.png]

[23] Kemp, G. Capillary Electrophoresis. Biotechnology and Applied Biochemistry, 27: 9-17.

https://doi.org/10.1111/j.1470-8744.1998.tb01369.x Retrieved 10 May 2023.

[Available at https://doi.org/10.1111/j.1470-8744.1998.tb01369.x]

[24] ScienceDirect. Capillary Gel Electrophoresis Retrieved 22 May 2023.

[Available at https://www.sciencedirect.com/topics/chemistry/capillary-gel-electrophoresisCGE]

[25] ScienceDirect. Capillary Zone Electrophoresis Retrieved 19 May 2023.

[Available at https://www.sciencedirect.com/topics/chemistry/capillary-zone-electrophoresis]

[26] MERCK. Hexadimethrine bromide. Retrieved 8 May 2023.

[Available at https://www.sigmaaldrich.com/SE/en/product/sigma/h9268]

[27] PubChem. Aminocaproic acid. Retrieved 8 May 2023.

[Available at Aminocaproic acid | C6H13NO2 - PubChem (nih.gov)

[28] MERCK. (Hydroxypropyl)methyl cellulose. Retrieved 8 May 2023.

[Available at (Hydroxypropyl)methyl cellulose - (sigmaaldrich.com)

[29] PubChem. Lithium acetate dihydrate. Retrieved 8 May 2023.

[Available at Lithium acetate dihydrate | C2H7LiO4 - PubChem (nih.gov)

[30] Harvey, David. (2003). Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Carbohydrates and Glycoconjugates. International Journal of Mass Spectrometry. 226. 1-35. 10.1016/S1387-3806(02)00968-5. Retrieved 21 Mars 2023.

[Available at (PDF) Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Carbohydrates and Glycoconjugates (researchgate.net)]