

**THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)**

**NEW ASPECTS IN ANALYSIS OF N-GLYCANS USING  
CAPILLARY GEL ELECTROPHORESIS**

by Márta Zsuzsa Kerékgyártó

UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF MOLECULAR MEDICINE

DEBRECEN, 2017

**THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PhD)**

**NEW ASPECTS IN ANALYSIS OF N-GLYCANS USING  
CAPILLARY GEL ELECTROPHORESIS**

by Márta Zsuzsa Kerékgyártó

Supervisor: Prof. András Guttman



UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF MOLECULAR MEDICINE

DEBRECEN, 2017

# TABLE OF CONTENTS

<b>1. INTRODUCTION.....</b>	<b>4</b>
<b>2. THEORETICAL BACKGROUND OF CGE.....</b>	<b>5</b>
2.1. INSTRUMENTATION.....	5
2.2. BASIC THEORY OF CAPILLARY GEL ELECTROPHORESIS.....	6
2.2.1. ELECTROPHORETIC MOBILITY .....	6
2.2.2. COLUMN EFFICIENCY AND RESOLUTION .....	9
2.3. MOST POPULAR GEL FORMULATIONS IN CGE .....	10
2.3.1. ACRYLAMIDE-BASED SIEVING POLYMERS.....	11
2.3.2. AGAROSE AS SIEVING MEDIUM.....	12
2.3.3. POLYETHYLENE OXIDE, POLYVINYLPYRROLIDONE, AND OTHER SIEVING MATRICES .....	12
2.4. CAPILLARY COATINGS .....	13
2.5. DETECTION OPTIONS FOR CGE .....	14
2.5.1. UV-VISIBLE DETECTOR.....	14
2.5.2. DIODE ARRAY DETECTOR.....	15
2.5.3. LASER-INDUCED FLUORESCENCE DETECTOR .....	15
2.5.4. LIGHT-EMITTING DIODE-INDUCED FLUORESCENCE DETECTION .....	17
2.6. NOVEL CGE INSTRUMENT DESIGNS.....	18
2.7. CAPILLARY GEL ELECTROPHORESIS OF NUCLEIC ACIDS.....	21
2.8. PROTEIN SEPARATION BY CAPILLARY GEL ELECTROPHORESIS.....	23
2.9. CAPILLARY GEL ELECTROPHORESIS OF COMPLEX CARBOHYDRATES .....	24
2.9.1. THEORY OF GLYCAN STRUCTURE ELUCIDATION BY CGE.....	26
2.10. FUTURE TRENDS: MINIATURIZATION, LAB-ON-A-CHIP SYSTEMS .....	28
<b>3. AIMS AND MOTIVATION .....</b>	<b>31</b>
<b>4. MATERIALS AND METHODS .....</b>	<b>33</b>
4.1. LARGE-SCALE GENOMIC DNA ANALYSIS .....	33
4.1.1. ANALYSIS OF gDNA BY CGE.....	33
4.1.2. ANALYSIS OF gDNA BY AGAROSE GEL ELECTROPHORESIS .....	34
4.2. MOLECULAR HAPLOTYPE ANALYSIS .....	34
4.2.1. NON-INVASIVE DNA SAMPLING AND DNA EXTRACTION.....	34
4.2.2. ALLELE-SPECIFIC PCR .....	34
4.2.3. ANALYSIS OF PCR-PRODUCT BY CGE.....	35
4.2.4. ANALYSIS OF PCR-PRODUCT BY AGAROSE GEL ELECTROPHORESIS .....	36
4.3. IMMUNE RESPONSE AGAINST NEOGLYCOPROTEINS .....	36
4.3.1. GENERAL PROCEDURES .....	36
4.3.2. SYNTHESIS OF NEOGLYCOPROTEINS .....	37
4.3.3. ANALYSIS OF NEOGLYCOPROTEINS BY CGE.....	40
4.3.4. ANALYSIS OF NEOGLYCOPROTEINS BY MALDI-TOF MS.....	41
4.3.5. IMMUNIZATION WITH NEOGLYCOPROTEINS.....	41
4.3.6. ANALYSIS OF NEOGLYCOPROTEINS BY ELISA.....	42

<b>4.4. HIGH-PERFORMANCE CGE ANALYSIS OF COMPLEX CARBOHYDRATES .....</b>	<b>43</b>
4.4.1. SAMPLE PREPARATION FOR CGE .....	43
4.4.2. N-GLYCAN SEPARATION BY CGE .....	44
<b>4.5. STUDY OF THE ACTIVATION ENERGY BY CGE .....</b>	<b>44</b>
4.5.1. SAMPLE PREPARATION FOR CGE .....	44
4.5.2. CGE-BASED SEPARATION OF LINEAR AND BRANCHED CARBOHYDRATES.....	45
<b><u>5. RESULTS .....</u></b>	<b><u>46</u></b>
<b>5.1. LARGE-SCALE GDNA ANALYSIS BY CGE .....</b>	<b>46</b>
<b>5.2. MOLECULAR HAPLOTYPE ANALYSIS .....</b>	<b>49</b>
5.2.1. ALLELE-SPECIFIC PCR .....	49
5.2.2. HAPLOTYPE DETERMINATION BY CGE .....	53
5.2.3. LIMIT OF DETECTION AND DETECTOR LINEARITY STUDY FOR CGE.....	55
<b>5.3. IMMUNE RESPONSE AGAINST CARBOHYDRATE ANTIGENS.....</b>	<b>57</b>
5.3.1. SYNTHESIS OF CARBOHYDRATE ANTIGENS .....	57
5.3.2. ANALYSIS OF NEOGLYCOPROTEINS BY CGE AND MALDI-TOF MS .....	59
5.3.3. POLYCLONAL ANTIBODY RESPONSE TO NEOGLYCOPROTEINS MEDIATED IMMUNIZATION .....	60
5.3.4. INHIBITION OF MOUSE ANTIBODY BINDING IN ELISA .....	62
<b>5.4. GENERATION OF AN ANTS-LABELED N-GLYCAN DATABASE FOR CGE ANALYSIS OF CARBOHYDRATES .....</b>	<b>64</b>
5.4.1. GENERATION OF AN ANTS-LABELED N-GLYCAN DATABASE .....	64
5.4.2. VALIDATION OF THE ANTS-LABELED N-GLYCAN DATABASE .....	67
<b>5.5. EFFECT OF SEPARATION TEMPERATURE AND BACKGROUND ELECTROLYTE COMPOSITION ON STRUCTURE SPECIFIC GLYCAN MIGRATION IN CGE.....</b>	<b>71</b>
5.5.1. ACTIVATION ENERGY ASSOCIATED WITH THE ELECTROMIGRATION OF LINEAR SUGAR OLIGOMERS.....	71
5.5.2. EFFECT OF VISCOSITY-MODIFIER (ETHYLENE GLYCOL, EG) .....	74
5.5.3. THE EFFECT OF POLYMERIC ADDITIVE (LINEAR POLYACRYLAMIDE, LPA).....	75
5.5.4. ACTIVATION ENERGY ASSOCIATED WITH THE ELECTROMIGRATION OF BRANCHED OLIGOSACCHARIDES .....	77
5.5.5. THE EFFECT OF BACKGROUND ELECTROLYTE ADDITIVES .....	81
<b><u>6. DISCUSSION .....</u></b>	<b><u>83</u></b>
<b><u>7. SUMMARY .....</u></b>	<b><u>87</u></b>
<b><u>9. BIBLIOGRAPHY .....</u></b>	<b><u>91</u></b>
<b><u>10. ABBREVIATIONS.....</u></b>	<b><u>101</u></b>
<b><u>11. KEYWORDS.....</u></b>	<b><u>104</u></b>
<b><u>12. ACKNOWLEDGEMENTS.....</u></b>	<b><u>105</u></b>

## 1. INTRODUCTION

Electrophoresis covers a number of bioseparation tools where charged materials are separated due to differences in migration speed under the influence of an electrical field. This separation and chemical analysis technique was born in 1937 with the work of a Swedish biochemist called Arne Tiselius [1], who won the Nobel Prize for Chemistry in 1948. Since the 1950s, slab gel electrophoresis has been routinely used by analysts to separate and identify large biopolymers such as nucleic acids, proteins and complex carbohydrates. Although it is one of the most widely used separation techniques in molecular biology and biochemistry labs, slab gel electrophoresis generally suffers from long analysis times, low efficiencies and difficulties in detection and automation.

Capillary gel electrophoresis (CGE), on the other hand, is fast becoming the separation and characterization technique of choice in the bioanalytical field, which combines the advantages of liquid chromatography and ‘conventional’ slab/rod gel electrophoresis. In addition, this technique is not only used for the separation of biopolymers, but also in other areas, where has not earlier used the electrophoresis (e.g., food analysis, forensics and environmental protection, etc.). In addition, CGE is a reliable, fully-automated approach, offering rapid separations with high sensitivity, excellent resolution, ruggedness and ease of operation for the analysis of biologically important macromolecules, which contributed to the earlier completion of the Human Genome Project (HUGO) than originally planned. Although capillary electrophoresis (CE) first emerged as a free solution method, sieving media for size selective separations were developed soon after (such as sieving matrices made of agarose, polyacrylamide, polyethylene oxide, polyvinylpyrrolidone, sodium dodecyl sulfate, and so on). In case of CGE, the narrow-bore capillaries are filled with cross-linked gels or linear polymer sieving matrices for the separation. In addition, CGE technique can be used for the investigation of small molecule interactions and enzyme kinetic reactions.

At present, existing CGE methods can be readily transferred from the capillary format to electrophoresis microchips, i.e. lab-on-a chip, which offers rapid and effective separations for the biologically important macromolecules.

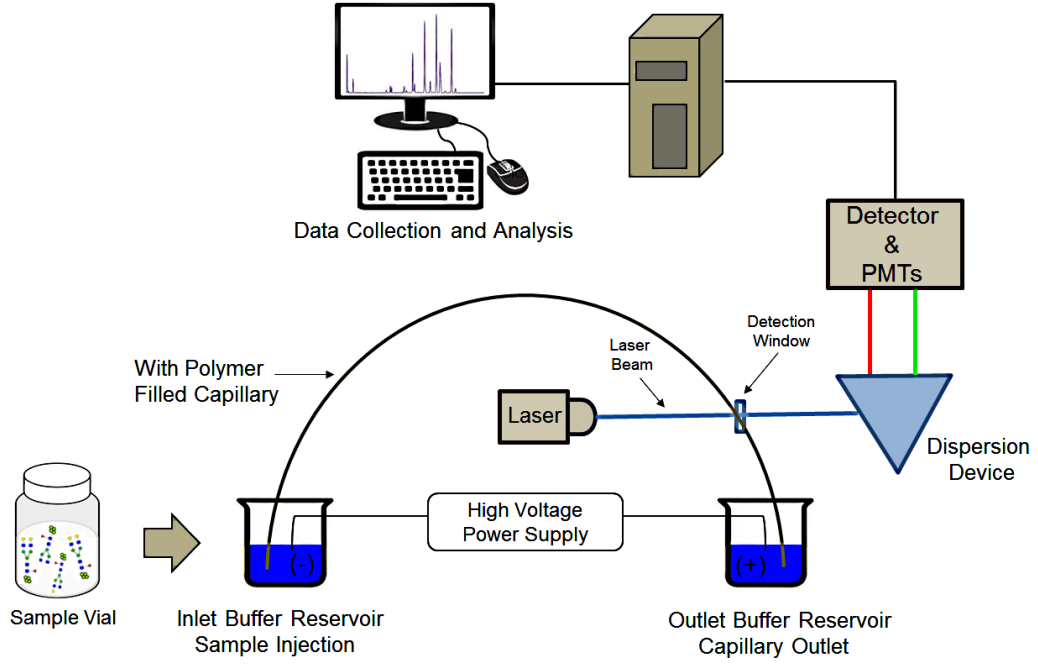
## **2. THEORETICAL BACKGROUND OF CGE**

### **2.1. Instrumentation**

A typical CGE instrument consists of a separation capillary, a high-voltage power supply, a sample introduction system, a detection assembly and data acquisition device as illustrated in Figure 1.

In case of CGE, the overall size of the separation capillary is in the range of 10-100 cm (usually 20-100  $\mu\text{m}$  i.d. with 150-360  $\mu\text{m}$  o.d.). In single- or multicapillary systems, the inside wall of the fused silica capillary can be coated (covalent or dynamic) to eliminate electroosmotic flow (EOF) and prevent possible adsorption of the analytes. For sample introduction, CGE instruments mainly apply electrokinetic method, because the higher viscosity of the gel-buffer usually does not support pressure injection mode. In the electrokinetic injection method, the charged analyte molecules are forced into the capillary tubing by an applied electric field, while in case of pressure injection, a sample vial is temporarily pressurized to allow the sample flowing into the separation capillary. Electrokinetic injection from aqueous samples (with no/or little salt) results in large sample intake, because the buffer co-ions do not compete with the analyte molecules resulting excellent limits of detection (LOD) [2]. It is also important that CGE systems have temperature control to ensure reproducible results.

According to CGE principle, during the injection process, both ends of the capillary and the electrodes are dipped into the respective sample and outlet buffer reservoirs, and the applied voltage or pressure drives the analyte molecules into the capillary tubing. During the separation process, the analyte molecules migrate within the separation gel-buffer system filled capillary tubing towards the detection area. Most of modern CGE systems allow separation voltages up to 30 kV, which is orders of magnitudes higher than in conventional slab gel electrophoresis. Detection can be accomplished by UV/UV-Vis light absorbance including scanning diode array, laser- (LIF) and light-emitting diode-induced fluorescence (LEDIF), or other special systems such as electrochemical detection (Figure 1). In optical detection systems, a photomultiplier tube (PMT) is placed close to the outlet end of the capillary. If mass spectrometry detection is used, the end of the capillary may act as a sprayer tip [3].



**Figure 1.** Schematic illustration of a CE-LIF instrument with reversed polarity (cathode at injection side).

## 2.2. Basic Theory of Capillary Gel Electrophoresis

### 2.2.1. Electrophoretic Mobility

In electric field-mediated separation methods, when a uniform electric field ( $E$ ) is applied to a solute molecule with a net charge of  $Q$ , an electrical force ( $F_e$ ) is defined as:

$$F_e = Q \cdot E \quad (\text{eq 1})$$

When using a gel-buffer matrix in the separation capillary, a frictional force ( $F_f$ ) acts in the opposite direction, where  $f$  is the translational friction coefficient,  $dx$  and  $dt$  are the distance and time increments as delineated in equation 2:

$$F_f = f \left( \frac{dx}{dt} \right) \quad (\text{eq 2})$$

The migration of the charged analytes under the applied electric field strength is expressed according to Newton's second law of motion as:

$$m \left( \frac{d^2 x}{dt^2} \right) = EQ - f \left( \frac{dx}{dt} \right) \quad (\text{eq 3})$$

In equation 3,  $m$  is the mass and  $d^2x/dt^2$  is the acceleration. When  $F_e$  and  $F_f$  are counterbalanced the analytes migrate with a steady-state velocity ( $v$ ):

$$v = \frac{dx}{dt} = E \cdot \frac{Q}{f} \quad (\text{eq 4})$$

The translational friction coefficient ( $f$ ) is proportional to the viscosity ( $\eta$ ) of the background electrolyte as delineated in equation 5 [4].

$$f = c \cdot \eta \quad (\text{eq 5})$$

where  $c$  is a proportional constant, which is influenced by molecular configuration, for example, for small spherical molecules, and  $c$  is defined as  $6\pi r$ , in which  $r$  is the radius of the analyte ion [5]. The electrophoretic mobility ( $\mu$ ) is defined as the velocity per unit field strength, which represents the field normalized velocity,  $v/E$ , can be expressed as:

$$\mu = \frac{v}{E} = \frac{Q}{c \cdot \eta} \quad (\text{eq 6})$$

When using a separation sieving matrix in the capillary tubing, the retardation of the analytes is defined by the sieving polymer concentration ( $P$ ) and its physical interaction with the analyte molecule is determined by the retardation coefficient ( $K_R$ ). Equation 7 represents the mobility of the solute as the function of the retardation coefficient at a given polymer concentration, where  $\mu_0$  is the free solution electrophoretic mobility of the analyte [6].

$$\mu = \mu_0 \cdot e^{-K_R \cdot P} \quad (\text{eq 7})$$

In this instance, the  $K_R$  values are the actual slopes of the Ferguson plots of logarithmic electrophoretic mobility vs. sieving polymer concentration:



$$\ln \mu = \ln \mu_0 - K_R P \quad (\text{eq 8})$$

In CGE, molecular sieving can be delineated by the Ogston theory [7] when the hydrodynamic radius of the migrating solute is in the same range as that of the average pore size of the sieving matrix. The retardation coefficient ( $K_R$ ) depends on the molecular weight (MW) of the migrating molecule at constant polymer concentration [8] and the mobility of the analyte molecule is a logarithmic function of the MW [6] as defined in equation 9.

$$\mu \sim \exp(-\text{MW}) \quad (\text{eq 9})$$

The Ogston theory also assumes that the migrating particles behave as spherical objects with similar size to the gel pores. On the other hand, large biopolymers containing flexible chains (e.g., DNA or SDS protein complexes) can migrate through the pores of the polymer sieving that has significantly smaller pores than the size of the solute [9]. These large biomolecules behave as “snake-like” objects when migrating through the much smaller gel pores [10-12]. This phenomenon is described by the reptation model, where the size of the migrating solute (e.g., the chain length,  $n$  in case of DNA molecules) is inversely proportional to the mobility as shown in equation 10.

$$\mu \approx \frac{E}{n} \quad (\text{eq 10})$$

In CGE, the biased reptation models describe electromigration at very high electric field strengths. In equation 11,  $b$  is a function of the size of the network polymer as well as the charge, and the segment length of the migrating biomolecules.

$$\mu \approx \frac{E}{n} + bE^2 \quad (\text{eq 11})$$

According to the Eyring-Polanyi equation [13], the viscosity of the background electrolyte can be expressed as:

$$\eta = A \cdot e^{E_a/RT} \quad (\text{eq 12})$$

where  $A$  is a pre-exponential factor,  $E_a$  is the activation energy of the viscous flow in the sieving matrix,  $R$  is the universal gas constant, and  $T$  is the absolute temperature. Please note that the electrophoretic mobility values have to be corrected with the temperature-influenced viscosity change of the solution (1.1% per °C) [14]. The electrophoretic mobility ( $\mu$ ) can also be expressed by combining equations 6 and 12 ( $const$  indicates a collection of constant including  $c$  and  $A$  from equations above).

$$\mu = \frac{Q}{const} \cdot e^{-E_a / RT} \quad (\text{eq 13})$$

In practice, the  $E_a$  values are usually derived from the slopes of the Arrhenius plots [15], when the logarithm of the electrophoretic mobility is plotted against the reciprocal absolute temperature according to equation 14:

$$\ln \mu = \ln \frac{Q}{const} - \frac{E_a}{RT} \quad (\text{eq 14})$$

In CGE, different electrophoretic mobilities are due to differences in sample properties such as shape, size, or net charge, all of which affect the electromigration. When viscosity modifiers and/or polymeric additives are added to the background electrolyte, the activation energy can be considered as the energy required by the migrating analyte ion to overcome the barriers created by the separation gel-buffer. Retardation of the migrating ions, on the other hand, depends on the viscosity modifier/polymeric additive of the background electrolyte, possible interactions between the solute and additive (earlier described as retardation coefficient,  $K_R$ ) as well as the concentration of viscosity and/or polymeric additive ( $P$ ) as characterized by equation 7 [6].

### 2.2.2. Column Efficiency and Resolution

In CGE, the theoretical plate number ( $N$ ) can be characterized by equation 15, where  $\mu$  is the electrophoretic mobility of the analyte,  $l$  is the effective length of the capillary tubing, and  $D$  is the diffusion coefficient of the migrating analytes.

$$N = \mu \frac{E \cdot l}{2D} \quad (\text{eq 15})$$

Resolution ( $R_s$ ) in CGE can be expressed by equation 16, calculated from the differences of the electrophoretic mobilities ( $\Delta\mu$ ) between the two electrophoretic peaks [16] and the electrophoretic mobility of the analyte molecule of interest ( $\mu_m$ ).

$$R_s = 0.18 \cdot \Delta\mu \sqrt{\frac{E \cdot l}{D \cdot \mu_m}} \quad (\text{eq 16})$$

Equations 15 and 16 show that higher efficiency ( $N$ ) and higher resolution ( $R_s$ ) would be resulted by lower solute diffusion coefficient and higher applied electric field. According to equation 17, the limiting factor in the achievement of the high resolution is mainly the so-called Joule heat ( $Q_j$ ) generated by the applied power ( $P=V \times I$ ) [17], where  $r$  means the internal radius of the column,  $L$  is the total length of the capillary tubing, and  $I$  is the current.

$$Q_j = \frac{P}{r^2 \cdot I \cdot L} \quad (\text{eq 17})$$

To achieve adequate migration reproducibility and decrease Joule heat change-related problems in CGE, state-of-the-art capillary electrophoresis units utilize temperature control systems (i.e. liquid or air cooling) during the separation process.

### 2.3. Most Popular Gel Formulations in CGE

In the early 1990s, high resolving power was achieved with the use of narrow-bore-fused silica tubings filled with various gel formulations for the analysis of the biopolymers such as nucleic acids, proteins and complex carbohydrates. Two types of polymers were utilized in CGE: cross-linked and non-cross-linked. Today, the most frequently used gels are linear non-cross-linked polymers such as linear polyacrylamide (LPA), polyethylene oxide (PEO), polydimethylacrylamide (PDMA), polydiethylacrylamide (PDEA), polyvinylpyrrolidone (PVP) and cellulose derivatives like hydroxypropyl methylcellulose (HPMC) [18]. The most important properties of these polymer matrices are discussed in the following section.

### **2.3.1. Acrylamide-Based Sieving Polymers**

In CGE, polyacrylamide (PA)-based sieving matrix can be utilized as linear and/or cross-linked polyacrylamide for the analysis of biologically important macromolecules, mostly DNA. Cross-linked PA gels (also known as chemical gels), have been useful for the separation of shorter single-stranded DNA (ssDNA) and small proteins by SDS-based CGE [19]. These gels are covalently attached to the inside wall of the fused silica capillary tubing to improve their stability. In that case, samples can only be introduced into cross-linked gel-filled capillaries by electrokinetic injection, which results in sharp peaks with some sample pre-concentration at the interface of the gel and the sample buffer. However, this method leads to biased injection as smaller or highly charged molecules preferably enter the capillary tubing. In addition, higher separation temperatures may lead to deterioration of the gel structure (shrinking or bubble formation), if cross-linked PA-gels are used as sieving matrices during the separation process. On the other hand, chemical gels offer high resolving power, especially for the analysis of low-molecular-weight oligomers.

Non-cross-linked gels (so-called physical gels or low-viscosity LPA solutions) are widely used in CGE. These linear polymer gels are not attached to the inner surface of the capillary, thus sample injection are possible both by electrokinetic and pressure modes. In addition, physical gels are not sensitive to high temperature, pH or salt concentration changes during the separation process. Linear polymer matrices also permit simple replacement of the separation matrix in the capillary by positive or negative pressure. Contrary to cross-linked PA, linear polyacrylamide does not become cross-linked during the gelation process, so the inner surface of the capillary must be coated with a non-charged material to minimize the effect of EOF. Such polymer matrices have been useful in coated capillaries for the separation of double-stranded DNA (dsDNA), SDS–protein complexes, and complex carbohydrates [20]. In addition, the use of these non-cross-linked gels in DNA sequencing was a very important step toward the introduction of large-scale DNA sequencing by CGE, which was capable to sequence the Human Genome within a reasonable time frame [21]. Table 1 summarizes the main differences between cross-linked and non-cross-linked polyacrylamide sieving matrices.

**Table 1.** Features of polyacrylamide gels used in CGE. Adapted from [22].

	<i>Cross-linked polyacrylamide gel (Chemical gel)</i>	<i>Non-cross-linked polyacrylamide gel (Physical gel)</i>
<b>Attachment</b>	linked to the inside wall of the capillary	not linked to the inside wall of the capillary
<b>Pore Structure</b>	well-defined, rigid	flexible, dynamic, polymer networks with linear/branched structures
<b>Pore Size</b>	cannot be varied after polymerization	Variable
<b>Viscosity</b>	high-viscosity	Low to high viscosity
<b>Heat Tolerance</b>	sensitive	insensitive
<b>Sample Injection</b>	electrokinetic	electrokinetic and pressure
<b>Application</b>	ssDNA fragments, small proteins	dsDNA fragments, large proteins, complex carbohydrates

### 2.3.2. Agarose as Sieving Medium

Contrary to polyacrylamide, agarose has a larger pore size and usually applied for the separation of relatively large dsDNA fragments [23]. CGE compared to manual slab gel electrophoresis provides better resolution, especially in the <600 base pair range during the separation process. In contrast with PA, agarose does not become cross-linked during the gelation process, therefore the inside wall of the capillary must be coated with an adequate non-charged material, e.g., LPA, to eliminate the effect of EOF. Agarose-gel-filled capillaries were used primarily for the separation of larger dsDNA fragments, but the effective size range was limited up to ~12kb [19]. Unlike polyacrylamide, agarose-gel-filled capillaries has not been widely used in CGE, therefore prepacked capillaries or gel kits with agarose are not available commercially [19].

### 2.3.3. Polyethylene Oxide, Polyvinylpyrrolidone, and Other Sieving Matrices

Although the most commonly used sieving matrices are acrylamide-based polymers in CGE, PEO and PVP are also preferred sieving polymers, offering good separation for biologically important biomolecules [18]. In the 1990s, low-viscosity PEO-based separation polymer solutions were introduced in capillary array electrophoresis for non-denaturing DNA fragment analysis [24].

Hu *et al.* utilized PEO to study the protein contents from HT29 human colon adenocarcinoma cells by SDS-CE [25]. Barta *et al.* analyzed 21-hydroxylase deficiency with a primer extension technique using a PVP sieving matrix and detected the most common mutations in the gene [26]. SDS-protein complexes were separated with PEG and dextran sieving matrices by Ganzler *et al.* [27, 28], utilizing one of the advantages of polysaccharides namely their low UV absorption. Such capillary gel separation kits are commercially available for CGE from companies such as, SCIEX ([www.sciex.com](http://www.sciex.com)), BiOptic Inc. ([www.biopic.com.tw](http://www.biopic.com.tw)), Qiagen ([www.qiagen.com](http://www.qiagen.com)) and Agilent Technologies ([www.agilent.com](http://www.agilent.com)). SDS-MW Gel Buffer Kit (SCIEX) has been extensively applied for quantitative and qualitative analysis of membrane proteins [29], protein biotoxins [30], and antibodies [31]. Another gel-buffer system from the same company has been routinely used for standard *N*-linked glycan analysis.

## 2.4. Capillary Coatings

In the early 1980s, special capillary coating methods (dynamic or covalent) were introduced for CGE to eliminate/modify the effects of electroosmotic flow (EOF) and decrease possible wall absorption of the analyte molecules. Firstly, Hjerten *et al.* used a  $\gamma$ -methacryloxypropyl-trimethoxysilane bifunctional reagent in the capillary tubing, and then cross-linked the surface-bound methylacryl groups with acrylamide (without crosslinker) polymerization solution to suppress EOF [32]. One of the problems with this coating technique was that linear polymers could not completely cover the inside wall of the capillary [28] and thus was not stable for a longer time, particularly at higher pH values. In 2004, Gao *et al.* introduced a cross-linked PA coating technique, which was successfully used in SDS-CGE analysis. Bruin *et al.* attached the same bifunctional of  $\gamma$ -methacryloxypropyl-trimethoxysilane to covalently coat fused silica capillaries for the separation of small carbohydrate molecules [33], but the Si-O-Si bond was tended to hydrolyze at higher pH values. To address this problem, Cobb, Dolnik and Novotny developed a more stable direct Si-C bond-based capillary coating that improved stability over a wide separation pH range of 2-10 [34]. Covalently, coated capillaries reduced interactions between analytes and capillary and also provided high-speed separation without the requirement of pre-separation equilibration with the separation background electrolyte. From the 1990s, in addition to covalently coated capillaries, the noncovalent so-called „dynamic coatings” have rapidly emerged in CGE. Dynamic coatings can be easily prepared by simply rinse the capillary with polymer solution, detergent, or multivalent ions, but these coatings do not completely eliminate the effect of EOF.

Polymers used as a dynamic coating substances involve polydimethylacrylamide [35], poly-*N*-hydroxyethylacrylamide [36], and epoxy-poly(dimethylacrylamide) [37-39]. Further efforts in CGE were that the advantages of covalent and dynamic coatings were combined, which showed a stable performance for more than 200 sample injections (acid rinses between runs) [18, 40]. Today in CGE, both covalently and dynamic coating techniques are successfully utilized in conjugation with polymer solutions and gels in bioanalysis [41, 42]. To improve separation efficiency and reduce adsorption effects, similar coatings are also applied to microchip electrophoresis [43, 44].

## **2.5. Detection Options for CGE**

### **2.5.1. UV-Visible Detector**

UV-visible absorbance detectors are the most widely used in CE, included CGE. Two special types of light sources are available for UV-Vis, i.e., single line and continuous source. Atomic lamps are the simplest UV light source, which produce strong emission lines at well-defined wavelengths, e.g., the As lamp at 200 nm, the Zn lamp at 214 nm, the Cd lamp at 229 nm or the low-pressure Hg lamp at 254 nm. A medium-pressure Hg lamp is also used as a light source providing emission lines at 254, 280, 313, 365, 405, 436, and 546 nm wavelengths. In addition, two special types of light sources are commonly employed as continuous light in UV-Vis detection. The first source, the deuterium arc lamp, yields good continuous intensity in the UV region, ranging from 180 to 350 nm. However, noise from the lamp often caused problems in the overall detection performance of the instrument, so modern deuterium arc lamps have lower noise [45]. The intensity of such light sources is consistently decreasing with usage time, with typical half-lives of approximately 1,000h. The second continuous wavelength light source is the tungsten-halogen lamp that provides high intensity over the upper part of the UV spectrum and the full visible range from 280 to 1000 nm. In addition, this type of lamp has very low noise, short drift, and long life with half-lives of approximately 10,000h.

In this “forward optics” design, a polychromatic light from the source is focused onto the entrance slit of a monochromator, which selectively transmits a narrow band of light. After that, this light is transported through the capillary tubing to a photodiode detector. Contrary to other liquid-based separation methods such as LC, the sensitivity of the UV-VIS detectors is limited, because the path length is a direct function of the capillary diameter, that is, rather short.

Indirect UV detection can also be used, if the analyte components do not have UV absorbing properties. Indirect UV detection is universal and useful for the simultaneous detection of absorbing and non-absorbing samples. On the other hand, the disadvantages of these detectors are that they cannot be applied for additional identity confirmation and samples can exhibit positive and negative peaks [46].

### **2.5.2. Diode Array Detector**

Multi-wavelength detectors, including diode array detectors (DAD), use a “reverse optics” design. In 1989, the first diode array detector in CE was introduced for on-capillary photometric detection [47]. In this “reverse optics” design, a polychromatic light from the source mentioned above is first transported through the capillary tubing, and then focused onto the entrance slit of a monochromator. This monochromator consists of a diffraction grating, which transmits the light onto an array system consisting of numerous photodiodes that measures a narrow band of the spectrum. The bandwidth of the light detected by a diode is defined by the size of the diode and monochromator entrance slit. In this array detection, the number of the individual diodes determines the wavelength range and resolution of the detector. Using a diode-array detector, the maximum absorbance in DADs can be calculated by software, and the data can be visualized in a three-dimensional form. Thus, a whole wavelength range can be used, i.e., all light-absorbing components of the analyte can be detected within a single analysis providing UV-Vis spectra at any time points of the separation, thus DADs offer spectral analysis of each sample component along with qualitative information [46].

### **2.5.3. Laser-Induced Fluorescence Detector**

The laser-induced fluorescence (LIF) detector is one of the most sensitive on-capillary detection designs for CE, even capable to detecting a single-molecule, but most commonly requires the use of chemical derivatizations. Various charged and uncharged fluorescent dyes are commercially available for protein analysis such as fluorescein isothiocyanate (FITC) [48], 3-(2-furoyl)quinoline-2-carboxaldehyde (FQ) [25], and naphthalene-2,3-dicarboxaldehyde (NDA) [49]. For the analysis of carbohydrate, mostly 8-aminopyrene-1,3,6-trisulfonic acid (APTS) [50, 51], 8-aminonaphthalene-1,3,6-trisulfonic acid (ANTS) [52], 5-amino-2-naphthalenesulfonic acid (5-ANSA), 7 aminonaphthalene-1,3-disulfonic acid (ANDSA), and 2-aminonaphthalene-1-sulfonic acid (2-ANSA) [53] are used. In addition, labeling of DNA with

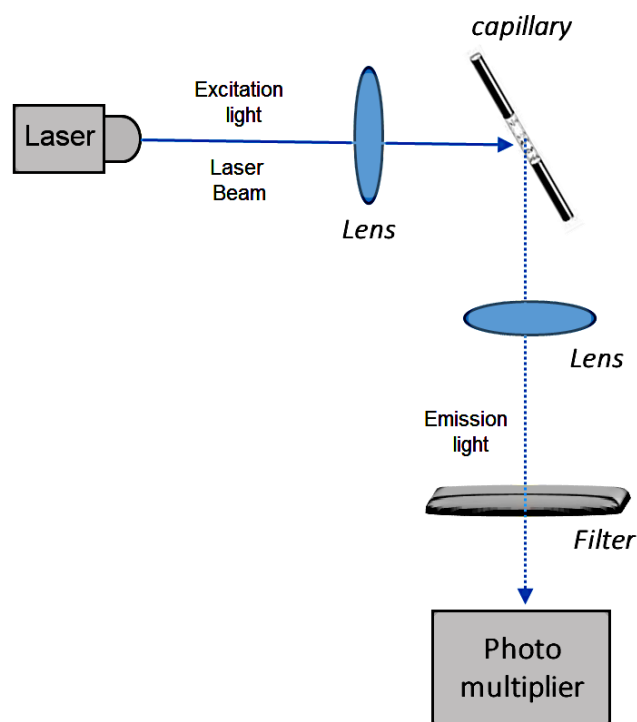


different dyes based on Sanger method provides the advantage of high-throughput and has been applied to DNA sequencing in capillary electrophoresis. Thus, the method has played a crucial role in completing the human genome sequences. Two types of labeling dyes have been utilized for the analysis of nucleic acids in CGE-LIF: non-intercalating dyes and intercalating dyes (e.g., mono- and bis-intercalating dye). For example, non-intercalating dyes such as 5-carboxyfluorescein (FAM), 6-carboxy-*N,N,N',N'*-tetramethyl-rhodamine (6-TAMRA), 6-carboxyrhodamine (ROX) and sulfoindocyanine succinimidyl esters (Cy5) have been commonly used for DNA sequencing by CE-LIF. Common intercalating dyes include ethidium bromide (EtBr), ethidium homodimer (EthD), benzoxazolium-4-pyridinium dimer (POPO-3), benzoxazolium-4-quinolinium dimer (YOYO-1) and benzothiazolium-4-quinolinium dimer (TOTO-1), and so on that can be inserted between adjacent base pairs of dsDNA [54].

In LIF detection, the two special types of lasers commonly used are the continuous wavelength or pulsed ones. The first CE system with LIF detection was introduced by Zare *et al.* [55] with largely improved limit of detection (LOD) [56]. Later, Mathies *et al.* described a confocal fluorescence detection setting using very same lens set to focus the illuminating laser beam into the center of the capillary and to collect the fluorescent light emitted by the fluorophore-labeled components (either covalently or non-covalently tagged) [57, 58]. This detection design proved to be relatively simple and very efficient, but some re-optimization of the alignment required on a daily basis [59, 60]. In 1990s, Dovichi *et al.* introduced a sheath-flow fluorescence detection cuvette to achieve high sensitivity with low light losses and/or scatter [61, 62]. In addition, CE system with LIF detection was used to sequence the Human Genome. This detector type neared the end of the capillary recorded the fluorescent signal in four different spectral channels to resolve the fluorescence signature from the four dyes [63]. Today in CE, commonly used lasers are the helium-neon, which generates light at 543 nm, 594 nm, 604 nm, 612 nm and 633 nm; helium–cadmium laser, which emits light at 325 nm, 354 nm and 442 nm; and argon–ion laser (air cooled), which radiates light at 457 nm, 472 nm, 476 nm, 488 nm, 496 nm, 501 nm and 514 nm.

Using an LIF detector, the excitation light from the laser is focused to the detection window, close to the outlet end of the separation capillary, and the emitted light is measured as shown in Figure 2. Laser-quality mirrors and lenses are used to orient the laser beam to the detection area of the separation capillary. In CGE-LIF detection, the use of these mirrors with good reflectance is not usually a limiting factor, because most of the lasers have more power than necessary to generate the fluorescent signal in the narrow bore columns.

However, the disadvantages of these mirrors are that owing to the long path of the rays from the laser source to the capillary, noise is often caused problems in the detector by small vibrations of the instrument [64]. Most of these limiting factors can be avoided if optical fibers or ball lenses are used, which collect the light from the laser source and focus onto the separation capillary [65]. In LIF detection, very low light emissions can be measured by using such devices as photomultiplier tubes, avalanche photodiodes, or image detectors like charge coupled devices (CCD) and intensified charge coupled devices [64].



**Figure 2.** Laser-induced fluorescence (LIF) detection in CGE. Adapted from [22].

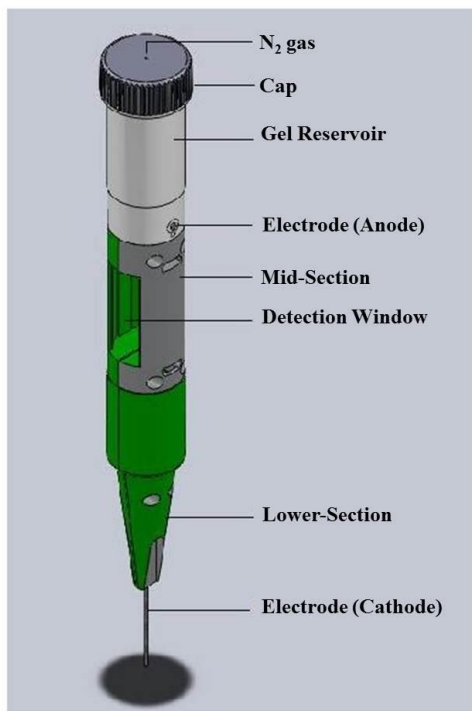
#### **2.5.4. Light-Emitting Diode-Induced Fluorescence Detection**

For high detection sensitivity, the use of a strong light source is required to achieve good signal intensity. Although for many years, laser detectors were exclusively used for this purpose, lately LEDIF detections have emerged [66]. Today in CGE, new-generation LEDs provide an alternative light source to lasers, because these small solid-state light sources are less expensive and consume very low energy [67]. For the separation of amino acids or proteins, some fluorescent labeling dyes are routinely utilized in CGE-LEDIF detection, such as FITC, FQ, NDA and 5-carboxy-tetramethylrhodamine N-succinimidyl ester (5-TAMRA) [68]. For CGE-LEDIF-based glycan profiling experiments, APTS and ANTS are routinely used [69].

Carboxyfluorescein dyes such as fluorescein, 6-carboxyfluorescein and rhodamine B are used for the separation of ssDNA oligonucleotides [70]. LEDs have also been used from 390 nm to 440 nm in few cases, e.g. an ultraviolet-emitting laser diode based on gallium nitride [71]. LEDs can also be utilized in integrated optical detectors in microchip designs [46].

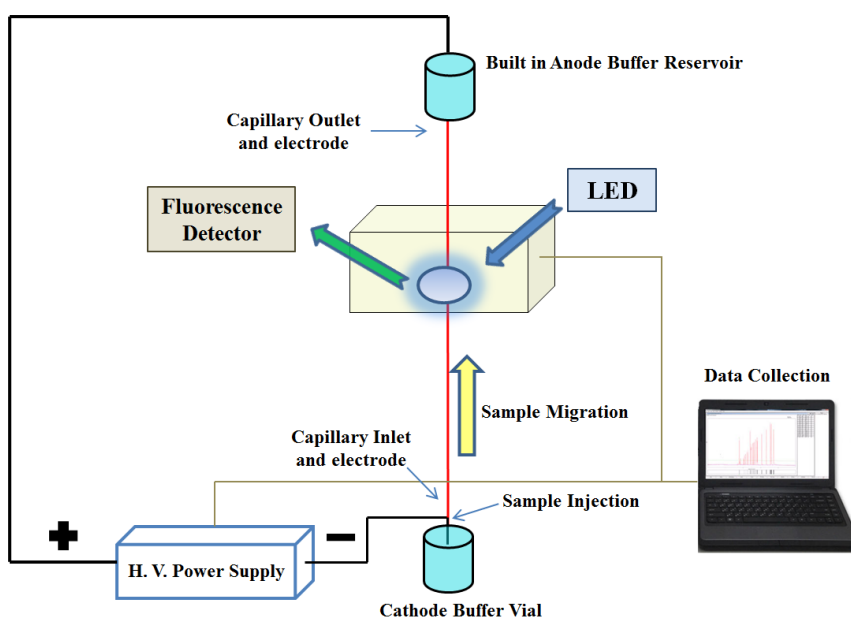
## 2.6. Novel CGE Instrument Designs

In recent years, a novel single-channel instrument has been introduced in CGE with real-time LED-induced fluorescent detection (LEDIF), using a disposable, pen-shaped cartridge that incorporated single separation capillary tubing, a lower (cathode) and upper (anode) electrode and a detection window (Figure 3). To accommodate separation matrix replacement, the pen-shaped cartridge consists of a compact injection molded body with an integrated running gel-buffer reservoir that directly coupled to a modular nitrogen pressure source as illustrated in Figure 3. Depending on the viscosity of the sieving matrix, pressures of up to 60 psi can be applied to the capillary tubing through the top gel-buffer reservoir for gel replacement. The overall length of this separation capillary is in the range of 15-20 cm (generally 50-75  $\mu\text{m}$  i.d. with 360  $\mu\text{m}$  o.d.). The external detection optics can automatically clamp onto the capillary cartridge, not requiring any fine alignment in respect to the detection zones after the installation. This approach offers simplicity in the use of the capillary cartridge and accommodates easy cartridge replacement [72].



**Figure 3.** The front view of the pen-shaped capillary cartridge. Adapted from [72].

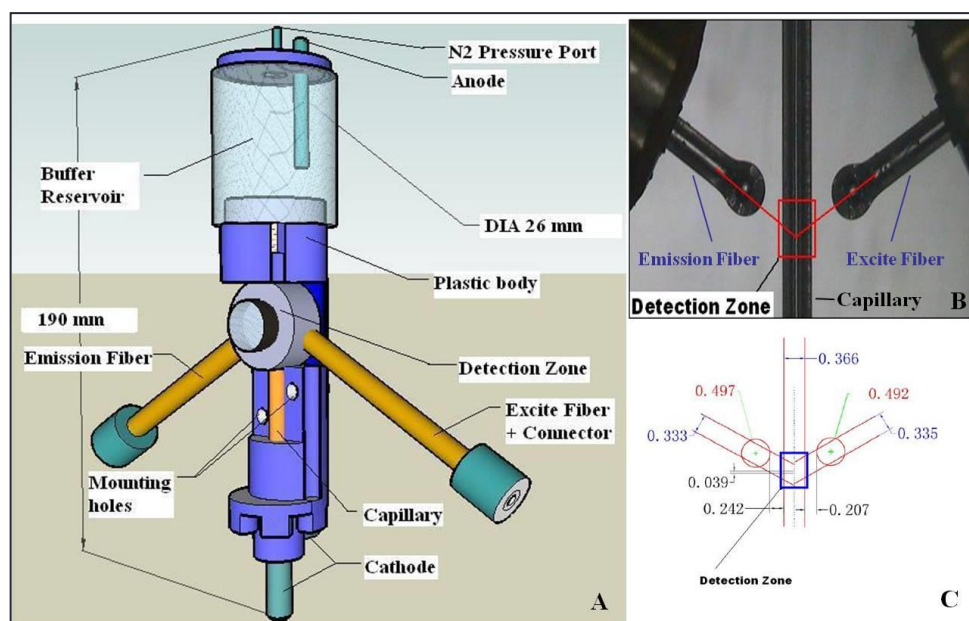
Figure 4 depicts the schematics of this CGE-LEDIF system, including the separation platform, the high-voltage power supply and the LEDIF detection. The instrument utilizes electrokinetic injection for sample introduction. During the injection process, the lower electrode (the cathode) with its embedded capillary is dipped into the sample well, the negatively charged sample components are electrokinetically forced by the applied voltage into the capillary toward the upper electrode (the anode). When the fluorescently labeled sample molecules migrate through the detection window of the separation capillary, the excitation light transports through by the excitation fiber from the LED source. The resulting fluorescent signal is collected by the emission fiber from the detection area and transferred to the PMT for detection and data acquisition. For sample injection and separation, a high-voltage power supply is utilized to provide 500-20 000 V of electrical potential to the separation capillary. The detection optics includes a super-bright royal blue LED (505 nm with FWHM about 30 nm for dsDNA or 380-390 nm without any filter for fluorophore-labeled glycan analysis) as excitation radiation source. Furthermore, the PMT detector contains an emission filter, 590-650 nm long pass filter for dsDNA or 525±45 nm band pass filter for ANTS-labeled glycan analysis to improve the detection sensitivity. To move the sample and buffer tray, the CGE instrument also comprises a fully automated modular X-Y-Z translation stage mechanism with a hybrid linear actuator and a stepper motor. The sample tray is able to accept a single microfuge tube, a 12-well sample strip or a 96-well microtiter plate interfacing with the disposable single-capillary gel-cartridge [72].



**Figure 4.** Schematic representation of the CGE instrument with LED-induced fluorescent detection design. Adapted from [72].

In addition, the detection optics consists of microball ended incident (excitation fiber) and output (emission fiber) optical fibers. Figure 5A shows the configuration of this detection optics. Using microball ended optical fibers, the incident radiation is transferred from the LED to the detection window and the output radiation is collected from the detection window as illustrated in Figure 5B, which is located at approximately  $140^\circ$  apart from each other to increase detection sensitivity. To reduce the background fluorescence, both excitation and emission fibers with microball tips are positioned at the opposite sides of the capillary tubing in a noncontact mode. The tips of the microball end of the optical fibers are spaced at approximately 200-250  $\mu\text{m}$  from the external wall of the capillary to prevent any physical damage to either the separation capillary or the microballs. Figure 5C illustrates the dimensions of the microball ended optical fibers, which have 200- $\mu\text{m}$ -diameter cores to orient the light within an external sheathing, and 350- $\mu\text{m}$ -diameter microball shaped tips (i.e. the ratio of the fiber core diameter to the ball diameter is 1:1.75).

The microball lenses establish a very strongly-built optical fiber assembly and effective optical alignment with respect to the separation capillary. By having the microball at the end of the optical fibers there are no additional micro lenses necessary simplifying in this way the optical alignment. The microballs also offer higher numerical aperture and simultaneously higher power density focusing on the inside of the separation capillary, and besides provide strength to the tip of the fiber to protect from possible mechanical damages. The angles of the optical fibers are about  $70^\circ$ . In contrast with flat-end fibers (i.e. bare fiber, without microball lenses), the microball ended fibers offer good focusing of incident radiation for the excitation fiber and high collection efficiency (high numerical aperture,  $> 0.3$  NA) for the emission fiber. To obtain high fluorescence detection sensitivity, the distal output end of the excitation fiber provides good coupling efficiency inside the capillary. Finally, the fiber optic and modular design coupling allow easy exchange of the excitation radiation sources, i.e. to a laser module or other type of light sources [72].



**Figure 5.** Optical track design. (A) Schematic representation of the detection optics of the gel-cartridge including detection optic configuration, (B) The actual center plane sectional view at the detection region in the gel-cartridge in panel A, (C) Dimensions of the microball ended incident (excitation fiber) and output (emission fiber) detection optics as well as the separation capillary. Adapted from [72].

## 2.7. Capillary Gel Electrophoresis of Nucleic Acids

Currently, most molecular biology laboratories still utilize manual PA or agarose slab gel electrophoresis methods for the size separation of oligonucleotides, which require time-consuming and labor-intensive manual steps, poorly automated, and also lack adequate resolving power in one-dimensional format [21]. CGE is an emerging tool used in the molecular biology and biotechnology industry laboratories, especially for the analysis of nucleic acids and protein purity studies. The method provides a reliable, easy-to-use and fully-automated approach with rapid separation times, excellent resolution, high sensitivity and ruggedness [72]. High-throughput and large-scale analysis of nucleic acids by CGE involves polymerase chain reaction (PCR) product sizing [73], DNA restriction fragment analysis [74], genotyping [75], and last but not least DNA sequencing [21].

The investigation of single nucleotide polymorphism (SNP) is one of the most commonly used techniques in the nucleic acid field, which can be detected by regular CGE format [76]. Barta *et al.* reported on rapid molecular diagnostics of 21-hydroxylase deficiency by detecting the most common mutation in the 21-hydroxylase gene using primer extension and CE with a PVP

matrix [26]. In this study, first the region with the site of interest was amplified by PCR, followed by separation of the two possible primer extension products (mutant and wild-type) and the Cy5-labeled primer by CGE within 90 s. Constant denaturant capillary electrophoresis (CDCE) is another CGE approach, which yields large-scale analysis of DNA variations based on their differential melting temperature for applications including mutation analysis and SNP discovery in individual and pooled samples [77]. To identify SNPs in *Scnn1a* and *Scnn1b* genes, Xue *et al.* also used the CDCE for the separation of pooled blood samples, and a two-point LIF detection setup was introduced to improve mutation identification [78].

Allele-specific PCR is another frequently used genotyping method, also known as the amplification-refractory sequence variant system [79]. In this amplification method, allele-specific primers are used to hybridize at the 3'-end of the SNP site, followed by amplification using a DNA polymerase enzyme lacking 3'-exonuclease activity, thus amplification in that case occurs only if the primer sequence perfectly matches with the template sequence. Szántai *et al.* reported on a novel haplotyping method to investigate two SNPs (-616CG and -521CT) in the Dopamine D4 receptor gene by CGE [80]. Our laboratory reported successful haplotyping of two adjacent miRNA-binding SNPs in the Wolframin gene (*WFS1*) by combining allele-specific amplification and rapid CGE analysis with LEDIF detection [2]. DNA sequencing by CGE became another mature field in a remarkably short period of time. The first paper on DNA sequencing option by CGE is appeared in the 1980s suggesting that CGE can solve the DNA sequencing problems [73, 81]. Later, Mathies *et al.* introduced a novel multicapillary array system for large-scale analysis of DNA sequencing fragments [44]. These capillary array tools also required sensitive detection design due to the use of capillary bundles. Dovichi *et al.* reported on a high-sensitivity post-column LIF detector, based on a sheath-flow cuvette to minimize background signal arising from the light diffraction of the gel-filled capillary (as mentioned in section 2.5.3.) [61]. At present, despite of the introduction of second-and third-generation DNA sequencing techniques, fully-automated, easy-to-use multicapillary array systems are still in use to analyze hundreds of DNA samples with long read lengths [82].

The use of microfabricated capillary arrays offers new opportunities in multiplexing nucleic acids research. Easley *et al.* developed a fully-integrated microfluidic genetic analysis device with sample-in-answer-out capability [83]. To analyze DNA fragments, some experiments have been accomplished in nanolits, nanochannels, channels with nanosized entropic traps and pillars [84].

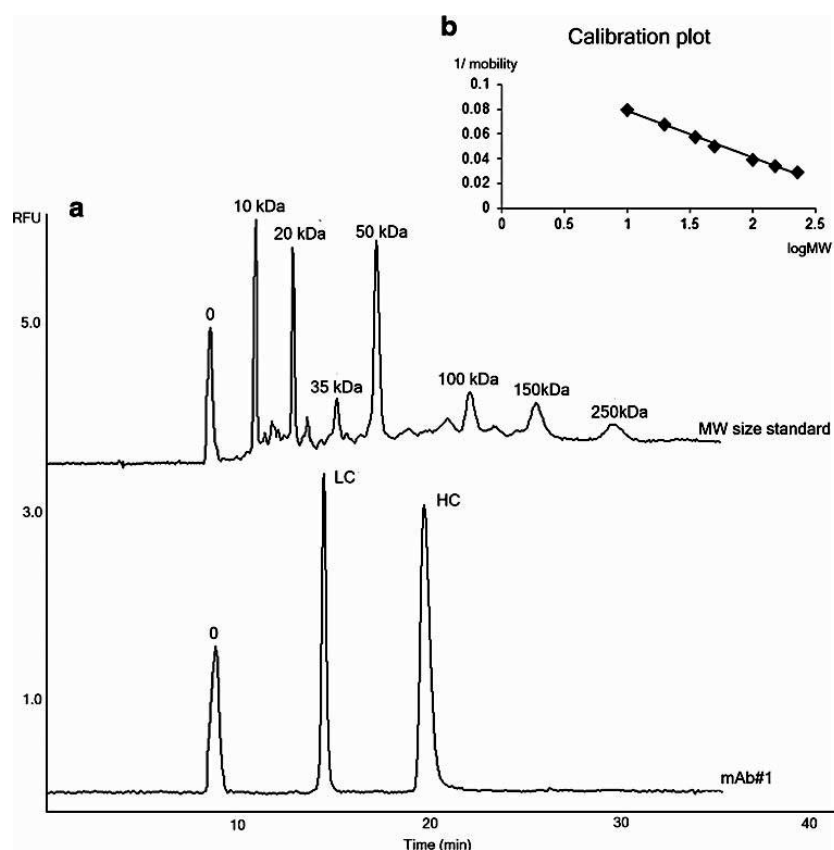
## **2.8. Protein Separation by Capillary Gel Electrophoresis**

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) is a commonly used technique in the biopharmaceutical field for size separation and quality control of proteins [28]. Traditional SDS-PAGE is time-consuming and labor intensive, also requiring improvements in terms of resolving power. SDS-capillary gel electrophoresis (SDS-CGE) offers many advantages over classical SDS-PAGE by providing on-column detection with excellent sensitivity, high-resolution and the capability of accurate quality control, fully-automated operation, and molecular mass evaluation [68, 85].

As early as in the 1980s, the first papers on SDS-CGE were reported by applying cross-linked PA or agarose as sieving matrices [86, 87]. These polymers were directly filled into the capillary tubing before the separation process, either as polymerization reaction mixture (cross-linked PA) or as hot solutions (agarose). In the early 1990s, cross-linked PA was polymerized in the capillary tubing but the lifespan of these capillaries was limited, i.e. <10 runs per capillary, with poor run-to-run reproducibility. Later, water-soluble and replaceable linear or slightly branched polymers were developed such as low concentration cross-linked PA [88-90] or linear PA [27, 91, 92], PEO [93], PEG [27], dextran [27], and hexosyl polymer (pullulan) [94, 95] as sieving gels for the separation of SDS-protein complexes. Currently, SDS-CGE kits are commercially available for quality control of therapeutic proteins such as recombinant monoclonal antibodies [96-99]. Our laboratory reported on high-throughput analysis of therapeutic monoclonal antibodies by SDS-multicapillary CGE (SDS-mCGE) as shown in Figure 6 [98].

SDS-CGE was successfully conjugated with matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) MS by using a moving poly(tetrafluoroethylene) membrane past the outlet end of the capillary tubing to continuously collect fractions for MS analysis [100]. CGE-MALDI-MS has also been used for the quality control of monoclonal antibodies in the biopharmaceutical industry [101].





**Figure 6.** High-resolution analysis of a therapeutic monoclonal antibody by SDS-multicapillary gel electrophoresis (SDS-mCGE). (a) Upper trace: covalently fluorophore tagged molecular mass ladder; lower trace: reduced mAb sample (0 remaining labeling dye, LC and HC represent the light and the heavy chains of mAb). (b) The calibration plot for molecular mass assessment in the 10–250 kDa mass range vs. reciprocal electrophoretic mobility [ $\text{cm}^2/\text{Vs}$ ]. Adapted from [98].

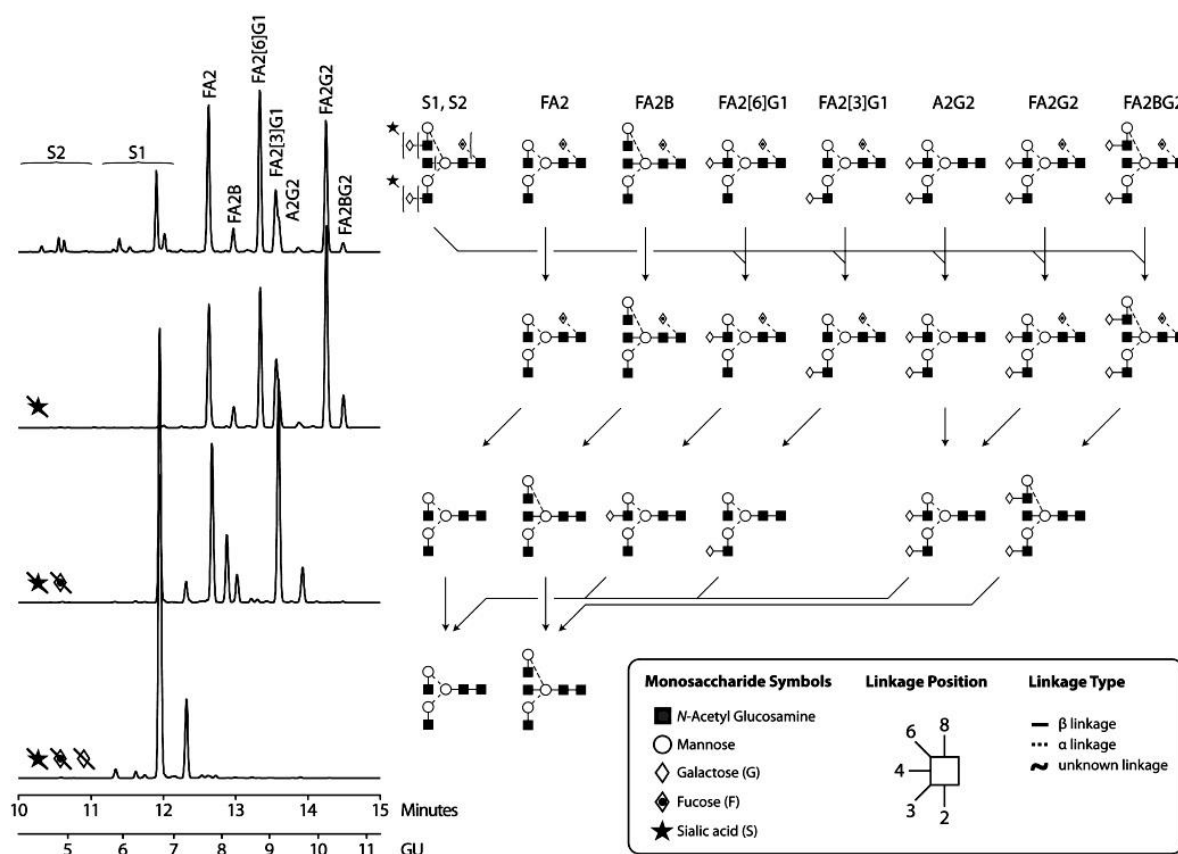
## 2.9. Capillary Gel Electrophoresis of Complex Carbohydrates

Glycosylation is one of the most common and structurally diverse post-translational modifications, which modulates the physical, chemical and biological properties of proteins. In addition, glycosylation plays crucial roles in different biochemical processes including protein stability, folding, localization, and cellular communications [102]. Some changes in biological activities are often the result of alterations in glycosylation, i.e., the same glycoproteins may possess by diversified carbohydrate moieties, even at the same glycosylation site (microheterogeneity), or exhibit different glycosylation site occupancy (macroheterogeneity) [103]. Aberrant glycosylation and the discovery of associated potential oligosaccharide-biomarkers have been noted in many disease including immune deficiencies, cardiovascular syndromes and cancer [104].

Currently, structural elucidation of protein derived *N*-linked oligosaccharides poses a substantial analytical challenge, because most native glycans lack chromophore or fluorophore features in their structures that would enable their UV-VIS or LIF detection. Due to the discovery of glycan-based biomarkers, MALDI-TOF MS and high-performance liquid chromatography (HPLC) of fluorescently labeled oligosaccharides have emerged as techniques for the analysis of biological fluids or tissue samples [105, 106]. Prior to MALDI-TOF MS, permethylation of *N*-glycans is often a necessary sample preparation step. In spite of these advances, MS is unable to differentiate between isobaric glycan structures without using tandem MS, which leads to problems with quantitation originating from competitive ionization of different glycan structures within the ion source of the mass spectrometer [107]. The great advance of CGE in carbohydrate separation is the ability to distinguish both linkage and positional isomers [108, 109]. In addition, CGE-LIF detection also offers glycan sequencing options by top-down digestion and bottom-up identification using sugar-specific exoglycosidases as illustrated in Figure 7 [109].

Typically, the carbohydrate moiety of a glycoprotein is enzymatically released with an endoglycosidase, such as peptide-N4-(*N*-acetyl- $\beta$ -glucosaminyl)asparagine amidase (PNGase F), endo- $\beta$ -*N*-acetylglucosaminidase H (endo-H), etc. While PNGase F liberates all classes of *N*-linked glycans from mammalian glycoproteins, endo-H specifically released high-mannose and hybrid type oligosaccharides [110]. In most cases, the liberated glycosylamines are converted to aldehyde form (reducing sugars) at acidic pH, and reacted with primary amino group containing dyes via reductive amination [70, 111, 112]. In most instances, for CGE the carbohydrates are labeled with APTS fluorescent dye. For glycan labeling, the use of other charged fluorescent reagents was also discussed in section 2.5.3 and 2.5.4. Aldose labeling with 2,3-naphthalenediamine to produce highly fluorescent naphthimidazole derivatives were introduced by Kuo *et al.* [113]. Our laboratory reported the successful establishment of an ANTS-labeled *N*-glycan database for rapid (<200 s) CGE-LEDIF-based glycan profiling [114].

CGE-LIF is an excellent orthogonal technique to LC or MS based methodologies for the investigation of carbohydrates based on their different separation mechanisms [109]. In addition, CE was successfully conjugated with MS for the analysis of glycans for glycoprofiling of biotherapeutics [101, 115, 116]. Another advantage of CGE-based techniques is the option of easy multiplexing up to 12, 48, or even 96 capillaries for high-throughput applications [117, 118].



**Figure 7.** Overview of the experimental strategy for exoglycosidase digestion-based carbohydrate sequencing of IgG N-glycans by CGE-LIF. Adapted from [109].

### 2.9.1. Theory of Glycan Structure Elucidation by CGE

**Oligosaccharide Standards.** Co-injection of purified glycan standards with an unknown sample mixture poses a simple and effective way for glycan structure elucidation. The analysis of glycans by CGE is based on their different hydrodynamic volume to mass ratios. This proved to be decisive in the separation of glycans built of different monosaccharide blocks and featuring diverse positional and linkage patterns. For correct structural assignment, the use of glycan sequencing is a crucial step [108]. Anticipation of molecular size from analyte migration time can be achieved by the ancillary separation of gradually sized oligomeric sugar standards. Optimal reference standards involve equally distributed oligosaccharide pools showing a linear relationship between molecule size and migration time. Thus, linear homooligosaccharide ladders with degrees of polymerization (DP) 1, 2, ..., n, of glucose ( $\alpha 1 \rightarrow 4$  or 6), glucose ( $\beta 1 \rightarrow 4$ )<sub>n</sub>, created by hydrolyzing starch, chitin-type polysaccharide chains, and cellulose are

frequently applied as standards. Hydrolysates of branched structures, such as high-mannose type glycans, are usually recommended due to the limited DP range and the occurrence of positional isomers. To assure the highest possible precision of migration time measurement, molecular size standards can either be co-injected with the sample or analyzed in a separate experimental run, where alignment standards are generally introduced in both standard and sample run.

**Migration Time Normalization.** High accuracy CGE-based glycan analysis can be achieved by normalizing the migration times of the peaks of interest with a lower (migrating faster than any peak of interest) and an upper bracketing standard (migrating slower than any sugar structure of interest in the sample mixture) [119]. The migration times of the peaks within the boundaries can be converted to relative migration times ( $RMT_x$ ) by applying the following equation [120]:

$$RMT_x = \left( \frac{MT_{upper}}{MT_x} \right) \times \left( \frac{MT_x - MT_{lower}}{MT_{upper} - MT_{lower}} \right) \quad (\text{eq 18})$$

According to equation 18, migration times normalize into the interval between 0 and 1, where  $MT_x$  is the migration time of the analyte of interest,  $MT_{lower}$  and  $MT_{upper}$  are the migration times of the lower and the upper bracketing standards, respectively. Thus, the analytical precision increases significantly, not only between consecutive repetitive runs, but also between oligosaccharide ladder standard and sample runs.

**Calculation of Glucose Unit (GU) Values.** Glucose unit (GU) values are derived from the relative migration times of separately run maltooligosaccharide oligomers and calculated for the sample peak of interest based on its adjacent glucose oligomers peaks as characterized by equation 19:

$$GU_x = n + \left( \frac{RMT_x - RMT_n}{RMT_{n+1} - RMT_n} \right) \quad (\text{eq 19})$$

where  $GU$  is the sugar unit value of the unknown oligosaccharide,  $n$  is the degree of polymerization (DP) of the preceding maltooligosaccharide ladder peak.  $RMT_x$  is the relative migration time of the unknown sample peak.  $RMT_n$  and  $RMT_{n+1}$  are the relative migration times of the homooligomers immediately preceding and following the peak of interest, respectively.

The generation of GU values of purified oligosaccharide standards with known structural identity helps the establishment of a GU database, which can support elucidation of unknown sugar structures [108].

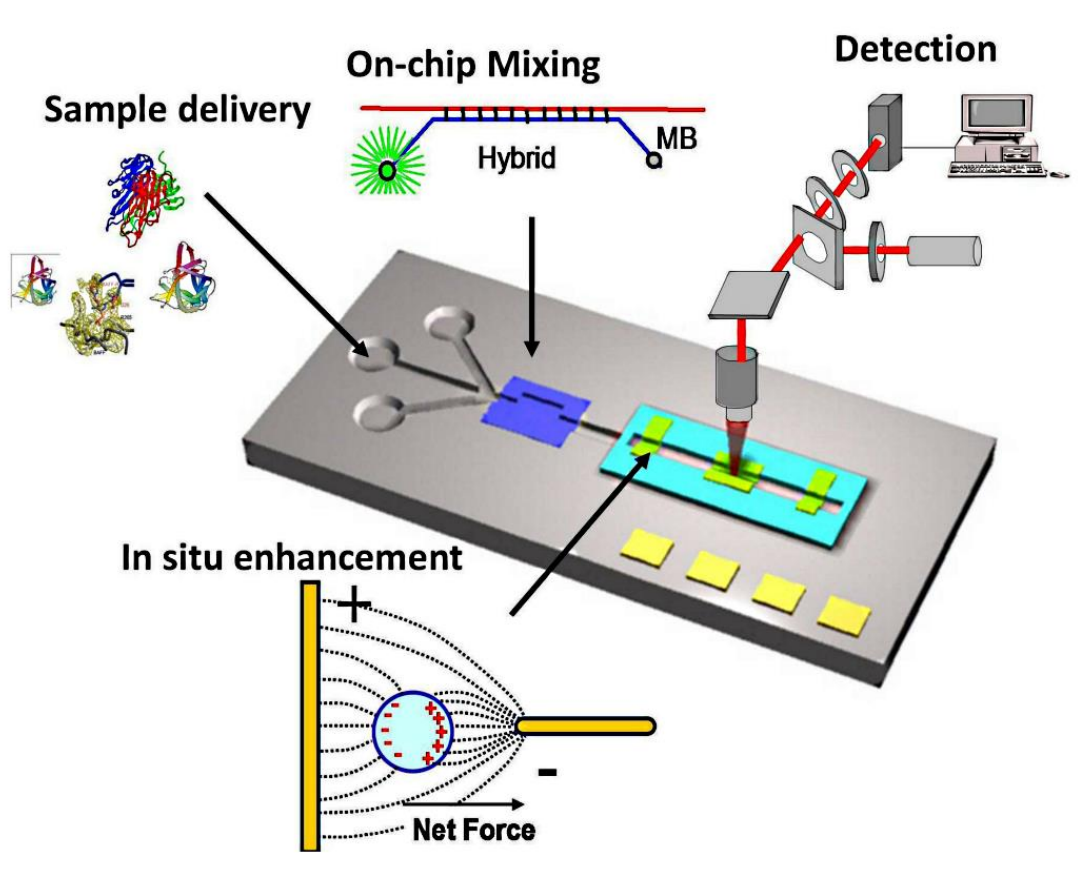
## **2.10. Future Trends: Miniaturization, Lab-On-a-Chip Systems**

To integrate multiple analytical processes on monolithic chip platform, a microfabricated lab-on-a-chip (LOC) devices have been developed, comprising sample pretreatment, solution distribution/mixing, separation, detection, and all other necessary compartments as illustrated in Figure 8 [121]. The major purpose of the utilization of microfluidic analysis systems is to apply standard laboratory steps in a miniaturized format. Microfabricated CGE systems are very fast, separations take place in the seconds scale, because of the short length of the separation channels [28]. Capillary electrophoresis microchips utilize channels etched into planar substrates by applying common microfabrication methods from the semiconductor industry such as photolithography and chemical etching. In the miniaturized format, nanoliter sample volumes can be introduced with a simple cross or double T-injector structures [122]. The separated analyte molecules are typically detected by confocal microscopy with LIF detection. These features make CE microchips attractive technology platforms for next-generation instrumentation. CE microchips have been utilized for the separation of biopolymers such as DNA fragments [123, 124], PCR-product sizing [125], genotyping [126], DNA sequencing [127], protein separation [28, 128, 129], carbohydrate profiling [130], cell and tissue studies [131], and environmental monitoring [132].

Ultrathin-layer gel electrophoresis is another approach of miniaturization, which combines slab gel electrophoresis and CGE techniques, providing multilane separation to achieve rapid and efficient analysis of biomolecules [133]. Lab-on-a-chip systems can also be integrated with applications of micromachined chromatographic methods, i.e., high-performance micro- and nanovolume liquid chromatography (LC) [134]. Toward fully-automated lab-on-a-chip systems, various microfluidic lab-on-a-chip platforms have emerged from the effective integration of multiple microfluidic components for sophisticated biomedical analyses [135]. In the field of genomic research, Rech *et al.* first introduced a portable microchip electrophoresis (ME)-based genotyping system including a four-color solid-state single photon avalanche photodiode (SPAD) array detector [136].

In recent years, Rhazi *et al.* reported on a high-throughput microchip-based protein assay for quantitation of high-molecular-weight glutenin subunits [137], which provides comparable resolution and sensitivity to conventional reversed-phase HPLC (RP-HPLC), but in approximately 100 times faster time scale (45 s per sample analysis vs. 80 min for RP-HPLC). In 2012, He *et al.* introduced a tunable thick polymer coating for on-chip electrophoretic protein and peptide separation to control the effect of EOF and eliminate non-specific protein adsorption [138]. Furthermore, ME technology is also utilized for miniaturized bioanalytical devices in clinical analysis [139]. In the field of cancer research, Zhang *et al.* introduced a microfluidic beads-based immunosensor for the sensitive detection of a cancer biomarker protein,  $\alpha$ -fetoprotein, using multienzyme amplification and quantum dot labeling [140]. Wang *et al.* established an application of poly(dimethylsiloxane)/glass microchip for fast electrophoretic separation of small, dense low-density lipoprotein in the field of clinical research [141]. In recent years, Kamaruzzaman *et al.* reported on a sensitive chemiluminescence microfluidic system using gold nanoparticles enhanced luminol silver nitrate for the identification of vitamin B12 [142]. For screening monoclonal antibody product quality, microchip CE-SDS systems are commonly used in the pharmaceutical industry, which offer sufficient resolution and sensitivity for this purpose, but faster [143]. Primack *et al.* established a high-throughput microchip-based glycan screening assay for antibody cell culture samples, including clone selection and cell culture process optimization [144]. With this microchip CE, the relative levels of high-mannose, fucosylated and galactosylated glycoforms in the Fc domain were measured for hundreds of crude cell culture samples in a few hours (i.e., less than 1 min per sample).

In recent years, microchip devices have been successfully coupled to electrospray ionization (ESI) and MALDI-TOF MS for the analysis of biological samples [145]. The main advantage ME-ESI-MS is the low flow rate that results in seamless interfacing without disrupting the electrophoretic separation.



**Figure 8.** Lab-on-a-chip electrophoresis design. Reproduced with permission from Ref. [146].

### 3. AIMS AND MOTIVATION

Capillary gel electrophoresis (CGE) is a rapidly developing technique, which is routinely used mostly in biopharmaceutical and biomedical analyses. This dissertation is started with a summary of the theoretical and practical aspects of CGE including separation principles, instrument design, capillary coatings, and the most frequently used gel-buffers. It is followed by the description of various detection options, such as UV-Vis, diode-array, laser-induced fluorescence, and light-emitting diode-induced fluorescence types. Then, an overview of the key application areas is given with the main emphasis on nucleic acids, proteins, and complex carbohydrate analysis, as well as the latest developments in microchip electrophoresis devices. The aims of this study were as follows:

1. To introduce a novel single-channel capillary gel electrophoresis system with microball ended emission optical fiber based LED-induced fluorescence detection design. To evaluate the suitability of this separation system, the degradation level of a large number (cc 1000) gDNA samples were analyzed to assess their quality, focusing on migration time reproducibility, limit of detection, as well as detector linearity in comparison to conventional agarose slab gel electrophoresis.
2. After the reproducibility studies, the aim was to develop a haplotyping (i.e. simultaneous multiple genotyping) method of two adjacent miRNA-binding miR-SNPs (*rs1046322* and *rs9457*) in Wolfram syndrome-1 gene (WFS1) by combining double-tube allele-specific amplification and rapid capillary gel electrophoresis with LED-induced fluorescent detection to analyze the resulting DNA fragments.
3. The increasing use of glyco-markers and the huge market of therapeutic monoclonal antibodies requires the generation of synthetic carbohydrate antigens. In the third part of this study, the aim was to investigate polyclonal antibody response for newly synthesized maltose-BSA conjugates. First of all, a simple carbohydrate, maltose, was conjugated to a carrier protein (BSA) by reductive amination to improve their immunogenicity. After that, the synthesized neoglycoproteins were investigated by SDS-CGE as well as MALDI-TOF MS, and then induced immune response against themselves.



This elaborated method may be utilized in order to generate monoclonal antibody libraries for the discovery of mAbs in normal and disease models with unique sugar specificities, which can be readily applied to biomedical research and clinical diagnostics.

4. In the fourth part of this study, the aim was to introduce an 8-aminonaphthalene-1,3,6-trisulfonic acid (ANTS) oligosaccharide database for capillary gel electrophoresis using LED-induced fluorescent detection. The validation of the generated GU values in the database was verified with *N*-glycans released from human immunoglobulin G and bovine pancreatic ribonuclease B. Currently, there are no structural databases for CGE-LEDIF separations, which contributed to the novelty and importance of this topic. In the absence of database, the migration properties must be calculated with the use of purified *N*-linked carbohydrate standards or in conjugation with other analytical techniques (e.g., NP-HPLC, MS, etc.). However, this generated database may prove useful for fellow separationists in the biomarker research and drug development. It is not necessary to separate large carbohydrate standards before each evaluation, thus, the unknown structures can be identified based on the migration times of the peaks of standards.
5. Finally, the activation energy concept related to the electromigration of carbohydrates were studied in CGE-LIF using linear (sugar oligomers with  $\alpha$ 1-4 linked glucose) and branched (sialylated, neutral and core fucosylated IgG) *N*-glycans. The aim of this study was to determine the electrophoretic mobility shift between the different structure oligosaccharides at different temperature/background electrolyte. The APTS-labeled oligosaccharides were separated in the temperature range of 20 – 50°C, using either viscosity modifier (ethylene glycol) or polymeric additive (linear polyacrylamide, polyethylene oxide) containing BGEs. Previous reports have not noticed the effect of separation temperature on structure specific glycan migration using viscosity modifiers and/or polymeric additives containing BGEs. These factors may influence the stability of structural database, which contributed to the novelty and importance of this topic.

## 4. MATERIALS AND METHODS

### 4.1. Large-Scale Genomic DNA Analysis

For large-scale genomic DNA (gDNA) analysis, the gDNAs were from the Clinical Genomics Center (Debrecen, Hungary) and kept at -20°C in TE buffer (10 mM Tris, 1 mM EDTA, adjusted to pH 8.0 with HCl) until use.

#### 4.1.1. Analysis of gDNA by CGE

For CGE-LEDIF separation, the GeneRuler 1kb Plus DNA sizing ladder (BioCenter, Szeged, Hungary, size range of 75 – 20 000 bases) was diluted to 50 ng/μL and the gDNA samples were diluted to 25 ng/μL final concentration with HPLC-grade water (Sigma-Aldrich, St. Louis, MO, USA). For the migration time reproducibility study, the 100 bp Plus DNA molecular mass marker (BioCenter, size range of 100 – 3 000 bases) was used for this evaluation in a final concentration of 25 ng/μL with HPLC-grade water. After dilutions, all gDNA samples and DNA molecular mass markers were aliquoted and stored at -20°C until use. CGE-LEDIF separations were accomplished by a novel single-channel DNA-CE instrument (Qsep100, BiOptic, New Taipei City, Taiwan) containing a 75 μm i.d. (360 μm o.d.) bare fused silica capillary with the effective separation length of 11 cm (total length: 15 cm). The optical detection system comprised a super-bright royal blue LED (Cree XLamp, New York, NY, USA) as excitation radiation source. The CGE instrument also included a high sensitivity multialkali PMT for data acquisition (Hamamatsu Photonics, Bridgewater, NJ, USA). DNA-CE high-resolution gel-buffer and DNA-CE running buffer (Qsep100, BiOptic) was used for the CGE separations. To accommodate fluorescent detection, the gel-buffer system contained ethidium-bromide. Prior to use, new capillary was washed first with 5 mL approximately 70°C MilliQ-grade water (Millipore, Billerica, MA, USA) and the capillary was rinsed for 500 s followed by the transfer of 5 mL DNA-CE gel buffer into the gel reservoir and purging for 2 × 1 000 s. To avoid any possible carryover, the sieving gel was replaced in the capillary by means of a 10 s purge step between injections, followed by the injection of an HPLC-grade water plug (4 kV for 3 s), and the electrokinetic sample introduction from a 96-well plate at 4 kV for 6 s. Each separation was carried out at ambient temperature applying 8 kV for 300 s with a high-voltage power supply (EMCO, Sutter Creek, CA, USA). Data analysis was performed using the Q-Analyzer software package (BiOptic).

#### **4.1.2. Analysis of gDNA by Agarose Gel Electrophoresis**

For agarose slab gel electrophoresis, the 1 kb Plus DNA sizing marker (Biocenter) was diluted to 0.1 µg/µL final concentration prior to use. A total of 500 ng gDNA samples (in TE buffer) were injected into the sampling wells of the agarose gels. Agarose slab gel electrophoresis was performed in an E-Gel iBas Power System (Thermo Fisher Scientific, Darmstadt, Germany) for separation and the E-Gel Safe Imager Real-Time Transilluminator (Thermo Fisher Scientific). Precast agarose gels containing 2% agarose (2 × 8 wells, size range of 100 – 2 000 bases) E-Gel 2% double comb with ethidium bromide (Thermo Fisher Scientific) was utilized for the separation of the gDNA samples and DNA sizing marker. The molecular mass marker and gDNA samples were applied into the sampling wells of the agarose gels and electrophoresed in the E-Gel system for 8 min, after that, the separated bands were visualized by a blue light transilluminator in real time. Data analysis was performed using the Alpha DigiDoc AD-1201 imaging system (Bio-Science, Budapest, Hungary) equipped with an Olympus digital camera (C-4000ZOOM) with appropriate filters (Olympus, Budapest, Hungary).

### **4.2. Molecular Haplotype Analysis**

#### **4.2.1. Non-Invasive DNA Sampling and DNA Extraction**

The gDNA samples were from the Department of Medical Chemistry, Molecular Biology and Pathobiochemistry (Budapest, Hungary). gDNAs were acquired using non-invasive DNA sampling with the use of buccal swabs from healthy Hungarian volunteers. The study protocol was approved by the Local Ethical Committee (ETT-TUKEB ad.328/KO/2005, ad.323-86/2005-1018EKG from the Scientific and Research Ethics Committee of the Medical Research Council). DNA samples were purified by standard procedure as described [147].

#### **4.2.2. Allele-Specific PCR**

For ultrafast molecular haplotype analysis, the amplification of the *rs1046322* and *rs9457* SNPs was performed by allele-specific PCR. The HotStarTaq DNA polymerase lacking 3' exonuclease activity together with the 10 × reaction buffer, and the Q-solution (Qiagen, Valencia, CA, USA) were applied for the allele-specific PCR and each DNA sample was investigated in two separate reactions.

Both PCR reaction mixtures included approximately 4 ng gDNA template, 200  $\mu$ M deoxyadenosine triphosphate (dATP), deoxycytidine triphosphate (dCTP), deoxyguanosine triphosphate (dGTP), deoxythymidine triphosphate (dTTP), 0.5 U HotStar Taq DNA polymerase, and 1  $\mu$ M of each outer oligonucleotide primer (Sigma Genosys, Woodlands, TX, USA; forward: 5' TCT GTC CAC TCT GAA TAC 3' and reverse: 5' CAG GCT CTT CTA AAC ACT 3'). Reaction mixture-I was employed to investigate the presence of the *rs1046322A* and *rs9457C* alleles, as well as their haplotype, thus it comprised the *rs1046322A* specific sense (5' GAG CCT GAC CTT TCT GAA 3') and the *rs9457C* specific antisense (5' CCA CTA CCT GCT GGA G 3') primers. Reaction mixture-II was used to analyze the other possible variants (*rs1046322G*-specific forward primer: 5' GAG CCT GAC CTT TCT GAG 3', *rs9457G*-specific reverse primer: 5' CCA CTA CCT GCT GGA C 3'). PCR amplification reactions were performed in a total volume of 10  $\mu$ L. The sequences of the primers were designed by the Oligo 5.0 software (Molecular Biology Insights, Cascade, CO, USA). Thermocycling conditions were started at 95°C for 15 min to activate the hotstart DNA polymerase enzyme. This was followed by 40 cycles of denaturation (94°C for 30 s), annealing (55°C for 30 s), and then extension (72°C for 1 min). The last procedure of the PCR was a final extension at 72°C for 10 min after that the PCR samples were stored at 8°C.

For the detection limit and linearity studies by CGE-LEDIF, the PCR mixture included approximately 4 ng gDNA template, 200  $\mu$ M of each deoxyribonucleotide triphosphate (dATP, dCTP, dGTP and dTTP), 0.05 U/ $\mu$ L HotStarTaq DNA polymerase with 1 $\times$  reaction buffer and 1 $\times$  Q-solution, as well as 1  $\mu$ M of each primer (forward: 5' GCC CTT CTC GAG TCT TGC AGC GCC GGA ATA GGC 3' and reverse: 5' GCA GAA GCT TAA GTT GTT CGG GAG CAG CTG AAC G 3'). The amplification reaction was carried out in a total volume of 100  $\mu$ L. The thermocycling was initiated at 95°C for 15 min to denature gDNA and to activate the enzyme; it was followed by 40 cycles of denaturation at 94°C for 30 s, annealing at 65°C for 30 s, and then extension at 72°C for 1 min. The last step of the amplification was a final extension at 72°C for 10 min after that the PCR-product was kept at 8°C.

#### **4.2.3. Analysis of PCR-Product by CGE**

In CGE-LEDIF separations, and the DNA mass ladder (BiOptic; size range of 50 – 3 000 bases, 10.5 ng/ $\mu$ L) as well as the lower and upper DNA alignment markers (BiOptic; 20 base pair, 1.442 ng/ $\mu$ L and 5000 base pair, 1.852 ng/ $\mu$ L) were used for this evaluation.

The WFS1 PCR-fragments (576 base pair, 253.19 ng/ $\mu$ L) were diluted to the appropriate concentrations with dilution buffer (BiOptic) or MilliQ-grade water (Millipore) for the detection limit and linearity studies. Rapid CGE analysis of the PCR-fragments was also accomplished by the DNA-CE instrument using the DNA-CE gel-buffer and DNA-CE running buffer (Qsep100, BiOptic). The capillary was conditioned under the same procedures as mentioned above. Samples were injected electrokinetically at 4 kV for 10 s. To avoid any sample cross-contamination, the capillary tubing was immersed into MilliQ-grade water (Millipore) at 0 kV for 1 s as a washing step. To improve precisions, two DNA alignment markers were used (20 bp and 5000 bp). Separations were carried out at ambient temperature by applying 6 kV electric potential. Data analysis was also performed using the Q-Analyzer software package (BiOptic).

#### ***4.2.4. Analysis of PCR-Product by Agarose Gel Electrophoresis***

The GeneRuler 100 base pair DNA ladder (Thermo Fisher Scientific, FL, USA) was diluted to a final concentration of 0.5  $\mu$ g/ $\mu$ L and kept at -20°C until use. Agarose powder was mixed with 1 $\times$  TAE electrophoresis buffer (40 mM Tris, 20 mM acetic acid, and 1 mM EDTA, pH 8.0) in a final concentration of 2%, and then heated until the agarose powder dissolved completely. Ethidium bromide as an intercalating dye was added to the melted gel in a final concentration of 0.5  $\mu$ g/mL. After jellification at room temperature, 20 ng of PCR samples and the 100 base pair DNA molecular mass marker (100 – 1 000 bp, 0.5  $\mu$ g/ $\mu$ L) containing DNA loading dye (6 $\times$  loading dye: 10 mM Tris-HCl (pH 7.6), 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol and 60 mM EDTA) were injected into the sample wells and electrophoresed for 45 min at 100V (Bio-Rad PowerPac 300, Hercules, CA, USA). The separated DNA bands were detected in a UV light box (Bio-Rad Gel-Doc XR System).

### **4.3. Immune Response against Neoglycoproteins**

#### ***4.3.1. General Procedures***

Thin-layer chromatography (TLC) was performed on Kieselgel 60 F<sub>254</sub> precoated aluminum plates (Merck Millipore, Darmstadt, Germany). Sample compounds were detected by treating the plates with 10% H<sub>2</sub>SO<sub>4</sub> and heating to 140°C. Column chromatography was accomplished with Kieselgel 60 (Merck Millipore, 0.063 – 0.2 mm) with solvents specified under the relevant sections.

All organic phases were concentrated in a rotary vacuum evaporator (Büchi Rotavapor R-114, Flawil, Switzerland) after extraction. NMR spectra were performed using a Bruker (Bremen, Germany) DRX-360 instrument ( $^1\text{H}$  360.13 MHz,  $^{13}\text{C}$  90.56 MHz) at room temperature in  $\text{CDCl}_3$  solvent, with a  $\text{Me}_4\text{Si}$  internal standard.

The uncorrected melting points were carried out by a Kofler apparatus (Dresden, Germany). All chemicals for the synthesis of neoglycoproteins and the general reagents for CE, MALDI-TOF MS, and ELISA tests were purchased from Sigma-Aldrich (St. Louis, MO, USA).

#### 4.3.2. Synthesis of Neoglycoproteins

Neoglycoproteins are carbohydrate-modified proteins. The synthesis of neoglycoproteins can be random or specific using chemical and enzymatic methods. Reducing oligosaccharides can be conjugated directly to amino groups or using functionalized glycan bearing spacer arms between the carbonyl group of sugar and the amino groups of proteins [148]. In this study, a formyl-heptil spacer was used to minimize the risk of introducing a stronger immunogen close to the sugar residue, i.e. this spacer contains no functional groups in the vicinity of the carbohydrate.

*Acetylation step:* 6.8 mg (105 mmol) of sodium acetate was dissolved in 60 mL acetic anhydride and the reaction mixture was stirred for 15 min at 60°C. After that 10 g (29 mmol) maltose (**1**, numbers indicate the structures of compounds in section 5.3.1.) was added to the mixture and stirred for 1.5 h at 110°C. The reaction solution was mixed with ~200 mL ice-water, diluted in 2 to 1 ratio with methylene chloride and washed twice with  $\text{dH}_2\text{O}$ . The combined organic phases were dried over  $\text{MgSO}_4$ , filtered and concentrated in a rotary vacuum evaporator. Finally, the sample was crystallized from methanol resulting in 11.61 g of 2,3,4,6-tetra-*O*-acetyl- $\alpha$ -D-glucopyranosyl-(1 $\rightarrow$ 4)-1,2,3,6-tetra-*O*-acetyl- $\beta$ -D-glucopyranose (**2**) compound (59%) as white crystals.

NMR data for compound **2** (melting point= 158°C):  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  5.85 (d,  $J$  = 8 Hz, 1H, H-1), 5.51 (d,  $J$  = 4 Hz, 1H, H-1'), 5.46 (t,  $J$  = 10, Hz, 1H), 5.40 (t,  $J$  = 9 Hz, 1H), 5.16 (t,  $J$  = 10 Hz, 1H), 5.08 (t,  $J$  = 9 Hz, 1H), 4.96 (dd,  $J$  = 10, 4.00 Hz, 1H, H-2'), 4.56 (dd,  $J$  = 12, 2 Hz, 1H, H-6a'), 4.34 (dd,  $J$  = 12, 3.5 Hz, 1H, H-6b'), 4.18-4.11 (m, 1H, H-6a, H-6b), 4.08-4.03 (m, 1H), 3.99-3.92 (m, 1H), 2.24, 2.21, 2.20, 2.15, 2.13, 2.12, 2.11 and 2.10 (8s, 3H, 8  $\text{COCH}_3$ );  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  170.37, 170.30, 170.23, 169.87, 169.68, 169.40, 169.25 and 168.60 ( $\text{COCH}_3$ ), 95.59 and 91.13 (C-1 and C-1'), 75.06, 72.87, 72.38, 70.81, 69.89, 69.17, 68.45 and 67.86 (C-2, C-2', C-3, C-3', C-4, C-4', C-5, C-5'), 62.40 and 61.34 (C-6 and C-6'), 20.70, 20.62, 20.50 and 20.41 ( $\text{COCH}_3$ ).

*Deprotection step:* The resulted 2 g of maltose octaacetate (**2**) (2.95 mmol) was dissolved in 30 mL anhydrous *N,N*-dimethylformamide, the solution was cooled to 0°C and 494 mg hydrazine-acetate (5.34 mmol) was added to the mixture. The reaction solution was stirred at room temperature for 3 h and concentrated in a rotary vacuum evaporator. The residue was diluted in 2 to 1 ratio with methylene chloride and washed twice with dH<sub>2</sub>O. The combined organic phases were dried over MgSO<sub>4</sub>, filtered, concentrated in a rotary vacuum evaporator, and then purified by Kieselgel 60 column chromatography using 3:2 hexanes-acetone eluent to yield 1.56 g of 2,3,4,6-tetra-*O*-acetyl- $\alpha$ -D-glucopyranosyl-(1 $\rightarrow$ 4)-2,3,6-tri-*O*-acetyl- $\alpha,\beta$ -D-glucopyranose (83%) anomers (3 $\alpha$ :3 $\beta$  ~ 1:3, NMR).

NMR data for compound **3 $\beta$** : <sup>1</sup>H NMR (CDCl<sub>3</sub>):  $\delta$  5.58 (t, *J* = 9.5 Hz, 1H), 5.43 (d, *J* = 4 Hz, 1H, H-1'), 5.41-5.32 (m, 2H), 5.08 (d, *J* = 10 Hz, 1H, H-1), 4.86 (dd, *J* = 10, 4 Hz, 1H, H-2'), 4.76 (dd, *J* = 10, 3.5 Hz, 1H, H-6a'), 4.52-4.46 (m, 1H, H-6b'), 4.30-4.17 (m, 3H), 4.10-3.91 (m, 3H), 2.14, 2.10, 2.06, 2.05, 2.02, 2.01 and 2.00 (7s, 3H, 7 COCH<sub>3</sub>); <sup>13</sup>C NMR (CDCl<sub>3</sub>):  $\delta$  170.64, 170.26, 169.94 and 169.44 (COCH<sub>3</sub>), 95.43 and 89.87 (C-1 and C-1'), 72.57, 72.26, 71.54, 69.95, 69.31, 68.33, 67.93 and 67.58 (C-2, C-2', C-3, C-3', C-4, C-4', C-5, C-5'), 62.75 and 61.36 (C-6 and C-6'), 20.89, 20.79, 20.61 and 20.53 (COCH<sub>3</sub>). MALDI TOF-MS calc. for C<sub>26</sub>H<sub>36</sub>O<sub>18</sub>: 636.19 [M]. Found: 659.22 [M+Na]<sup>+</sup>.

*Spacer addition step:* The methylene chloride (50 mL) solution of the resulted 1.27 g hemiacetal (**3**) (2 mmol) was mixed with 5.83 mL trichloro-acetonitrile (57.4 mmol). 2 g potassium carbonate was added and the mixture was stirred overnight at room temperature, filtered through a Celite pad and concentrated in a rotary vacuum evaporator. The resulted imidate (**4**) and 496 mg 7-(1,3-dioxan-2-yl)-heptan-1-ol (**5**) (2.42 mmol) was dissolved in 30 mL methylene chloride and 5 g of 4 Å molecular sieving beads were added to the mixture. The reaction solution was dried at room temperature for 1 h under Ar atmosphere, cooled to -40°C followed by dropwise addition of 95  $\mu$ L trimethylsilyl trifluoromethanesulfonate (0.52 mmol), stirred at room temperature for 1 h, neutralized with triethylamine, diluted with methylene chloride and washed twice with saturated aqueous NaHCO<sub>3</sub>. The combined organic phases were dried over MgSO<sub>4</sub>, filtered, concentrated in a rotary vacuum evaporator and the residue was purified by column chromatography using Kieselgel 60 with 3:2 hexanes-acetone eluent to yield 595 mg (36% in two steps) of [7-(1,3-dioxan-2-yl)-heptyl]-2,3,4,6-tetra-*O*-acetyl- $\alpha$ -D-glucopyranosyl-(1 $\rightarrow$ 4)-2,3,6-tri-*O*-acetyl- $\beta$ -D-glucopyranose (**6**) compound.

NMR data for compound **6**: <sup>1</sup>H NMR (CDCl<sub>3</sub>):  $\delta$  5.41 (d, *J* = 4 Hz, 1H), 5.36 (t, *J* = 10 Hz, 1H), 5.25 (t, *J* = 9 Hz, 1H), 5.05 (t, *J* = 10 Hz, 1H), 4.90-4.78 (m, 2H), 4.57-4.43 (m, 3H), 4.31-4.21 (m, 2H),

4.15-3.93 (m, 5H), 3.88-3.63 (m, 4H), 3.54-3.39 (m, 1H), 2.14, 2.10, 2.05, 2.03, 2.01 and 2.00 (5s, 21H, COCH<sub>3</sub>), 1.65-1.47 (m, 5H), 1.47-1.21 (m, 11H); <sup>13</sup>C NMR (CDCl<sub>3</sub>): δ 170.50, 170.23, 169.92, 169.57 and 169.39 (COCH<sub>3</sub>), 102.34 (acetal C), 100.25 and 95.48 (C-1 and C-1'), 75.45, 72.78, 72.21, 72.03, 69.97, 69.34, 68.44 and 68.02 (C-2, C-2', C-3, C-3', C-4, C-4', C-5, C-5'), 70.13 and 66.86 (CH<sub>2</sub>), 62.90 and 61.50 (C-6 and C-6'), 35.16, 29.36, 29.11, 25.83, 25.63 and 23.84 (CH<sub>2</sub>), 20.89, 20.82 and 20.58 (COCH<sub>3</sub>). MALDI TOF-MS calc. for C<sub>37</sub>H<sub>56</sub>O<sub>20</sub>: 820.34 [M]. Found: 843.42 [M+Na]<sup>+</sup>.

Then the protected maltose with the 387 mg aglycone spacer (**6**) (0.47 mmol) was dissolved in 20 mL methanol, and then sodium methylate catalyst was added (pH~8-9) in the catalytic amount.

The mixture was stirred overnight and quenched with Amberlite IR 120 hydrogen ion exchange resin until pH 7.0 was reached. Ultimately, the residue was filtered, concentrated in a rotary vacuum evaporator, and the mixture was purified by Kieselgel 60 column chromatography using 4:1 methylene chloride-methanol eluent to yield 178 mg (72%) of [7-(1,3-dioxan-2-yl)-heptyl]-α-D-glucopyranosyl-(1→4)-β-D-glucopyranose (**7**) compound.

NMR data for compound **7**: <sup>1</sup>H NMR (CD<sub>3</sub>OD): δ 5.15 (d, *J* = 4 Hz, 1H, H-1), 4.53 (t, *J* = 5 Hz, 1H), 4.26 (d, *J* = 8 Hz, 1H, H-1'), 4.07-4.00 (m, 2H), 3.94-3.73 (m, 6H), 3.72-3.48 (m, 6H), 3.43 (dd, *J* = 10, 4 Hz, 1H), 3.39-3.16 (m, 6H), 2.08-1.89 (m, 1H), 1.67-1.47 (m, 4H), 1.43-1.26 (m, 9H); <sup>13</sup>C NMR (CD<sub>3</sub>OD): δ 104.35 (acetal C), 103.69 and 102.95 (C-1 and C-1'), 81.37, 77.90, 76.62, 75.12, 74.81, 74.74, 74.21 and 71.24 (C-2, C-2', C-3, C-3', C-4, C-4', C-5, C-5'), 70.94 and 67.95 (CH<sub>2</sub>), 62.80 and 62.24 (C-6 and C-6'), 36.21, 30.78, 30.54, 30.48, 27.01 and 24.95 (CH<sub>2</sub>). MALDI TOF-MS calc. for C<sub>23</sub>H<sub>42</sub>O<sub>13</sub>: 526.26 [M]. Found: 549.24 [M+Na]<sup>+</sup>.

*Conjugation step:* After removing the protecting acetyl groups, the 13.15 mg compound (**7**) (0.025 mmol) was dissolved in 96% acetic acid (1.18 mL) and stirred for 2 h at 50°C, and then concentrated in a rotary vacuum evaporator. After heating, the residue was dissolved in 2 mL phosphate buffer (0.1 M, pH=7.0) and 16.75 mg (250 nmol) bovine serum albumin (BSA) was added to the mixture. Then 3.17 mg (0.050 mmol) sodium cyanoborohydride (NaCNBH<sub>3</sub>) was added and allowed to stand at room temperature for 1 day. The resulting neoglycoprotein (carbohydrate antigen) was dialyzed against water and lyophilized (Heraeus-Christ, Osterode, Germany) to yield 4.5 mg of neoglycoprotein (**9a**). The synthesis of other neoglycoproteins was performed similar as described above (**9b**, **9c** and **9d**). The molar ratio of compound **7** with aglycone spacer and the reagents to BSA was varied in the coupling reaction as shown in Table 2.



**Table 2.** Analysis of the number of conjugated sugar residues in the synthesized antigens by MALDI-TOF MS and parameters of the conjugation step.

	NEOGLYCOPROTEINS			
	BSA-32- maltose (9a)	BSA-50- maltose (9b)	BSA-59- maltose (9c)	BSA-66- maltose (9d)
<b>Maltose containing spacer (mg)</b>	13.15	26.3	39.5	78.9
<b>Maltose containing spacer (mmol)</b>	0.025	0.05	0.075	0.15
<b>Acetic acid (mL)</b>	1.18	2.36	3.54	7
<b>Water (mL)</b>	0.3	0.6	0.9	1.8
<b>Phosphate buffer (mL)</b>	2	2	2	2
<b>BSA (mg)</b>	16.75	16.75	16.75	16.75
<b>BSA (nmol)</b>	250	250	250	250
<b>NaCNBH<sub>3</sub> (mg)</b>	3.17	6.35	9.51	19
<b>NaCNBH<sub>3</sub> (mmol)</b>	0.050	0.1	0.15	0.3
<b>Amount of antigens (mg)</b>	4.5	19.2	9.9	18.6
<b>MALDI-TOF MS (Da)<sup>a)</sup></b>	80 851.1	88 854.5	93 171.6	96 258.5
<b>Carbohydrate units</b>	<b>32</b>	<b>50</b>	<b>59</b>	<b>66</b>

a) Sugar-spacer (C<sub>20</sub>H<sub>35</sub>O<sub>11</sub>):451 Da; BSA: 66 401 Da; Calculation: (80 851 – 66 401)/451 = 32 units.

#### 4.3.3. Analysis of Neoglycoproteins by CGE

For fluorophore labeling of neoglycoproteins, the fluorescent reagent Chromeo P503 (Active Motif, Rixensart, Belgium) was diluted with methanol to 1 mg/mL final concentration and the stock solution was kept at 4°C until use. Samples were diluted in MilliQ-grade water (Millipore) in a final concentration of 2 mg/mL. For the SDS-CGE separation, a total of 5 µL of each glycoconjugate was transferred to 200 µL PCR tubes followed by the addition of 0.5 µL of 1 mg/mL dye stock solution, 2.5 µL of 10% SDS, 2 µL of 0.5 M dithiothreitol (DTT) reducing agent and 25 µL of 25 mM NaHCO<sub>3</sub> buffer (final concentration: 17.86 mM, pH 9.20) to a total volume of 35 µL and incubated at 90°C for 30 min. Prior to the SDS-CGE separation, the glycoconjugates were allowed to cool to ambient temperature and centrifuged at 3000 × g for 30 s (Centrifuge 5424, Eppendorf, Hamburg, Germany).

In all electrophoretic analysis, 60 mM tris–borate containing 0.1% SDS buffer (pH 8.45) was used as background electrolyte (BGE). The sample buffer consisted of 25 mM NaHCO<sub>3</sub> and the pH was adjusted to 9.2 by the addition of 0.1 M NaOH. All separations were carried out in a P/ACE MDQ instrument (SCIEX, Brea, CA, USA), which equipped with a 488 nm Ar-ion laser having a strong 514 nm line and the emission signal was collected through a 600 nm long pass filter. Before the first use, new capillary (75 µm i.d., 360 µm o.d., 10 cm effective length and 30 cm total length) was washed with MilliQ-grade water, then 0.1 M HCl, and rinsed with MilliQ-grade water before the SDS-MW gel buffer (SCIEX) was transferred into the capillary. Prior to each sample injection, the gel-buffer was replaced in the capillary tubing. The sample injection process was preceded by the introduction of an HPLC-grade water plug (12.5 kV for 5 s). The glycoconjugates were introduced electrokinetically at the short side of the capillary (10 cm) from a 96-well plate (25 kV for 10 s). All separations were carried out at 25°C by applying 500 V/cm electric field strength. Data analysis was performed using the Karat 32 version 7.0 software package (SCIEX).

#### **4.3.4. Analysis of Neoglycoproteins by MALDI-TOF MS**

MALDI-TOF MS analysis of the synthesized glycoconjugations was performed in positive reflectron mode using a BIFLEX III mass spectrometer (Bruker) with delayed-ion extraction. Spectra from multiple (>100) laser shots (N<sub>2</sub> laser, 337 nm) using 19 kV accelerating and 20 kV reflectron voltage were summarized. BSA was used (6 – 8 mg/mL in 0.1% TFA) as external calibration solution. TFA solution was prepared by dissolving 0.1% TFA in a mixture of 2:1 acetonitrile–water. A total of 10 µL of sample, 25 µL matrix (3,5-dimethoxy-4-hydroxycinnamic acid), and 15 µL TFA solution were mixed and 0.5 µL was applied to the target plate and allowed to dry at room temperature before analysis.

#### **4.3.5. Immunization with Neoglycoproteins**

For each immunization step (BALB/c (♀): inbred white albino mouse, Charles River Hungary, Isaszeg, Hungary), both the lowest (32 units/mol) and highest (66 units/mol) amount of sugar-incorporated synthetic glycoconjugates (50 µg/mL in phosphate buffered saline, pH 7.4) were mixed with equal volume of Freund's complete adjuvant and incomplete Freund's adjuvant for the first and consecutive immunization injections, respectively. After that, the mixture was injected subcutaneously into the hind legs of the mice three times in two week intervals.

Seven to ten days after the third injection, blood was taken from the eye corners of the mice. The serum was separated from the red blood cells and tested by enzyme-linked immunosorbent assay (ELISA).

#### **4.3.6. Analysis of Neoglycoproteins by ELISA**

Immunoassay plates (Corning 96-well plates, half-area clear polystyrene, high-binding, nonsterile, Sigma-Aldrich) were coated with synthetic immunogens, i.e. 30  $\mu$ L per well in the coating buffer (14.7 mM  $\text{Na}_2\text{CO}_3$ , 34.9 mM  $\text{NaHCO}_3$ , pH 9.6), and then incubated at 37°C for 1 h. To detect the antibody response, a total of 10  $\mu$ g/mL and a twofold serial dilution from 1.25 to 0.08  $\mu$ g/mL of immunogen coatings were used, as well as 0.2  $\mu$ g/mL BSA and synthetic neoglycoprotein as immunogen coatings were prepared for the experiments, which were done in the presence of inhibitors. The plates were washed twice with PBS-Tween wash buffer (123.2 mM NaCl, 3.2 mM  $\text{KH}_2\text{PO}_4$ , 10.5 mM  $\text{Na}_2\text{HPO}_4$ , 0.05 v/v% Tween-20, pH 7.2), the remaining protein-binding sites were blocked by the addition of 60  $\mu$ L PVP blocking buffer (376.5 mM NaCl, 0.5 m/v% polyvinyl-pyrrolidone (Mw 10 000) in the PBS-Tween buffer) and incubated at 37°C for 30 min. After that, the plates were washed twice with wash buffer and 30  $\mu$ L of appropriately diluted immune sera was added to each well followed by incubation at 37°C for 1 h. To check the polyclonal antibody response, serum dilutions of 1 000 – 128 000 $\times$  were used, and then the plates were washed four times with the wash buffer. This was followed by washing the plates four times with wash buffer before 30  $\mu$ L of appropriately (8000-fold) diluted horseradish peroxidase (HRP-labeled) conjugated secondary antibody (goat-anti-mouse IgG, Southern Biotechnology, Birmingham, AL, USA) was added to each well and incubated at 37°C for 30 min. To measure the bound antibody, 30  $\mu$ L of 3,3',5,5'-tetramethylbenzidine substrate solution was added to each well. This solution comprised 3,3',5,5'-tetramethylbenzidine tablet (2 mg) dissolved in the mixture of 1 mL DMSO and 9 mL citrate-phosphate buffer (52.0 mM  $\text{Na}_2\text{HPO}_4$ , 25.5 mM citric acid, pH 5.0). A total of 2  $\mu$ L of fresh 30% hydrogen peroxide was added per 10 mL of substrate buffer solution immediately prior to use. The estimated incubation times for the enzyme-substrate reaction ranged from 10 to 15 min, after which 30  $\mu$ L stop solution (4 M  $\text{H}_2\text{SO}_4$ ) was added to the reaction and the plates were interrogated at 450 nm in a Multiscan Ascant reader (Thermo Fisher Scientific, Hudson, NH, USA). In the inhibition studies, 0.1 – 1.6 g/mL BSA, neoglycoproteins, and 2.00 – 0.06 mg/mL maltose, glucose, isomaltose, lactose, galactose, and maltodextrin were used to inhibit the binding of specific antibodies from the suitably diluted (2 000-fold) plasma preparations.

4 000-fold dilutions of sera were prepared without inhibitors as control. After coating and blocking, the serum was mixed with an equal volume of the inhibitor. The mixture was placed on the well and incubated for 1 h at 37°C. This was followed by washing the plates four times with wash buffer before 30 µL of appropriately (8 000-fold) diluted horseradish peroxidase (HRP labeled) conjugated secondary antibody (goat-anti-mouse IgG, Southern Biotechnology) was added to each well and incubated at 37°C for 30 min. Then the plates were washed four times with the wash buffer. Antibody binding was measured as described above.

## **4.4. High-Performance CGE Analysis of Complex Carbohydrates**

### **4.4.1. Sample Preparation for CGE**

Glycoprotein digestions, i.e. 50 µg of hIgG and bovine RNase B (Sigma-Aldrich), were performed using the GlykoPrep Digestion Module according to the manufacturer's protocol (Prozyme, Hayward, CA, USA). The released *N*-glycans were dried in a centrifugal vacuum evaporator (Thermo Fisher Scientific, Asheville, USA). The maltodextrin ladder (SCIEX), dried sugars, D-(+)-maltose monohydrate (Sigma-Aldrich), and carbohydrate standards (0.2 µg of each *N*-glycan standard) were labeled with ANTS-fluorescent dye (Prozyme) as described in [52]. Briefly, the dried sugars were fluorescently labeled via reductive amination by the addition of 6.0 µL of 0.1 M ANTS ( $\lambda_{\text{ex}}$  356 nm and  $\lambda_{\text{em}}$  512 nm) in 15% v/v acetic acid and 2.0 µL of 1 M sodium cyanoborohydride (in THF). The reaction mixture was incubated at 37°C overnight. After that, the unreacted fluorescent dye was subsequently removed by Cleanup cartridges according to the protocol of the manufacturer (Prozyme). Briefly, the binding step was performed in 200 µL 1× cleanup solution to each glycan sample followed by washing off the unreacted dye from the cartridges, and finally the elution of glycans by water. In CGE separation, ANTS-labeled-maltose (ANTS-G2) was used as lower bracketing standard and the upper bracketing standard was prepared as follows: maltose (G2) was fluorescently labeled via reductive amination by the addition of 6.0 µL of 0.1M 5-amino-2-naphtalenesulfonic acid (5-ANSA, Sigma-Aldrich) in 15% v/v acetic acid and 2.0 µL of 1 M sodium cyanoborohydride (in THF) at 37°C overnight. The upper bracketing standard was chosen based on the electrophoretic mobility of the 5-ANSA-maltose conjugate [149].

#### **4.4.2. *N*-Glycan Separation by CGE**

Rapid CGE separations of the ANTS-labeled *N*-glycans were performed by a single-channel Glycan Analyzer (GL1000, BiOptic) instrument containing a 50  $\mu\text{m}$  i.d. (360  $\mu\text{m}$  o.d.) N-CHO capillary with an effective separation length of 11.5 cm (total length: 15.5 cm). To match the ANTS excitation/emission spectra, the instrument contained an LED source of 350 – 390 nm range (Cree XLamp). The glycan separation gel-buffer and running-buffer (BiOptic) was used for CGE-LEDIF separations. Prior to each sample injection, the gel-buffer was replaced in the separation capillary by means of a 30 s purge step, followed by injection of the mixture of the lower (ANTS-G2) and upper (ANSA-G2) bracketing standards by applying 2 kV for 3 s. For migration time correction, the bracketing standards were used as mentioned earlier in section 2.9.1. To avoid any cross-contamination, the inlet end of the capillary tubing was immersed into MilliQ-grade water (0 kV for 1 s) as a washing step. This cleaning step was carried out after injecting each of the *N*-glycans. After that, the samples were introduced electrokinetically from a 96-well plate at 2 kV for 3 s. All CGE separations were accomplished at ambient temperature by applying 6 kV electric potential with a high-voltage power supply (EMCO). Data analysis was also performed using the Q-Analyzer software package (BiOptic). All samples were run six times and the GU values were calculated from the average. Glycan nomenclature and symbolic representations are as previously suggested by Harvey *et al.* [150].

### **4.5. Study of the Activation Energy by CGE**

#### **4.5.1. *Sample Preparation for CGE***

For the release of the *N*-linked glycans, 100  $\mu\text{g}$  of hIgG (Sigma-Aldrich, stock solution: 10 mg/mL) was dissolved in 10  $\mu\text{L}$  of 20 mM sodium-bicarbonate solution (pH 7.0), reduced with 1  $\mu\text{L}$  of 50 mM DTT for 15 min at 65°C and then alkylated with 1  $\mu\text{L}$  of 50 mM iodoacetamide (IAA) for 30 min at 37°C in dark. After incubation, 1  $\mu\text{L}$  of 1 U/ $\mu\text{L}$  PNGase F (Prozyme) was added to the solution and incubated overnight at 37°C. The released *N*-linked glycans were partitioned with CleanSeq magnetic beads (SCIEX) as described in Ref. [151]. The maltooligosaccharide ladder (SCIEX), the maltose internal standard (Sigma-Aldrich) and the purified sugars were fluorescently labeled via reductive amination by the addition of 6.0  $\mu\text{L}$  of 20 mM APTS (SCIEX) in 15% v/v acetic acid and 2.0  $\mu\text{L}$  of 1 mM cyanoborohydride (in THF) and incubated for 1 h at 50°C. After the labeling reaction, the unreacted fluorescent dye was

removed with the CleanSeq magnetic beads as above. Cleanup procedure was not necessary, because the maltooligosaccharide was in 3x excess to the fluorescent dye. The samples were analyzed directly after solubilization or kept at -20°C until use. Viscosity measurements of the buffer systems, i.e. 25 mM lithium acetate buffer containing i) no additives, ii) 10 – 60% ethylene glycol iii) 0.5 – 3.0% linear polyacrylamide iv) 0.4% 300kDa polyethylene oxide, were made at 20, 30, 40 and 50°C on an AR 550 model rheometer (TA Instruments, New Castle, Germany) equipped with a 60 mm diameter, 2° angle steel cone geometry and a built-in Peltier heating/cooling plate.

#### ***4.5.2. CGE-based Separation of Linear and Branched Carbohydrates***

Rapid CGE separations of the APTS-labeled oligosaccharides were performed in a PA 800 Plus Pharmaceutical Analysis System (SCIEX, Brea, CA, USA) containing a 50 µm i.d. N-CHO capillary (360 µm o.d.) with an effective separation length of 21 cm (total length: 31 cm). The CGE system was equipped with a solid state laser based fluorescent detector ( $\lambda_{\text{ex}}$  488 nm and  $\lambda_{\text{em}}$  520 nm). To alleviate for any variations caused by temperature changes, the temperature of the gel-filled separation capillary was controlled by the liquid cooling system of the CE instrument ( $\pm 0.1^\circ\text{C}$ ). All samples were analyzed at 20, 30, 40 and 50°C using the corresponding buffer systems. The background electrolyte was 25 mM lithium acetate buffer (Sigma-Aldrich, pH 4.75) containing: i) no additives, ii) ethylene glycol (Sigma-Aldrich, 10 – 60% v/v), iii) linear polyacrylamide (Polysciences, Warrington, PA, USA, 10 kDa, 0.5 – 3.0% v/v) and iv) polyethylene oxide (Sigma-Aldrich, 300 kDa 0.4% v/v). Prior to use, the separation capillary was conditioned by subsequent rinsing with MilliQ-grade water and the BGE (both with and without modifiers/additives) for 10 min at 30 psi. The capillaries were also rinsed with the respective BGE for 5 min at 30 psi between injections to avoid any possible carryover issues. Samples were pressure injected (0.5 psi for 5.0 s). The applied electric field strengths were 400 V/cm, with the cathode at the injection side and the anode at the detection side (reversed polarity). For migration time correction and normalization, APTS-labeled maltose (APTS-G2) was co-injected with each sample. All sample runs were done in triplicates. Data analysis was performed using the 32 Karat, ver 9.1 software package (SCIEX). GU values were calculated according to Eq. 19 as described above, using the 5<sup>th</sup> order polynomial approach [120]. Hydrodynamic volumes of relevant (DP 7-12) and major IgG N-glycans (FA2, FA2[6]G1, FA2[3]G1, and FA2G2) were calculated using TURBOMOLE 6.3 quantum chemical program package and COSMOtherm suite (COSMOlogic GmbH, Leverkusen, Germany).

## 5. RESULTS

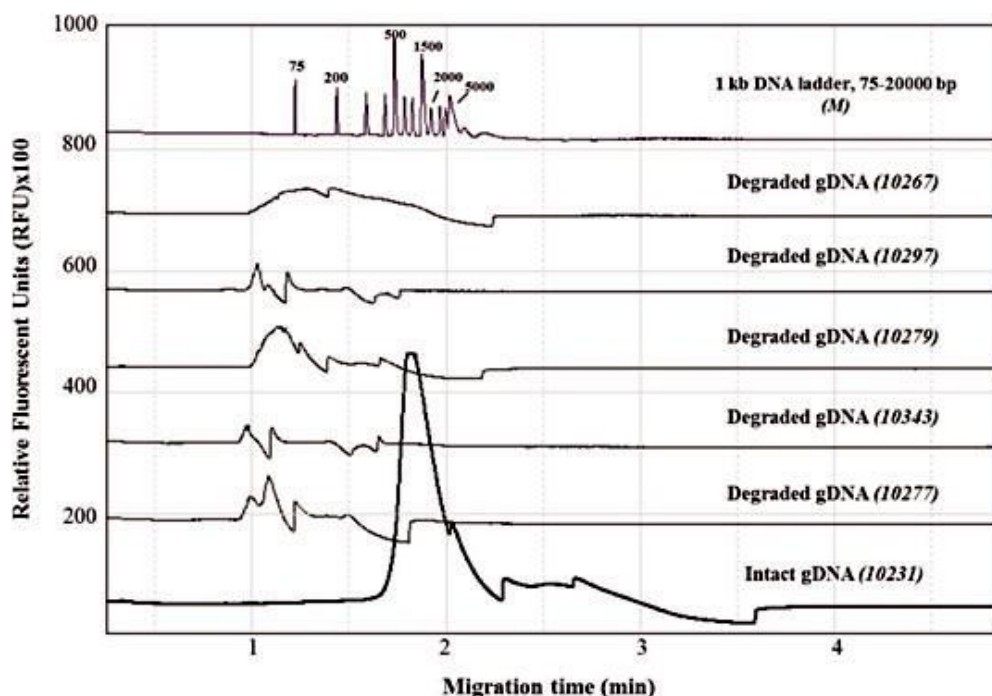
Here, this section starts with the introduction of a novel single-channel capillary gel electrophoresis in order to evaluate the suitability of this separation system. Therefore, the degradation level of a large number (about 1000) gDNA samples were analyzed focusing on migration time reproducibility, limit of detection and detector linearity in comparison to conventional technique. DNA analysis was chosen at the beginning, because this system was first commercially available in the market. It followed by the development of a haplotyping method in WFS1 gene by allele-specific amplification and rapid capillary gel electrophoresis with the same CGE system in order to analyze the resulted DNA fragments. After that the synthesis of neoglycoproteins was introduced as carbohydrate antigens to elicit specific immune response, which was analyzed then by CGE and MALDI-TOF MS. In the next part of this study, the establishment of a novel ANTS-labeled database was generated for CGE analysis of complex *N*-linked carbohydrates. The validation of the generated GU values in the database was verified with *N*-glycans released from human IgG and bovine pancreatic RNase B. Finally, the activation energy concept related to the electromigration of complex *N*-linked carbohydrates was analyzed by capillary electrophoresis in order to determine the electrophoretic mobility shift between the different structure oligosaccharides at different temperature/background electrolyte. The results of this study are discussed in more detail in the following sections.

### 5.1. Large-Scale gDNA Analysis by CGE

In the first part of this study, an efficient micoball ended fiber optic based CGE-LEDIF detection system was introduced for large-scale gDNA analysis, also utilizing a single-channel separation capillary. To assess the suitability of this novel system, the degradation level of a close to a thousand gDNA samples were analyzed. The capability to quickly distinguish between intact and degraded gDNA samples was a crucial step for downstream quantitative PCR fragment analysis.

Figure 9 compares the results of the analysis of representative intact and variously degraded gDNA samples. For fragment size determination, a DNA molecular mass ladder was used in the base pair range of 75 – 20 000 bp (upper trace). Non-degraded intact gDNA sample showed a large peak pattern at the 2 000 base pair range of the separation area (lower trace: intact gDNA, sample ID: 10231).

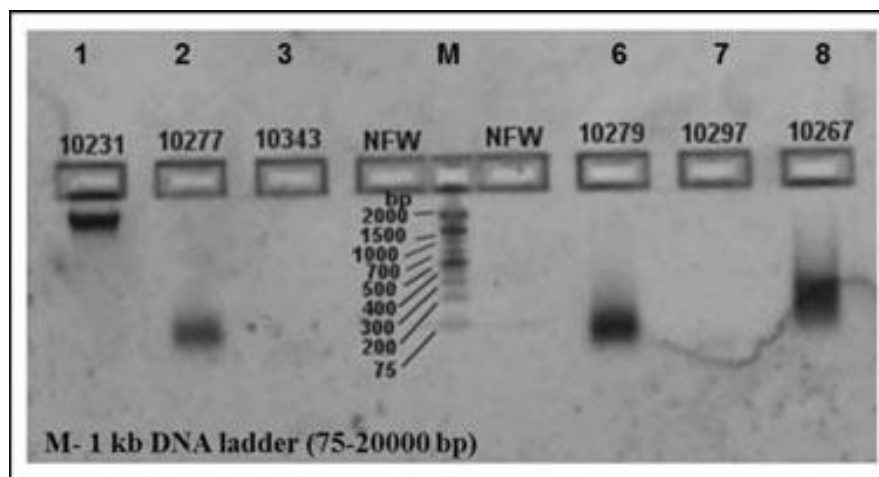
The variously degraded gDNA samples featured patterns at the lower base pair range of several hundreds, i.e. strongly suggesting sample degradation (middle traces: degraded gDNAs, sample IDs: 10267, 10297, 10279, 10343, 10277).



**Figure 9.** Typical CGE-LEDIF analysis of selected gDNA samples. Upper trace: DNA sizing ladder (75 bases to 20 000 bases), middle traces: degraded gDNA samples (sample IDs 10267, 10297, 10279, 10343 and 10277), lower trace: non-degraded gDNA sample (sample ID 10231). Separation conditions: capillary: 75  $\mu\text{m}$  i.d., effective separation length of 11 cm (total separation length of 15 cm); sample injection: 4 kV/6 s; separation voltage 8 kV/300 s; ambient temperature. Adapted from [72].

Figure 10 shows the conventional agarose slab gel electrophoresis based separation of the same gDNA samples for comparative purposes. Please note that compared to CGE analysis, several degraded DNA bands were not visible on 2% agarose slab gel (sample IDs: 10343, 10297), while all degraded DNA peak patterns were clearly detected by CGE.





**Figure 10.** Agarose slab gel electrophoresis based analysis of the same gDNA samples as in Figure 9. Lanes 1: non-degraded gDNA sample (sample ID 10231); Lanes 2– 8: degraded gDNA samples (sample IDs 10267, 10297, 10279, 10343, and 10277); M: DNA mass ladder (75 bases to 20 kilobases). Separation conditions: samples were separated on the E-Gel system power source containing 2% agarose gel ( $2 \times 8$  wells) for 8 min; injection: 20  $\mu$ L per well containing 500 ng gDNA. Adapted from [72].

The migration time reproducibility of this CGE system was determined by a 10-run repeatability study. For this evaluation, a 100 bp Plus DNA mass ladder was used in the base pair range of 100 – 3 000 bp, which was diluted in a final concentration of 25 ng/ $\mu$ L. Each sample was injected from the same well of the 96-well plate. Table 3 summarizes the CGE-LEDIF results, showing that the average migration times of the DNA mass ladder were between 88.22 and 143.22 s in the base pair range of 100 – 3 000 bp. The average percent RSD (RSD%) of the migration times were between 0.5118 – 0.7477%, which exhibited an excellent reproducibility.

**Table 3.** Ten-run migration time reproducibility study using the 100 – 3 000 bp DNA mass ladder. Adapted from [72].

sec \ bp	100	200	300	400	500	600	700	800	900	1000	1200	1500	2000	3000
First run (s)	87.6	101.8	114.3	122.2	125.9	128.3	130.0	130.8	131.0	132.2	134.8	137.6	140.2	143.8
Second run (s)	88.6	102.5	114.8	122.7	126.4	128.8	130.4	131.2	131.9	132.7	135.2	137.9	140.4	144.0
Third run (s)	89.5	102.5	115.0	123.1	125.0	127.5	132.2	130.0	131.6	132.5	135.0	137.8	140.3	144.0
Fourth run (s)	88.9	102.9	115.3	123.4	127.2	129.7	130.3	132.2	132.9	133.6	136.1	138.9	141.3	143.0
Fifth run (s)	88.5	102.3	114.7	122.6	126.4	128.8	130.4	131.2	131.4	132.6	135.2	137.9	140.4	144.0
Sixth run (s)	88.2	101.9	114.2	122.1	125.8	128.2	129.9	130.6	131.7	132.0	134.5	137.2	139.7	143.3
Seventh run (s)	88.0	101.7	114.0	121.8	125.5	127.9	129.6	130.3	130.4	131.7	134.2	136.9	139.4	142.9
Eighth Run (s)	87.9	101.7	114.0	121.8	125.5	127.9	129.5	130.3	131.0	131.7	134.2	136.9	139.3	142.9
Ninth run (s)	87.6	101.3	113.5	121.3	125.0	127.5	129.0	129.8	130.4	131.2	133.7	136.4	138.8	142.4
Tenth Run (s)	87.4	101.0	113.2	121.0	124.7	127.1	128.7	129.4	130.0	131.8	133.3	135.9	138.3	141.9
tM (s)	88.2	102.0	114.3	122.2	125.7	128.2	130.0	130.6	131.2	132.2	134.6	137.3	139.8	143.2
SD	0.7	0.6	0.7	0.8	0.8	0.8	1.0	0.8	0.9	0.7	0.8	0.9	0.9	0.7
RSD (%)	0.7	0.6	0.6	0.6	0.6	0.6	0.7	0.6	0.6	0.5	0.6	0.6	0.6	0.5

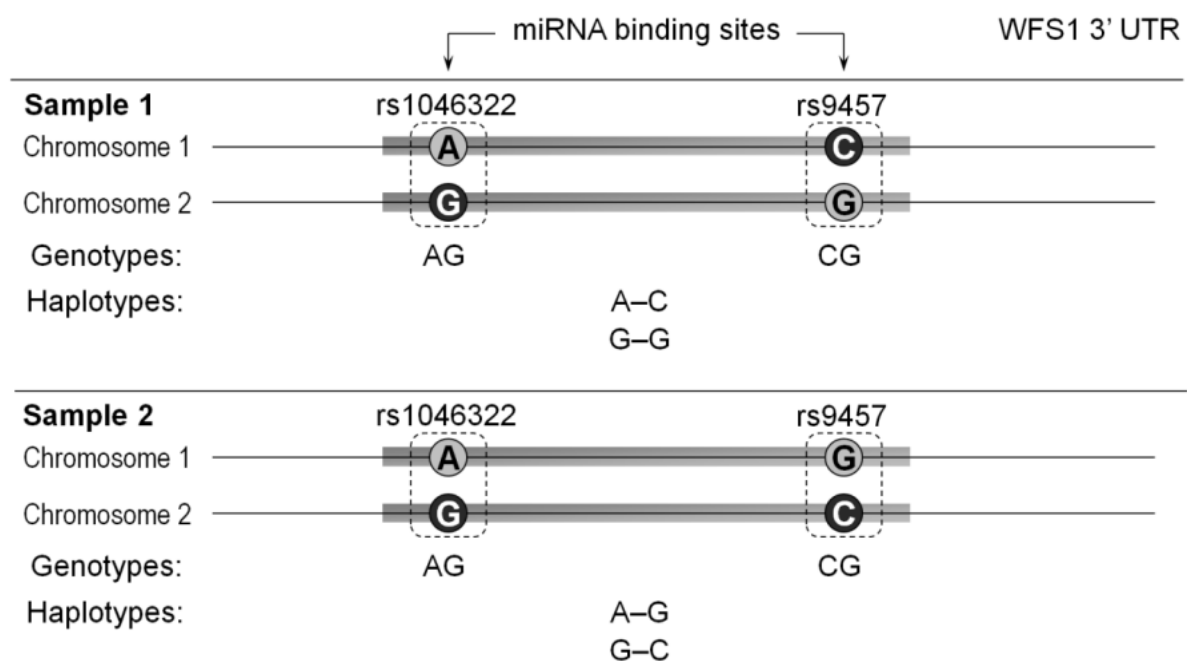
## 5.2. Molecular Haplotype Analysis

### 5.2.1. Allele-Specific PCR

In the second part of this study, simultaneous multiple genotyping of two adjacent putative miRNA-binding SNPs in the WFS1 gene was introduced by an efficient double-tube allele-specific amplification method in conjunction with a novel CGE-LEDIF detection system to investigate the resulting DNA fragments.

Wolframin (WFS1) is a transmembrane protein in the endoplasmic reticulum (ER), which is generated at higher levels in pancreatic beta-cells and specific neurons in the central nervous system [152]. It plays a crucial role in the ER calcium homeostasis [152-154] and in the ER stress response [155]. Riggs *et al.* reported on the inactivation of the Wolframin gene (WFS1) in the beta-cells of rodents, which caused ER stress and death of the beta-cells by accelerated apoptosis [156]. Mutations in the WFS1 gene are causing the so-called Wolfram syndrome [157] that comprises young onset non-autoimmune insulin dependent diabetes mellitus, diabetes insipidus, optic atrophy, deafness or other neurological and endocrine abnormalities [158, 159]. In addition, SNPs in the WFS1 gene has recently been presented to be associated with diabetes mellitus type 2 in populations of European descents [160-162]. Therefore, simultaneous study of haplotyping these abnormalities is getting more and more attention to investigate the genetic background of complex diseases [163].

Direct haplotype determination by allele-specific amplification (ASA), i.e. molecular haplotyping, offers fast and reliable genotyping data of many SNP variations in a single tube multiplex polymerase chain reaction (mPCR) followed by electrophoresis analysis [79, 164]. In this molecular haplotype study, the *rs9457* and *rs1046322* SNPs are presumably miR-SNPs, which are located in the WFS1 gene 3' UTR. In addition, their *in silico* data analysis suggested that they may alter the binding of miR-185 and miR-668. The haplotype determination of double heterozygote samples (*rs1046322AG* and *rs9457CG*) is significant, since otherwise it is uncertain if the two allelic variants possibly affecting miRNA-binding are located on the same mRNA ("cis") or can be found on two different chromosomes ("trans") as illustrated in Figure 11.



**Figure 11.** Schematic illustration of the putative effect of the two SNPs on miRNA binding. Genotypes and haplotypes can be specified by allele-specific amplification using sense *rs1046322*- and antisense *rs9457*-specific primers in different combination in case of double heterozygote samples (samples 1 and 2). The four thick gray lines mark the four haplotypes on the same chromosomes: A–C, G–G, A–G and G–C. Adapted from [2].

An allele-specific amplification based approach was introduced for the haplotype determination of the two SNPs of interest, which was comprised simultaneously two outer and two allele-specific primers in a multiplex reaction as shown in Figure 12. These allele-specific primers were designed to anneal to the SNP by their 3' end. For the chromosomal localization, a sense *rs1046322*- and an antisense *rs9457*-specific primer were used in the amplification reaction.

One reaction mixture examined the presence of one allele at each loci as well as one haplotype combination, consequently two reaction mixtures were necessary to determine the genotype and the haplotype (Figure 12).

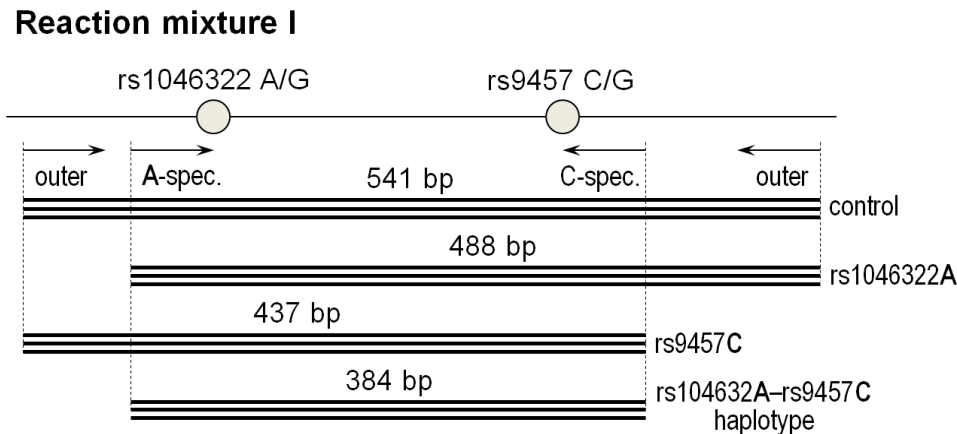
Figure 12A illustrates the analysis of the reaction mixture-I using the sense *rs1046322A*- and the antisense *rs9457C*-specific primers in combination with the outer oligos. A 488-bp-long fragment was generated by the *rs1046322A*-specific and the antisense outer primers in case of the presence of an A allele at the *rs1046322* site. Similarly, if the sample possessed the C allele at the *rs9457* locus, the primer specific for this variant together with the sense outer primer could amplify a 437-bp-long fragment. More importantly, a 384-bp-long PCR fragment could be produced, if the *rs1046322A* and *rs9457C* alleles are located on the same chromosome, which is generated by the two allele specific primers and suggested the presence of the A–C haplotype. The longest, 541-bp outer product is a control PCR fragment synthesized independently of the genotype and haplotype of the sample of interest.

Figure 12B depicts the analysis of the reaction mixture-II using the *rs1046322G*- and *rs9457G*-specific primers in conjugation with the outer oligos. If the *rs1046322G* allele was present, a 488-bp-long product could be observed, whereas a 437-bp-long product produced in case of the presence of the *rs9457G* allele and the 384-bp-long product suggested the G–G haplotype. Thus, genotype and haplotype information could be unambiguously determined by these two amplification reactions. In addition, two redundant combinations were also utilized in a subset of 24 samples for additional validation, i.e. *rs1046322G* allele–, *rs9457C* allele– and thus G–C haplotype specific reaction and *rs1046322A* allele–, *rs9457G* allele– and consequently A–G haplotype specific mixture. Results of these analyses confirmed the data obtained by the original setup. For additional optimization effect of partial (25% and 50%) substitution of dGTP with deoxyinosine-triphosphate (dITP) in the PCR was tested. This optimization step is particularly important if longer PCR-products tend not to be efficiently amplified, as deoxyinosine forms a base pair with deoxycytidine (similarly to deoxyguanosine) containing only two hydrogen-bonds, therefore the melting temperature of the long amplicons is decreased. The amount of allele-specific primers has been increased to 2.5  $\mu$ M in a pilot experiment. However, in case of our haplotyping technique, neither modification has resulted in any further improvement, consequently, the original protocol (described above) has been utilized in our subsequent experiments. 95 healthy Hungarian individuals were investigated by the method described above, and the genotype results were in 98.9% concordance with the data determined

earlier using sequence specific TaqMan probes, i.e. by an independent approach (data not shown). The single discordant result could be resolved by repeated genotype and haplotype determination. This latter method provided the possibility of single genotype analysis only, not capable to determine the haplotype of the two SNPs.

Figure 12C delineates genotype and haplotype determination of the *rs1046322* and the *rs9457* SNPs by conventional agarose slab gel electrophoresis in case of double heterozygote samples. To determine the size of the double allele-specific amplicons in the case of both haplotypes, a 100 bp DNA molecular mass marker (M) was used with the PCR products (Lanes 1-2). PCR fragments generated using reaction mixture-I (*rs1046322A*- and *rs9457C*-specific primers) can be seen on the left (“A”), whereas reaction mixture II (*rs1046322G*- *rs9457G*-specific primers) are shown on the right side (“B”) in Figure 12C.

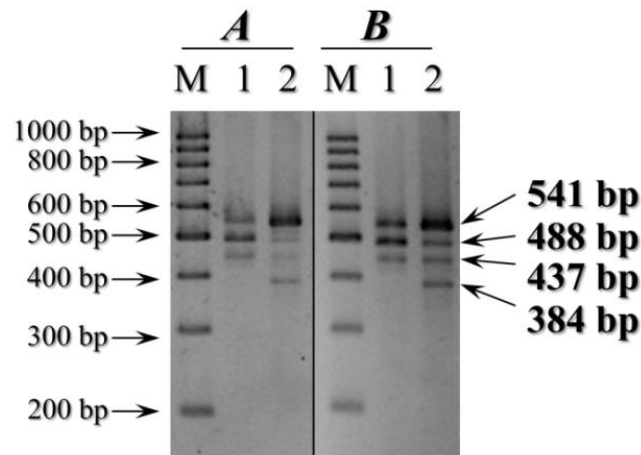
**A**



**B**



C

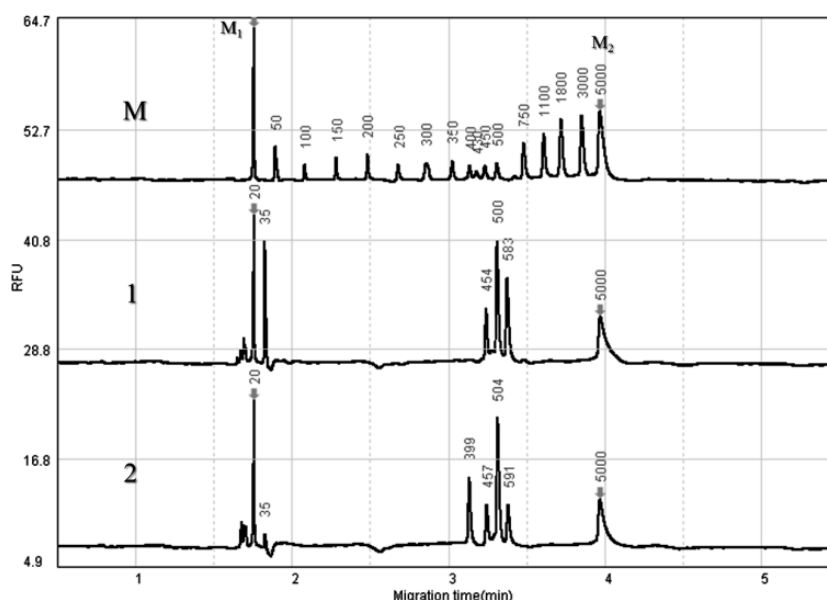


**Figure 12.** Direct haplotype determination of the rs1046322 and rs9457 SNPs in the *WSF1* gene by allele-specific multiplex PCR. **(A)** Products were generated in the presence of the sense rs1046322A- and the antisense rs9457C-specific primers in conjugation with the outer oligos. The 384-bp-long fragment was produced by rs1046322A and rs9457C primers and suggested the presence of the A–C haplotype. **(B)** PCR products were obtained in the presence of the rs1046322G- and rs9457G-specific primers in combination with the outer oligos. The 384-bp-long fragment demonstrated the presence of the G–G haplotype. **(C)** Genotype and haplotype readings by agarose slab gel electrophoresis. M: 100 base pair DNA mass ladder; Lanes 1–2: PCR samples: A: 437, 488, 541 bp dsDNA fragments; B: 384, 437, 488, 541 bp dsDNA fragments from the multiplex amplification reaction. Separation conditions: 2% agarose gel in  $1\times$  TAE containing 0.5  $\mu\text{g/mL}$  ethidium bromide;  $U = 100\text{ V}$ ;  $t = 45\text{ min}$ ; room temperature. Adapted from [2].

### 5.2.2. Haplotype Determination by CGE

The final step of the haplotyping protocol was the CGE-based size determination of the dsDNA fragments from the multiplex amplification reaction. Figure 13 depicts the CGE traces of the PCR fragments created by the haplotyping protocol. For fragment size assessment, the DNA mass ladder was extended with fragments in the size range of 50 – 3 000 bp (upper trace). The CGE separation of the mPCR samples is represented in the middle and lower traces. To obtain high fragment sizing accuracy, the mPCR samples were co-injected with the lower and upper alignment markers ( $M_1 = 20\text{ bp}$  dsDNA and  $M_2 = 5\text{ 000 bp}$  dsDNA).

The middle trace in Figure 13 represents the CGE separation of three dsDNA fragments from the reaction mixture-I with calculated sizes of 454, 500 and 583 bp fragments, corresponding to 437, 488 and 541 bp of the actual PCRs (see variance data in Table 4). The lower trace in Figure 13 illustrates the CGE separation of four dsDNA fragments from the reaction mixture-II with calculated sizes of 399, 457, 504 and 591 bp fragments, corresponding to 384, 437, 488 and 541 bp of the actual PCRs (with better than 95% average accuracy) by the rapid CGE-LEDIF analysis (see variance data in Table 4). The bp accuracy of each DNA fragment was analyzed by the Q-Analyzer software package with the accuracy range of 2.4 – 9.2% as shown in Table 4. In addition, the concentration of each DNA fragment was calculated based on their peak areas as also listed in Table 4.



**Figure 13.** Representative electropherograms of selected mPCR amplicons. Upper trace: DNA mass ladder ( $M = 50 - 3\,000$  bp) co-injected with the lower ( $M_1 = 20$  bp) and upper ( $M_2 = 5\,000$  bp) alignment markers; middle and lower traces: representative mPCR samples of 1 and 2 were the same as in Figure 12 with the respective alignment markers. Separation conditions: capillary:  $75\ \mu\text{m}$  i.d., effective separation length of 11 cm length (total separation length of 15 cm); sample and marker injection: 4 kV/10 s; separation voltage: 6 kV; ambient temperature. Adapted from [2].

**Table 4.** Base pair accuracy determination and calculated concentrations of the mPCR samples by CGE. Adapted from [2].

1. PCR sample (Figure 13 middle trace)					2. PCR sample (Figure 13 lower trace)				
Fragment (bp)	Measured (bp)	Variance (bp)	Accuracy (%)	Concentration (ng/ $\mu$ L)	Fragment (bp)	Measured (bp)	Variance (bp)	Accuracy (%)	Concentration (ng/ $\mu$ L)
–	–	–	–	–	<b>384</b>	399	15	3.9	1.65
<b>437</b>	454	17	3.8	1.34	<b>437</b>	457	20	4.5	0.90
<b>488</b>	500	12	2.4	4.16	<b>488</b>	504	16	3.2	4.04
<b>541</b>	583	42	7.7	2.38	<b>541</b>	591	50	9.2	1.09

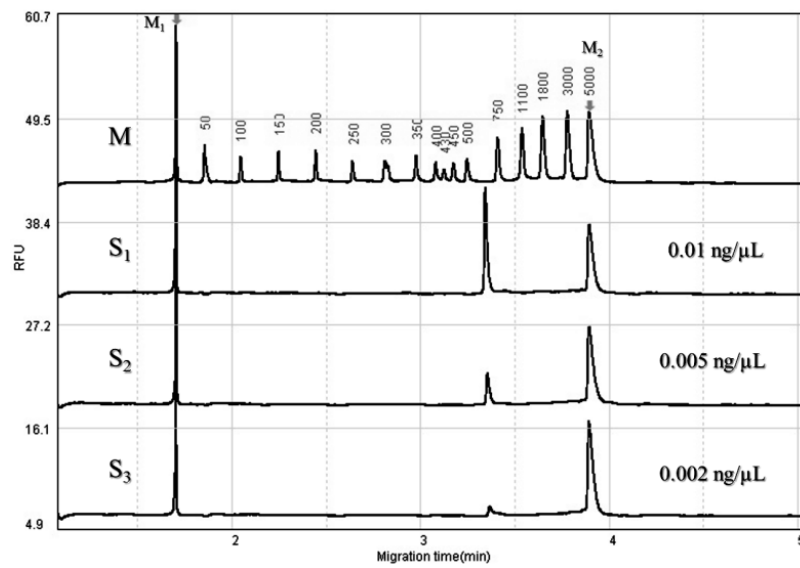
### 5.2.3. Limit of Detection and Detector Linearity Study for CGE

To determine the detection limit of the system, Figure 14A compares the resulting CGE traces after the injection of different concentration samples from 0.01 ng/ $\mu$ L to as low as 0.002 ng/ $\mu$ L. In this study, the 576 bp PCR fragment was diluted in water. When the detector linearity analysis was performed with the same water diluted samples, the linear detection range was quite narrow, i.e. 1.5 orders of magnitude, probably due to the effect of field amplification. Therefore, the detection linearity study was accomplished by using a dilution series in sample buffer in that case a linear detector response was obtained in a large interval of 0.08 – 10.0 ng/ $\mu$ L with an  $R^2 = 0.9997$ , as shown in Figure 14B and Table 5. Injection from water diluted samples resulted in much larger sample intake, because the buffer co-ions did not compete with the sample molecules, resulting in excellent LOD. On the other hand, measurements of the sample concentration were more precise from buffer diluted samples as shown in Table 5.

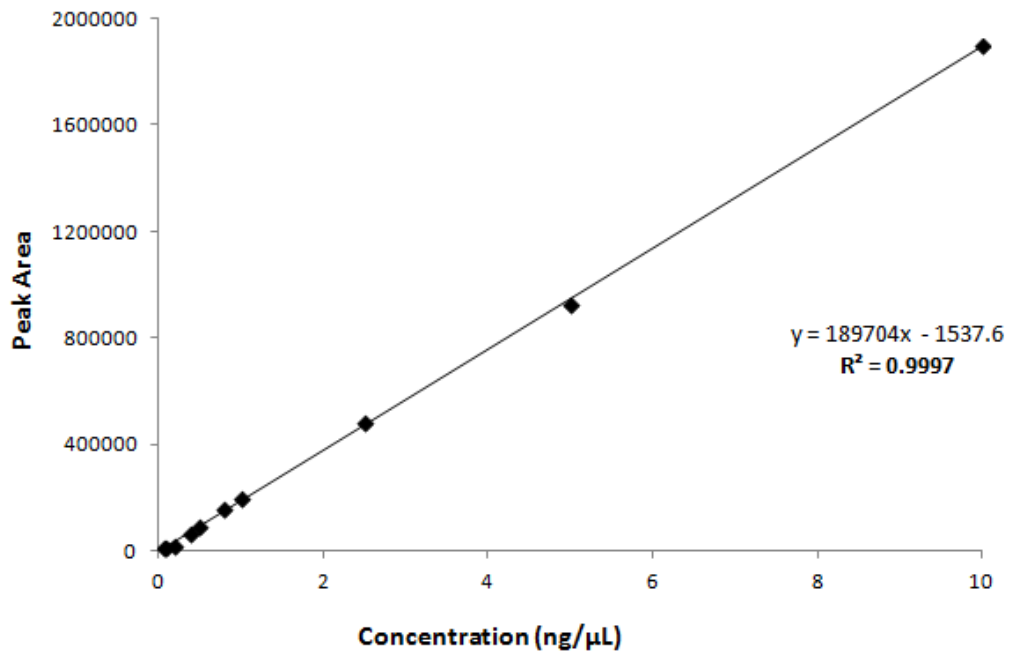
These results clearly present the numerous advantages of the capillary gel electrophoresis unit (Qsep100, BiOptic) in comparison to traditional techniques. The detection limit of slab gel electrophoresis was in the nanogram scale, which – even loading relatively high volume of a diluted sample – was ca. 2 orders of magnitude worse than that of CGE. Furthermore, automated quantification is readily available in a broad concentration range by this CGE instrument exceeding the accuracy, reliability, throughput and sensitivity of slab gel densitometry.



**A**



**B**



**Figure 14.** Limit of detection and detector linearity study for CGE. **(A)** Determination of the limit of detection (LOD) with the 576-bp-long PCR product serially diluted in water compared to the DNA mass ladder. **(B)** Detection linearity study using the 576 bp PCR product serially diluted in the sample buffer. Separation conditions, mass ladder as well as lower and upper alignment markers were the same as in Figure 13. Adapted from [2].

**Table 5.** CGE detection linearity study using the 576-bp-long PCR product in the 10.00 - 0.08 ng/ $\mu$ L concentration range. Adapted from [2].

Detector linearity of the 576 bp PCR sample										
	No 1.	No 2.	No 3.	No 4.	No 5.	No 6.	No 7.	No 8.	No 9.	No 10.
Concentration (ng/ $\mu$ L) <sup>a)</sup>	10.0	5.00	2.50	1.00	0.80	0.50	0.40	0.20	0.10	0.08
Average peak area <sup>b)</sup>	1898510	930498	487797	199921	160618	96405	65104	21230	15755	12894
SD	21552	7487	23358	8014	748	2960	11572	2038	1589	1007
RSD%	1.13	0.80	4.78	4.00	0.46	3.07	17.77	9.60	10.08	7.81

<sup>a)</sup> WFS1 PCR samples were diluted with dilution buffer. <sup>b)</sup> Average peak area was determined from triplicate measurements for each concentration.

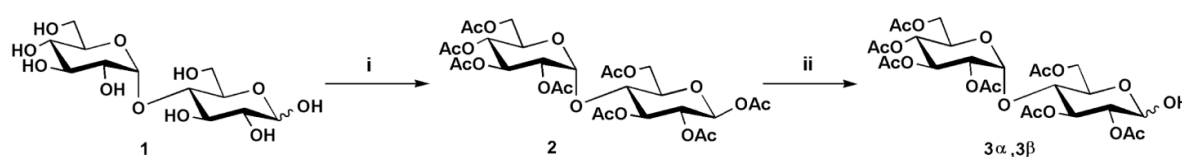
### 5.3. Immune Response against Carbohydrate Antigens

#### 5.3.1. Synthesis of Carbohydrate Antigens

Today, the study of glycoproteins is an area of intensive research for chemical biologist in the glycobiology field [165]. Antibody development against carbohydrates is of increasing importance, because carbohydrate antigens play a crucial role in biomarker research and discovery [166]. For detecting carbohydrate antigens, the anti-carbohydrate antibodies have been utilized as diagnostic markers in several diseases [167] as well as it may occur as neoglycoprotein-based antibacterial vaccines or neoglycoprotein-based antitumor vaccines for cancer [168]. As early as in 1931, neoglycoproteins as sugar-specific antigens were noted as glycoconjugate vaccines by Avery and Goebel [169]. Since then neoglycoproteins have been utilized as immune response triggers antitumor-associated sugar antigens [170]. Cancer-associated carbohydrates are mostly located on the surface of cancer cells, which demonstrate potential targets for new diagnostic assays and therapeutic treatments [171]. The synthesis of neoglycoproteins particularly incorporates random or defined coupling sites on the surface of the carrier protein, which is covalently modified with glycans via their reducing end or functionalized glycan bearing spacer arms [168, 172, 173].

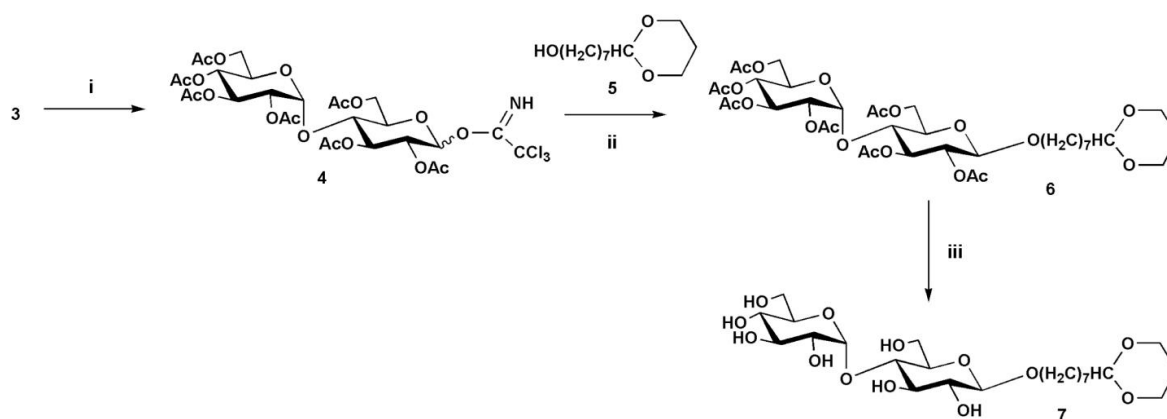
In the third part of this study, a simple sugar-specific antigen (maltose) was conjugated to a carrier protein (namely BSA) by reductive amination. BSA was chosen as a carrier protein since

in its natural form it voids glycosylation. To conserve the intact annular maltose structure, an aglycone spacer was utilized for the synthesis of neoglycoproteins. Neoglycoproteins were synthesized under controlled conditions with different numbers of oligomeric/dimeric maltose molecules (between 32 and 66 units/mol) linked to the lysine  $\epsilon$ -amino residues of BSA. After that the maltose-conjugated BSA samples were investigated by SDS-CGE as well as MALDI-TOF MS and the resulted glycoconjugates were presented on their ability to elicit specific immune response. The first step in the synthesis, maltose **1** was converted into maltose octaacetate **2** with acetic anhydride, followed by the removal of the anomeric acetyl group with hydrazine acetate to yield the **3 $\alpha$** , **3 $\beta$**  hemiacetal anomers as shown in Scheme 1.



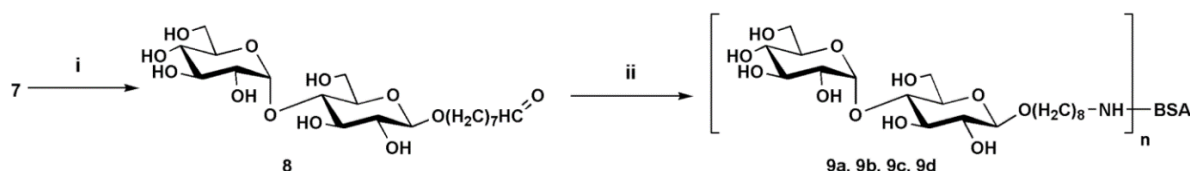
**Scheme 1.** Synthesis of neoglycoproteins I.: Maltose (**1**) was converted into maltose octaacetate (**2**) and then into **3 $\alpha$** , **3 $\beta$**  hemiacetal anomers (**3**). Reagents and conditions: (i)  $\text{Ac}_2\text{O}$ ,  $\text{NaOAc}$ ,  $110^\circ\text{C}$  (ii)  $\text{NH}_2\text{NH}_2 \cdot \text{AcOH}$ , DMF. Adapted from [174].

After that the resulted hemiacetal was converted to a trichloroacetimidate donor **4** with trichloro-acetonitrile, and then reacted with the aglycone spacer of 7-(1,3-dioxan-2-yl)-heptan-1-ol **5** to obtain a protected maltose with the formyl-heptyl spacer **6**. Ultimately, the protecting acetyl groups were removed with sodium methylate **7** as illustrated in Scheme 2.



**Scheme 2.** Synthesis of neoglycoproteins II.: The resulted hemiacetal was converted (**3**) to a trichloroacetimidate donor (**4**), and then was reacted with an aglycone spacer (**5**) to obtain a protected maltose with a formyl-heptyl spacer (**6**). Reagents and conditions: (i)  $\text{Cl}_3\text{CCN}$ ,  $\text{K}_2\text{CO}_3$ ,  $\text{CH}_2\text{Cl}_2$  (ii)  $\text{TMSOTf}$ ,  $\text{CH}_2\text{Cl}_2$ ,  $-40^\circ\text{C}$  (iii)  $\text{NaOCH}_3$ ,  $\text{CH}_3\text{OH}$ . Adapted from [174].

After removing the acetal protecting groups with acid hydrolysis, the resulting aldehyde **8** was attached to the lysine  $\epsilon$ -amino groups of BSA to form the Schiff base, which was consequently reduced with  $\text{NaCNBH}_3$  to obtain a stable conjugate **9**. Different sugar-containing neoglycoproteins were synthesized (**9a**, **9b**, **9c** and **9d**) by means of the conjugation methods as illustrated in Scheme 3. The molar ratio of compound **7** (and the reagents) to BSA was varied in the coupling reaction as mentioned in Table 2.



**Scheme 3.** Synthesis of neoglycoproteins III.: After removing the protecting acetyl groups (**7**), carbohydrate antigens (**9a**, **9b**, **9c** and **9d**) were synthesized by reductive amination. Reagents and conditions: (i) 96%  $\text{AcOH}$ ,  $50^\circ\text{C}$  (ii) BSA,  $\text{NaCNBH}_3$ , pH 7.0. Adapted from [174].

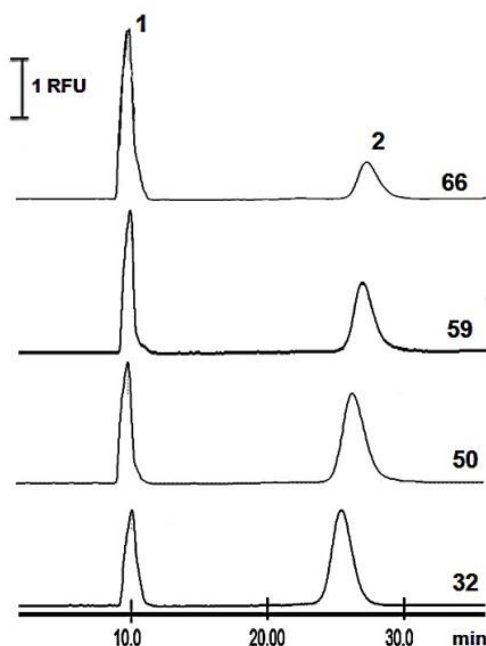
### 5.3.2. Analysis of Neoglycoproteins by CGE and MALDI-TOF MS

After modifying BSA with many different sugars, the reaction products were first analyzed by SDS-CGE for purity and homogeneity. Figure 15 depicts the resulting SDS-CGE traces of the unconjugated dye (first peaks at 10 min) and the resulted neoglycoproteins (second peaks), having 32, 50, 59, and 66 units/mol maltose units. The number of conjugated sites was determined by MALDI-TOF mass spectrometry, as shown in Table 2.

In this study, a decrease was observed in signal intensity of the glycoconjugates with increasing sugar unit content and the longer migration times during the separation. The former was due to the effect of the larger molecular mass of the conjugates with the attached maltose residues and the concomitantly suppressed SDS binding to these hydrophilic patches, both causing migration time increase. The decreasing signal intensity of the conjugate (and the increasing signal intensity of the unconjugated free dye) with the increasing maltose content was due to the reduced number of available conjugation sites for the fluorescent dye, i.e. remaining free amino residues on BSA.

It should be emphasized that the incorporation level of sugars increased from 32 to 66 units by increasing the molar ratio of the maltose to BSA. This result was surprising, because BSA only contained 60 free amino groups but can be explained by the studies of Schwartz *et al.* [175],

who demonstrated that tertiary amines can be also formed during the reductive amination reaction possessing two sugar structures, consequently increasing the maximum number of conjugated glycans beyond 60.

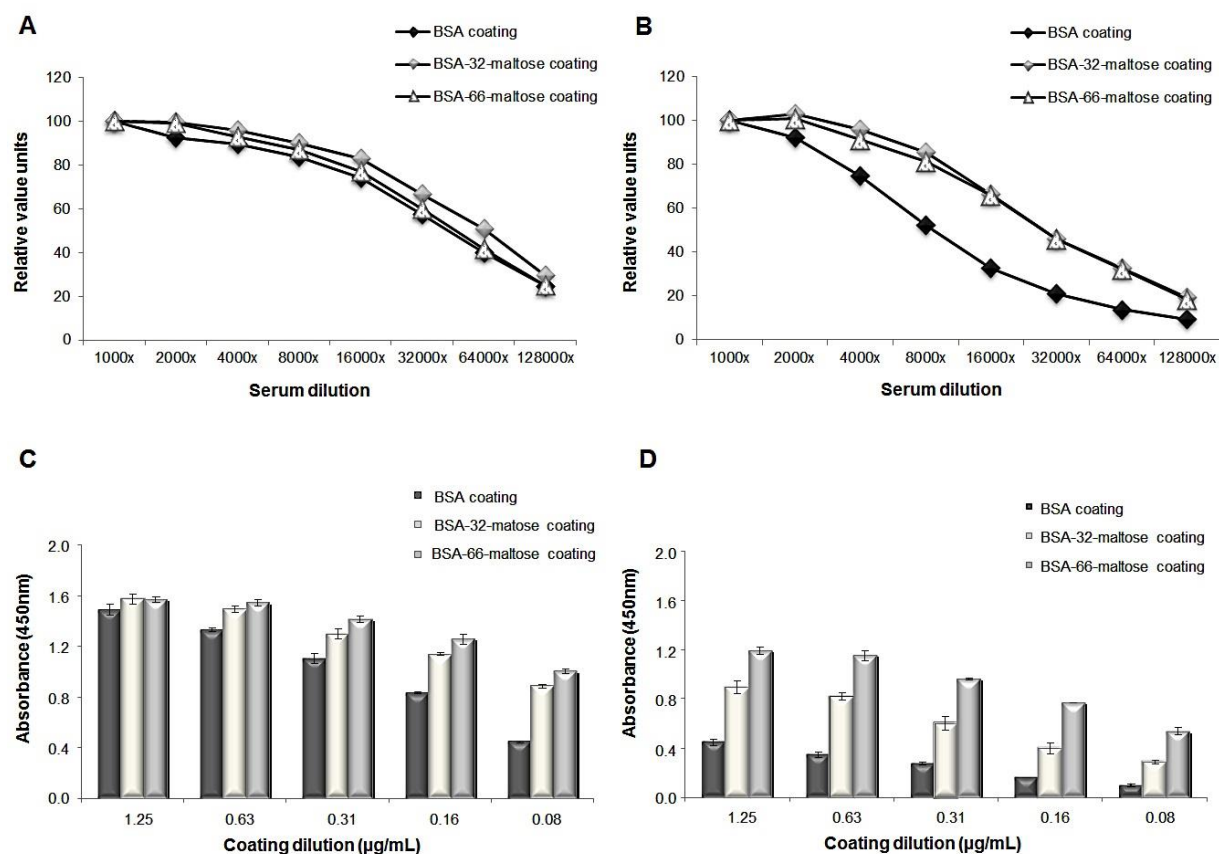


**Figure 15.** Representative electropherograms of covalently fluorophore tagged and reduced glycoconjugate samples by SDS-CGE. Peaks: (1) remaining labeling dye, (2) sugar-conjugated BSA. Separation conditions: capillary: 75  $\mu\text{m}$  i.d., effective length of 10 cm (total separation length of 30 cm); BGE: SDS-MW gel buffer (SCIEX);  $E=500$  V/cm; electrokinetic injection: 25 kV for 10 s; ambient temperature. Adapted from [174].

### 5.3.3. Polyclonal Antibody Response to Neoglycoproteins Mediated Immunization

To generate polyclonal antibody response, BALB/c mice were immunized using neoglycoproteins with low and high sugar incorporation levels as immunogens (i.e. 32 and 66 maltose units/BSA) and the sugar components of the conjugates were investigated by ELISA. In mice, immunized with the neoglycoprotein of low sugar incorporation level, a rather unified response was observed towards the different antigens, as illustrated in Figure 16A. In mice immunized with the high sugar incorporated neoglycoprotein the IgG response was more intense towards the modified antigens and was distinguishable to that of the significantly weaker carrier specific response (Figure 16B). When the experiment was monitored under immunogen limiting conditions, this difference was even more pronounced.

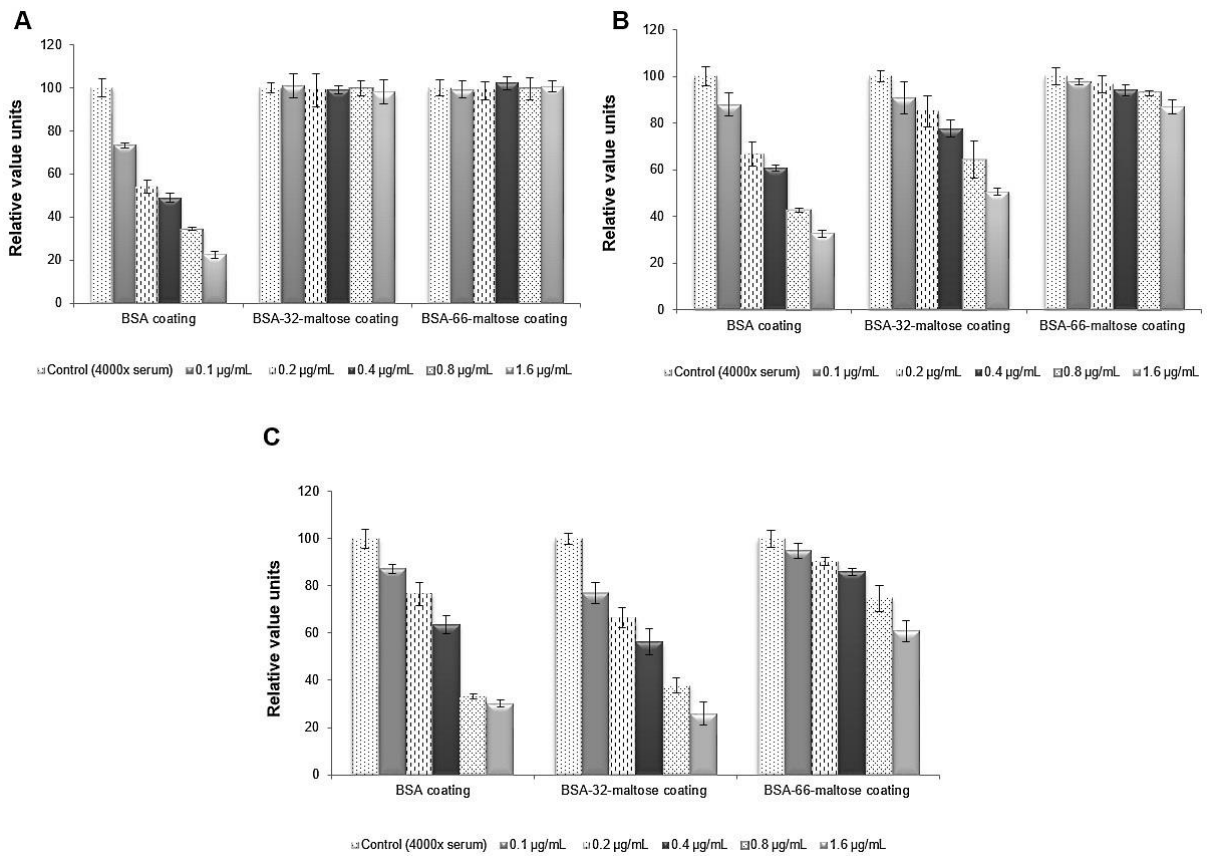
Thus, a more significant antibody response could be observed against neoglycoproteins than against the carrier proteins in the presence of limited amount of immunogen as shown in Figure 16C and D. Although, this difference might vary depending on the immunogen used for the immunization or even as the response of the individual animals. However, the recognition of the carrier and the glycan modified BSA by the immune sera of some mice suggested the presence of antibody populations specific for the neoglycoproteins.



**Figure 16.** Polyclonal antibody response towards the protein carrier and the synthesized neoglycoproteins. Antiglycan antibodies generated by immunization with BSA-32-maltose (**A** and **C**) and BSA-66-maltose (**B** and **D**) incorporated sugar-containing neoglycoproteins detected by ELISA. Experiments were done either by using a fixed amount of immobilized antigens (10 µg/mL); and serial (1 000 – 128 000×) dilutions of immune-sera (**A** and **B**), or by using serial dilution of immobilized antigens (from 1.25 to 0.08 µg/mL) and 4 000-fold dilution of the sera. Adapted from [174].

#### **5.3.4. Inhibition of Mouse Antibody Binding in ELISA**

In order to test the specificity of the polyclonal antibody response to synthetic neoglycoproteins, some inhibition experiments were analyzed by ELISA. In this study, the polyclonal antisera were reacted with immobilized BSA-32-maltose and BSA-66-maltose in the presence of inhibitors. Response to each immunogen, that is BSA, BSA-32-maltose, and BSA-66-maltose, was specifically inhibited by the corresponding immunogens, i.e. BSA-specific response was inhibited by BSA (Figure 17A), BSA-32-maltose-specific response was inhibited by BSA-32-maltose (Figure 17B), and BSA-66-maltose specific response was inhibited by BSA-66-maltose (Figure 17C). In addition, BSA-32-maltose inhibited BSA-specific response (Figure 17B), while BSA-66-maltose inhibited both BSA- and BSA-32-specific responses (Figure 17C). On the other hand, no IgG antibody response was detected in the presence of free maltose and other sugars such as glucose, isomaltose, lactose, galactose and maltodextrin (data not shown). This result suggests that the synthesized glycoconjugates contained specific epitopes, which were unique to the presence of the defined number of conjugated maltose units per BSA molecules. Overall, these results suggest that immunogenic neoepitopes were induced by maltose conjugation and increase in the maltose units from BSA-32-maltose to BSA-66-maltose resulted in the appearance of new epitopes not present on the BSA-32-maltose molecules.



**Figure 17.** Schematic illustration of the inhibition of antiglycan antibody binding to their antigens using BSA (A), BSA-32-maltose (B), and BSA-66-maltose (C) sugar incorporated glycoprotein inhibitors. Antiglycan antibodies (4 000× serum dilutions, immunized with high sugar incorporated antigens in mice) were examined with all three coatings (0.2 µg/mL) in the presence of BSA inhibitor (0.1–1.6 µg/mL), BSA-32-maltose and BSA-66-maltose (0.1–1.6 µg/mL) glycoprotein inhibitors. Adapted from [174].



## 5.4. Generation of an ANTS-labeled *N*-Glycan Database for CGE Analysis of Carbohydrates

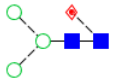
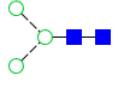
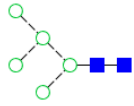
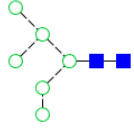
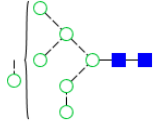
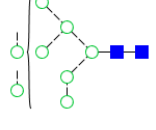
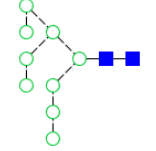
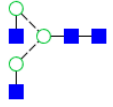
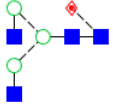
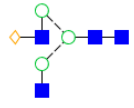
### 5.4.1. Generation of an ANTS-labeled *N*-Glycan Database

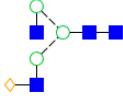
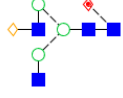

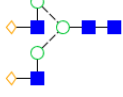
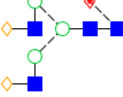
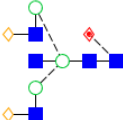
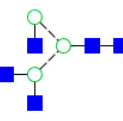

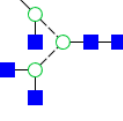
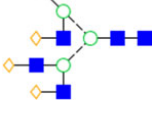
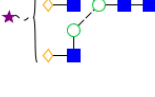
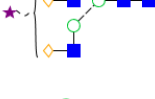

In this part of the work, an 8-aminonaphtalene-1,3,6-trisulfonate (ANTS)-labeled *N*-glycan database was developed for the LED-induced fluorescent (LEDIF) CGE system. At this time, the establishment of ANTS-labeled *N*-glycan database includes 25 oligosaccharides of mostly biopharmaceutical interest, such as for *N*-glycosylation profiling of therapeutic antibodies.

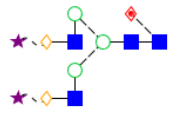
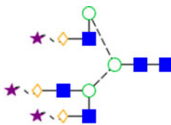
For the generation of this comprehensive *N*-glycan database, the individual glycan standards were labeled via reductive amination using the ANTS labeling reagent, and then separated by CGE, each possessing only one charged fluorophore. This fluorescent dye provided the necessary charges for the carbohydrate molecules to support their electric-field mediated separation, and it also assisted as a fluorophore for high-sensitivity LEDIF detection. In addition, this labeling reagent offers 1:1 stoichiometry, meaning that only one fluorophore label, possessing three strong negative charges, is attached to the glycosidic OH group of all sugar structures, making accurate quantitation possible. Relative migration times (RMTs) of the oligosaccharides were calculated based on the migration times of the bracketing standards according to equation 18, which was introduced in section 2.9.1.

Table 6 illustrates the commercially available individual sugar standards with their GU values, separated one-by-one by CGE-LEDIF. The database represents the abbreviated names of the individual glycan standards with their graphic representation following the proposed symbol nomenclature of Rudd *et al.* [150], the precise molecular mass and the glucose unit (GU) values with the corresponding standard deviations (SD). These GU values in Table 6 show the migration position of each sugar structure, which is dependent on the number of oligosaccharide residues as well as positional and linkage related conformational changes, i.e. shape of the molecule. Each GU entry is the average of a total of six runs and associated with their SD values. The resulted GU values can be applied as reference values for rapid structural elucidation of unknown glycan samples as explained below in reference to Figures 18 and 19.

**Table 6.** Abbreviated names, structures, molecular masses, and CGE-LEDIF GU values of ANTS-labeled individual N-glycan standards. Adapted from [114].

Peak No.		Name	Structure <sup>a)</sup>	Mass Average <sup>b)</sup>	GU values <sup>c)</sup>
1	<i>NEUTRAL</i>	F(6)M3		1056.9651	<b>5.90 ± 0.05</b>
2		M3		910.8236	<b>5.25 ± 0.03</b>
3		M5		1235.1053	<b>7.10 ± 0.01</b>
4		M6		1397.2462	<b>7.97 ± 0.01</b>
5		M7		1559.387	<b>8.94 ± 0.02</b>
6		M8		1721.5278	<b>9.69 ± 0.14</b>
7		M9		1883.6687	<b>10.28 ± 0.08</b>
8		A2		1317.2094	<b>7.34 ± 0.07</b>
9		F(6)A2		1463.3508	<b>8.02 ± 0.06</b>
10		A2[6]G(4)1		1479.3502	<b>8.31 ± 0.01</b>

11		A2[3]G(4)1		1479.3502	<b>8.51 ± 0.01</b>
12		F(6)A2[6]G(4)1		1625.4917	<b>8.91 ± 0.04</b>
13		F(6)A2[3]G(4)1		1625.4917	<b>9.13 ± 0.05</b>
14		A2G(4)2		1641.4911	<b>9.44 ± 0.02</b>
15		F(6)A2G(4)2		1787.6326	<b>10.26 ± 0.01</b>
16		F(6)A2BG(4)2		1990.8254	<b>10.41 ± 0.18</b>
17		A3		1520.4023	<b>8.22 ± 0.01</b>
18		A3G3		2006.8248	<b>11.78 ± 0.04</b>
19		A4		1723.5952	<b>9.27 ± 0.01</b>
20		A4G(4)4		2372.1587	<b>13.72 ± 0.07</b>
21	<i>MONO-SIALO</i> (S1)	A2G(4)2S1		1932.7462	<b>6.81 ± 0.19</b>
22		F(6)A2G(4)2S1		2078.8877	<b>7.04 ± 0.02</b>
23	<i>DI-SIALO</i> (S2)	A2G(4)2S2		2224.0012	<b>5.33 ± 0.10</b>

24		F(6)A2G(4)2S2		2370.1426	<b>5.69 ± 0.01</b>
25	<b>TRI-SIALO (S3)</b>	A3G(4)3S3		2880.5900	<b>5.63 ± 0.02</b>

Symbols: ○ Mannose; ■ *N*-acetyl glucosamine; ◆ Fucose; ◇ Galactose; ★ Sialic acid

<sup>a)</sup> The glycan symbol nomenclature proposed by Harvey *et al.* was used for drawing structural diagrams of *N*-linked glycans [150]. <sup>b)</sup> Mass averages of the glycan standards were applied from the NIBRT GlycoBase 3.2.4 suite of databases. <sup>c)</sup> Each GU value was calculated from six runs.

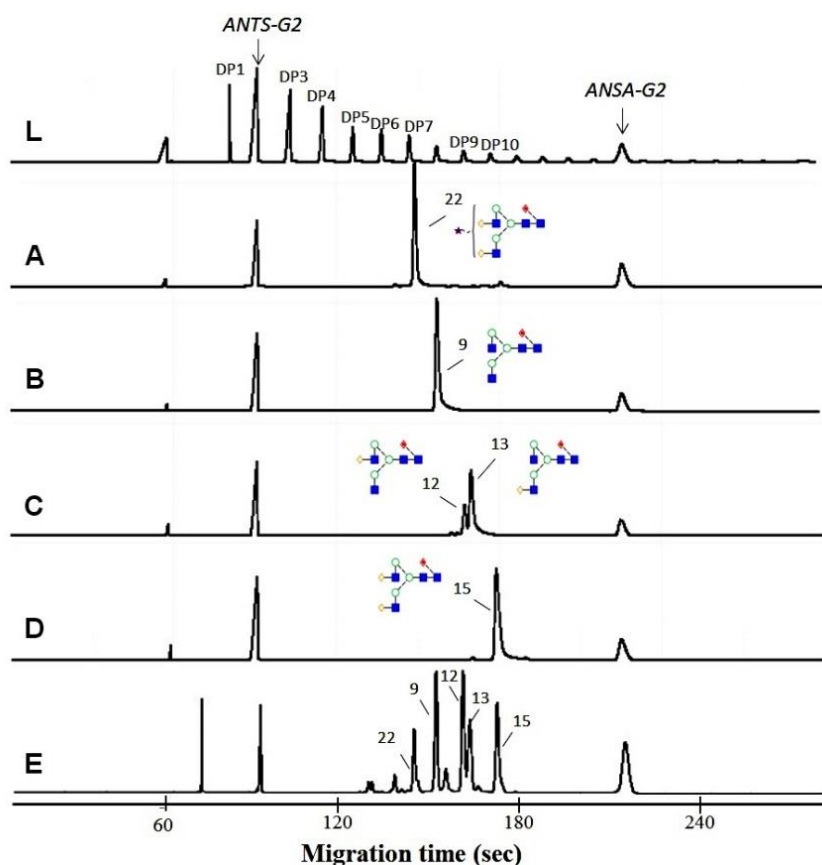
#### 5.4.2. Validation of the ANTS-labeled *N*-Glycan Database

Finally, the generated GU values in this database were verified with *N*-glycans released from human IgG and bovine pancreatic RNase B. Briefly, glycoproteins were first released by the endoglycosidase PNGase F and then the liberated sugars were labeled via reductive amination using the ANTS labeling reagent. After that, the ANTS-labeled *N*-glycans were separated by CGE.

Immunoglobulins are highly abundant glycosylated serum proteins, readily available for glycan analysis, containing two light (LC) and two heavy (HC) chains, held together and folded through intra- and inter-chain disulfide bounds. Each heavy chain possesses a single covalently attached *N*-glycan at the highly conserved Asn-297 site in each of the C<sub>H</sub>2 domains of the Fc portion of the molecule [176]. The attached oligosaccharides are structurally important for the stability of the antibody [177].

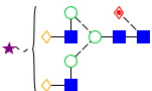
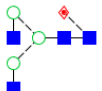
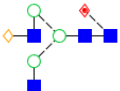
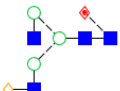
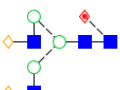
Another commonly used glycoprotein for glycan analysis is ribonuclease B (bovine pancreatic, *M<sub>r</sub>*: 15 500) with one *N*-linked glycosylation site at the Asn-34 position, which occupies at least five to nine different high-mannose (Man) type carbohydrate structures of varying size linked to the side chain of asparagine through an *N*-acetylglucosamine (GlcNAc) residue [178]. Please note that CGE can differentiate even isobaric oligosaccharide structures holding the same charge just based on their differences in shape. Oligosaccharides differing by one or more sugar residues and/or with different linkages/positions shift the migration time therefore readily separated by CGE. To support the highest possible precision of migration time measurement, a lower (ANTS-G2) and a higher (ANSA-G2) bracketing standard were co-injected with the sample in all CGE separations.

Figure 18 compares the CGE separation traces of the ANTS-labeled maltooligosaccharide ladder (trace L) with the individual glycan standards of F(6)A2G(4)2S1 (trace A, peak 22), F(6)A2 (trace B, Peak 9), F(6)A2[6]G(4)1 and F(6)A2[3]G(4)1 (trace C, peaks 12 and 13, respectively), and F(6)A2G(4)2 (trace D, peak 15). Trace E shows the CGE separation of the ANTS-labeled human IgG glycan pool, including the individual peaks of 22, 9, 12, 13, and 15 as the standards above. The measured GU values of these peaks are compared to the database values in Table 7 and show very little differences (<0.14 GU). It should be emphasized that the CGE analysis times were around 200 s. All peak numbers correspond to the entries in Table 6.



**Figure 18.** CGE-LEDIF separation of ANTS-labeled glycan standards and N-glycans released from human IgG. Traces: (L) maltooligosaccharide ladder, (A) F(6)A2G(4)2S1, (B) F(6)A2, (C) F(6)A2[6]G(4)1, and F(6)A2[3]G(4)1, as well as (D) F(6)A2G(4)2 and (E) hIgG glycans. All samples were co-injected with lower (ANTS-G2) and upper (ANSA-G2) bracketing standards. Numbers on the peaks in the upper trace correspond to the degree of polymerization of the standard glucose ladder. Numbers on the peaks in traces A–E correspond to the structures in Tables 6 and 7. Separation conditions: capillary: 50  $\mu$ m i.d., effective length of 11.5 cm (total separation length: 15.5 cm); sample and marker injection: 2 kV/3 s; separation voltage: 6 kV; ambient temperature. Adapted from [114].

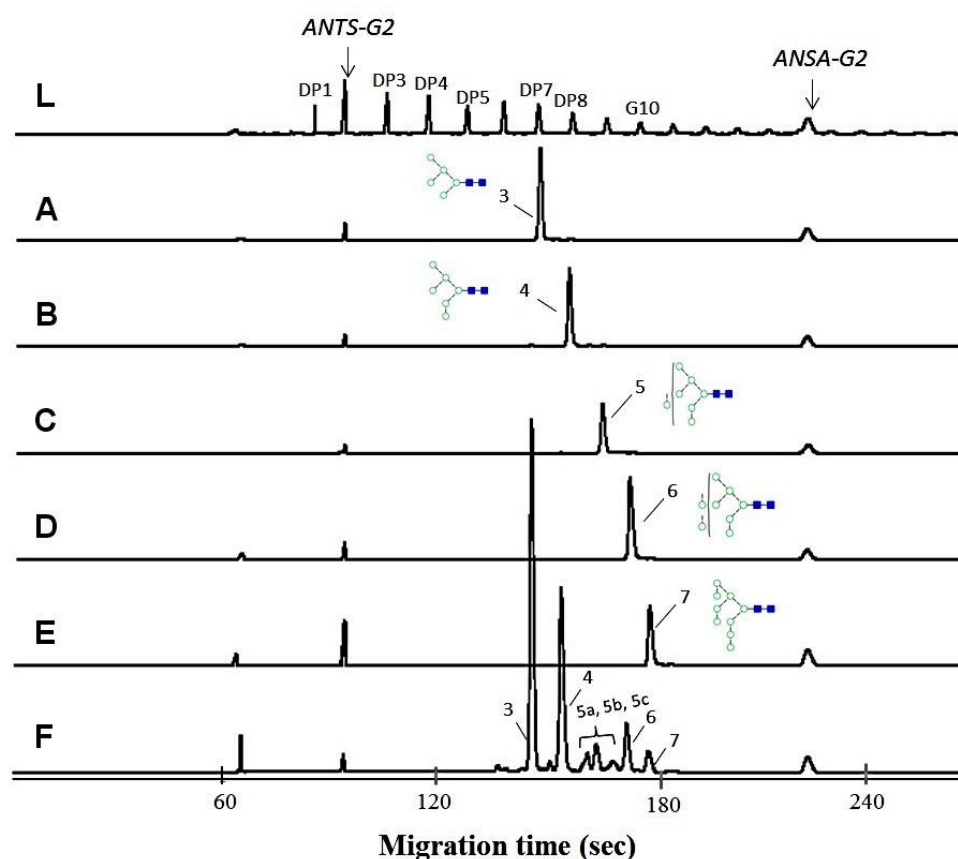
**Table 7.** Abbreviated names, structures and CGE-LEDIF GU values (measured and database values) of ANTS-labeled *N*-glycans released from human IgG. Adapted from [114].

Peak No.	Name	Structure <sup>a)</sup>	Database Values <sup>b)</sup>	Measured GU Values	Differences
22	FA2G(4)2S(6)1		$7.04 \pm 0.02$	$6.94 \pm 0.05$	-0.10
9	F(6)A2		$8.02 \pm 0.06$	$7.88 \pm 0.05$	-0.14
12	F(6)A2[6]G(4)1		$8.91 \pm 0.04$	$8.95 \pm 0.06$	+0.04
13	F(6)A2[3]G(4)1		$9.13 \pm 0.05$	$9.23 \pm 0.06$	+0.1
15	F(6)A2G(4)2		$10.26 \pm 0.01$	$10.27 \pm 0.06$	+0.01

Symbols are the same as in Table 6.




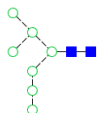

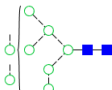
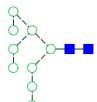
<sup>a)</sup> The glycan symbol nomenclature proposed by Harvey *et al.* was used for drawing structural diagrams of *N*-linked glycans [150]. <sup>b)</sup> The database values of the glycan standards arose from Table 6.

In a similar manner, Figure 19 compares the CGE analysis of the ANTS-labeled maltooligosaccharide ladder (trace L) with the individual high-mannose type glycan structures of M5 (trace A, peak 3), M6D3 (trace B, peak 4), M7D3 (trace C, peak 5), M8 (trace D, peak 6), and M9 (trace E, peak 7). Trace F shows the CGE separation trace of the ANTS-labeled RNase B glycan pool, including the individual peaks of 3, 4, 5a (M7D2), 5b (M7D3), 5c (M7D1), 6, and 7 as the standards in Table 8. Very little differences can also be observed between the GU values of the corresponding glycoprotein derived glycan peaks and the individual standards (<0.13 GU). All peak numbers correspond to the entries in Table 6.



**Figure 19.** Representative CGE-LEDIF electropherograms of ANTS-labeled sugar standards and high-mannose type oligosaccharides released from bovine RNase B. Traces: (L) maltooligosaccharide ladder, (A) M5, (B) M6, (C) M7, (D) M8, (E) M9, and (F) RNase B glycan pool. All samples were co-injected with lower (ANTS-G2) and upper (ANSA-G2) bracketing standards. Numbers on peaks in traces A–F correspond to the structures in Tables 6 and 8. Separation conditions were the same as in Figure 18. Adapted from [114].

**Table 8.** Abbreviated names, structures and CGE-LEDIF GU values (measured and database values) of ANTS-labeled N-glycans released from bovine RNase B. Adapted from [114].

Peak No.	Name	Structure <sup>a)</sup>	Database Values <sup>b)</sup>	Measured GU Values	Differences
3	M5		$7.10 \pm 0.01$	$6.99 \pm 0.03$	-0.11
4	M6		$7.97 \pm 0.01$	$7.84 \pm 0.03$	-0.13
5a	M7 a (M7 D2)		N/A	$8.58 \pm 0.03$	N/A
5b	M7 b (M7 D3)		$8.94 \pm 0.02$	$8.83 \pm 0.03$	-0.11
5c	M7 c (M7 D1)		N/A	$9.31 \pm 0.03$	N/A
6	M8		$9.69 \pm 0.14$	$9.71 \pm 0.03$	+0.02
7	M9		$10.28 \pm 0.08$	$10.34 \pm 0.03$	+0.06

Symbols are the same as in Table 6.

<sup>a)</sup> The glycan symbol nomenclature proposed by Harvey *et al.* was used for drawing structural diagrams of N-linked glycans [150]. <sup>b)</sup> The database values of the glycan standards arose from Table 6.

## 5.5. Effect of Separation Temperature and Background Electrolyte Composition on Structure Specific Glycan Migration in CGE

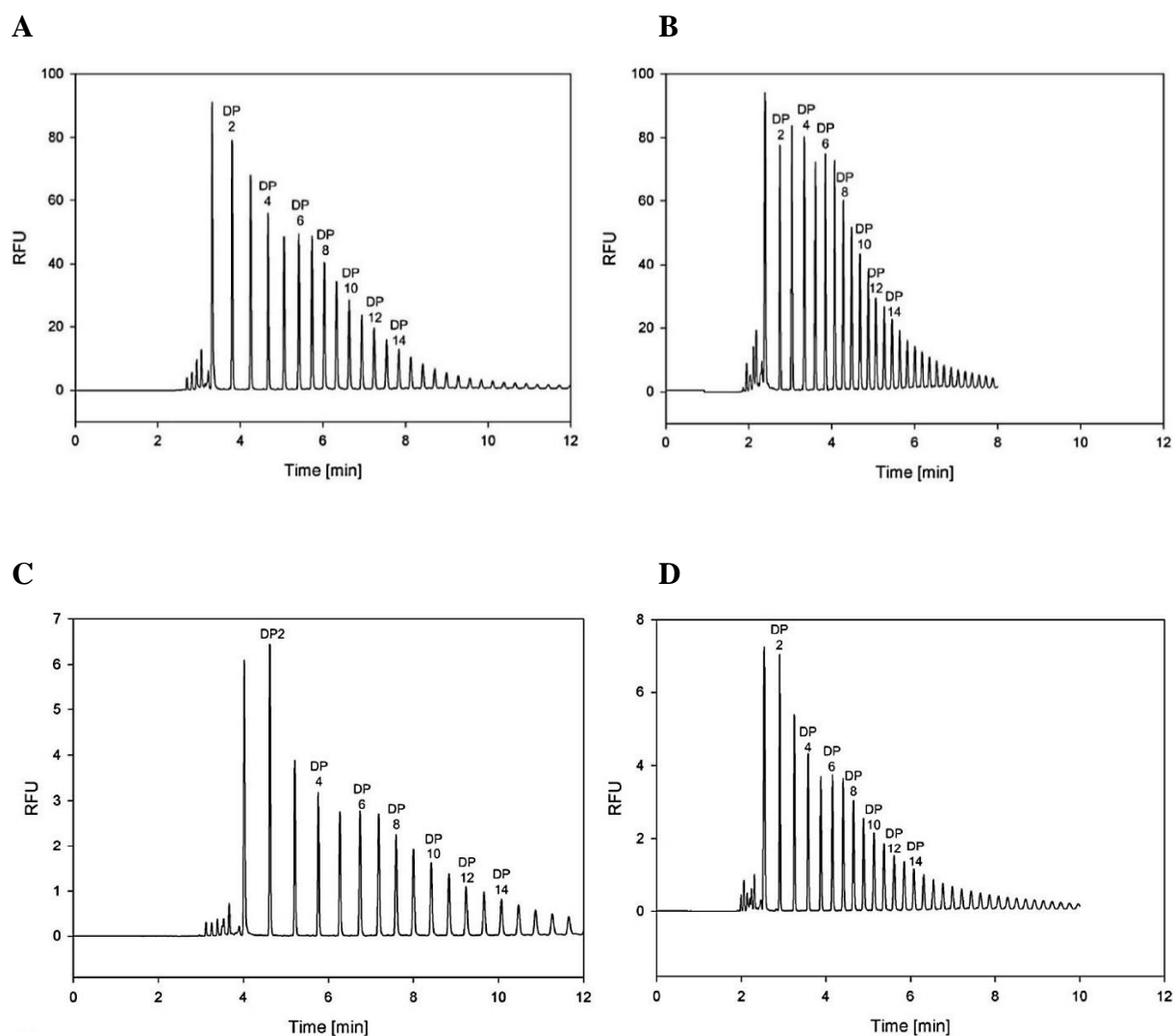
### 5.5.1. Activation energy associated with the electromigration of linear sugar oligomers

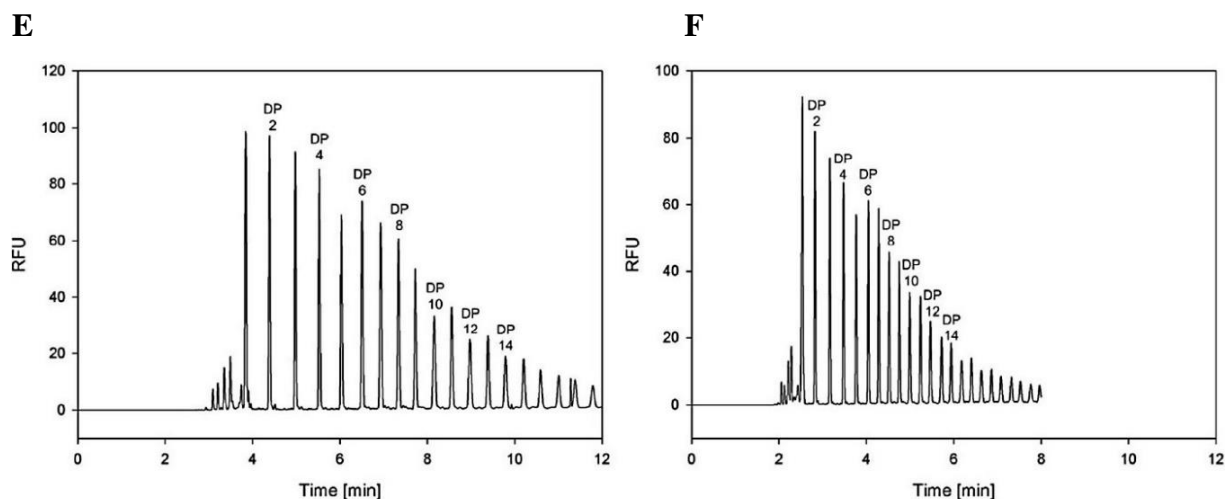
In the last part of the study, the activation energy ( $E_a$ ) associated with the electromigration of linear sugar oligomers was determined from the electrophoretic mobilities of the sample components at different temperatures.



The effects of a monomeric viscosity modifier (ethylene glycol) and a polymeric additive (linear polyacrylamide) containing BGE were investigated with respect to the electrophoretic migration of increasing chain length  $\alpha$ 1-4 linked glucose units (maltooligosaccharide, DP 1→15) using the activation energy concept.

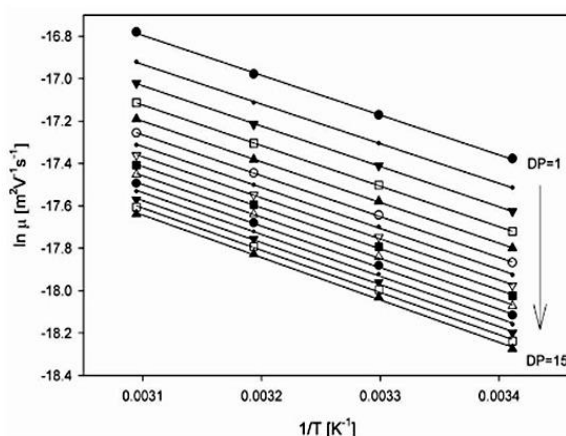
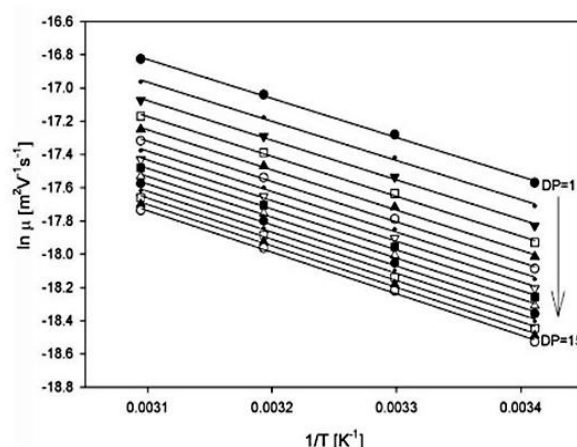
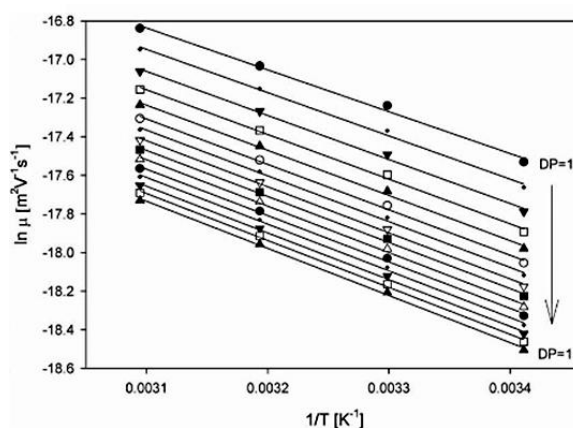
First, the electrophoretic separation of APTS-labeled linear sugar oligomers was monitored by CGE-LIF detection in the temperature range of 20 – 50°C, using a simple BGE of 25 mM lithium acetate (pH 4.75) without any viscosity modifier or polymeric additive. The results suggested that the electromigration of the sample components significantly increased with the increase of the separation temperature, i.e. resulted in faster migration as shown in Figure 20, panel A and B.





**Figure 20.** CGE-LIF separation traces of APTS-labeled linear maltooligosaccharides at 20°C (panels A, C and E) and 50°C (panels B, D and F). Panels A and B: no viscosity modifier or polymeric additive containing BGE; Panels C and D: 10% EG containing BGE; Panels E and F: 2% LPA (MW 10 000) containing BGE. Separation conditions: capillary: 50  $\mu$ m i.d., effective length of 21 cm (total length of 31 cm); BGE: 25 mM lithium acetate (pH 4.75) with corresponding viscosity modifier (C and D) and polymeric additive (E and F);  $E=400$  V/cm; Sample injection: 0.5 psi/5 s; Separation temperature: as indicated. Adapted from [179].

In Figure 21A, the  $E_a$  curves were plotted based on the slopes of the Arrhenius plots obtained from glucose (DP 1) to maltopentadecaose (DP 15) using equation 14, as discussed in section 2.2.1. Detailed investigation of the Arrhenius plots suggested that the  $E_a$  requirement associated with the electrophoretic mobility of linear maltooligosaccharides changed with their degree of polymerization (DP). In Figure 22, the  $E_a$  values were plotted based on the function of DP using lithium acetate (plot of 0% EG) as background electrolyte, where a monotone increase of  $E_a$  with growing size of the maltooligomers was observed with 73 J/mol per glucose unit ( $r^2=0.96$ ) in the range of between maltose (DP 2) to maltopentadecaose (DP 15). It should be emphasized that the monomeric glucose unit (DP 1) was excluded from the slope calculation, because it did not maintain the ring structure after reductive amination based derivatization.

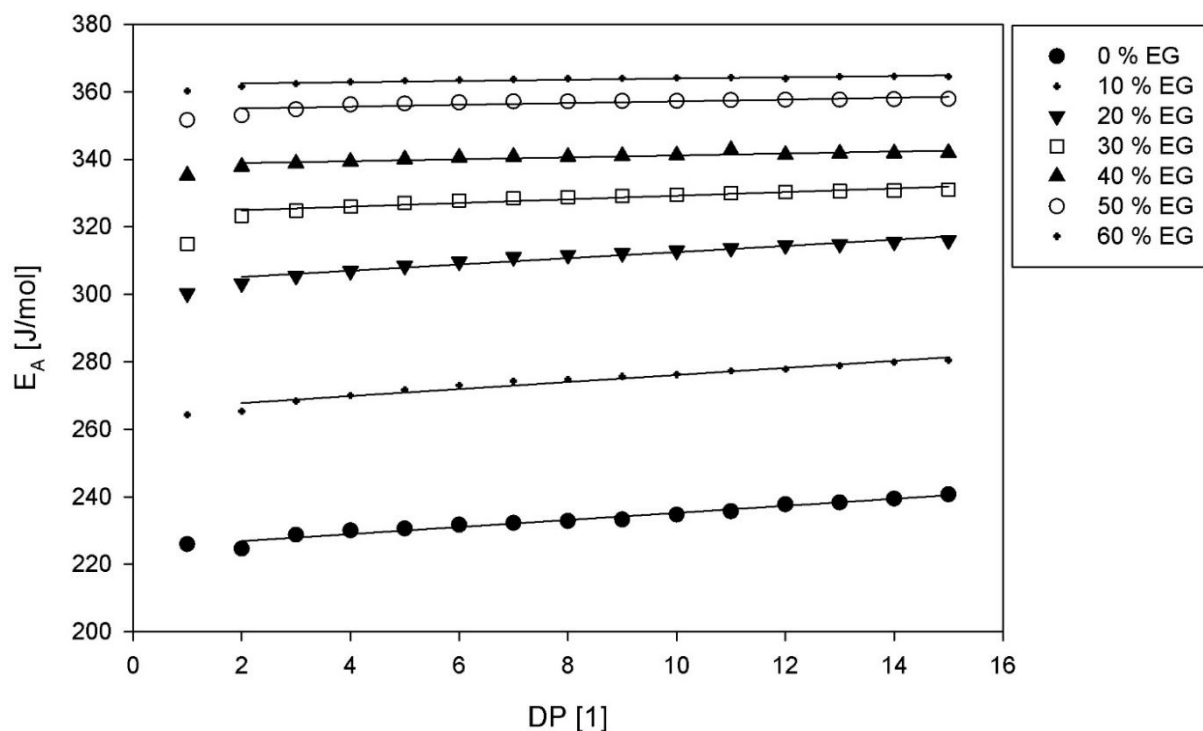
**A****B****C**

**Figure 21.** Arrhenius plots of logarithmic electrophoretic mobility vs. reciprocal absolute temperature. Panel A: no viscosity modifier or polymeric additive containing BGE; Panel B: 10% EG viscosity modifier containing BGE; Panel C: 2% LPA (MW 10 000) polymeric additive containing BGE. The solid lines illustrate the linear least-squares fit of the data. Adapted from [179].

### 5.5.2. Effect of Viscosity-Modifier (Ethylene Glycol, EG)

The separation of APTS-labeled maltooligosaccharides was accomplished by CGE at 20°C and 50°C, using a simple BGE of 25 mM lithium acetate (pH 4.75) with 0 – 60% ethylene glycol (EG) in order to investigate the effect of viscosity modifier on structure specific glycan migration. Figure 20, panel C and D represent the CGE-LIF electropherograms with 10% EG containing BGE in the temperature range of 20°C to 50°C, respectively. Similarly to panel A and B, faster migration were observed at 50°C in both instances. To calculate the  $E_a$  values (Figure 21B), the Arrhenius plots were plotted similarly as mentioned above.

Figure 22 shows the  $E_a$  changes as the function of the DP of the maltooligosaccharides with increasing EG concentration from 10 – 60% delineating a monotonous increase of about 72-13 J/mol per glucose unit, respectively. It should be emphasized that the  $E_a$  values slightly increased with the size of the analyte molecules, when the BGE of 25 mM lithium acetate contained more than 40% EG.



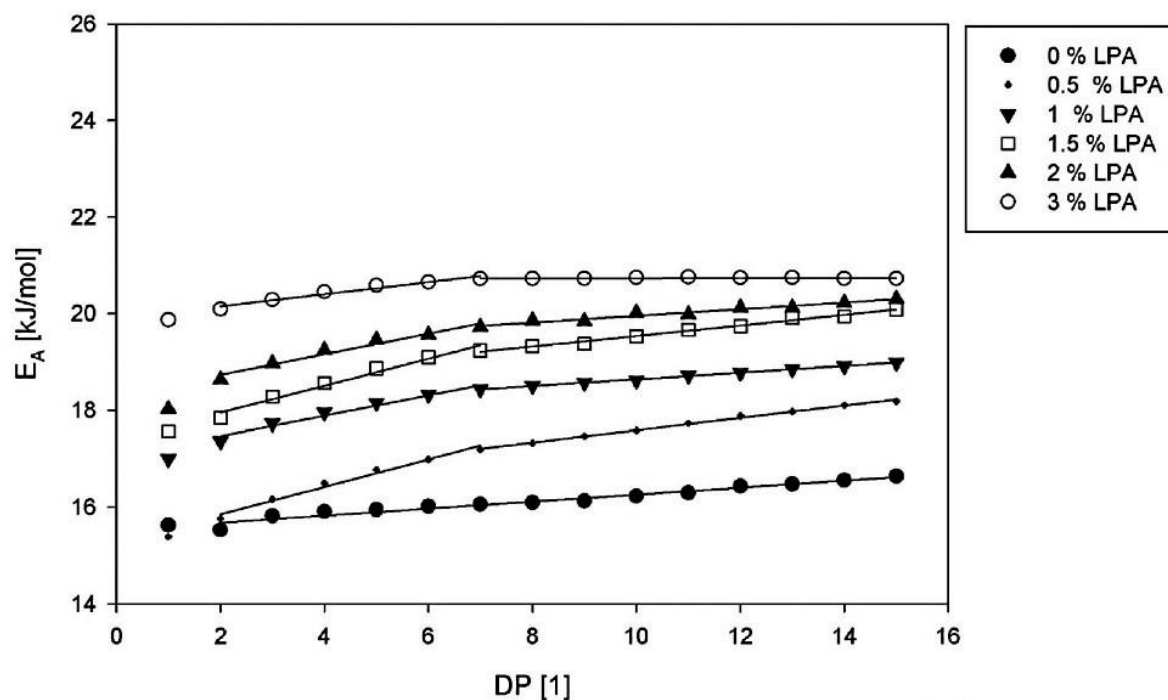
**Figure 22.** Calculated  $E_a$  values as a function of the DP of the linear maltooligosaccharides in increasing EG viscosity modifier (0 – 60%) containing BGE of 25 mM lithium acetate (pH 4.75). Adapted from [179].

### 5.5.3. The Effect of Polymeric Additive (Linear Polyacrylamide, LPA)

First the entanglement threshold value of the LPA solution was measured as published earlier [52] and found to be 1.5%. Based on this information, LPA polymeric additive was used in the concentration range of 0.5 – 3.0%, i.e., below (0.5 and 1.0%, considered as diluted solution), at (1.5%) and above (2.0% and 3.0% entangled polymer network) the threshold. Molecular sieving, in the traditional sense of gel electrophoresis based separation system [7], was not considered in these experiments because of the mass to charge ratios of the growing DP maltooligosaccharides were monotonically increasing, supporting zone electrophoresis-based migration.

Please note that the viscosity of 2% LPA solution was very similar to the 10% EG containing BGE at 20°C and 50°C, i.e. 2% LPA (20°C) = 1.30 mPa.s; 2% LPA (50°C) = 0.82 mPa.s and 10% EG (20°C) = 1.28 mPa.s; 10% EG (50°C) = 0.74 mPa.s.

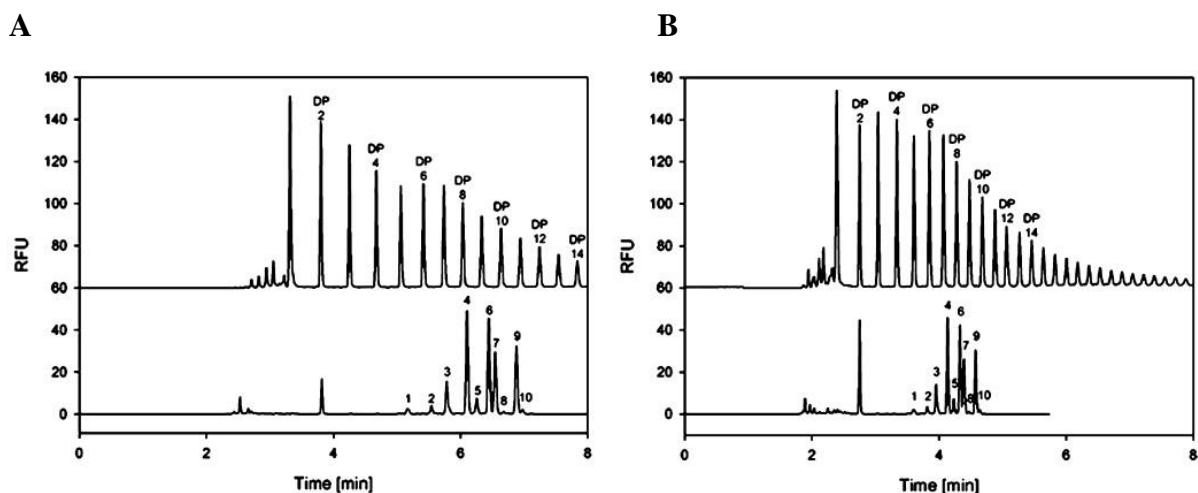
Figure 20, panel E and F compare the CGE-LIF separation traces of the APTS-labeled linear maltooligosaccharides using 2% LPA polymeric additive containing lithium acetate BGE in the temperature range of 20°C and 50°C. To calculate the  $E_a$  values (Figure 21C), the Arrhenius plots were plotted similarly as mentioned above. Figure 23 represents the  $E_a$  changes as the function of the DP of the maltooligosaccharides with increasing LPA concentration from 0 – 3.0%. Interestingly, clear breakpoints were noted in practically all the plots around DP 7 (maltoheptaose). For oligosaccharides shorter than DP 7 the  $E_a$  values increased by the rate of 212 J/mol per glucose unit, while the  $E_a$  values only increased with the rate of 68 J/mol per glucose unit for oligosaccharides longer than DP 7. It has been reported earlier, that the glucose building blocks in maltooligosaccharides linked by  $\alpha$ 1-4 linkage form a full turn of a helix around 7 glucose units (DP 7) [120], therefore these breakpoints in the  $E_a$  plots suggest a conformational change in the linear carbohydrate chains. This conformation change was considered to be the reason for this 144 J/mol per glucose unit transition in the  $E_a$  values. Therefore, oligosaccharides with DP < 7 migrate in random coil configuration, while at and above DP 7, they migrate as elongated helical structures under the conditions applied. However, the shorter oligosaccharides (DP < 7) may have their less hydrophilic side exposed (no helix yet) that might interact with the  $-(CH_2-CH(CO-NH_2)-CH_2)$ -LPA chains in the background electrolytes, resulting in higher  $E_a$  requirement. In case of higher polymeric additive concentrations (2.0 – 3.0%), shallower  $E_a$  slopes were observed for both segments, because the polymer network should be more entangled, therefore less non-networked LPA chains are available to interact with the solute molecules. At the entire range of 0 – 3.0% polymeric additive concentration, the slope values were positive, suggesting no analyte and/or no network deformation.



**Figure 23.** Calculated  $E_a$  values as a function of the DP of the linear maltooligosaccharides in increasing LPA polymeric additive (0 – 3.0%) containing BGE of 25 mM lithium acetate (pH 4.75). Adapted from [179].

#### 5.5.4. Activation Energy Associated with the Electromigration of Branched Oligosaccharides

Next, the activation energy associated with the electromigration of branched oligosaccharides was determined from the electrophoretic mobilities of the sample components in the temperature range of 20 – 50°C. Figure 24 represents the CGE-LIF separation traces for the two sample types (upper trace: maltooligosaccharides; lower trace: IgG glycans) at the separation temperatures of 20°C (panel A) and 50°C (panel B). Interestingly, the migration time differences between all the branched (lower trace: peaks 1-10 showing sialylated, neutral and core fucosylated biantennary IgG glycans) and linear structures (upper trace: peaks DP1 – DP15) decreased with the increase of the separation temperature. For example, F(6)A2 glycan derived from IgG (see peak 4) migrated close to the maltooctose (DP 8) at 20°C, but migrated slightly slower than the maltoheptaose (DP 7) at 50°C as shown in Figure 24.



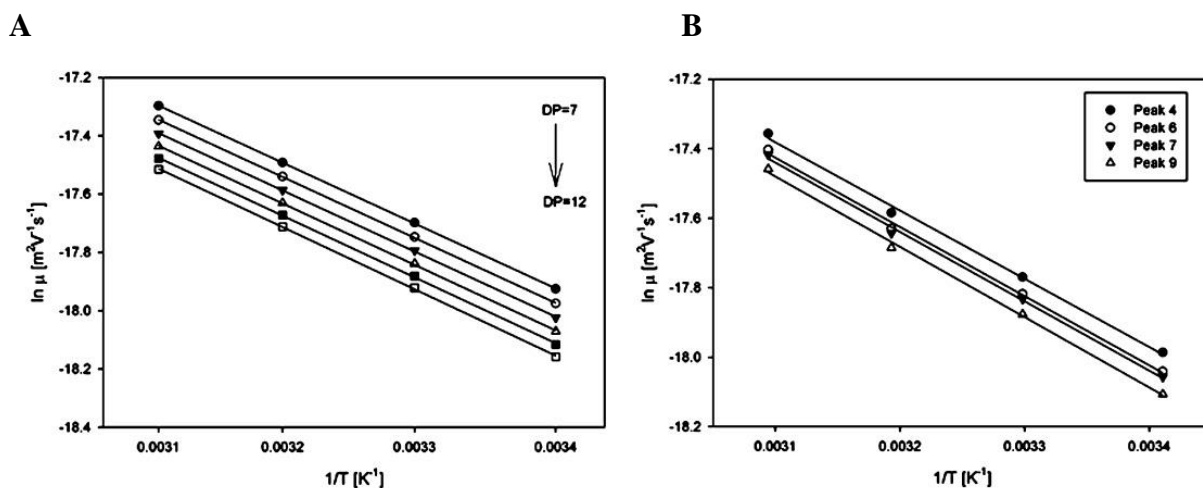
**Figure 24.** Representative electropherograms of APTS-labeled linear maltooligosaccharides (upper traces) and branched IgG N-glycans (lower traces) at the separation temperatures of 20°C (panel A) and 50°C (panel B). Numbered peaks correspond to the abbreviated structures listed in Table 9. Separation conditions: capillary: 50  $\mu\text{m}$  i.d., effective length of 21 cm (total length of 31 cm); BGE: 25 mM lithium acetate (pH 4.75);  $E=400$  V/cm; Sample injection: 0.5 psi/5 s; Separation temperature: as indicated. Adapted from [180].

The corresponding GU values of the IgG N-glycans were calculated based on equation 19 (see in section 2.9.1.) and were shown in the upper left panel of Table 9 (column: 25 mM Li-acetate). In the separation temperature range of 20 – 50°C, the average GU values shift of the IgG N-glycans was -0.89 GU, i.e., the complex biantennary structures migrated that much faster than their corresponding linear maltooligosaccharide counterparts at higher temperatures. As a first approximation, the  $E_a$  requirement for the electromigration of these structurally different sample types was considered to play an important role. To calculate the  $E_a$  values, the Arrhenius diagrams were plotted for the relevant DP range of the linear maltooligosaccharides (DP 7–12) and the major biantennary IgG N-glycans (peaks 4, 6, 7, 9) using equation 13 as shown in Figure 25, panel A and B, respectively. It is important to note that the glucose building blocks in maltooligosaccharides with  $\alpha$ 1-4 linkage form a full helical turn at DP 7 as mentioned above, therefore, their structure was considered as linear-helical at the range of this comparative study with the IgG N-glycans [120].

**Table 9.** GU values shifts of the IgG N-glycans as the function of separation temperature using 25 mM lithium acetate BGE with and w/o monomeric (10% EG viscosity modifier) as well as polymeric additives (2% LPA and 0.4% PEO). N-glycan nomenclature as proposed by Harvey et al. [150], and the major structures are highlighted in bold. Adapted from [180].

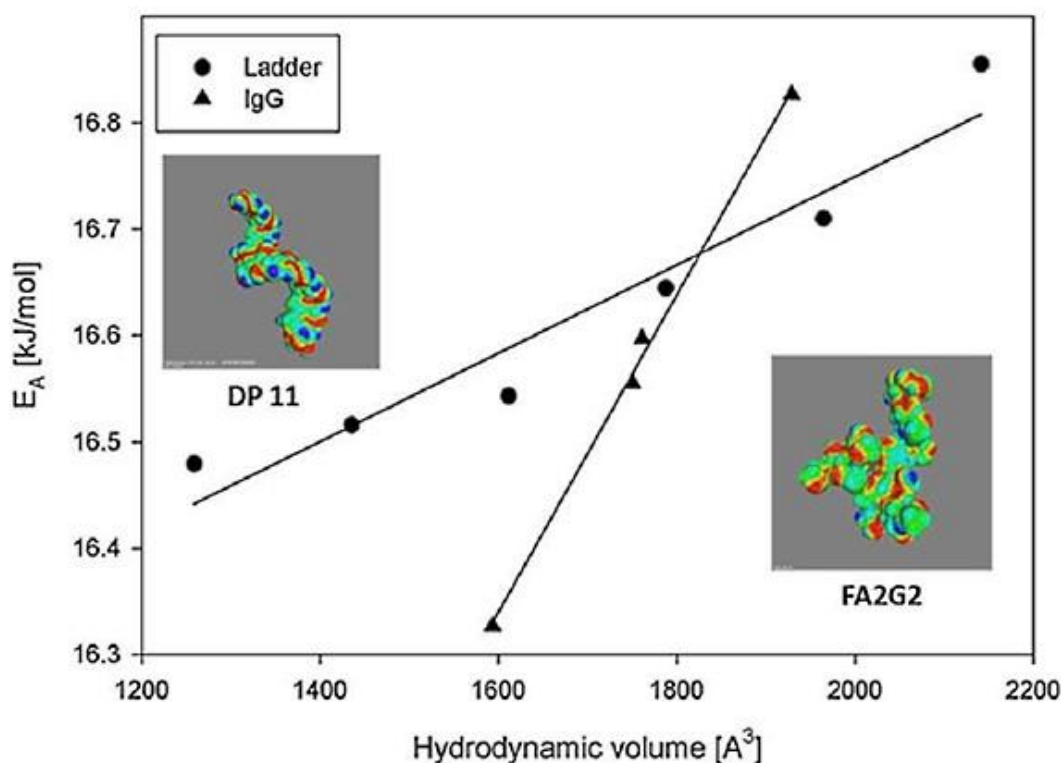
GU values		25 mM Li-acetate							25 mM Li-acetate + 10% EG						
Peak No.	Sample ID	20°C	30°C	40°C	50°C	RSD%	$\Delta$		20°C	30°C	40°C	50°C	RSD%	$\Delta$	
1	F(6)A2G(4)2S2	5.22	4.95	4.86	4.95	3.12	-0.27		6.29	6.29	6.20	6.09	1.55	-0.20	
2	F(6)A2[3]G(4)1S1	6.31	5.91	5.75	5.83	4.23	-0.48		6.48	6.50	6.41	6.30	1.41	-0.18	
3	F(6)A2G(4)2S1	7.07	6.58	6.34	6.44	4.90	-0.63		7.00	7.01	6.89	6.76	1.66	-0.24	
<b>4</b>	<b>F(6)A2</b>	<b>8.14</b>	<b>7.50</b>	<b>7.17</b>	<b>7.28</b>	<b>5.74</b>	<b>-0.86</b>		<b>8.00</b>	<b>7.99</b>	<b>7.82</b>	<b>7.62</b>	<b>2.24</b>	<b>-0.37</b>	
5	F(6)A2B	8.65	7.97	7.62	7.74	5.78	-0.91		8.50	8.50	8.33	8.11	2.21	-0.39	
<b>6</b>	<b>F(6)A2[6]G(4)1</b>	<b>9.27</b>	<b>8.53</b>	<b>8.11</b>	<b>8.23</b>	<b>6.12</b>	<b>-1.04</b>		<b>9.07</b>	<b>9.04</b>	<b>8.85</b>	<b>8.61</b>	<b>2.41</b>	<b>-0.47</b>	
<b>7</b>	<b>F(6)A2[3]G(4)1</b>	<b>9.61</b>	<b>8.85</b>	<b>8.40</b>	<b>8.54</b>	<b>6.10</b>	<b>-1.07</b>		<b>9.38</b>	<b>9.35</b>	<b>9.16</b>	<b>8.91</b>	<b>2.34</b>	<b>-0.47</b>	
8	A2BG(4)2	10.03	9.25	8.79	8.93	5.98	-1.10		9.81	9.78	9.59	9.33	2.28	-0.48	
<b>9</b>	<b>F(6)A2G(4)2</b>	<b>10.71</b>	<b>9.82</b>	<b>9.32</b>	<b>9.45</b>	<b>6.37</b>	<b>-1.26</b>		<b>10.41</b>	<b>10.37</b>	<b>10.15</b>	<b>9.86</b>	<b>2.48</b>	<b>-0.55</b>	
10	F(6)A2BG(4)2	11.02	10.12	9.61	9.77	6.25	-1.26		10.73	10.70	10.49	10.19	2.34	-0.53	
<i>11</i>	<i>Average</i>					<b>5.46</b>	<b>-0.89</b>						<b>2.09</b>	<b>-0.39</b>	
GU values		25 mM Li-acetate + 2% LPA							25 mM Li-acetate + 0.4% PEO						
Peak No.	Sample ID	20°C	30°C	40°C	50°C	RSD%	$\Delta$		20°C	30°C	40°C	50°C	RSD%	$\Delta$	
1	F(6)A2G(4)2S2	4.98	5.14	5.25	5.26	2.49	0.28		4.52	4.70	4.81	4.90	3.44	0.38	
2	F(6)A2[3]G(4)1S1	5.94	6.13	6.23	6.22	2.18	0.28		5.49	5.67	5.74	5.93	3.23	0.45	
3	F(6)A2G(4)2S1	6.72	6.91	7.01	7.00	1.94	0.28		6.17	6.36	6.43	6.65	3.11	0.48	
<b>4</b>	<b>F(6)A2</b>	<b>7.57</b>	<b>7.67</b>	<b>7.79</b>	<b>7.76</b>	<b>1.30</b>	<b>0.19</b>		7.25	7.42	7.45	7.64	2.15	0.39	
5	F(6)A2B	8.02	8.20	8.35	8.32	1.82	0.30		7.79	7.96	7.98	8.17	1.96	0.38	
<b>6</b>	<b>F(6)A2[6]G(4)1</b>	<b>8.67</b>	<b>8.85</b>	<b>8.92</b>	<b>8.88</b>	<b>1.26</b>	<b>0.21</b>		8.36	8.52	8.51	8.70	1.64	0.34	
<b>7</b>	<b>F(6)A2[3]G(4)1</b>	<b>9.03</b>	<b>9.20</b>	<b>9.27</b>	<b>9.21</b>	<b>1.12</b>	<b>0.18</b>		8.73	8.89	8.86	9.07	1.58	0.34	
8	A2BG(4)2	9.43	9.63	9.71	9.64	1.22	0.21		9.18	9.33	9.31	9.53	1.57	0.35	
<b>9</b>	<b>F(6)A2G(4)2</b>	<b>10.16</b>	<b>10.39</b>	<b>10.40</b>	<b>10.30</b>	<b>1.05</b>	<b>0.14</b>		9.76	9.90	9.87	10.07	1.28	0.31	
10	F(6)A2BG(4)2	10.42	10.66	10.72	10.65	1.22	0.22		10.10	10.26	10.23	10.45	1.40	0.35	
<i>11</i>	<i>Average</i>					<b>1.56</b>	<b>0.23</b>						<b>2.13</b>	<b>0.38</b>	





**Figure 25.** Arrhenius plots of logarithmic electrophoretic mobility vs. reciprocal absolute temperature for (A) maltooligosaccharides of DP 7-12 and (B) the four major IgG N-glycans (peak 4, 6, 7 and 9). BGE: 25 mM lithium acetate (pH 4.75) with no viscosity modifier or polymeric additive. The solid lines illustrate the linear least-squares fit of the data ( $r^2 = 0.99$ , each). Adapted from [180].

To shed light on the possible effect of separation temperature mediated electromigration differences on structure specific glycan migration, the  $E_a$  values were plotted against the hydrodynamic volumes of the analyte molecules as shown in Figure 26. To define the size of the branched sugar oligomers, the hydrodynamic volumes were calculated, because the simple DP value was not able to represent the shape of the oligosaccharides. Molecular mechanics optimized structures were derived from the Glycoscience.de database [181]. In order to gain the realistic molecular volume, a subsequent single-point calculation was accomplished by the Density Functional Theory (DFT)/ Conductor-like Screening Model (COSMO) [182]. Figure 26 shows the changes in  $E_a$  requirement of the analyte molecules to migrate through the background electrolyte for the maltoundecaose (G11) and the F(6)A2G(4)2 glycan. The  $E_a$  increment was 0.4 J/mol per  $\text{\AA}^3$  for the linear maltooligosaccharides, in contrast to the biantennary branched IgG N-glycans increased by almost four times as much, i.e. 1.5 J/mol per  $\text{\AA}^3$ . It can be considered that behind these differences is the change of GU values with increasing separation temperature during CE analysis of these structurally different oligosaccharides. In the 1990s, Dovichi *et al.* introduced a similar secondary structure change-induced difference theory in  $E_a$  requirement for DNA sequencing fragments, but in the case of nucleic acids the change was reportedly due to segment differences [183].



**Figure 26.**  $E_a$  plots as a function of the hydrodynamic volume of the linear maltooligosaccharides (DP 7–12) (●) and the major biannetary IgG N-glycans (peak 4, 6, 7 and 9) (▲). Solid lines represent the linear least-squares fit of the data (maltooligosaccharides  $r^2 = 0.93$ ; IgG  $r^2 = 0.99$ ). Molecular structures of the linear maltoundecaose (DP 11, left figure) and branched F(6)A2G(4)2 N-glycan (right figure) represent the charge distribution on the surface of the molecular cavity calculated by COSMO. Adapted from [180].

#### 5.5.5. The Effect of Background Electrolyte Additives

CGE-LIF separation of APTS-labeled linear and branched carbohydrates was also accomplished in the temperature range of 20 – 50°C using 10% EG, 2% LPA (MW 10 kDa) and 0.4% PEO (300 kDa) additive containing background electrolytes. The viscosity values of the additive containing separation buffers at 20°C and 50°C were mentioned above.

Similar to as described above, faster migration was noticed in all three additive containing separation buffers with increasing separation temperature, and various temperature-dependent differential migration of both structures as illustrated in the respective panels in Table 9.

While in the presence of additive-free background electrolyte the corresponding average GU values shift of the branched glycans were -0.89 GU, this value changed to -0.39, +0.23 and +0.38 GU unit by the addition of 10% EG, 2% LPA and 0.4% PEO, respectively, in the temperature range of 20 – 50°C. When background electrolytes containing no or only monomeric (ethylene glycol) additives were used, the GU values of the branched biantennary IgG *N*-glycans decreased with increasing separation temperature, however, with background electrolytes containing polymeric additives the GU values increased with increasing temperature, probably due to possible solute-network interactions or deformations effects in addition to the influence of viscosity.

To calculate the  $E_a$  values, the Arrhenius diagrams were plotted for all additive types. After that, the  $E_a$  vs hydrodynamic volume diagrams were plotted (not shown) and the  $E_a$  changes were derived for the both the linear and branched structures as 0.3 J/mol per Å<sup>3</sup> and 0.9 J/mol per Å<sup>3</sup>, 0.4 J/mol and 1 J/mol per Å<sup>3</sup>, as well as 0.06 J/mol and 0.3 J/mol per Å<sup>3</sup> for the background electrolytes containing EG, LPA and PEO. In the presence of the monomeric (ethylene glycol) additive, the  $E_a$  requirement was presumable associated with the passage of the increasing size analyte molecules through the higher viscosity background electrolyte. In the instance of polymeric additive, in addition to the viscosity component, other physical interactions between the analyte molecules and the polymer network and/or deformation effects should also be considered.

## 6. DISCUSSION

### *Large-Scale Genomic DNA Analysis by CGE*

Currently, most bioanalytical laboratories still utilize manual slab gel based electrophoresis techniques, which have routinely been used for checking DNA properties [184]. However, these methods are time consuming and labor intensive, also requiring improvements in terms of resolving power and analysis throughput [184]. On the other hand, rapid analysis of ssDNA or dsDNA fragments is a need in most molecular biology laboratories such as restriction fragment mapping, PCR fragment analysis, mutation detection and DNA sequencing. Combining the sensitivity of fluorescence detection with the variability of replaceable polymer sieving and the opportunity of automation in both sample handling and data acquisition, CGE has become an attractive alternative to traditional slab gel electrophoresis [185]. In the first part of this study, the aim was to introduce a novel single-channel CGE system with LED-induced fluorescence detection utilizing a pen-shaped capillary cartridge for automatic separation of samples from a 96-well plate. The separation performance of this instrument was demonstrated by rapid and large-scale purity analysis of a thousand gDNA samples, exhibiting excellent migration time reproducibility (RSD <0.75%) and detection limit of ~0.1 ng/μL in comparison to manual agarose slab gel electrophoresis. In addition, the instrument quickly distinguished between degraded and intact gDNA samples, thus provided crucial information if they could be used for downstream quantitative PCR processing where high-quality intact gDNA was the key.

### *Molecular Haplotype Analysis by CGE*

Wolframin (WFS1) is a transmembrane protein in the ER, which is generated at higher levels in pancreatic beta cells and specific neurons in the central nervous system [152]. Genetic variations in the WFS1 gene have been described to be associated with the Wolfram syndrome or type 2 diabetes mellitus. Simultaneous study of multiple polymorphisms (i.e. haplotyping) is getting more and more attention in the analysis of genetic variations of complex diseases. Direct molecular haplotyping is of particular importance in the case of double heterozygote samples, since in these instances the haplotype structure cannot be constructed simply based on genotype data. Our group successfully reported on genotyping and haplotyping in the dopamine D4 receptor gene by CGE-LIF separations [80]. However, currently, there are no sensitive and powerful platforms in most bio-analytical laboratories for genotype and haplotype analysis by CGE-LEDIF system.

Thus in the second part of this study, the aim was to elaborate an efficient double-tube allele-specific PCR-based approach in conjunction with ultrafast CGE-LEDIF system for direct haplotyping of the SNPs in two important miRNA-binding sites (*rs1046322* and *rs9457*) in the WFS1 gene. In addition, the separation performance was demonstrated by ultrafast (<240 s) and accurate (2.4 – 9.2%) sizing analysis of multiplex PCR samples, exhibiting excellent detector linearity ( $R^2=0.9997$ ) with a dynamic quantitation range of 0.08 – 10.0 ng/ $\mu$ L. Moreover, the detection limit was 0.002 ng/ $\mu$ L using field amplified injection from water diluted samples. In conclusion, this CGE-LEDIF system offers a sensitive and easy to use bio-analytical tool for automated haplotyping of a large number samples in clinical settings.

### ***Immune Response against Carbohydrate Antigens***

Interactions of proteins or lipids with glycans play a significant role in the antigen and/or pathogen recognition machinery, malignant transformation and neurological disorders, etc. Today, antiglycan antibody development against carbohydrate antigens is of increasing importance since glycosylation is recognized as a significant player in biomarker research and discovery [166]. Therefore, the aim was to investigate polyclonal antibody response for newly synthesized maltose-BSA conjugate neoglycoproteins. In the third part of this study, selective antibody binding was demonstrated to the synthesized carbohydrate antigens with different glycosylation degrees (low and high) suggesting the possible use of this approach to generate antiglycan antibodies. Moreover, the polyclonal antibody response was not inhibited by maltose or other simple carbohydrates such as glucose, isomaltose, lactose, galactose, and maltodextrin to confirm presence of the neoglycoprotein-specific antibodies. It remains to be determined, however, whether the new immunogenic epitopes included the sugar component or merely represented conformational changes of the core polypeptide chain induced by the sugar conjugation process. To the best of my knowledge there is no report suggesting that the core protein structure was changed by glycosylation changes or some polyclonal antibody response to glycans with detectable specificity directed strictly to the sugar moiety. In the future, this elaborated method may be utilized for conjugation of complex disease specific sugar structures to carriers in order to generate monoclonal antibody libraries.

### ***Generation of an ANTS-labeled N-Glycan Database for CGE Analysis***

There is an increasing trend to develop therapeutic glycoproteins, such as monoclonal antibodies, which require sensitive and high-resolution bioanalytical techniques for comprehensive carbohydrate characterization. Protein derived glycans in most instance have a large variety of structural diversity such as positional and/or linkage isomers, which calls for the need of state-of-the-art instrumentation for their analysis. Thus, high-performance analytical techniques including CGE is a well-established tool, which plays a significant role in the structural elucidation of glycosylation in the biomedical and pharmaceutical fields. On the other hand, the time consuming carbohydrates sequencing step in most *N*-glycan profiling studies can be avoided with the use of a comprehensive glycan glucose unit (GU) database [186], which is specific for a given instrument and gel-buffer system. Our laboratory reported the successful establishment of an APTS-labeled *N*-glycan database released from human serum polyclonal IgG using LIF detection [108]. However, currently there are no well-established, automated, and high-throughput bioanalytical platforms for CGE-LEDIF systems to provide rapid analysis of biopharmaceutical or biotechnology samples of interest without the need of additional carbohydrate sequencing steps. Therefore, the aim was to introduce an initial version of a novel GU database for ANTS-labeled *N*-glycans by CGE-LEDIF detection. Comparison of the normalized migration times of the peaks of interest of glycoprotein derived *N*-glycans with carbohydrate standards with known GU values in this database poses a simple and effective way for rapid structural assessment. This database provided 25 *N*-linked glycan structures of mostly biopharmaceutical interest such as for *N*-glycosylation profiling of therapeutic antibodies with rapid (around 200 s) CGE separation times. The validation of the generated GU values in the database was verified with *N*-glycans released from human IgG and bovine ribonuclease B, which showed very little differences between the database values and measured values (i.e. hIgG: <0.14 GU and RNase B: <0.13 GU). On the other hand, glycans are structurally diverse, therefore in some rare instances there are possible co-migrations of species, in which cases exoglycosidase enzyme array based carbohydrate sequencing can be applied for correct structural elucidation [187]. This novel application may provide a broadly applicable bio-analytical tool for rapid glycan analysis of biotechnology and clinical samples.

### ***Effect of Separation Temperature and Background Electrolyte Composition on Structure Specific Glycan Migration in CGE***

For the physicochemical characterization of biopolymers, CGE has been extensively utilized using cross-linked gels and/or linear polymer sieving matrices. However, today almost exclusively linear polymer gels are used in CGE due to the difficulties of working with cross-linked sieving matrices in narrow bore capillary tubings. On the other hand, with the use of linear polymer gels, network dynamics should be considered both below and above their entanglement [188]. In addition, the activation energy concept is often used to investigate possible temperature induced deformations affecting the analyte, the network, or both. Therefore, in the last part of this study, the aim was to investigate the effect of separation temperature on the differential electromigration shift between linear (maltooligosaccharides) and branched (sialylated, neutral and core fucosylated biantennary IgG glycans) carbohydrates in narrow bore capillaries in the range of 20 – 50 °C. To understand the structure specific electrophoretic migration of the different sugar oligomers, the activation energy concept was used in this study. In addition, viscosity modifiers (e.g., 0 – 60% ethylene glycol) and/or polymer additives (e.g., 0 – 3% linear polyacrylamide, 0.4% 300 kDa polyethylene oxide) were added to the background electrolytes in order to investigate this phenomenon. The results have shown that glucose unit value shift were observed with increasing temperature between the linear and branched sugar structures caused by the temperature-dependent activation energy requirement in order to migrate through the polymer network. Therefore, this emphasizes the high importance of tight temperature control during glycan analysis by capillary electrophoresis if glucose unit values from existing databases are used for structural elucidation.

## 7. SUMMARY

In the first part of this study, the design of a single capillary CGE system was introduced with a pen-shaped compact capillary cartridge comprising a novel microball ended fiber optic-based LED-induced fluorescence detection setting. This automated system offered a good and easy-to-use alternative to labor intensive slab gel electrophoresis systems, also featured excellent detection sensitivity, high-resolving power and rapid analysis times for quantitative or qualitative analysis of biopolymers. The separation performance of this instrument was demonstrated by rapid and large-scale purity analysis of close to a thousand gDNA samples, exhibiting excellent migration time reproducibility (RSD <0.75%).

CGE-LEDIF-based method was also applied to the analysis of multiplex PCR amplification for genotyping and haplotyping of two important, adjacent miRNA-binding sites (*rs1046322* and *rs9457*) in the WFS1 gene. The separation performance was also demonstrated by ultrafast (<240 s) and accurate (2.4 – 9.2%) sizing analysis of multiplex PCR samples, exhibiting excellent detector linearity ( $R^2 = 0.9997$ ) with a dynamic quantitation range of 0.08–10.0 ng/ $\mu$ L. The LOD of the system was 0.08 ng/ $\mu$ L for samples in buffer and 0.002 ng/ $\mu$ L for samples in water, this latter was due to the field amplified injection effect.

In the third part of this study, the analysis and polyclonal antibody response for newly synthesized maltose-BSA conjugates was described. First of all, a simple carbohydrate, maltose, was linked to BSA by reductive amination. To conserve the intact annular maltose structure, an aglycone spacer was utilized for the synthesis of neoglycoproteins. The synthesized neoglycoproteins were analyzed by SDS-CGE and the number of conjugated maltose residues was determined by MALDI-TOF MS. The carbohydrate antigens were then evaluated by immunization of BALB/c mice and the polyclonal antibody response was analyzed by ELISA as evidence for the presence of sugar-containing epitope-specific antibodies. Selective antibody binding was demonstrated to the synthesized neoglycoproteins with different (low and high) glycosylation degrees offering the possible use of this approach to generate antibodies. In addition, the polyclonal antibody response to these neo-epitopes was not inhibitable by maltose or other simple and oligomeric sugars like glucose, isomaltose, lactose, galactose, and maltodextrin.



It should be emphasized that I found no report in the literature suggesting that the polyclonal antibodies response would be generated against core protein structure (BSA).

In the fourth part of this study, the establishment of a novel ANTS-labeled *N*-glycan database was introduced for rapid (around 200 s) CGE analysis of complex *N*-linked carbohydrates. The validation of the generated GU values in this database was accomplished by the use of *N*-glycans released from human IgG and bovine pancreatic RNase B. The corresponding database values were in good agreement with the results of glycoprotein derived glycans. For rapid assessment of *N*-glycan profiles, these 25 database entries will prove useful for fellow separationists to analyze their glycan profiles.

In the final section of this study, the activation energy concept with electromigration of linear and branched oligosaccharides was investigated in CGE using viscosity modifier (ethylene glycol) or polymeric additive (linear polyacrylamide, polyethylene oxide) containing BGE in the temperature range of 20 – 50°C. The relationship between the  $E_a$  and the DP of linear maltooligosaccharides in the range of 1–15 was closely scrutinized with special respect to the temperature-dependent GU values shift of branched *N*-glycans. The GU value shifts were probably caused by the temperature-dependent  $E_a$  requirement for the different structures (linear vs. branched), therefore this emphasizes the high importance of tight temperature control during glycan analysis by CE, if GU values from existing databases are used for structural elucidation.

## 8. ÖSSZEFOGLALÁS

A kapilláris gélelektroforézis napjaink egyik legdinamikusabban fejlődő analitikai módszerének tekinthető, melyet legfőképpen nukleinsavak, fehérjék és szénhidrátok elválasztására használnak. Együttműködő partnereink segítségével egy új, egycsatornás, LED-alapú fluoreszcens detektorral felszerelt, kapilláris gél elektroforézis rendszer került bevezetésre a bioanalitika területére. A kapilláris elektroforetikus elválasztást rövid mérési idő, nagy elválasztású hatékonyság, modern detektálási technika és teljes automatizálhatóság jellemzi, szemben a hagyományos lap gélelektroforézissel. Ezen okokból dolgozatom első részében a rendszer megbízhatóságának ellenőrzésére nagy molekulásúlyú genomiális DNS minta (~1000 darab) tisztaságát és degradációs fokát vizsgáltam. Relatív szórása RSD <0,75% volt, amely kiváló reprodukálhatósági paramétereket mutatott.

A dolgozatom második részében haplotipizáló módszerek validálását végeztem el a CGE műszeren a WFS1 gén 3' szabályzó régiójában található *rs1046322* és *rs9457* mikro-RNS kötő SNP-k használatával. A migrációs idő reprodukálhatóság, a kimutatási határ és a detektor linearitás értékeit is meghatároztam összehasonlítva a hagyományos lap gélelektroforézis technikával. A CGE-LEDIF igen hatékony, gyors elemzést tett lehetővé, gyakran 4 percnél is rövidebb elválasztásokkal, valamint pontos mPCR minta bázispár értékeinek számolásával (2.4 – 9.2 %). A készülék érzékenysége hígító pufferben 0,08 ng/μl és desztillált vízben 0,002 ng/μl volt. Fontos megjegyezni, hogy injektáláskor a vízben hígított mPCR minták sokkal nagyobb mintabevitelt eredményeztek, mivel a puffer-ionok nem versenyeztek a mintamolekulákkal, így nagyobb érzékenységet lehetett elérni. Ezzel szemben a pufferrel történt mintahígítással pontosabb CGE méréseket kaptam. Továbbá a detektor linearitási értékek meghatározásakor egyértelmű lineáris detektor választ kaptam 0,08 – 10,0 ng/μl közötti koncentráció tartományban, melynek értéke  $R^2 = 0.9997$  volt.

Dolgozatom harmadik részében a gliko-biomarker kutatás céljából neoglikoproteinek (szénhidrát antigén) szintézisét és azok vizsgálatát végeztem elektroforetikus és tömegspektrometriás módszerekkel. Annak érdekében, hogy megőrizsem a maltóz szerkezetet, a szénhidrátot formil-heptil [7-(1,3-dioxán-2-il)-heptán-1-ol] híd molekulán keresztül kapcsoltam a hordozó fehérjéhez (BSA) redukív aminálással. Az így előállított mesterséges szénhidrát antigének ellen egerekben megtermelt új cukorhoz specifikusan kötődő antitestek kimutatása ELISA teszttel történt.

Az eredmények azt mutatták, hogy immunizálást követően a magasabb szénhidrát tartalmú antigénekre (66 szénhidrát egység/mol) jóval nagyobb immunválasz jött létre, mint az alacsonyabb cukorszámú (32 szénhidrát egység/mol) glikokonjugátum ellen. Ha az antigén mennyiségét lecsökkentettem felező hígításban ( $1,25 - 0,08 \mu\text{g/ml}$ ), akkor nagyobb immunválasz jött létre a szénhidrát antigénre, mint a hordozó fehérjére. Másrésről, nem történt IgG antitesttermelés maltóz és egyéb cukrok, mint például glükóz, izomaltóz, laktóz, galaktóz és maltodextrin hozzáadásával. Így megállapítható, hogy makromolekuláris hordozóhoz kötött szénhidrát molekulák olyan specifikus epitópokat tartalmaznak, amelyek egyediek és alkalmasak specifikus ellenanyag termelésének kiváltására.

A dolgozat negyedik részében fehérjékhez kapcsolódó ANTS fluoreszcens festékkel jelölt *N*-glikán szerkezeti és GU adatbázist vezettem be egy új, egycsatornás, LED- alapú kapilláris gélelektroforézis készülék használatához. Az adatbázis jelenleg 25 oligoszacharidot tartalmaz, ezek többnyire a gyógyszeripar területén használt, mint például terápiás antitestek *N*-glikozilációs szerkezeti vizsgálataihoz szükségesek. Az adatbázis validálását humán immunglobulin G és szarvasmarha ribonukleáz B glikoproteinekről lehasított *N*-glikánokkal végeztem el. Az eredmények megmutatták az elhanyagolható különbséget a glikoproteinekből származó *N*-glikánok és a standard *N*-glikánok GU értékei között (hIgG:  $<0.14$  GU és RNáz B:  $<0.13$  GU). Így a rendelkezésre álló adatbázis megkönnyítheti a kutatók további munkáját, mivel nem szükséges nagyszámú standardot minden elemzés során megfuttatni a kiértékeléshez. Ugyanakkor a CGE-LEDIF igen hatékony, gyors analízist tett lehetővé, gyakran 240 másodpercnél is rövidebb elválasztásokkal.

A dolgozatom utolsó részében lineáris, glükóz egységekből felépülő maltooligoszacharid (DP1 – 15), valamint elágazó láncú – IgG glikoproteinről lehasított – komplex *N*-glikán szerkezetek elektroforetikus mobilitás változásait, nevezetesen a hőmérséklet-függő GU értékek eltolódását, vizsgáltam kapilláris elektroforézissel. Különböző hőmérsékleten ( $20 - 50^\circ\text{C}$ ) és eltérő összetételű puffer rendszerekkel (0 – 60 % etilén-glikol, 0 – 3 % lineáris poliakrilamid, 0,4 % 300 kDa polietilén-oxid) tanulmányoztam az APTS fluoreszcens festékkel jelölt *N*-glikánok elektroforetikus vándorlását. Az eredmények azt mutatták, hogy a GU érték eltolódása feltehetően a különböző mintakomponensek hőmérséklet-függő aktiválási energia eltéréseiből erednek, ezért hangsúlyozni kell a hőmérséklet kontrol használatát minden egyes glikán szerkezetének vizsgálatakor.

## 9. BIBLIOGRAPHY

- [1] Tiselius, A., *Trans. Faraday Soc.* **1937**, 33, 524-531.
- [2] Kerekgyarto, M., Nemeth, N., Kerekes, T., Ronai, Z., Guttman, A., *J. Chromatogr. A* **2013**, 1286, 229-234.
- [3] Li, S.E.Y., Detection Techniques, in *Journal of Chromatography Library*, 1st edn (eds S.E.Y. Li), Elsevier, **2009**, pp. 55-154.
- [4] Andrews, A.T., *Electrophoresis: theory, techniques, and biochemical and clinical applications*, 2nd edn, Clarendon Press, Oxford University Press, Oxford Oxfordshire, New York, **1986**.
- [5] Stokes, G.G., *Trans. Camb. Phil. Soc.* **1985**, 8, 287-305.
- [6] Ferguson, K.A., *Metabolis.* **1964**, 13 Suppl, 985-1002.
- [7] Ogston, A.G., *Trans. Faraday Soc.* **1958**, (54), 1754-1757.
- [8] Grossman, P.D., Menchen, S., Hershey, D., *Genet. Anal. Tech. Appl.* **1992**, 9 (1), 9-16.
- [9] Lumpkin, O.J., Dejardin, P., Zimm, B.H., *Biopolymers* **1985**, 24 (8), 1573-1593.
- [10] de Gennes, P.G., *Scaling concepts in polymer physics*. Cornell University Press, Ithaca, N.Y., **1979**.
- [11] Slater, G.W., Noolandi, J., *Biopolymers* **1989**, 28 (10), 1781-1791.
- [12] Viovy, J.L., Duke, T., *Electrophoresis* **1993**, 14 (4), 322-329.
- [13] Eyring, H., Polanyi, M., *Z. Phys. Chem. Abt. B* **1931**, 12, 279-311.
- [14] Szoke, M., Sasvari-Szekely, M., Guttman, A., *J. Chromatogr. A* **1999**, 830 (2), 465-471.
- [15] Arrhenius, S.A., *Z. Physik. Chem.* **1889**, 4, 96-116.
- [16] Karger, B.L., Cohen, A.S., Guttman, A., *J. Chromatogr.* **1989**, 492, 585-614.
- [17] Nelson, R.J., Paulus, A., Cohen, A.S., Guttman, A., Karger, B.L., *J. Chromatogr.* **1989**, 480, 111-127.
- [18] Guttman, A., *J. Chromatogr. Sci.* **2003**, 41 (9), 449-459.
- [19] Guttman, A. and Schwartz, H.E., Separation of DNA by capillary electrophoresis, in *Capillary Electrophoresis, Theory and Practice*, 2nd edn(eds P. Camilleri), CRC Press, Boca Raton, **1998**, pp. 397-441.
- [20] Ettre, L.S. and Guttman, A., *Lc Gc N. A.* **2004**, 22 (9), 896-904.
- [21] Karger, B.L. and Guttman, A., *Electrophoresis* **2009**, 30 Suppl 1, S196-202.

- [22] Kerekgyarto, M., Guttman, A., Capillary Gel Electrophoresis, in *Analytical Separation Science*, 1st edn (eds J.L. Anderson, A. Berthod, V. Pino Estévez, and A.M. Stalcup), Wiley-VCH Verlag GmbH & Co. KGaA., Weinheim, **2015**. pp. 555–580.
- [23] Stellwagen, N.C., Stellwagen, E., *J. Chromatogr. A* **2009**, 1216 (10), 1917-1929.
- [24] Madabhushi, R.S., Vainer, M., Dolnik, V., Enad, S., Barker, D.L., Harris, D.W., and Mansfield, E.S., *Electrophoresis* **1997**, 18 (1), 104–111.
- [25] Hu, S., Jiang, J., Cook, L.M., Richards, D.P., Horlick, L., Wong, B., Dovichi, N.J., *Electrophoresis* **2002**, 23 (18), 3136-3142.
- [26] Barta, C., Ronai, Z., Sasvari-Szekely, M., Guttman, A., *Electrophoresis* **2001**, 22 (4), 779-782.
- [27] Ganzler, K., Greve, K.S., Cohen, A.S., Karger, B.L., Guttman, A., Cooke, N.C., *Anal. Chem.* **1992**, 64 (22), 2665-2671.
- [28] Zhu, Z., Lu, J.J., Liu, S., *Anal. Chim. Acta* **2012**, 709, 21-31.
- [29] Lin, C., Cotton, F., Boutique, C., Dhermy, D., Vertongen, F., Gulbis, B., *J. Chromatogr. B Biomed. Sci. Appl.* **2000**, 742 (2), 411-419.
- [30] Fruetel, J.A., Renzi, R.F., Vandernoot, V.A., Stamps, J., Horn, B.A., West, J.A., Ferko, S., Crocker, R., Bailey, C.G., Arnold, D., et al., *Electrophoresis* **2005**, 26 (6), 1144-1154.
- [31] Zhang, J., Burman, S., Gunturi, S., Foley, J.P., *J. Pharm. Biomed. Anal.* **2010**, 53 (5), 1236-1243.
- [32] Hjerten, S., *J. Chromatogr.* **1985**, 347 (2), 191-198.
- [33] Bruin, G.J., Huisden, R., Kraak, J.C., Poppe, H., *J. Chromatogr.* **1989**, 480, 339-349.
- [34] Cobb, K.A., Dolnik, V., Novotny, M., *Anal. Chem.* **1990**, 62 (22), 2478-2483.
- [35] Madabhushi, R.S., *Electrophoresis* **1998**, 19 (2), 224-230.
- [36] Albarghouthi, M.N., Stein, T.M., Barron, A.E., *Electrophoresis* **2003**, 24 (7-8), 1166-1175.
- [37] Chiari, M., Cretich, M., Damin, F., Ceriotti, L., Consonni, R., *Electrophoresis* **2000**, 21 (5), 909-916.
- [38] Chiari, M., Cretich, M., Desperati, V., Marinzi, C., Galbusera, C., De Lorenzi, E., *Electrophoresis* **2000**, 21 (12), 2343-2351.
- [39] Cretich, M., Stastna, M., Chrambach, A., Chiari, M., *Electrophoresis* **2002**, 23 (14), 2274-2278.

- [40] Shieh, C.H., *Coated Capillary Columns and Electrophoretic Separation Methods for Their Use*. Beckman Instruments, Inc., U.S.A. Patent Application No. US5462646., **1995**.
- [41] Horvath, J., Dolnik, V., *Electrophoresis* **2001**, 22 (4), 644-655.
- [42] Polikarpov, N., Potolytsyna, V., Bessonova, E., Tripp, S., Appelhans, D., Voit, B., Kartsova, L., *J. Chromatogr. A* **2015**, 1378, 65-73.
- [43] Lu, J.J., Liu, S., *Electrophoresis* **2006**, 27 (19), 3764-3771.
- [44] Huang, X.C., Quesada, M.A., Mathies, R.A., *Anal. Chem.* **1992**, 64 (18), 2149-2154.
- [45] Wani, I. A., Nanomaterials, novel preparation routes, and characterizations, in *Nanotechnology Applications for Improvements in Energy Efficiency and Environmental Management* (eds. M.A. Shah, M.A. Bhat, J. Paulo Davim), Research Essentials Collection, Hershey, **2014**, pp. 478.
- [46] Crego, A.L. and Marina, M.L., UV-Vis absorbance detection in capillary electrophoresis, in *Analysis and Detection by Capillary Electrophoresis* (eds M.L. Marina, A. Ríos, C.-L. Mancha, M. Valcárcel), Elsevier Science, Oxford, **2005**, pp. 225-304.
- [47] Kobayashi, S., Ueda, T., Kikumoto, M., *J. Chromatogr.* **1989**, 480, 179-184.
- [48] Liu, Y., Wang, R., Gao, L., Jia, Z.P., Xie, H., Zhang, J.L., Ma, J., Zhang, A.M., Xie, X.H., *Acta Chim. Sinica* **2011**, 69 (5), 543-547.
- [49] Ye, M., Hu, S., Quigley, W.W., Dovichi, N.J., *J. Chromatogr. A* **2004**, 1022 (1-2), 201-206.
- [50] Guttman, A., *Nature* **1996**, 380 (6573), 461-462.
- [51] Guttman, A., Chen, F.T., Evangelista, R.A., Cooke, N., *Anal. Biochem.* **1996**, 233 (2), 234-242.
- [52] Guttman, A., Cooke, N., Starr, C.M., *Electrophoresis* **1994**, 15 (12), 1518-1522.
- [53] ElRassi, Z., Mechref, Y., *Electrophoresis* **1996**, 17 (2), 275-301.
- [54] Lin, Y.W., Chiu, T.C., Chang, H.T., *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2003**, 793 (1), 37-48.
- [55] Gassmann, E., Kuo, J.E., Zare, R.N., *Science* **1985**, 230 (4727), 813-814.
- [56] Zhang, X., Stuart, J.N., Sweedler, J.V., *Anal. Bioanal. Chem.* **2002**, 373 (6), 332-343.
- [57] Mathies, R.A., Huang, X.C., *Nature* **1992**, 359, 167-169.
- [58] Righetti, P.G., Gelfi, C., D'Acunto, M.R., *Electrophoresis* **2002**, 23 (10), 1361-1374.
- [59] Dodgson, B.J., Krylov, S.N., *Trends Anal. Chem.* **2012**, 33, 23-34.

- [60] Trost, P., Guttman, A., *Anal. Chem.* **1998**, 70 (18), 3930–3935.
- [61] Swerdlow, H., Wu, S.L., Harke, H., Dovichi, N.J., *J. Chromatogr.* **1990**, 516 (1), 61–67.
- [62] Zhang, J., Voss, K.O., Shaw, D.F., Roos, K.P., Lewis, D.F., Yan, J., Jiang, R., Ren, H., Hou, J. Y., Fang, Y., Puyang, X., Ahmadzadeh, H., Dovichi, N. J., *Nucleic Acids Res.* **1999**, 27 (24), e36.
- [63] Dovichi, N.J., Zhang, J., *Angew. Chem. Int. Ed. Engl.* **2000**, 39 (24), 4463–4468.
- [64] Veledo, M.T., Lara-Quintanar, P., de Frutos, M. and Diez-Masa, J.C., Fluorescence detection in capillary electrophoresis, in *Analysis and Detection by Capillary Electrophoresis* (eds M.L. Marina, A. Ríos, C.-L. Mancha, M. Valcárcel), Elsevier Science, Oxford, **2005**, pp. 305–374.
- [65] Bruno, A.E., Gassmann, E., Pericles, N., Anton, K., *Anal. Chem.* **1989**, 61, 876–883.
- [66] Swinney, K. and Bornhop, D.J., *Electrophoresis* **2000**, 21 (7), 1239–1250.
- [67] Malik, A.K. and Faubel, W., *Chem. Soc. Rev.* **2000**, 29 (4), 275–282.
- [68] Rodat-Boutonnet, A., Naccache, P., Morin, A., Fabre, J., Feurer, B., Couderc, F., *Electrophoresis* **2012**, 33 (12), 1709–1714.
- [69] Ruhaak, L.R., Zauner, G., Huhn, C., Bruggink, C., Deelder, A.M., Wuhrer, M., *Anal. Bioanal. Chem.* **2010**, 397 (8), 3457–3481.
- [70] Hurth, C., Lenigk, R., Zenhausern, F., *Appl. Phys. B-Lasers O* **2008**, 93 (2-3), 693–699.
- [71] Nakamura, S. and Kaenders, W., *Laser Focus World* **1999**, 35 (4), 69–75.
- [72] Kerekgyarto, M, Kerekes, T., Tsai, E., Amirkhanian, V.D., Guttman, A., *Electrophoresis* **2012**, 33 (17), 2752–2758.
- [73] Monstein, H.J., Tarnberg, M., Persis, S., Johansson, A.G., *J. Microbiol. Methods* **2014**, 96, 81–83.
- [74] Wang, R., Xie H., Xu, Y.B, Jia, Z.P., Meng, X.D., Zhang, J.H., Ma, J., Wang, J., Wang, X.H., *Biomed. Chromatogr.* **2012**, 26 (3), 393–399.
- [75] Pan, T.Y., Wang, C.C., Shih, C.J., Wu, H.F., Chiou, S.S., Wu, S.M., *Anal. Bioanal. Chem.* **2014**, 406 (22), 5447–5454.
- [76] Choi, H. and Kim, Y., *B. Kor. Chem. Soc.* **2003**, 24 (7), 943–947.
- [77] Bjorheim, J. and Ekstrom, P.O., *Electrophoresis* **2005**, 26 (13), 2520–2530.
- [78] Xue, M.Z., Bonny, O., Morgenthaler, S., Bochud, M., Mooser, V., Thilly, W.G., Schild, L., Leong-Morgenthaler, P.M., *Clin. Chem.* **2002**, 48 (5), 718–728.
- [79] Szantai, E., Ronai, Z., Sasvari-Szekely, M., Bonn, G., Guttman, A., *Clin. Chem.* **2006**, 52 (9), 1756–1762.

- [80] Szantai, E., Szilagyi, A., Guttman, A., Sasvari-Szekely, M., Ronai, Z., *J. Chromatogr. A* **2004**, 1053 (1-2), 241-245.
- [81] Guttman, A., Cooke, N., *Anal. Chem.* **1991**, 63 (18), 2038-2042.
- [82] Kircher, M. and Kelso, J., *Bioessays* **2010**, 32 (6), 524-536.
- [83] Easley, C.J., Karlinsey, J.M., Bienvenue, J.M., Legendre, L.A., Roper, M.G., Feldman, S.H., Hughes, M.A., Hewlett, E.L., Merkel, T.J., Ferrance, J.P., et al., *Proc. Natl. Acad. Sci. U S A* **2006**, 103 (51), 19272-19277.
- [84] Salieb-Beugelaar, G.B., Dorfman, K.D., van den Berg, A., Eijkel, J.C., *Lab. Chip* **2009**, 9 (17), 2508-2523.
- [85] Guttman, A. and Nolan, J., *Anal. Biochem.* **1994**, 221 (2), 285-289.
- [86] Hjerten, S., *J. Chromatogr. A* **1983**, 270, 1-6.
- [87] Cohen, A.S., Karger, B.L., *J. Chromatogr.* **1987**, 397, 409-417.
- [88] Han, J., Singh, A.K., *J. Chromatogr. A* **2004**, 1049 (1-2), 205-209.
- [89] Hatch, A.V., Herr, A.E., Throckmorton, D.J., Brennan, J.S., Singh, A.K., *Anal. Chem.* **2006**, 78 (14), 4976-4984.
- [90] Lo, C.T., Throckmorton, D.J., Singh, A.K., Herr, A.E., *Lab Chip* **2008**, 8 (8), 1273-1279.
- [91] Okada, H., Kaji, N., Tokeshi, M., Baba, Y., *Anal. Sci.* **2008**, 24 (3), 321-325.
- [92] Widhalm, A., Schwer, C., Blaas, D., Kenndler, E., *J. Chromatogr.* **1991**, 549 (1-2), 446-451.
- [93] Verhelst, V., Mollie, J.P., Campeol, F., *J. Chromatogr. A* **1997**, 770 (1-2), 337-344.
- [94] Griebel, A., Rund, S., Schonfeld, F., Dorner, W., Konrad, R., Hardt, S., *Lab Chip* **2004**, 4 (1), 18-23.
- [95] Michels, D.A., Hu, S., Dambrowitz, K.A., Eggertson, M.J., Lauterbach, K., Dovichi, N.J., *Electrophoresis* **2004**, 25 (18-19), 3098-3105.
- [96] Salas-Solano, O., Tomlinson, B., Du, S., Parker, M., Strahan, A., Ma, S., *Anal. Chem.* **2006**, 78 (18), 6583-6594.
- [97] Szekely, A., Szekrenyes, A., Kerekgyarto, M., Balogh, A., Kadas, J., Lazar, J., Guttman, A., Kurucz, I., Takacs, L., *Electrophoresis* **2014**, 35 (15), 2155-2162.
- [98] Szekrenyes, A., Roth, U., Kerekgyarto, M., Szekely, A., Kurucz, I., Kowalewski, K., Guttman, A., *Anal. Bioanal. Chem.* **2012**, 404 (5), 1485-1494.
- [99] Tous, G.I., Wei, Z., Feng, J., Bilbulian, S., Bowen, S., Smith, J., Strouse, R., McGeehan, P., Casas-Finet, J., Schenerman, M.A., *Anal. Chem.* **2005**, 77 (9), 2675-2682.



- [100] Lu, J.J., Zhu, Z., Wang, W., Liu, S., *Anal. Chem.* **2011**, 83 (5), 1784-1790.
- [101] Haselberg, R., de Jong, G.J., Somsen, G.W., *Electrophoresis* **2013**, 34 (1), 99-112.
- [102] Dwek, R.A., *Chem. Rev.* **1996**, 96 (2), 683-720.
- [103] Mechref, Y., *Electrophoresis* **2011**, 32 (24), 3467–3481.
- [104] Lowe, J.B., Marth, J.D., *Annu. Rev. Biochem.* **2003**, 72, 643-691.
- [105] Taylor, A.D., Hancock, W.S., Hincapie, M., Taniguchi, N., Hanash, S. M., *Genome Med.* **2009**, 1 (6), 57.
- [106] Novotny, M.V., Soini, H.A., Mechref, Y., *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2008**, 866 (1-2), 26-47.
- [107] Alley, W.R., Jr., Madera, M., Mechref, Y., Novotny, M.V., *Anal. Chem.* **2010**, 82 (12), 5095-5106.
- [108] Mittermayr, S., Bones, J., Guttman, A., *Anal. Chem.* **2013**, 85 (9), 4228-4238.
- [109] Mittermayr, S., Bones, J., Doherty, M., Guttman, A., Rudd, P.M., *J. Proteome Res.* **2011**, 10 (8), 3820-3829.
- [110] O'Neill, R.A., *J. Chromatogr. A* **1996**, 720, 201–215.
- [111] Anumula, K.R., *Anal. Biochem.* **2006**, 350 (1), 1-23.
- [112] Harvey, D.J., *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2011**, 879 (17-18), 1196-1225.
- [113] Kuo, C.Y., Wang, S.H., Lin, C., Liao, S.K., Hung, W.T., Fang, J.M., Yang, W.B., *Molecules* **2012**, 17 (6), 7387-7400.
- [114] Kerekgyarto, M., Guttman, A., *Electrophoresis* **2014**, 35 (15), 2222-2228.
- [115] Liu, Y., Salas-Solano, O., Gennaro, L.A., *Anal. Chem.* **2009**, 81 (16), 6823-6829.
- [116] Thakur, D., Rejtar, T., Karger, B.L., Washburn, N.J., Bosques, C.J., Gunay, N.S., Shriver, Z., Venkataraman, G., *Anal. Chem.* **2009**, 81 (21), 8900-8907.
- [117] Laroy, W., Contreras, R., Callewaert, *Nat. Protoc.* **2006**, 1 (1), 397-405.
- [118] Ruhaak, L.R., Hennig, R., Huhn, C., Borowiak, M., Dolhain, R.J., Deelder, A.M., Rapp, E., Wührer, M., *J. Proteome Res.* **2010**, 9 (12), 6655-6664.
- [119] Khandurina, J., Anderson, A.A., Olson, N.A., Stege, J.T., Guttman, A., *Electrophoresis* **2004**, 25 (18-19), 3122–3127.
- [120] Mittermayr, S., Guttman, A., *Electrophoresis* **2012**, 33 (6), 1000–1007.
- [121] Breadmore, M.C., *J. Chromatogr. A* **2012**, 1221, 42-55.
- [122] Effenhauser, C.S., Bruin, G.J., Paulus, A., *Electrophoresis* **1997**, 18 (12-13), 2203-2213.

- [123] Jang, Y.C., Jha, S.K., Chand, R., Islam, K., Kim, Y.S., *Electrophoresis* **2011**, 32 (8), 913-919.
- [124] Shi, N., Ugaz, V.M., *Methods Mol. Biol.* **2014**, 1094, 13-24.
- [125] Liu, P., Li, X., Greenspoon, S.A., Scherer, J.R., Mathies, R.A., *Lab Chip* **2011**, 11 (6), 1041-1048.
- [126] Barta, C., Ronai, Z., Nemoda, Z., Szekely, A., Kovacs, E., Sasvari-Szekely, M., Guttman, A., *J. Chromatogr. A* **2001**, 924 (1-2), 285-290.
- [127] Meltzer, R.H., Krogmeier, J.R., Kwok, L.W., Allen, R., Crane, B., Griffis, J.W., Knaian, L., Kojanian, N., Malkin, G., Nahas, M.K., et al., *Lab Chip* **2011**, 11 (5), 863-873.
- [128] Nagata, H., Tabuchi, M., Hirano, K., Baba, Y., *Electrophoresis* **2005**, 26 (14), 2687-2691.
- [129] Spisak, S., Tulassay, Z., Molnar, B., Guttman, A., *Electrophoresis* **2007**, 28 (23), 4261-4273.
- [130] Donczo, B., Kerekgyarto, J., Szurmai, Z., Guttman, A., *Analyst* **2014**, 139 (11), 2650-2657.
- [131] Nie, F.Q., Yamada, M., Kobayashi, J., Yamato, M., Kikuchi, A., Okano, T., *Biomaterials* **2007**, 28 (27), 4017-4022.
- [132] Ge, R., Allen, R.W., Aldous, L., Bown, M.R., Doy, N., Hardacre, C., MacInnes, J.M., McHale, G., Newton, M.I., *Anal. Chem.* **2009**, 81 (4), 1628-1637.
- [133] Guttman, A., Ronai, Z., *Electrophoresis* **2000**, 21 (18), 3952-3964.
- [134] He, B., Tait, N., Regnier, F., *Anal. Chem.* **1998**, 70 (18), 3790-3797.
- [135] Mark, D., Haeberle, S., Roth, G., von Stetten, F., Zengerle, R., *Chem. Soc. Rev.* **2010**, 39 (3), 1153-1182.
- [136] Rech, I., Marangoni, S., Gulinatti, A., Ghioni, M. Cova, S., *Sensor. Actuat. B-Chem.* **2010**, 143 (2), 583-589.
- [137] Rhazi, L., Bodard, A.L., Fathollahi, B., Aussenac, T., *J. Cereal Sci.* **2009**, 49 (2), 272-277.
- [138] He, M., Zeng, Y., Jemere, A.B., Jed Harrison, D., *J. Chromatogr. A* **2012**, 1241, 112-116.
- [139] Rios, A., Zougagh, M., Avila, M., *Anal. Chim. Acta* **2012**, 740, 1-11.
- [140] Zhang, H., Liu, L., Fu, X., Zhu, Z., *Biosens. Bioelectron.* **2013**, 42, 23-30.
- [141] Wang, H., Wang, D., Wang, J., Wang, H., Gu, J., Han, C., Jin, Q., Xu, B., He, C., Cao, L., Wang, Y., Zhao, J., *J. Chromatogr. A* **2009**, 1216 (35), 6343-6347.

- [142] Kamruzzaman, M., Alam, A.M., Kim, K.M., Lee, S.H., Kim, Y.H., Kabir, A.N., Kim, G.M., Dang, T.D., *Biomed. Microdevices* **2013**, 15 (1), 195-202.
- [143] Chen, X., Tang, K., Lee, M., Flynn, G.C., *Electrophoresis* **2008**, 29 (24), 4993-5002.
- [144] Primack, J., Flynn, G.C., Pan, H., *Electrophoresis*, **2011** 32 (10), 1129-1132.
- [145] He, X.W., Chen, Q.S., Zhang, Y.D., Lin, J.M., *Trac-Trends Anal. Chem.* **2014**, 53, 84-97.
- [146] Sin, M.L.Y., Gao, J., Liao, J.C., Wong, P.K., *J. Biol. Eng.* **2011**, 5 (6), 1-21.
- [147] Sasvari-Szekely, M., Gerstner, A., Ronai, Z., Staub, M., Guttman, A., *Electrophoresis* **2000**, 21 (4), 816-821.
- [148] Lee, Y.C., Chemistry of Neoglycoproteins, in Wiley Encyclopedia of Chemical Biology 1st edn (ed T.P. Begley), Wiley-Interscience, New Jersey, **2008**, pp. 1–10.
- [149] Chiesa, C., O'Neill, R.A., *Electrophoresis* **1994**, 15 (8-9), 1132–1140.
- [150] Harvey, D.J., Merry, A.H., Royle, L., Campbell, M.P., Dwek, R.A., Rudd, P.M., *Proteomics* **2009**, 9 (15), 3796–3801.
- [151] Varadi, C., Lew, C., Guttman, A., *Anal. Chem.* **2014**, 86 (12), 5682-5687.
- [152] Takeda, K., Inoue, H., Tanizawa, Y., Matsuzaki, Y., Oba, J., Watanabe, Y., Shinoda, K., Oka, Y., *Hum. Mol. Genet.* **2001**, 10 (5), 477-484.
- [153] Osman, A.A., Saito, M., Makepeace, C., Permutt, M.A., Schlesinger, P., Mueckler, M., *J. Biol. Chem.* **2003**, 278 (52), 52755-52762.
- [154] Takei, D., Ishihara, H., Yamaguchi, S., Yamada, T., Tamura, A., Katagiri, H., Maruyama, Y., Oka, Y., *FEBS Lett.* **2006**, 580 (24), 5635-5640.
- [155] Fonseca, S.G., Fukuma, M., Lipson, K.L., Nguyen, L.X., Allen, J.R., Oka, Y., Urano, F., *J. Biol. Chem.* **2005**, 280 (47), 39609-39615.
- [156] Riggs, A.C., Bernal-Mizrachi, E., Ohsugi, M., Wasson, J., Fatrai, S., Welling, C., Murray, J., Schmidt, R.E., Herrera, P.L., Permutt, M.A., *Diabetologia* **2005**, 48 (11), 2313-2321.
- [157] Wolfram, D.J., Wagener, H.P., *Mayo Clin. Proc.* **1938**, 13, 715-718.
- [158] Barrett, T.G., Bunday, S.E., Macleod, A.F., *Lancet* **1995**, 346 (8988), 1458-1463.
- [159] Rigoli, L., Lombardo, F., Di Bella, C., *Clin. Genet.* **2011**, 79 (2), 103-117.
- [160] Franks, P.W., Rolandsson, O., Debenham, S.L., Fawcett, K.A., Payne, F., Dina, C., Froguel, P., Mohlke, K.L., Willer, C., Olsson, T., Wareham, N.J., Hallmans, G., Barroso, I., Sandhu, M.S., *Diabetologia* **2008**, 51 (3), 458-463.

- [161] Lyssenko, V., Jonsson, A., Almgren, P., Pulizzi, N., Isomaa, B., Tuomi, T., Berglund, G., Altshuler, D., Nilsson, P., Groop, L., Engl. N., *J. Med.* **2008**, 359 (21), 2220-2232.
- [162] Sandhu, M.S., Weedon, M.N., Fawcett, K.A., Wasson, J., Debenham, S.L., Daly, A., Lango, H., Frayling, T.M., Neumann, R.J., Sherva, R., Blech, I., Pharoah, P.D., Palmer, C.N., Kimber, C., Tavendale, R., Morris, A.D., McCarthy, M.I., Walker, M., Hitman, G., Glaser, B., Permutt, M.A., Hattersley, A.T., Wareham, N.J., Barroso, I., *Nat. Genet.* **2007**, 39 (8), 951-953.
- [163] Lewin, B. *Genes VI*, Oxford University Press, Oxford, New York, **1997**.
- [164] Ronai, Z., Szantai, E., Szmola, R., Nemoda, Z., Szekely, A., Gervai, J., Guttman, A., Sasvari-Szekely, M., *Am. J. Med. Genet. B: Neuropsychiatr. Genet.* **2004**, 126B (1), 74-78.
- [165] Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W., Etzler, M.E. (Eds.), *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), **2009**.
- [166] Heimbürg-Molinaro, J., Rittenhouse-Olson, K., *Methods Mol. Biol.* **2009**, 534, 341–357.
- [167] Pazur, J.H., *Adv. Carbohydr. Chem. Biochem.* **1998**, 53, 201–261.
- [168] Monsigny, M., Roche, A.-C., Duverger, E., Srinivas, O., *Carbohydrate-mediated Interactions. 3.23. Neoglycoproteins. Comprehensive Glycoscience. From Chemistry to Systems Biology*, Elsevier, Amsterdam, **2007**.
- [169] Avery, O.T., Goebel, W.F., *J. Exp. Med.* **1931**, 54 (3), 437–447.
- [170] Helling, F., Shang, A., Calves, M., Zhang, S., Ren, S., Yu, R.K., Oettgen, H.F., Livingston, P.O., *Cancer Res.* **1994**, 54 (1), 197-203.
- [171] Fuster, M.M., Esko, J.D., *Nat. Rev. Cancer* **2005**, 5 (7), 526–542.
- [172] Gray, G.R., *Arch. Biochem. Biophys.* **1974**, 163 (1), 426–428.
- [173] Astronomo, R.D., Burton, D.R., *Nat. Rev. Drug Discovery* **2010**, 9 (4), 308-324.
- [174] Kerekgyarto, M., Fekete, A., Szurmai, Z., Kerekgyarto, J., Takacs, L., Kurucz, I., Guttman, A. *Electrophoresis* **2013**, 34 (16), 2379-2386.
- [175] Schwartz, B.A., Gray, G.R., *Arch. Biochem. Biophys.* **1977**, 181 (2), 542-549.
- [176] Nezlin, R., Ghetie, V., *Adv. Immunol.* **2004**, 82, 155-215.
- [177] Raju, T.S., Scallan, B., *Biotechnol. Prog.* **2007**, 23 (4), 964-971.
- [178] Guttman, A., Pritchett, T., *Electrophoresis* **1995**, 16 (10), 1906-1911.

- [179] Kerekgyarto, M., Jarvas, G., Novak, L., Guttman, A. *Electrophoresis* **2016**, 37 (4), 573-578.
- [180] Guttman, A., Kerekgyarto, M., Jarvas, G., *Anal. Chem.* **2015**, 87 (23), 11630-11634.
- [181] Lutteke, T., Bohne-Lang, A., Loss, A., Goetz, T., Frank, M., von der Lieth, C.W. *Glycobiology* **2006**, 16 (5), 71R–81R.
- [182] Klamt, A., *From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, Elsevier: Amsterdam, The Netherlands, **2005**, pp. ix–xi.
- [183] Lu, H., Arriaga, E., Da, Y. C., Figeys, D., Dovichi, N. J. *J. Chromatogr. A* **1994**, 680 (2), 503–510.
- [184] Liu, M. S., Amirkhanian, V. D., *Electrophoresis* **2003**, 24 (1-2), 93–95.
- [185] Muller, O., Minarik, M., Foret, F., *Electrophoresis* **1998**, 19 (8-9), 1436–1444.
- [186] Campbell, M.P., Royle, L., Radcliffe, C.M., Dwek, R.A., Rudd, P.M., *Bioinformatics* **2008**, 24 (9), 1214-1216.
- [187] Guttman, A., *Electrophoresis*, **1997**, 18 (7), 1136–1141.
- [188] Barron, A. E., Soane, D. S., Blanch, H.W., *J. Chromatogr.* **1993**, 652 (1), 3-16.

## 10. ABBREVIATIONS

ANSA	5-amino-2-naphtalenesulfonic acid
ANTS	8-aminonaphthalene-1,3,6-trisulfonic acid
APTS	8-aminopyrene-1,3,6-trisulfonic acid
ASA	allele-specific amplification
BGE	background electrolyte
BSA	bovine serum albumin
CDCE	constant denaturant capillary electrophoresis
CE	capillary electrophoresis
CGE	capillary gel electrophoresis
CPA	cross-linked polyacrylamide
DAD	diode array detector
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
dITP	deoxyinosine triphosphate
DP	degree of polymerization
dsDNA	double-stranded deoxyribonucleic acid
dTTP	deoxythymidine triphosphate
EG	ethylene glycol

ELISA	enzyme-linked immunosorbent assay
EOF	electroosmotic flow
G2	maltose
gDNA	genomic DNA
GlcNAc	<i>N</i> -acetylglucosamine
GU	glucose unit
hIgG	human immunoglobulin G
HILIC	hydrophilic interaction liquid chromatography
HPLC	high-performance liquid chromatography
LEDIF	light-emitting diode-induced fluorescence
LIF	laser-induced fluorescence
LOC	lab-on-a-chip
LOD	limit of detection
LPA	linear polyacrylamide
MALDI-TOF	matrix-assisted laser desorption/ionization
mCGE	multicapillary gel electrophoresis
ME	microchip electrophoresis
miRNA	microRNA
mPCR	multiplex PCR
MS	mass spectrometry

NMR	nuclear magnetic resonance
PA	polyacrylamide
PCR	polymerase chain reaction
PEO	polyethylene oxide
PNGase F	peptide-N4-(N-acetyl- $\beta$ -glucosaminyl)asparagine amidase
PVP	polyvinylpyrrolidone
RMT	relative migration time
RNA	ribonucleic acid
RNase B	ribonuclease B
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SNP	single nucleotide polymorphism
ssDNA	single-stranded deoxyribonucleic acid
TFA	trifluoroacetic acid
TLC	thin layer chromatography
WFS1	Wolframin gene



## **11. KEYWORDS**

Capillary gel electrophoresis

Fluorescence detection

Genomic DNA

Haplotyping

Neoglycoprotein

Glycan analysis

Database

Activation energy

Kapilláris gélelektroforézis

Fluoreszcens detektálás

Genomiális DNS

Haplotipizálás

Neoglikoprotein

Glikán elemzés

Adatbázis

Aktiválási energia

## 12. ACKNOWLEDGEMENTS

First of all, I would like to thank to my parents for their encouragement, patience and above all for giving me the opportunity of the study.

I would thankful to my supervisor **Prof. Dr. András Guttman** for his excellent training and support during my Ph.D. education. Thank you for giving me the opportunity to work in his well-equipped laboratory in a pleasant scientific atmosphere, as well as for providing me to visit pioneer institutes and companies in the United States or Czech Republic.

I am thankful to **Dr. Zoltán Szurmai**, **Dr. János Kerékgyártó** and **Dr. Anikó Fekete** at Faculty of Sciences and Technology, University of Debrecen for sharing their knowledge's with me regarding the carbohydrate synthesis. Special thank **Dr. István Kurucz** and **Prof. Dr. László Takács** at BioSystems International Ltd. for giving me the opportunity to work in their labs, and for teaching me some interesting methods. Furthermore, I am thankful to **Dr. Zsolt Rónai** and **Nóra Németh** at Department of Medical Chemistry, Semmelweis University for sharing their knowledge's with me regarding the haplotype analysis.

I would also like to gratefully acknowledgement all my colleagues at the **Horváth Csaba Laboratory of Bioseparation Sciences** for their teamwork in an excellent atmosphere.

Furthermore, I want to thank to BiOptic Inc., SCIEX and ProZyme for their kind supports and providing me with consumables.

The financial support of the University of Debrecen, the K-81839 OTKA grant, the MTA-PE translation glycomics grant (#97101) and the TÁMOP-4.2.4.A/2-11/1-2012-0001 "National Excellence Program", co-financed by the European Social Fund, is also gratefully acknowledged.



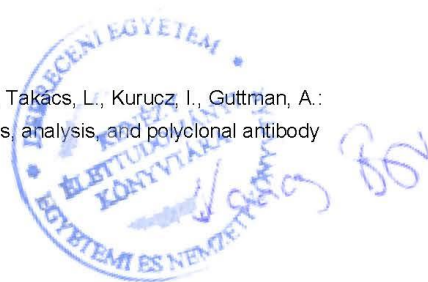
Registry number:  
Subject:

DEENK/222/2016.PL  
PhD Publikációs Lista

Candidate: Márta Kerékgyártó  
Neptun ID: F10PDA  
Doctoral School: Doctoral School of Molecular Medicine  
MTMT ID: 10037473

### List of publications related to the dissertation

1. **Kerékgyártó, M.**, Járvas, G., Novák, L., Guttman, A.: Activation energy associated with the electromigration of oligosaccharides through viscosity modifier and polymeric additive containing background electrolytes.  
*Electrophoresis*. 37 (4), 573-578, 2016.  
DOI: <http://dx.doi.org/10.1002/elps.201500394>  
IF: 2.482 (2015)
2. **Kerékgyártó, M.**, Guttman, A.: Capillary Gel Electrophoresis.  
In: Analytical Separation Science. Ed.: Jared L. Anderson, Alain Berthod, Verónica Pino Estévez, Apryll M. Stalcup, Wiley-VCH Verlag, Weinheim, 555-580, 2015.
3. Guttman, A., **Kerékgyártó, M.**, Járvas, G.: Effect of Separation Temperature on Structure Specific Glycan Migration in Capillary Electrophoresis.  
*Anal. Chem.* 87 (23), 11630-11634, 2015.  
DOI: <http://dx.doi.org/10.1021/acs.analchem.5b03727>  
IF: 5.886
4. **Kerékgyártó, M.**, Guttman, A.: Towards the generation of an aminonaphthalene trisulfonate labeled N-glycan database for capillary gel electrophoresis analysis of carbohydrates.  
*Electrophoresis*. 35 (15), 2222-2228, 2014.  
DOI: <http://dx.doi.org/10.1002/elps.201400054>  
IF: 3.028
5. **Kerékgyártó, M.**, Fekete, A., Szurmai, Z., Kerékgyártó, J., Takács, L., Kurucz, I., Guttman, A.: Neoglycoproteins as carbohydrate antigens: synthesis, analysis, and polyclonal antibody response.  
*Electrophoresis*. 34 (16), 2379-2386, 2013.  
DOI: <http://dx.doi.org/10.1002/elps.201300052>  
IF: 3.161

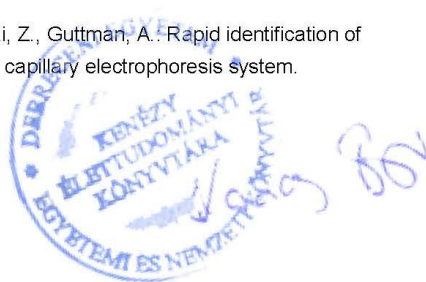




6. **Kerékyártó, M.**, Németh, N., Kerekes, T., Rónai, Z., Guttman, A.: Ultrafast haplotyping of putative microRNA-binding sites in the WFS1 gene by multiplex polymerase chain reaction and capillary gel electrophoresis.  
*J. Chromatogr. A.* 1286, 229-234, 2013.  
DOI: <http://dx.doi.org/10.1016/j.chroma.2013.02.062>  
IF: 4.258
7. **Kerékyártó, M.**, Kerekes, T., Tsai, E., Amirkhanian, V. D., Guttman, A.: Light-emitting diode induced fluorescence (LED-IF) detection design for a pen-shaped cartridge based single capillary electrophoresis system.  
*Electrophoresis.* 33 (17), 2752-2758, 2012.  
DOI: <http://dx.doi.org/10.1002/elps.201200139>  
IF: 3.261

#### List of other publications

8. Járvas, G., **Kerékyártó, M.**, Guttman, A.: On the electromigration of charged-fluorophore labeled oligosaccharides in polyethylene oxide solutions.  
*Electrophoresis. [Epub ahead of print]*, 2016.  
DOI: <http://dx.doi.org/10.1002/elps.201600183>  
IF: 2.482 (2015)
9. Székely, A., Szekrényes, Á., **Kerékyártó, M.**, Balogh, A., Kádas, J., Lázár, J., Guttman, A., Kurucz, I., Takács, L.: Multi Capillary SDS-Gel Electrophoresis for the Analysis of Fluorescently Labeled mAb preparations: a high throughput quality control process for the production of QuantiPlasma and PlasmaScan mAb libraries.  
*Electrophoresis.* 35 (15), 2155-2162, 2014.  
DOI: <http://dx.doi.org/10.1002/elps.201400208>  
IF: 3.028
10. Németh, N., **Kerékyártó, M.**, Sasvári-Székely, M., Rónai, Z., Guttman, A.: Rapid identification of human SNAP-25 transcript variants by a miniaturized capillary electrophoresis system.  
*Electrophoresis.* 35 (2-3), 379-384, 2014.  
DOI: <http://dx.doi.org/10.1002/elps.201300221>  
IF: 3.028





11. Szekrényes, Á., Roth, U., **Kerégyártó, M.**, Székely, A., Kurucz, I., Kowalewski, K., Guttman, A.:  
High-throughput analysis of therapeutic and diagnostic monoclonal antibodies by  
multicapillary SDS gel electrophoresis in conjunction with covalent fluorescent labeling.  
*Anal. Bioanal. Chem.* 404 (5), 1485-1494, 2012.  
DOI: <http://dx.doi.org/10.1007/s00216-012-6213-2>  
IF: 3.659

**Total IF of journals (all publications): 34,273**

**Total IF of journals (publications related to the dissertation): 22,076**

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.

24 August, 2016

