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

Development of Capillary Electrophoresis and Mass Spectrometry Methods for the Characterization of Therapeutic Proteins, as well as N-linked and Free Glycans

by Dániel Sárközy

Supervisor: Dr. András Guttman

UNIVERSITY OF DEBRECEN
DOCTORAL SCHOOL OF MOLECULAR MEDICINE

DEBRECEN, 2024

Development of Capillary Electrophoresis and Mass Spectrometry Methods for the Characterization of Therapeutic Proteins, as well as N-linked and Free Glycans

By Dániel Sárközy, biochemical engineer (MSc)

Supervisor: Dr. András Guttman

Doctoral School of Molecular Medicine, University of Debrecen

Head of the **Defense Committee**:	László Csernoch, PhD, DSc
Reviewers:	Csaba Ágoston, PhD
	Gergő Kalló, PhD
Members of the Defense Committee:	Krisztina Frey, PhD
	János Kádas, PhD

The PhD Defense takes place at the Lecture Hall of Bldg. A, Department of Internal Medicine, Faculty of Medicine, University of Debrecen, 13 December 2024, 13:30.

1. Background and Objectives of the Doctoral Dissertation

The routine application of analytical techniques has become fundamentally important in recent years. This shift is driven by the exponential increase in current biotechnological, food industry, and medical research projects, along with the emergence of strict quality assurance regulations for new active drug ingredients and food components. These developments have posed new challenges for representatives of analytical and molecular biology fields. Next-generation biological products - whether considering large-molecule therapeutic antibodies or small-molecule pharmaceuticals and dietary supplements - exhibit natural structural variability, leading to a certain degree of heterogeneity in these products. To characterize these variants and ensure their safe and effective use, there has been a demand for the development of orthogonally applicable analytical methods in addition to the known and accepted standard approaches. Moreover, continuous development and optimization steps are required to integrate these methods into everyday analytical practice [1, 2].

Based on the available educational materials and scientific publications in the analytical field, it is known that there is a simple, rapid, and high-resolution technique orthogonal to the standard chromatography-based approaches. With its low sample requirement and unparalleled efficiency, this technique is gaining increasing significance in the field of separation science. It is particularly targeted towards industrial applications (e.g., pharmaceutical purity testing) and medical diagnostics (e.g., biomarker identification), as well as providing alternative solutions to analytical problems that classical HPLC-based methods cannot address or can only address with difficulties. This technique is capillary electrophoresis (CE), surrounded by many questions regarding its advantageous properties and applicability for the analysis of biological molecules [3]. During this doctoral research, my goal was to demonstrate the applicability of today's most significant capillary gel electrophoresis (CGE) methods for two major groups of biomolecules: proteins and oligosaccharides. Additionally, I aimed to outline the fundamental method development criteria that address questions related to the reproducibility and efficiency of CGE analysis of the examined biological samples. I also consider it important to present the developed electrophoretic separation media optimized for the analytes studied in my research, as these enabled more efficient analysis for different types of samples.

To investigate certain proteins and carbohydrates using the CGE technique, it is not only necessary to find the appropriate separation principle and buffer systems, but also to select a sensitive detection method. A sufficient detector, in an optimal case, provides a well-defined image of the analytes, enabling their accurate qualitative and/or quantitative determination. Besides the classic ultraviolet absorbance (UV) detection methods, indirect detection methods (e.g., fluorescence-based approaches) have gained preference in recent times, allowing for the highly sensitive detection and characterization of derivatized samples. In my dissertation, I also address detection possibilities based on certain labeling reactions, as well as coupled techniques that provide high-sensitivity and high-resolution structural information directly [4, 5]. Among these, I detail electrospray ionization mass spectrometry (ESI-MS) technique, which is commonly used in modern analytical practice for its standard chromatographic connectivity and precise structural identification capabilities for both small and large molecules. In contrast, the concept I have studied explores the complex yet highly promising coupling of the ESI-MS technique with CGE. The applicability of this combination primarily depends on the composition of the separation media. With this understanding, I aimed to address a long-standing analytical challenge in the biotechnology industry: the "on-line" size-based capillary gel electrophoretic separation of proteins followed by the structural identification via mass spectrometry. The essence of this new approach is to optimize SDS-capillary gel electrophoresis (SDS-CGE) coupled with MS detection, analogous to the classical sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) technique, but allowing for faster and on-line detection with accurate mass-to-charge ratio information on the molecules of interest.

The separation media, methods, and detection modes developed during my work were primarily examined through the characterization of new-generation, protein-based biotherapeutics, as well as N-linked and free (non-protein-bound) oligosaccharides.

2. Novel Scientific Findings of the Dissertation

2.1. Determination of Protein Subunit Integrity and N-Glycan Profile of Therapeutic Antibodies

In my work, I demonstrated the purity analysis of two therapeutic proteins (omalizumab and bamlanivimab) using sodium dodecyl sulfate capillary gel electrophoresis (SDS-CGE) in both their intact and reduced forms. This workflow was performed to verify the purity of the molecules and the integrity of their subunits, with particular focus on the presence and levels of glycosylated and non-glycosylated forms of the heavy chain fragment (HC and ngHC). For the asthma treatment therapeutic containing omalizumab, I established that the drug did not show any presence of the non-glycosylated heavy chain (ngHC) fragment, thus indicating that the entire mAb content carries N-glycosylation modifications on its heavy chains. Both omalizumab and bamlanivimab in their intact forms exhibited longer migration times compared to expected results, which could be explained by the extensive Y-shape of the molecules and their unique migration properties in the dextran and borate-containing separation gel. I also observed the presence of free subunits and their aggregation in each sample, a phenomenon likely resulting from incomplete chemical reactions during product manufacturing. For the bamlanivimab protein, developed against Covid-19, I supplemented my purity tests with charge variant analysis using capillary isoelectric focusing (cIEF) technique. Additionally, I subjected my results to more detailed reproducibility analysis to demonstrate the suitability of the CGE-based mAb measurements.

In case of omalizumab, based solely on the subunit integrity analysis, and supplemented with the results of charge variant analysis for bamlanivimab, I concluded that comprehensive N-glycosylation analysis is necessary to identify their major glycan structures, since the presence or absence of these structures linked to the protein can regulate certain adaptive immune responses. During N-glycan analysis of the anti-asthma product, I observed that only high-mannose sugar structures were afucosylated. Generally, galactose-free IgG molecules have a longer circulation time, attributed to their lower affinity for binding to FcγRI. Furthermore, I found that the quantity of monogalactosylated glycans in the sample correlates with longer serum half-life, while the presence of N-acetyl-glucosamine is associated with faster clearance, indicating the crucial roles of FA2 and FA1 type glycans. Additionally, the presence of high-

mannose glycans, consistent with literature data, likely increases the drug molecule's clearance rate via the mannose receptors [6, 7].

Based on the fucosylation level of the core structures of glycans from the Covid-19 mAb, I inferred that the approximately 8.2% afucosylated glycan content in the product exhibits some ADCC function, which aligns with analytical and immunological conclusions reached through other techniques. Furthermore, based on the obtained results and literature data, it can be assumed that the glycosylation of the FcγRIIIa protein (which provides a binding site for IgG molecules), and the lack of core structure fucosylation in the IgG Fc region are correlated resulting in greater carbohydrate-carbohydrate interactions, which increase the binding strength [8, 9]. Additionally, based on my measurements, a significant portion of the fucosylated biantennary core structures exhibited terminal galactosylation, which typically enhances the mentioned ADCC effector function and also prolongs the molecule's serum half-life [10, 11]. Finally, it was determined that the presence of very low levels (~1.3%) of high mannose structures detected during the glycosylation analysis did not significantly contribute to the reduction in serum half-life, based on literature information [12]. Overall, my CGE-based glycan characterization results associated with therapeutic proteins demonstrate that alongside their known inactivating or neutralizing effects, the inhibition or presence of ADCC function related to glycosylation can be confirmed by this orthogonal analytical technique.

2.2. Development of an Agarose-Based SDS-CGE-UV Method for Characterizing Separation Gel and Migration Properties, as well as Determining Therapeutic Protein Size via the Ferguson Principle

During my agarose-gel development experiments, I elucidated the effect of the agarose and borate concentration ratios on the electrophoretic properties of SDS-protein complexes, using the light and heavy chain subunits of the previously analyzed model therapeutic protein, omalizumab. To understand the migration behavior of the examined fragments, I compared the behavior of the protein's light chain (LC), non-glycosylated heavy chain (ngHC), and heavy chain (HC) using Tris-acetate-EDTA (TAE) and Tris-borate-EDTA (TBE) buffer systems containing 0.2-1.0% agarose. This evaluation involved the utilization of 15 internally developed agarose-gel formulations. Simultaneously, during the optimization process, I created three-dimensional Ferguson diagrams to facilitate the understanding of migration properties and to

explore changes in analytical parameters such as theoretical plate number or efficiency (N), resolution (Rs), and selectivity (α).

Drawing upon the resolving power of the classic dextran and borate-based gel systems in sodium dodecyl sulfate-capillary gel electrophoresis (SDS-CGE) measurements, the applicability of the developed agarose-containing separation media proved to be similar to conventional gels, but their performance and analytical parameters varied concerning the separation of monoclonal antibody fragments. Evaluation of the analytical characteristics of the agarose-SDS-CGE method confirmed that resolution (Rs) could be practically defined as a result of the relationship between selectivity (α) and efficiency (N). For agarose-based sieving matrices, the significant counter-flow electroosmotic flow (EOF) mainly showed dependence on boric acid concentration, while viscosity typically varied with agarose concentration. In light of the calculated parameters, I exemplified how the Ferguson principle could be used to estimate the molecular weight of an unknown-sized component solely from the calculated migration time, EOF, and temperature values obtained from gel electrophoretic runs.

Furthermore, I demonstrated that this gel system, besides ensuring short separation times, requires less preparation between parallel measurements. This was due to the time-consuming process of capillary washing and regenerating steps between every analytical runs, which is required when using the conventional SDS-based separation matrices. These steps can be avoided by using the agarose-based gel composition. Therefore, it can be concluded that refilling the separation capillary with the borate-stabilized agarose media without the necessity for repetitive washing steps, can substantially decrease the downtime between measurements, that are already considered extremely fast (~ 10 minutes). Additionally, regarding the reproducibility of migration times, RSD values below 0.75% can be achieved even without the capillary regeneration steps. Moreover, the method ensures excellent resolution between the antibody subunits.

Development of a CGE Method Coupled with MS Detection for the Analysis of Peptides, Native Proteins and SDS-Protein complexes

Alongside the development of capillary gel electrophoresis methods and separation matrices, I also investigated their compatibility with mass spectrometry (MS) detection. As a result, I developed a simple and widely applicable Coaxial Sheath Flow Reactor Interface (CSFRI) operating on the coaxial principle. This interface enables the coupling of CE and CGE-based separation methods with MS, allowing for more detailed structural information to be obtained from electrophoretically separated components (even using viscous gel matrices). Primarily, this method can facilitate the high-resolution analysis of proteins and peptides using CGE, including SDS-protein complexes, of which molecules size-based separation was previously interpretable only with UV detection. During the development of this interface, I outlined a novel principle for transferring samples from the separation capillary to a conventional ESI source designed by any manufacturer. While the electrical connection in this setup is closed at the outlet of the separation capillary, the separated sample components exit the separation gel and carried towards to the ion source by the sheath liquid through a closed-loop (~200 μm internal diameter) capillary in the flow reaction zone, which provides space for potential post-capillary reactions. Consequently, the interface allows for the coupling of any commercial MS ion source with the CE instrument, requiring only the optimization of the flow rate of the sheath fluid, drying gas, and the temperature of the latter. Due to the design of this interface, the higher liquid flow rate (up to 75 μL/min) compared to the nanospray-type connecting units (enabling nL/min flow rates) results in sample dilution and hence lower sensitivity with the application of the CSFRI unit. Nonetheless, it ensures a much more stable electrospray process throughout the entire separation process (even over 60+ minutes).

2.3. Analysis of Free Oligosaccharides via CGE-LIF based methods

During my doctoral research, I presented a particularly rapid and efficient separation system for the analysis of human milk oligosaccharides (HMOs) using CGE-LIF. In addition to the analysis of HMOs, I also conducted a comprehensive analysis on a galactooligosaccharide (GOS) composition in a newly developed infant formula.

For the HMO measurements, I employed a 12-channel multicapillary gel electrophoresis device (which had not been previously used for the analysis of free oligosaccharides), using two commercially available gel compositions that complement each other's separation capability. First, I optimized a traditional gel matrix used in the industry for carbohydrate separation, followed by a dextran and borate containing gel system, enabling the qualitative and quantitative determination of certain low-abundance oligosaccharides in a human milk sample with the help of the provided HMO standards. My results showed that if structural isomer pairs of 2'-fucosyllactose (2'FL) - 3-fucosyllactose (3FL), and lacto-N-tetraose (LNT) - lacto-N-neotetraose (LNnT) were present together in the sample mixture, the application of both optimized gel systems was justified due to the co-migration phenomenon of the respective isomers. The different properties of the separation matrices allowed for the resolution of co-migration issues. With the traditional carbohydrate separation gel composition, oligosaccharides were separated based on their charge-to-hydrodynamic volume ratios in less than 10 minutes using capillary zone electrophoresis (CZE) mode. However, as for the dextran and borate-containing gel, a different migration order was observed. Due to the high borate content, it was assumed that the gel system could form complexes with the vicinal hydroxyl groups of the sugar molecules, thereby imparting additional charge to the oligosaccharides. Consequently, using this latter gel system, the separation took only 6.5 minutes and mutually corrected the deficiencies observed with the first composition. Specifically, in the latter scenario, the co-migrating 2'FL and 3FL isomer pairs were effectively resolved, whereas for LNT and LNnT, the opposite outcome was observed.

Considering that every HMO molecule has a reducing end, it was observed that the investigated sugar structures reacted with only one labeling (APTS) molecule, facilitating their quantitative analysis with CGE-LIF. This way the calibration curve prepared using the pLNnH standard was easily applicable to other HMO structures as well. With the help of the calibration curve, the concentrations of all known low-abundance (RFU <0.1) oligosaccharides, including five HMO structures (LNT2, DFL, LNnT, LNT, and pLNnH), were determined in a human breast milk sample. The determined values corresponded to concentrations measurable in the third month of lactation according to literature data and the origin of the sample [13].

During my measurements aimed at determining the total GOS content of the complex infant formula containing free oligosaccharides, I presented a rapid and orthogonal CE-based analytical method that offers an alternative solution to the standard LC-based oligosaccharide composition determination, applicable even to samples with extremely high lactose content.

The industrial GOS containing sample was analyzed using four different approaches as follows:

(1) - untreated reference galactooligosaccharide sample,

(2) - sample spiked with lactose,

(3) - sample digested with amyloglucosidase,

(4) - sample digested with amyloglucosidase + β-galactosidase.

During the process, maltodextrins were first removed by enzymatic digestion with amyloglucosidase, and then the GOS content was determined as the difference between the analyses of steps 3 and 4, namely from the results obtained with and without β-galactosidase treatment. Alongside the measurements, a standard lactose calibration curve was prepared for the quantitative determination of identified structures.

Based on the results of my first measurement aiming to determine the complete oligosaccharide profile, it can be concluded that the obtained reference electropherogram yielded the same number of peaks as the previously conducted HPLC analysis on the same sample [14], with the difference that the CGE runs proved to be three times faster between the injections. Furthermore, the suitability of the method is evidenced by the reproducibility of average migration times and peak areas, which showed RSD values below 2.5%. From the observed peak shifts and changes in peak ratios during the enzymatic reactions, I was able to draw clear conclusions about the positions of GOS structures on the electropherograms, enabling to calculate the total GOS content of the sample. Comparing the obtained value of 37.23 g/100 g with the officially accepted value of 37.0 g/100 g determined by the HPLC approach [15] it could be stated that the presented CGE-based method is suitable for the accurate determination of GOS content in infant formulas. Based on my findings, it is reasonable to consider that the electrophoretic technique presented in this study might serve as a viable alternative to chromatographic methods, which typically involve longer analysis times. Alternatively, it could function as an orthogonal method for accurately detecting prebiotic GOS content in a variety of infant formulas and other lactose-containing samples.

3. References

1. István, D.B., *Műszeres analitikai technikák a gyógyszerészi és bioanalitikai vizsgálatokban.* 2011, Debreceni Egyetem Orvos- és Egészségtudományi Centrum Gyógyszerésztudományi Kar Gyógyszerhatástani Tanszék Gyógyszerészi műszeres- és bioanalitikai részleg.
2. Bizottság, E.G.E., *A biológiailag hasonló gyógyszerek helyzete az EU-ban.* 2019.
3. Attila, D.G., *A kapilláris zónaelektroforézis*. 2000, Debreceni Egyetemi Kiadó.
4. Guttman, A., *High-resolution carbohydrate profiling by capillary gel electrophoresis.* Nature, 1996. **380**(6573): p. 461-462.
5. De Hoffmann, E. and V. Stroobant, *Mass spectrometry: principles and applications*. 2007: John Wiley & Sons.
6. Liu, L., *Antibody glycosylation and its impact on the pharmacokinetics and pharmacodynamics of monoclonal antibodies and Fc-fusion proteins.* J Pharm Sci, 2015. **104**(6): p. 1866-1884.
7. Ryman, J.T. and B. Meibohm, *Pharmacokinetics of Monoclonal Antibodies.* CPT: Pharmacometrics & Systems Pharmacology, 2017. **6**(9): p. 576-588.
8. Hsieh, Y.T., et al., *Characterization of FcγRIIIA effector cells used in in vitro ADCC bioassay: Comparison of primary NK cells with engineered NK-92 and Jurkat T cells.* J Immunol Methods, 2017. **441**: p. 56-66.
9. Wang, X., M. Mathieu, and R.J. Brezski, *IgG Fc engineering to modulate antibody effector functions.* Protein Cell, 2018. **9**(1): p. 63-73.
10. Raju, T.S. and R.E. Jordan, *Galactosylation variations in marketed therapeutic antibodies.* MAbs, 2012. **4**(3): p. 385-91.
11. Aoyama, M., et al., *Effects of terminal galactose residues in mannose α1-6 arm of Fc-glycan on the effector functions of therapeutic monoclonal antibodies.* MAbs, 2019. **11**(5): p. 826-836.
12. Filep, C., et al., *Multilevel capillary gel electrophoresis characterization of new antibody modalities.* Anal Chim Acta, 2021. **1166**: p. 338492.
13. Coppa, G.V., et al., *Oligosaccharides in human milk during different phases of lactation.* Acta Paediatr Suppl, 1999. **88**(430): p. 89-94.
14. Austin, S., et al., *Determination of β -Galactooligosaccharides by Liquid Chromatography.* Int J Anal Chem, 2014. **2014**: p. 768406.
15. *Official Methods of Analysis*. 2019, AOAC International: Rockville, MD. p. Method 2001.02.

4. List of publications

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/17/2024.PL
Subject: PhD Publication List

Candidate: Dániel Sárközy
Doctoral School: Doctoral School of Molecular Medicine
MTMT ID: 10064668

List of publications related to the dissertation

1. **Sárközy, D.**, Guttman, A.: Analysis of Peptides and Proteins by Native and SDS Capillary Gel Electrophoresis Coupled to Electrospray Ionization Mass Spectrometry via a Closed-Circuit Coaxial Sheath Flow Reactor Interface.
   Anal. Chem. 95 (18), 7082-7086, 2023.
   DOI: http://dx.doi.org/10.1021/acs.analchem.2c04332
   IF: 7.4 (2022)

2. **Sárközy, D.**, Farsang, R., Szigeti, M., Austin, S., Lock, S., Guttman, A.: Capillary electrophoresis analysis of industrial galactooligosaccharides.
   J. Pharm. Biomed. Anal. 233, 1-5, 2023.
   DOI: http://dx.doi.org/10.1016/j.jpba.2023.115434
   IF: 3.4 (2022)

3. **Sárközy, D.**, Guttman, A.: Capillary Sodium Dodecyl Sulfate Agarose Gel Electrophoresis of Proteins.
   Gels. 8 (2), 1-10, 2022.
   DOI: http://dx.doi.org/10.3390/gels8020067
   IF: 4.6

4. Szabó, M., **Sárközy, D.**, Szigeti, M., Farsang, R., Kardos, Z., Kozma, Á., Csánky, E., Chung, D. S., Szekanecz, Z., Guttman, A.: Introduction of a Capillary Gel Electrophoresis-Based Workflow for Biotherapeutics Characterization: size, Charge, and N-Glycosylation Variant Analysis of Bamlanivimab, an Anti-SARS-CoV-2 Product.
   Front. Bioeng. Biotechnol. 10, 1-9, 2022.
   DOI: http://dx.doi.org/10.3389/fbioe.2022.839374
   IF: 5.7

5. Szabó, M., Filep, C. B., Nagy, M., **Sárközy, D.**, Szigeti, M., Sperling, E., Csánky, E., Guttman, A.: N-glycosylation structure - function characterization of omalizumab, an anti-asthma biotherapeutic product.
   J. Pharm. Biomed. Anal. 209, 1-5, 2022.
   DOI: http://dx.doi.org/10.1016/j.jpba.2021.114483
   IF: 3.4

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

6. **Sárközy, D.**, Borza, B., Domokos, A., Váradi, E., Szigeti, M., Mészáros-Matwiejuk, Á., Molnár-Gábor, D., Guttman, A.: Ultrafast high-resolution analysis of human milk oligosaccharides by multicapillary gel electrophoresis.
Food Chem. 341 (2), 1-8, 2021.
DOI: http://dx.doi.org/10.1016/j.foodchem.2020.128200
IF: 9.231

List of other publications

7. **Sárközy, D.**, Guttman, A.: CE and CE-MS approaches for glycan analysis.
In: Capillary Electrophoresis - Mass Spectrometry for Proteomics and Metabolomics: Principles and Applications. Ed.: Rawi Ramautar, David D. Y. Chen, Wiley, Weinheim, 313-334, 2022.

8. Szarka, M., Szilasi, S., Döncző, B., **Sárközy, D.**, Rajta, I., Guttman, A.: The effect of simulated space radiation on the N-glycosylation of human immunoglobulin G1.
Electrophoresis. 39 (22), 2872-2876, 2018.
DOI: http://dx.doi.org/10.1002/elps.201800151
IF: 2.754

Total IF of journals (all publications): 36,485
Total IF of journals (publications related to the dissertation): 33,731

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

18 January, 2024

DEBRECENI EGYETEM * Egyetemi és Nemzeti Könyvtár *