Trends in Analytical Chemistry 176 (2024) 117762

Contents lists available at ScienceDirect

Trends in Analytical Chemistry

journal homepage: www.elsevier.com/locate/trac

# Glucose unit computation in capillary zone electrophoresis of carbohydrates using a numerical approximation based search for a virtual EOF marker. A tutorial

Robert Farsang [a], Anna Farkas [b], Gabor Jarvas [a], Andras Guttman [a,b,*]

[a] Translational Glycomics Group, Research Institute of Biomolecular and Chemical Engineering, University of Pannonia, Veszprem, Hungary
[b] Horváth Csaba Memorial Laboratory of Bioseparation Sciences, Research Center for Molecular Medicine, Faculty of Medicine, University of Debrecen, Hungary

ARTICLE INFO

*Keywords:*
Carbohydrates
Capillary electrophoresis
Virtual EOF marker
Numerical approximation
Maltodextrin ladder
Glucose unit

ABSTRACT

Several high-resolution carbohydrate separation methods are currently available for structural elucidation in the analytical glycomics field, but the associated glyco-informatics support is lagging. Identifying protein-bound glycan structures represents a multifaceted challenge with elements of separation science, bioinformatics, and sophisticated computational methods. In electroosmotic flow (EOF) driven separations with counter currently migrating sample components, the apparent migration times of the analyte molecules are the result of the interplay between their effective electrophoretic mobilities and the EOF. The resulting continuously increasing migration time distances between the consecutively migrating maltodextrin oligomers make glucose unit value calculation and the subsequent structural interpretation of unknown glycans extremely challenging. To correct this irregularity, a numerical approximation-based search-derived virtual electroosmotic flow marker is introduced and utilized. This tutorial guides the reader through the computation process for structural elucidation of N-linked carbohydrates by using a detailed worked example with a monoclonal antibody of biopharmaceutical interest.

## 1. Introduction

Analysis of protein-bound carbohydrates is of great importance in the biopharmaceutical and diagnostic fields [1]. High-resolution separation and the downstream structural elucidation of complex oligosaccharides represent a major challenge in the method development stage of process analytical endeavors. In addition, the large number of potentially occurring structural glycoforms along with their positional and linkage isomers also carry important information but are difficult to decipher [2]. Capillary electrophoresis-based N-glycan analysis methods have gained increasing importance in recent decades due to their excellent separation selectivity and detection sensitivity [3]. Among the various methods explored to identify glycans utilizing publicly available databases, the conventional glucose unit (GU) determination based approach is the most frequently used for structure determination [4]. In ambiguous cases, exoglycosidase digestion-mediated glycan sequencing can be applied [5]. Such specific enzymes usually include but are not limited to sialidase, fucosidase, galactosidase, hexosaminidase, and mannosidase along with their linkage and anomer-specific forms. It is also possible to determine glycan structures using capillary electrophoresis coupled to mass spectrometry, but that approach requires a high level of expertise and expensive instrumentation [6].

Currently used capillary gel electrophoresis (CGE) methods separate glycans labeled with a charged fluorophore (e.g., 8-Aminopyrene-1,3,6-trisulfonic acid trisodium salt (APTS)) under close to zero electroosmotic flow conditions and utilize three internal standards to normalize the migration times of the sample components to the maltodextrin ladder for GU value calculation [7]. That method assumes constant migration time ratios between the subsequently electromigrating maltodextrin oligomers in the reported application domain. The resulting GU values can then be used for database search to suggest the possible glycan structures corresponding to the peaks of interest, similar to that of the Kovats

*Abbreviations:* EOF, electroosmotic flow; GU, glucose unit; CZE, capillary zone electrophoresis; CGE, capillary gel electrophoresis; APTS, 8-Aminopyrene-1,3,6-trisulfonic acid trisodium salt; MD, maltodextrin.

* Corresponding author. Translational Glycomics Group, Research Institute of Biomolecular and Chemical Engineering, University of Pannonia, Veszprem, Hungary.
*E-mail address:* guttmanandras@med.unideb.hu (A. Guttman).

https://doi.org/10.1016/j.trac.2024.117762
Received 20 April 2024; Accepted 14 May 2024
Available online 15 May 2024

index in the petrochemical industry [8], in which case unknown sample components are defined relative to normal alkane homologs.

Electroosmotic flow-assisted high-resolution capillary zone electrophoresis (CZE) separation of APTS labeled oligosaccharides enabled to skip the costly and labor-intensive purification process to remove the non-binding excess fluorophore after the labeling reaction [9]. However, using that approach, the electrophoretic separation of the maltodextrin ladder peaks did not follow a linear migration time vs degree of polymerization (DP) pattern. The apparent mobility of the individual maltooligosaccharides is the result of the interplay between their effective electrophoretic mobility and the counter-current electroosmotic flow (EOF). Therefore, peak identification is not possible by directly normalizing the apparent migration times using the conventional approach, since the migration time differences between the ladder peaks do not vary linearly. To alleviate this issue, in this tutorial, we guide the reader through a novel computational method with a worked example, based on the utilization of a virtual EOF marker to determine the glucose unit (GU) value of any separated glycans with excellent repeatability. This approach utilizes numerical approximation-based search to obtain a virtual EOF marker to be utilized to compute the GU value of the separated glycans with high reproducibility, independently of the length of the capillary.

## 2. Theory

The concept of glucose unit determination using a numeric search-based approximation to obtain a virtual EOF Marker was based on the observation that in conventional no electroosmotic flow-based methods (i.e., CGE) a linear plot with $r^2 > 0.9999$ can be fitted between the apparent migration times and the degree of polymerization (DP) of the maltodextrin ladder only from DP8 to DP15. Please note that at > DP8, all maltooligosaccharide components have complete helical twists, and their structure is not that well defined below DP8. Using EOF-assisted separation in CZE mode, on the other hand, it is possible to numerically search for the migration time of a virtual EOF marker (vEOF) and utilize that to fit a linear and a second-degree polynomial plot (both with $r^2 > 0.9999$) between the effective migration time (i.e., vEOF-corrected, equation (3) below) of the maltodextrin ladder peaks and their degree of polymerization between DP8-DP15 and DP3-DP8, respectively.

The apparent electrophoretic mobility ($\mu_{app}$) of an electromigrating species is the algebraic sum of its effective mobility ($\mu_{eff}$) and the EOF mobility ($\mu_{EOF}$) as equation (1) shows.

$$\mu_{app} = \mu_{eff} + \mu_{EOF} \tag{1}$$

Thus, the migration time of the virtual EOF marker can be readily determined in all separations as shown below. In this tutorial, GU computations are based on the vEOF-corrected ($v'$) migration times of the APTS labeled maltodextrin ladder components $\left(\left|t_{mig;v'}^{MD(n)}\right|\right.$, the absolute values are used to make visualization more manageable), which is proportional to the reciprocal vEOF-corrected mobility under isorheic separation conditions, i.e., at constant applied electric field strengths (equation (2)).

$$\frac{1}{\mu_{mig;v'}} \propto t_{mig;v'} \tag{2}$$

The vEOF-corrected migration time of the electromigrating sample components $\left(\left|t_{mig;v'}^{peak}\right|\right)$ can then be determined by using the migration time of the vEOF marker $\left(t_{mig}^{vEOF}\right)$, and the apparent (measured) migration time of the sample component of interest $\left(t_{mig;app}^{peak}\right)$.

$$\left|t_{mig;v'}^{peak}\right| = -\left(\frac{1}{t_{mig;app}^{peak}} - \frac{1}{t_{mig}^{vEOF}}\right)^{-1} \tag{3}$$

Equation (3) is used for all GU calculations in this tutorial.

### 2.1. Numerical search for the migration time of the virtual EOF marker to compute the vEOF-corrected migration times of the ladder peaks

Calculation of the effective migration times of the sample components in EOF-assisted capillary zone electrophoresis analysis of APTS-labeled oligosaccharides requires the determination of the migration time of the electroosmotic flow as equation (3) depicts. The electroosmotic flow front is measured at the detection window with the use of an EOF marker.

Fig. 1 shows the unprocessed CE trace of the maltodextrin ladder components (upper panel) along with its current trace (lower panel). As one can observe, the current trace is not stable until the EOF front reaches the outlet-end of the separation capillary. This phenomenon is due to the non-consistent EOF rate caused by pH and conductivity differences between the sample plug and the background electrolyte. The migration time distances between the consecutive peaks are non-linearly

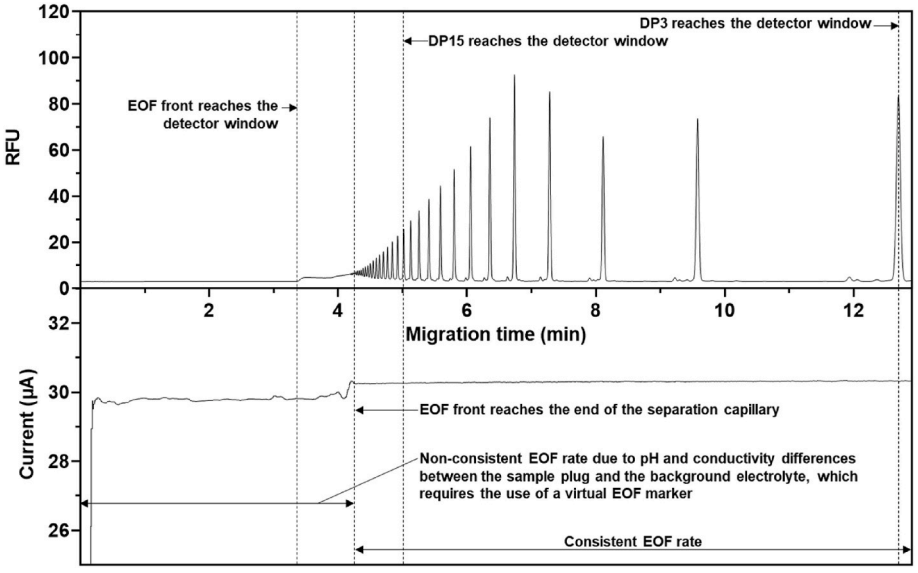


**Fig. 1.** EOF-assisted capillary zone electrophoresis separation of APTS labeled maltooligosaccharides (upper panel) with the associated current trace (lower panel).

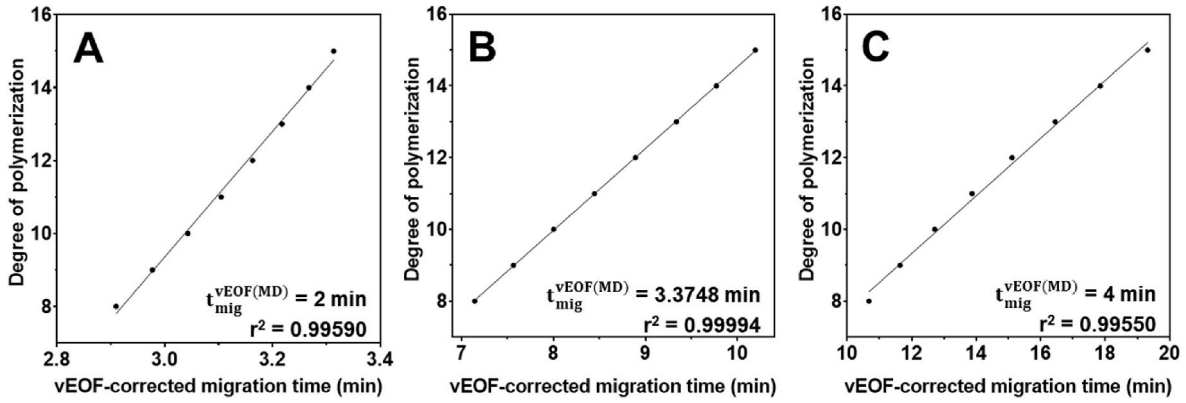


**Fig. 2.** Fitting examples between vEOF-corrected migration times and degrees of polymerization of the maltodextrin ladder components evaluating different vEOF values. Panel A: $t_{mig}^{vEOF(MD)}$ = 2 min, $r^2$ = 0.99590. Panel B: $t_{mig}^{vEOF(MD)}$ = 3.3748 min, $r^2$ = 0.99994 (best fit). Panel C: $t_{mig}^{vEOF(MD)}$ = 4 min, $r^2$ = 0.99550. Also refer to Table 2.

increasing, posing difficulties in GU value determination. To address this issue, we introduced the virtual EOF marker concept that allows GU value calculation with high precision and good repeatability in EOF-driven separations.

The migration time of the virtual EOF marker for the maltodextrin ladder separation $\left(t_{mig}^{vEOF(MD)}\right)$ was determined by a numeric approach using equation (3) to find the best linear fit ($r^2 > 0.9999$) to the coordinates of $i$ and $\left|t_{mig;v}^{DP(i)}\right|$, where $i$ = 8 to 15, illustrated in an example in Fig. 2.

The relative vEOF-corrected migration time of all maltodextrin ladder peaks can be calculated using $t_{mig}^{vEOF(MD)}$ considering a unity value for $\left|t_{mig;v}^{DP15}\right|$ (thus obtaining numbers between zero and unity). Plotting the calculated value of all data points of the maltodextrin ladder electropherogram shows that the relative (compared to DP15) vEOF-corrected migration time distances between the peaks in the range of DP8-DP15 are close to equal (Fig. 3).

In other words, one can say that the relative vEOF-corrected migration time ratios of the ladder peaks in the system are invariant. Therefore, the presented method eliminates the error that might be originating from conductivity and pH variation-mediated current changes when the sample plug passes through the capillary as one can observe between 0 and 4.2 min in the lower trace of Fig. 1.

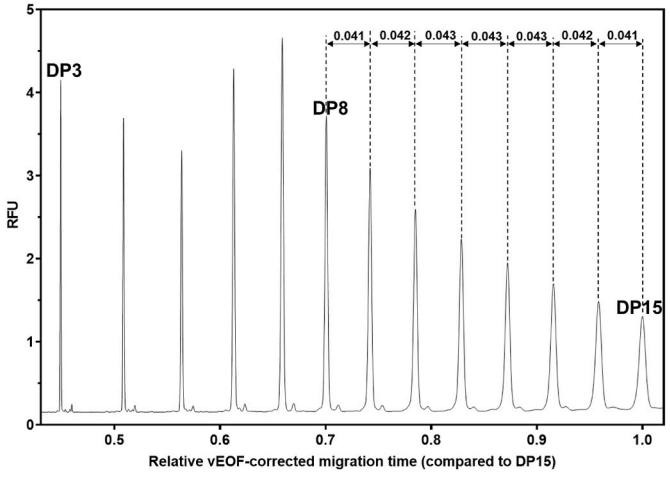


**Fig. 3.** Computed relative (normalized to DP15) vEOF-corrected migration time distances between the maltooligosaccharide peaks using the numeric search based virtual EOF marker.

### 2.2. Application of the virtual EOF marker concept to the analysis of an unknown glycan sample

In a given separation system (with unchanged capillary length, separation temperature etc.), the relative vEOF-corrected migration times of the maltooligosaccharide ladder peaks are not changing, i.e., constant. Thus, the migration time of the virtual EOF marker $\left(t_{mig}^{vEOF(S)}\right)$ in the case of all individual unknown sample (S) measurements with proper internal standards can be numerically determined by using equation (3). The relative vEOF-corrected migration time of DP3 (compared to DP15) must be equal to the determined relative vEOF-corrected migration time of DP3 in the MD ladder measurement (as discussed in the previous chapter). To achieve this, one needs to numerically find the $\left(t_{mig}^{vEOF(S)}\right)$ value for which the vEOF-corrected migration time ratio of DP3 and DP15 will be the same value as calculated earlier. With $t_{mig}^{vEOF(S)}$, $\left|t_{mig;v}^{DP15}\right|$ and the relative (vEOF-corrected) migration times of all ladder peaks (determined by MD ladder measurement calculation), a virtual MD ladder can be created between DP3 – DP15 for an unknown sample injected together with the internal standards. Using the $n$ and $\left|t_{mig;v}^{DP(n)}\right|$ coordinates a straight line can be fitted in the DP range of $n$ = 8 to 15 ($r^2 > 0.9999$) and a second-degree polynomial in the DP range of $n$ = 3 to 8 ($r^2 > 0.9999$).

Next, the vEOF-corrected migration times of the unknown sample peaks were calculated using the apparent migration times and the $t_{mig}^{vEOF(S)}$ values according to equation (3). Calculation of the GU values from the virtual MD ladder graphs was as follows:

substitution into the linear equation if $\left|t_{mig;v}^{peak}\right| \geq \left|t_{mig;v}^{DP8}\right|$,

substitution into the second-degree polynomial equation if $\left|t_{mig;v}^{peak}\right| < \left|t_{mig;v}^{DP8}\right|$.

The GU calculation method is illustrated in the flowchart in Fig. 4.

### 2.3. Worked example

First, the measured migration times of the maltodextrin ladder peaks were recorded in the range of DP15-DP3 from the corresponding electropherogram (Fig. 1, upper trace; Table 1).

Then, the migration time of each ladder peak between DP15 and DP8 was substituted into equation (3) to obtain the $t_{mig}^{vEOF(MD)}$ values with submillisecond precision from zero to the optimal value, i.e., until the degree of fit (coefficient of correlation, $r^2$) started to decrease.

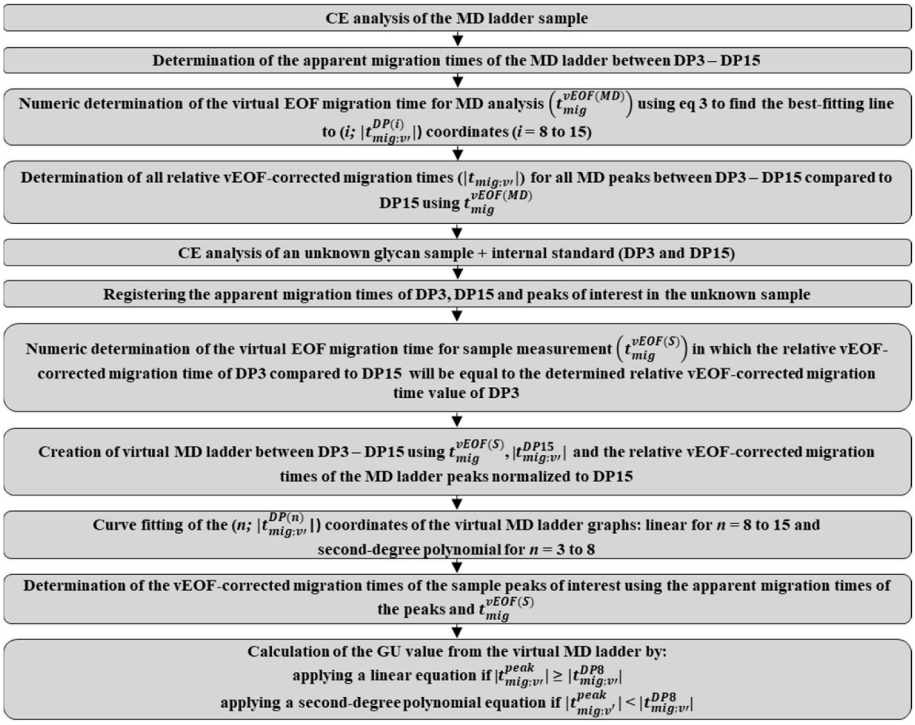


**Fig. 4.** Workflow of the GU computation method.

**Table 1**
Measured (apparent) migration times of the maltodextrin ladder components.

| MD Peak | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Migration time (min) | 5.045 | 5.155 | 5.285 | 5.440 | 5.621 | 5.837 | 6.094 | 6.398 | 6.783 | 7.339 | 8.180 | 9.671 | 12.846 |

$$\left|t_{mig;v'}^{DP(n)}\right| = -\left(\frac{1}{t_{mig;app}^{DP(n)}} - \frac{1}{t_{mig}^{vEOF(MD)}}\right)^{-1} \quad (4)$$

In each case, a straight line can be fitted between the polymerization degrees of the maltodextrin ladder components (n) and the resulting $t_{mig;v'}^{DP(n)}$. Then the best fit value regarding the coefficient of correlation ($r^2$) must be sought for and the corresponding value will be the $t_{mig}^{vEOF(MD)}$ (Table 2).

The dots represent the continuously increasing $t_{mig}^{vEOF(MD)}$ values but their corresponding fit values are not shown, i.e., only focusing on the important information part. In this actual instance, the best fit was $r^2$ = 0.9999424, with $t_{mig}^{vEOF(MD)}$ = 3.3748 min. Using this $t_{mig}^{vEOF(MD)}$, one can compute the $\left|t_{mig;v'}^{DP(n)}\right|$ values for all ladder peaks between DP15 and DP3 using equation (4), which is a modified version of equation (3) for MD ladder measurement (Table 3).

Using these $|t_{mig;v'}^{DP(n)}|$ values, one can calculate the relative vEOF-corrected migration times of the ladder peaks normalized to DP15 between DP15 and DP3 (Table 4).

The next step was to obtain the migration times of two randomly chosen unidentified peaks in the electropherogram during the analysis of the released and APTS-labeled N-glycans of a sample of biopharmaceutical interest, as well as the two internal standard peaks of DP15 and DP3 as shown in Fig. 5 (Table 5).

For the system with unchanged separation parameters used throughout the worked example, the relative vEOF-corrected migration

**Table 2**
Examples of the best fit values for the different virtual EOF migration times calculated by equation (4) during the numerical search.

| $t_{mig}^{vEOF(MD)}$ | … | 3.1378 | … | 3.37480 | … | 3.8320 | … |
|---|---|---|---|---|---|---|---|
| Fit ($r^2$) | | 0.9996830 | | 0.9999424 | | 0.9980192 | |

**Table 3**
Calculated vEOF-corrected migration times of the maltodextrin ladder components using the best fit 3.3748 min virtual EOF migration time.

| MD Peak | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\left\|t_{mig;v'}^{DP(n)}\right\|$ (min) | 10.195 | 9.772 | 9.336 | 8.891 | 8.446 | 8.001 | 7.564 | 7.142 | 6.716 | 6.248 | 5.745 | 5.184 | 4.577 |

**Table 4**
Relative vEOF-corrected migration times of the ladder components normalized to DP15 (shown in Fig. 2).

| MD Peak | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Relative $\left\|t_{mig;v'}^{DP(n)}\right\|$ | 1.000 | 0.959 | 0.916 | 0.872 | 0.828 | 0.785 | 0.742 | 0.701 | 0.659 | 0.613 | 0.564 | 0.508 | 0.449 |

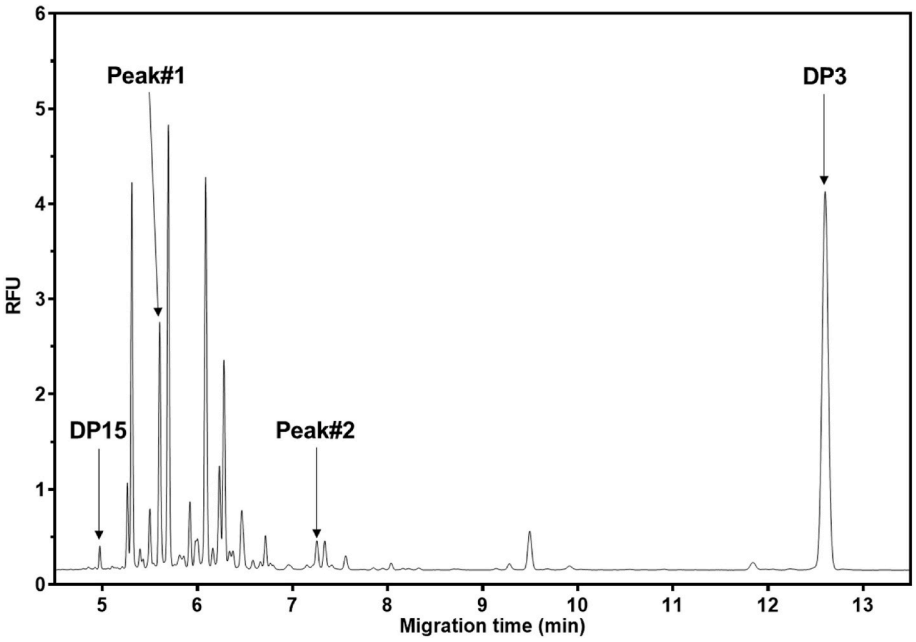


**Fig. 5.** Capillary electrophoresis separation of PNGase F released and APTS labeled hIgG1 glycans with two peaks randomly selected for exemplary computation.

**Table 5**
The apparent migration times of the two internal standards and the two randomly selected peaks in Fig. 5.

| Peak ID | DP15 | Peak#1 | Peak#2 | DP3 |
|---|---|---|---|---|
| Migration time (min) | 4.973 | 5.603 | 7.256 | 12.601 |

time of the DP3 did not change, so it was possible to find the $t_{mig}^{vEOF(S)}$ value for the sample analysis. To do this again, values were substituted into eq (3) in submillisecond increments to numerically search for the ratio closest that was already determined in Table 4 for DP3, i.e., 0.4490 (Table 6).

**Table 6**
Relative vEOF-corrected migration time of DP3 for different virtual EOF values.

| $t_{mig}^{vEOF(S)}$ $(min)$ | … | 3.0931 | … | 3.3303 | … | 3.6835 | … |
|---|---|---|---|---|---|---|---|
| Relative $\lvert t_{mig;v}^{DP3} \rvert$ | | 0.5010 | | 0.4490 | | 0.3664 | |

With this step, the $t_{mig}^{vEOF(S)}$ value was determined as 3.3303 min, which can be used to calculate the vEOF-corrected migration time of DP15, again, based on equation (3). Using this and the $\sqrt{\text{relative}}$ vEOF-corrected migration times of the ladder peaks between DP3-DP15 (Table 4) a virtual ladder can be established as shown in Table 7.

Fitting a straight line between the DP8-DP15 datapoints, and a second-degree polynomial between DP3-DP8 using the value pairs of Table 7 resulted in the plots shown in Fig. 6.

It was then necessary to examine how the vEOF-corrected migration times of the unknown peaks compare to the vEOF-corrected migration time of DP8 because the calculation system changes the equation at this point. If this value is greater than the $\left|t_{mig;v}^{DP(8)}\right|$ of the virtual MD ladder (e.g., unknown Peak#1 in the hIgG glycan trace), it is substituted into the linear equation. If less, it is substituted into the second-degree equation (e.g., unknown Peak#2 in the hIgG glycan trace). The resulting GU values for the two randomly selected peaks were $GU_{(Peak\#1)} = 10.665$ and $GU_{(Peak\#2)} = 5.953$, corresponding to the FA2(6)G1 and FA2BG2S2 structures based on searching the relevant databases (Table 8).

**Table 7**
vEOF-corrected migration times of the virtual ladder components.

| MD Peak | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lvert t_{mig;v}^{vDP(n)} \rvert$ (min) | 10.082 | 9.664 | 9.233 | 8.792 | 8.352 | 7.913 | 7.480 | 7.063 | 6.642 | 6.179 | 5.681 | 5.126 | 4.527 |

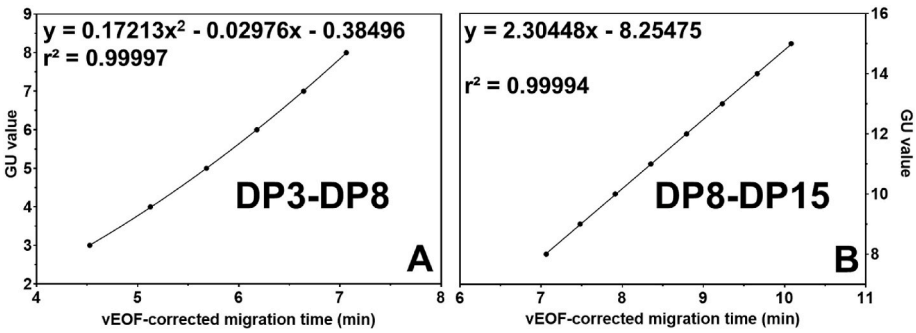


**Fig. 6.** Calibration plots between the vEOF-corrected migration times and GU values. Panel A: DP3-DP8. Panel B: DP8-DP15.

**Table 8**
GU database of hIgG1 N-glycan structures.

| Structure | GU value | Structure | GU value | Structure | GU value |
|---|---|---|---|---|---|
| FA2BG2 | 12.62 | A2B(6)G1 | 9.53 | A2G2S1 | 7.78 |
| FA2G2 | 12.32 | FA2B | 9.28 | FA2(3)G1S1 | 7.52 |
| FA2B(3)G1 | 11.79 | A2(6)G1 | 8.99 | FA2(6)G1S1 | 6.93 |
| A2G2 | 11.60 | FA2 | 8.69 | FA2BG2S2 | 5.95 |
| FA2(3)G1 | 11.20 | A2B | 8.44 | FA2G2S2 | 5.83 |
| FA2(6)G1 | 10.66 | FA2BG2S1 | 8.22 | A2G2S2 | 5.54 |
| A2B(3)G1 | 10.23 | FA2G2S1 | 8.08 | | |
| A2B(6)G1 | 9.71 | A2BG2S1 | 7.88 | | |

In case the GU calculation was performed on all the data points of the electropherograms, a GU-normalized visualization was possible as shown in Fig. 7. The lower trace shows the GU-normalized representation of MD the ladder, the upper trace the GU-normalized representation of the hIgG1 trace. For both electropherograms, the equations were generated separately, with different pairs for each measurement. It shows the accuracy of the system, i.e., there was only a very minimal discontinuity at the GU8 value in both cases, despite the fact that the second-degree polynomial was used to the left of this value and the straight line to the right. If there were a large breakage in Fig. 7 around GU = 8, the fit of the lines would not be acceptable, which would lead to inaccuracies in the GU calculation. If there were not a sufficient fit of the two equations to the corresponding points, the endpoints of the two equations would not approximate GU8 properly, so would follow each other with a discontinuous transition at the breakpoint as depicted by the arrow in Fig. 7.

### 2.4. Robustness (40 cm total length capillary)

The applicability and reproducibility of the above-described GU computing method were tested first using a 30 cm effective length capillary with one MD ladder and 6 hIgG1 samples with pre-injected internal standard measurements per day for 6 days as shown in Fig. 8 (the MD measurements are not shown in the figure). Please note that DP4 was not used in the calculation method, it was only used to check the accuracy of the obtained GU values. Panel A shows the obtained electropherograms of 6 injections of hIgG1 sample (unpurified) per day for 6 days. Since in this experiment a new capillary was used, the EOF rate was somewhat higher compared to the fully equilibrated condition. By the second day, the capillary was in full equilibrium, here the small decrease in EOF during the measurements was probably due to a tiny change in the pH of the CATRIS buffer, which was the result of the electrolysis during the separation process. Panel B shows the GU-normalized representations of the electropherograms generated by GU value computation of each and every single data point. It can be clearly seen that, regardless of the state of the capillary and the separation buffer, very high GU value reproducibility of the separated peaks was obtained.

To quantify robustness, the GU values of seven randomly chosen hIgG1 structures and the DP4 were calculated for the 36 measurements, which are shown in supplementary in Tables 1S, 2S and 3S, and the summary statistics are shown in Table 9.

### 2.5. Reproducibility in case of using 50 and 60 cm total length capillaries

Similar measurements were then performed in triplicate measurements using 50 cm and 60 cm total length capillaries. Fig. 9 depicts the resulting traces before (Panel A) and after GU normalization (as shown in Fig. 7) (Panel B). Fig. 9 also features three randomly chosen measurements from Fig. 8 (i.e., 6-1, 6-2, and 6-3 runs) for comparative purposes with the 40 cm total length capillary.

Evaluating the electropherograms as above, the statistical results obtained are shown in Table 10.

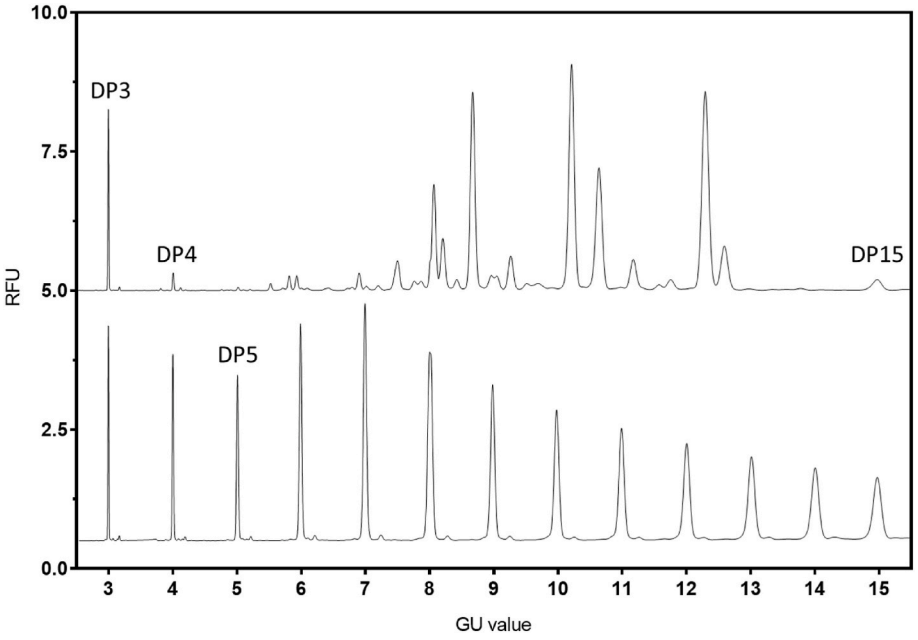


**Fig. 7.** GU normalized maltodextrin ladder (lower trace) and hIgG1 N-glycan $\sqrt{profile}$ (upper trace) electropherograms using the GU computing method. All data points of the electropherogram have been substituted into the corresponding equation, i.e., linear if the vEOF-corrected migration time of the data point was greater than that of DP8, and into the second-degree equation if it was less.

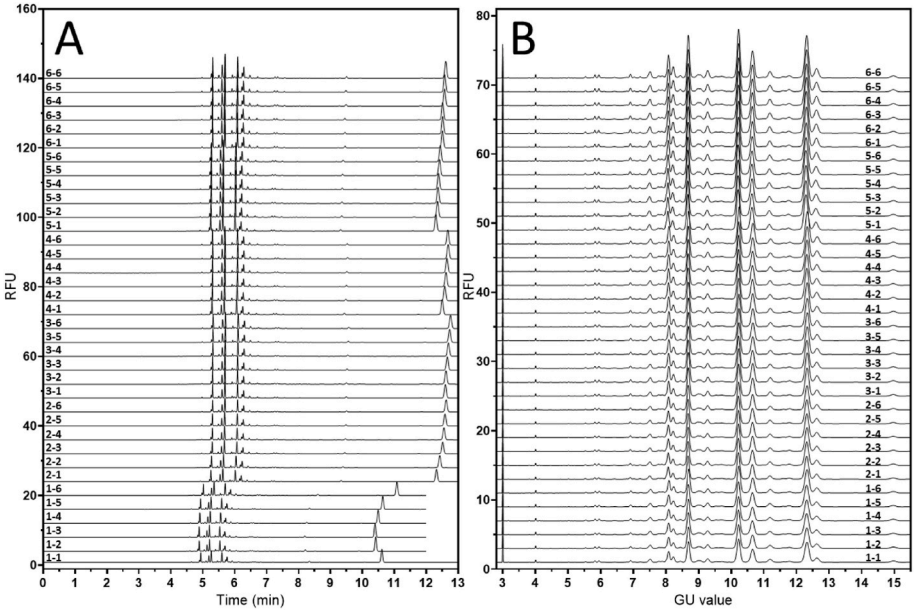


**Fig. 8.** Panel A: Six consecutive injections of the released and APTS labeled hIgG1 glycan pool per day for six days with pre-injected internal standards using a 30 cm effective length capillary (40 cm total length). Traces: the first and second numbers on the electropherograms depict the day and the actual measurement at that day, respectively. Panel B: The GU normalized separations of Panel A using the GU computing method for each and every single data points in the electropherograms (as described in Fig. 7).

**Table 9**
Overall statistics of the GU values of seven randomly chosen hIgG1 structures and DP4 for the 36 measurements shown in Fig. 8 over six days.

| | Average | SD | %RSD |
|---|---|---|---|
| FA2(3)G1 | 11.196 | 0.012 | 0.103 |
| FA2(6)G1 | 10.659 | 0.008 | 0.075 |
| A2B(3)G1 | 10.229 | 0.008 | 0.076 |
| FA2B | 9.281 | 0.007 | 0.074 |
| FA2 | 8.686 | 0.007 | 0.079 |
| FA2(6)G1S1 | 6.931 | 0.007 | 0.098 |
| FA2BG2S2 | 5.953 | 0.006 | 0.108 |
| DP4 | 3.991 | 0.002 | 0.041 |

### *2.6. GU values with DP4 and DP15 as internal standards*

Using all calculation steps as above, but using DP4 instead of the DP3 internal standard, the GU values were obtained for a capillary with a total length of 50 cm are shown in Table 11.

In this case, DP4 has about 4 minutes shorter migration time than DP3, so a faster analysis time with practically identical GU values obtained is apparently sufficient if there are no structures in the sample with a migration time longer than DP4.

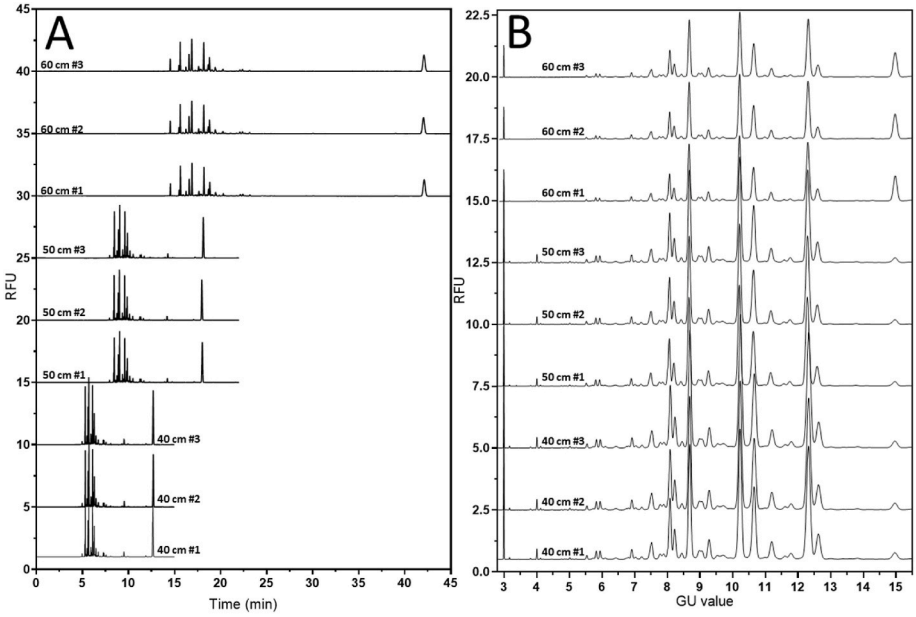


**Fig. 9.** Panel A: Triplicate injections of the hIgG1 glycans with pre-injected internal standard on a 30 cm effective length (40 cm total length, lower triplet), 40 cm effective length (50 cm total length, middle triplet) and 50 cm effective length (60 cm total length, upper triplet) capillary columns. Panel B: GU normalized electropherograms of Panel A using the GU computing method for each and every single data points in the electropherograms.

**Table 10**
Overall statistics of the GU values in case of 40 cm, 50 cm and 60 cm total length capillaries with triplicate injections.

| 40 cm (n = 3) | | | | 50 cm (n = 3) | | | | 60 cm (n = 3) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Average | SD | %RSD | | Average | SD | %RSD | | Average | SD | %RSD |
| FA2(3)G1 | 11.188 | 0.003 | 0.029 | FA2(3)G1 | 11.177 | 0.009 | 0.076 | FA2(3)G1 | 11.185 | 0.007 | 0.058 |
| FA2(6)G1 | 10.655 | 0.002 | 0.022 | FA2(6)G1 | 10.643 | 0.008 | 0.071 | FA2(6)G1 | 10.646 | 0.004 | 0.038 |
| A2B(3)G1 | 10.225 | 0.003 | 0.026 | A2B(3)G1 | 10.216 | 0.007 | 0.064 | A2B(3)G1 | 10.218 | 0.003 | 0.030 |
| FA2B | 9.280 | 0.002 | 0.025 | FA2B | 9.267 | 0.004 | 0.044 | FA2B | 9.262 | 0.003 | 0.032 |
| FA2 | 8.684 | 0.002 | 0.018 | FA2 | 8.674 | 0.006 | 0.063 | FA2 | 8.671 | 0.004 | 0.040 |
| FA2(6)G1S1 | 6.928 | 0.001 | 0.014 | FA2(6)G1S1 | 6.922 | 0.005 | 0.073 | FA2(6)G1S1 | 6.919 | 0.004 | 0.058 |
| FA2BG2S2 | 5.948 | 0.001 | 0.010 | FA2BG2S2 | 5.942 | 0.005 | 0.076 | FA2BG2S2 | 5.940 | 0.003 | 0.042 |
| DP4 | 3.991 | 0.000 | 0.000 | DP4 | 3.991 | 0.002 | 0.038 | DP4 | 3.986 | 0.001 | 0.014 |

**Table 11**
Overall statistics of the GU values in case of 50 cm length capillary using triplicate runs with DP3 + DP15 (left panel) and DP4 + DP15 (right panel) for GU calculation.

| | 50 cm (n = 3; DP3 + DP15 internal standards) | | | | 50 cm (n = 3; DP4 + DP15 internal standards) | | |
|---|---|---|---|---|---|---|---|
| | Average | SD | %RSD | | Average | SD | %RSD |
| FA2(3)G1 | 11.177 | 0.009 | 0.076 | FA2(3)G1 | 11.172 | 0.007 | 0.065 |
| FA2(6)G1 | 10.643 | 0.008 | 0.071 | FA2(6)G1 | 10.639 | 0.006 | 0.058 |
| A2B(3)G1 | 10.216 | 0.007 | 0.064 | A2B(3)G1 | 10.211 | 0.005 | 0.052 |
| FA2B | 9.267 | 0.004 | 0.044 | FA2B | 9.262 | 0.003 | 0.030 |
| FA2 | 8.674 | 0.006 | 0.063 | FA2 | 8.668 | 0.004 | 0.047 |
| FA2(6)G1S1 | 6.922 | 0.005 | 0.073 | FA2(6)G1S1 | 6.915 | 0.003 | 0.043 |
| FA2BG2S2 | 5.942 | 0.005 | 0.076 | FA2BG2S2 | 5.936 | 0.003 | 0.044 |

## 3. Conclusions

In this tutorial, we presented a novel GU determination method utilizing a numerical approach for EOF assisted CZE analysis of APTS labeled carbohydrates. As described in our earlier paper [9], electroosmotic flow assisted capillary electrophoresis separation offers high resolution of APTS labeled glycans without the need to remove the excess labeling dye before the analysis, therefore, saving time and resources. However, to identify the glycan structures with the use of this approach in the highest precision, a new numerical approximation based method was required because of the non-constant migration time differences between the maltodextrin ladder peaks. It was important to alleviate the effect of the interplay between the EOF and the effective electrophoretic mobility of the sample components, also considering the inconsistency originated form the actual state of the capillary wall and the conductivity of the sample plug. The theoretical background for the use of a virtual EOF marker-corrected (vEOF-corrected) migration time was described first, which was used to fit a straight line with $r^2 > 0.9999$ to the maltodextrin ladder components between DP8 and DP15 and a second-degree polynomial between DP3 and DP8 with $r^2 > 0.9999$ for the EOF-driven glycan separation method to allow the GU values to be determined with the highest accuracy and reproducibility. We presented a step-by-step tutorial workflow of calculating the GU values of two exemplary unknown hIgG1 glycan structures and identifying them by the calculated GU values using the correspondingly developed database (Table 8). A visualization approach, based on the computational method was introduced, where the electropherograms were normalized and linearized to the GU in the case of the maltodextrin ladder and hIgG1 measurements. The outstanding robustness of the proposed method was also demonstrated by performing 6 injections per day for 6 days, which were used to calculate the GU of the detected glycan structures with the average GU value deviation of 0.1 RSD%. It was also demonstrated that the calculated GU values of several structures of interest with the use of 40 cm, 50 cm, and 60 cm capillaries were independent of the capillary length. Finally, it was shown that by using DP4 instead of DP3 as internal standard but keeping all other parameters of the calculation method unchanged, practically the same GU value was obtained for all structures but with shorter analysis time.

## 4. Experimental section

### 4.1. Materials

Glacial acetic acid, glycerol, hexanoic acid, acetonitrile, Tris-base, sodium dodecyl sulfate (SDS) and tetrahydrofuran were from VWR (Radnor, PA, USA). The human IgG1, 8-Aminopyrene-1,3,6-trisulfonic acid trisodium salt (APTS), maltotriose, maltotetraose, sodium-cyanoborohydride (1 M in THF) and dithiothreitol (DTT) were purchased from Sigma-Aldrich (St. Louis, MO, USA). The PNGase F enzyme was made in-house following the procedure described in Ref. [10]. The maltopentadecaose was from Elicityl S.A. (Crolles, France) and the M040 maltooligosaccharide ladder was a kind gift of the Grain Processing Co. (Muscatine, IA). The 30 μm ID (365 μm OD) fused silica capillary was from Polymicro Technologies (Phoenix, AZ, USA).

### 4.2. Carbohydrate release by PNGase F

Ten microliters of 1 mg/mL hIgG1 was first denatured by the addition of 4 μL denaturing solution (400 mM DTT and 5 % SDS) and incubated for 15 min at 70 °C. This step was followed by the addition of 2 μL PNGase F (1.35 mg/mL in 50 % glycerol) and incubated for 2.0 h at 37 °C to release the asparagine linked carbohydrates from the hIgG1.

### 4.3. APTS labeling of carbohydrates

The released oligosaccharides as well as 10 μL of a 1 mg/mL maltooligosaccharides ladder, maltotriose (DP3), maltotetraose (DP4) and maltopentadecaose (DP15) were all subject to fluorophore labeling. This process involved the addition of 20 μL of reductive amination reaction mixture containing 1 μL of 120 mM APTS, 10 % of 1 M sodium cyanoborohydride (dissolved in THF), 15 % HPLC-grade water, 35 % glacial acetic acid, and 40 % tetrahydrofuran. The reaction mixture was then incubated overnight at 37 °C with evaporative derivatization [11]. Because of the use of EOF assisted CZE separation mode [9], the excess dye was not removed from the samples before analysis.

### 4.4. Capillary zone electrophoresis

All capillary zone electrophoresis separations were conducted using a PA800 Plus instrument (Beckman Coulter, Brea, CA), equipped with laser-induced fluorescent detection (excitation at 488 nm and emission at 520 nm). Bare fused silica (BFS) capillaries (30 μm ID/365 μm OD) with 30 cm (40 cm total), 40 cm (50 cm total) and 50 cm effective lengths (60 cm total) were used. The background electrolyte contained 253 mM Tris and 150 mM hexanoic acid (pH 8.1, CATRIS buffer [9]. The applied electric potential was 30 kV at 20 °C, with the anode positioned at the injection side and the cathode at the detection side (normal polarity). The separation capillary was rinsed for 3 min between runs using 0.5 N NaOH, followed by a 3-min rinse with the separation background electrolyte. The APTS labeled internal standard mixture (DP3, DP4 and DP15 in the ratio of 5:1:1, diluted by 1,000x) was pressure-injected at 1 psi for 5 s followed by sample injection (1 psi for 5 s). System control and data acquisition were managed by using the 32 Karat software (version 10.1, Beckman Coulter).

## CRediT authorship contribution statement

**Robert Farsang:** Conceptualization, Formal analysis, Investigation, Writing – original draft. **Anna Farkas:** Validation. **Gabor Jarvas:** Project administration. **Andras Guttman:** Funding acquisition, Supervision, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgment

The authors gratefully acknowledge the support of the ATBG Korea V4 joint project of the National Research, Development and Innovation Office of Hungary #2023-1.2.1-ERA_NET-2023-00015, the University of Debrecen Program for Scientific Publication and the TudFin grant #1G3DBK0CTUF247. This is contribution #210 from the Horváth Csaba Memorial Laboratory of Bioseparation Sciences.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trac.2024.117762.

## References

[1] B. Byrne, G.G. Donohoe, R. O'Kennedy, Drug Discov. Today 12 (2007) 319.
[2] C.D. Gutierrez Reyes, P. Jiang, K. Donohoo, M. Atashi, Y.S. Mechref, J. Separ. Sci. 44 (2021) 403.
[3] A. Guttman, L. Hajba. Capillary Gel Electrophoresis, Elsevier Science, 2021.
[4] S. Mittermayr, A. Guttman, Electrophoresis 33 (2012) 1000.
[5] A. Guttman, K.W. Ulfelder, J. Chromatogr. A 781 (1997) 547.
[6] C.M. Snyder, X. Zhou, J.A. Karty, B.R. Fonslow, M.V. Novotny, S.C. Jacobson, J. Chromatogr. A 1523 (2017) 127.
[7] G. Jarvas, M. Szigeti, J. Chapman, A. Guttman, Anal. Chem. 88 (2016) 11364.
[8] E. Kovats, Helv. Chim. Acta 41 (2004) 1915.
[9] R. Farsang, K. Hogyor, G. Jarvas, A. Guttman, Anal. Chem. 95 (2023) 16459.
[10] N. Kovács, R. Farsang, M. Szigeti, F. Vonderviszt, H. Jankovics, Mol. Biotechnol. 64 (2022) 914.
[11] B. Reider, M. Szigeti, A. Guttman, Talanta 185 (2018) 365.