LAPPEENRANTA UNIVERSITY OF TECHNOLOGY
Faculty of Technology
Department of Chemical Technology
Bachelor's Degree Program in Chemical Engineering

Jere Elfving

In-House Validation of Capillary Electrophoresis Method

Examiners: Satu-Pia Reinikainen
Maaret Paakkunainen
Laura Kaijanen
Supervisor: Satu-Pia Reinikainen

Contents

Abstract

Capillary electrophoresis method designed originally for the analysis of monosaccharides was validated using reference solutions of polydatin. The validation was conducted by studying and determining the concentration levels of LOD and LOQ and the range of linearity and by determining levels of uncertainty in respect to repeatability and reproducibility. The reliability of the gained results is also discussed. A guide with recommendations considering the validation and overall design of analysis sequences with CE is also produced as a result of this study.

1. Introduction

Capillary electrophoresis in analytic chemistry is a widespread analysis method for many different applications for example in food industry [1] [2] and pharmaceutical industry [3]. As previous studies have shown, CE is fast and fairly accurate compared to other analysis methods, for example HPLC, *High Performance Liquid Chromatography* [4] [5] or SELDI, *Surface Enhanced Laser Desorption/Ionization* [6]. However, as altering different condition parameters (pH, temperature etc.) can radically affect the gained results, it is advisable that every applied CE method, with a different set of conditions, is validated separately and sufficiently.

Different variations of CE have been introduced, but the basic principle remains the same: Compounds are separated and migrated in a capillary between two electrodes in the influence of a high electric field and osmotic flow. The analysis method used in this study was capillary zone electrophoresis (CZE) - type, in which the electro-osmotic flow (EOF) results in separation of compounds by an orderly migration of anions, neutral compounds and cations. [7],[8]

A specific CE run method (condition parameters examined in chapter 2.) has been used in Lappeenranta University of Technology to analyze mostly monosaccharides. Although this method has been in use for some time, no proper validation has been reported. The accurate ranges of examined concentrations have been estimated by user experience, only. CE is used in LUT also in other analyses and applications by many users. Thus, an in-house validation of this specific CE method is required, as well as a guide considering general directions in the use of CE.

In this paper, results of an in-house validation of a CE method are represented and discussed, and based on these results, a recommendation considering the planning of a running sequence with this CE method is represented. The validation was conducted viewing parameters of limit of detection, limit of quantification, reproducibility, repeatability and linearity using different sets of concentrations of reference solutions.

2. Experimental

2.1. Instrumentation

The CE system under examination was HP 3D CE (Agilent). The used detector was an UV diode array detector capable of detecting wavelengths in the range of 190-600nm. In this study, detected wavelengths of 210nm, 254nm and 270nm were of interest. Agilent ChemStation software was used to integrate the gained electropherograms. The capillary used was polyimide-coated silica of length 70 cm (61.5 cm to the detector) and diameter 50 µm. Background electrolyte (buffer) solution composition was 130 mM sodium hydroxide (NaOH) and 36 mM disodium hydrogen phosphate (Na_2HPO_4) in purified water. Sample injection was performed at 40 mbar for 8.0 s.

The method used was named "LKSUG70C1.M". The capillary was operated at a constant temperature of +25 ^{o}C and with a separation voltage of 17kV with positive polarity. After each run the column was rinsed with the buffer for 8 minutes.

2.2. Reference material

The working reference solution was prepared by diluting 50.0 milligrams of polydatin (>95% pure, Sigma-Aldrich) in 100ml of deionized water purified with Elgan Centra-R 60/120 (Veolia Water). The 500 ppm working solution was then diluted in purified water to gain solutions ranging from 2 ppm to 120 ppm.

Every reference solution was prepared by filtering through syringe to vial. The detection time for each injection was 50 min, followed by the 8 min run of buffer.

Polydatin was used as a reference material due to it being one of the chemicals that the analysis method in question was designed for. Resveratrol and its glycoside, polydatin, have been studied previously because of their health benefitting properties for many plant species and for mammals [9]. Other stilbenes, the group in which the two compounds are included in, and their beneficial effects have also been studied in Finnish pine tree samples [10] [11]. Thus, the accurate analysis of resveratrol and its derivative polydatin from extraction samples, among many other organic compounds, is under close examination at LUT Chemistry.

2.3 Validation parameters

The definitions for the determined validation parameters were gained from an Agilent method validation guide [12].

Defined by ICH (*International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use*), the limit of detection (LOD) is the point at which a measured value is larger than the uncertainty associated with it. It is the lowest concentration of analyte in a sample that can be detected but not necessarily quantified.

ICH defines the limit of quantitation (LOQ) of an individual analytical procedure as the lowest amount of analyte in a sample which can be quantitatively determined with suitable precision and accuracy.

ICH defines linearity of an analytical procedure as its ability (within a given range) to obtain test results that are directly proportional to the concentration (amount) of analyte in the sample. A linear regression equation applied to the results should have an intercept not significantly different from zero.

Reproducibility, or ruggedness, is defined by the USP as the degree of reproducibility of results obtained under a variety of conditions, such as different laboratories, analysts, instruments, environmental conditions, operators and materials. In the case of in-house validation, however, the results do not cover reproducibility consisting comparison between different laboratories. Therefore, in this case, it is safer to address the term within-laboratory reproducibility. Within-laboratory reproducibility is defined as variation in the results when the same sample is analyzed at different points in time in the laboratory [13].

ICH defines the precision of an analytical procedure as the closeness of agreement (degree of scatter) between a series of measurements obtained from multiple sampling of the same homogeneous sample under the prescribed conditions. Repeatability expresses the precision under the same operating conditions over a short interval of time.

3. Validation results and discussion

All responses were integrated manually with the Agilent Chemstation tool, with a minor average error of about 1% caused by the method. The average error was determined by repeating the manual integration procedure. Microsoft Excel software was used in the plotting procedures. If not mentioned separately, all results below were gained from the electropherograms detected at 210nm.

3.1 LOD and LOQ

Limits of detection and quantification were determined by both calculatory and visual analysis from the electropherograms. Calculations were conducted by first integrating approximately the same segment in each of the zero-sample-electropherograms. An average value was then calculated from these responses, which was used as the zero-value-point when plotting a curve from 0 ppm to 60 ppm with the corresponding responses. The response correlating 60 ppm concentration was gained as an average value area from the first runs of series of 30 repeats of 60 ppm polydatin solution. Slope was then determined from the received graph.

The LOD-response was gained by using the zero-average-value and multiplying it by signal-to-noise-ratio of **3**, which is commonly used in the determination of the LOD. The LOD concentration was then determined by dividing the LOD-response value with the slope mentioned above. Thus was determined **LOD=2.28 ppm** (see App. 2). Equations used are presented below.

$$A_{LOD} = 3\,A_{blank}$$

$$A = bc$$

$$c_{LOD} = \frac{A_{LOD}}{b}$$

The defined LOD-value can be further justified by examining the two electropherograms below (Fig. 1); the first shows an electropherogram of a zero-polydatin sample, whereas the latter was gained with a 2 ppm- polydatin solution. The response peak of the polydatin from the 2 ppm-solution is fairly easy to spot and can be separated from the noise.

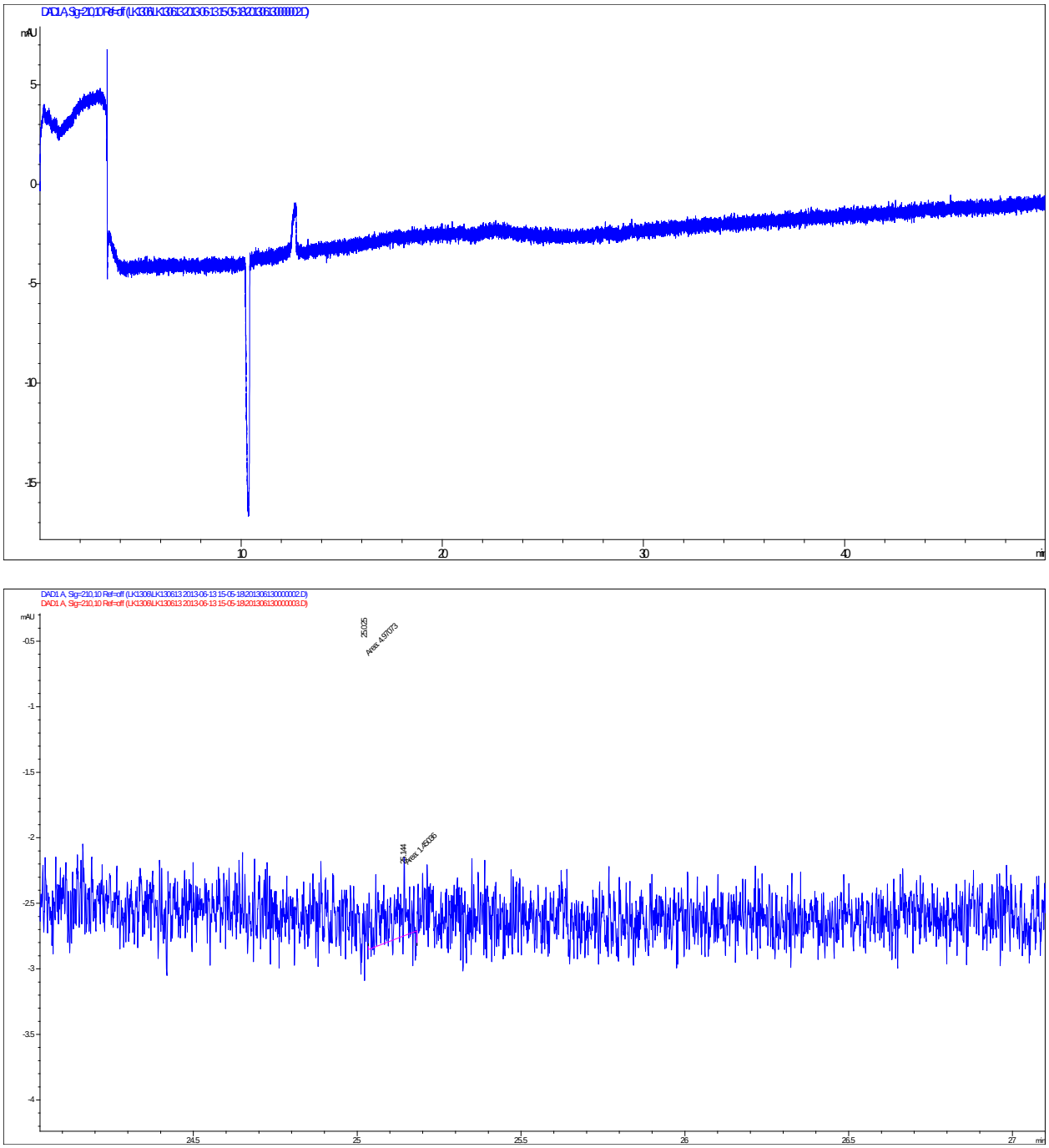


Figure 1. The upper figure shows an electropherogram of the first zero-sample run in the calibration series. No major peaks can be observed, apart from the one caused by buffer solution at migration time of 12 min. The lower figure is a close-up from the zero-sample with integration of noise.

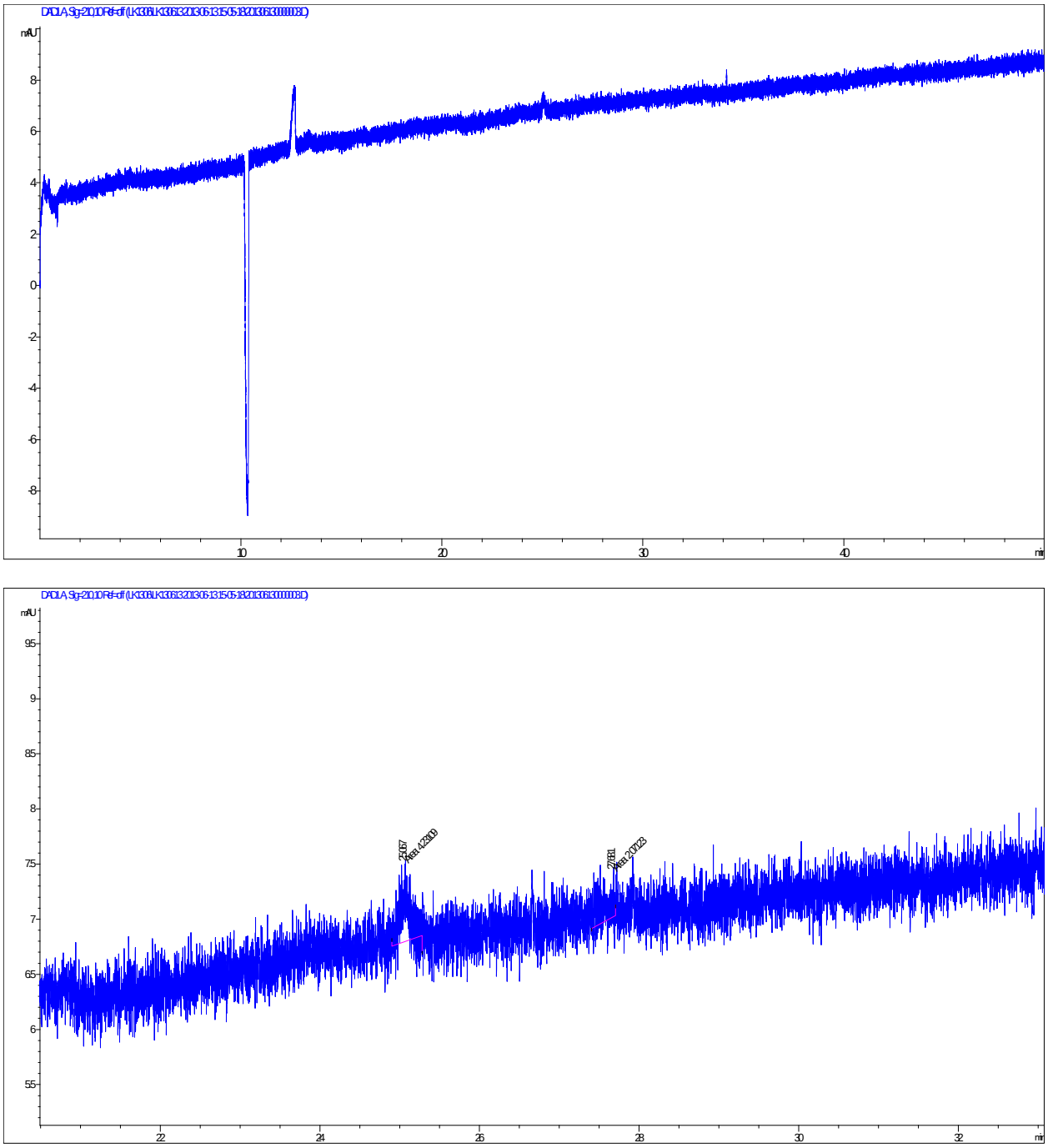


Figure 2. The upper figure represents an electropherogram of the first 2 ppm run in the calibration series. The polydatin peak at 25 min is easily detected. The lower figure is a close-up of the same figure, showing the integrated peak of the 2 ppm response.

Limit of quantification was gained by using a coefficient of **3** to multiply the gained LOD. Thus was gained **LOQ = 3 LOD = 6.84 ppm**. The used multiplier coefficient can be explained by examining the error related to both difficulty of manual integration and the correctness of the gained responses.

The relative error caused by manual integration is usually at its highest with the low concentrations, and in this case, down from 10 ppm. This error was determined by conducting the integration of the peaks two times, and comparing the average standard deviations in each concentration point gained from the two

sets of integrations. Thus was found, that these average errors sublimed radically in 5 ppm and 8 ppm integrations, ranging from 6.5% to 8.5%, when compared to the relative standard deviation of over 30 % in the case of 2 ppm integrations (see App. 2: Table VIII).

The correctness of the low concentration responses was calculated using the same slope as in determining LOD above. The relative errors of these experimental concentrations were then calculated, and averages from these were determined. The gained average relative errors in each concentration point were then compared with each other, heeding the difference in results of the third series of injections. Close observation revealed that the 5 ppm results were not accurate enough, with over 20% average relative error. The 8 ppm sample, however, gave an acceptable average relative error of 5 %, and only 2 %, if the third injection was ignored.

Considering the examinations above, it can be stated that the LOQ was best valuated by using a LOD-multiplier between 3 and 4, with the sample concentration ranging from 7 ppm to 9 ppm.

3.2 Linearity

A crude estimation of the linear zone was known beforehand to be approximately in the range of 5-100 ppm based on previous in-house studies at LUT Chemistry. Thus, linearity was determined by examining the calibration curves gained by first performing the calibration series run of 11 points in the concentration range of 0-120 ppm. The examined concentration range was then narrowed and a calibration run of 7 injections in the range of 5-100 ppm was performed. The first calibration series of injections was repeated three times, and the second series was repeated two times in a row.

The first rough estimates concerning the starting and ending points of the linear concentration range were determined by visual examination based on figure 3.

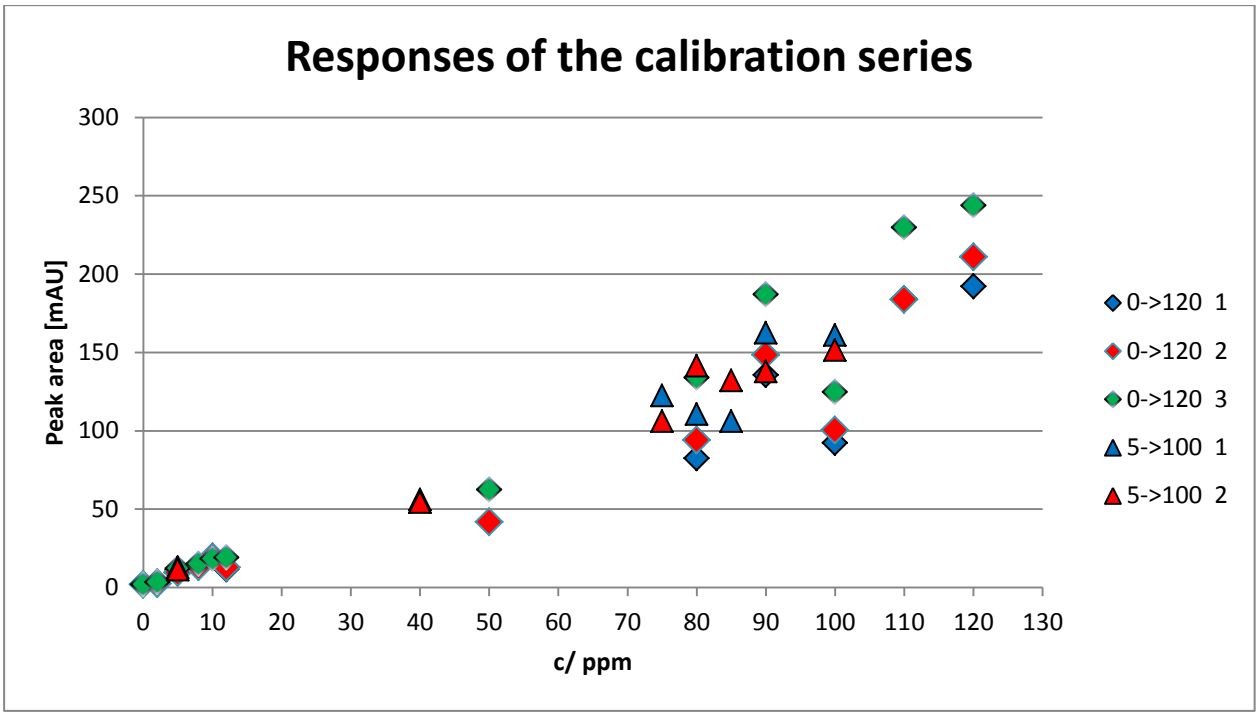


Figure 3. The responses of the five calibration series plotted against sample polydatin concentration. Visual examination reveals that divergence between the responses of the same series' repeat runs starts to grow rapidly when the concentration reaches 80 ppm.

A concentration range of **5-75 ppm** was visually determined to be the linear concentration range for polydatin using this CE-method. This range could further be explained by examining correlations of the calibration curves when plotted in different ranges of concentrations. Table I shows, that the average of the R^2- values, the correlations of the responses and the sample concentrations, is highest in the mentioned concentration zone. Since the LOD = 8 ppm, the working range is 8-75 ppm in the example case (see App.3).

Table I. The R^2- values gained from the trend lines of the calibration curves plotted in different zones of concentrations. The calculated averages reveal the best correlation in the zone of 5-75 ppm.

$c_{[ppm]}$	0-75	2-75	5-75	0-80	2-80	5-80
5-100 $_{1.}$	0.9856	0.9856	**0.9856**	0.9721	0.9721	0.9721
5-100 $_{2.}$	0.9973	0.9973	**0.9973**	0.9523	0.9523	0.9523
0-120 $_{3.}$	0.989	0.9902	**0.9995**	0.9747	0.9726	0.9693
0-120 $_{2.}$	0.9256	0.9231	0.946	**0.9559**	0.9518	0.9465
0-120 $_{1.}$	0.9236	0.9265	0.9394	**0.969**	0.9667	0.9638
average	0.96422	0.96454	**0.97356**	0.9648	0.9631	0.9608
$c_{[ppm]}$	0-90	2-90	5-90	0-100	2-100	5-100
5-100 $_{1.}$	0.8975	0.8975	0.8975	0.9154	0.9154	0.9154
5-100 $_{2.}$	0.9659	0.9659	0.9659	0.9702	0.9702	0.9702
0-120 $_{3.}$	0.9578	0.9549	0.9503	0.9138	0.9056	0.8936
0-120 $_{2.}$	0.9195	0.9138	0.9045	0.8989	0.8897	0.8757
0-120 $_{1.}$	0.9175	0.9104	0.9005	0.8992	0.8891	0.9005
average	0.93164	0.9285	0.92374	0.9195	0.914	0.91108

3.3 Repeatability

Typically repeatability is determined by repeating the injection 3-5 times, and the uncertainty component is estimated as standard deviation of the measurements divided by the square root of number of repeats. However, this procedure would have given too optimistic results (0.5-1.5 %) hiding systematic drift of the result. The response areas tend to increase systematically over the time and therefore the repeatability was determined over longer segments.

Repeatability was studied by examining the series consisting repeats of a 60 ppm calibration solution. To study the peak area variation, the 30-repeat series was divided into three 10 repeat segments, that were compared with each other by performing Student's t-tests (see App. 5.1). From these three segments, uncertainty components were also determined (Table II). To determine an approximation for a suitable amount of repeated injections, variation coefficients were also determined from the peak areas. Migration time variation was also studied by calculating relative error in respect of time. In the following, the series that were tested together, will be addressed as series A, B and C, corresponding to the three groups of samples numbering 1-10, 11-20 and 21-30 in the 60 ppm series.

A confidence interval of 95% was used in two-tailed t-distribution to determine t-test values. F-tests were first performed to test the significance in the difference between the two series' standard deviation. The calculated F- value was greater than the table value with series A versus B, but not with series B vs. C. Thus, degrees of freedom were gained by calculation with the Satterthwait equation for the first tested pair, but for the latter, the degrees of freedom had to be united additively. T-test values were then calculated using the calculated mean standard deviations and the gained degrees of freedom, and compared to the t-test table values. Statistically, in both cases the comparison showed that the mean values of these series were insignificantly different.

However, the calculated standard uncertainty (Table II) with series A was significantly greater than with series B and C, and therefore, the first 10 sample results were not included in the analysis of suitable amount of injections (see chapter **3.5**). The calculated variation coefficients were plotted against sample number and a polynomial correlation was used to determine the curve equation, which is showed in the figure 4 (see App. 5.2). Based on the curve, the increase point for variation coefficient was at 17 repeats.

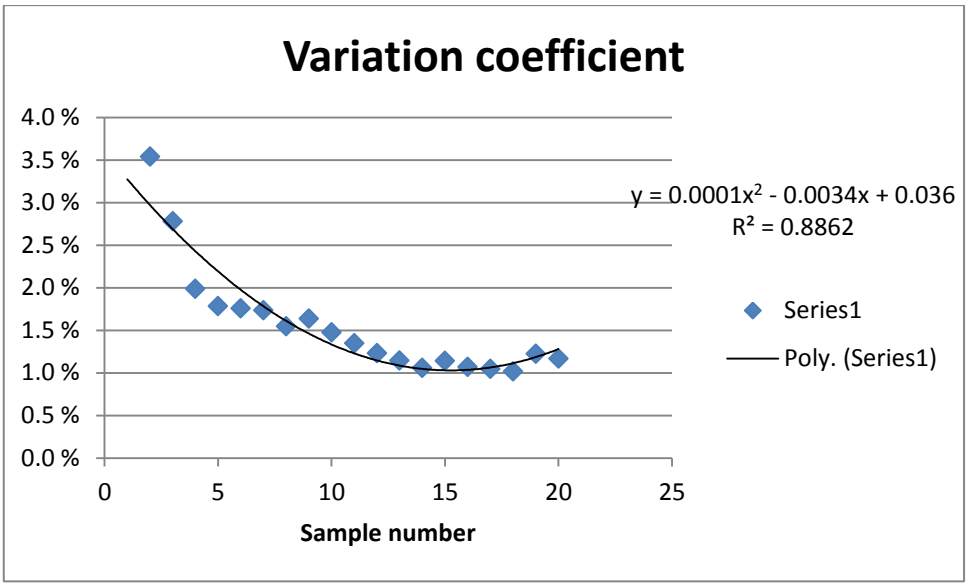


Figure 4. Variation coefficients gained from the 60 ppm polydatin sample CE injections. The curve shows that after about 15 injections, the CV starts to increase again. Sample number 11 was used as the origin value.

When determining repeatability based on migration time variation, a maximum acceptable relative error of 10% was used. In the 60 ppm series, the relative error of migration time remained acceptable in 15 repeated injections. The said amount of injections was determined by plotting the relative errors against sample number, with six values excluded from the correlation (figure 5, see App. 5.3).

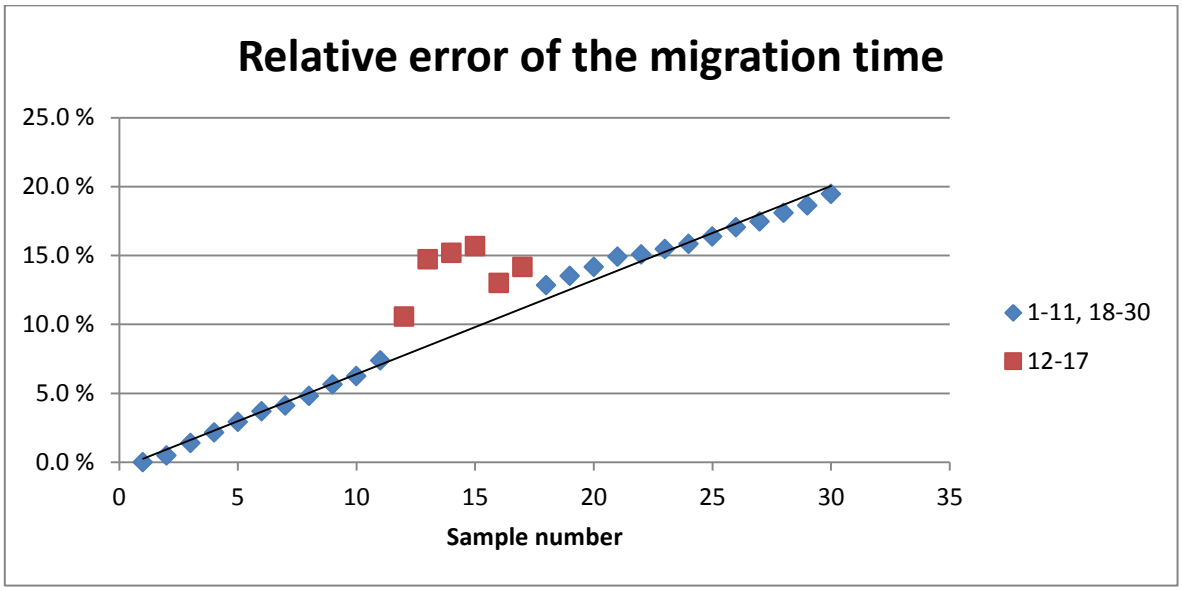


Figure 5. Relative error of the migration time against sample number in the 60 ppm polydatin sample CE series. The linear correlation curve intersects 10% acceptable value at over 15 injections. The sample repeats 12-17 were out of limits of the correlation and were not included in it.

3.4 Reproducibility

Reproducibility was studied by examining the difference of the repeated calibration series in each of the two calibration sample batches using the ANOVA- method [14] (see App. 4). The CE runs of repeated calibration series were lengthy, and the operating conditions could not be determined as constant throughout the entire series of runs. For example, the electrical current during the runs increased from ~66 µA to over 80 µA after over 30 hours of injections. Therefore, it is justifiable to address the term reproducibility, and in this case, within-laboratory reproducibility, technically.

When the whole concentration range was studied, the relative standard uncertainties were in the range of 10-14 %, and no systematic variation in difference between the results could be found. When the studied concentration range was reduced to the linear range of 5-75 ppm, the results were in correlation with the assumptions made from the plotted curves.

As can be seen from Table II, the series comparison with fewest injections, the 5-100 ppm series, gave the best reproducibility within the linear range. In the 0-120 ppm series, the uncertainty of the difference of the first and the second injection series was still somewhat acceptable. When a third series of injections was run, however, the uncertainty increased far too significantly; over 9 percentage-points.

3.5 Uncertainty components

Table II. The calculated relative standard uncertainty components from all of the conducted analyses of variance in different polydatin CE runs.

uncertainty component	c-range [ppm] /series	u(x)
manual integration	60 / 30 injections	1.0 %
repeatability	60 / 1&2	4.5 %
	60 / 2&3	3.2 %
	60 / 1&3	1.3 %
	60 / 10 vs 30 injections	6.8 %
reproducibility	0-100	12.1 %
	5-120/ 1&2	12.2 %
	5-120/ 2&3	10.0 %
	5-120/ 1&3	14.0 %
reproducibility In the linear range of 5-75 ppm	0-100	4.3 %
	5-120/ 1&2	7.3 %
	5-120/ 2&3	13.0 %
	5-120/ 1&3	16.6 %

Due to device issues in the 60 ppm runs, the uncertainties of the series containing the first third of the series of 30 repeats should be examined with caution. These difficulties were first encountered when the first attempt to produce a 60 ppm polydatin series failed and only the 9 first injections were run

successfully (see App: Table IV). This device malfunction seemed to impact on the following runs also, as the first third of the otherwise successful series of 30 repeats of the 60 ppm polydatin solution had a very high variance. The uncertainty caused by manual integration was calculated from two series of integrations (App. Table II). Therefore, if these unreliable results were sorted out, a combined uncertainty of the optimal repeatability was gained $U(x) = \sqrt{(1\ )^2 + (3.2)^2} = 3.35\ \%$.

4. Conclusions

The validation procedure with the particular CE method was found to be even more complex task than one could easily predict. The method's sensitivity to different factors both known and unknown caused uncertainty to the results and complicated the analysis of the results somewhat. Constant experimental conditions put aside, the most significant factors that could be controlled were the amount of conducted experiments and the concentrations of the injected solutions. Therefore, the design of the validation experiments was found to be a crucial part of the whole validation, which should not be rushed through.

From the results presented in this report, it can easily be stated, that the worst results are gained when massive amounts of repeats with too many levels of concentrations are conducted; the experiments should be designed to be as compact as possible, still producing the needed data for a full validation. The reason for the worsening of the results in a lengthy sequence was possibly the tarnishing of the buffer solution and the capillary, which alters the electrical current and finally the detector's functioning. Therefore, more than 24 hours consuming exhaustive experiments could not be recommended in any circumstances with the method in question.

Even with the discussed different issues along the validation procedure, the needed validation parameters and their uncertainties in different circumstances were ultimately gained. The adequacy of the reliability and the quantity of uncertainties of the gained results could be argued, but the picked optimal results were within reasonable uncertainty. Also, with the practical knowledge and the experience gained from this operation, a recreation of this kind of a study would be less of a "shot in the dark" procedure. With a well-designed procedure, at least 3-4 days of ongoing laboratory work, plus many working hours with the results analysis, a validation of this magnitude could be conducted.

References

[1] Analysis of resveratrol in wine by capillary electrophoresis, University of Nebraska, X. Gu, Q. Chu, M. O'Dwyer, M. Zeece, 2000
available: http://www.sciencedirect.com/science/article/pii/S0021967300002119

[2] Development of a fast capillary electrophoresis method for determination of carbohydrates in honey samples, Federal University of Santa Catarina, Rizelio VM, Tenfen L, da Silveira R, Gonzaga LV, Costa AC, Fett R., 2012
available: http://www.ncbi.nlm.nih.gov/pubmed/22483877

[3] Validation of a capillary electrophoresis method for the determination of cephradine and its related impurities, P. Emaldi, S. Fapanni, A. Baldini, 1995
available: http://www.sciencedirect.com/science/article/pii/002196739500520W

[4] Comparison of CE and HPLC Methods for Determining Lovastatin and Its Oxidation Products after Exposure to an Oxidative Atmosphere, S. J. Rajh, S. Kreft, B. Štrukelj, F. Vrečer, 2003
available: http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=151747

[5] Capillary Electrophoresis Determination, Synthesis, and Stability of Resveratrol and Related 3-*O*-β-d-Glucopyranosides, V. Brandolini , A. Maietti , P. Tedeschi , E. Durini , S. Vertuani , and S. Manfredini, 2002
available: http://pubs.acs.org/doi/abs/10.1021/jf0256384

[6] Mass spectrometry for the detection of differentially expressed proteins: a comparison of surface-enhanced laser desorption/ionization and capillary electrophoresis/mass spectrometry, N. Neuhoff, T. Kaiser, S. Wittke, R. Krebs, A. Pitt, A. Burchard, A. Sundmacher, B. Schlegelberger, W. Kolch, H. Mischak
available:
http://onlinelibrary.wiley.com/doi/10.1002/rcm.1294/abstract;jsessionid=B26D8B160AEE273BBAEF134D5290BB5E.f04t03?deniedAccessCustomisedMessage=&userIsAuthenticated=false

[7] Capillary Electrophoresis Guidebook: Principles, Operation, and Applications, K .D. Altria, 1996 ,p. 3-10
available:
http://www.google.fi/books?hl=fi&lr=&id=bW4YZYTEIIUC&oi=fnd&pg=PR5&dq=Capillary+Electrophoresis+Guidebook:+Principles,+Operation,+and+Applications&ots=9AhZ5e3sNV&sig=n6SiDOstV3B05KLZ9TDRemqifdo&redir_esc=y#v=onepage&q=Capillary%20Electrophoresis%20Guidebook%3A%20Principles%2C%20Operation%2C%20and%20Applications&f=false

[8] CE Separation Techniques, Agilent Technologies, 2013,
available: http://www.chem.agilent.com/en-US/Products-Services/Instruments-Systems/Automated-Electrophoresis/pages/gp833.aspx

[9] Resveratrol: A molecule whose time has come? And gone?, G. J. Soleas, E. P. Diamandis, D. M. Goldberg, 1997
available: http://www.sciencedirect.com/science/article/pii/S0009912096001555

[10] The antimicrobial effects of wood-associated polyphenols on food pathogens and spoilage organisms, C. P. Ferrer, K. Väkeväinen,H. Komulainen, M. Rautiainen, A. Smeds, J.-E. Raitanen, P. Eklund,S. Willför, H.-L. Alakomi, M. Saarela, A. von Wright, 2013
available: http://www.sciencedirect.com/science/article/pii/S0168160513001736

[11] Antimicrobial and cytotoxic knotwood extracts and related pure compounds and their effects on food-associated micro-organisms, A.-L. Välimaa, U. Honkalampi.-Hämäläinen, S. Pietarinen, S. Willför, B. Holmbom, A. von Wright, 2007
available: http://www.sciencedirect.com/science/article/pii/S0168160506005794

[12] Validation of Analytical Methods, Agilent Technologies, L. Huber, 2010, p. 14-28
available: http://www.chem.agilent.com/Library/primers/Public/5990-5140EN.pdf

[13] Internal Quality Control - Handbook for Chemical laboratories (Trollboken - Troll book), H. Hovind, B. Magnusson, M. Krysell, U. Lund, I. Mäkinen, 2011, p. 2-4
available: http://www.nordtest.info/images/documents/nt-technical-reports/nt%20tr%20569_ed4_en%20internal%20quality%20controll%20%20handbook%20for%20chemical%20laboratories.pdf

[14] EURACHEM / CITAC Guide CG 4 Quantifying Uncertainty in Analytical Measurement (Third Edition), S. L. R. Ellison, A. Williams, 2012
available: http://eurachem.org/images/stories/Guides/pdf/QUAM2012_P1.pdf

[15] Beta Mathematics Handbook for Science and Engineering, L. Råde, B. Westergren, 2004, p. 472-476

Appendix

App. 1, results

Table III. Results of a 50 ppm polydatin solution CE sequence.

Sample no.	t/h	$t_{mig.}$/min	A/mAU min
1	0.966667	22.207	90.7742
2	1.933333	22.612	93.9208
3	2.9	23.046	92.4784
4	3.866667	23.482	93.408
5	4.833333	23.783	96.3371
6	5.8	24.015	96.6837
7	6.766667	24.152	93.3022
8	7.733333	24.328	95.2779
9	8.7	24.478	94.5383

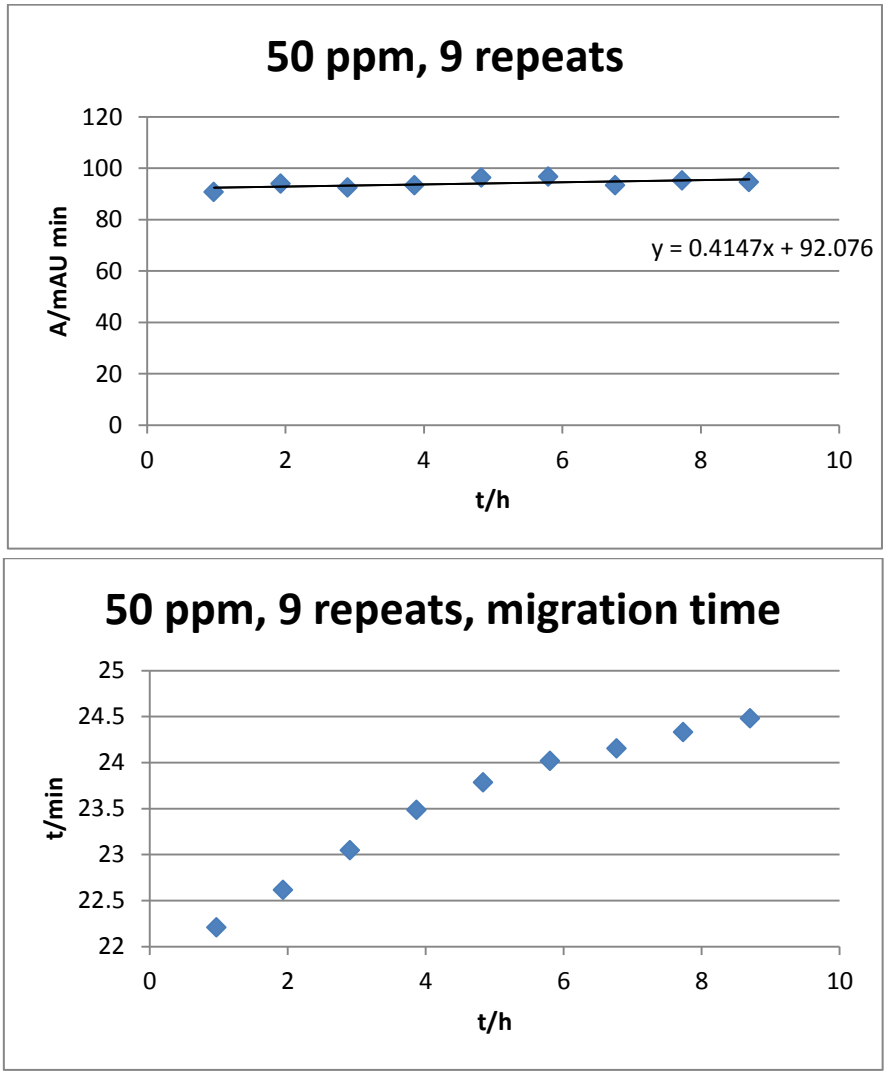
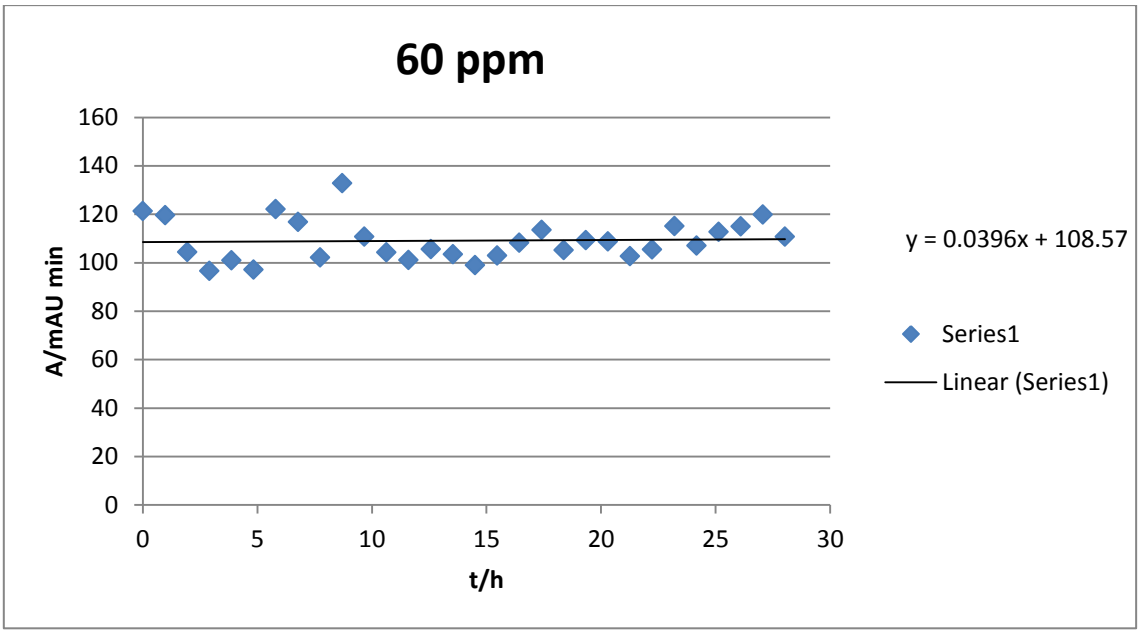


Figure 6. 50 ppm polydatin solution sequence peak areas and migration times plotted against time.

Table IV. Results of a 60 ppm polydatin solution CE sequence with integrations conducted twice.

		First integration		Second integration	
Sample nr.	t/h	$t_{mig.}$/min	A/mAU min	$t_{mig.}$/min	A/mAU min
1	0.966667	23.977	120.414	23.976	120.137
2	1.933333	24.088	113.607	24.088	112.058
3	2.9	24.31	97.6511	24.31	101.817
4	3.866667	24.496	90.9956	24.494	91.8527
5	4.833333	24.677	97.4329	24.677	93.5333
6	5.8	24.859	89.2389	24.859	93.7769
7	6.766667	24.957	125.435	24.957	120.364
8	7.733333	25.114	108.572	25.125	109.783
9	8.7	25.313	98.6108	25.324	95.8158
10	9.666667	25.473	129.497	25.473	125.356
11	10.63333	25.735	101.125	25.746	106.283
12	11.6	26.512	96.6607	26.506	99.0172
13	12.56667	27.498	92.6966	27.504	97.0835
14	13.53333	27.614	99.6449	27.614	102.156
15	14.5	27.731	97.58	27.733	96.7403
16	15.46667	27.09	99.4905	27.091	94.5388
17	16.43333	27.365	96.502	27.369	93.2674
18	17.4	27.061	98.4097	27.054	101.493
19	18.36667	27.207	105.39	27.212	107.112
20	19.33333	27.371	99.729	27.371	101.79
21	20.3	27.548	98.1038	27.548	102.172
22	21.26667	27.596	98.3983	27.585	101.053
23	22.23333	27.677	93.3408	27.684	98.3794
24	23.2	27.771	95.4034	27.771	100.141
25	24.16667	27.905	101.495	27.901	108.875
26	25.13333	28.056	101.7	28.059	101.693
27	26.1	28.158	106.518	28.162	106.074
28	27.06667	28.305	107.013	28.313	105.607
29	28.03333	28.44	116.916	28.44	116.118
30	29	28.636	100.424	28.64	105.336



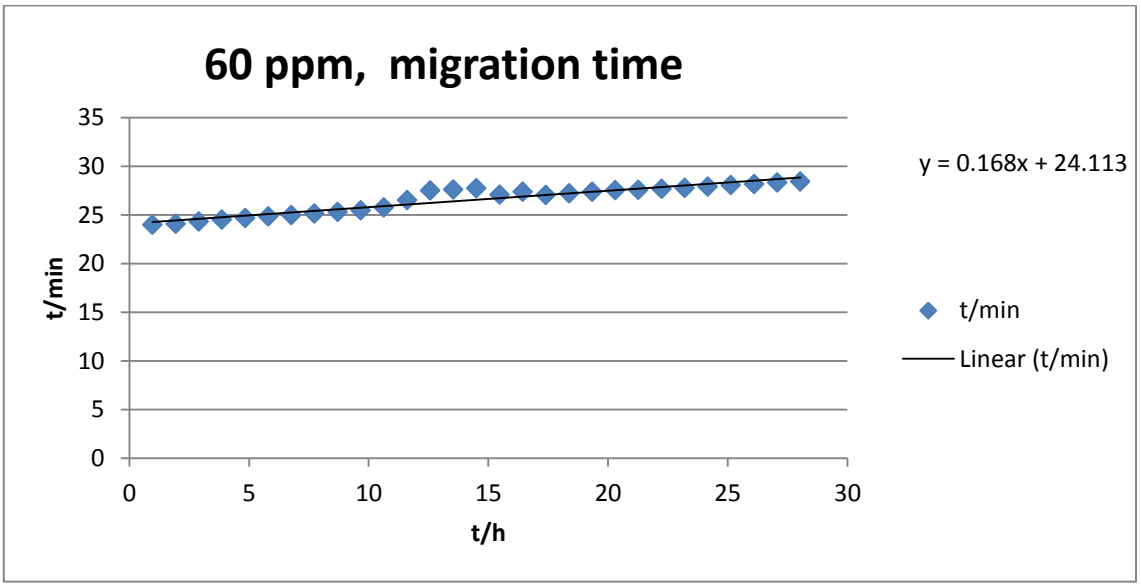
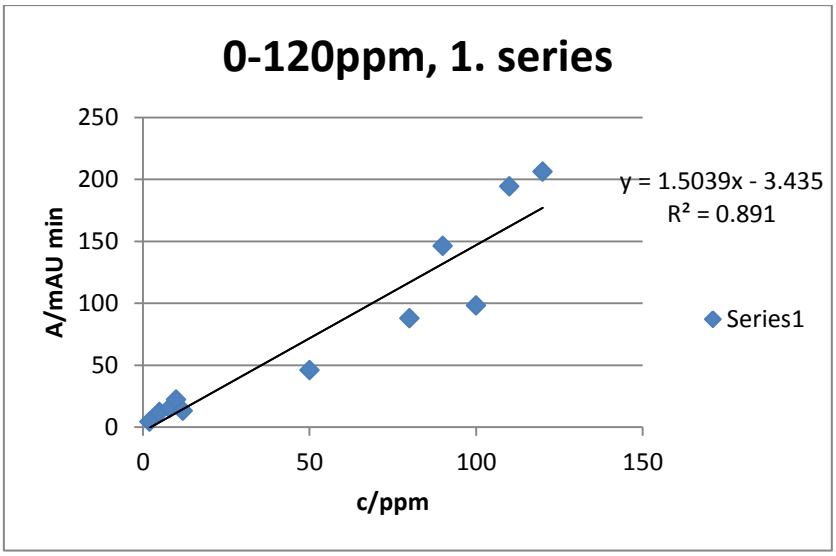
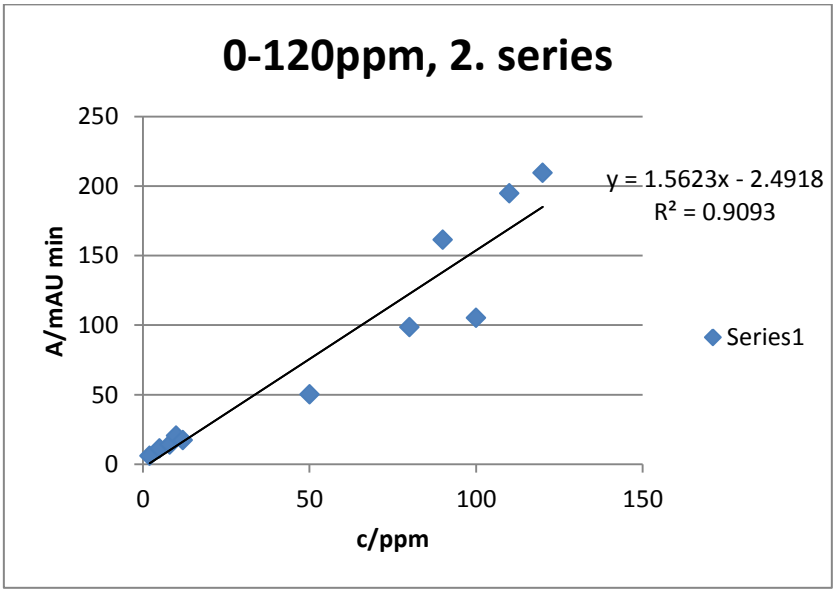


Figure 7. 60 ppm polydatin solution sequence peak areas and migration times plotted against time.

Table V. Results of the CE sequence with three series of injections with concentration levels ranging 2-120 ppm polydatin with two integrations.

1. integration	**1**		**2**		**3**	
c/ppm	$t_{mig.}$/min	A/mAU min	$t_{mig.}$/min	A/mAU min	$t_{mig.}$/min	A/mAU min
2	24.994	3.66922	26.855	2.27161	28.94	3.36453
5	25.064	11.1931	26.975	9.58417	29.014	12.1425
8	25.166	13.6651	27.136	13.1021	29.238	15.0651
10	25.316	19.208	27.335	19.4945	29.491	18.2279
12	25.563	11.5478	27.44	12.7553	29.681	19.2127
50	25.679	42.0019	28.256	41.8235	29.876	62.4952
80	26.112	82.5162	27.803	94.1543	30.152	134.02
90	26.326	135.478	27.921	148.4	30.389	186.996
100	26.233	92.1625	27.957	100.466	30.502	124.777
110	26.419	184.344	28.247	183.811	30.8	229.679
120	26.577	192.078	28.476	211.018	30.996	243.934
2. integration	**1**		**2**		**3**	
c/ppm	$t_{mig.}$/min	A/mAU min	$t_{mig.}$/min	A/mAU min	$t_{mig.}$/min	A/mAU min
2	25.067	4.23109	26.811	6.02671	28.94	4.44529
5	25.055	11.9853	26.987	11.3522	29.014	12.639
8	25.166	15.1874	27.134	14.0927	29.238	18.0586
10	25.328	22.1687	27.335	20.6723	29.489	24.6994
12	25.63	13.2295	27.442	17.3124	29.681	20.5082
50	25.695	45.8052	28.256	50.3951	29.876	61.0369
80	26.114	87.7167	27.792	98.6617	30.158	132.886
90	26.335	146.161	27.92	161.469	30.389	196.412
100	26.236	98.1881	27.946	105.367	30.51	123.238
110	26.42	194.147	28.247	194.781	30.812	226.209
120	26.579	206.17	28.476	209.519	30.985	240.76





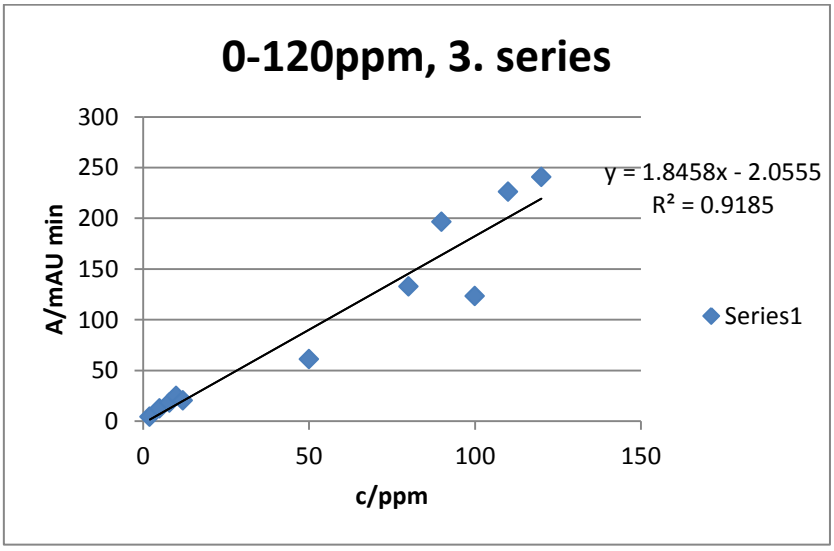


Figure 8. Peak areas of the three 2-120 ppm polydatin series plotted against concentration, with correlation coefficients and linear equations.

Table VI. Results of the CE sequence with two series of injections with concentration levels ranging 5-100 ppm polydatin.

c/ppm	$t_{mig.}$/min	A/mAU min	$t_{mig.}$/min	A/mAU min
5	26.989	12.8078	29.351	11.2769
40	27.638	56.1941	29.523	54.4132
75	27.673	122.57	30.458	106.076
80	28.136	110.626	29.625	141.562
85	28.269	106.169	30.774	132.171
90	28.323	162.464	30.784	137.74
100	28.647	161.067	31.399	151.585

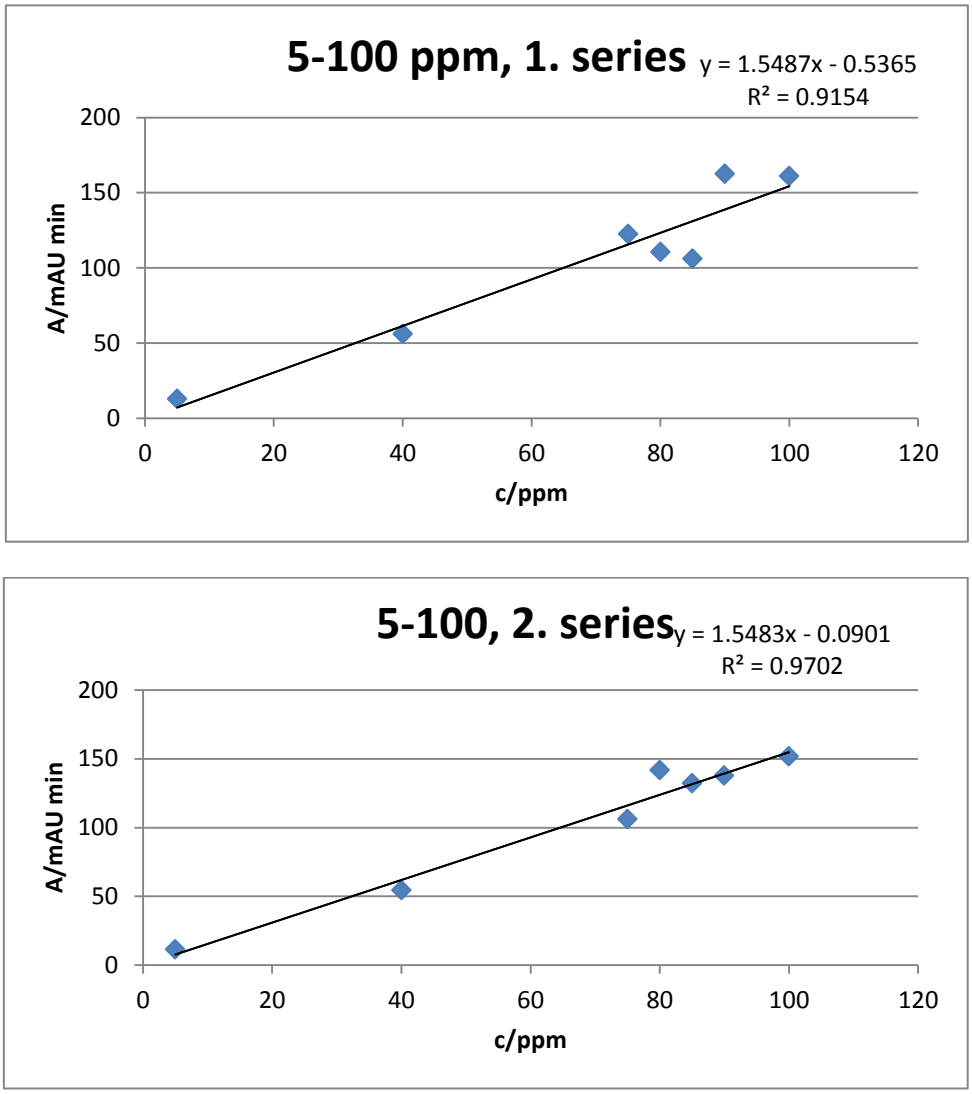


Figure 9. Peak areas of the two 5-100 ppm polydatin series plotted against concentration, with correlation coefficients and linear equations.

App. 2, LOD and LOQ calculation

Table VII. Results of the zero-sample integrations from the 0-120 ppm sequence.

sample no.	$t_{mig.}$/min	A/mAU min
2	25.143	1.6355
14	25.055	2.5413
15	25.087	1.54531
28	25.075	2.00469
average		1.9317

A simple calibration curve intersecting the average of the zero sample integrations (table VII) and the average of the 60 ppm results (table IV) was first determined (fig. 10).

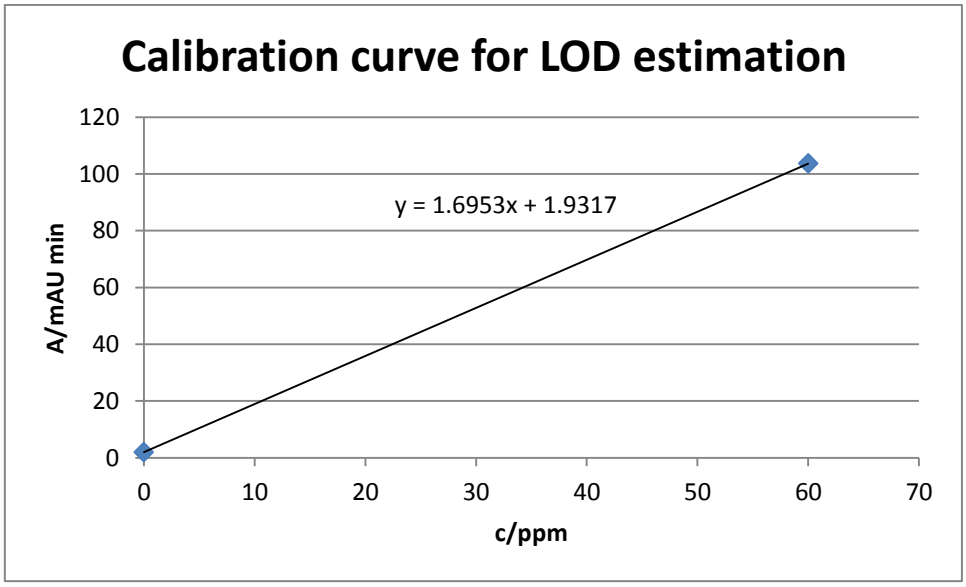


Figure 10. A linear curve intersecting the points of noise level average and the 60 ppm sequence average.

Now the LOD could be determined using the gained slope and the signal/noise-ratio (k) of 3.

$$LOD = \frac{(k\ A_{av.noise} - \ A_{av.noise})}{slope} = \frac{(3 * \ 1{,}9317 - \ 1{,}9317)}{1.6953} \approx 2{,}279\ ppm$$

LOQ could be determined by equation

$$LOQ = 3\ LOD$$

For further study, differences of two integrations from the 0-120 ppm sequence were determined.

Table VIII. The averages from 2 different integrations of the same 0-120 ppm sequences (table V) and their relative errors of the standard deviations. The smallest variation in the results was found to occur in the range of 5-8 ppm.

series	**1**			**2**			**3**		
c [ppm]	avg.	stdev	rel. stdev	avg.	stdev	rel. stdev	avg.	stdev	rel. stdev
2	3.950155	0.397302	10.06 %	4.14916	2.655257	64.00 %	3.90491	0.764213	19.57 %
5	11.5892	0.56017	4.83 %	10.46819	1.250186	11.94 %	12.39075	0.351079	2.83 %
8	14.42625	1.076429	7.46 %	13.5974	0.70046	5.15 %	16.56185	2.116724	12.78 %
10	20.68835	2.093531	10.12 %	20.0834	0.83283	4.15 %	21.46365	4.576042	21.32 %
12	12.38865	1.189141	9.60 %	15.03385	3.222356	21.43 %	19.86045	0.916057	4.61 %
50	43.90355	2.689339	6.13 %	46.1093	6.061036	13.14 %	61.76605	1.031174	1.67 %

rel. error	**from**	**series**
c [ppm]	**1,2,3**	**1,2**
2	31.21 %	37.03 %
5	**6.54 %**	**8.39 %**
8	**8.46 %**	**6.31 %**
10	11.86 %	7.13 %
12	11.88 %	15.52 %
50	6.98 %	9.64 %

App. 3, Estimating the linear concentration range

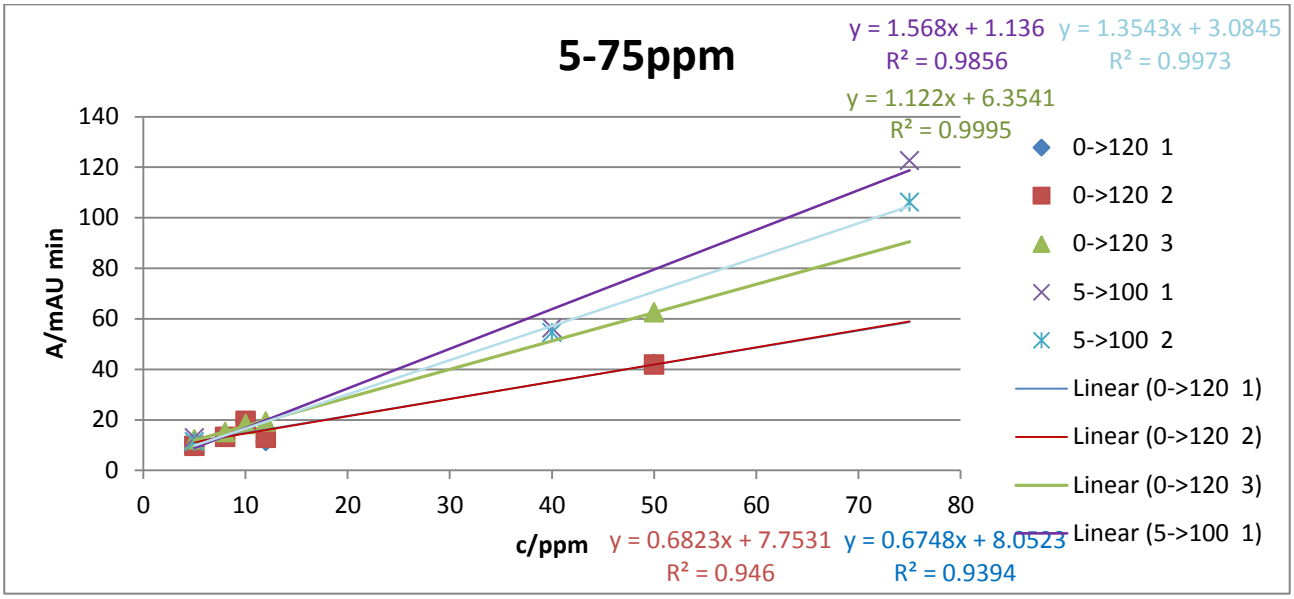


Figure 11. Calibration series' results plotted in the range of approximated linearity, which gave the best correlation coefficients.

App. 4, Reproducibility studied by ANOVA method

Table IX. The calculation and the results of ANOVA analysis from every series comparison that was conducted and included in the report. u(x) depicts the uncertainty in each case.

0-120 series 1&2						
c/ppm	x_1	x_2	$x_{k.a.}$	$d = x_1 - x_2$	$d_r = d/x_{k.a.}$	$(d_{ri}-d_{r,ka})^2$
2	3.66922	2.27161	2.970415	1.39761	0.47051	0.2081148
5	11.1931	9.58417	10.38864	1.60893	0.1548741	0.0197571
8	13.6651	13.1021	13.3836	0.563	0.0420664	0.0007702
10	19.208	19.4945	19.35125	-0.2865	-0.014805	0.0008479
12	11.5478	12.7553	12.15155	-1.2075	-0.09937	0.0129241
50	42.0019	41.8235	41.9127	0.1784	0.0042565	0.0001012
80	82.5162	94.1543	88.33525	-11.6381	-0.131749	0.0213345
90	135.478	148.4	141.939	-12.922	-0.091039	0.0110993
100	92.1625	100.466	96.31425	-8.3035	-0.086213	0.0101056
110	184.344	183.811	184.0775	0.533	0.0028955	0.0001304
120	192.078	211.018	201.548	-18.94	-0.093973	0.011726
avg					0.014314	
s_{dr}						0.1723111
u(x)						0.1218423

Within linear range						
c/ppm	x_1	x_2	$x_{k.a.}$	$d = x_1 - x_2$	$d_r = d/x_{k.a.}$	$(d_{ri}-d_{r,ka})^2$
5	11.1931	9.58417	10.388635	1.60893	0.1548741	0.02635059
8	13.6651	13.1021	13.3836	0.563	0.0420664	0.00245233
10	19.208	19.4945	19.35125	-0.2865	-0.0148052	5.4032E-05
12	11.5478	12.7553	12.15155	-1.2075	-0.09937	0.00844845
50	42.0019	41.8235	41.9127	0.1784	0.0042565	0.00013715
80	82.5162	94.1543	88.33525	-11.6381	-0.1317492	0.01544915
avg					-0.0074546	
s_{dr}						0.10285106
u(x)						0.07272668

0-120 series 2&3						
c/ppm	x_1	x_2	$x_{k.a.}$	$d = x_1 - x_2$	$d_r = d/x_{k.a.}$	$(d_{ri}-d_{r,ka})^2$
2	2.27161	3.36453	2.81807	-1.09292	-0.387826	0.0213747
5	9.58417	12.1425	10.86334	-2.55833	-0.235501	3.75E-05
8	13.1021	15.0651	14.0836	-1.963	-0.139382	0.0104536
10	19.4945	18.2279	18.8612	1.2666	0.0671537	0.0953442
12	12.7553	19.2127	15.984	-6.4574	-0.403991	0.026363
50	41.8235	62.4952	52.15935	-20.6717	-0.396318	0.0239301
80	94.1543	134.02	114.0872	-39.8657	-0.349432	0.0116224
90	148.4	186.996	167.698	-38.596	-0.230152	0.0001316
100	100.466	124.777	112.6215	-24.311	-0.215865	0.0006636
110	183.811	229.679	206.745	-45.868	-0.221858	0.0003907
120	211.018	243.934	227.476	-32.916	-0.144701	0.0093942
avg					-0.241625	
s_{dr}						0.1413172
u(x)						0.0999264

Within linear range						
c/ppm	x_1	x_2	$x_{k.a.}$	$d = x_1 - x_2$	$d_r = d/x_{k.a.}$	$(d_{ri}-d_{r,ka})^2$
5	9.58417	12.1425	10.863335	-2.55833	-0.2355013	5.4916E-05
8	13.1021	15.0651	14.0836	-1.963	-0.139382	0.01071844
10	19.4945	18.2279	18.8612	1.2666	0.0671537	0.09614068
12	12.7553	19.2127	15.984	-6.4574	-0.4039915	0.02594664
50	41.8235	62.4952	52.15935	-20.6717	-0.3963182	0.0235335
80	94.1543	134.02	114.08715	-39.8657	-0.349432	0.01134653
avg					-0.2429119	
s_{dr}						0.18316152
u(x)						0.12951476

0-120 series 1&3

c/ppm	x_1	x_2	$x_{k.a.}$	$d = x_1 - x_2$	$d_r = d/x_{k.a.}$	$(d_{ri}-d_{r,ka})^2$
2	3.66922	3.36453	3.516875	0.30469	0.0866366	0.0975814
5	11.1931	12.1425	11.6678	-0.9494	-0.081369	0.020844
8	13.6651	15.0651	14.3651	-1.4	-0.097458	0.0164571
10	19.208	18.2279	18.71795	0.9801	0.0523615	0.0773425
12	11.5478	19.2127	15.38025	-7.6649	-0.49836	0.0743196
50	42.0019	62.4952	52.24855	-20.4933	-0.392227	0.0277167
80	82.5162	134.02	108.2681	-51.5038	-0.475706	0.0624812
90	135.478	186.996	161.237	-51.518	-0.319517	0.0087935
100	92.1625	124.777	108.4698	-32.6145	-0.300678	0.0056152
110	184.344	229.679	207.0115	-45.335	-0.218997	4.551E-05
120	192.078	243.934	218.006	-51.856	-0.237865	0.0001469
avg					-0.225744	
s_{dr}						0.1978241
u(x)						0.1398828

Within linear range

c/ppm	x_1	x_2	$x_{k.a.}$	$d = x_1 - x_2$	$d_r = d/x_{k.a.}$	$(d_{ri}-d_{r,ka})^2$
5	11.1931	12.1425	11.6678	-0.9494	-0.0813692	0.02803079
8	13.6651	15.0651	14.3651	-1.4	-0.0974584	0.02290222
10	19.208	18.2279	18.71795	0.9801	0.0523615	0.09069417
12	11.5478	19.2127	15.38025	-7.6649	-0.4983599	0.06228353
50	42.0019	62.4952	52.24855	-20.4933	-0.3922272	0.02057329
80	82.5162	134.02	108.2681	-51.5038	-0.4757061	0.05148947
avg					-0.2487932	
s_{dr}						0.23493551
u(x)						0.16612449

5-100 series						
c/ppm	x_1	x_2	$x_{k.a.}$	$d = x_1 - x_2$	$d_r = d/x_{k.a.}$	$(d_{ri}-d_{r,ka})^2$
5	12.8078	11.2769	12.04235	1.5309	0.1271263	0.0127266
40	56.1941	54.4132	55.30365	1.7809	0.0322022	0.00032
75	122.57	106.076	114.323	16.494	0.1442754	0.01689
80	110.626	141.562	126.094	-30.936	-0.245341	0.0674206
85	106.169	132.171	119.17	-26.002	-0.218192	0.0540593
90	162.464	137.74	150.102	24.724	0.1647147	0.0226204
100	161.067	151.585	156.326	9.482	0.0606553	0.0021475
avg					0.0093487	
s_{dr}						0.1713594
u(x)						0.1211694

Within linear range						
c/ppm	x_1	x_2	$x_{k.a.}$	$d = x_1 - x_2$	$d_r = d/x_{k.a.}$	$(d_{ri}-d_{r,ka})^2$
5	12.8078	11.2769	12.04235	1.5309	0.1271263	0.00067211
40	56.1941	54.4132	55.30365	1.7809	0.0322022	0.00476088
75	122.57	106.076	114.323	16.494	0.1442754	0.00185538
avg					0.1012013	
s_{dr}						0.06036706
u(x)						0.04268595

App. 5 Repeatability

5.1 T-test

The t-test analysis calculation and results for 1. and 2. series from the 0-120 ppm sequence is presented as an example. All table values were gained from Beta statistics tables [15].

$\bar{x}_1 = 106{,}44937$

$\bar{x}_2 = 99{,}94812$

$s_1 = 12{,}71923$

$s_2 = 4{,}659057$

$\nu_1 = \nu_2 = 9$

Combined standard deviation: $s_c = \sqrt{\frac{\nu_1 {s_1}^2 + \nu_2 {s_2}^2}{\nu_1 + \nu_2}} = \sqrt{\frac{9 * 12{,}71923^2 + 9 * 4{,}659057^2}{9+9}} \approx 9{,}578$

Combined mean: $s_{\bar{x}_1 - \bar{x}_2} = s_c \sqrt{\frac{1}{\nu_1 + 1} + \frac{1}{\nu_2 + 1}} \approx 4{,}284$

F-test for the determination of whether variances are equal or not:

$$F = \left(\frac{s_1}{s_2}\right)^2 = \left(\frac{12{,}71923}{4{,}659057}\right)^2 \approx 7{,}45$$

Comparison to table value: $7{,}45 > 3{,}18$ The difference is significant.

In this case, the combined degree of freedom (Satterthwait equation):

$$\nu_c = \frac{{s_{\bar{x}_1 - \bar{x}_2}}^4}{\frac{\left(\frac{s_1}{\sqrt{\nu_1 + 1}}\right)^4}{\nu_1} + \frac{\left(\frac{s_2}{\sqrt{\nu_2 + 1}}\right)^4}{\nu_1}} = \frac{4{,}284}{\frac{\left(\frac{12{,}71923}{\sqrt{9+1}}\right)^4}{9} + \frac{\left(\frac{4{,}659057}{\sqrt{9+1}}\right)^4}{9}} \approx 11$$

The t-test value can now be computed: $t = \frac{(\bar{x}_1 - \bar{x}_2) - \mu}{s_c} = \frac{(106{,}44937 - 99{,}94812) - 0}{9{,}578} \approx 1{,}5$

With confidence level of 95%, the two-tailed t-test value with $\nu_c = 11$ $t_{0,95} = 2{,}3646 > 1{,}5$

The t-value comparison shows, that the two series have no significant statistical difference, and the sequence is repeatable after two series of injections.

5.2 Relative standard deviations

Table X. Relative standard deviation in relation to the number of samples. The data was used in figure 4.

n_i	A/mAU min	s_i	$\bar{x}_i$	$\left(\frac{s_i}{\bar{x}_i}\right)/\sqrt{n}$
1	106.283			
2	99.0172	5.13769645	102.6501	3.54 %
3	97.0835	4.85046124	100.794567	2.78 %
4	102.156	4.01846049	101.134925	1.99 %
5	96.7403	3.99669417	100.256	1.78 %
6	94.5388	4.26925998	99.3031333	1.76 %
7	93.2674	4.51587356	98.4408857	1.73 %
8	101.493	4.31789731	98.8224	1.54 %
9	107.112	4.89376969	99.7434667	1.64 %
10	101.79	4.65905717	99.94812	1.47 %
11	102.172	4.47054092	100.150291	1.35 %
12	101.053	4.27045173	100.225517	1.23 %
13	98.3794	4.12058167	100.083508	1.14 %
14	100.141	3.95895663	100.087614	1.06 %
15	108.875	4.43865858	100.67344	1.14 %
16	101.693	4.2957203	100.737163	1.07 %
17	106.074	4.35606349	101.051094	1.05 %
18	105.607	4.36030071	101.3042	1.01 %
19	116.118	5.43193499	102.083874	1.22 %
20	105.336	5.33683342	102.24648	1.17 %

5.3 Migration times

Table XI. Relative error of the migration time in 60 ppm sequence in relation to sample number. The data was used in figure 5.

n	t/min	$\frac{t - t_0}{t_0}$
1	23.976	
2	24.088	0.467 %
3	24.31	1.393 %
4	24.494	2.160 %
5	24.677	2.924 %
6	24.859	3.683 %
7	24.957	4.092 %
8	25.125	4.792 %
9	25.324	5.622 %
10	25.473	6.244 %
11	25.746	7.382 %
12	26.506	10.552 %
13	27.504	14.715 %
14	27.614	15.174 %
15	27.733	15.670 %
16	27.091	12.992 %
17	27.369	14.152 %
18	27.054	12.838 %
19	27.212	13.497 %
20	27.371	14.160 %
21	27.548	14.898 %
22	27.585	15.053 %
23	27.684	15.465 %
24	27.771	15.828 %
25	27.901	16.371 %
26	28.059	17.030 %
27	28.162	17.459 %
28	28.313	18.089 %
29	28.44	18.619 %
30	28.64	19.453 %

App. 6 Recommendations for the validation of a CE method

1. Introduction

The following recommendations are based on an in-house validation performed with polydatin solutions with known concentrations, purified water and a known capillary electrophoresis method. Therefore, recommendations made from these results may not apply fully in operating conditions that differ from the described validation. These conditions may include for example different type of molecule under examination, increased concentration, a sample solution with organic impurities, a different buffer solution used and therefore altered pH, etc. Thus, these recommendations are best to be followed when studying molecules similar to polydatin in chemical respect, using the exact same CE method with concentrations not significantly higher than 100 ppm and not lower than 5 ppm. Of course, many of the following guidelines are universal considering a typical CE sequence, but still, care should be taken in strict application of these directives. Practical utilization of the Agilent CE sequence procedure and the use of the monitoring software are not discussed here. The following recommendations are meant for sequences with the single run lasting 60 minutes, and therefore, a 30 min run, for instance, would probably mean totally different design of experiments.

2. Validation

As usually, the experiments need to be designed to correspond to the requirements of the results especially in respect of validation parameters. Determination of at least approximate values of Limit of Detection (LOD), Limit of Quantification (LOQ) and Linearity range is universally necessary in chromatography. Also, when using this particular CE method, it is highly recommendable to determine repeatability and within-laboratory reproducibility to design and produce successful and accurate experiments.

Bearing the above in mind, experiments should start with series dedicated to the validation of the method, conducted with a reference chemical in a pure matrix.

2.1 Linearity, LOD, LOQ and reproducibility

To study linearity, design a calibration series with a concentration range spanning around 80-120 % of the expected minimum and maximum of the linearity zone. It is recommendable not to conduct more than 7-8 different concentration levels. For example, in an expected linearity zone of 5-100 ppm, a calibration series could consist of concentrations of 2, 5, 8, 10, 50, 90, 100 and 110 ppm.

In respect to reproducibility, the calibration series should be run no more than three times. Also, a sample consisting of only purified water should be injected preferably 2 times before each of the calibration series runs.

With a calibration series of the example above, the full run procedure would consist of a zero- sample in the beginning followed with a calibration series run, after which the zero-sample would be injected 2 times, then a calibration series run, 2 zeros and finally the third repeat of the same calibration series.

2.2 Repeatability

To determine repeatability, design a separate series consisting of repeats of the same concentration level that is inside the linearity range. In the expected linearity range in the example above, the repeated solution could be 60 ppm, for example. In terms of migration time repeatability, this series should consist of no more than 10-15 repeated injections. Also, a zero- sample should be injected before the series, preferably 2 times.

3. Preparations

Every step from now on can produce some level of error in the final results, and therefore, care has to be taken in every step of the CE runs.

3.1 Reference samples

Calculate the needed mass of the reference chemical to produce a working solution of a desired concentration. A concentration of 500-1000 ppm ($mg/l_{solvent}$) in a 100-1000 ml volumetric flask is recommendable. Use only purified water and flasks cleaned with purified water. If the reference solution is difficult to dilute, use lower concentrations, for example 200 ppm, or use an ultrasonic bath.

Prepare the needed solutions using clean micro-pipettes. To minimize error from the dispensing of small quantities of the working solution, it is recommendable to use volumetric flasks at least 20 ml. Use vials and caps that are specifically designed for the Agilent CE device. The use of a fresh Millipore membrane filter and a fresh dispensing syringe for each solution is recommendable. Also, the syringe and the filter should be rinsed thoroughly with the solution in question before injecting it into the vial. This is especially important when samples with unknown concentrations in impure solvents are analyzed.

3.2 The buffer solution

For carbohydrate compounds analysis with the CE run method in question, an alkaline buffer solution of 130 mM NaOH and 36 mM Na_2HPO_4 in purified water (pH 12,6) has been used. Prepare the buffer solution in an ultrasonic bath for 20 min before inserting it into the vials.

3.3 CE equipment

Prepare a fresh polyimide-coated silica capillary before any validation series are run. In this example, 70 cm should be adequate. Take care not to tarnish the capillary when preparing the point of detection by burning off the outer layer.

Any other piece of apparatus in direct contact with the capillary or the detector should also be checked and cleaned before use. Per se, the CE device should be in “out of the box” shape when conducting any experiments.

4. Electropherogram analysis

To gain large enough responses (peak areas), and therefore to keep error minimal, use the best possible detector wavelength for the desired compound. Manual integration of the electropherogram peaks may cause some error to the results, especially in the case of an inexperienced user. Therefore, it is recommendable to conduct the integration at least twice, and determine the difference of these results. Always make sure that the peak is the correct one, and not one caused by for example the buffer solution. This is easily verified by comparing the migration times in the calibration curves. Attention should also be paid to the similarity in the way that the peaks are being integrated: always try to draw the baseline between the same points of the curve.

To gain approximate values for the response of the background noise, several zero-samples should be analyzed in different points of the runs. Bring up the wanted zero-sample electropherograms side by side, and integrate the same migration time interval in every gram in the similar way (figure 1).

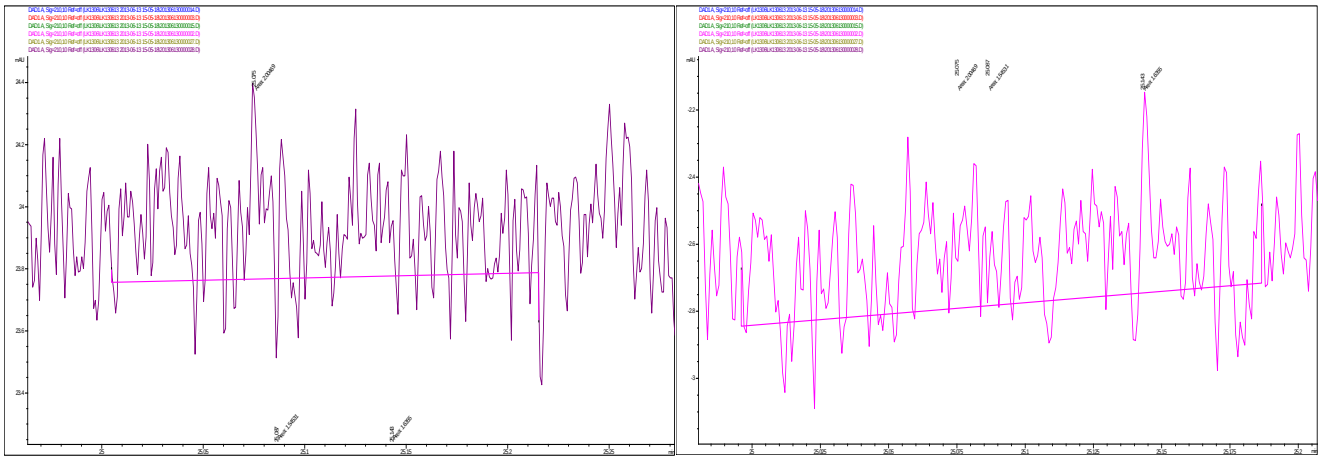


Figure 1. Two electropherograms of different zero-samples integrated in the similar way.

In addition to determining the migration times and the peak areas, a visual analysis of the peaks themselves is important especially in LOD and LOQ determination.

5 Calculations

Recommendations for the determination of the validation parameters by calculative means are presented in the following. Rather than expressing the exact means to determine these parameters, references to appropriate sources are given.

5.1 LOD and LOQ

A definition for LOD can be found for example in the IUPAC Goldbook [1], although other definitions are known. If LOD is calculated by using the mean value of the zero-sample responses above, it is recommendable to use a signal-to-noise ratio of 3 to gain an approximate value for the LOD response value. This value is then converted into concentration by using the slope gained from a curve beginning not from the origin, but from a value gained as a mean from the zero samples, and intersecting a point of a mean of responses that were gained by repeating the same concentration. LOQ can then be determined by multiplying the LOD by an appropriate factor, which is recommended as 4.

5.2 Repeatability and reproducibility

It is highly recommended to study repeatability from two different angles: the increasing migration time, and the peak area variation. To study the repeatability of the migration time, calculate their relative errors in relation to the number of repeats. As long as the relative error stays under 10%, the series are repeatable. In the case of a run that has no significant outliers, the peak area repeatability can be studied by determining relative standard deviation in relation to time.

Reproducibility is best studied by carrying out the ANOVA-method [2] to compare the calibration run series. Also, student's t-tests can be carried out, although this method does not properly reveal the magnitude of the error caused by reproducibility. If plenty of repeats (>10-15) of the mid-level concentration (60 ppm) were carried out, repeatability can also be studied using the ANOVA method by splitting the series in 2 or more parts that are compared with one another.

5.3 Linearity

No exact standard of the calculative methods to define linearity exists, although recommendations can be found for example in the Agilent validation guide [3]. Plot the peak areas from the calibration curve series against the input concentrations, and using appropriate software (Excel), conduct least-squares fit through the data points. The user should be able to determine an approximate range of linearity by observing the data points. To further study the linearity range, several LSQ- fits can be conducted through different sets of data points, and the best fit by the R-value usually gives the best approximation. Take care not to eliminate other than clear outlier data points from the analysis.

References (App. 6)

[1] IUPAC Compendium of Chemical Terminology (Goldbook), 2012, p.839
available: http://goldbook.iupac.org/PDF/goldbook.pdf

[2] Qualifying Uncertainty in Analytical Measurement, Eurachem, S L R Ellison, A Williams, 2012, p. 19
available: http://eurachem.org/images/stories/Guides/pdf/QUAM2012_P1.pdf

[3] Validation of Analytical Methods (Agilent), Ludwig Huber, 2010, p. 20-22
available: http://www.chem.agilent.com/Library/primers/Public/5990-5140EN.pdf