5 Chapter Five: Validation of the glycoanalytical methods

To evaluate the analytical limits of the methods by a rational statistical approach, a validation procedure was performed. This was also necessary in order to transformate the methods to routine analytical techniques for batch-to-batch-consistency-analysis at ProBioGen.

In dependence of the clinical phase the company is producing API (active pharmaceutical ingredient) for, the validation effort of the process has to be adapted. The further the product is going into the different development stages, the more validation effort is necessary to make the whole process more secure and reliable. It is like a 'validation gradient' that the company has to pass in order to convince the regulatory authorities that the production process including all quality control methods is working properly under the pre-defined conditions.

In this work, I have validated the analytical methods described above for the pre-clinical stage up to clinical phase 1. Validation data shown in the following chapter corresponded to a pre-validation study to determine the acceptance criteria for the formal validation of the methods. The confirmation of these acceptance criteria is generally understood as the formal validation process but is not shown here because it would go beyond the scientific scope of this thesis. For this phase of development the validation criteria are significantly different from those for the higher stages of development and the phase after market-approval [150]. According to ICH and FDA guidance, the objective of method validation is to demonstrate that analytical procedures are suitable for their intended purpose. In initial clinical trials the purpose is to determine the safe dosing range and key pharmacological data. The purpose of the pharmaceutical product in early phases is to deliver a known dose that is bioavailable. Bioavailability corresponds directly to the pharmacokinetic characteristics of the drug molecule. As mentioned in the literature review (see Chapter Two), the pharmacokinetic characteristics of recombinant glycoproteins often depend on their glycosylation patterns and therefore glycosylation is one of the most important topics for examinations before clinical phases should start.

In general, before the validation of an analytical method can be started, the devices used for the method have to be validated. Validation of devices is defined as qualification. Qualification data of the HPLC-device is shown in the Appendix.

5.1 Validation guidelines

In this chapter, the applicable official recommendations for validation of chromatographical methods are introduced and judged regarding their relevance for this study. General definitions of validation parameters are shown in the Appendix.

Within the European Union, the Pharmacopeia commission has published a technical guide for the elaboration of monographs which includes, among other things, different proposals for validation. It was published in 1993 for the first time and is constantly actualized.

The aim is the development of a guide for analytical methods and procedures which is also able to be harmonized internationally. That is why the commission refers their definitions to the proposals of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, in which drug application procedures in the EU, USA and Japan shall be adapted to each other.

In the document of the Ph.Eur.-commission, test procedures, reagents, quantities and equipment are described. In this chapter, the general recommendations for the validation of analytical procedures are only considered insofar as they are crucial for liquid chromatographical procedures. In principle, liquid chromatographical procedures are distinguished by their purpose, especially identity-, purity-, or quantitative analysis. The purpose defines the extent of validation (Table 18).

Table 18: Validation of liquid chromatographical procedures

Validation parameter	Identification	Impurities, quantitative	Impurities, qualitative	Quantitative analysis
Accuracy	no	**yes**	no	**yes**
Reproducibility	no	**yes**	no	**yes**
Comparability	no	if applicable	no	if applicable
Specificity [1]	**yes**	**yes**	**yes**	**yes**
Detection limit	no	no	**yes**	no
Quantification limit	no	**yes**	no	no
Linearity	no	**yes**	no	**yes**
Linear Range (Working area)	no	**yes**	no	**yes**
Robustness	no	**yes**	if applicable	**yes**

[1] If specificity cannot be shown, a second (specific) analytical method can be used.

As mentioned above, the most important guidelines for analytical method validation in pharmaceutical industry, are the ICH (International Conference on Harmonisation) guidelines, especially the guidelines Q2a and Q2b. These guidelines help to understand the validation principles and how to calculate the most important validation parameters, but because of their guideline-character, they do not go into detail too much.

For drug substance method validation in early development, the table has to be adapted as follows (Table 19) [150].

Definitions of individual validation parameters are shown in the Appendix.

Table 19: Recommended drug substance method validation in early development

Validation parameter	Assay	Organic Impurities
Accuracy	Inferred from precision, linearity & specificity	Inferred from precision, linearity & specificity
Precision	From 3 sample preparations at 100% of test concentration	From 3 sample preparations at 100% of test concentration
Intermediate Precision	Delay*	Delay
Specificity	Show resolution of drug substance from likely impurities	Show resolution of drug substance from likely impurities
Quantification Limit	N/A**	Confirmed to be no greater than the reporting limit
Detection Limit	N/A	Delay
Linearity	From impurity linearity or 3 levels between 80 - 120% of test concentration	Determine from 3 sample concentrations (e.g. at test concentration, 1% of test concentration, and quantification limit)
Range	Defined by the linearity work	Defined by the linearity work and quantification limit work
Robustness	Solution stability and information gathering	Solution stability and information gathering

* Delay = experiments can be delayed until a later stage of development

** N/A = Not applicable according to ICH 2A/B

5.2 Validation strategy

To find out the critical aspects with an influence on certain validation parameters, Ishikawa-diagrams (cause-effect-diagrams, fishbone-diagrams) were created. This is a routine procedure in management analysis since 1960 but in analytical chemistry and HPLC-analysis, it is a relative new tool [151].

Applied to the glycoanalytical HPLC-methods including all sample preparation procedures described above, the following Ishikawa-diagrams resulted.

For GlycoSepC-analysis (Figure 33):

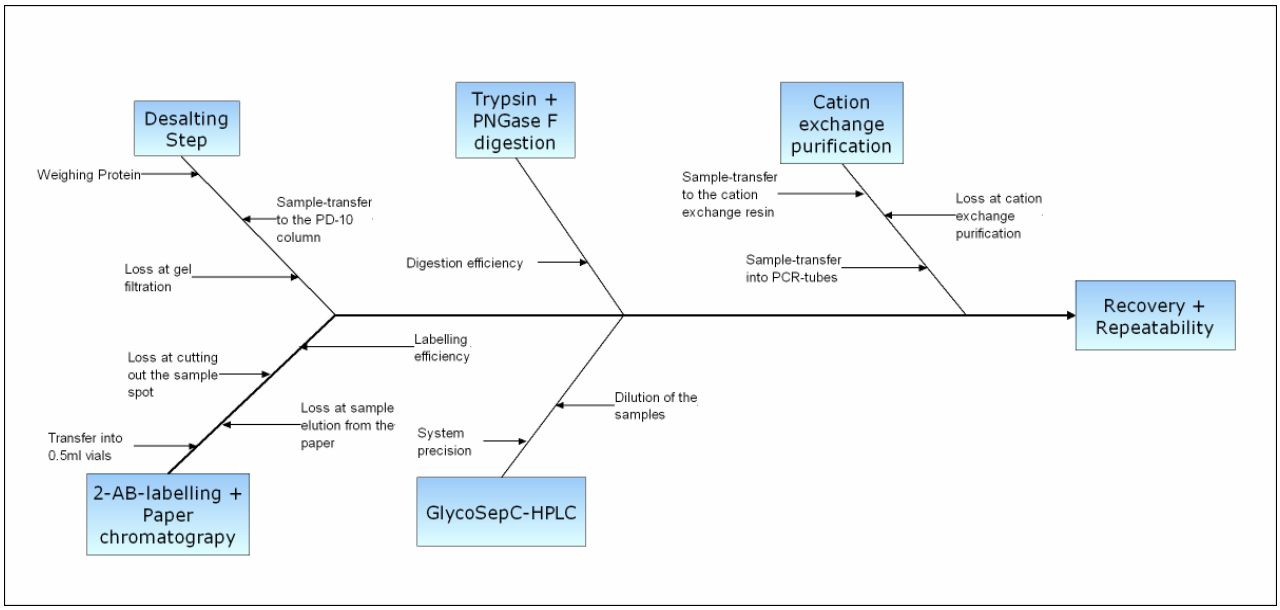


Figure 33: Ishikawa-diagram for GlycoSepC-analysis

And for Aminophase-analysis (Figure 34):

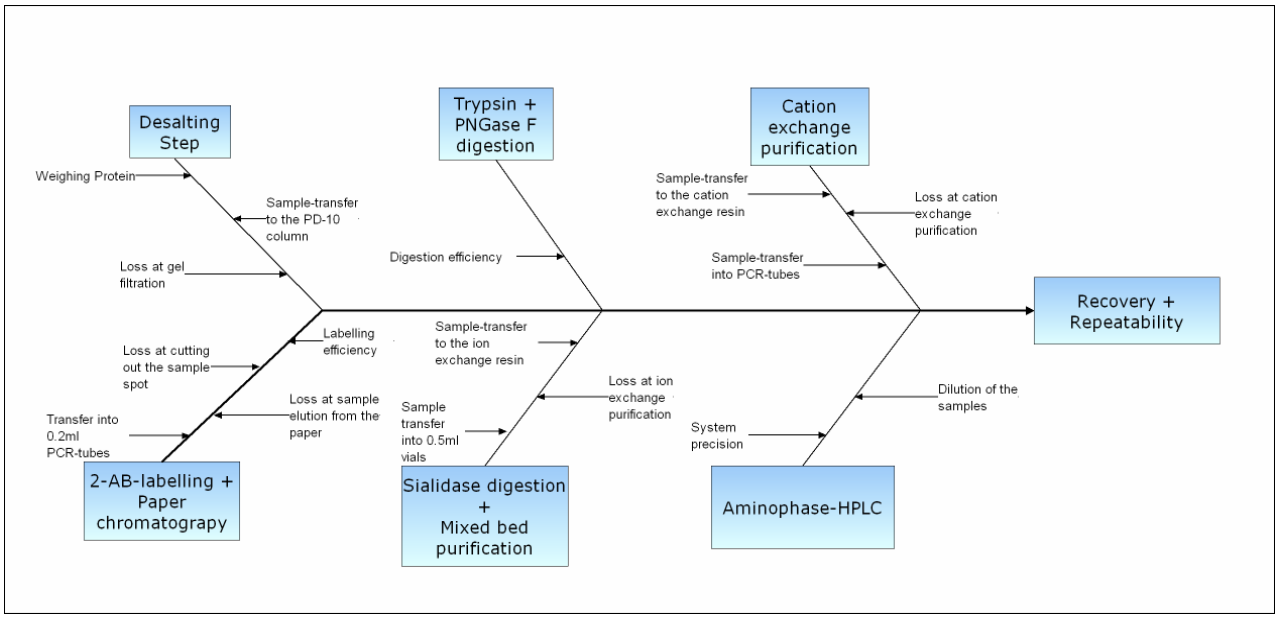


Figure 34: Ishikawa-diagram for Aminophase-analysis

To examine the critical effect parameters recovery and repeatability described in the Ishikawa-diagram, two proposals for the examination were possible. The first possibility (procedure one) was to determine each root parameter alone and to sum up all individual standard deviations to get the total standard deviation of the method. After several experiments, it became clear that this procedure was not universally practicable. For example, it was not possible to determine the 2-AB-labelling efficiency, although MALDI-TOF-MS- and H-NMR-examinations were performed. Labelled and unlabelled glycans should be quantified and the quotient of both should be created. But both methods failed because of a deficient sensitivity regarding unlabelled glycans. The most interfering problem was the huge amount of 2-AB-label-excess in the labelling mixture. The other possibility (procedure two) was to treat the GlycoSepC- and the Aminophase-method as "black-box-methods" and to examine only the amount of protein which was given into the black-boxes and the amount of oligosaccharides (HPLC-chromatograms) which resulted from each method. Finally, although some parameters from procedure one could be determined, it turned out that it was only possible to determine the overall repeatability and recovery by procedure two.

5.3 Results of validation

5.3.1 Validation raw data

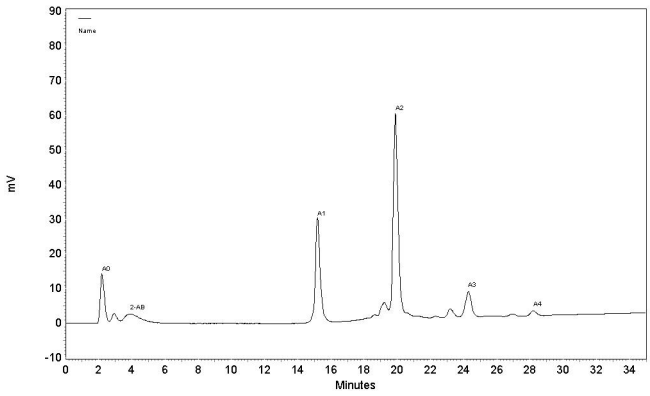
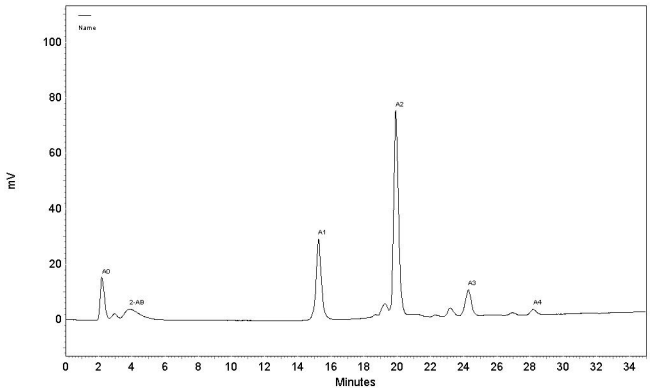
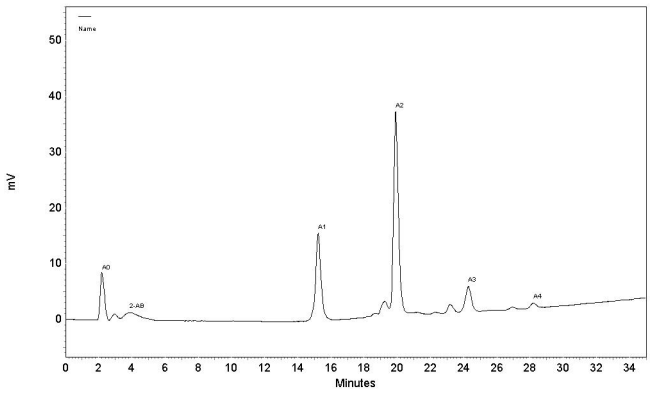
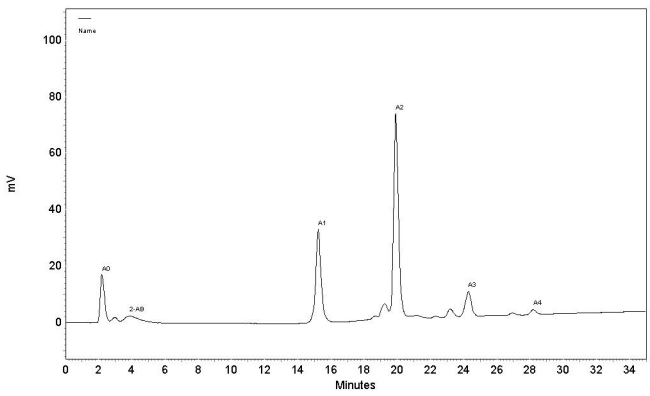
Linearity, repeatability, specificity and accuracy of GlycoSepC- as well as Aminophase-analysis were determined using a specific CHO-GP-sample. Five concentration levels of the protein, each with a four time repetition, were analyzed by the whole analytical procedure and absolute peak area determinations of the individual chromatograms were used for the calculation of the individual linearity curves (Table 20).

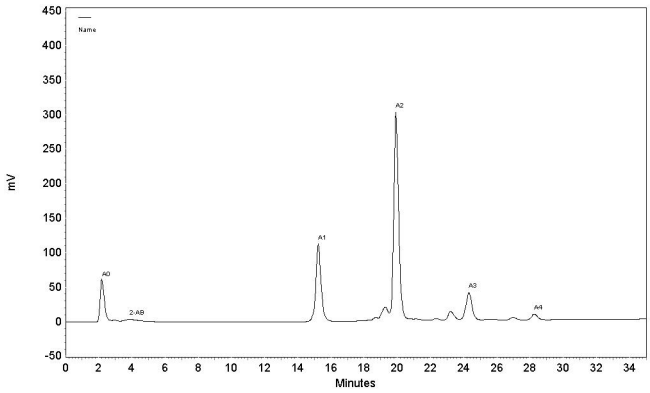
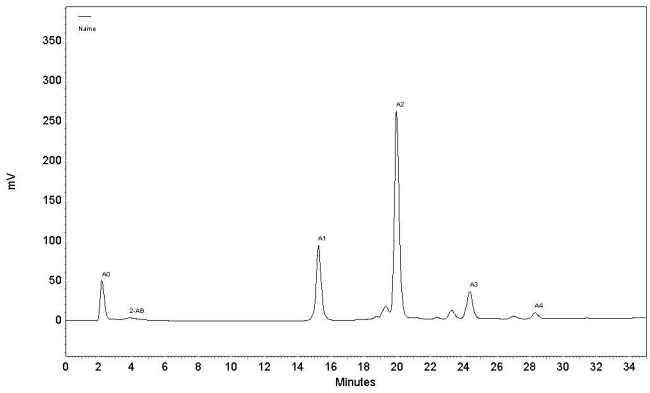
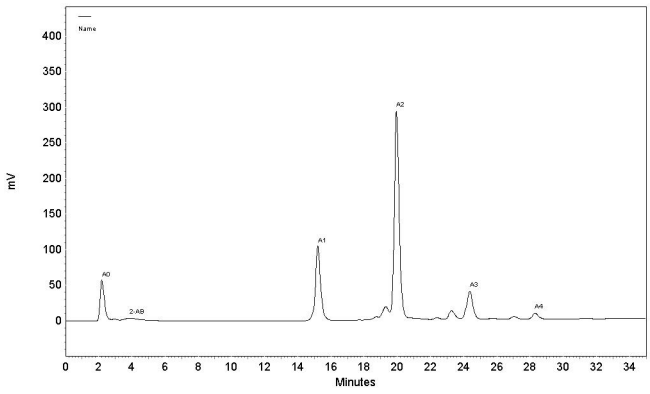
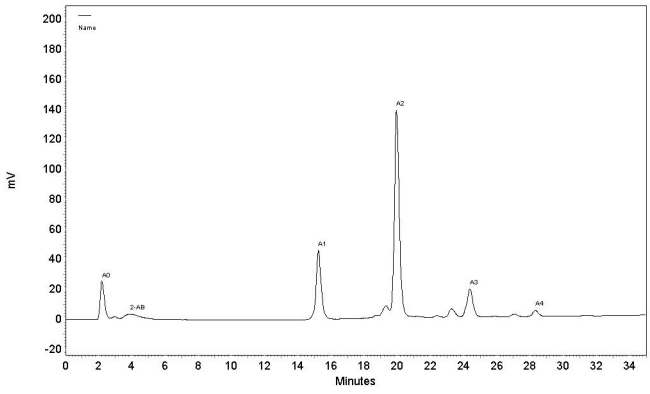
Table 20: Concentrations of specific CHO-GP-sample and number of repetitions for linearity determination

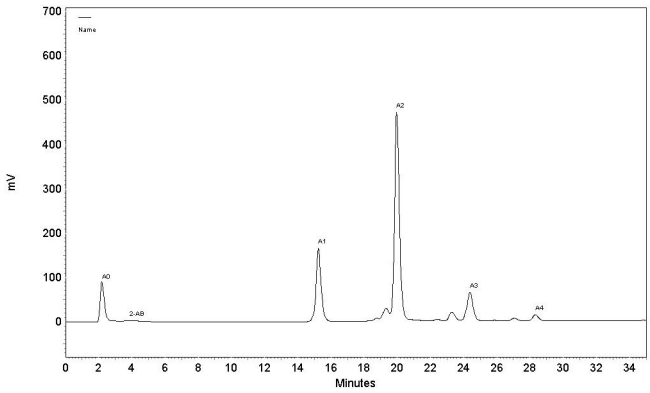
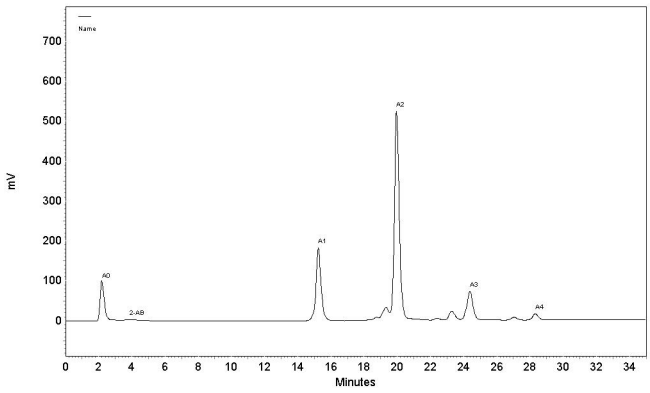
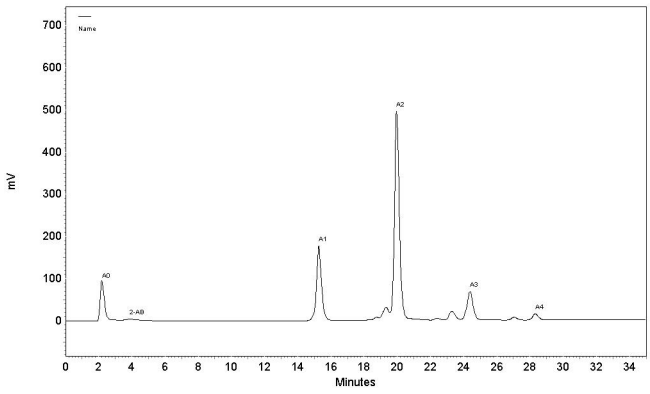
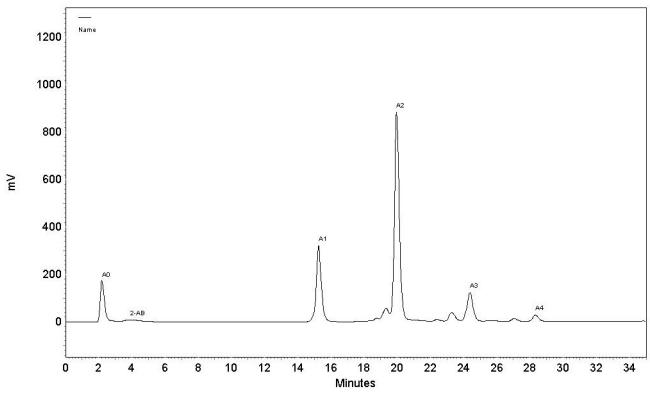
Concentration of CHO-GP-sample [nmol]:	**Repetition:**
3.5	4
14	4
24.5	4
35	4
45.5	4

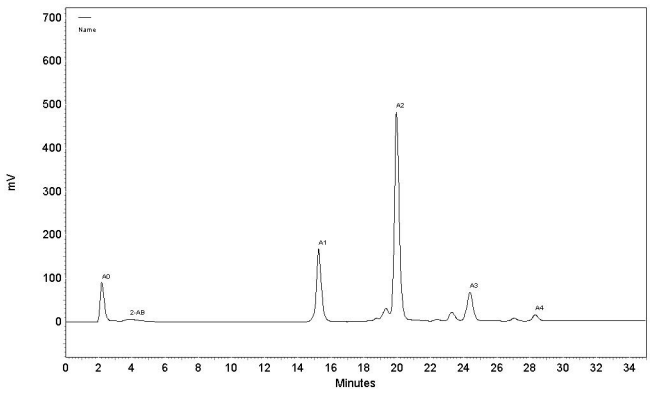
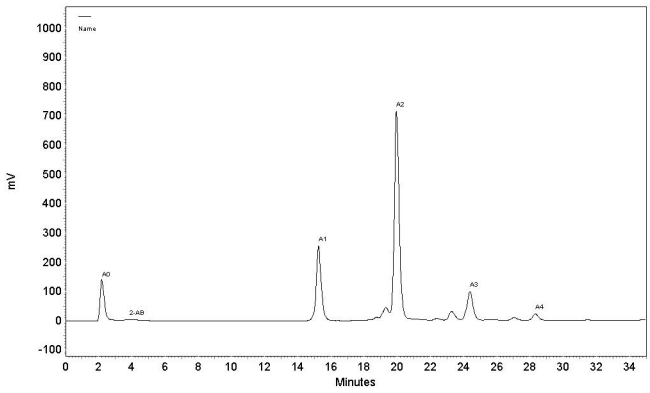
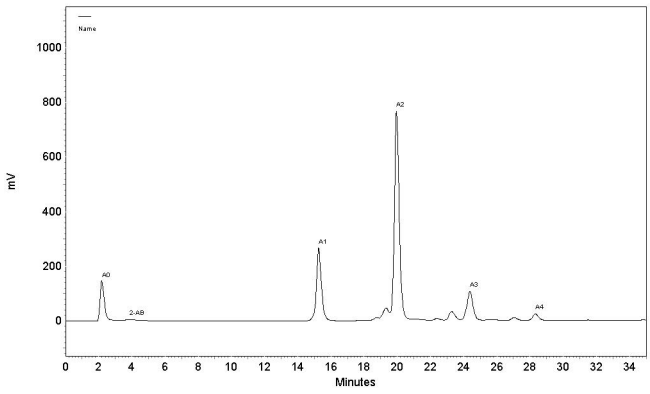
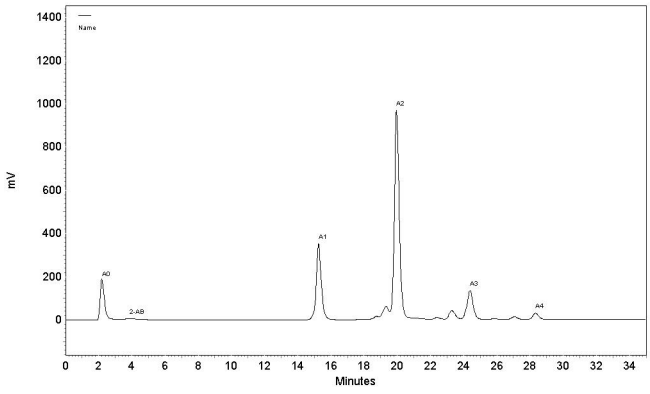
The resulting chromatograms are shown below (GlycoSepC-analyses: Table 21, Aminophase-analyses: Table 22).

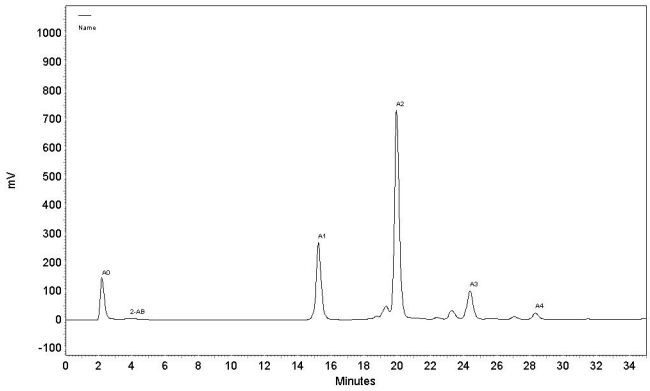
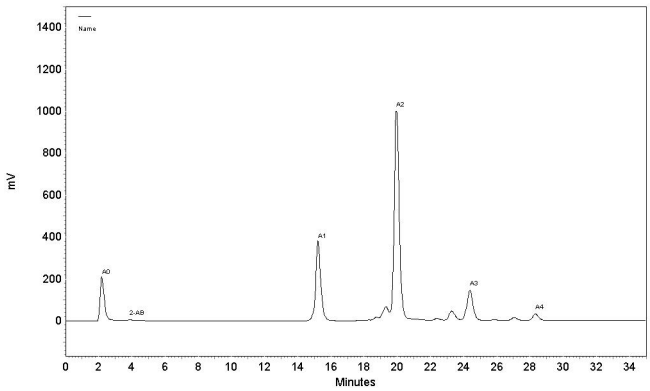
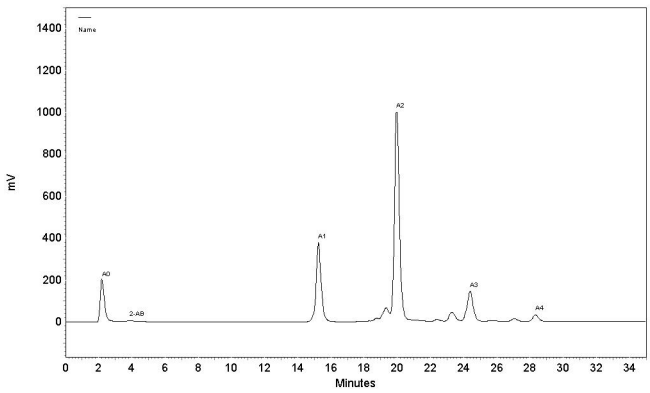
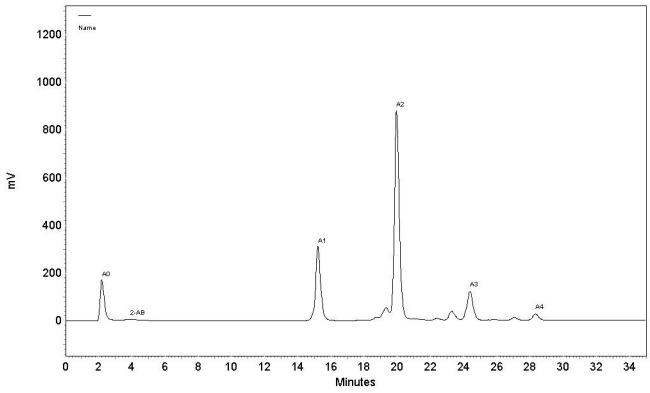
Table 21: GlycoSepC-chromatograms for linearity-determination

	<u>CHO-GP-sample A1a:</u> 3.5 nmol
	<u>CHO-GP-sample A1b:</u> 3.5 nmol
	<u>CHO-GP-sample A1c:</u> 3.5 nmol
	<u>CHO-GP-sample A1d:</u> 3.5 nmol

	<u>CHO-GP-sample A2a:</u> 14.0 nmol
	<u>CHO-GP-sample A2b:</u> 14.0 nmol
	<u>CHO-GP-sample A2c:</u> 14.0 nmol
	<u>CHO-GP-sample A2d:</u> 14.0 nmol

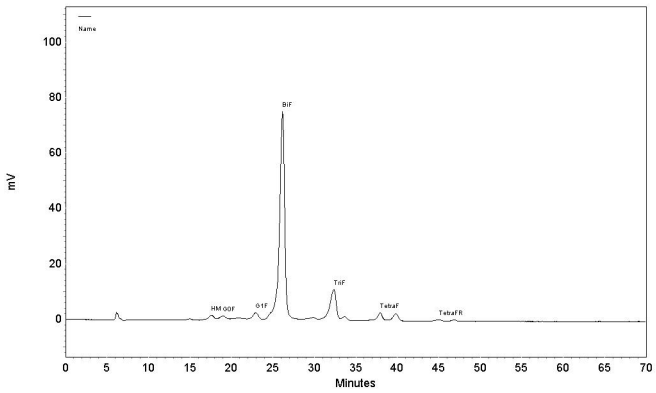
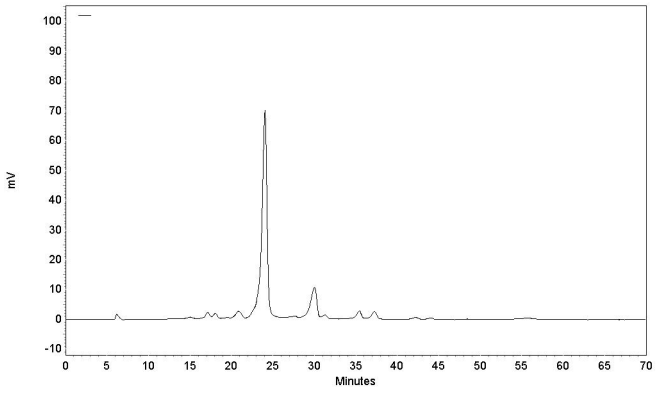
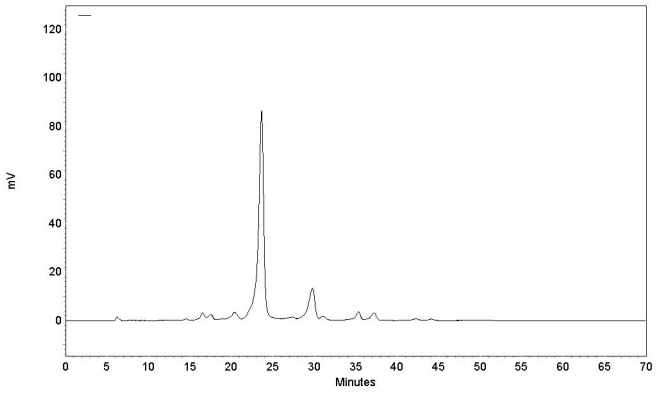
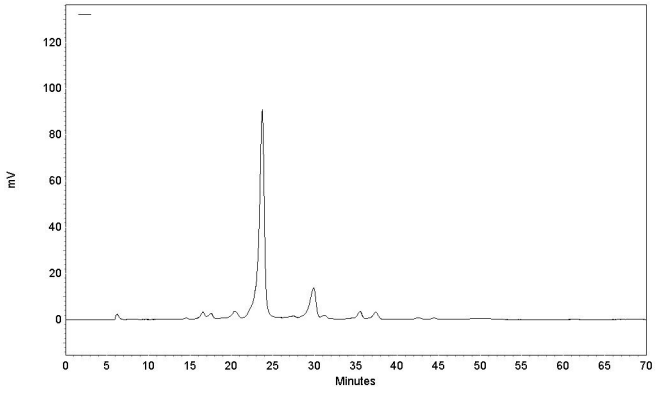
	<u>CHO-GP-sample A3a:</u> 24.5 nmol
	<u>CHO-GP-sample A3b:</u> 24.5 nmol
	<u>CHO-GP-sample A3c:</u> 24.5 nmol
	<u>CHO-GP-sample A3d:</u> 24.5 nmol

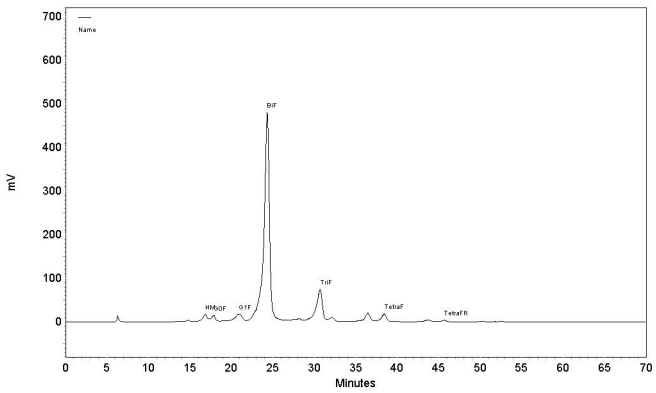
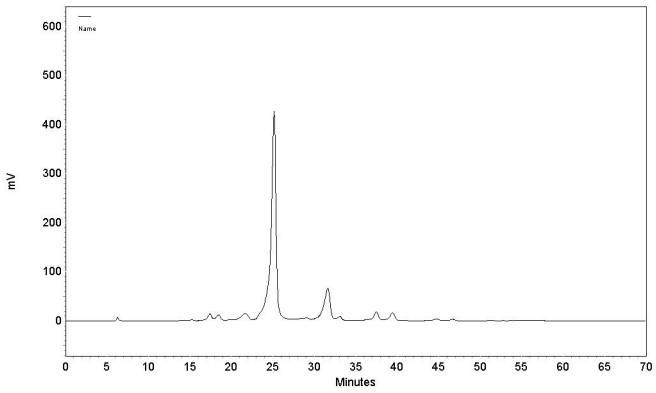
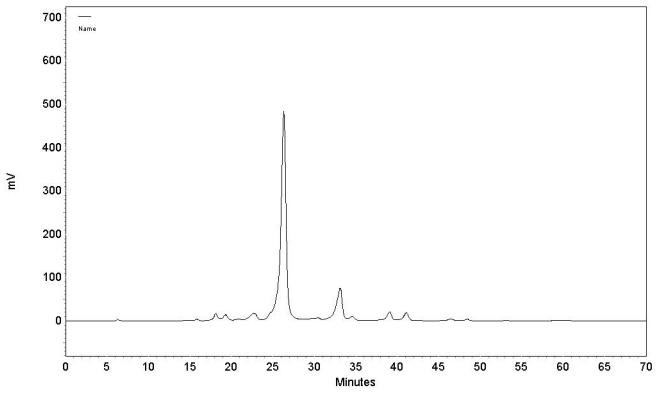
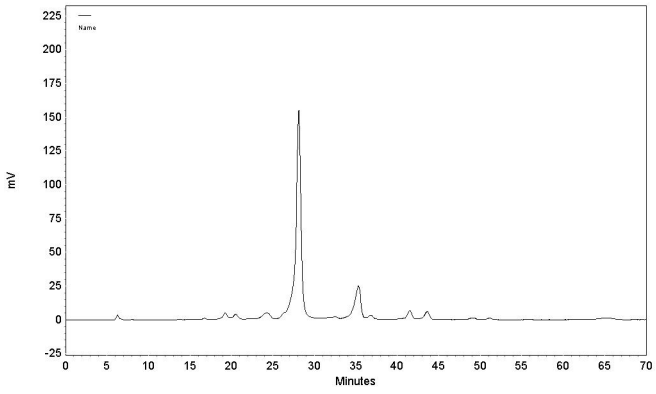
	<u>CHO-GP-sample A4a:</u> 35.0 nmol
	<u>CHO-GP-sample A4b:</u> 35.0 nmol
	<u>CHO-GP-sample A4c:</u> 35.0 nmol
	<u>CHO-GP-sample A4d:</u> 35.0 nmol

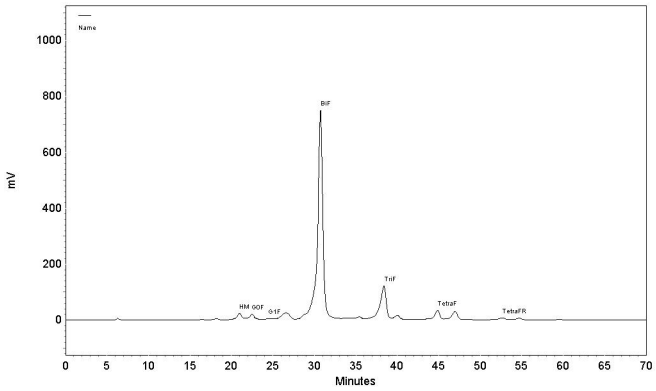
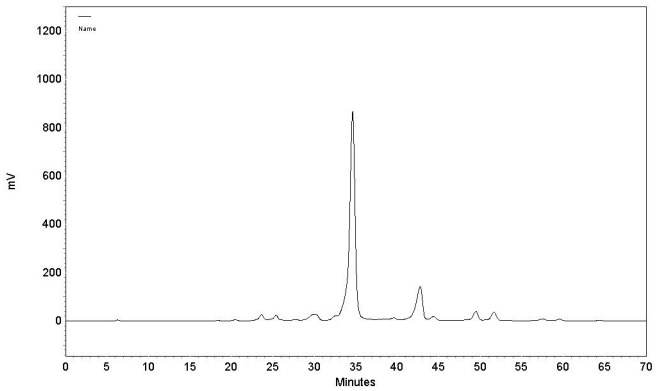
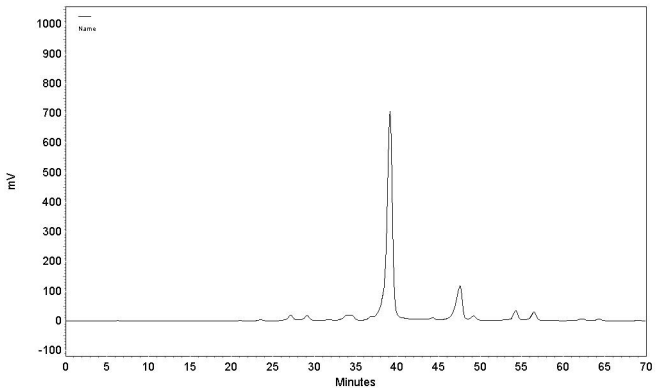
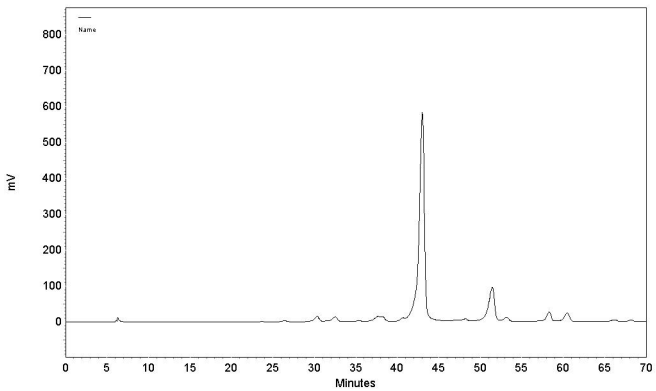
	<u>CHO-GP-sample A5a:</u> 45.5 nmol
	<u>CHO-GP-sample A5b:</u> 45.5 nmol[17]
	<u>CHO-GP-sample A5c:</u> 45.5 nmol
	<u>CHO-GP-sample A5d:</u> 45.5 nmol

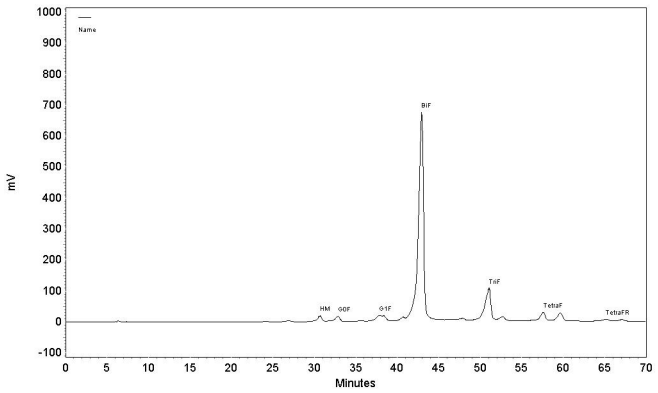
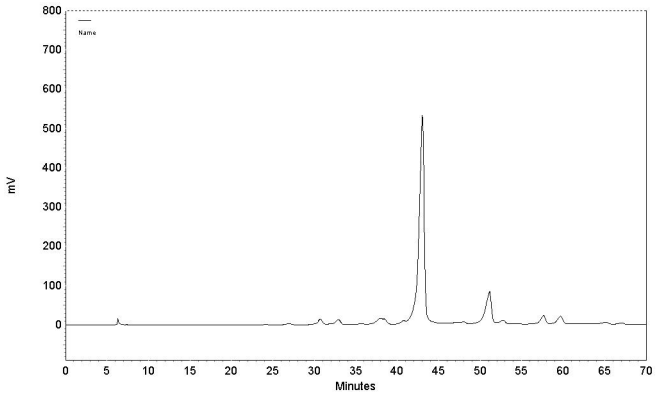
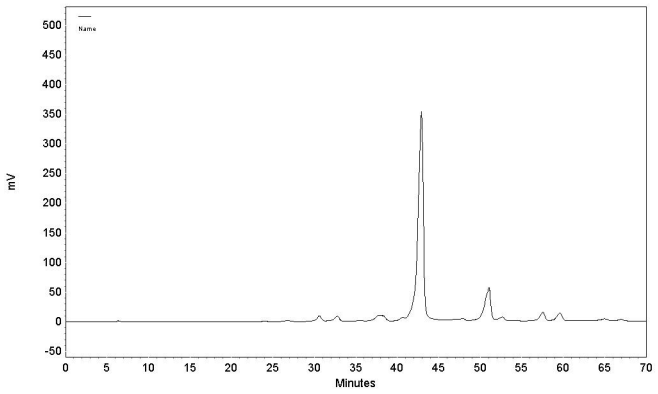
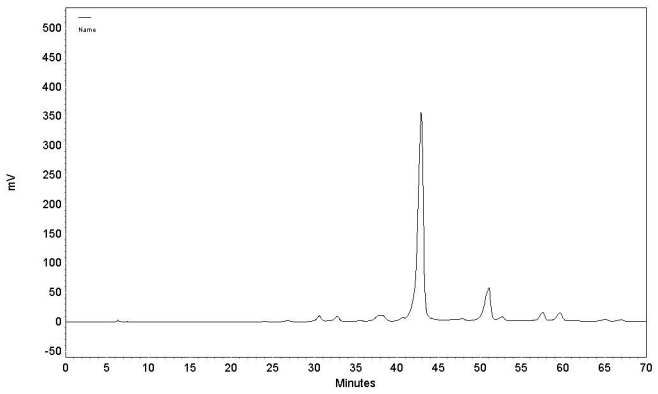
[17] *Because of incompleteness of A2, quantification was performed by injection of 10 µl instead of 20 µl and multiplication of the peak areas with a factor of 2*

Table 22: Aminophase-chromatograms for linearity-determination

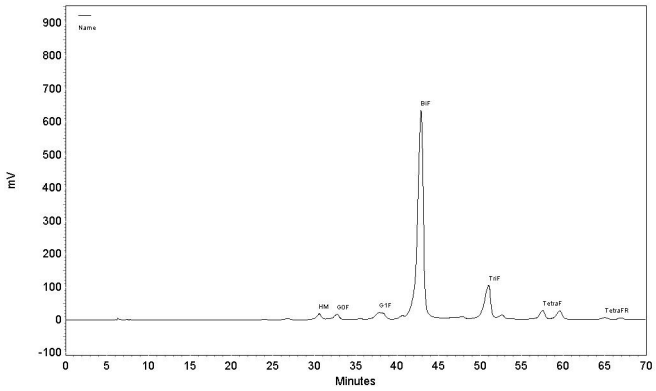
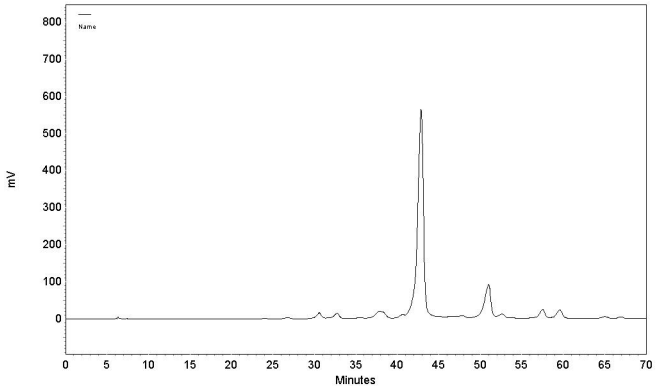
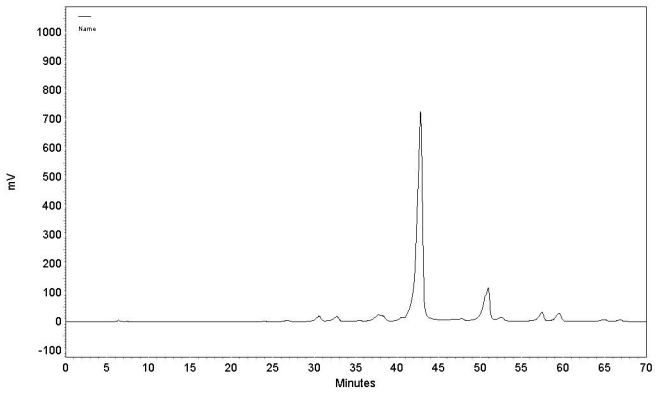
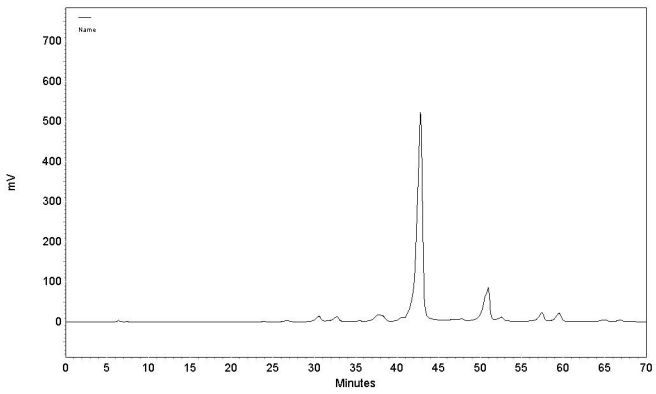
	<u>CHO-GP-sample B1a:</u> 3.5 nmol
	<u>CHO-GP-sample B1b:</u> 3.5 nmol
	<u>CHO-GP-sample B1c:</u> 3.5 nmol
	<u>CHO-GP-sample B1d:</u> 3.5 nmol

	<u>CHO-GP-sample B2a:</u> 14.0 nmol
	<u>CHO-GP-sample B2b:</u> 14.0 nmol
	<u>CHO-GP-sample B2c:</u> 14.0 nmol
	<u>CHO-GP-sample B2d:</u> 14.0 nmol

	<u>CHO-GP-sample B3a:</u> 24.5 nmol
	<u>CHO-GP-sample B3b:</u> 24.5 nmol
	<u>CHO-GP-sample B3c:</u> 24.5 nmol
	<u>CHO-GP-sample B3d:</u> 24.5 nmol

	CHO-GP-sample B4a: 35.0 nmol[18]
	CHO-GP-sample B4b: 35.0 nmol
	CHO-GP-sample B4c: 35.0 nmol
	CHO-GP-sample B4d: 35.0 nmol

[18] *Because of incompleteness of BiF and strong peak drift quantification was performed by injection of 10 µl instead of 20 µl and multiplication of the peak areas with a factor of 2*

	<u>CHO-GP-sample B5a:</u> 45.5 nmol
	<u>CHO-GP-sample B5b:</u> 45.5 nmol
	<u>CHO-GP-sample B5c:</u> 45.5 nmol
	<u>CHO-GP-sample B5d:</u> 45.5 nmol

As it becomes obvious in the chromatograms, the two methods differed in the reproducibility of retention times. In the Aminophase-chromatograms, a strong drift to higher retention times at higher sample concentrations was visible. This was due to the different separation mechanisms. GlycoSepC-HPLC was an anion exchange chromatography (AEX), whose major separation principle was based on adsorption on the stationary phase of the chromatography column. In contrast to that, the Aminophase-HPLC was a hydrophilic interaction chromatography (HILIC) which was based on distribution of the analytes between two fluid phases, one located at the stationary phase and the other located in the eluent. The distribution of the analytes between the two fluid phases was characterized by a distribution coefficient and the adjustment of this balance was a function of time. This meant that the more analyte was injected, the more analyte had to be distributed between the two phases and the more time was needed for this process. This was the reason for the drift to higher retention times at higher sample concentrations in the Aminophase-HPLC.

5.3.2 Determination of accuracy and repeatability

The linearity curves from the resulting peak area determinations are shown below for each charged glycan fraction (GlycoSepC-HPLC - Table 23) and for each uncharged structure (Aminophase-HPLC - Table 24). Accuracy can be inferred from the fact that the prolonged linearity curve at sample amounts of 0 nmol does not have a significant difference from 0 in the y-axis-direction [152]. Therefore linear regression was performed and the estimated 95% confidence interval was laid over x = 0. If the confidence interval enclosed y = 0, then the method would be accurate by a probability of 95%, implied that precision, linearity and specificity were taken for granted. For GlycoSepC-analysis, accuracy could be inferred with this procedure, whereas for Aminophase-analysis the confidence interval did not overlap x=0. This was due to the first data point (concentration x=3.5 nmol) which did not only narrow the confidence interval extremely, but which seemed to show specific loss in glycan structures. As a consequence of this problem, the analytical range for Aminophase-analysis was shortened so that the first data point was

outside this new linear range. Within the new range (14 nmol - 45.5 nmol) accuracy could be inferred (Table 25).

Repeatability was determined for each structure fraction (GlycoSepC) and each structure (Aminophase) at all concentration levels and was expressed as relative standard deviation (RSD).

Before linear regression was performed, an F-test was used to examine if the standard deviations of the different concentration levels were comparable. Therefore the highest and the lowest standard deviations determined at all concentration levels were used to calculate the F-test statistic.

$$F = \left(\frac{sdv_{high}}{sdv_{low}} \right)^2$$

To test only representative data for both methods instead of the individual fractions (GlycoSepC) and the individual structures (Aminophase), only total peak area linearity curves were used.

GlycoSepC:

$$F = \left(\frac{sdv_{high}}{sdv_{low}} \right)^2 = \left(\frac{8334867}{869527} \right)^2 = \frac{7E13}{7.6E11} = 92$$

Aminophase:

$$F = \left(\frac{sdv_{high}}{sdv_{low}} \right)^2 = \left(\frac{18460729}{742508} \right)^2 = \frac{3.4E14}{5.5E11} = 618$$

F* (f1 = n - 1 = 3; f2 = n - 1 = 3; p = 0.05) = 9.277

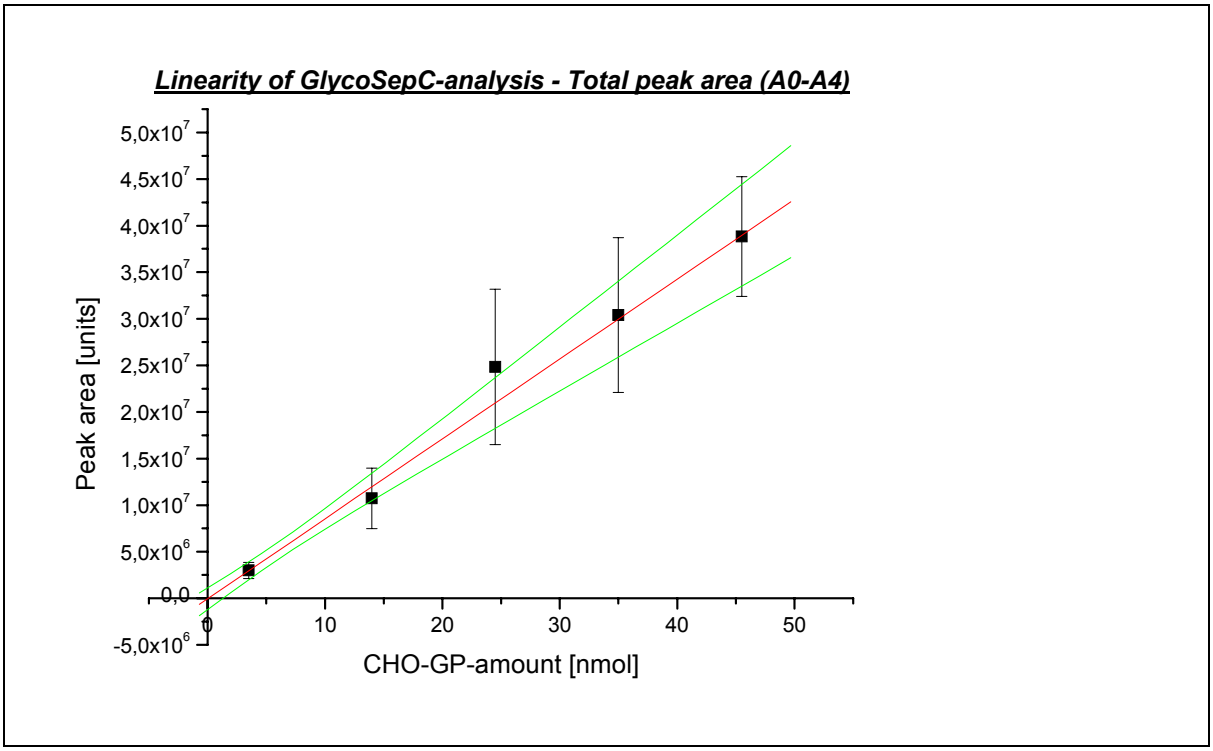
F > F*, standard deviations were not homogenous. As consequence, normal linear regression could not be performed, but linear regression which considered the different standard deviations. That was by using weighted linear regression, where a weighing factor was used which normalized the

different standard deviations. In this case, a weighing factor of 1/sdv(y) was used.

An assumption for the F-test is normally a test of gaussian distribution of the samples. But in this study I deliberately waived this test because analytical samples are usually gaussian distributed. Besides the minimum sample size for test procedures to reach a significance level of 95% is higher than n = 4 which I used (e.g. test by David recommended $n \geq 5$).

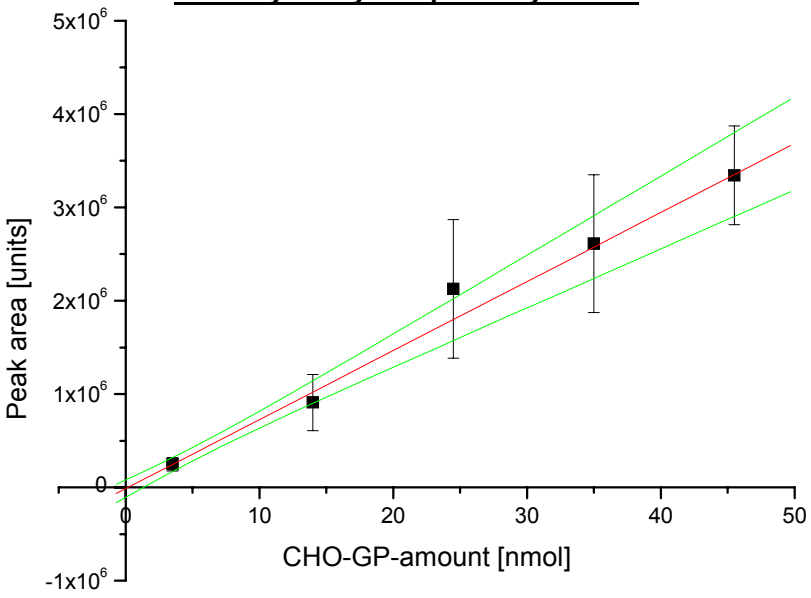
I also did not perform outlier tests because newer tendencies by the FDA and other regulatory administrations are to prohibit any forms of deleting raw data [152]. And by reducing the sample size again the significance of the results would have been even more reduced.

Table 23: Linearity curves of GlycoSepC-fractions of CHO-GP[19]

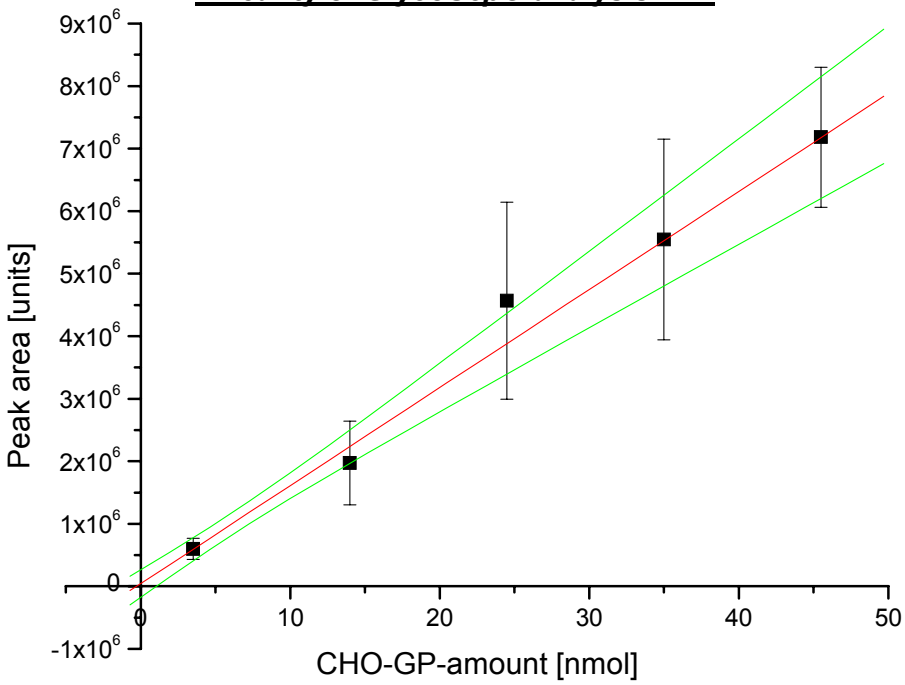


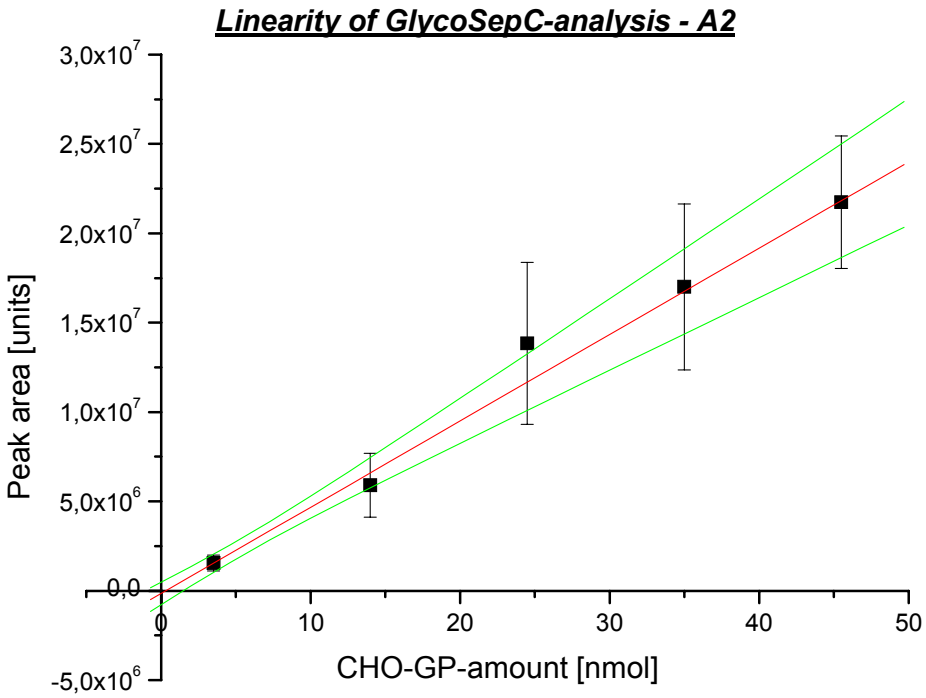
[19] *Error bars = sdv, red line = f(y, sdv_y) = weighted linear regression curve, green line = 95% confidence interval, UCI(y) = upper confidence interval, LCI(y) = lower confidence interval*

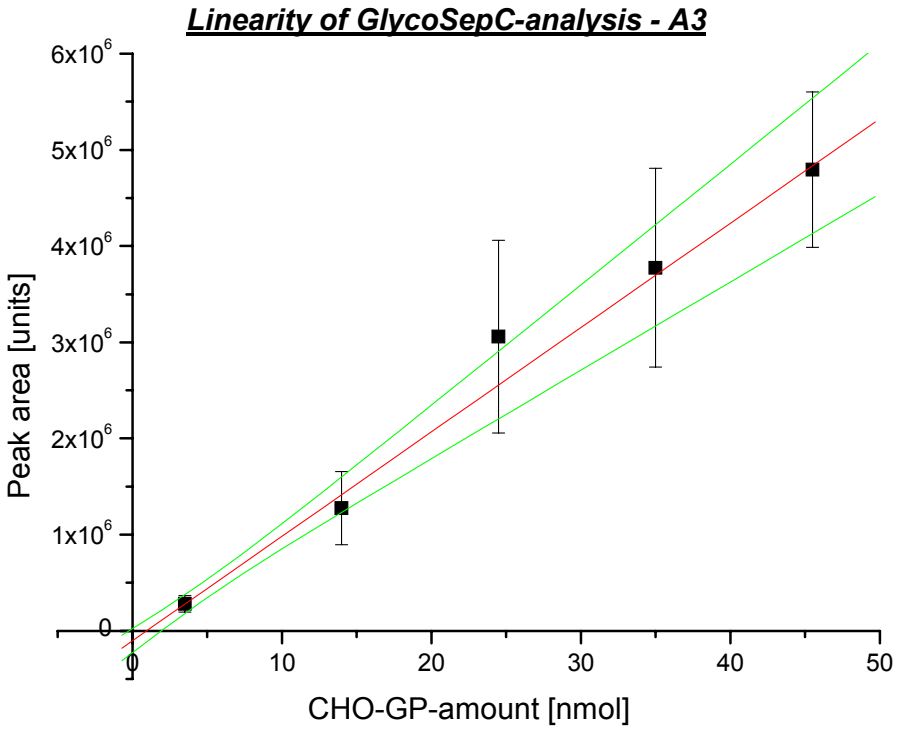
Linearity of GlycoSepC-analysis - A0



Linearity of GlycoSepC-analysis - A1




Linearity of GlycoSepC-analysis - A2
3,0x10^7
2,5x10^7
2,0x10^7
1,5x10^7
1,0x10^7
5,0x10^6
0,0
-5,0x10^6
Peak area [units]
0
10
20
30
40
50
CHO-GP-amount [nmol]


Linearity of GlycoSepC-analysis - A3
6x10^6
5x10^6
4x10^6
3x10^6
2x10^6
1x10^6
0
Peak area [units]
0
10
20
30
40
50
CHO-GP-amount [nmol]

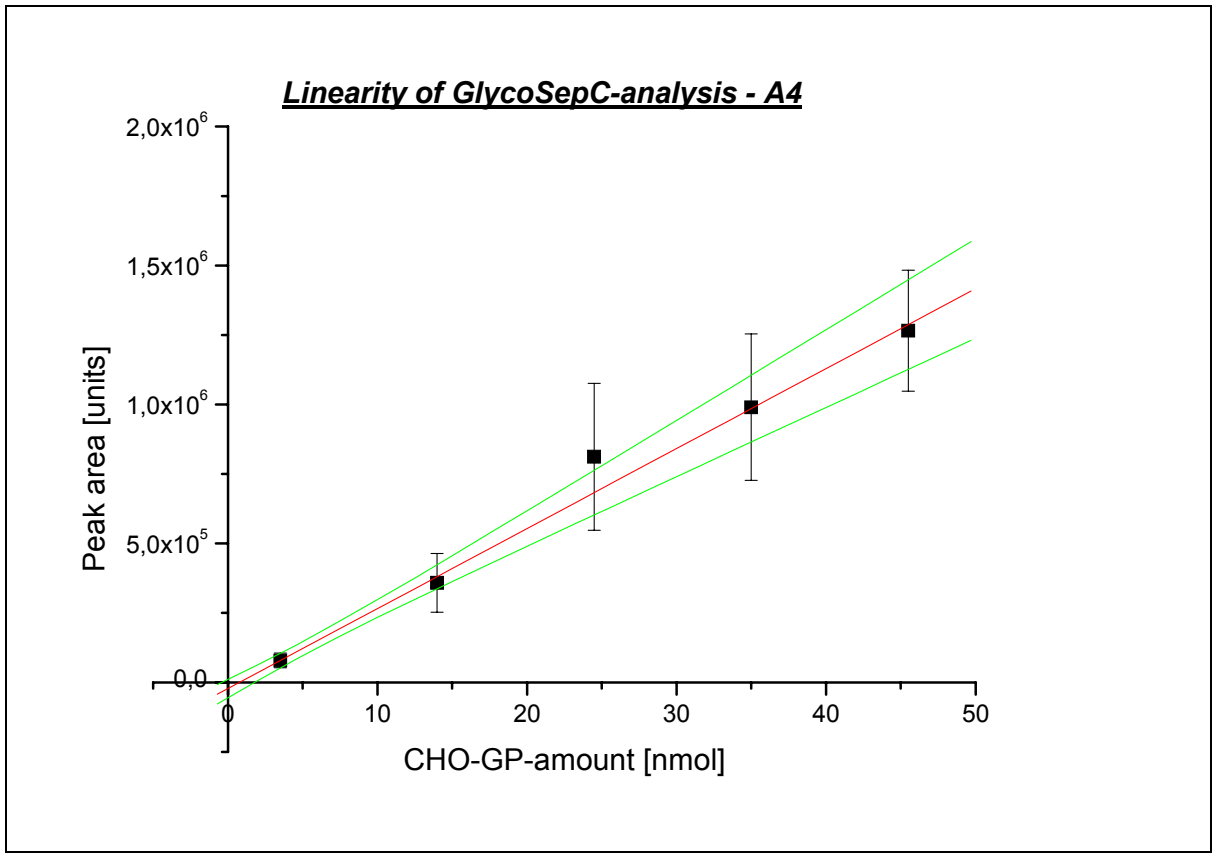
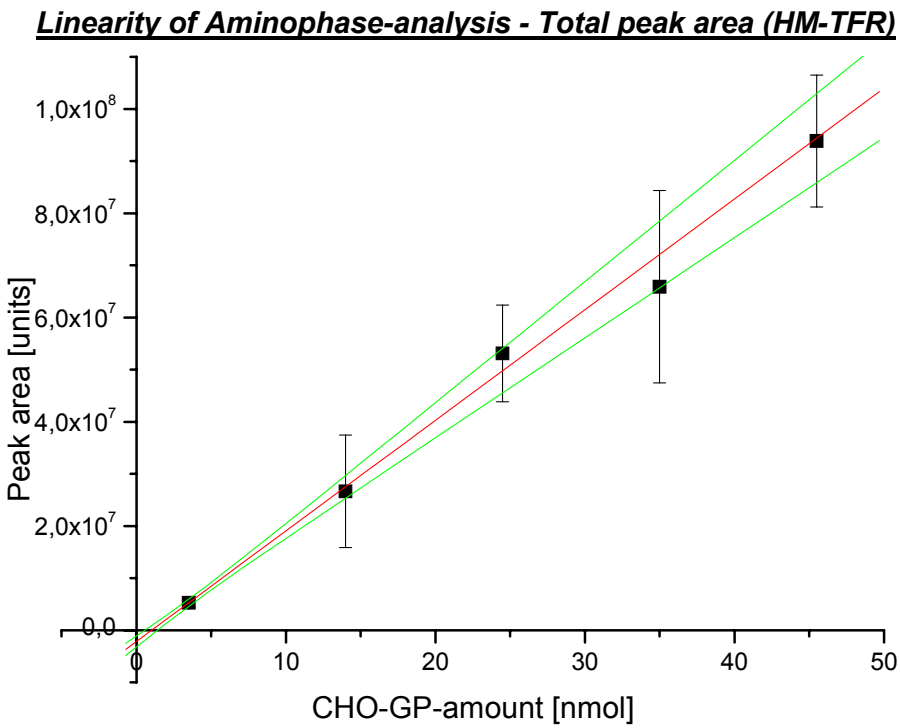

Linearity of GlycoSepC-analysis - A4
2,0x10^6
1,5x10^6
1,0x10^6
5,0x10^5
0,0
Peak area [units]
0
10
20
30
40
50
CHO-GP-amount [nmol]

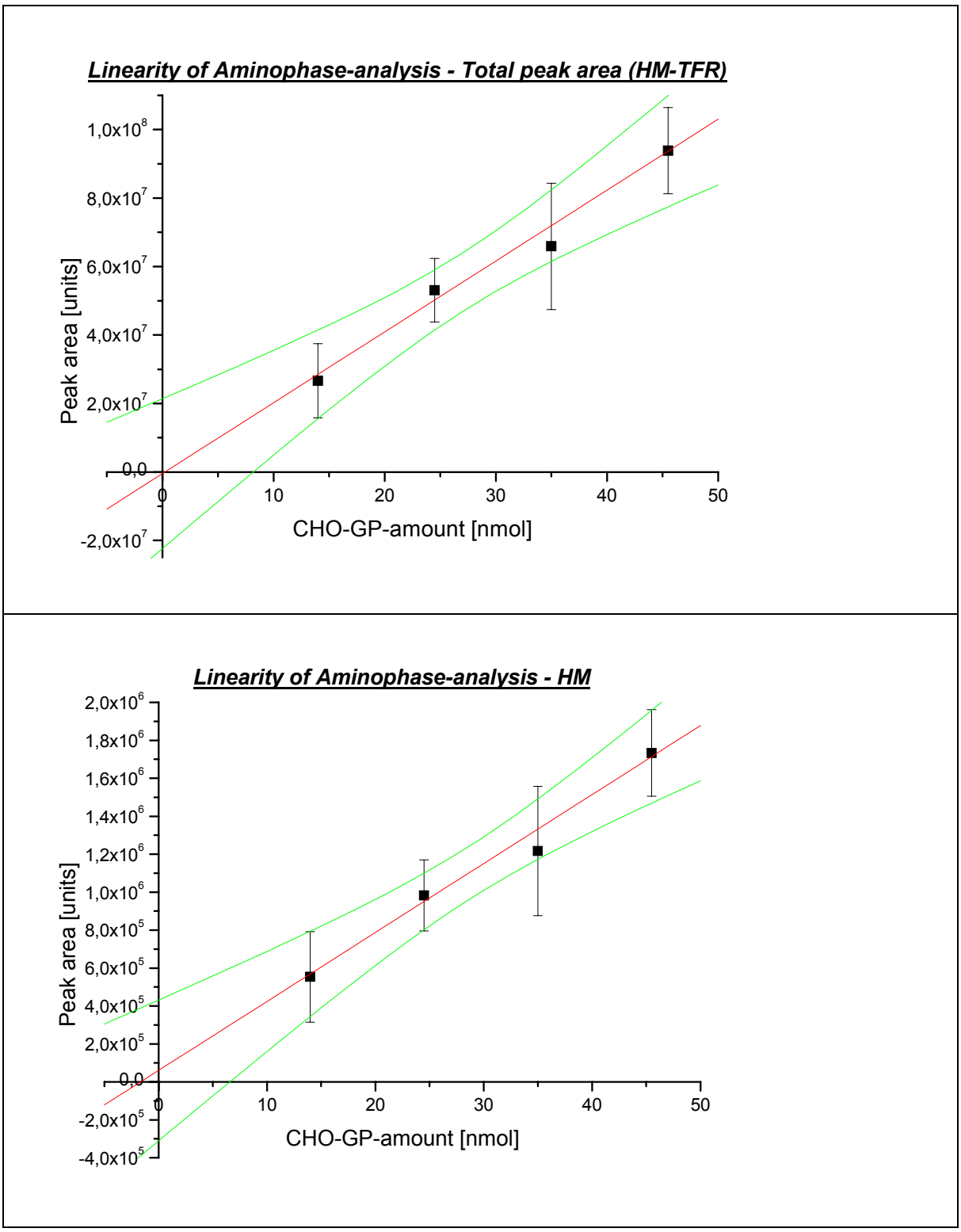
Table 24: Linearity curves of Aminophase-structures of CHO-GP[20]



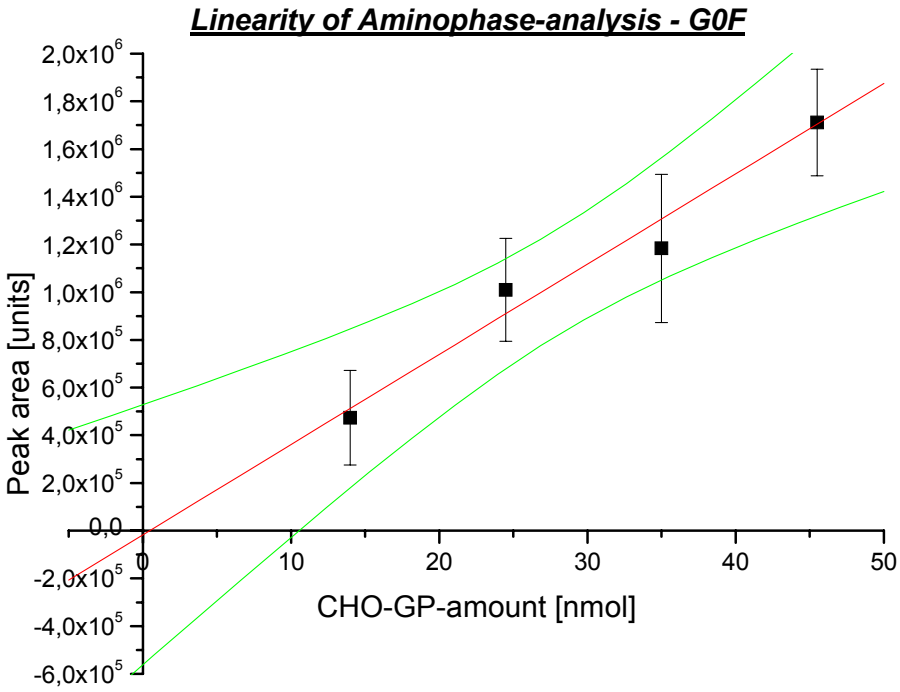
CHO-GP-amount [nmol]	Relative standard deviation (RSD) [%]	$f(y,sdv_y)$-value at x=0	UCI(y)-value at x=0	LCI(y)-value at x=0	Accuracy inferred
3.5	14.2	-2191210	-1131420	-3251000	no
14.0	40.6				
24.5	17.5				
35.0	28.0				
45.5	13.4				

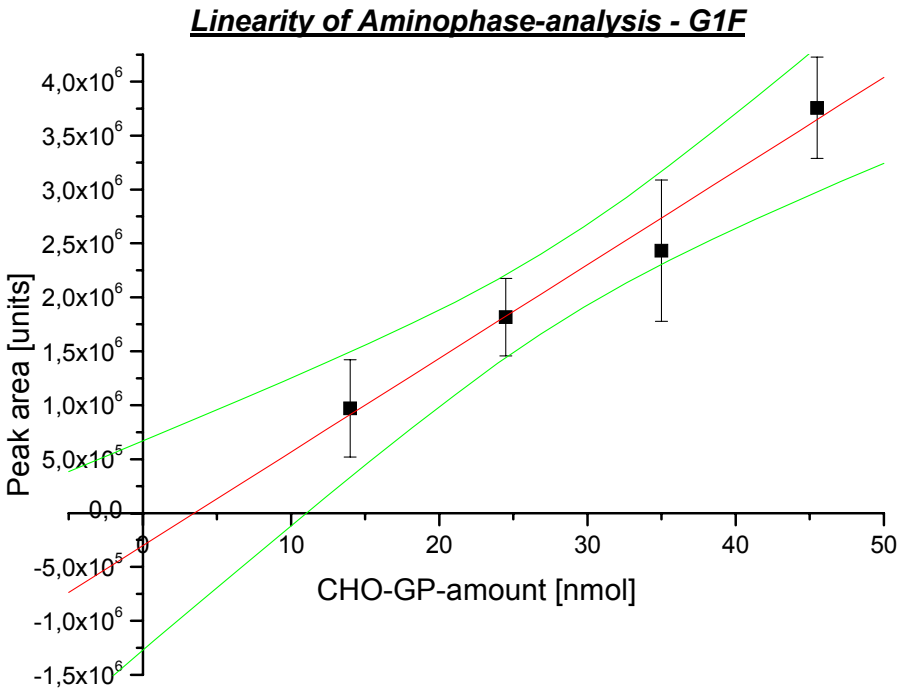
[20] *Error bars = sdv, red line = f(y, sdv_y) = weighted linear regression curve, green line = 95% confidence interval, UCI(y) = upper confidence interval, LCI(y) = lower confidence interval; calibration curve does not cover the zero-y-axis-section because of the first data point at 3.5 nmol; single structure curves (HM - TFR) not shown*

Table 25: Linearity curves of Aminophase-structures without first data point at x = 3.5 nmol[21]

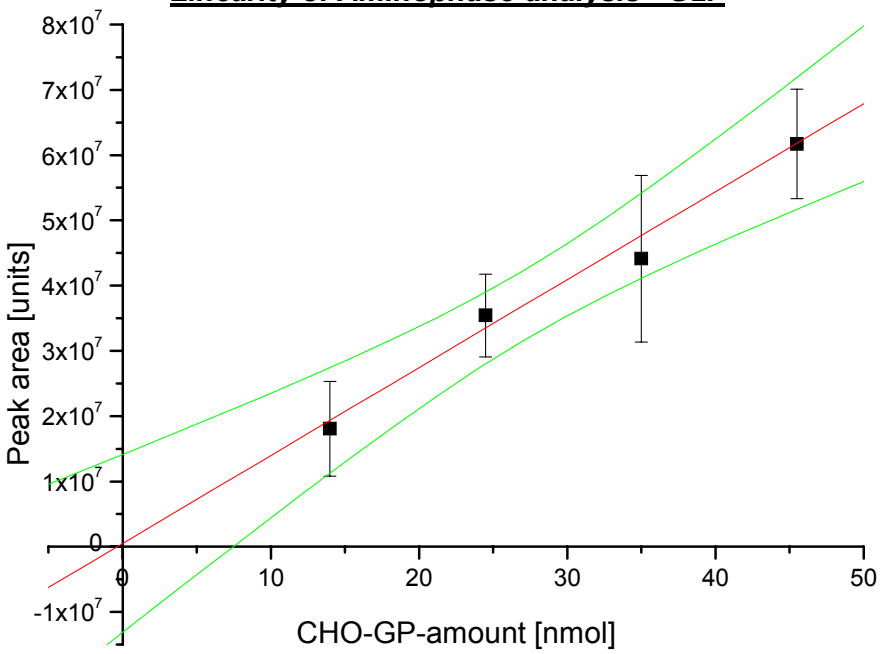


[21] *Error bars = sdv, red line = f(y, sdv$_y$) = weighted linear regression curve, green line = 95% confidence interval, UCI(y) = upper confidence interval, LCI(y) = lower confidence interval*

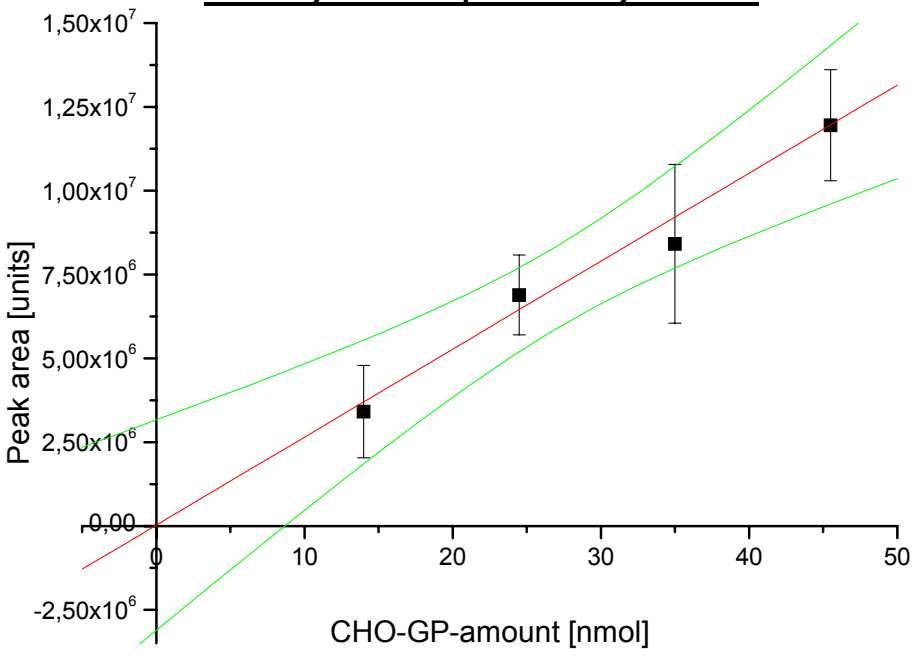

Linearity of Aminophase-analysis - G0F
Peak area [units]
2,0x10^6
1,8x10^6
1,6x10^6
1,4x10^6
1,2x10^6
1,0x10^6
8,0x10^5
6,0x10^5
4,0x10^5
2,0x10^5
0,0
-2,0x10^5
-4,0x10^5
-6,0x10^5
0
10
20
30
40
50
CHO-GP-amount [nmol]

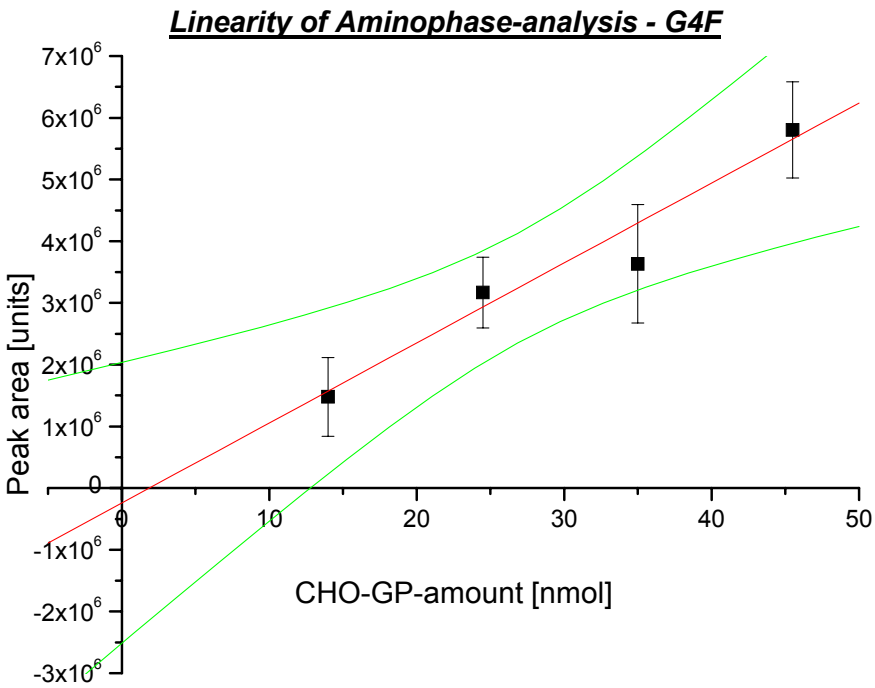

Linearity of Aminophase-analysis - G1F
Peak area [units]
4,0x10^6
3,5x10^6
3,0x10^6
2,5x10^6
2,0x10^6
1,5x10^6
1,0x10^6
5,0x10^5
0,0
-5,0x10^5
-1,0x10^6
-1,5x10^6
0
10
20
30
40
50
CHO-GP-amount [nmol]

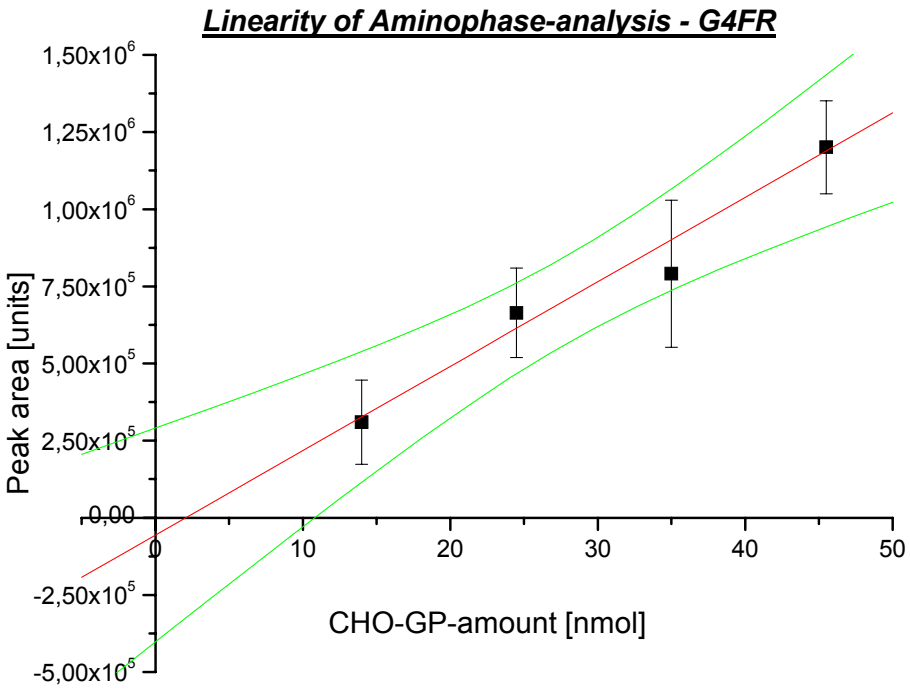
Linearity of Aminophase-analysis - G2F



Linearity of Aminophase-analysis - G3F




Linearity of Aminophase-analysis - G4F
Peak area [units]
7x10^6
6x10^6
5x10^6
4x10^6
3x10^6
2x10^6
1x10^6
0
-1x10^6
-2x10^6
-3x10^6
0
10
20
30
40
50
CHO-GP-amount [nmol]


Linearity of Aminophase-analysis - G4FR
Peak area [units]
1,50x10^6
1,25x10^6
1,00x10^6
7,50x10^5
5,00x10^5
2,50x10^5
0,00
-2,50x10^5
-5,00x10^5
0
10
20
30
40
50
CHO-GP-amount [nmol]

As described in the introduction of the glycoanalytical methods, AEX- and HILIC-analyses were already used for a long time in basic research at universities all over the world to examine differences in glycosylation patterns caused by parameters which could have influenced glycosylation. To overcome variability in the sample preparation procedure, chromatographic peak areas were normalized to 100%. Glycosylation patterns were compared regarding increases and decreases in individual structures and peak areas were used for conclusions that specific parameters had a positive or negative influence on specific glycan structures. What was often not recognized, were the standard deviations of the glycoanalytical methods themselves. These were examined in the following tables for GlycoSepC- and Aminophase-analysis. Therefore the relative standard deviations (variation coefficients) of the above CHO-GP-runs were calculated over the whole working range, for both, the relative quantification approach (Table 26 and Table 27) and also the absolute quantification approach (Table 28 and Table 29).

GlycoSepC-analysis:

Table 26: Relative standard deviations [%] of GlycoSepC-peak-areas over the whole working range with normalization (Relative Quantification)

CHO-GP [nmol]	A0	A1	A2	A3	A4
3.5	3.5	4.5	1.4	3.7	9.0
14.0	6.0	4.4	1.4	2.3	2.7
24.5	1.4	1.0	0.4	0.4	0.8
35.0	1.5	1.7	0.5	0.7	1.1
45.5	1.1	1.6	0.5	0.7	1.1

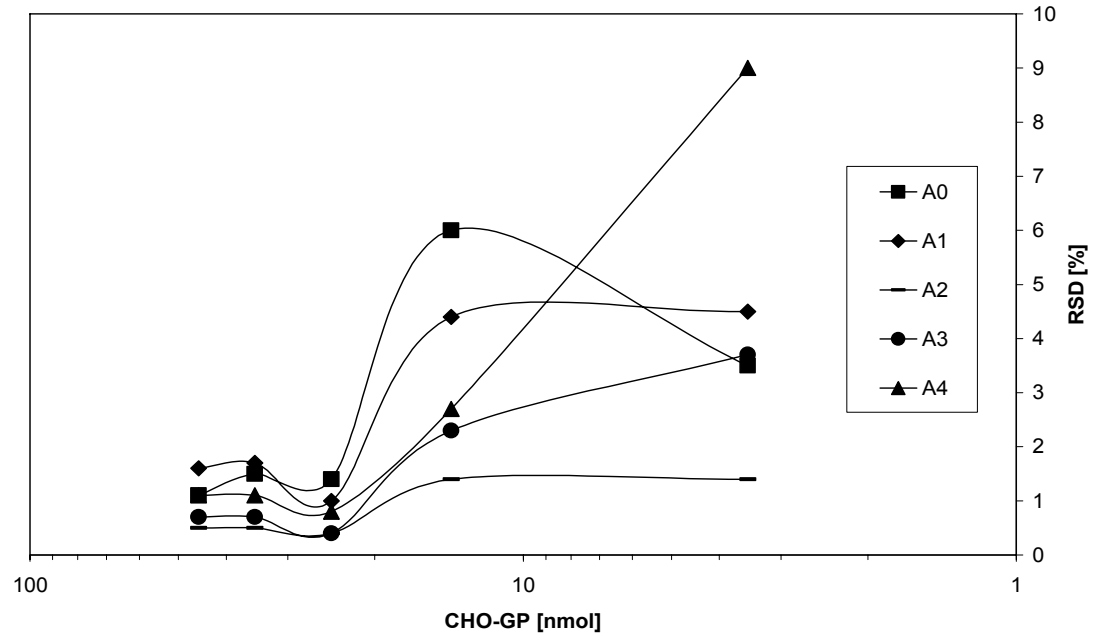
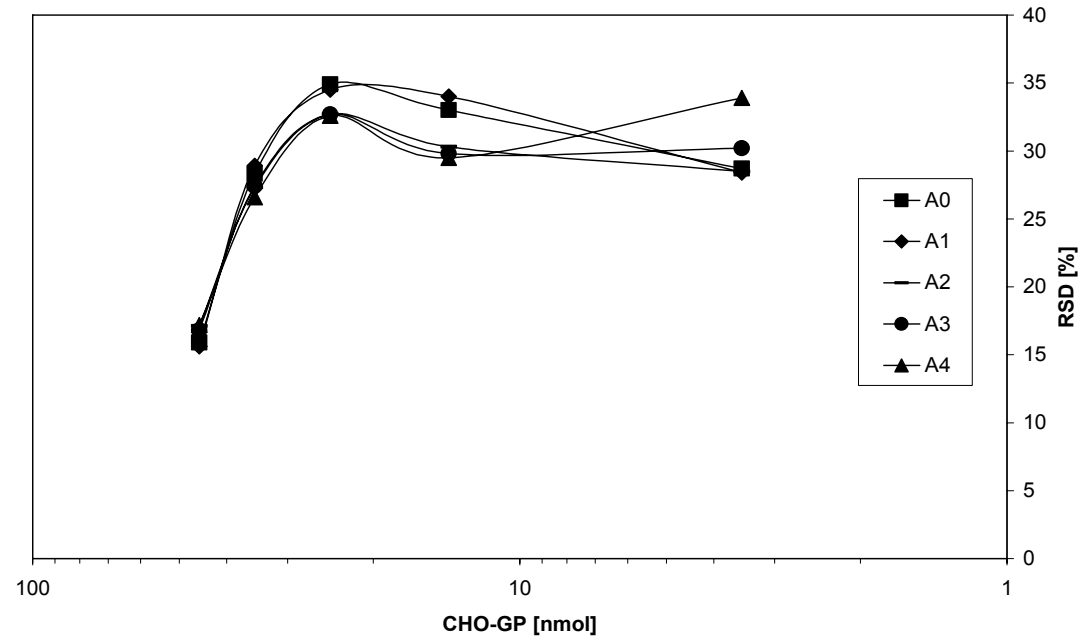


Table 27: Relative standard deviations [%] of GlycoSepC-peak-areas over the whole working range without normalization (Absolute Quantification)

CHO-GP [nmol]	A0	A1	A2	A3	A4
3.5	28.7	28.4	28.5	30.2	33.9
14.0	33.0	34.0	30.3	29.8	29.5
24.5	34.9	34.5	32.7	32.7	32.6
35.0	28.3	28.9	27.3	27.4	26.6
45.5	15.9	15.6	17.1	16.8	17.2



Aminophase-analysis:

Table 28: Relative standard deviations [%] of Aminophase-peak-areas over the whole working range with normalization (Relative Quantification)

CHO-GP [nmol]	HM	G0F	G1F	G2F	G3F	G4F	G4FR
3.5	/	/	/	/	/	/	/
14.0	5.2	3.1	10.6	1.1	0.6	5.3	6.2
24.5	1.6	5.0	2.2	0.3	1.0	1.9	5.5
35.0	1.5	2.3	2.3	1.1	0.4	12.3	10.0
45.5	0.8	1.3	2.8	0.6	0.7	3.4	1.2

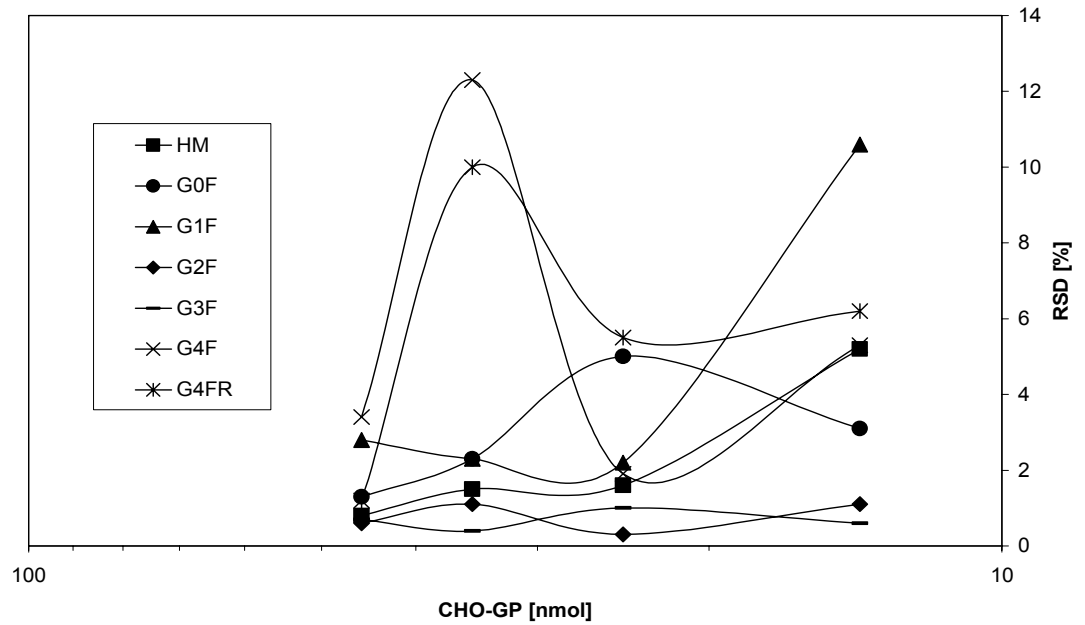
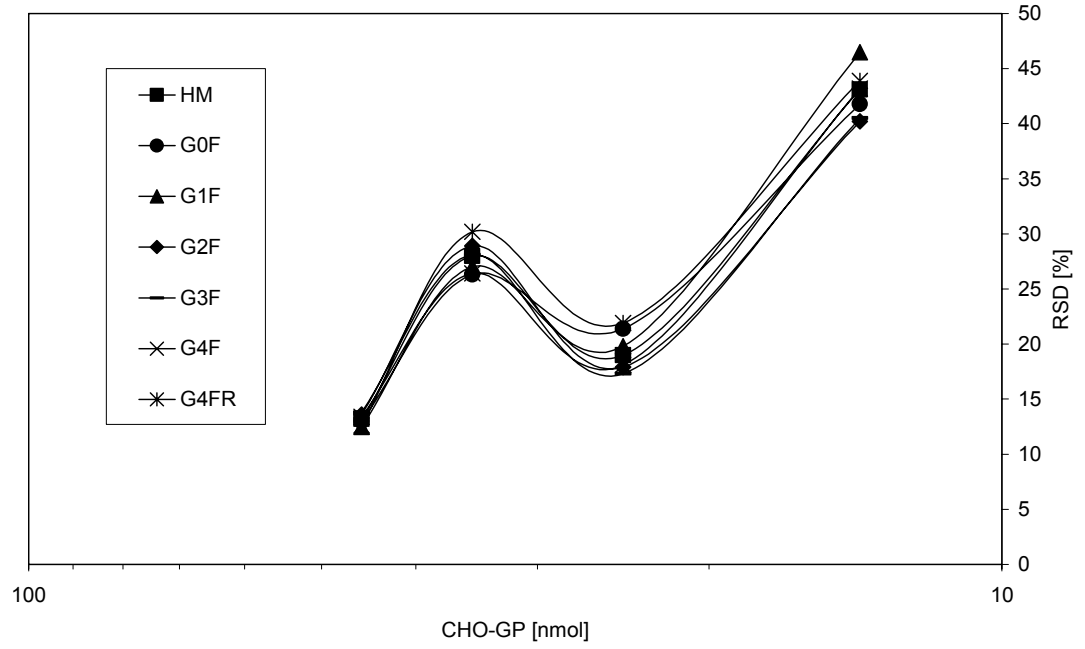


Table 29: Relative standard deviations [%] of Aminophase peak areas over the whole working range without normalization (Absolute Quantification)

CHO-GP [nmol]	HM	G0F	G1F	G2F	G3F	G4F	G4FR
3.5	/	/	/	/	/	/	/
14.0	43.1	41.8	46.5	40.2	40.5	43.2	43.9
24.5	19.0	21.4	19.8	17.9	17.3	18.1	21.9
35.0	28.0	26.3	27.0	28.9	28.1	26.4	30.2
45.5	13.2	13.0	12.5	13.6	13.8	13.4	12.6



The variation coefficients of most analytical methods correspond to the Horwitz-equation [153]

$$CV(\%) = 2^{(1-0.5\log C)},$$

where C is the concentration expressed as powers of 10 (Figure 35).

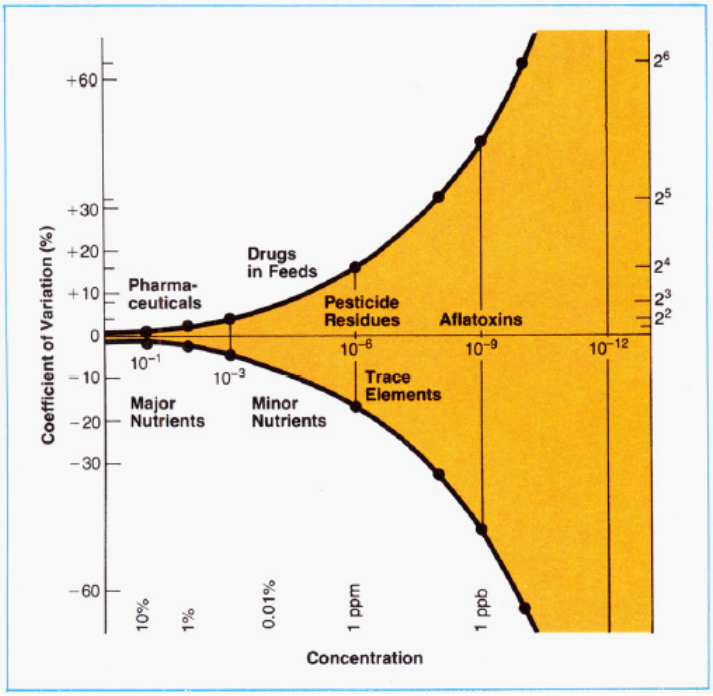


Figure 35: The general curve relating interlaboratory coefficients of variation with concentration along the horizontal center axis[22]

Regarding the RSD-graphs above, no clear trend comparable with the Horwitz-graph (Figure 35) was observed. For relative quantification, a CHO-GP-amount of 24.5 nmol showed the lowest RSD in GlycoSepC- (average: 0.8%) and Aminophase-analysis (average: 2.5%). For absolute quantification, CHO-GP-amounts of 45.5 nmol gave the most precise results. As result from this validation study, the amount of protein used for the application experiments in Chapter Six was adapted to 25 nmol to guarantee the lowest RSD in relative quantification possible. As mentioned before, CHO-Mab and CHO-GP were not appropriate for absolute quantification approaches, so that this RSD-value was only interesting for the absolute quantification strategy itself.

[22] *copied from 153. Horwitz, W., Evaluation of Analytical Methods Used for Regulation of Foods and Drugs. Analytical Chemistry, 1982.* ***54****(1): p. 67 - 76.*

5.3.3 Determination of linearity

For the test of linearity a statistical F-test was performed [154]. Therefore linear and quadratic regression curves were calculated and the residual standard deviations of both regressions were determined. For Aminophase-analysis, data without x = 3.5 nmol was chosen to calculate whether the accurate Aminophase-analysis was also linear.

Linear residual standard deviations

$$S_{y(L)} = \sqrt{\frac{\sum (r_i - r_o)^2}{N-2}}$$

Quadratic residual standard deviations:

$$s_{y(Q)} = \sqrt{\frac{\sum (r_i - r_o)^2}{N-3}}$$

N = number of concentration levels

r_i = individual residuals

r_o = optimal regression value

Test statistic:

$$test\ statistic = \left(\frac{s_{y(L)}}{s_{y(Q)}} \right)^2$$

This test statistic was compared with the F-table (p = 0.05, f1 = N - 2, f2 = N - 3). If the test statistic was higher than the critical value in the F-table, a significant difference between both residual standard deviations would exist and the quadratic regression would be preferred. If the test statistic was lower than the critical value in the F-table, linear regression would be preferred (Table 30 + Table 31).

Table 30: GlycoSepC-analysis - Test of linearity

Fraction:	Test statistic:	Critical test statistic:	Result:
All (A0 - A4)	0.662	19.16	linear
A0	0.665	19.16	linear
A1	0.667	19.16	linear
A2	0.660	19.16	linear
A3	0.661	19.16	linear
A4	0.670	19.16	linear

Table 31: Aminophase-analysis - Test of linearity

Structure:	Test statistic:	Critical test statistic:	Result:
All (HM - TFR)	0.499	199.5	linear
HM	0.500	199.5	linear
G0F	0.493	199.5	linear
G1F	0.503	199.5	linear
G2F	0.497	199.5	linear
G3F	0.499	199.5	linear
G4F	0.500	199.5	linear
G4FR	0.497	199.5	linear

As it became obvious from the data above, linearity for both methods could be shown for all structures.

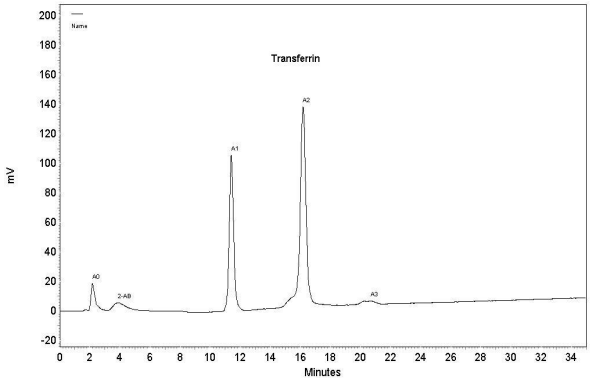
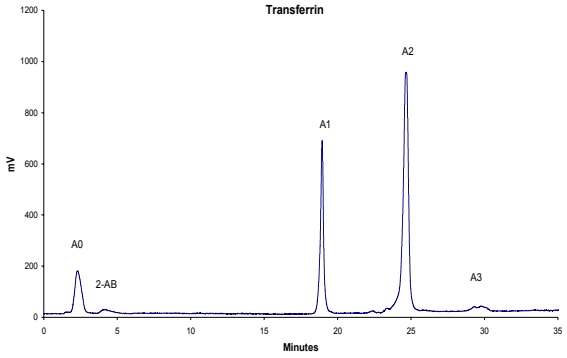
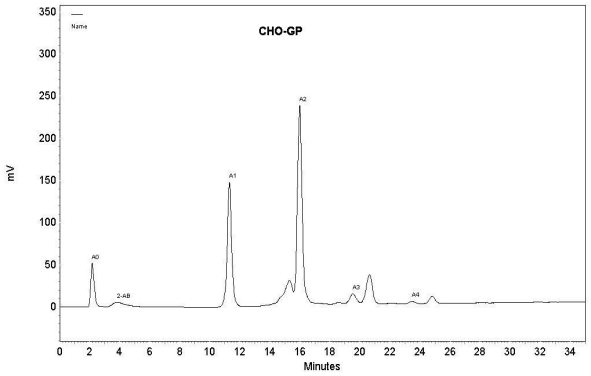
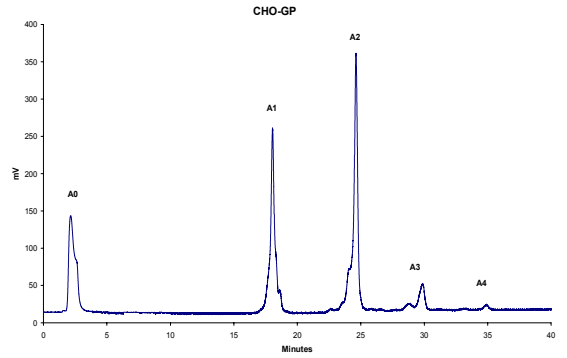
5.3.4 Determination of specificity

Specificity is the ability to assess unequivocally the analyte in the presence of components which may be expected to be present. Typically these might include impurities, degradants, matrix, etc. The best way to examine the specificity of a chromatographic method is to combine it with another method which can identify the analytes, for example mass spectrometry which is very specific regarding identification of substances. For Aminophase-analysis, this approach was chosen. Peaks were fractionated and analyzed by MALDI-TOF-MS. For GlycoSepC-analysis, this was not possible, because of the restriction of the applied MALDI-method which could only measure positive charged oligosacharides (desialylated, neutral glycans). Therefore GlycoSepC-chromatograms were compared with reference chromatograms of another laboratory and peak resolution was calculated.

GlycoSepC-analysis:

The GlycoSepC-column generally separates charged glycans into charge classes [155, 156]. The specificity of GlycoSepC-analysis was demonstrated by comparison of chromatograms of two reference substances with the chromatograms of the same method from another laboratory at the Freie Universität Berlin [140]. The reference standards were transferrin, a human, well examined glycoprotein and CHO-GP (Table 32). Transferrin has been described to have two N-linked glycosylation sites with more than 90% of biantennary structures [157].

Table 32: Comparison of GlycoSepC-chromatogram at ProBioGen with the corresponding chromatogram at university

GlycoSepC-chromatogram at ProBioGen:	Reference-chromatogram (university):
	
	

Specificity of a method can also be supported by determination of the peak resolution. Well resolved peaks are not implicitly necessary when the accuracy of the method can be shown. But for the sake of completeness, resolution of the peaks was calculated (Table 33) by the following formula [158] (Figure 36).

$$R_S = \frac{1.18 * \left({}^{2}t_{m+s} - {}^{1}t_{m+s} \right)}{{}^{2}b_{0.5} + {}^{1}b_{0.5}}$$

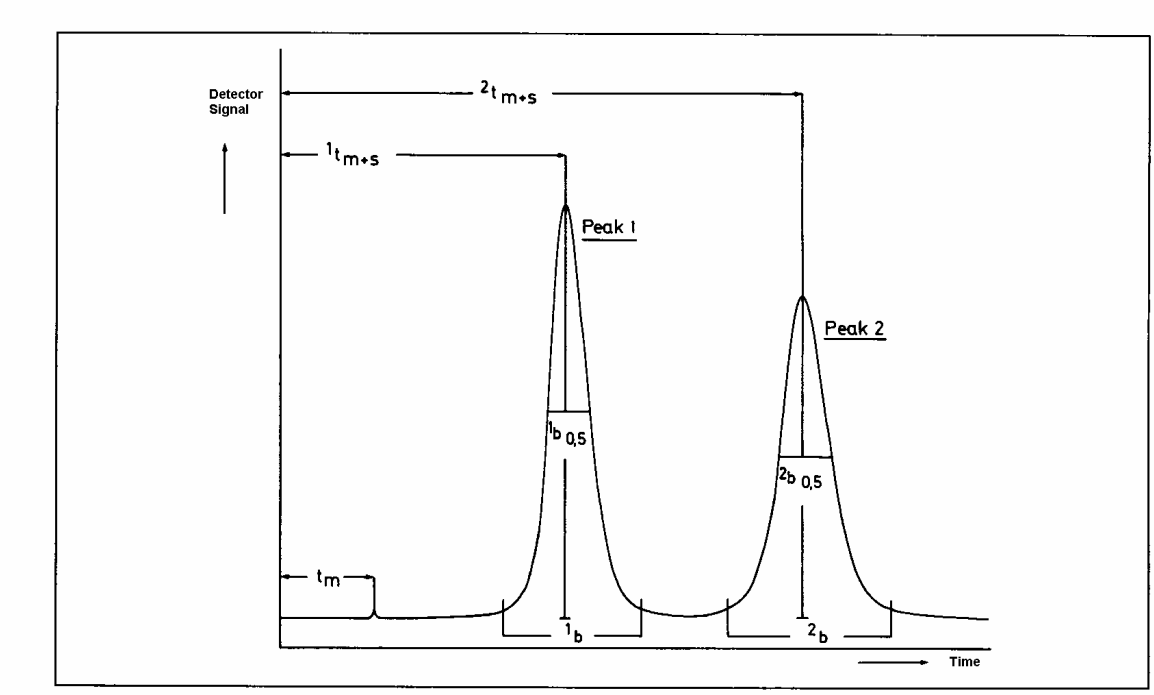


Figure 36: Legend to the formula above *(modified from [158])*

Table 33: Resolution of GlycoSepC-peaks for transferrin and for CHO-GP[23]

Sample	A0	2-AB	A1	A2	A3	A4
Transferrin	N/A	1.6	7.1	8.3	3.3	N/A
CHO-GP	N/A	2.1	8.7	4.5	8.3	3.4

For a baseline-separation a peak resolution higher than R_s = 1.4 should be reached. As it becomes obvious from the data, this criteria is reached for all charge classes. However, within the charge classes (e.g. CHO-GP: A2, A3, A4) peak resolution could not always be reached perfectly. This was due to microheterogeneity of the charge classes caused by the huge amount of possible glycan isomers in each charge class.

[23] *CHO-GP-resolution was calculated between the two nearest peaks of two charge classes, e.g. last peak of A3 and first peak of A4.*

Aminophase-analysis:

To show specificity for Aminophase-analysis, HPLC-peaks of a CHO-GP-sample (Figure 37) were fractionated, desalted and subjected to MALDI-TOF-MS-analyses (Table 34). Because of the restricted detection limit of the MALDI-TOF-MS, not all fractionated peaks resulted in useful spectrograms. Therefore, a spectrum of the whole sample was taken to get an overview of all available CHO-GP-structures (Figure 38). The aminophase column generally separated oligosaccharides by molecular weight, although other separation principles seemed to be important as well, e.g. regarding the double peak in the G4F-fraction which probably resulted from linkage differences within G4F-glycans, maybe due a fucose not bound at the core-structure, but at the antennae of the N-glycan-structure. But this phenomenon was not further examined in this work.

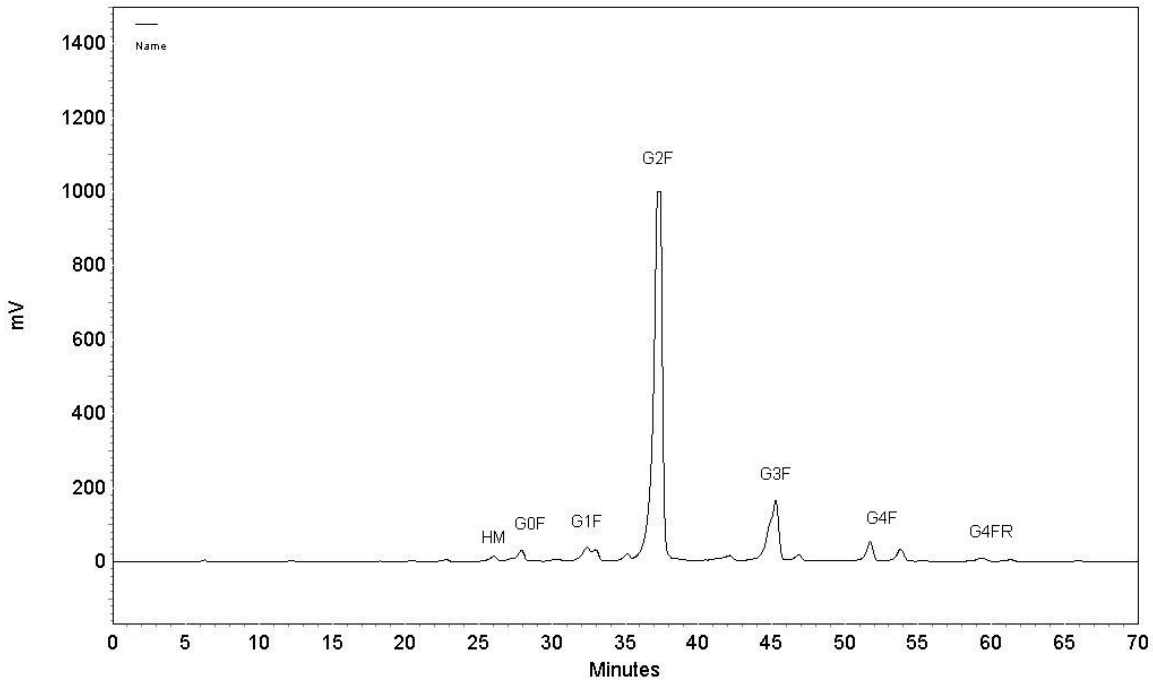


Figure 37: CHO-GP-aminophase chromatogram used for peak fractionation and MALDI-TOF-MS-analyses

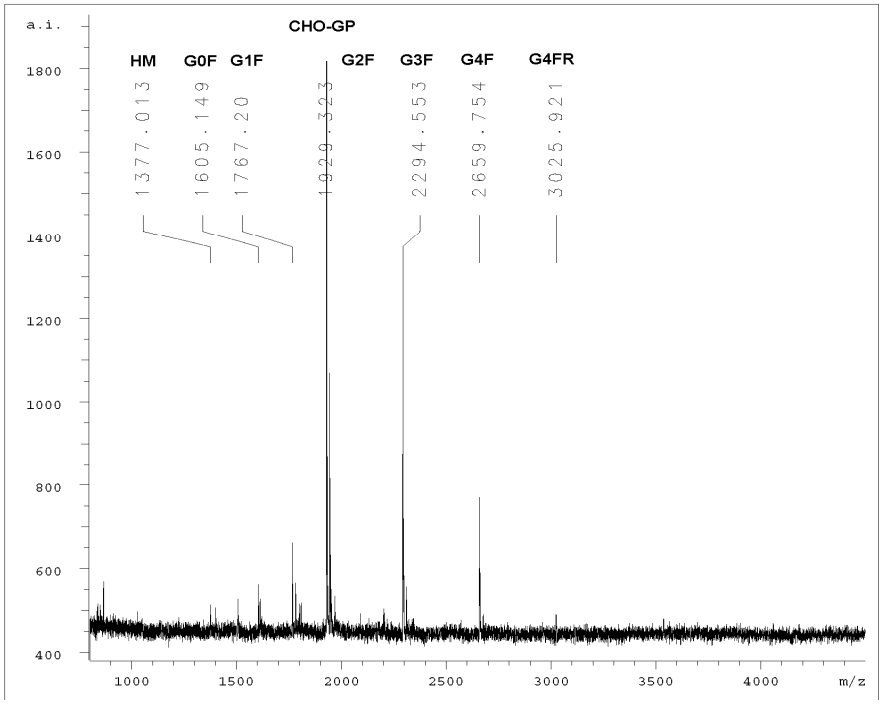
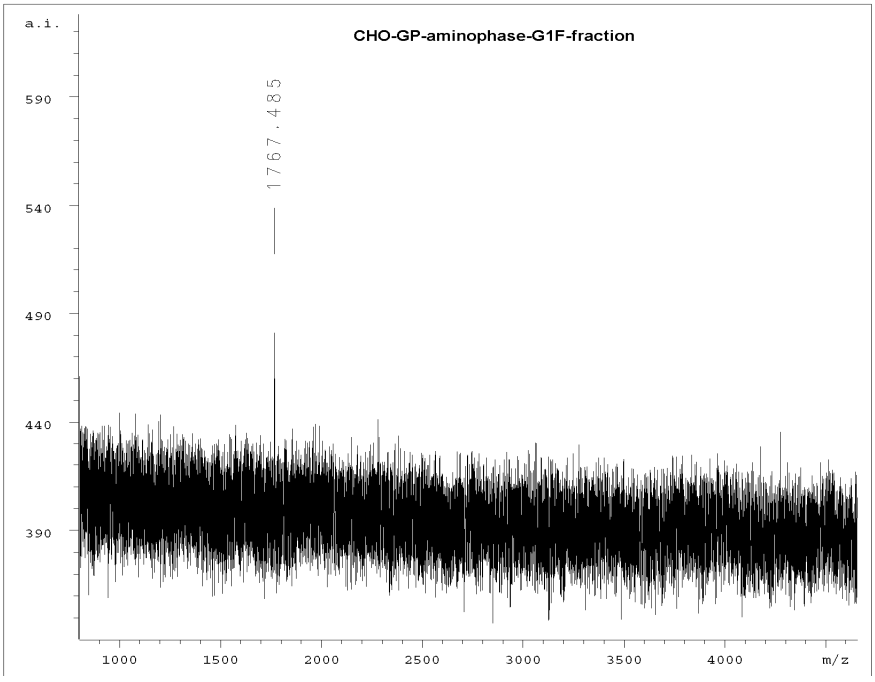
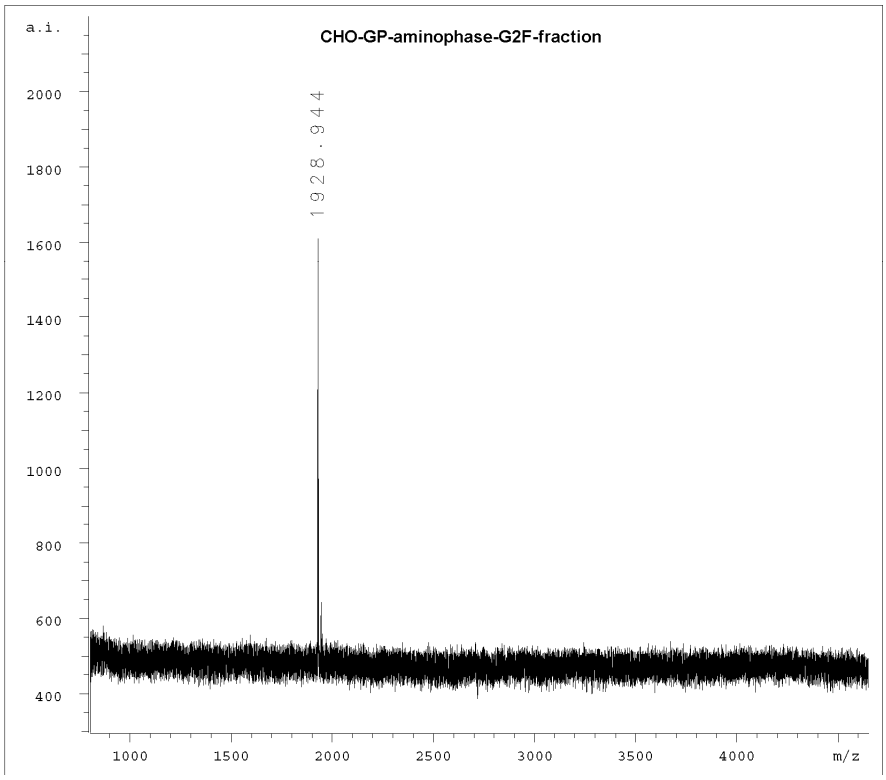
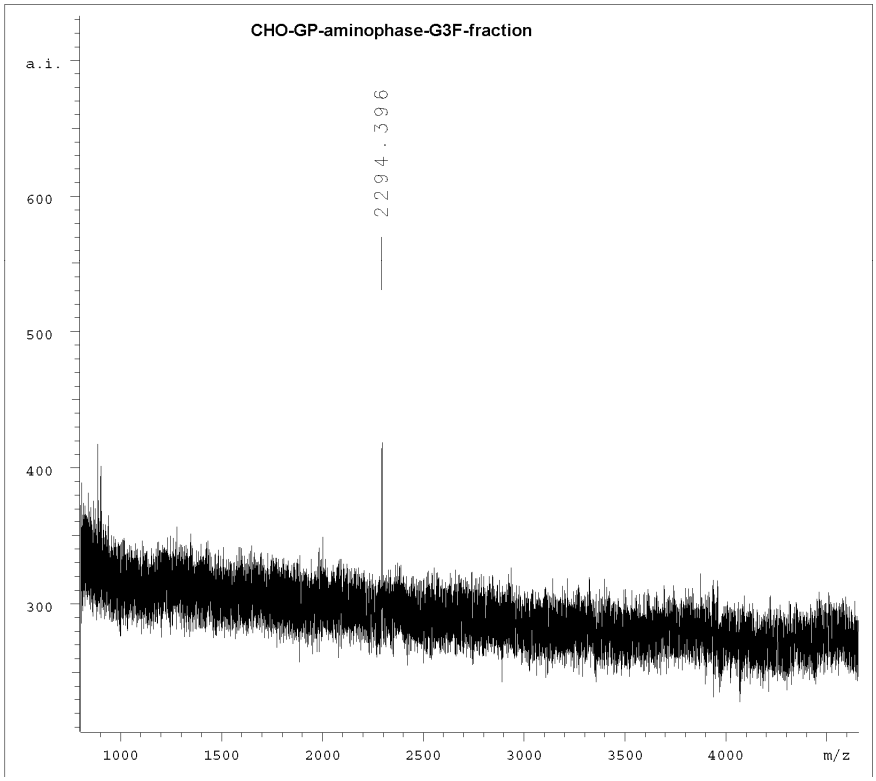


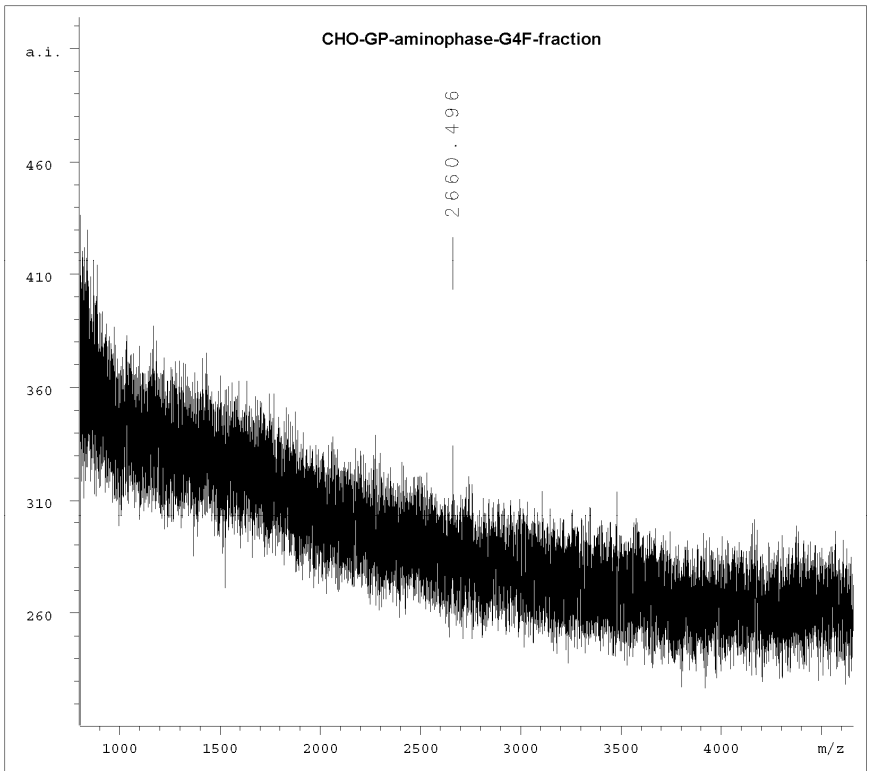
Figure 38: MALDI-TOF-MS of whole desialylated CHO-GP-sample

Table 34: MALDI-TOF-spectra of fractionated aminophase peaks[24]

Peak	MALDI-TOF-MS-spectra of Aminophase-HPLC-fractions	2-AB-structures
G1F		m/z = 1767 = G1F

[24] *HM-, G0F- and G4FR-aminophase fractions did not result in evaluable MALDI-patterns because of the restricted detection limit of MALDI-TOF-MS compared to fluorimetric HPLC-detection*

G2F		m/z = 1929 = G2F
G3F		m/z = 2294 = G3F

G4F		m/z = 2660 = G4F

As it got obvious from the MS-spectra above, each peak fraction showed only one structure, representing high specificity for the Aminophase-HPLC.

5.3.5 Determination of the detection and quantification limit

The determination of the detection and quantification limit was performed by examining the linearity data of the GlycoSepC- and the Aminophase-analyses. Because of the weighted linear regression approach, specialized mathematical calculations for LOD and LOQ had to be used [159, 160].

In this case, the approach of Zorn and Oppenheimer was chosen [160, 161].

$$LOD = LC + \frac{t_{(1-\beta,n-p-2)} s_w}{b_{1w}} \left[\frac{1}{w_{LOD}} + \frac{1}{\sum w_i} + \frac{\left(LOD - \bar{X}_w \right)^2}{Sxx_w} \right]^{\frac{1}{2}}$$

$LC = Critical\ Level\ in\ concentration\ units$
$b_{1w} = weighted\ slope$
$s_w = weighted\ residual\ s\tan dard\ deviation$
$w_{LOD} = weight\ at\ the\ \det ection\ \lim it$
$w_i = weight\ at\ concentration\ i$

$\bar{X}_w = weighted\ arithmetic\ average\ of\ x$
$Sxx_w = weighted\ sum\ of\ x - deviation - squares$
$t_{(1-\beta,n-p-2)} = t - factor\ in\ dependence\ of\ \beta - failure\ and\ p$

with

$$LC = \frac{Y_C - b_{0w}}{b_{1w}}$$

$Y_C = Critical\ Level\ in\ response\ units$
$b_{0w} = weighted\ \text{int}ercept$

and

$$Y_C = b_{0w} + t_{(1-\alpha,n-p-2)} s_w \left[\frac{1}{w_0} + \frac{1}{\sum w_i} + \frac{\bar{X}_w^2}{Sxx_w} \right]^{\frac{1}{2}}$$

Based on the weighted least-squares model

$$Y = b_{0w} + b_{1w} X + \varepsilon ,$$

the variables above were defined as follows:

$$b_{1w} = \frac{Sxy_w}{Sxx_w}$$

$$Sxy_w = \sum w_i \left(X_i - \bar{X}_w \right) Y_i$$

$$Sxx_w = \sum w_i \left(X_i - \bar{X}_w \right)^2$$

$$\bar{X}_w = \frac{\sum w_i X_i}{\sum w_i}$$

$$\bar{Y}_w = \frac{\sum w_i Y_i}{\sum w_i}$$

p = number of parameters used to model the weights, in this case p = 3, due to a quadratic model to calculate the standard deviation at an unknown concentration x:

$$s_x = a_0 + a_1 X + a_2 X^2$$

To avoid an iterative solution for LOD, Oppenheimer's conservative assumptions were used, resulting in larger estimates of LOD:

$$w_0 \approx 1$$
$$w_{LOD} \approx 1$$
$$\left(LOD - \bar{X}_w \right)^2 \approx \bar{X}_w^2$$

LOQ was calculated by the approach of Gibbons et al. [162].

$$LOQ = \frac{Y_Q - b_{0w}}{b_{1w}}$$

with

$$Y_Q = 10 s_{LOC} + b_{0w}$$

Alpha- and beta-level were set to 95% according to the ICH-Guidelines [159].

GlycoSepC-analysis (Table 35 and 36):

Raw data: Linearity curve GSC - all peak areas (A0 - A4)

Table 35: LC, LOD and LOQ for GlycoSepC-analysis[25]

LC	LOD	LOQ
0,88	***1,75***	***3,72***

Aminophase-analysis:

Raw data: Linearity curve AP - all peak areas (HM - TFR) without x = 3.5 nmol to infer accuracy from this dataset.

Table 36: LC, LOD and LOQ for Aminophase-analysis

LC	LOD	LOQ
2,02	***4,03***	***11,76***

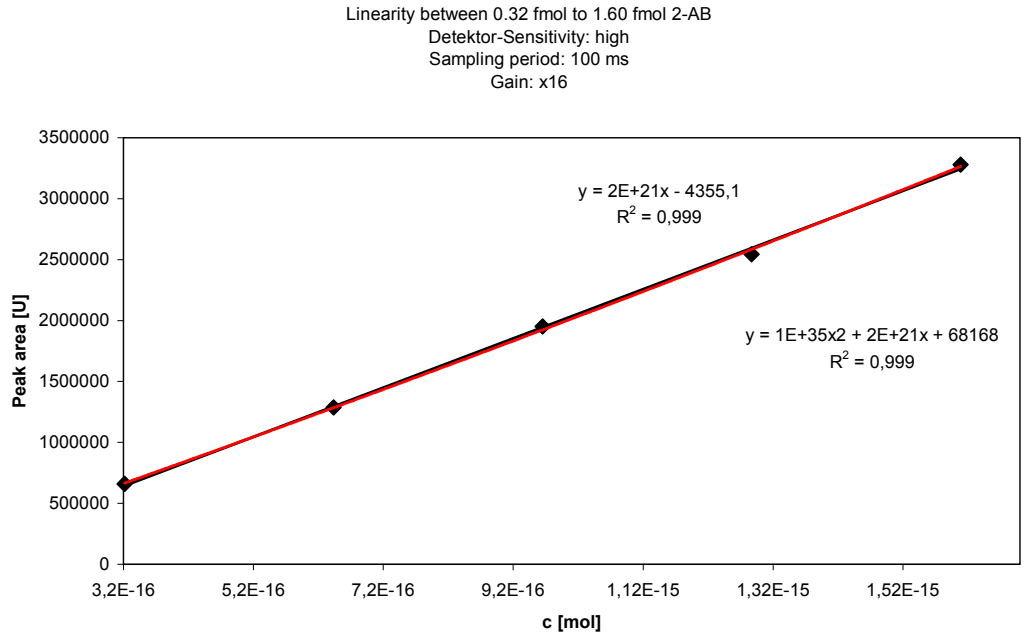
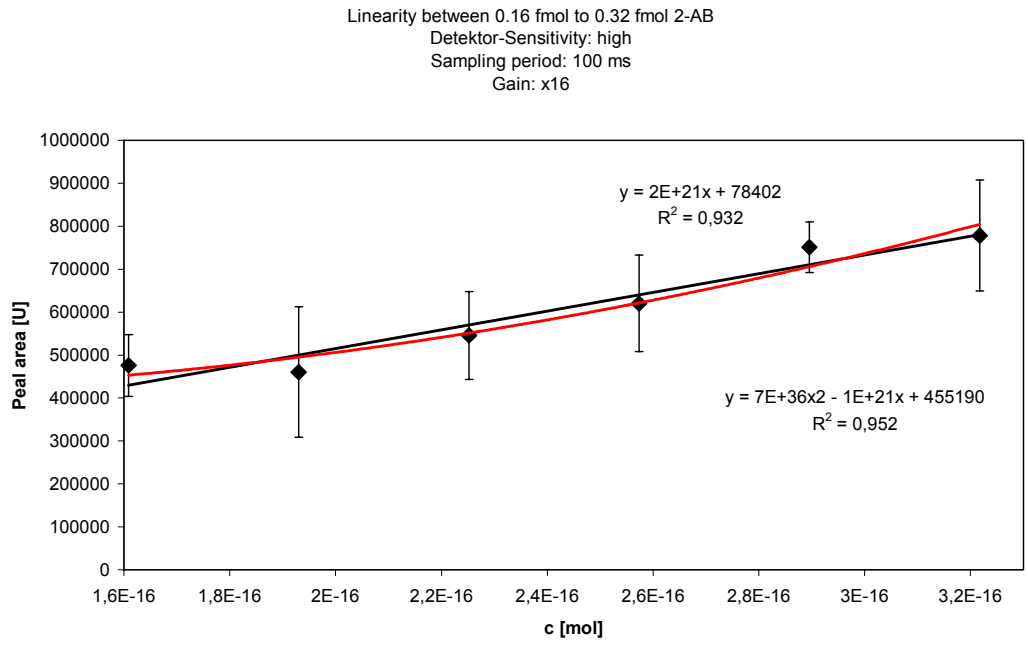
As it was not surprising, LC, LOD and LOQ laid higher for Aminophase-analysis than for GlycoSepC-analysis.

5.3.6 Determination of detector-LOD and -LOQ

Independent from the calculations above, the overall detector sensitivity was determined to get an idea of the LOD and the LOQ of the labelled samples without regarding the preparation steps. Therefore, the pure 2-AB-label was diluted continuously and peak areas were examined. When linearity became worse, repetitions of runs were performed at this specific concentration level (Table 37).

[25] *in nmol CHO-GP*

Table 37: Determination of the lower concentration limit of linearity of diluted 2-AB-samples

Black line = linear regression curve
Red line = quadratic regression curve



To test the goodness of linearity, correlation coefficients of linear regression versus quadratic regression were calculated. Below 0.32 fmol 2-AB the quadratic regression showed higher correlation coefficients than the linear regression. Within this range of concentration, linearity could be confirmed until 0.19 fmol 2-AB were reached. The residual standard deviation of the peak areas were determined from 0.32 - 0.19 fmol (linear range) and LOD and LOQ could be calculated by the following formula (Table 38) [159]:

$$LOD = \frac{3.3 * s_y}{S}$$

$$LOQ = \frac{10 * s_y}{S}$$

with

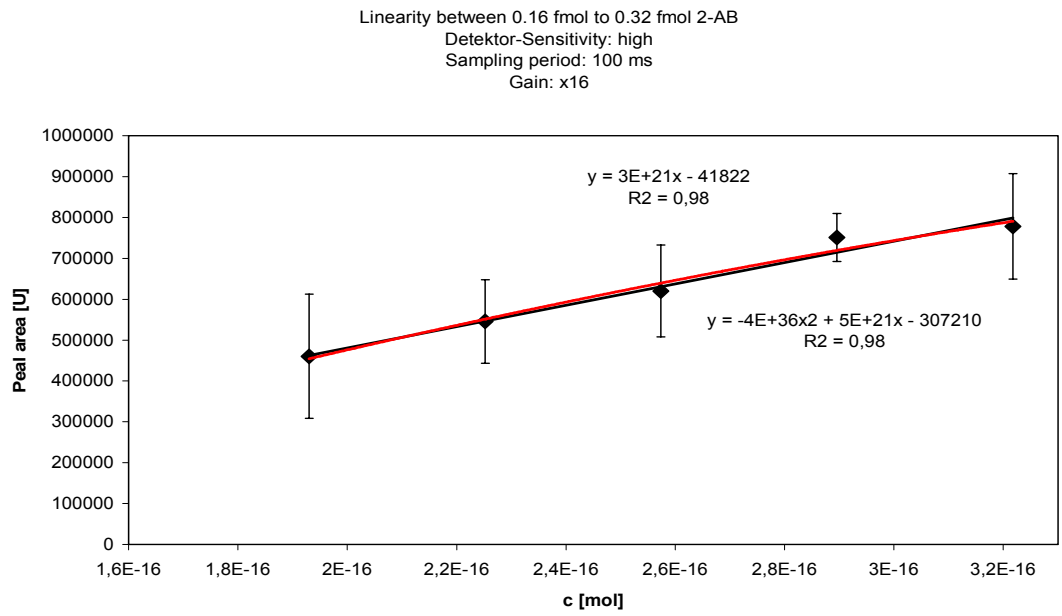
$$s_y = \sqrt{\frac{RSS}{n-2}}$$

$$RSS = \sum \left(y_i - \left(a + bx_i\right)\right)^2$$

S = slope of regression line

Because of variance homogeneity, no weighing factor was needed. GlycoSepC-HPLC was used as HPLC-method.

Table 38: Determination of $LOD_{detector}$ and $LOQ_{detector}$ by analyzing diluted 2-AB-samples

Black line = linear regression curve
Red line = quadratic regression curve


Results:

Detector-LOD = $5.3*10^{-17}$ mol 2-AB = 53 amol 2-AB

Detector-LOQ = $1.6*10^{-16}$ mol 2-AB = 160 amol 2-AB

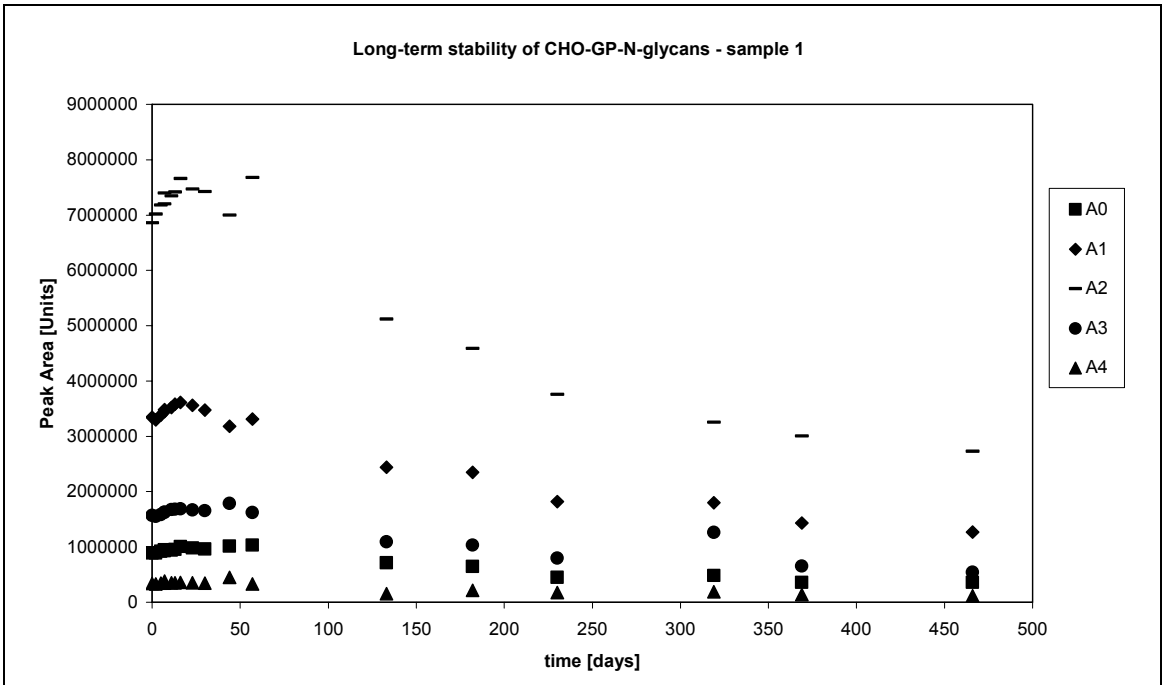
The data above demonstrate the high sensitivity of the flourimetric detection principle, undershooting LODs and LOQs of electrochemical and mass spectrometrical detection by a factor of 1000!

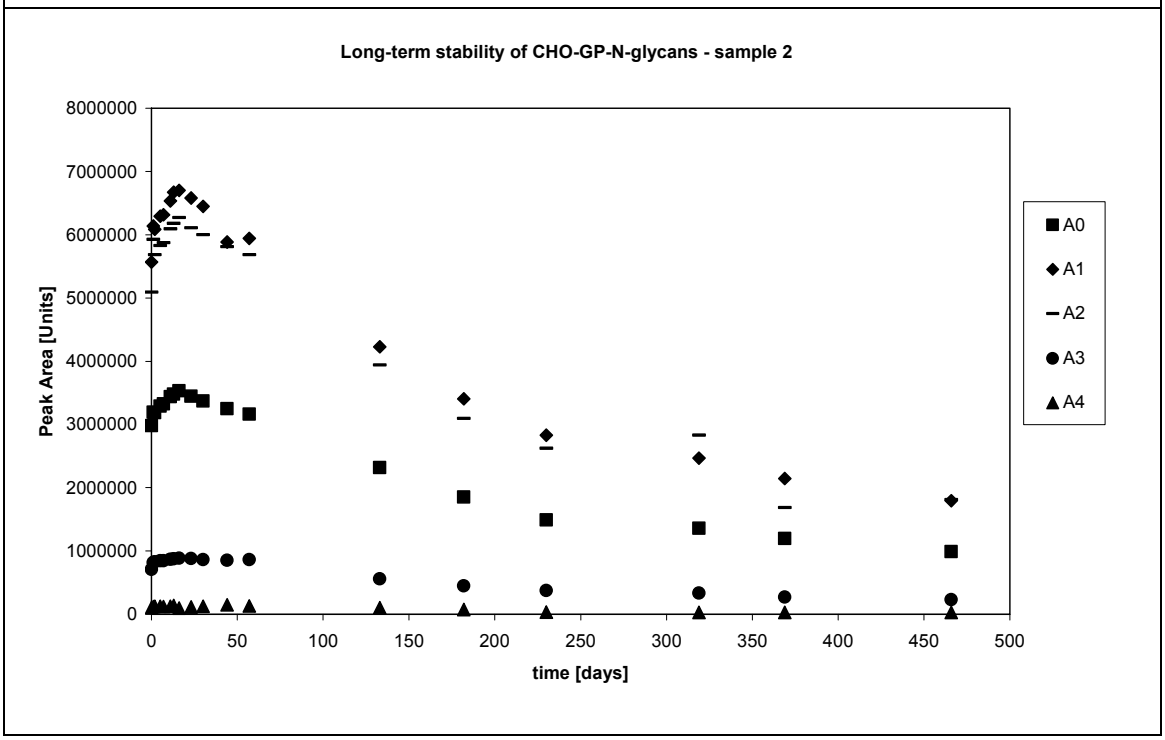
5.3.7 Determination of robustness

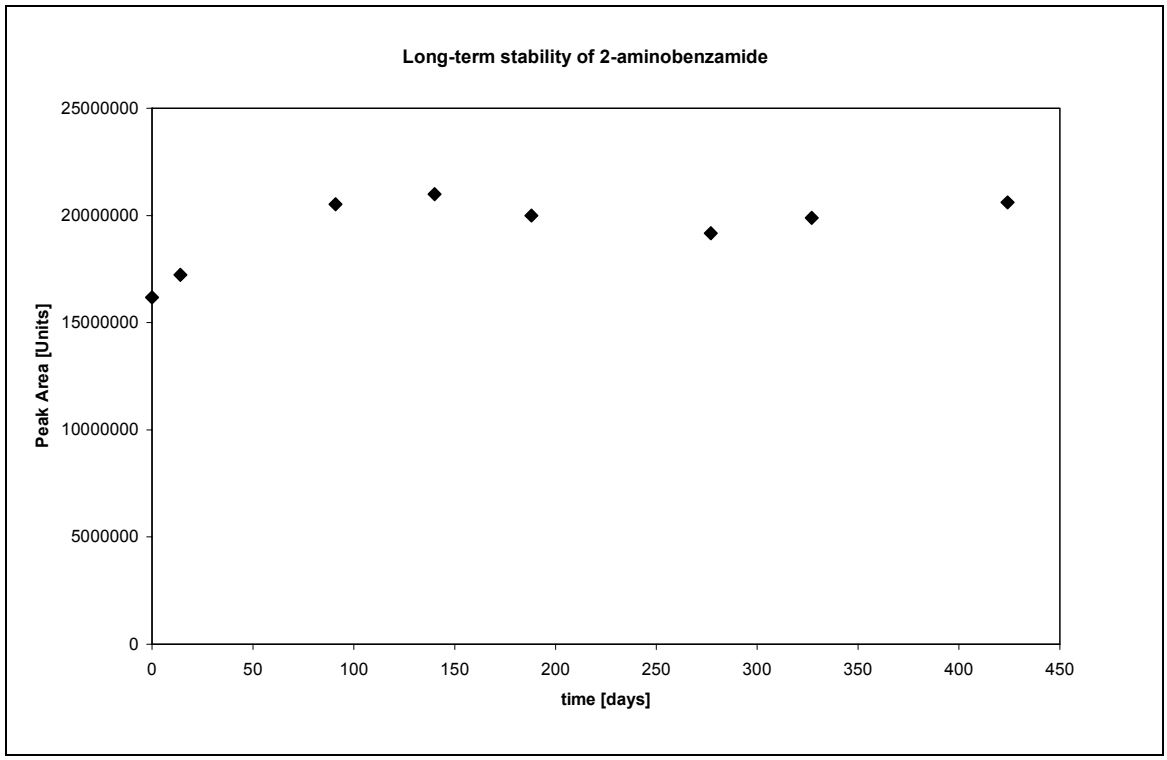
Robustness of the methods was evaluated for early development by examining the stability of the 2-AB-labelled oligosaccharide sample solutions. Samples should normally be stored at -20°C in the darkness by manufacturers' instructions. It should be evaluated how long the samples

would be stable under normal HPLC-conditions. Therefore, the long-term stability of different samples in HPLC-vials at room temperature and daylight was tested over a period of 15 month. GlycoSepC-HPLC-runs were performed at specific points and peak areas were evaluated (Table 39).

Table 39: Long-term stability testing of different 2-AB-labelled oligosaccharide samples and the label 2-AB itself



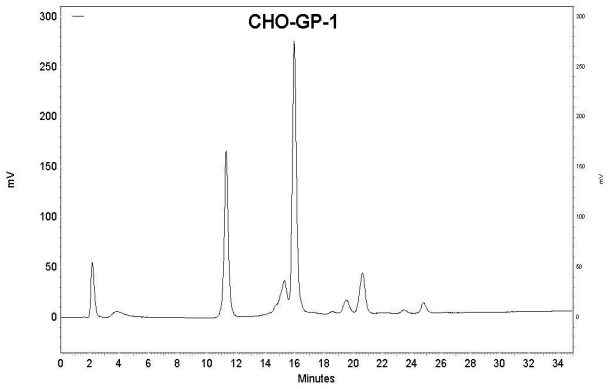
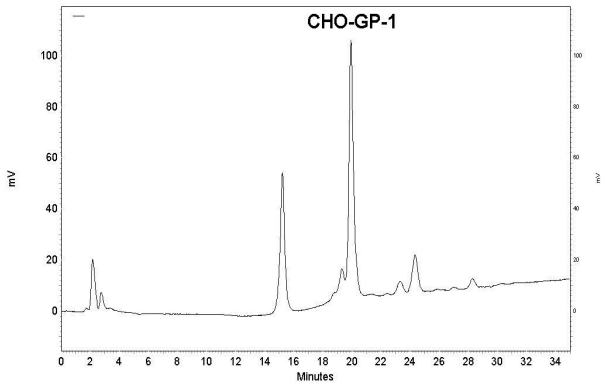
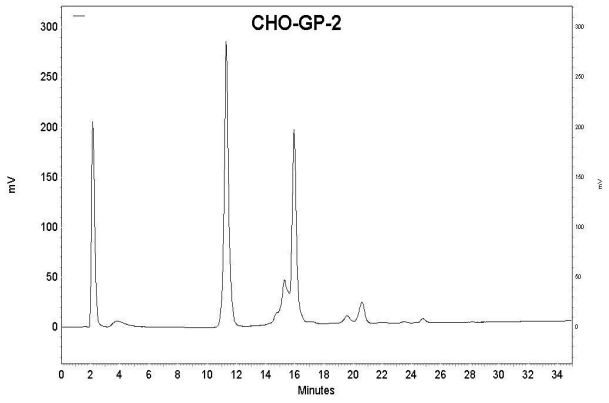
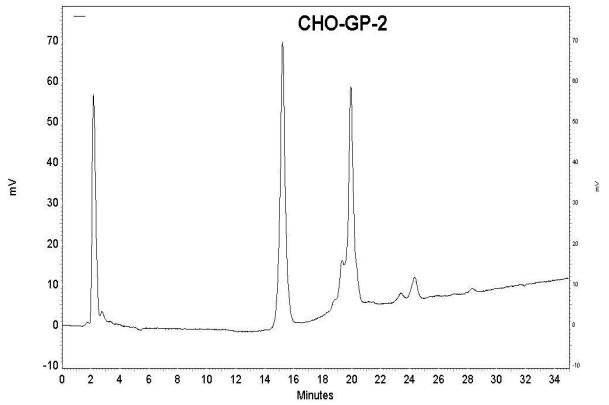




As it became obvious from the data above, the samples were stable over a period of 50 days. After that, a significant decrease of glycan structures could be observed. A temporary increase in peak areas in the beginning of the study was probably explainable by evaporation of the solvent out of the HPLC-vials. The 2-AB-label itself showed no instability. As a result, it could be confirmed that the labelled oligosaccharides were stable under normal HPLC-working conditions (max. 2-3 days storage in the HPLC-device).

Besides, a potential decrease of glycan structures would not have had a disturbing influence on resulting HPLC-patterns analyzed by relative quantification, because there were observed only minimal losses of specific glycan structures. These differences resulted from variable integration parameter settings (Table 40).

Table 40: Comparison of the first and the last long-term stability HPLC run of the CHO-GP-samples above

First GlycoSepC-chromatogram:	GlycoSepC-chromatogram after 15 month:
	
A0 = 6.8% A1 = 25.7% A2 = 52.7% A3 = 12.1% A4 = 2.6%	A0 = 7.2% A1 = 25.3% A2 = 54.4% A3 = 10.8% A4 = 2.3%
	
A0 = 20.6% A1 = 38.6% A2 = 35.2% A3 = 4.9% A4 = 0.7%	A0 = 20.4% A1 = 36.9% A2 = 37.3% A3 = 4.8% A4 = 0.6%