TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

D I P L O M A R B E I T

Procedure design and development of chromatographic test procedures according to the analytical method lifecycle approach

durchgeführt
in Zusammenarbeit mit Lonza AG (Visp, Schweiz)
eingereicht am
Institut für chemische Technologien und Analytik der Technischen Universität Wien

Unter Supervision von
Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Erwin Rosenberg

Und unter der Leitung von
Mag.rer.nat. Dr.rer.nat. Norbert Schelosky

Laura BALTA, BSc
01325292

Wien am

_____
Laura BALTA, BSc

Acknowledgement

At first, I want to thank my supervisor at Lonza AG Mag.rer.nat. Dr.rer.nat. Norbert Schelosky, who made it possible to write my master thesis at the company. And also I want to thank him for all your input he gave me, all the discussions we made about new theories and all the time he took for me and my work.
Next I want to thank my professor Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Erwin Rosenberg, who supervised the master thesis from afar and later for all the input regarding the finalisation of this thesis.

A special thanks goes to my family, especially my parents, who always believed in me, supported me and gave me the strength to make it through my studies even while working at the same time.

Zusammenfassung

Quality by Design ist bereits seit Jahren ein wichtiges Thema in der pharmazeutischen Entwicklung, wobei die Qualität des Produkts über Prozesswissen und kontinuierliche Verbesserungen in einem ständigen Kreislauf gesichert wird.

Dieses Konzept ist zum Beispiel durch die USP (United States Pharmacopeia) <1220> etabliert, hingegen gibt es für die Entwicklung analytischer Verfahren nur Entwürfe von Richtlinien, die zukünftig zu rechtlichen Vorgaben werden könnten.

Um ein besseres Verständnis der Strategie zu erhalten, wurden die Entwürfe des analytischen Lifecycles mit den bereits etablierten Richtlinien der Produkt- und Prozessentwicklung abgestimmt, um eine schnelle und effiziente Methodenentwicklung zu ermöglichen, die für eine große Range an pharmazeutischen Produkten verwendet werden kann.

Das Ergebnis ist ein Leitfaden einer Methodenentwicklung am Beispiel der RP-HPLC (Reverse Phase Hochleistungsflüssigkeitschromatographie), einer der komplexesten und am häufigsten verwendeten Analysetechniken zur Analyse von Arzneimitteln. Basierend auf einem besseren Verständnis des Ziels der Methode im Vorfeld und während der Entwicklung, sowie Risikobewertungen und Robustheitsuntersuchungen mit Hilfe der Software DryLab und Design Expert für eine möglichst effiziente und effektive Nutzung der gesammelten Daten, wurde die Methodenentwicklung durchgeführt.

Dadurch kann zukünftig schneller auf eine mögliche gesetzliche Vorgabe der Richtlinien reagiert werden und gleichzeitig sichert das Konzept die schnellere Umsetzung der Entwicklung neuer analytischer Methoden.

Abstract

Quality by Design has been an important topic in Pharmaceutical Development for the last years.
It is about assuring quality by understanding the process and sources of variation and controlling variables through the knowledge gathered in the lifecycle of the product.
This concept is already formally established in product and process development, but there is only a draft for the implementation in the analytical procedure development [1].

For a better understanding of the strategy (*Analytical) Quality by Design*, this thesis addresses the theory behind the question *What is the analytical method lifecycle?* and describes the commonly used abbreviations and terms in the draft guidelines like the USP <1220> and their implementation in the analytical approach.
The outcome is a method development guideline for the example of reversed phase high performance liquid chromatography, one of the more complex and very commonly used analytical techniques in the analysis of small molecules drug substances, with tools for better understanding the aim of the method beforehand and throughout the development, as well as descriptions of the softwares DryLab and Design Expert, for a most efficient and effective use of the gathered data.

So, this thesis is mostly for readers aiming for a faster analytical method development with more robustness understanding, to get more reliable data. And also, for understanding the (analytical) method lifecycle and how to implement the idea behind it in the laboratory.

Contents

List of Abbreviations

α ..... Selectivity
(A)CS ..... (Analytical) Control Strategy
(A)QbD ..... (Analytical) Quality by Design
ATP ..... Analytical Target Profile
C18..... Octadecyl Carbon chains
CPP..... Critical Process Parameter
CMV ..... Critical Method Variable
CMA ..... Critical Method Attribute
CQA ..... Critical Quality Attribute
CMM..... Continuous Method Monitoring
DAD..... Diode-Array Detector
DoE ..... Design of Experiment
DS ..... Design Space
ECV..... Extra Column Volume
FMEA ..... Failure Mode and Effect Analysis
FL ..... Fluorescence radiation
HPLC..... High-Pressure Liquid Chromatography
ICH ..... International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
I.D..... Inner Diameter
k ..... Retention factor
MODR ..... Method Operable Design Region
MS..... Mass Spectrometry
MV..... Method Variable
N..... Plate count
NP ..... Normal Phase
PFP ..... Pentaflourophenylpropyl
PDA..... Photodiode-Array
PMV ..... Potential (critical) Method Variable
QTPP ..... Quality Target Product Profile
QSM ..... Quaternary Solvent Manager
RA ..... Risk Assessment

r ..... Radius

R..... Resolution

RI..... Refractive Index

RP ..... Reverse Phase

RT ..... Retention time

SM-FTN..... Sample Manager with a Flow Through Needle

SST ..... System Suitability Test

SPP ..... Superficially Porous Particle

t ..... Time

T ..... Temperature

TFA ..... Trifluoroacetic Acid

TPP/FPP ..... Total/Fully Porous Particle

TMU ..... Target Measurement Uncertainty

USP..... the United States Pharmacopeia

UV/VIS ..... Ultraviolet-Visible light

U(H)PLC..... Ultra High-Pressure Liquid Chromatography

V ..... Volume

1 Introduction

The main purpose of analytical quality by design (A)QbD is the establishment of a robust analytical method with the capability of producing high quality results, to guarantee the obtained data is fit for purpose.

QbD is already commonly used in pharmaceutical product development and the application is more and more required in the analytical department as well.

Already the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) guidelines 8, 9 and 10, which address the pharmaceutical lifecycle, quality risk and quality system, are adapted to fit for an analytical method lifecycle as well. [2] [3] [4]

The United States Pharmacopeia USP <1220>, discussed in a Stimuli article [5], contains the guidance to the analytical method lifecycle, which has been further elaborated in the draft *⟨1220⟩ Analytical Procedure Lifecycle* [1].

These guidelines state that the (A)QbD start with setting up the analytical target profile (ATP) and therefore the objective of a method before method development and validation is carried out. Then all the knowledge generated through the development, validation, and routine use of the method is used for its adjustment, which sometimes leads back to earlier steps and therefore result in a lifecycle (see Figure 1). A more detailed version is shown and discussed in section 2.4.

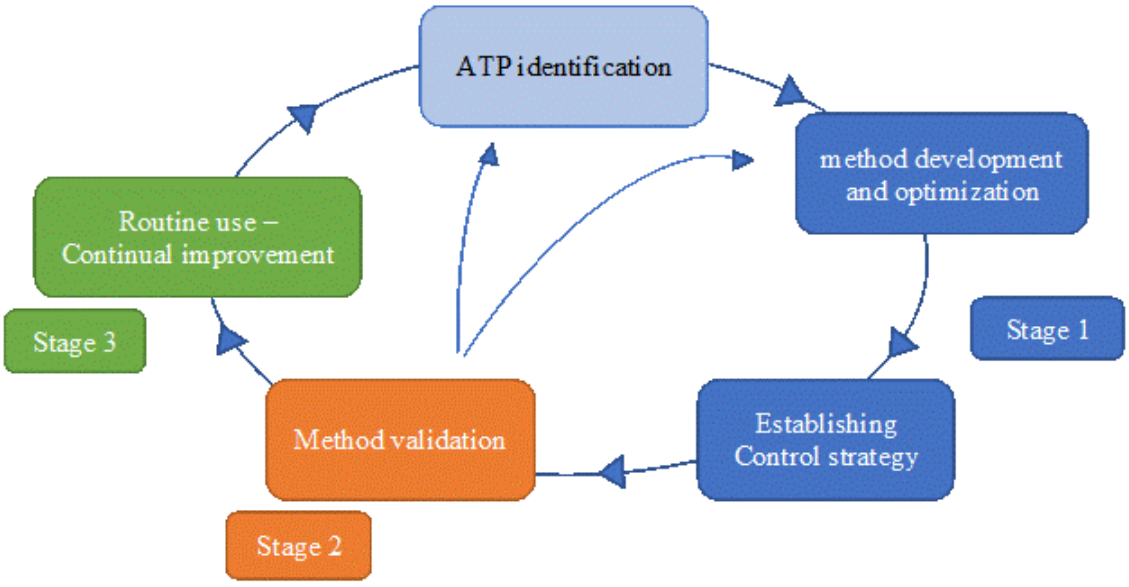

Figure 1: Overview of the analytical lifecycle stages.

The difference between the traditional way of developing a method and the new method lifecycle approach is the handling of data. While traditionally one factor at a time approaches were common, now the variables are investigated with help of know-how and literature through a design space and tools.
Robust methods are important to assure that the results are meeting the ATP criteria at any given time. Additionally, the process must be monitored through a control strategy, so errors can be detected as fast as possible.

The aim of the master thesis is the construction of a method development guideline for a fit for purpose method, based on the understanding of the analytical method lifecycle, its fundamental steps, with focus on the risk assessment and generating a method operable design region.

This method development was created for a high-pressure liquid chromatography (HPLC), but it gives also the base for other analytical techniques.

The experiments were made on a Waters H-Class UPLC and the substance *1-(α-amino benzyl)-2-naphthol hydrochloride* was used as a sample.
This compound was chosen as a proxy for the kind of compounds, for which methods often need to be developed by pharma laboratories.
For optimisation and robustness studies the softwares DryLab and Design Expert 11 were used.

1.1 Motivation

As already explained, this master thesis discusses the approach of an analytical method development based on the application of the criteria from the Quality by Design concept.
The reason for the implementation of the guidelines for an Analytical Quality by Design is the prospect of a robust method, which can be even more optimised throughout routine use.
Another motivation behind this thesis is, that guidelines like the USP <1220> are introduced to the industry and it is only a matter of time, when the Quality by Design approach will be implemented into the analytical method development.

Therefore, this thesis addresses the realisation of the guidelines from the lifecycle into the analytical lifecycle to be better prepared for prospective challenges and new guidelines.

2 Theory

This chapter summarises the theory behind the method development, including the operating principle of an HPLC with UV/VIS detection. It also describes the mobile and stationary phases used for HPLC measurements and the importance of choosing the right conditions for an analyte. Also, the significance of the resolution equation is discussed.

The second part of this chapter contains a summary of the Analytical Method Lifecycle and explains all the acronyms frequently used in this area.

2.1 HPLC

2.1.1 General Principle

HPLC is a separation technique based on the interaction of the analyte with a liquid mobile phase and a solid stationary phase. Depending on the type of analyte, it remains adsorbed on the stationary phase for a specific amount of time, which leads to different retention times for different substances. Therefore, an identification is possible, depending on the time the substance needs to pass the column. The sample must be dissolved and after injection it is pumped through the column containing the solid phase with a stream of solvent, the mobile phase.

The affinity to the solid phase varies with the compound, therefore they have different retention times.

If it has a higher affinity to the stationary phase, it stays longer in the system, while it elutes faster with a higher affinity to the mobile phase.

The different elution time of the compounds is the basis of the HPLC separation. After passing the column, the fluid passes through a detector, which detects, depending on its type, a substance specific property as UV/VIS adsorption, refractive index or conductivity. The measured signal varies in intensity based on the amount of substance, which passes through in the measured moment. These signals are recorded as peaks.

The most common configuration is the Reversed Phase (RP)-HPLC, where the stationary phase is nonpolar, and the mobile phase is polar.

The used detectors are based on measurement of specific substance properties which can be the refractive index (RI), ultraviolet-visible light (UV/VIS) adsorption, fluorescence radiation (FL), electrochemistry (EC) properties, conductivity (CD) and

the substance mass (mass spectrometry MS). In this thesis a photodiode array detector (PDA) was used, which in principle is a UV/VIS absorption detection (see section 2.1.5).

The data gathered from HPLC is commonly used for quantification of known compounds, for checking the purity of samples or in combination with selective detectors as mass spectroscopy, for identification of unknown compounds or confirmation of the identity of substances. There is also the possibility of purifying products through preparative chromatography, but this is not further enlarged upon in this thesis.

2.1.2 The stationary phase

This section follows mainly the information given of K. Stavros in HPLC Made to Measure [6].

Most commonly the column consists of a heavy-walled stainless-steel tube with compression fittings on both ends. On either side of the column a steel disc filter, also called frit serves as protection for the column. On the inlet side it stops too big particles to enter the column, while at the outlet it prevents particles of the column packing to be pushed out and possibly enter the detector cell.

The column is filled with the stationary phase, which can consist of a monolith, superficial porous, total porous or non-porous particles.

Typically, the dimension of a column varies between 18 and 300 mm with an inner diameter of 2 to 4.6 mm for analytical separations and up to 25 mm for semi preparative systems.

In RP-HPLC the stationary phase is commonly silica-based with surface alkyl-chains.

The most popular type of stationary phase is a silica type with bonded octadecyl carbon chains (C18).

Other surface modifications are for example C8-, cyano- and phenyl-bonded phases, which all have different types of affinity.

C18 and C8 phases are both used in environmental science, pharmaceutical industries and for chemical analysis in a wide spectrum. The primary difference is the longer chain of C18 and therefore its more hydrophobic manner and has a bigger steric hindrance for the analyte, which is therefore longer retained inside the column. Larger molecules like proteins and recombinant peptides, which also behave rather hydrophobic are better separated by a C8 or C4 configuration.

The problem with both ligands is, that their size prevents some silanol groups from being derivatised, which leads to the presence of free hydroxy groups on the surface, where ion exchange effects can take place, too.
This can be averted by an “end-capping” process at which smaller alkyl silane reagents replace the remaining silanol group and therefore reduce the negative side effects.

Another type of surface modifications are the phenyl phases, which are most effective for aromatic-samples.
On rare occasions columns packed with Pentaflourophenylpropyl (PFP) are used to separate halogenated compounds.

2.1.2.1 Particle type

The terms fully or total porous particle (FPP/TPP) and superficial porous particle (SPP) describe the type of particle used for the column packing.
While the TPP are wholly made of porous particles, the SPP are made of a solid nonporous silica core, coated with a porous shell similar to the TPP phase.
The difference between the two types of particles is shown in Figure 2. The greatest advantage of SPP is, that particles for the column packing of larger size can be used and still the same efficiency as with smaller totally porous particles can be obtained. The bigger size reduces the back pressure for columns of the same length. While the analyte can pass through the TPP arbitrarily, the SPP has only a small diffusion distance where the compound can pass. Mostly this leads to a smaller loading capacity of the SPP, but it decreases the disadvantage by the compounds passing randomly through the total porous particle with a bigger diffusion distance.

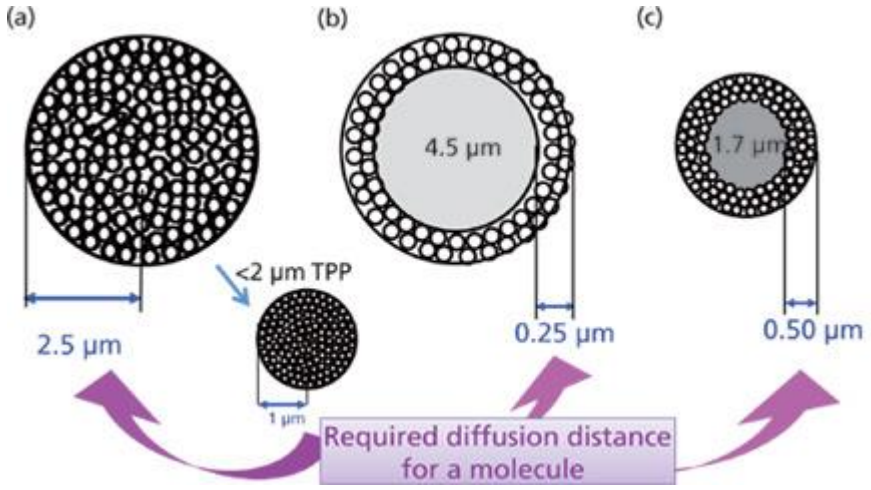

Figure 2: Comparison of (a) 2 and 5 µm total porous particles (TPP) with (b) 5 µm and (c) 2.7 µm superficially porous particles (SPP) and their diffusion distances. [6]

2.1.2.2 Choosing suitable columns for a screening [7]

The type of ligand is the main influence on how strong the analyte is retained. But columns with the same type of stationary phase have still differences. Depending on their packing characteristics like particle distribution, they vary in their efficiency.

To compare the columns on their separation properties a comparison function (also *F*-Factor) is used. It considers the Column hydrophobicity (*H*), steric selectivity (*S*), column hydrogen-bond acidity (*A*), column hydrogen-bond basicity (*B*) and the column cation exchange (*C*) of both columns through the following equation called the hydrophobic subtraction model. These parameters are experimentally examined through testing of different solutions as described in *The hydrophobic-subtraction model of reversed-phase column selectivity* from Snyder, Dolan and Carr [8] and gathered in databases like the *USP database* [9].

$$F = \sqrt{w_H(H_2 - H_1)^2 + w_S(S_2 - S_1)^2 + w_A(A_2 - A_1)^2 + w_B(B_2 - B_1)^2 + w_C(C_2 - C_1)^2} \qquad (1)$$

- ***H* … hydrophobicity**; the higher H the longer the retention time of hydrophobic compounds
- ***S* … steric selectivity**; decreases for phases with higher steric hindrance, as they prevent the interaction of bulky solutes with the stationary phase.
- ***A* … column hydrogen-bond acidity**; the larger A the longer the retention time of compounds with hydrogen-bond acceptor properties like aliphatic amides.
- ***B* … column hydrogen-bond basicity**; the larger B the longer the retention time of compounds with hydrogen-bond donor properties like phenols.
- ***C* … column cation exchange**; the higher the negative charge (more ionised silanol groups), the higher C and the longer the retention time of protonated bases.
- ***wx* … weighting factor** for X = H, S, A, B and C [8]

If the value of *F* is below 3, the columns are well matched, between 3 and 5 they are still adequate and if it is higher than 5 the columns do not match.

This means, if a column is needed to replace an old one, but needs to have the same characteristics, a low *F* value is desirable. For a bride screening of different phases for an analyte, the *F* values of the columns should be as wide apart as possible, to examine a wide range of different selectivities.

2.1.3 The mobile phase [10]

The most important step in choosing the right mobile phase for most pharmaceutic substances is selecting the right pH.

Most pharmaceutical analytes are bases, which are ionisable. Depending on the state of the bases (neutral or ionised) their retention time can differ by a significant amount. This can be controlled through the right pH value in the mobile phase.

If the pH value differs 2 units from the pK_a of the analyte, the compound is either almost wholly ionised or neutral. If their values vary by less than 2 pH units from the pK_A or pK_B, the analyte is only partially dissociated and therefore small changes in the pH can lead to large differences in the retention time.

Therefore, for a good reproducibility the pH value must be chosen with an appropriate distance to the pK_a of the analyte and it must be strictly controlled.

To ensure the stability of the pH in the mobile phase, buffers can be used. Buffers consists of a conjugated pair of proton donor and proton acceptor, which can compensate changes in the pH value caused by adding bases or acids. Examples are phosphoric acid (approx. pH 2 and 7.2), acetate (pH 4.75) and ammonium (pH about 9.24) buffers.

In addition to the right pH value a mobile phase must also show some other characteristics depending on the detector type.

When a UV/VIS detection is used, the mobile phase must have a low absorption at the measured wavelength. Buffers with phosphoric acid are particularly good for low wavelengths down to 210 nm.

For measuring with MS, a high volatility of the buffer substances is important. In this case phosphate or other non-volatile buffers are unsuitable.

The organic component of the mobile phase in RP-HPLC is commonly methanol or acetonitrile.

2.1.4 System Set Up

An HPLC can be set up in different configurations, but in simple terms, it always contains a solvent reservoir with solvent filters, a degasser, a pump, an injector, a column compartment, which is thermostatically controlled and a detector.

Depending on the set up, the solvent reservoir of this HPLC can vary in the amount of tubings and valves, which sets the number of solvents that can be used at the same run.

Every tubing is equipped with a solvent filter to minimise the possibility of particles in the solvent to pass and damage the columns.

The following degasser separates dissolved gases in the solvents, to reduce noise and cycling in the baseline of the later measured chromatogram and increases the compressibility of the eluent.

As in this analytical method a mobile phase passes through the system and the column, a pump is needed. It is important that the eluent is pumped with a constant flow rate.

The injector or autosampler applies the sample solutions into the system, where it passes through the column. The column is usually thermostatically controlled, and the solution is heated beforehand through a preheater to ensure a steady temperature throughout the separation.

After the separation a detector analyses the component peaks.

As explained above, there are a lot off possible configurations and set ups regarding the solvent reservoirs.

Depending on the structure of the HPLC tubing lengths can also vary. For the most functional configuration, the manufacturer guidelines were considered.

This is important, because all HPLC systems have an extra column volume (ECV), the volume without the column, which the sample passes from injection to detection. The ECV affects the peak shape in the chromatogram. A high ECV means a greater broadening of the peaks when detected. For an isocratic run all peaks have the same broadening as a consequence of the ECV, but for longer retention times it gets less significant because of the simultaneous peak broadening in the column with time. With a gradient there is still a broadening of the peaks, but it is a lot smaller, because the longer retained components are more compressed at the column head and any broadening happening before entering the column is more balanced out. Nevertheless, a high ECV can still lead to a poorer peak shape and efficiency and therefore to a

worse resolution when using a gradient. Also, the shorter the column and the smaller its inner diameter (ID), the higher the efficiency its packing can achieve, the more impact the ECV has on the results.
The ECV is measured with one eluent spiked with an UV active substance. For this measurement, the column is replaced by a zero dead volume union. The delayed time of the start of the use of the spiked mobile phase until the substance is measured in the detector is equate to the ECV:

The dwell volume is another important value of the specific device and depends on the type of pump and set up before the column.

It is the reason for the time delay of changes in the mobile phase.
It is the difference of volumes between the point of mixture of the mobile phases in a gradient and the head of the column and therefore the volume before the mixed gradient reaches the column.
This includes the volume of the eluent mixer, all tubings, the pump head, the injector, all valves and also the column inlet.
If known, the injection of a sample can be delayed or an isocratic hold introduced at the beginning of the gradient program, matching the dwell volume, to prevent any differences based on the gradient delay time.
This is especially important for method developments, which are done on another device than the routine use will take place, as the dwell volume is specific for every device.
To measure these volumes with UV/VIS detection, UV active substances are injected into the HPLC and depending on their delay in measurement the volumes can be detected.
For the dwell volume also one eluent is spiked with an UV active substance. Then a gradient of this spiked eluent from 0 to 100 % is measured. The dwell time can be read off the measured chromatogram. It is the time the system reaches half of the gradient change minus half of the actual programmed time in the gradient.
An example for an UV active substance used for these measurements is acetone, which has the highest absorption peak at 265 nm.

2.1.5 The detection

As already mentioned in section 2.1.1, different types of detectors are applied in the field of HPLC.

For this master thesis a photodiode-array (PDA)-, also known as diode-array detector (DAD), based on UV/Vis radiation was used.

Figure 3 shows the light path inside the detector.

The detection is based on the amount of light, which reaches the photodiode array after passing through the sample in the flow cell.

The photo diodes have a fixed initial amount of charge, which is discharged by contact of the light of the source after interacting with the sample. The amount of discharge of the photodiode depends on the intensity of the light and afterwards they are recharged. The detector measures the amount needed to recharge the photodiodes. This is measured with a beam through the sample and a reference beam. The ratio of light transmitted through the flow cell in comparison to the reference beam, in a specific exposure time is then used to calculate the absorption.

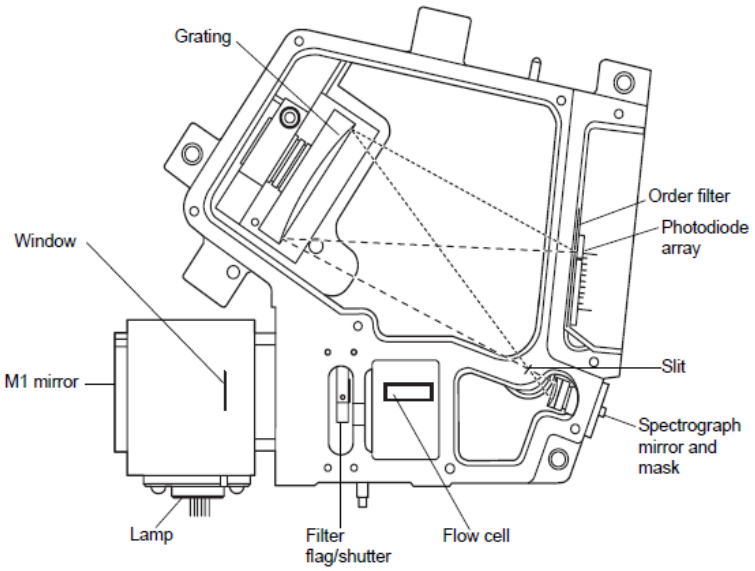

Figure 3: Light path inside the PDA detector used for this thesis according to the manufacturer. [11]

2.1.6 Resolution equation

A good resolution *R* is the primary aim of a suitable chromatographic method because it states if peaks are sufficiently separated.

There are two ways that can be used to calculate the resolution. First is the calculation through the *retention times* t_R and *peak widths w*, which can differ by the type of width (at the baseline, at 50 % height etc.) that is used. The USP resolution, which had been applied in this master thesis, uses the *peak widths w* at the baseline between the tangent lines drawn at 50 % peak height through the following equation:

$$R = \frac{2 \cdot (t_{R2} - t_{R1})}{(w_2 + w_1)} \tag{2}$$

Formula (2) is typically used to calculate the resolution of two peaks in a chromatogram. Formula (3) shows the relation of the three factors contributing to the resolution, which are the *plate count N,* also *efficiency*, *retention factor k* and *selectivity α,* also called *separation factor*.

$$R = \frac{k}{1+k} \cdot \frac{1}{4} \cdot \sqrt{N} \cdot \frac{\alpha - 1}{\alpha} \tag{3}$$

The *retention factor k* is one of these factors to consider when developing a method. It specifies how long a compound needs to pass through the system to the detection compared to the mobile phase. The retention factor is calculated through the following equation through the *retention time of the compound* t_R and of the *mobile phase* t_0. The subtraction of t_0 from t_R leads to the *reduced retention time* t'_R of the compound.

$$k = \frac{t_R - t_0}{t_0} = \frac{t'_R}{t_0} \tag{4}$$

Through this relation it becomes clear, that the retention factor is independent of the instrumental conditions like ECV or column dimensions, because they affect both, the retention of the compound and the passing of the mobile phase, through the system equally.

The value of *k* only changes if the interaction to the mobile and stationary phases varies.

Starting from a pair of unretained compounds, increasing *k* leads to a better resolution, but the influence of k on R decreases with a higher value. At a retention factor of approximately five, the impact on R is too small to be furthermore advantageous.

The *plate count N* is a measure of column efficiency. The higher the column efficiency, the smaller is the width of a given peak which leads to a better separation.
It can be calculated through the *retention time* t_R and *peak width at the baseline* w_B or *at half height* $w_{\frac{1}{2}}$ as followes.

$$N = 16 \cdot \left(\frac{t_R}{w_B}\right)^2 = 5.54 \cdot \left(\frac{t_R}{w_{\frac{1}{2}}}\right)^2 \quad (5)$$

The *plate count* is compound specific and is also influenced by the mobile phase, different column parameters (length and particle size) and the configuration of the HPLC. Other method parameters such as injection volume or temperature have an impact too.

The plate count is characteristic for columns but can be influenced by the system set up too. As with high ECVs (further discussed in section 2.1.4).
The smaller the ECV the less peak broadening occurs and the sharper the peak. A way to increase the plate count is to use a longer column, but this leads to higher back pressure and longer run times. Also, using smaller particle sizes can increase the plate count, but with the same problem of a higher back pressure.

Another factor, with the highest impact on the resolution, is the *selectivity α*. It represents how good an HPLC method separates two analytes from each other. this is shown in equation (6) with t'_{R1} and t'_{R2} being the retention times or k_1 and k_2 being the retention factors of compound 1 and 2.

$$\alpha = \frac{t'_{R2}}{t'_{R1}} = \frac{k_2}{k_1} \quad (6)$$

Like the retention factor, *α* is also independent of instrumental set up and changes solely with a change in the chemistry, like the affinity to the phases.
The selectivity can be influenced by changing the stationary phase type (more nonpolar, other substitutes like phenyl-groups etc.) as discussed with the F-value. Depending on the interaction of the analyte with the stationary phase, other retention times and resolutions are achieved.
Another way to change *α* is through variation of the mobile phase like adjustments of the pH through modifiers or by changing the main components.

For a better understanding of the impact of all three separation factors on the resolution, the variation of them and their influence on R calculated through equation (3) is shown in Figure 4.
The curves show the variation of one factor, while the other two remain constant.

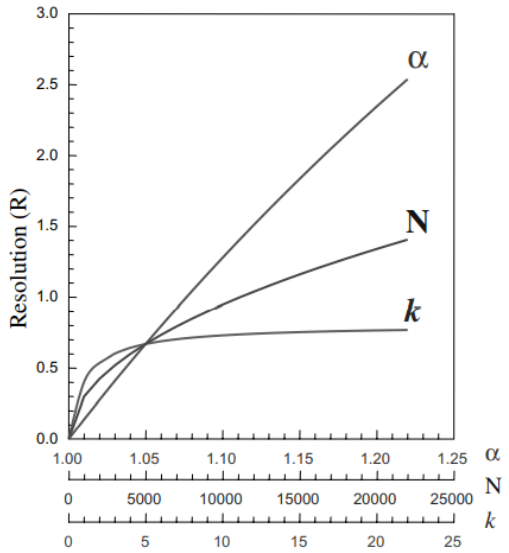

Figure 4: The influence on resolution by varying plate count, capacity factor and selectivity. [12]

Through Figure 4 it becomes clear, that *k* has almost no influence on the resolution, after reaching a value higher than five.
Doubling the plate count leads to a resolution increased by a factor of 1.4, which is still smaller than the influence of the selectivity, which has the highest influence on the resolution.

2.1.7 Characterisation of the peak purity

The purity angle is a mean to determine the purity of a peak. It is comprised out of the average angle of every spectrum of the peak, together with the spectrum at the top of the peak.
A value smaller than 0.2 confirms a pure peak, with no overlays, while angles higher than 1 implicate an overlay of peaks. The values in between must be researched more thoroughly. There are some misinterpretations because differences in shape are not always overlays of sample peaks, but they can also originate for example from the background noise. Mainly, this is the case if the sample concentration is low.
Another value to characterise the peak is the threshold angle. It outlines the effect of the background noise on the peak and is used in comparison to the purity angle.

If the latter is smaller than the purity threshold angle, it can be concluded that no co-elution is present.
Therefore, a higher purity angle than the threshold angle indicates the co-elution of a second substance.

2.2 Drylab

The analytical quality by design approach is based on processing gathered data more quickly and efficient. For this softwares are used, which can be either nonspecific and based on statistics only or they can be developed specifically for a type of analytical method.
In case of Drylab, the software has a focus on chromatographic methods and therefore considers interactions specific for these analytical techniques.
As is described on their website, the software is based on "Solvophobic Theory" of Csaba Horváth [13].
The software also applies considerations about interactions of the typical solvents in HPLC, methanol and acetonitrile with water as the polar eluent.
The gradient separation is based on the change of surface tension between the eluents according to Horvath and his team. Water is strongly lipophobic, because of a high surface tension, which is reduced by blending it with organic solvents as methanol and acetonitrile.
The software is based on data gathered for these conditions and with only two gradients measured experimentally the software can calculate chromatograms for different gradient times.
Depending on the optimisation, Drylab proposes different designs to optimise gradient, temperature and / or pH value.
For gradient and temperature only two set points are needed, while the pH value is mostly the pK_S or pK_B of the analyte and two further set points, one above and one lower than the middle set point.
But the software does not only consider the influence of the mobile phase, but also of the stationary phase.
As the selectivity of columns changes with their surface properties like the amount of free silanol groups, their influence must be considered too.

Starting an optimisation with Drylab, the column type has to be specified, so the software can gather data from the Snyder-Dolan hydrophobicity subtraction database, which has its focus on silanol effects for the separation selectivity.

When executing the design, the data of the beforehand proposed experiments, are uploaded to the software.
Based on the comparison of peak areas, the peaks are tracked for each chromatogram to see the change of retention times of the peaks because of parameter variations.
The peaks with their retention times and areas are listed in tables and above the corresponding chromatograms are modelled.
The first chromatogram gives the peaks identification numbers starting from one to the maximum number of peaks and for the following chromatograms, the order does not have to be the same, as the retention times of peaks can differ with variation of the process parameters.
If the areas of the different chromatograms are matching for one peak, Drylab marks them green, if they do not match, they are shown in read.
After the peak tracking was successful, the software calculates the resolution map.
The next step before the evaluation of data starts, is the entry of the theoretical plate count of the device set up, which must be measured beforehand.
To check, if the data is matching, Drylab shows overlays of the experimental data with calculated chromatograms based on the entered plate count together with the peak tracking information. An example is shown in Figure 5Figure 5, where the experimental data of run 3 of the design used for this thesis is overlayed with the calculated data of Drylab for these conditions.

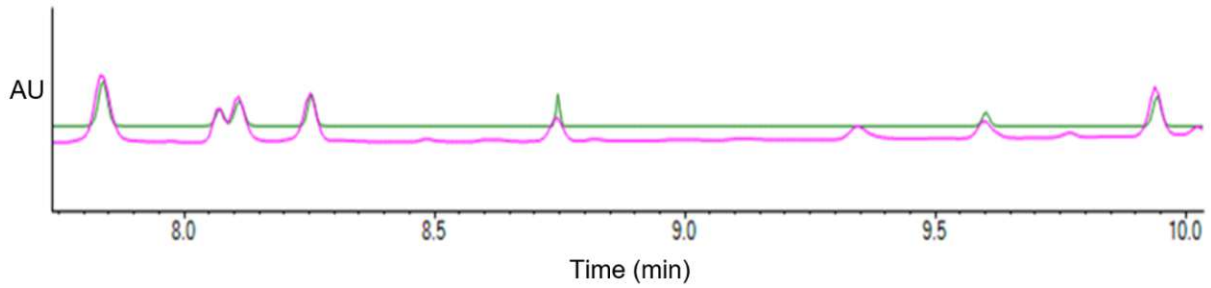

Figure 5: Overlay of the experimental data (pink line, below) with calculated chromatograms (green line, above) based on the entered plate count together with the peak tracking information from Drylab.

If the overlay is matching, it can be concluded, that the design calculation and peak tracking was successful, and the resolution map can be used.

The resolution map shows sectors of the parameter conditions, where the resolution of the peaks is good enough that they do not overlay.

To every spot in the resolution map a calculated chromatogram is shown, where the peak pair with the worst resolution is highlighted in red as is shown in **Fehler! Verweisquelle konnte nicht gefunden werden.**.

With this resolution map the optimum conditions can be calculated.

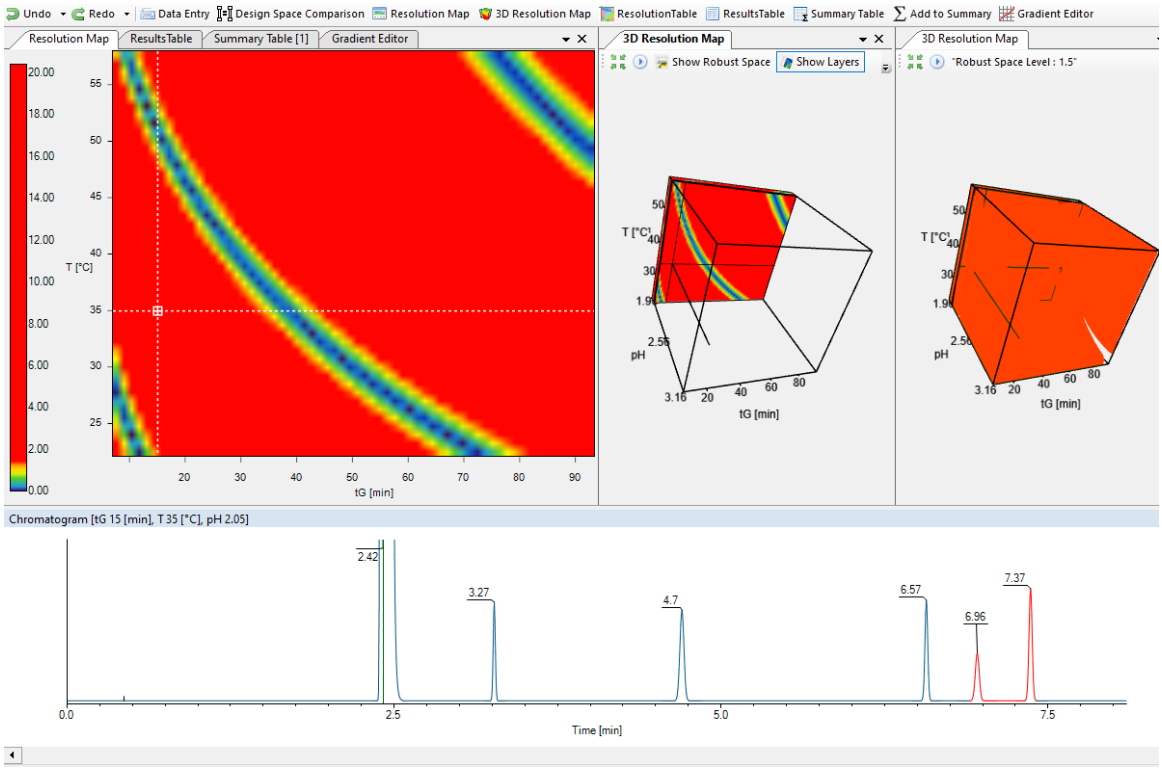

Figure 6: Resolution map of Drylab processed data.
To every spot in the resolution map a calculated chromatogram is shown, where the peak pair with the worst resolution is highlighted in red.

2.3 Definition of commonly used terms in Analytical Quality by Design

2.3.1 (A)QbD – (Analytical) Quality by Design

A systematic approach to development that begins with predefined objectives and emphasizes product and process understanding and process control, based on sound science and quality risk management. [2]

The analytical quality by design is an approach for developing an analytical method. Unlike the traditional approach, where the development proceeded by one factor at a time, the lifecycle examines more factors and their impact on each other simultaneously, for a better understanding of result variations and their reduction. The aim is a more robust method with a high understanding of the performance of the method outside of the set ranges.
AQbD starts with defining an analytical target profile and builds on tools for a more robust method like establishing a method operable design region, to get high quality results. This is one of the biggest advantages of Quality by Design.

2.3.2 QTPP – Quality Target Product Profile and ATP – Analytical Target Profile

A prospective summary of the quality characteristics of a drug product that ideally will be achieved to ensure the desired quality, taking into account safety and efficacy of the drug product. (QTPP) [2]
A fundamental component of the lifecycle approach to analytical procedures is having a predefined objective that stipulates the performance requirement for the analytical procedure. These requirements are described in the ATP. [5]

This means that the Analytical Target Profile is a defined objective, beforehand of the method. Here, the critical method attributes CMA, and their ranges are defined based on knowledge, device limits and guidelines.

Most important is the information about what should be analysed, in which range and which matrix should be used. By comparing the results later in the development with the ATP criteria, it can be confirmed that the produced method is fit for purpose.

ATP is not linked to a type of analytical method. Rather there can be more than one option possible for one ATP.

The main question to answer is what the acceptable values of the CMAs like accuracy, precision, range, specificity etc. are?

Besides these considerations the ATP can also include business requirements such as cycle time and throughput.

2.3.3 TMU – Target Measurement Uncertainty

TMU is the maximum uncertainty that can be associated with a reportable result while still remaining fit for its intended purpose. TMU is a consolidation of the uncertainty from all sources. [5]

The TMU combines the traditional terms precision (for random errors) and bias (for systematic errors or accuracy). Factors with an impact on the attributes and to be considered during AQbD are listed in the *Stimuli* Article [5] and shown in Table 1.

Table 1: Factors that are contributing to bias and precision and are therefore part of the TMU. [14]

Target Measurement Uncertainty (TMU)	
Bias	Precision
Selectivity	Sample preparation
Linearity	Weighing
Extraction efficiency	Instrument
Filter recovery	Integration
Detector Wavelength	Background noise
Solution Stability	Replicate Strategy
Analyte Solubility	Analyst

2.3.4 CQA – Critical Quality Attributes and CMA – Critical Method Attributes

A physical, chemical, biological, or microbiological property or characteristic that should be within an appropriate limit, range, or distribution to ensure the desired product quality. (Description for CQA) [2]

The critical quality attributes represent the quality of the product and their analytical method lifecycle counterpart are the CMAs, which portray the quality of the results of the test procedure.
They are defined in the ATP together with their limits.
In chromatography they include the TMU including the accuracy, precision, specificity, range, analysis time and many more.

2.3.5 PMV – Potential Method Variables

They include all variables of an analytical method regardless of their influence on the CMAs and are further split into variables with low or high impact, on the output. Often variables with a high impact are listed as medium impact if they can be adequately controlled. The classification of the variables starts through prior knowledge and literature and is later finalized and adjusted with experiments.

2.3.6 CPP – Critical Process Parameter and CMV – Critical Method Variables [2]

A process parameter whose variability has an impact on a CQA and therefore should be monitored or controlled to ensure the process produces the desired quality. (Description for CPP) [2]

For an analytical method lifecycle, the term critical method variable CMV is used, to describe the influence of the variation of method variables on the CMA. It is the same as critical process parameter in the Quality by Design, which describe the influence of variation on CQAs.
They contain method parameters, which vary for every analytical technique.
Possible examples for HPLC are mobile phase buffer, pH, eluent, column, organic modifier, gradient/isocratic operation, temperature.
CMVs are all variables, whose variability have a significant influence on CMAs.

Their impact depends on which analytical technique, for example RP-HPLC, is chosen, as well as on sample characteristics and therefore can only be finalized after experimental studies.
These are also the variables which are thoroughly tested in the method design and development.

2.3.7 DoE – Design of Experiments

It is a systematic method to determine the relationships between variables affecting a process, and it is used to find cause-and-effect relationships. [5]

DoE is a tool used for method development, optimisation, and robustness studies.
It builds on the literature and prior knowledge of PMVs and their possible influence on the method for a first assessment and then expands the knowledge and information further, for a more situation related classification into high and low impact.
DoE is used to find an optimum working range, which is at best also robust, of the examined method variables and to understand the impact of variations of CMV's.
The difference to the traditional approach is that the factors are not investigated one at a time, but more than one parameter can be changed at a time and therefore their combined influence on CMAs is observed, too.

2.3.8 MODR – Method Operable Design Region and DS – Design Space

MODR is the result from DoE and the analytical counterpart to the design space. It describes a region or area in which small variations of parameters still lead to robust results without loss of performance.
For example, the MODR states whether the results of a method obtained at 37°C instead of 40°C still fulfil the ATP criteria.
This leads to a wider operable robust area in which adjustments in method settings can be implemented easier, as mentioned in the ICH Q8 guideline, which describes the design space as:
"The multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality. Working within the design space is not considered as a change. Movement out of the design space is considered to be a change and would normally initiate a regulatory post approval change process." (Description of Design Space) [2]

2.3.9 (A)CS – (Analytical) Control Strategy [15]

A planned set of controls, derived from current product and process understanding, that assures process performance and product quality (Description of Control Strategy) [4]

The ACS is the equivalent to the Control Strategy in the analytical QbD approach. In this case the above-mentioned definition can be altered into:
The ACS is a planned set of controls derived from current analyte and method understanding, that assures method performance and the quality of the results.
In the Stimuli article *Analytical Control Strategy* [16] three distinct unit operations are mentioned:

- Sample Preparation
- Measurement
- Replicate Strategy

These three combined contribute to the Analytical Procedure Control Strategy.
The ACS builds on the knowledge of how CMVs affect the CMAs and how this influence can be controlled. The goal is, that there is little to almost no method related variation in the outcome of the routine use of the test procedure, so that the results always meet the ATP criteria.

2.3.10 CMM – Continuous Method Monitoring

Effective monitoring of an analytical procedure provides confidence that the reportable value generated is fit for purpose. This stage should include an ongoing program to collect and analyze data that relate to analytical procedure performance. [5]

CMM follows the method validation and ensures that the quality of results is maintained at an acceptable level, detects potential procedure performance issues and points to necessary changes to the analytical procedure.
The monitoring of the results is established through control charts or tracking systems and if they indicate a loss of control, an investigation to find the root cause is carried out.

2.3.11 RA – Risk Assessment

Risk assessment consists of the identification of hazards and the analysis and evaluation of risks associated with exposure to those hazards. [3]

Risk assessment is the general term for identifying, analysing, and evaluating risk. It starts with defining CMVs of a method based on prior knowledge, moves on to understanding their impact on CMAs and ends with evaluating the strength of their influence. This is followed by the risk mitigation in which actions are taken to minimize or control the impact.

Starting the RA, a precise problem description is needed, which is followed by choosing a suitable risk assessment tool. Three fundamental questions to identify the risk accurately are mentioned in the ICH guideline Q9 [3]:

1. *What might go wrong?*
2. *What is the likelihood (probability) it will go wrong?*
3. *What are the consequences (severity)?*

The **risk identification** answers the first and third question. It systematically uses information to identify potential variations in the results, to assess what might go wrong and the following consequences. [3]

Through **risk analysis** the risk from the potential hazards found in the identification step are estimated. It links the gravity of the harm with the probability it if happens. [3] Depending on the risk assessment tool the detectability is considered too.

In the **risk evaluation** the identified and analysed risks are compared against risk criteria. The weighting of all three questions mentioned above are considered. [3]

2.4 The Analytical method lifecycle

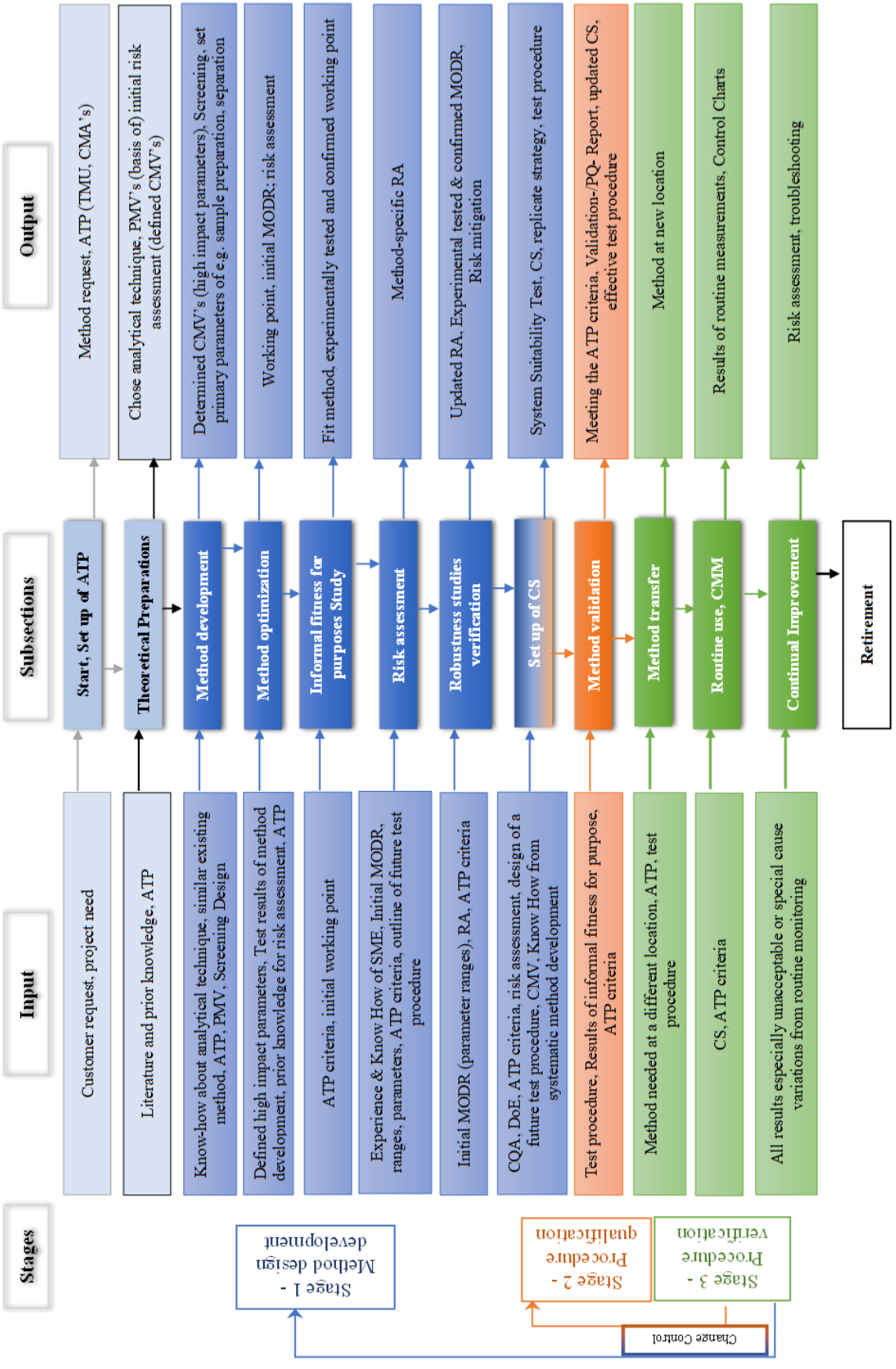

Figure 7: Overview of the stages and sub sections of the Analytical method lifecycle.

The flow chart in Figure 7 shows the three main stages *Method Design and Development, Procedure Qualification* and *Procedure Verification* of the analytical method lifecycle, as well as their subsections and inputs and outputs to them.

Stage 1, *Method Design and Development*, starts with the need for an analytical method, followed by a request, for which an ATP is defined.
For that, CMAs and their required limits can be set based on product requirements (specifications) and product strategy (achievement), product know-how, prior knowledge (e.g. of similar methods, similar cases), guidelines/literature.
This information, together with possible consideration of which materials need to be used and which by-products and other components will be present, lead to the selection of an analytical technique.
Because the knowledge gathered in prior projects can be a key part in choosing the right technique, it is important to have them on a good accessible database (knowledge management).

The next step focuses on PMVs and possible CMVs, either through a whole new assessment of the parameters or based on an already existing evaluation.
This initial examination of possible variables is not only important for the following screening but also for later risk assessment and DoE.
Firstly, the variables can be grouped in primary and secondary method variables.

The primary rated should be parameters with a higher impact on the method, which are in this example for a RP-HPLC pH-selection (acid, neutral, basic) and organic compound of the mobile phase and stationary phase.
They are experimentally tested in the screening of the development phase.
Secondary parameters like gradient slope, flowrate, temperature and pH (a smaller range around the set point) can be later investigated in the optimisation phase.
With the primary method variables, a rough working point can be chosen and then through a more accurate examination of the secondary parameters in the optimisation step, the set point can be confirmed.
This can be achieved through a DoE, which in this case (for a RP-HPLC method) was generated with the software DryLab.
It was used to calculate a set point with only few systematic experiments. This way the selectivity was examined, but also other attributes can be optimised.

After confirming it with an experimental run and comparison of the chromatograms, the informal fitness for purpose study can demonstrate that the ATP criteria are met. This should later ensure a successful method validation in stage 2 – procedure qualification.

During the following risk assessment, the influence of the variation of variables on the results of the method are examined. The aim is to understand and control the risk.
This step is divided into a first (theoretical) and a second (experimental) assessment.
Firstly, method variables are assessed theoretically, which for example can be achieved through prior knowledge, literature, know-how, information from prior results in the development and optimisation phase or an already roughly set guideline like a checklist.
Only the method variables, which are rated possible critical are then experimentally examined in the second assessment.
In this stage the variables are tested through a DoE, where the output is not only the understanding of critical method variables but also a robust area (MODR). Depending on the chosen design a defined set of experiments must be carried out and then a robust area can be calculated.
The output of the DoE is also important for a final assessment of the parameters.
A helpful tool for an accurate classification of parameters is a traffic light risk assessment, where variables are coloured based on if they are high (red), middle (yellow) or low (green) risk variables.
A more precise but therefore also more complicated tool is the Failure Mode and Effect Analysis (FMEA), where severity but also occurrence and detectability are considered more detailed.
Understanding the influence of the variability of method variables is an important step to mitigate and to be prepared for method failures.

All these steps, defining ATP, method development and optimisation, informal fitness for purposes study, risk assessment and robustness study, form a cycle. For example, if the method does not seem fit for purpose, more testing in the method development stage may generate suitable results.

The last step in the first stage is the establishment of a control strategy CS, which shows, when the ongoing performance of the method does not meet the set criteria.

With all the gathered data, including ATP criteria, information about the CMVs, the output of the risk assessment and the design of a future test procedure, the Control Strategy is developed.
Included in the CS can be a replicate strategy, a thoughtful selection of method variables and possible ways to make them less critical, as well as system suitability tests.

Stage 2 – the procedure qualification or validation builds on the test procedure established in Stage 1 and on the results of the informal fitness for purpose studies, as well as on the ATP criteria.
The method is tested if it meets the ATP criteria. If that is the case, it is confirmed in a validation or process qualification report.
If not, either the development and optimisation steps must be repeated and tested more thoroughly, or the ATP criteria or CS must be adjusted.

After a successful validation Stage 3 – the procedure verification follows. If the technique is not developed at the same location, where it will be used for routine measurements, this stage starts with a method transfer.
In the draft about the Analytical Lifecycle of USP from 2020 [1] the method transfer is the last step of stage 2, but it is more a formal question to which stage this step belongs and it could change until a final guideline is set.
Through the routine measurements the method generates a lot of results, which should be almost all replicable and robust.
To monitor them continuously control charts can be used. In few cases, where the results have a significant deviation and do not meet the ATP criteria, they are examined through troubleshooting tools mostly with the help of the knowledge gathered in the risk assessment in stage 1.
This circle of control and required adjustments leads to a continual improvement and forms the core of the analytical method lifecycle.
The analytical method lifecycle comes to an end, when the production of the analyte is ceased. Then the method is retired but can still be used as a base or for know-how for new methods.

3 Procedure

The goal of this thesis was to produce a guideline for developing an RP-HPLC method based on the Analytical Quality by Design approach.
This chapter contains the experiments made for this thesis, for an example of a method development based on the through this thesis composed guideline. The improvements are listed later in section 4.11.

A flow chart for a better understanding and reproduction can be found in Figure 76 to Figure 78 in the appendix.

Here are also measurements of dwell volume, extra column volume, dead volume, measurement of the systems back pressure and establishment of a system suitability test explained, which are mostly found in the manuals of the instrument or columns and can be measured for a new system to underline the given information. The experiments of this thesis were conducted on an UPLC, which had not been used regularly, so the above-mentioned volumes were tested to check the system suitability. This section also describes the set-up of the UPLC used for this thesis, all the used materials, as well as a description of sample preparation followed by the selection of the mobile and stationary phases and the screening of these conditions.
At the end of this section the implementation of optimising a method and robustness studies based on the Design of Experiment in DryLab and Design Expert 11 are explained.

3.1 Characterisation of UPLC system

The used device was an H-Class U(H)PLC from Waters with a Photodiode array (PDA) detector. The instrument was configured for four columns with preheaters.
The appearance of the H-Class is shown in Figure 8 and a detailed representation of the flow inside the HPLC system is shown in Figure 9.
The H-Class consist of a Quaternary solvent manager (QSM), a sample manager with a flow through needle (SM-FTN) design, two column heaters, a detector, and a bottle tray on top.

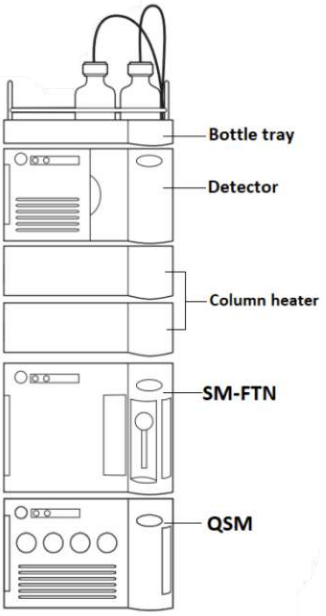

Figure 8: H-Class UPLC system configuration [17]

The QSM contains a degasser, a gradient proportional valve a primary and accumulator pump head, a vent valve, and a mixer. The SM-FTN consists of a sample tray, sample syringe, sample syringe valve and the injection valve. A particularity of the set up was a valve for position D before the degasser, where six different eluents could be selected.

For a more precise information of all the parts see the handbooks of Waters [17] [18] [19].

The column heaters have a switching valve before and after the column positions, with each heater having positions for two columns.

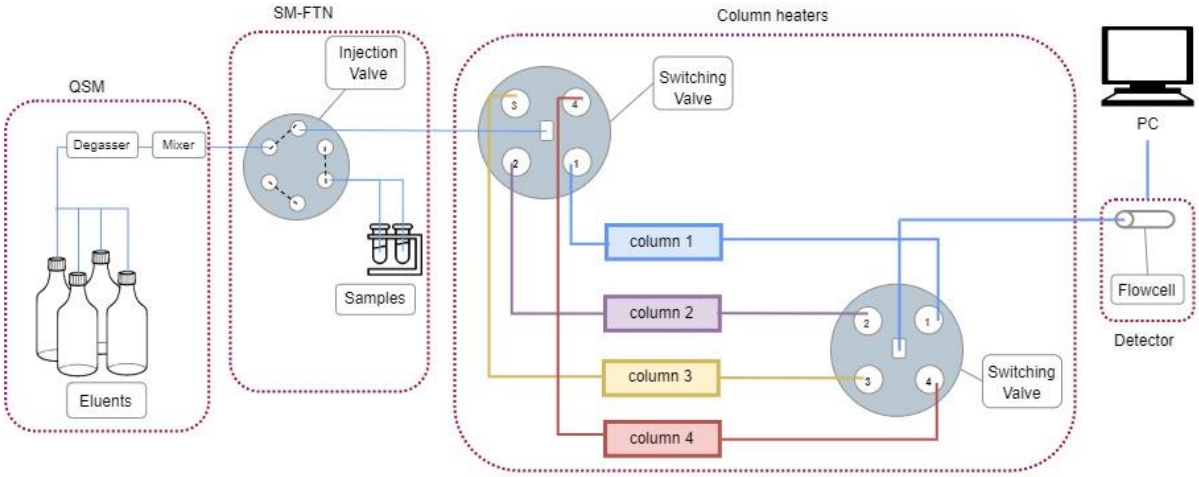

Figure 9: System set up of the U(H)PLC used at this thesis.

3.2 Measuring system and back pressure

The UPLC system operates at back pressure up to 1034 bar.

To exclude any problems of overpressure for the columns, because some of the used ones had only a pressure tolerance of 600 bar while running the screening with a flowrate of 1 mL min^{-1}, all the columns were tested with a gradient of 40-60 % methanol and water at 40 °C. Before that, the pressure was measured without column, to understand how high the pressure on the system was and therefore to have an accurate value of the actual pressure drop caused by the column.

3.3 Measuring dwell, extra column and void volume

Before developing a method on an HPLC, it is important to know if the system is working on a proper level. Therefore, the dwell and extra column volume can be tested. This step is to characterise the HPLC system and does not have to be repeated for every new method development. These volumes are mostly given in the manufacturer certificate. Only in special cases it is tested as it was for this thesis.

3.3.1 Extra column volume

The extra column volume *ECV* is the sum of all volumes from the injection point to the detection, excluding the column. So, it was calculated with formula (7). All volumes contributing to the ECV are also shown in Figure 10.

$$ECV = V_{tubing} + V_{Inj.Valve} + V_{sw,valves} + V_{detector} + \cdots .. \quad (7)$$

V_{tubing}... Volume of all tubings in the system (from autosampler, to column etc.)

$V_{Inj.Valve}$ Volume of the injector valve

$V_{sw.valves}$ Volume of the switching valves before and after the columns (only by configurations with more than one column)

$V_{detector}$. Volume of the detector cell

The *Volume V* of the tubings was calculated with equation (8).

$$V = l \cdot \pi \cdot r^2 \quad (8)$$

l..... length of tubing

r ..... radius of tubing

As the information of the volumes of valves, autosampler and other parts, which have a complicated geometrical structure, are not always precise or even known, the ECV was determined experimentally too, to verify the calculated value.

For this purpose, the column was substituted by a zero dead volume union (connector), then 1 µL pure acetonitrile was injected.

The run was isocratic with 50 % acetonitrile in water as the eluent, a temperature of 40 °C, a flowrate of 0.6 mL min^{-1} and the measurement took place at 200 nm.

The retention time or in this case time delay t_R of the measured peak was then multiplied with the flowrate *F* to get the ECV (see formula (9)).

$$ECV = F \times t_R \quad (9)$$

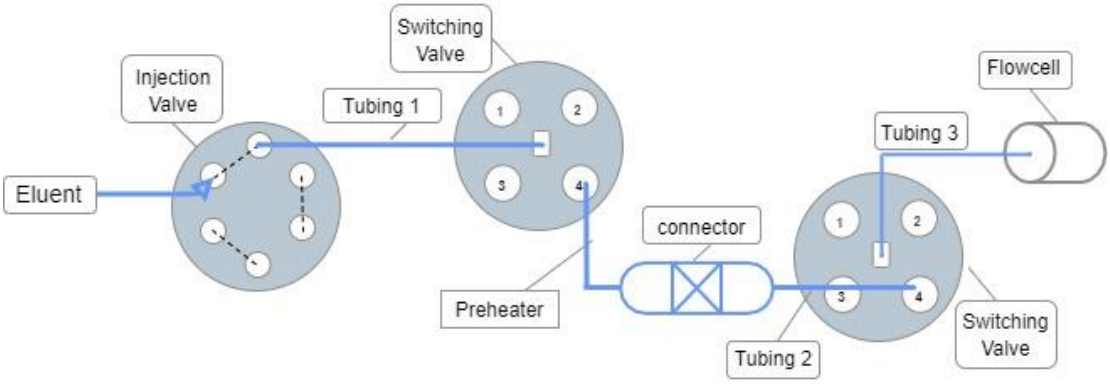

Figure 10: Tubings and valves set up for ECV measurement for the U(H)PLC set up used in this thesis.

3.3.2 Dwell volume or gradient delay volume

For measuring the dwell volume two eluents were needed, with the second containing an UV-active compound and then a step from 0-100 % of B in from 3 to 3.1 min was programmed.

Both mobile phases were 50 % acetonitrile in water, with eluent B spiked with 0.1 % acetone.

The other measurement parameters were a temperature of 40 °C, a wavelength of 265 nm and a flowrate of 1 mL min^{-1}.

Then the dwell volume V_D was calculated by multiplying the dwell time t_D with the flowrate F (see formula (10)).

$$V_D = F \cdot t_D \tag{10}$$

The dwell time is the time $t_{\frac{1}{2}}$ the system reached half of the gradient change minus half of the actual programmed time t_G of the step in the gradient at 3.05 min. (see formula (11)).

$$t_D = t_{\frac{1}{2}} - t_G \tag{11}$$

The middle between the start and end intensity lines, which are pictured in red in Figure 11 is $h_{\frac{1}{2}}$, from which $t_{\frac{1}{2}}$ can be read off this spectrum.

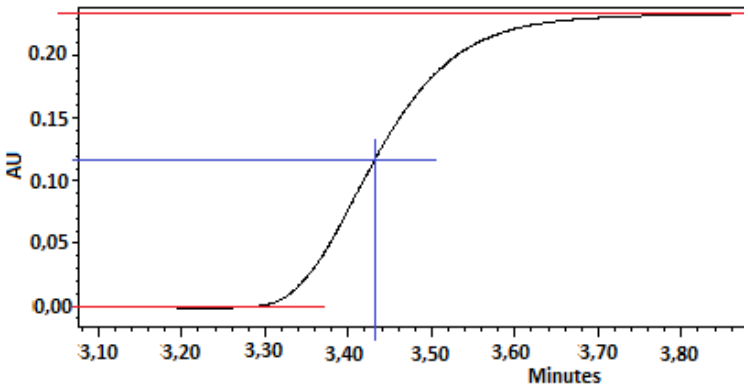

Figure 11: Chromatogram for evaluation of dwell volume of the U(H)PLC set up used in this thesis.

3.3.3 Void volume

The void volume, also called dead volume, is the time, a component with no interaction with the stationary phase of a column needs to pass the UPLC system.

The conditions of the testing are listed in the following section 3.4, as the substance for the void volume, was added to the SST sample.

For the measurement 0.1 mg Uracil were used.

3.4 System suitability test

To assure that the system is suitable the plate counts of the columns were measured with ethylbenzene and toluene.

The sample consisted of 100 µL of each of ethylbenzene and toluene and 0.1 mg uracil, as mentioned above for the void volume measurement. The compounds were dissolved in 20 mL of 30 % acetonitrile in water.

The method parameters are listed in Table 2.

Table 2: Method conditions for the system suitability test run.
The SPP columns have a bigger diameter (3 mm) than the TPP columns (2.1 mm), which is why the flow rates were adjusted appropriately.

Column type	SPP	TPP
Flow rate (mL min^{-1})	1	0.5
Injection volume (µL)	0.3	
Eluent A	0.1 %aqueous phosphoric acid	
Eluent B	acetonitrile	
Isocratic	40 % Eluent A; 60 % Eluent B	
Run time (min)	4	
Detection Wavelength (nm)	250	
Column temperature (°C)	23	

As the TPP columns had smaller particle sizes and therefore a higher back pressure, the flow rate was lowered for their measurement.

3.5 Stationary phases

For this thesis several columns with different properties were used. Depending on manufacturer there were differences in the dimensions and also all columns were selected to have different chemical properties. For a better overview, all the columns used throughout experiments for this thesis and their properties are summarised in the following Table 3.

.

Table 3: SPP and TPP columns used in this thesis and their dimensional, application limits, as well as chemical properties based on the manufacturer websites of the columns. [20] [21] [22] [23] [24]

Column name	Particle type	Particle size	Dimensions length x i.d. (mm)	Max. pressure (bar)	Max. Temperature (°C)	pH Range	Carbon load (%)	Surface modification	Specific surface (m^2 g^{-1})	Endcapped
Waters Cortecs C18+	SPP	2.7	100 x 3	1000	45	2-8	5.7	Positive charge	100	yes
Agilent Poroshell 120 EC-C18	SPP	2.7	100 x 3	1000	60	2-8	10	none	130	double
Agilent Poroshell 120 Bonus RP	SPP	2.7	100 x 3	600	60	2-9	9.5	none	130	triple
Phenomenex Kinetex C18 EVO	SPP	2.6	100 x 3	600	60	1-12	11	organo-silica grafting	200	no
Thermo Scientific Accucore Phenyl-X	SPP	2.6	100 x 3	1000	70	2-8	6	phenyl-X alkyl aromatic bonded	130	yes
AMT Halo PFP	SPP	2.7	100 x 3	600	60	2-9	5.5	Pentafluorophenyl-propylsilane	135	yes
Waters XSelect HSS PFP	TPP	1.8	100 x 2.1	1240	45	2-8	7	Pentafluorophenyl-propylsilane	230	no
Waters BEH Phenyl	TPP	1.7	100 x 2.1	1200	low pH: 80°C high pH: 60°C	1-12	15		185	yes

3.6 Loading capacity

To choose the right type of stationary phase, the loading capacity for bases was tested on different column types (see Table 3).

This is particularly important for impurity testing, because the concentrations of the sample contaminations need to be high enough for the detection limit. Consequently, the loading of the main peak is very high, because of the expected high concentration difference between main substance and contaminations.

In case of basic main components, which are charged under acidic conditions, a high loading can lead to a peak shape distortion. This can lead to a deterioration of the resolution for adjacent contamination peaks and therefore loss of selectivity.

Typical building blocks of pharmaceutical substances are often basic, which is why this consideration is particularly important for pharmaceutics methods.

For the loading capacity tests, the following columns were used:

- Waters Cortecs C18+
- Agilent Poroshell 120 EC-C18
- Agilent Poroshell 120 Bonus RP
- Phenomenex Kinetex C18 EVO
- Waters XSelect HSS PFP
- Waters BEH Phenyl

A solution of *1-(α-Aminobenzyl)-2-naphthol hydrochloride* dissolved in 1.5 % acetonitrile in water, with a concentration of 0.1503 g L^{-1} was injected with three different volumes, which different for the column types because of their dimensions, see section 3.5 to see if there is a shift of the retention time and if a higher concentration affects the peak symmetry significantly.

The conditions for these runs are listed in Table 4. Some values are varying because of the different column inner diameters of the TPP and SPP columns. The adjustment of the conditions according to the different cross-sectional area was made for a valid comparison between the columns.

Table 4: Conditions for the loading capacity runs, with adjusted parameters because of the different inner diameters of SPP (i.d. 3 mm) and TPP (i.d. 2.1 mm) columns.

Column Type	SPP	TPP
Flow rate (mL min^{-1})	1	0.5
Injection Volume - high (µL)	2	1
Inj. V. - medium (µL)	1	0.5
Inj. V. - low (µL)	0.5	0.3
Column temperature (°C)	40	
Detection Wavelength (nm)	226	
Sampling rate (points s^{-1})	20	
Run time (min)	21	
Gradient	**t (min)**	**% Eluent B**
	Initial	5
	15	95
	18	95
	18.01	5
	21	5

3.7 Analytical Target Profile

One of the foundations of the Analytical Quality by design approach is the theoretical preparation and processing of prior information and data.

Therefore, the targeted outcomes are considered thoroughly before starting with the method development.

The main goal was to analyse the purity of the substance *1-(α-Aminobenzyl)-2-naphthol hydrochloride*. Therefore, a maximum number of contaminations (peaks beside the main peak), with a minimum percentage area of 0.05 % from the main peak was searched for.

The values for the defined CMAs are listed in Table 5.

Table 5: Method Attributes listed for the establishment of the ATP and their values reporting threshold and % area.

Min. % area	0.05
RSD of %area for main peak	≤ 1 %
RSD for secondary components with an area smaller 0.5 % compared to the main peak	≤ 20 %
RSD for secondary components with an area between 1 and 0.5 % compared to the main peak	≤ 10 %
RSD for secondary components with an area larger 1 % compared to the main peak	≤ 5 %
Resolution of two peaks	> 2

3.8 Selection of the stationary phases

For the following screening, the listed columns had been chosen, based on the F-value discussed in section 2.1.2.2. Their limits of application according to the manufacturers, dimensions and chemical properties can be found in section 3.5. Also, based on expert statements of the manufacturers, the Phenyl- and PFP phases were only used with Methanol and not acetonitrile for elution.

- Waters Cortecs C18+
- Agilent Poroshell 120 EC-C18
- Agilent Poroshell 120 Bonus RP
- Phenomenex Kinetex C18 EVO
- Thermo Scientific Accucore Phenyl-X
- AMT Halo PFP

3.9 Selection of the mobile phases

For the screening, the organic eluents methanol and acetonitrile were chosen.
The selected aqueous eluents are 0.1 % phosphoric acid in water for acidic conditions, 4 mM ammonium hydrogen phosphate with a neutral pH-value and 10 mM ammonium carbonate for basic conditions.
The reason for the low concentration of ammonium hydrogen phosphate is that at higher concentration it precipitates, when the percentage of acetonitrile in the mobile phase reaches a higher amount than 85 %.

3.10 Sample preparation

The criteria for the chosen sample were a similar structure to typical building blocks of pharmaceutical products and that it was a catalogue product, to have no secrecy problems when publishing this thesis.
The *1-(α-aminobenzyl)-2-naphthole hydrochloride* is a product from Sigma-Aldrich and its specifications are listed in Table 6 and its structure is pictured in Figure 12.

Table 6: Specifications of the sample used for this thesis *1-(α-aminobenzyl)-2-naphthole hydrochloride*.

Product Number	Sigma Aldrich 732400
CAS-Number	219897-32-2
Formula	$C_{17}H_{16}ClNO$
Formula weight	285.77 g mol^{-1}
Purity	≥ 96.5 %
Melting point	195-202 °C

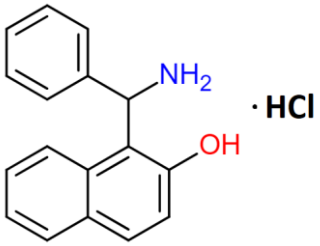

Figure 12: Structure of the main component *1-(α-aminobenzyl)-2-naphthole hydrochloride* of the sample used for this thesis.

As the sample was not wholly solvable in the eluent, solutions with different concentrations of acetonitrile in water were prepared and tested with different sample concentration.

The amount of *1-(α-aminobenzyl)-2-naphthole hydrochloride* was varied to examine the sample amount in which the peak of the main component and the peaks of the contaminations could be determined most precise.

The result was a sample solution of 3 mg 1-*(α-aminobenzyl)-2-naphthole hydrochloride* dissolved in 20 mL of 1.5 % acetonitrile in water.

3.11 Initial run

Before the screening started, an initial run was made with the prepared sample solutions mentioned above, to set a reasonable concentration and to find a fitting wavelength adjusted to detect all impurities.

The sample was measured with the following conditions (Table 7).

Table 7: Conditions of the initial run with the Waters Cortecs C18+ to set a reasonable concentration for main peak and impurities and to find a fitting wavelength adjusted to detect the impurities together with the main component.

Flow rate (mL min^{-1})	1	
Inj. V. (µL)	2	
Column temperature (°C)	40	
Detection Wavelength (nm)	200-400	
Sampling rate (points s^{-1})	20	
Run time (min)	21	
Eluent A	0.1 % H_3PO_4 in water	
Eluent B	acetonitrile	
Gradient	**t (min)**	**% Eluent B**
	Initial	5
	15	95
	18	95
	18.01	5
	21	5

3.12 Screening

The following conditions, listed in Table 8, were used for the screening and the combination of the mobile and stationary phases mentioned in section 3.8 are shown in Table 9.

Table 8: Conditions for the screening runs.

Flow rate(mL min^{-1})	1	
Inj. V. (µL)	2	
Column temperature (°C)	40	
Detection wavelength (nm)	190-400 spectrum with an additional channel at 226	
Sampling rate (points s^{-1})	20	
Run time (min)	21	
Gradient	**t (min)**	**% Eluent B**
	Initial	5
	15	95
	18	95
	18.01	5
	21	5

Table 9: Combination of mobile phases and columns for the screening runs.

Column	approx. pH 2	pH 6.7	pH 10	MeOH	ACN
Waters Cortecs C18+	x	x		x	x
Agilent Poroshell 120 EC-C18	x	x		x	x
Agilent Poroshell 120 Bonus RP	x	x		x	x
Phenomenex Kinetex C18 EVO			x	x	x
Thermo Scientific Accucore Phenyl-X	x	x		x	
AMT Halo PFP	x	x		x	

3.13 Optimisation with DryLab

In this step the gradient time, the column temperature, and the pH-value of the chosen conditions of the screening were optimised.

Therefore, three eluents of 0.1 % phosphoric acid with pH-values of 2; 2.4 and 2.8 were prepared. The pH-value was set with ammonium hydroxide.

For an U(H)PLC a gradient time of 55 min is very long, but the DryLab guideline recommended a four- to fivefold difference between the time set points, so the long gradient time was chosen.

In the following Table 10 are the conditions listed from DryLab for the 12 optimisation runs. All the other conditions matched the screening (see Table 8).

Table 10: Conditions of the optimisation runs, with a variation of the parameters based on a Drylab proposed design, where the gradient times were suggested to vary at least 40 minutes, the temperatures more than 30 °C and the pH-values approx. 0.4.

Run	Gradient time (min) (5 % to 95 % B)	Column temperature (°C)	pH-value
1	15	25	2.06
2	55	25	2.06
3	15	55	2.06
4	55	55	2.06
5	15	25	2.4
6	55	25	2.4
7	15	55	2.4
8	55	55	2.4
9	15	25	2.8
10	55	25	2.8
11	15	55	2.8
12	55	55	2.8

3.14 Robustness studies

To start the robustness studies an initial risk assessment was made, to categorize PMV into potentially and not critical parameters.

With this assessment a definitive screen design was made in Design Expert 11.

The eluents 0.09; 0.1 and 0.11 % phosphoric acids were prepared for the testing and the variation of the parameters are listed in Table 11.

Table 11: Conditions of the robustness studies proposed by the chosen design in Design Expert for five variables.

Run	Gradient slope (%B min^{-1})	Flow rate (mL min^{-1})	Detection Wavelength (nm)	Column temperature (°C)	Buffer concentration (%)
1	5.6	1	225	38	0.11
2	5.6	0.9	227	38	0.09
3	5.06	1.1	225	32	0.11
4	5.33	1	226	35	0.1
5	5.6	1.1	225	35	0.09
6	5.6	1.1	227	32	0.1
7	5.33	0.9	225	32	0.09
8	5.6	0.9	226	32	0.11
9	5.06	0.9	225	38	0.1
10	5.33	1.1	227	38	0.11
11	5.06	1	227	32	0.09
12	5.06	1.1	226	38	0.09
13	5.06	0.9	227	35	0.11
Centre point	**5.33**	**1**	**226**	**35**	**0.1**

3.14.1 Processing data in DesignExpert 11

Following is a description of processing outputs in the software DesignExpert 11, which was used for all the results.

Firstly, the type of design must be chosen, in this case a *definitive screen design*. The next step is entering the variables, see Figure 13, where the difference between categoric and numeric factors must be considered. Numeric factors are parameters with a value, while categoric factors have a condition, for example a column lot number is either A or B, but a temperature can be everything between 30 and 50 °C.

Block splits the test series, for example when you want the experiments made on two different HPLCs.

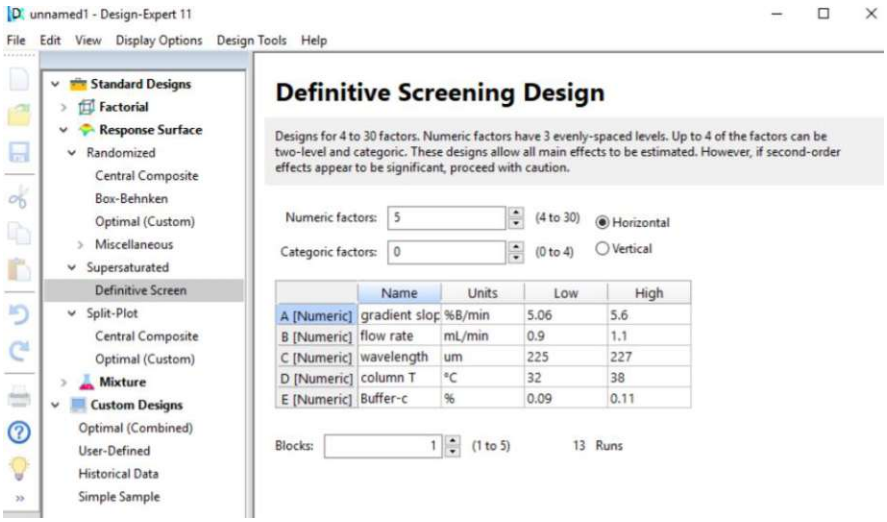

Figure 13: Entering parameters in the definitive screen design in Design Expert.

It is followed by the entry of the responses shown in Figure 14.

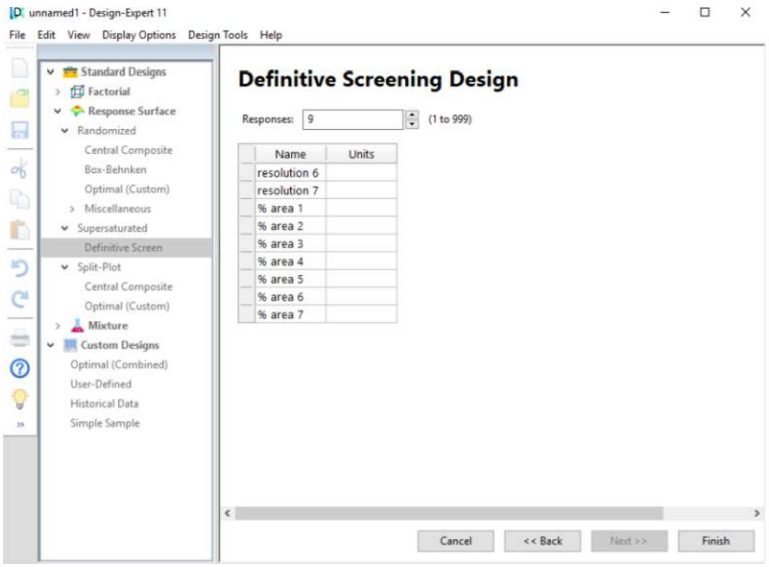

Figure 14: Entering responses in the definitive screen design.

A table with different method conditions is generated by the software. After testing these experimentally and transferring the gathered data, the processing of the individual output must be started.

In the navigation pane under *Analysis* the first output must be selected. Then the window *Transform* pops up, see Figure 15. Normally no transformation of data is needed, except for the examples further below.

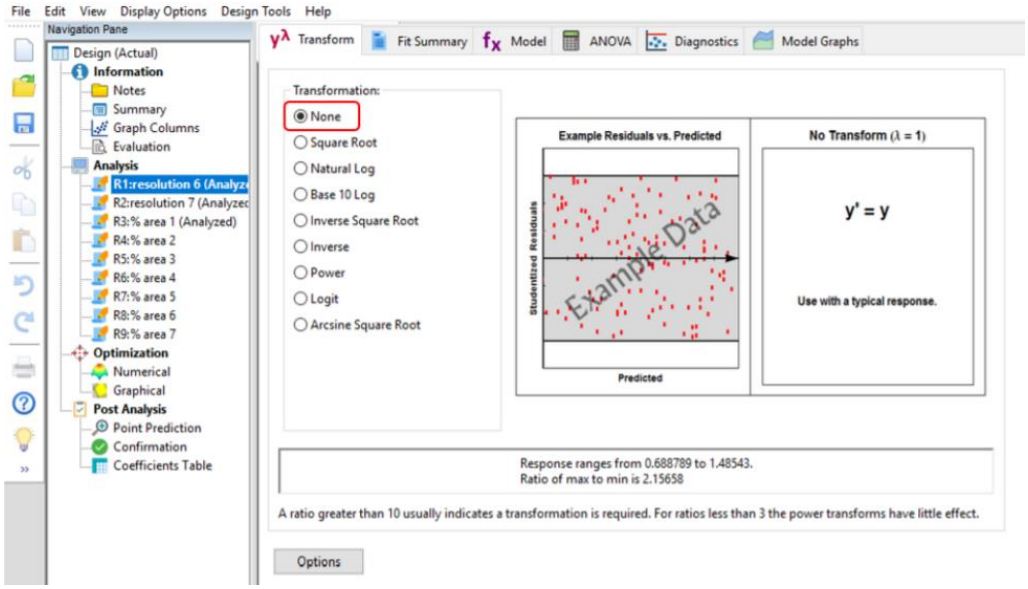

Figure 15: First window (Transform) for processing the data in DesignExpert 11.

The next window is *Fit Summary*. Here, statistical information of various possible models can be extracted. The table in Figure 16 on the left is particularly important because it suggests a model for the evaluation (printed in bold).

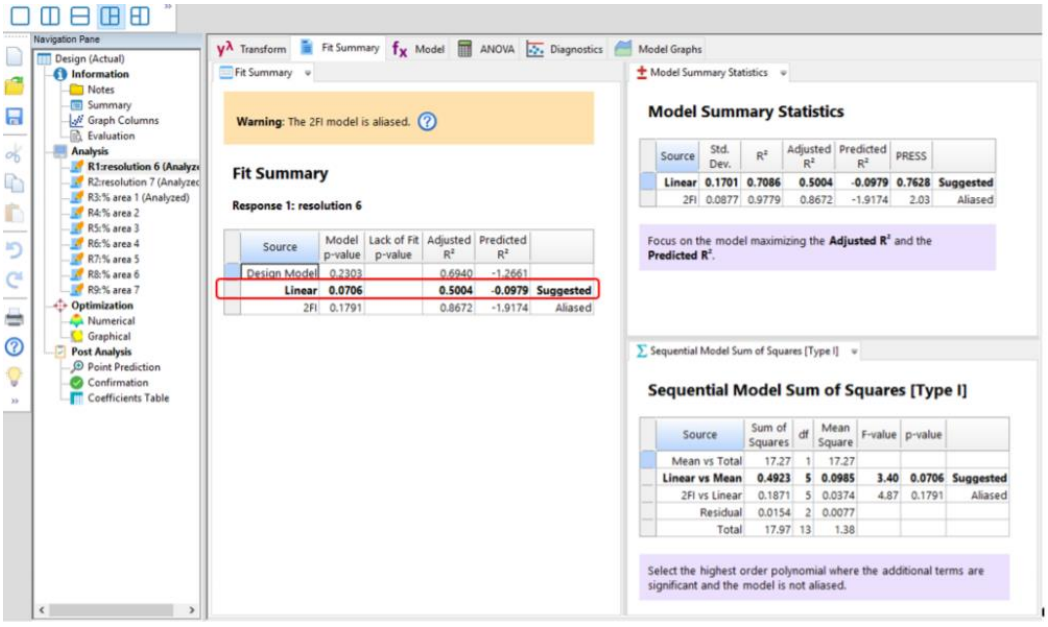

Figure 16: Fit Summary of the response and suggested model type.

This model must be selected in the following window as shown in Figure 17.

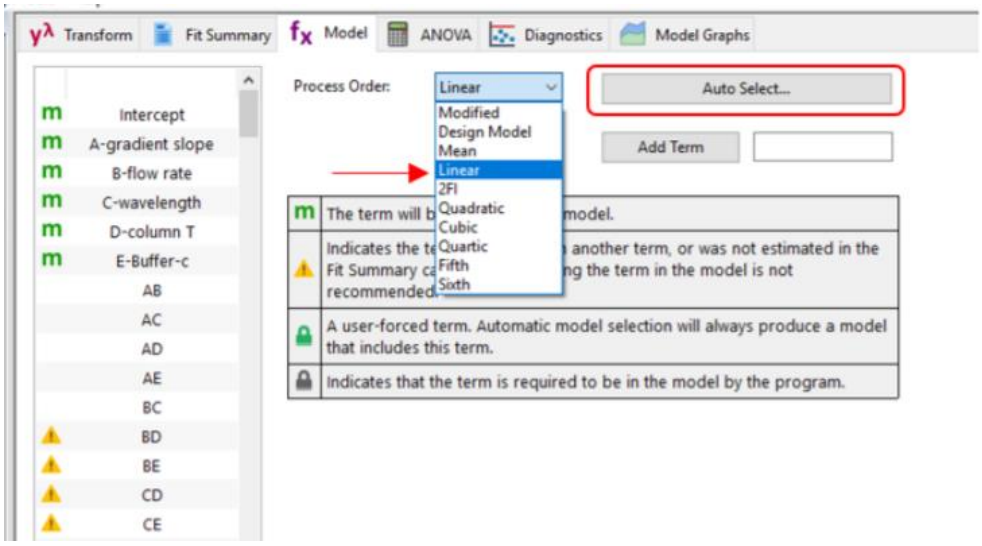

Figure 17: Selecting the right model.

Under *ANOVA* the statistical data is listed. Parameters with a p-value smaller than 0.05 are statistically significant and very likely critical, see Figure 18.

If the model is chosen correctly, it is marked as *significant*. If the table states *not significant* there can be two possibilities. Either the model is not appropriate, and another must be chosen, or all parameters are not significant and therefore none of them have a critical influence on the output.

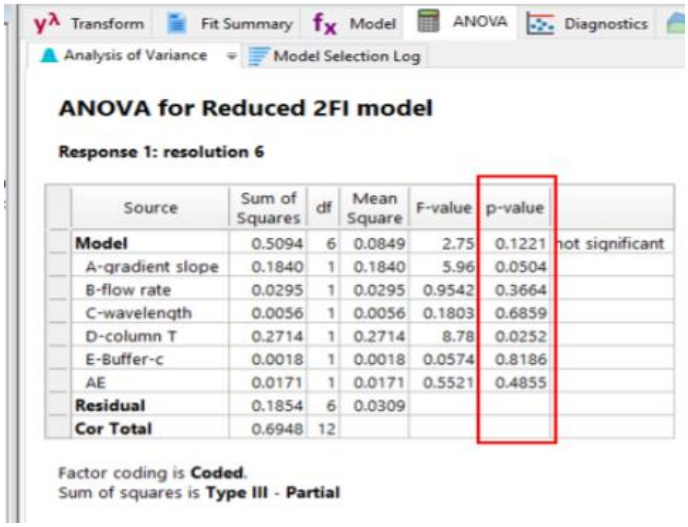

Figure 18: Listed p-values in the window “ANOVA”.

Under “Diagnostics” the ANOVA data must be verified. The following Figure 19 to Figure 26 compare the desired data distribution with undesired examples.

In the *normal plot,* the data should scatter around a straight line as shown in Figure 19.

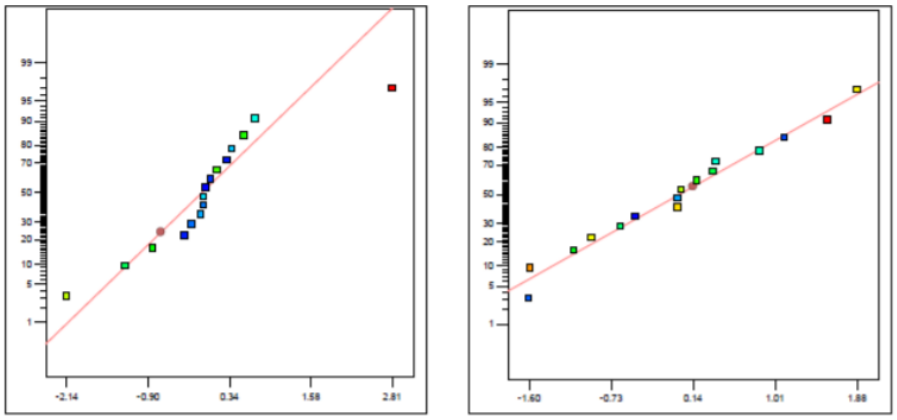

Figure 19: On the left the data is S – shaped, which is undesirable. On the right it is around a straight line as desired.

In the Plot *residuals vs. predicted*, the data should be scattered randomly.

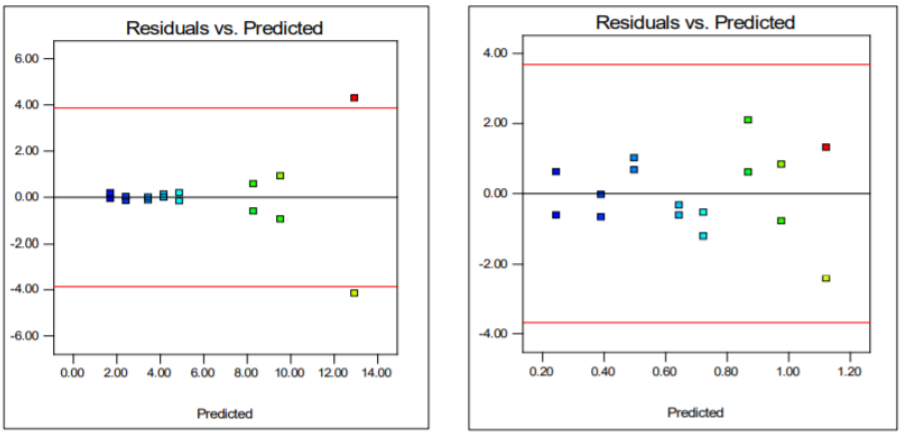

Figure 20: On the left the data is megaphone shaped, which is undesirable. On the right it is randomly scattered as desired.

Also, in the Plot *residuals vs. run*, the data should be scattered randomly too, as seen in Figure 21. Here, it is important to analyse data points outside of the red lines, which mark the statistical range in which the data points should be, (see Figure 21 to Figure 23) to detect possible problems or missing transformations.

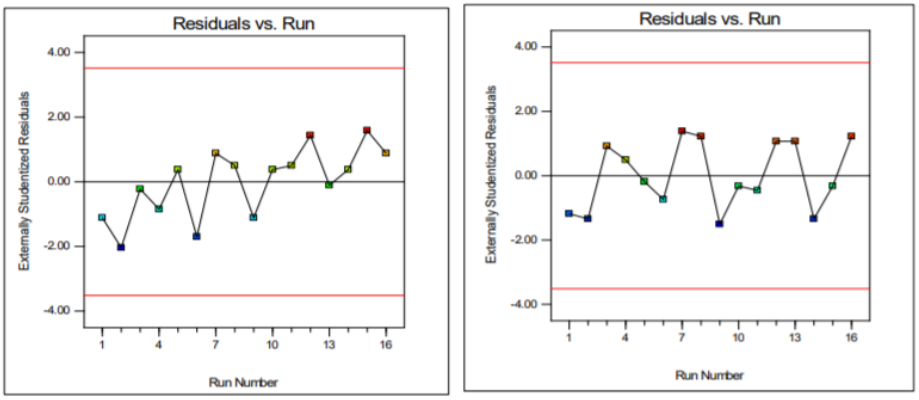

Figure 21: On the left the data shows an undesirable ascent. On the right it is randomly scattered as desired.

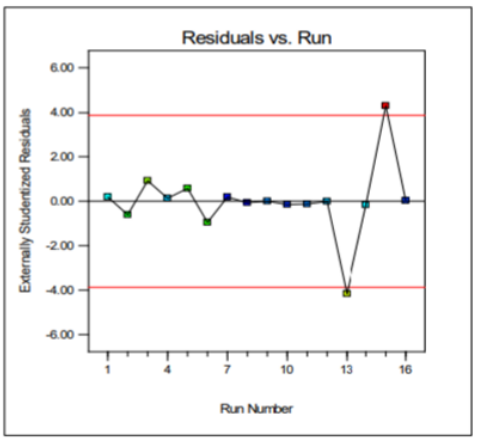

Figure 22: Results with an extreme deviation.

Another important plot to look at is the *predicted vs. actual*, where the data should be scattered around a 45-degree line, as seen on the right of Figure 23.

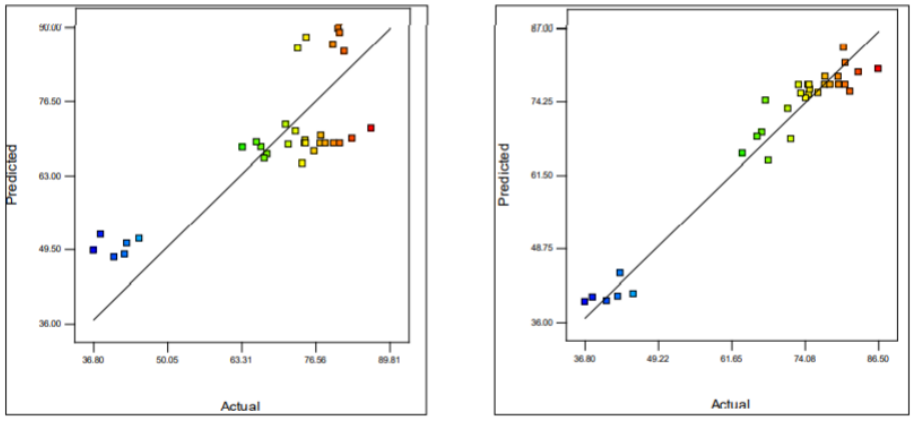

Figure 23: The left shows a bad prediction, on the right the desired scattering around a 45-degree line is displayed.

The *Box-Cox* plot indicates if a transformation is needed after all.

The blue line in the Figure 24 points to the current transformation (Lambda being one means no transformation) and the green line stands for the best value.

The red lines enclose a 95 % confidence interval around the green one.

It is recommended to use the standard transformation nearest to the calculated best lambda, except if one is included in the confidence interval, then no transformation is needed.

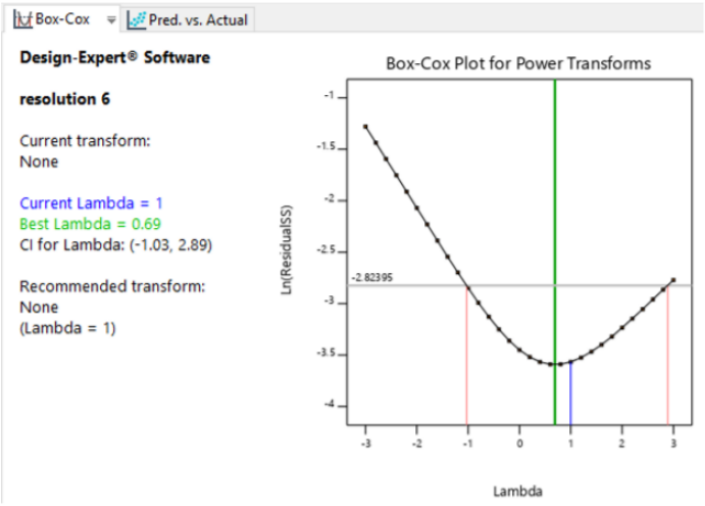

Figure 24: A Box-Cox plot, which indicates that no transformation is needed.

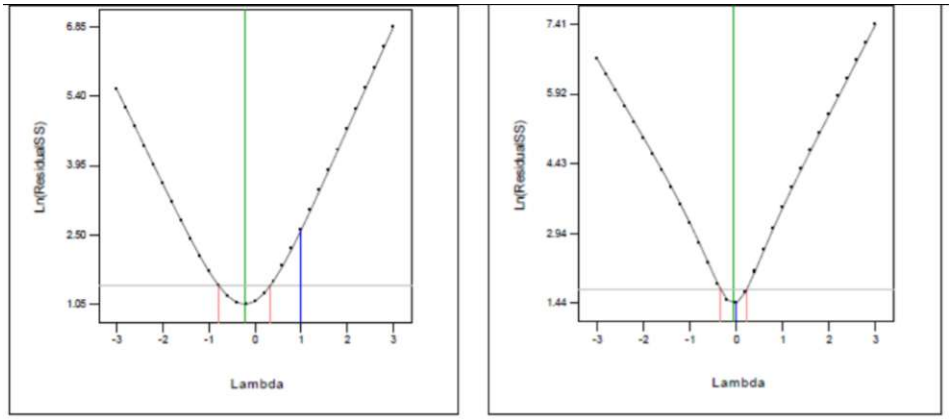

Figure 25: On the left is an example for a needed transformation, while the right shows the plot after the transformation.

The data in the *section residuals vs. factor* must be equally scattered over and under the zero line at the two end ranges.

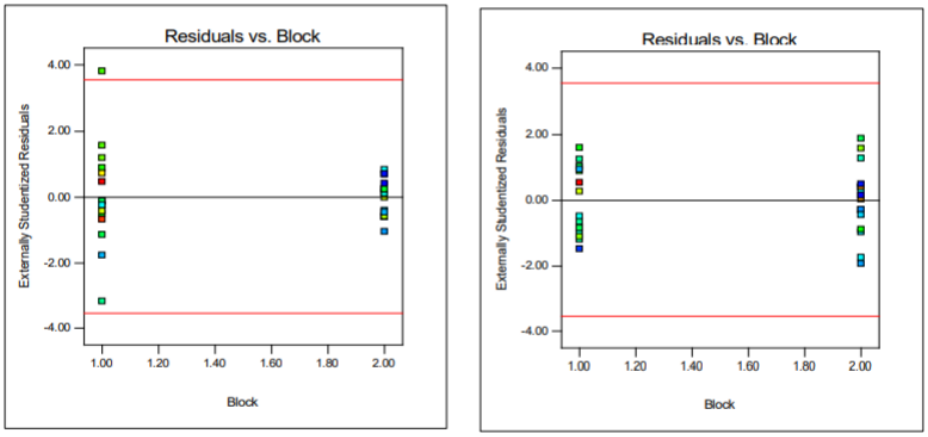

Figure 26: On the left plot, there is more variations between the scattering, while the left shows an equal distribution.

Beside the already mentioned ones, there are a lot more plots, which can be used for analysing and optimising the data, which should be used based on the individual problem solving.

After the data is processed, the influence of the parameters on the responses can be examined. In the window *Model Graphs*, under *Contour*, the value of the parameters can be changed and the possible influence of this on the response is then shown (see Figure 27)

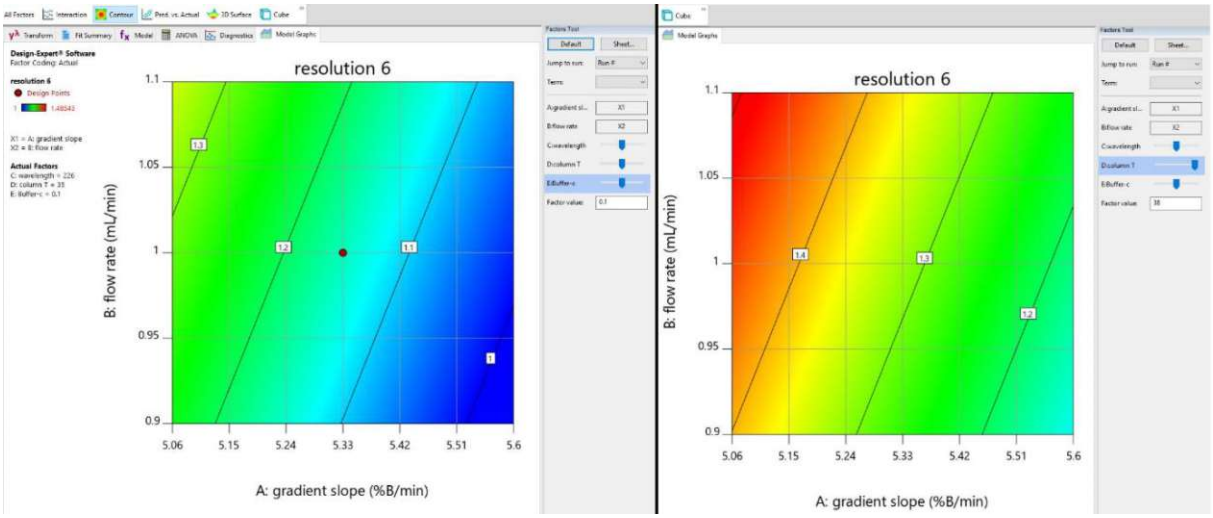

Figure 27: Graphical analysis of the influence of parameters on a response.

One of the last steps is categorising the parameters as critical or not. If a parameter is listed not critical, it should not have an influence on the outputs.
In the navigation pane under *Post Analysis – Coefficient Table* are all p-values listed, which point to the critical variables, where only the p-values smaller than 0.05 are statistically significant.

But this table only shows the direct influence of one parameter on the output, not really the interaction with other variables and is based only on the p-value.
For a graphical examination and investigating robust ranges the tab *Optimization* must be used. Ranges for every output must be set and then the overlay plot can be viewed (see Figure 28).
In this plot, yellow means inside and grey stands for outside the range. At best, the whole plot is yellow, which means both parameters, also when interacting, are robust for the method.

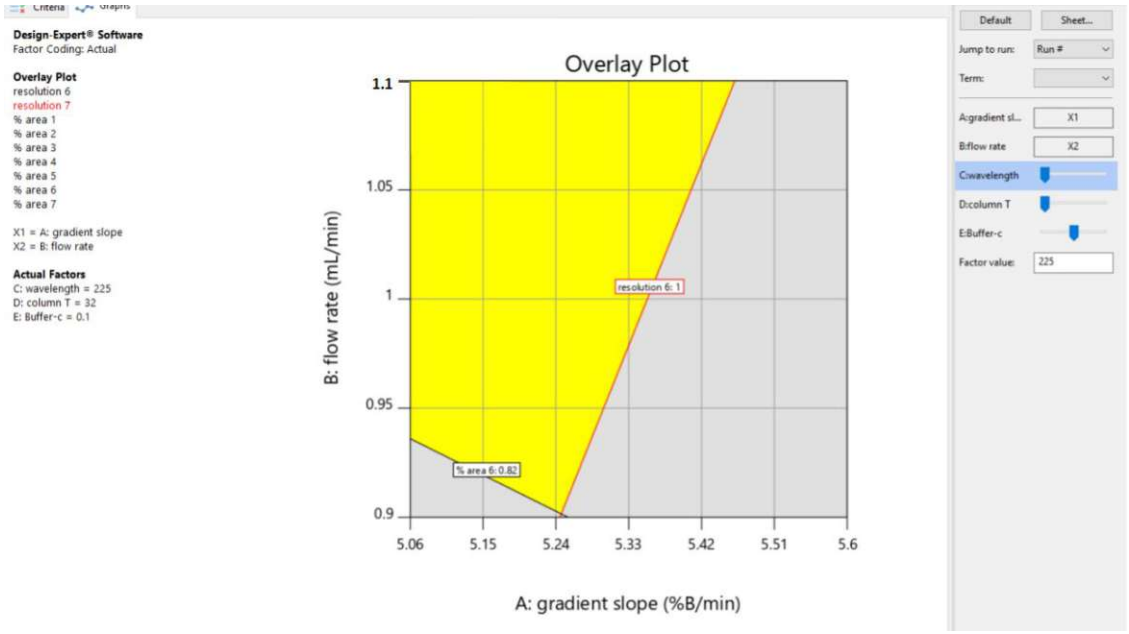

Figure 28: Overlay plot for robustness studies with flow rate over gradient slope.

4 Result and discussion

In this chapter the results from the master thesis are stated and discussed. Mostly, it includes the data, which was gathered through the method development described in chapter 0, but it also shows some problematic matters that came up, like poor baseline quality of some columns, which is the reason for some retrospective changes.
For a final procedure guideline, a flow chart was created and added to the appendix (Figure 76 to Figure 78).

4.1 Measuring system pressure

The back pressure never exceeded the limit of the columns. The system pressure was about 100 bar, so the limit of the methods was programmed at least 75 bar higher than the critical pressure given by the manufacturer of the columns.

4.2 Dwell, extra column and void volume

The dwell, extra column and void volume of the system were measured as mentioned in section 3.3.1 and their results are listed in Table 12.

Table 12: Values of dwell, extra column and void volume in mL for the set up and columns used in this thesis.

Dwell volume	$ECV_{measured}$	$ECV_{calculated}$	Void volume	
			I.D. of 3.0 mm	I.D. of 2.1 mm
0.4	0.033	0.013	0.43	0.23

As explained before the ECV was not only calculated but also measured. The calculated value was put together with the following data in Table 13 and the equation (7).
All tubings have an I.D. of 0.1016 mm, therefore a *radius r* of 0.0508 mm and their various *lengths l* are also shown in the table below. Their *Volumes V* were gathered through the equation (8).
The preheater was presumed as a tubing, because of missing data from Waters for a more precise value, although the volume can be expected to be higher.
The volumes of the detector cell, the switching and injection valves were given from Waters.

Table 13: Values of the contributing volumes to the ECV based on manufacturer informations for injection valve, preheater, switching valves, detector cells and calculations based on their dimension (length and diameter) for the tubings.

Contributor	Volume (µL)
Injection Valve	0.456
Tubing 1 after Injection valve (approx. 31 cm)	2.574
First switching valve	0.149
Preheater (31.75 cm)	2.574
Connecting element	approx. 0 (ideal)
Tubing 2 after connection (36.83 cm)	2.983
Second switching valve	0.149
Tubing 3 to detector (48.26 cm)	3.913
Detector flow cell	0.25

Because of the big difference between the calculated and measured ECV, possible errors of worn-out fittings needed to be eliminated.

New tubings were installed and the preheaters were switched.

After that, the difference was about the same, so the most likely reason is, that the approximation of the preheater as a tubing is not precise enough.

Normally an ECV of 33 µL seems too high for an UPLC, but because of the configuration with a lot of valves and column positions, it could be possible, so there was no further error testing.

4.3 Loading capacity

As explained in section 0 the loading capacity was measured with different stationary phases, to find an appropriate particle type.

The main goal was only a small to no drift of the retention time with increasing concentration and if possible, also no peak broadening.

From a theoretical point of view, the TPPs should have a better loading capacity because SPP cores are unavailable for analyte interaction with the stationary phase and therefore have a smaller specific surface area and can bind less analyte (graphically represented in Figure 2 of section 2.1.2).

But experimentally, it is already explored, that SPP columns can keep up with TPPs [25], which was confirmed in the following experiments.

As shown in Figure 29 to Figure 33 below, as well as in Table 14, the peak broadening of all the columns are alike and also the retention time drift is comparable.

So, the decision point (discussed in section 0 below) based on the loading capacity in favour of the TPP columns could not be verified.

Table 14: Width at the baseline and retention time of the loading capacity testing. As the inner diameter of the SPP (3 mm) and TPP columns (2.1 mm) vary, the conditions of the method as well as the injection volume was adapted, for a better comparison.

Column	Phase	Inj-V. (uL)	Width @ Baseline	RT	Plate count
Agilent Poroshell 120 EC-C18	SPP	2	0.155	7.32	$3{,}56 \cdot 10^4$
		1	0.120	7.35	$6{,}00 \cdot 10^4$
		0.5	0.095	7.38	$9{,}64 \cdot 10^4$
Waters Cortecs C18+		2	0.115	5.66	$3{,}87 \cdot 10^4$
		1	0.090	5.68	$6{,}38 \cdot 10^4$
		0.5	0.070	5.70	$1{,}06 \cdot 10^5$
Phenomenex Kinetex C18 EVO		2	0.125	5.43	$3{,}02 \cdot 10^4$
		1	0.095	5.46	$5{,}29 \cdot 10^4$
		0.5	0.070	5.48	$9{,}82 \cdot 10^4$
Waters BEH Phenyl	TPP	1	0.130	5.16	$2{,}52 \cdot 10^4$
		0.5	0.100	5.19	$4{,}31 \cdot 10^4$
		0.3	0.075	5.21	$7{,}73 \cdot 10^4$
Waters XSelect HSS PFP		1	0.145	5.63	$2{,}41 \cdot 10^4$
		0.5	0.110	5.66	$4{,}24 \cdot 10^4$
		0.3	0.075	5.69	$9{,}22 \cdot 10^4$

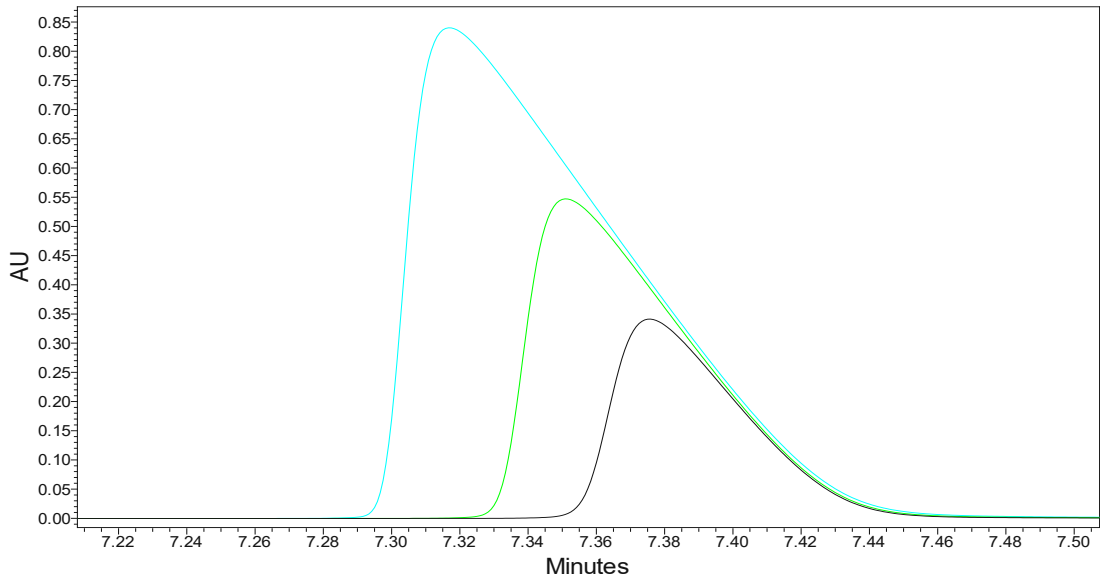

Figure 29: Loading capacity of Poroshell 120 EC-C18 with 2, 1 and 0.5 µL injection volume of the sample *1-(α-Aminobenzyl)-2-naphthol hydrochloride* dissolved in 1.5 % acetonitrile in water, with a concentration of 0.1503 g L^{-1} (left to right).

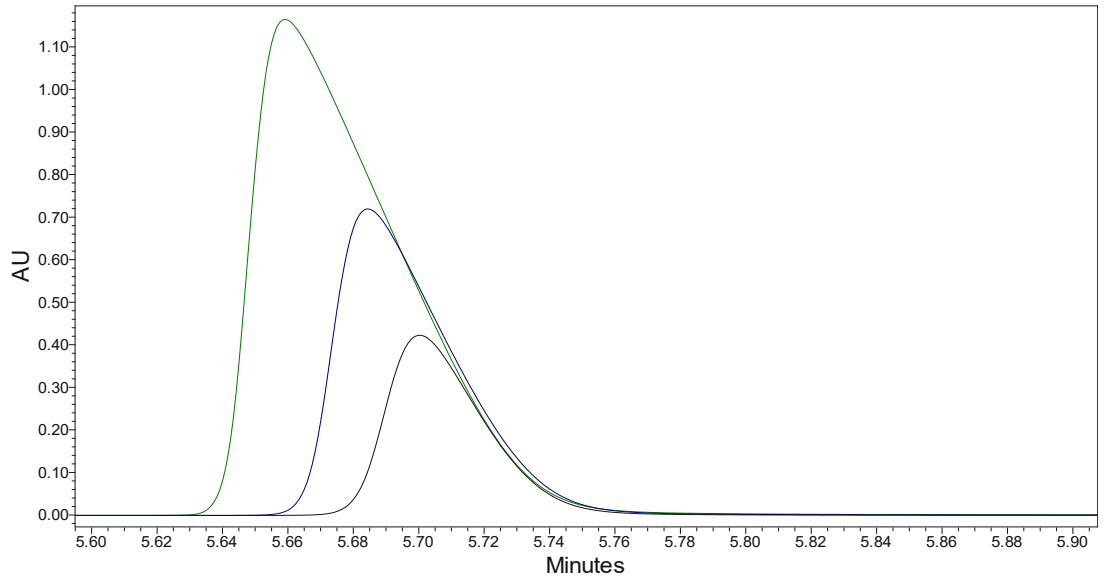

Figure 30: Loading capacity of Waters Cortecs C18+ with 2, 1 and 0.5 µL injection volume of the sample *1-(α-Aminobenzyl)-2-naphthol hydrochloride* dissolved in 1.5 % acetonitrile in water, with a concentration of 0.1503 g L^{-1} (left to right).

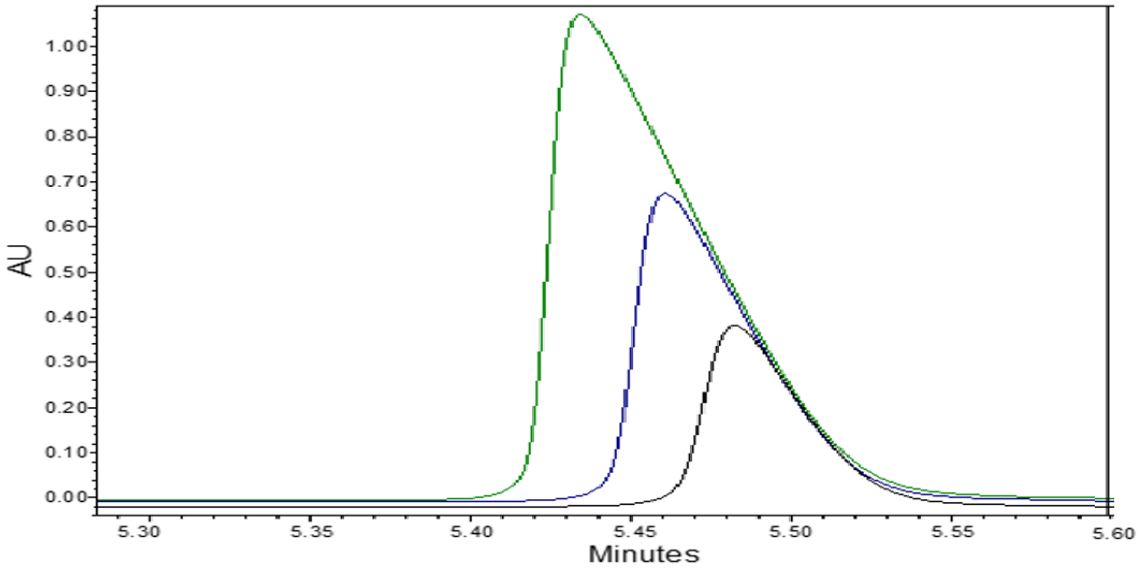

Figure 31: Loading capacity of EVO with 2, 1 and 0.5 µL injection volume of the sample *1-(α-Aminobenzyl)-2-naphthol hydrochloride* dissolved in 1.5 % acetonitrile in water, with a concentration of 0.1503 g L^{-1} (left to right).

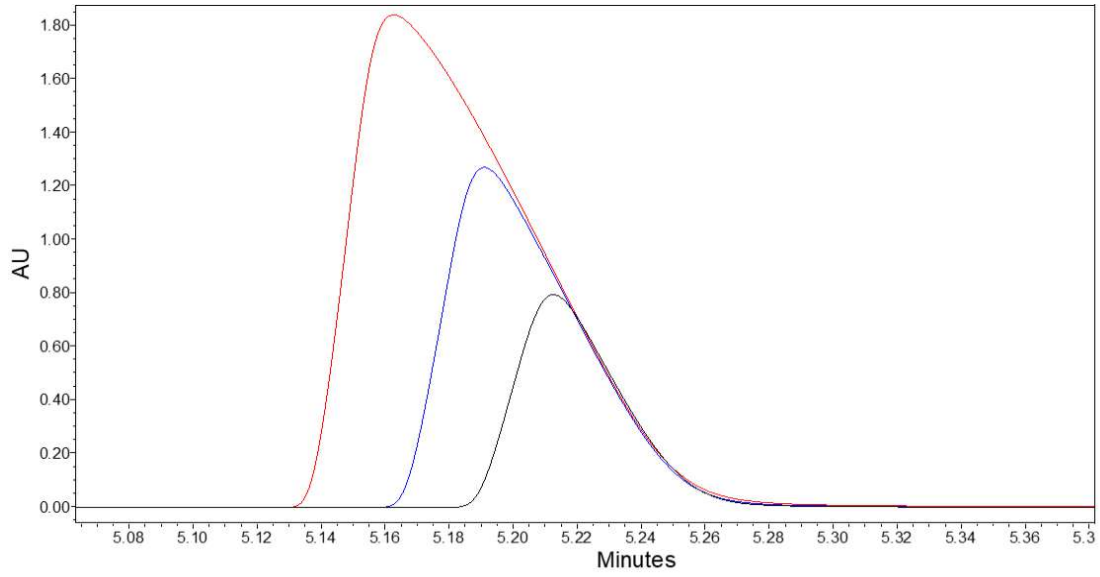

Figure 32: Loading capacity of BEH Phenyl with 1, 0.5 and 0.3 µL injection volume of the sample *1-(α-Aminobenzyl)-2-naphthol hydrochloride* dissolved in 1.5 % acetonitrile in water, with a concentration of 0.1503 g L^{-1} (left to right).

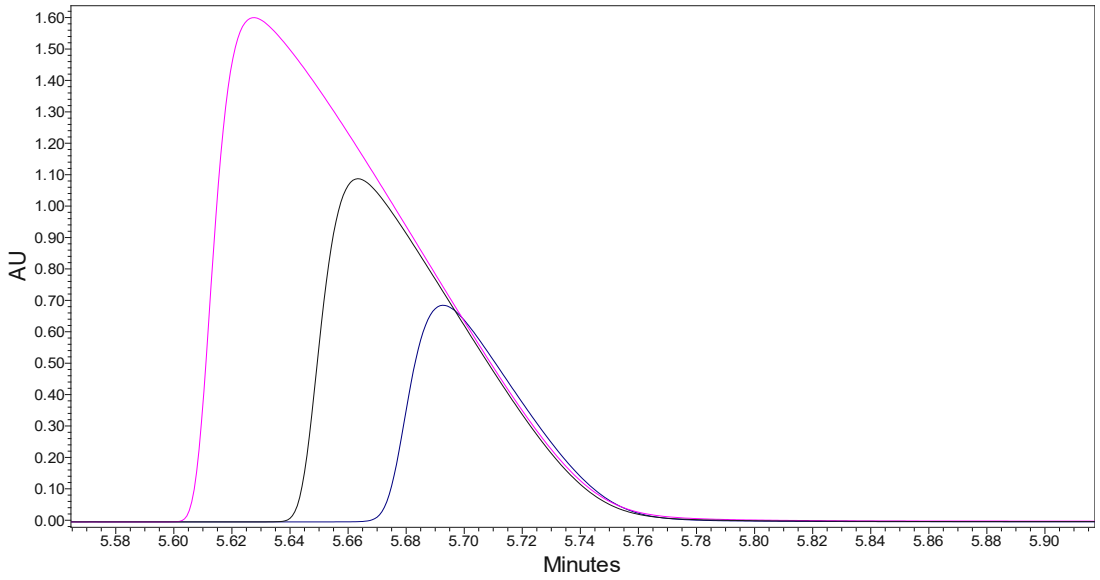

Figure 33: Loading capacity of HSS PFP with 1, 0.5 and 0.3 µL injection volume of the sample *1-(α-Aminobenzyl)-2-naphthol hydrochloride* dissolved in 1.5 % acetonitrile in water, with a concentration of 0.1503 g L^{-1} (left to right).

4.4 System Suitability Test

The results of the SST for the following columns are shown in Table 15 and Figure 34 to Figure 37. The reason of the SST is, to test the setup of the HPLC together with the columns on every column position. At first there were some symmetry problems, but after replacing the preheaters on the positions 1, 3 and 4, the shapes of the peaks of ethylbenzene and toluene were adequate.

Also, the plate count of the columns was compared to the statement of the manufacturer. The values were deemed close enough to keep on with the testing. Only the plate count of Waters Cortecs C18+ could not be compared, because of missing information of the manufacturer.

The high difference between the measured and the proposed values by the manufacturer is most probably based on the rather high extra column volume of the system set up.

Table 15: Summary of RT, Resolution, Plate Count and Tailing of the tested columns for the system suitability test runs.

Column	Retention Time	USP Resolution	USP Plate Count	USP Tailing	Plate count manufacturer
Phenomenex Kinetex C18 EVO	0.834		$1.72 \cdot 10^4$	1.26	$2.23 \cdot 10^4$
	1.035	6.86	$1.77 \cdot 10^4$	1.16	
Waters Cortecs C18+	1.097		$2.26 \cdot 10^4$	1.06	n.a.
	1.477	10.84	$2.23 \cdot 10^4$	1.06	
Waters BEH Phenyl	1.066		$1.76 \cdot 10^4$	1.09	$2.85 \cdot 10^4$ (tangent efficiency)
	1.29	6.36	$1.89 \cdot 10^4$	1.06	
Waters XSelect HSS PFP	1.143		$2.00 \cdot 10^4$	1.06	$2.51 \cdot 10^4$ (tangent efficiency)
	1.296	4.42	$2.10 \cdot 10^4$	1.02	

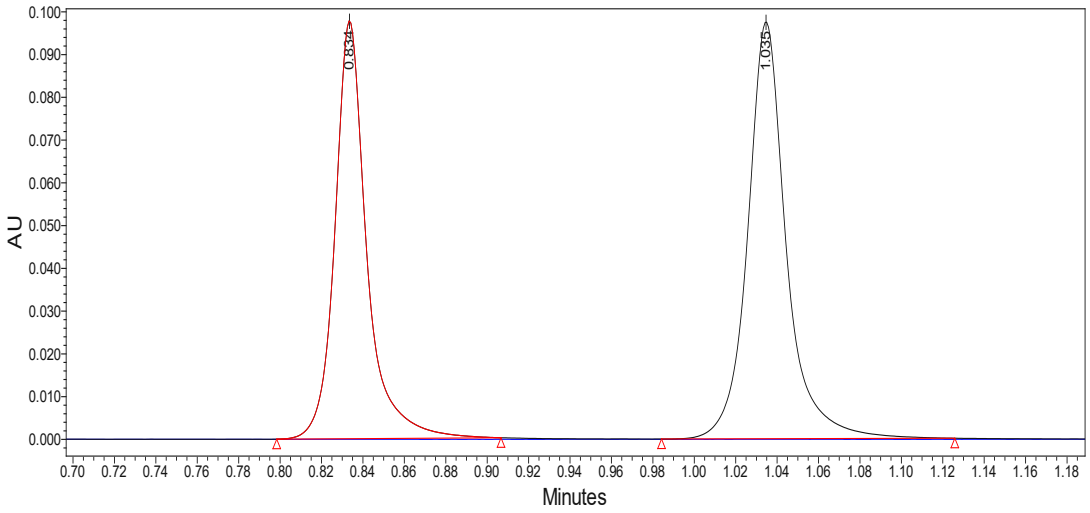

Figure 34: SST of the column Phenomenex Kinetex C18 EVO, with toluene (retention time of 0.834) ethylbenzene (retention time of 1.035).

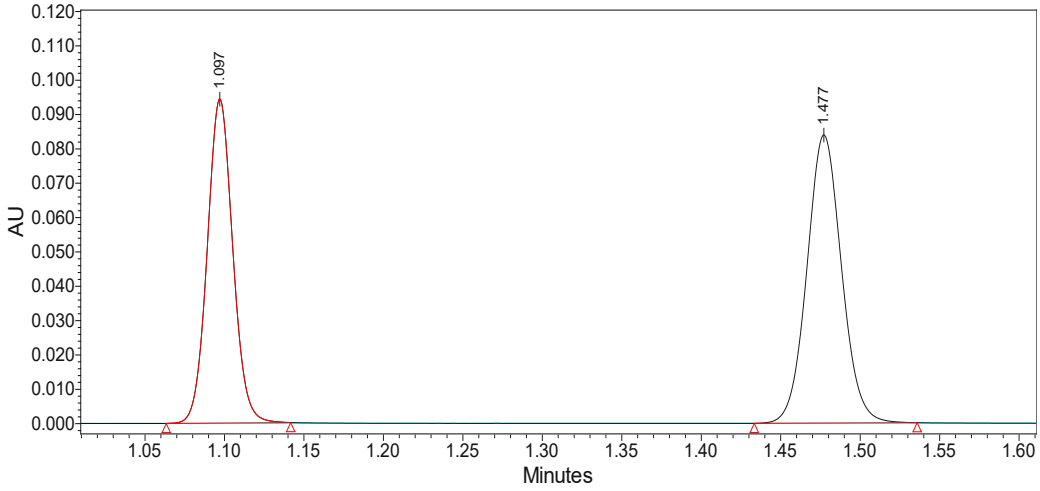

Figure 35: SST of the column Waters Cortecs C18+ with toluene (retention time of 1.097) ethylbenzene (retention time of 1.477)

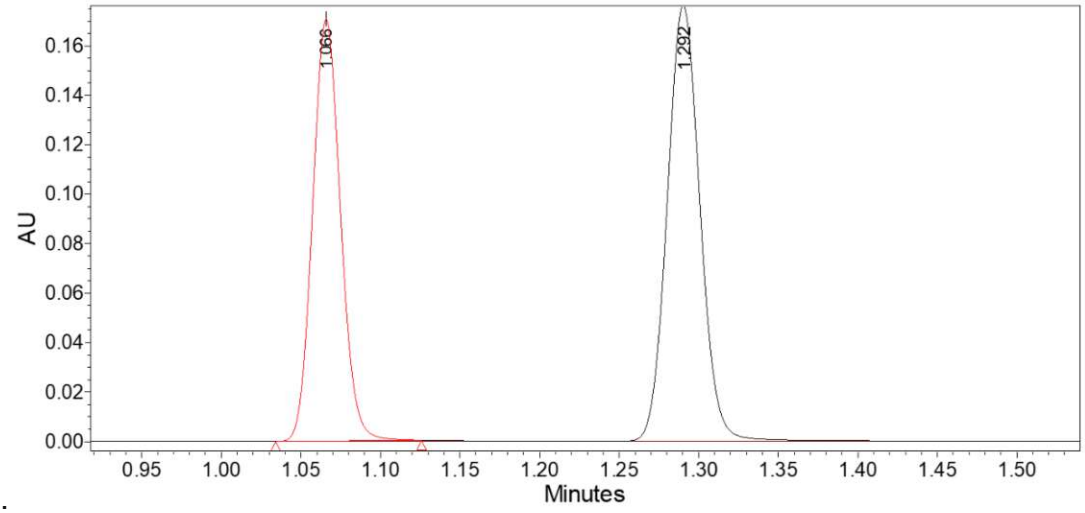

Figure 36: SST of the column Waters BEH Phenyl with toluene (retention time of 1.066) ethylbenzene (retention time of 1.292).

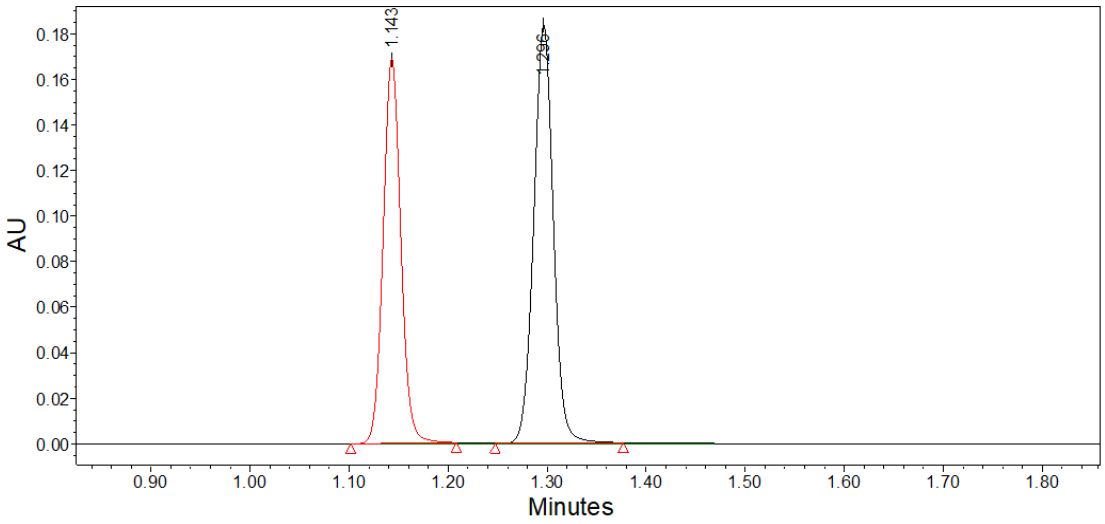

Figure 37: SST of the column Waters XSelect HSS PFP toluene (retention time of 1.143) ethylbenzene (retention time of 1.296).

4.5 Stationary phases

The decision of the particle type for a set of columns for a generally used screening setup was based on the requirements listed in Table 16. To get comparable results the TPP columns need a particle size of about 1.7 µm, while the SPP columns can be used with 2.7 µm. That is the reason for the high back pressure for the examined TPP columns listed in Table 16.

Also, most of the commercially TPP columns, which were experienced to be reliable, are only available with an inner diameter of 2.1 mm, which therefore led to a slower flow rate, because of higher back pressure.

Table 16: Comparison of the criteria of the for this thesis considered column types.

Column type	TPP (I.D. 2.1 mm)	SPP (I.D. 3.0 mm)
Particle size (µm)	approx. 1.7	approx. 2.7
Back pressure (at the chosen particle sizes)	high	low
Particle size distribution		narrower than TPP
Flow rate	approx. half of SPP	
Efficiency	similar	
Loading capacity	better (theoretically)	similar (practically)

The inner diameter of 3 mm has the benefit of a lower back pressure at higher flowrates in comparison to the 2.1 mm and can therefore be established more easily on an HPLC system too.

All in all, the columns used for this thesis with the SPPs had a smaller back pressure and a better resolution, because of their dimensions. But one of the most important demands was a good loading capacity, which was expected to be better with totally porous particles. Experimentally, in this case, the possible higher loading capacity of the TPPS columns, with the used dimensions was not sufficient for a better separation. Summarising all these points, the stationary phases listed below in Table 17 were chosen for the procedure for this master thesis.

For the final procedure, the originally tested Phenyl column *Phenomenex Kinetex Biphenyl* and the PFP column *AMT Halo PFP*, both with an SPP phase, were replaced with the *Waters BEH Phenyl* and *Waters XSelect HSS PFP,* both with a TPP phase. The reason were performance issues with the *Phenomenex Kinetex Biphenyl* and the *AMT Halo PFP* column, which the two TPP columns did not show.

More details about these performance issues are discussed in section 4.11.

Another decision-making tool to find the right stationary phases, was the F-value, which is discussed in section 2.1.2.2. The goal was achieving high F-values between the columns because it ensures the base of a most effective orthogonal phase system. Table 17 compares the F-values of the finally selected columns at different conditions. (For the comparison of all columns see Table 25 in the Appendix).

The conditions were top down at pH 2.8 with a base present in the sample, pH 2.8 with an acid present in the sample, pH 2.8 without acid or base present in the sample and at pH 6.7 with a base present in the sample.

The F-values with conditions of pH 6.7 with an acid present in the sample or without both, acid and base, present in the sample were not included in the table separately, because they were very similar to the values of pH 6.7 with base present in the sample.

Table 17: F-values of the columns top down at pH 2.8 with a base present in the sample, with an acid present in the sample, pH 2.8 without acid or base present in the sample and at pH 6.7 with a base present in the sample.

Columns	EC-C18	Bonus-RP	EVO	HSS PFP	BEH Phenyl
Waters Cortecs C18+	18.57	63.03	12.19	66.34	29.2
	10.19	31.84	12.68	18.1	12.3
	9.24	23	10.48	16.87	11.17
	32.12	51.85	43.87	101.33	12.55
Agilent Poroshell 120 EC-C18		70.47	22.41	51.99	13.49
		30.02	3.45	19.97	12.36
		14.45	1.93	19.83	7.98
		21.35	12	132.17	37.32
Agilent Poroshell 120 Bonus RP			48.39	119.76	82.07
			31.88	38.26	25.4
			12.86	25.42	18.99
			13.43	148.58	55.53
Phenomenex Kinetex C18 EVO				72.96	34.56
				19.22	15.59
				19.2	9.59
				143.81	49.25
Waters XSelect HSS PFP					44.76
					27.53
					24.91
					97.45

This table shows that the F-values of the columns show mostly high differences and therefore the columns have rather different separation properties, which was the goal of selecting the columns.

4.6 Mobile phases

As mentioned in section 3.9 phosphoric acid and ammonium phosphate were chosen for the aqueous eluents. The reason is the better signal to noise ratios of low-level impurity peaks at lower detection wavelengths than measurements with other eluents like trifluoroacetic acid (TFA) or formic acid.

The downside of the chosen eluents is the incompatibility with MS-measurements, but they were still chosen because of the good UV detection at low wavelength and the consequential good automatic integrability for an efficient data processing and data integrity for future routine measurements.

But if the product and its contaminations can be detected at higher wavelength TFA, formic acid and ammonium acetate should be considered, too, because of their better compatibility with MS-detections.

4.7 Initial run and sample preparation

Shown in Figure 38 is the initial run of *1-(α-amino benzyl)-2-naphthol hydrochloride*. All peaks with an area percent higher than 0.03 % were integrated to locate all possible impurities. Later in the Screening, only Peaks with an area percent value higher than 0.05 % are integrated, which is why at the initial run, there are more peaks integrated than later in the screening runs.

The peak height (about 1 AU) of the main peak indicates that the concentration of the dilution and the injection volume are suitable for the method, which confirms the appropriateness of the sample preparation.

The 1 AU was chosen based on the detector specification as well as out of experience with HPLC methods.

The UV spectra of all the integrated peaks show a high absorption at the wavelengths between 224 and 229 nm. So, for further tests the wavelength of 226 nm was used.

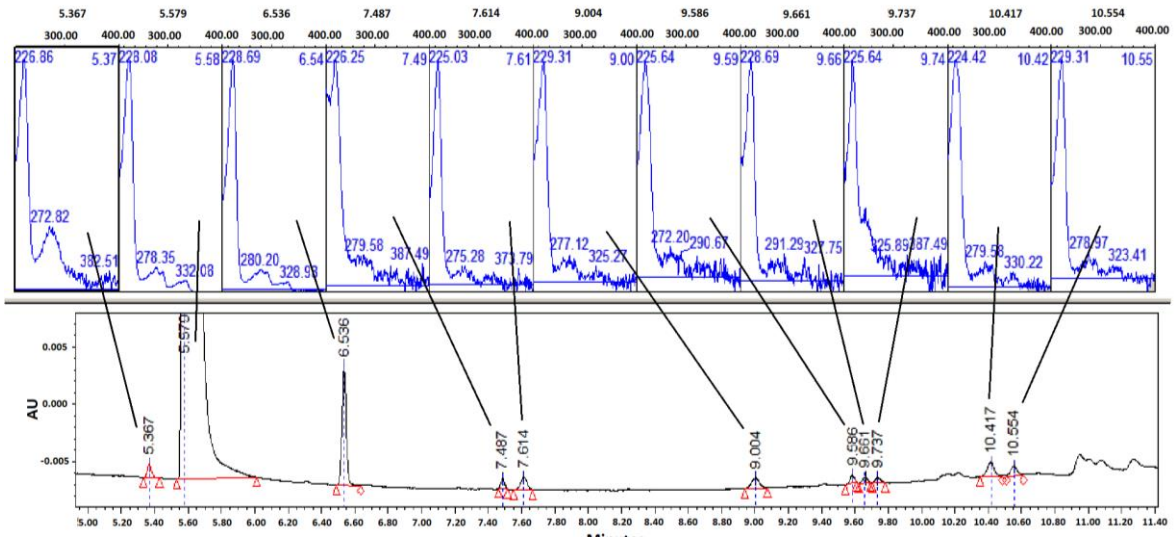

Figure 38: Chromatogram of the initial run on the Waters Cortecs C18+ to set a reasonable concentration for main peak and impurities and to find a fitting wavelength adjusted to detect the impurities together with the main component. and above the UV spectra corresponding to the peaks below.

4.8 Screening

In this section, only the mobile and stationary phase combinations that were rated useful for further potential development were thoroughly examined and discussed. Their chromatograms are shown below in Figure 39 to Figure 44. The residual chromatograms can be found in the Appendix (Figure 65 to Figure 75).

The selection of the chromatograms depended on the defined ATP. Therefore, all chromatograms with noticeable smaller peak amounts or with baseline problems were sorted out.

The outputs of the closer selection were all with 0.1 % aqueous phosphoric acid as the aqueous mobile phase. At a first view, the columns Waters Cortecs C18+, Agilent Poroshell 120 Bonus RP and Agilent Poroshell 120 EC-C18 show the most promising results.

The chromatograms of these columns with the 0.1 % aqueous phosphoric acid combined with each methanol and acetonitrile are displayed in Figure 39 to Figure 44. The peaks with a minimum of 0.05 % peak area were integrated as impurities in the sample. Beside the number of peaks, other criteria like their resolution and symmetry of the main peak were considered, too.

The combination of Poroshell 120 EC-C18 with methanol and the acidic aqueous eluent shows one peak more than the other chromatograms, because the two peaks at about 11 min overlap with the other conditions. It is the only condition where an

impurity peak elutes directly after the main peak, which is why it had been excluded from the beginning and the higher number of peaks was overlooked unintentionally.

All the chromatograms with methanol as organic phase have a positive baseline drift, while the baselines of the measurements with acetonitrile stay stable.

But there was a problem with the blanks measured with acetonitrile. The chromatograms of the screening run with acetonitrile all showed a peak, which looked conspicuous and could be found in the blank too.

After replacing the acetonitrile eluent with a new one and changing the tubings, the problem decreased, which points to a contamination in the mobile phase.

Also, the chromatograms with acetonitrile as mobile phase have a better resolution between the peaks, which is an important asset.

Therefore, the decision was made, to use acetonitrile as the organic phase.

So, the final decision was about, which column to use. The resolution between the peaks is better with Poroshell 120 EC-C18 and Waters Cortecs C18+ than with Poroshell 120 Bonus RP.

In the end Waters Cortecs C18+ was chosen because the interfering blank peaks are further away from the impurity peaks than with the Poroshell 120 EC-C18.

Summarised, the **final conditions are 0.1 % phosphoric acid with acetonitrile on Waters Cortecs C18+.**

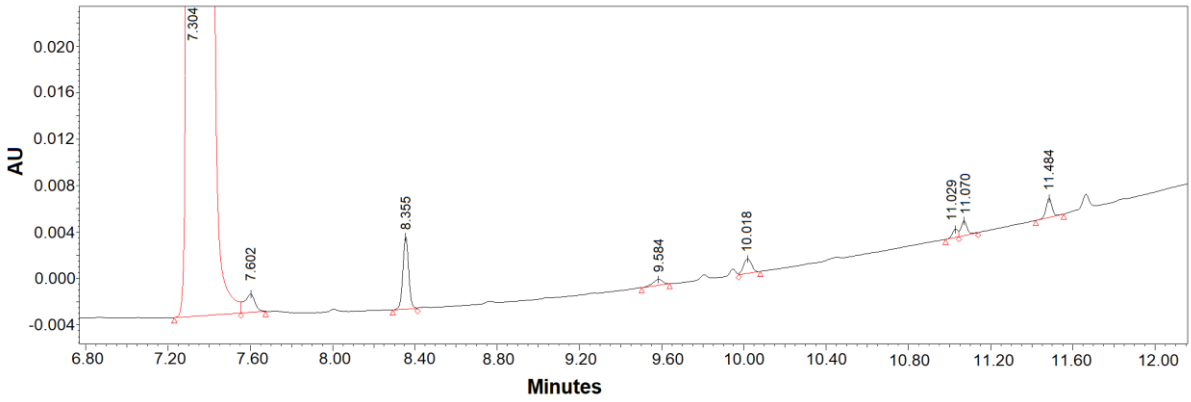

Figure 39: Chromatogram of the screening run with the column Poroshell 120 EC-C18 with 0.1 %aqueous phosphoric acid and methanol (zoomed).

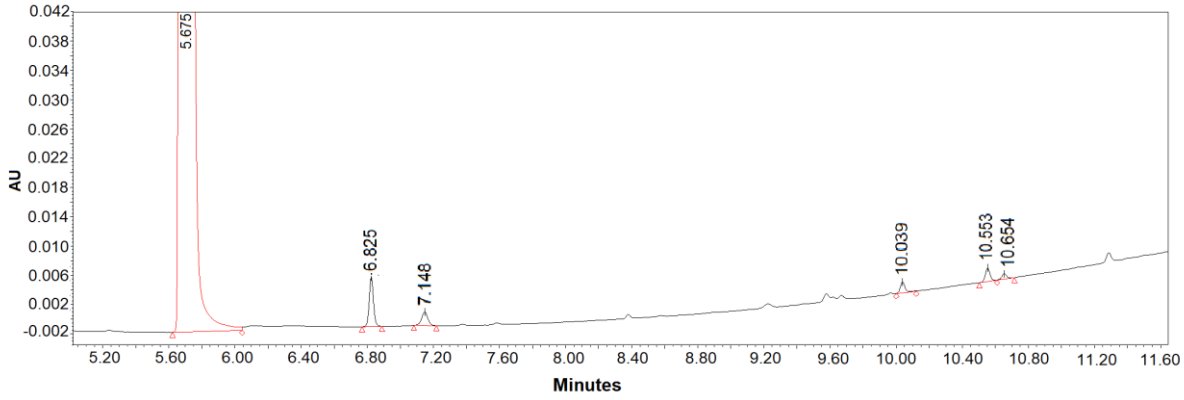

Figure 40: Chromatogram of the screening run with the column Waters Cortecs C18+ and the mobile phases 0.1 %aqueous phosphoric acid and methanol (zoomed).

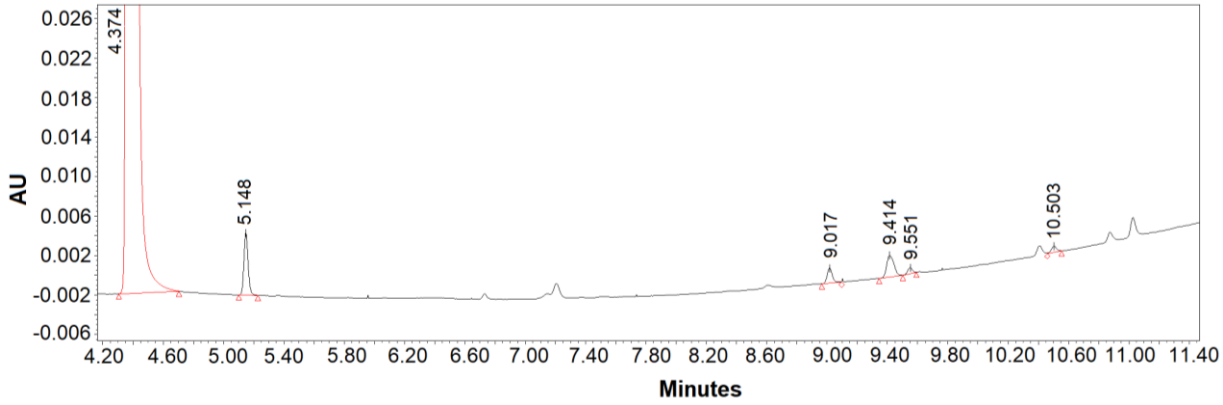

Figure 41: Chromatogram of Poroshell 120 Bonus RP with 0.1 % aqueous phosphoric acid and methanol (zoomed).

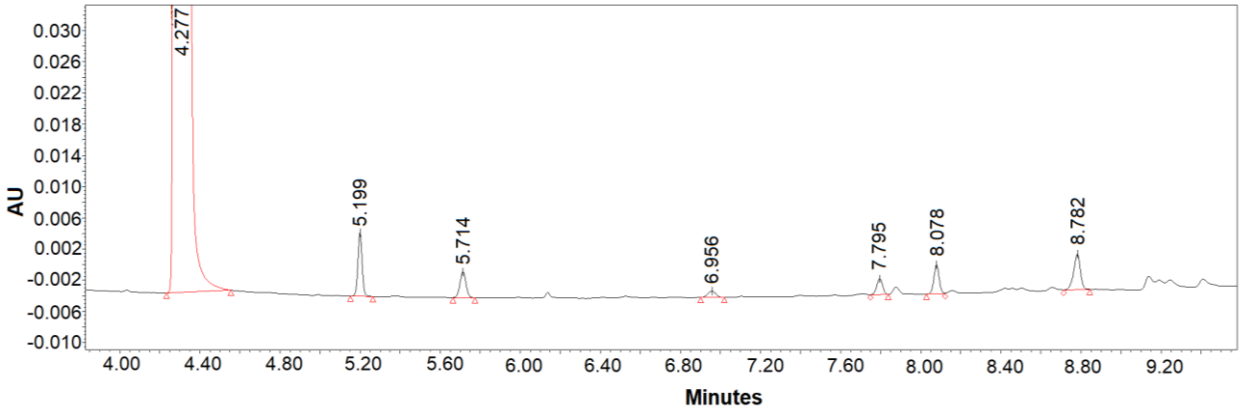

Figure 42: Chromatogram of Poroshell 120 EC-C18 with 0.1 % aqueous phosphoric acid and acetonitrile (zoomed).

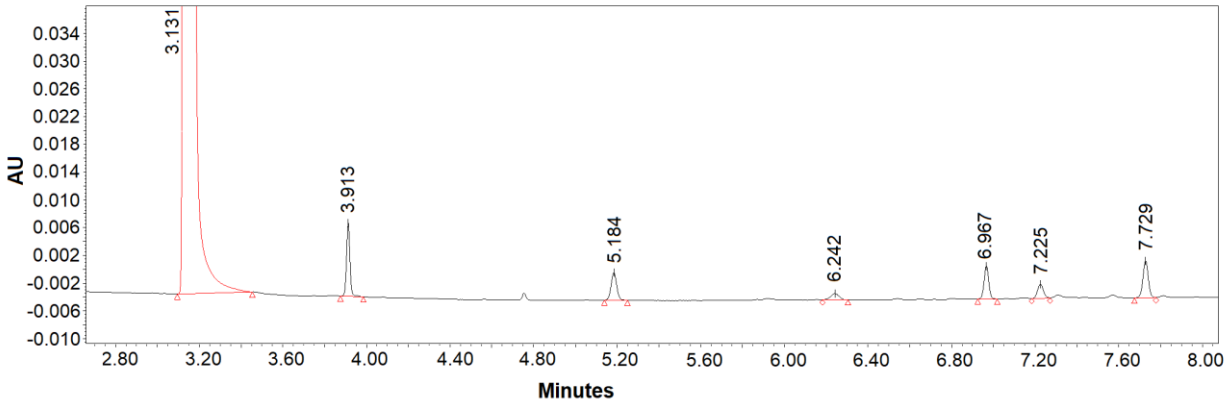

Figure 43: Chromatogram of Waters Cortecs C18+ with 0.1 % aqueous phosphoric acid and acetonitrile (zoomed).

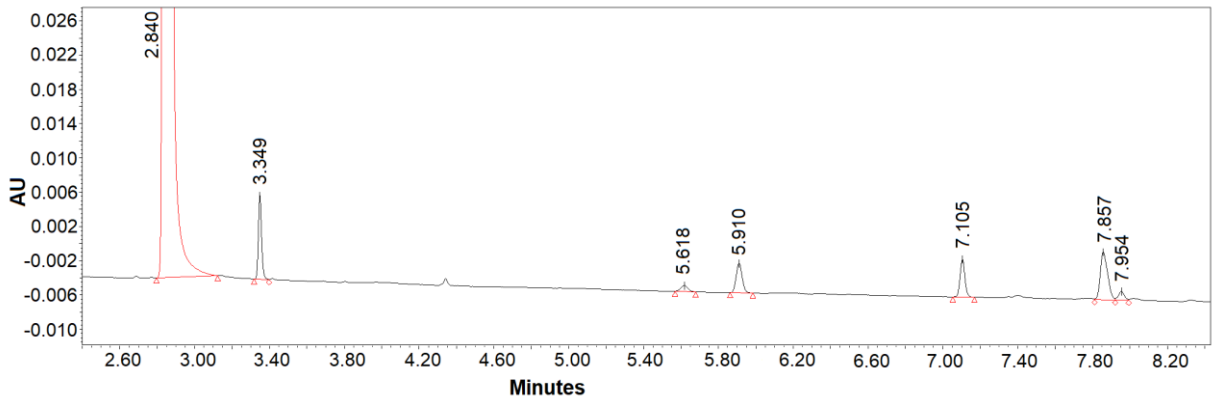

Figure 44: Chromatogram of Poroshell 120 Bonus RP with 0.1 % aqueous phosphoric acid and acetonitrile (zoomed).

Before continuing with the procedure, the purity of the main peak through the UV spectra was assessed (see Figure 45 and Table 18). The purity angle at 0.951 indicates an overlay with a second peak, but that is based on a known non-linearity of the purity angle around the peak maximum.

Table 18: Purity characteristics of the main peak with the final conditions. If the peak angle is higher than the purity threshold, it indicates possible overlapping of peaks.

Purity Angle	Purity Threshold	Maximum Impurity
0.951	0.648	3.150
0.054	0.648	3.212
0.017	0.648	3.126
0.011	0.648	3.136

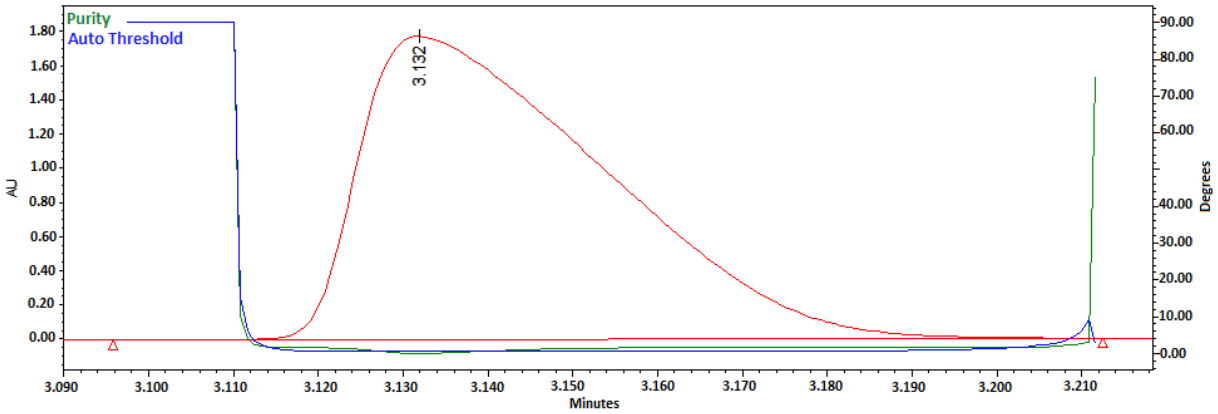

Figure 45: Purity check of the main peak of the chosen screening conditions showing no overlay with other peaks and therefore confirming the purity of the main peak.

4.9 Optimisation with DryLab

As mentioned in section 3.13 twelve runs were made for optimising the conditions. The parameters are also listed in section 3.13. The generated chromatograms together with the UV spectra of the peaks are shown in Figure 46 to Figure 57. The temperature of 55 °C is higher than the optimal temperature of 45 °C for the Cortecs C18+, which was overseen. But as there were only few runs with the high temperature and the 45 °C are a guidance for the longevity of the column it can be assumed, that the generated data is accurate, and the column was not damaged.

For a precise peak tracking the spectral data was used together with the peak area percentage.

As can be seen in Figure 46 some peaks like the fourth and sixth peak have very similar UV spectra percentage, which is why the comparison with the peak area percentage was important too.

Another complication was, that it became clear after some measurements, that the sample solution was not stable over a time period of one day.

Because the main goal of the thesis was to go through a method development based on the analytical quality by design and not to analyse the *1-(α-aminobenzyl)-2-naphthole hydrochloride* thoroughly, the inaccuracy of the result through time was tolerated.

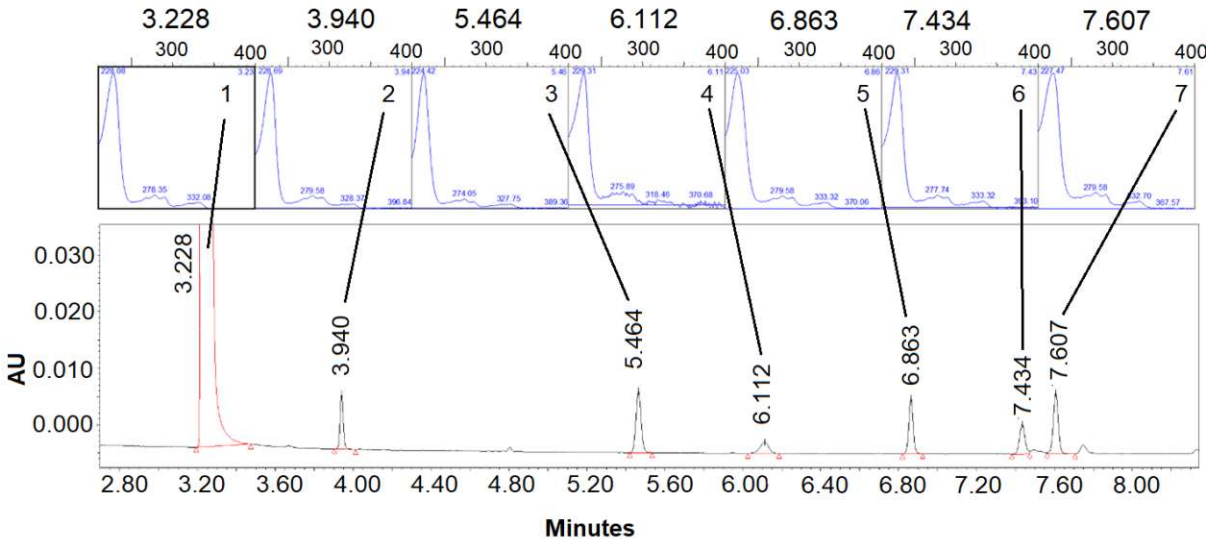

Figure 46: Optimisation run DryLab 1 at 25 °C, pH 2.06 and 6 % B min^{-1}.

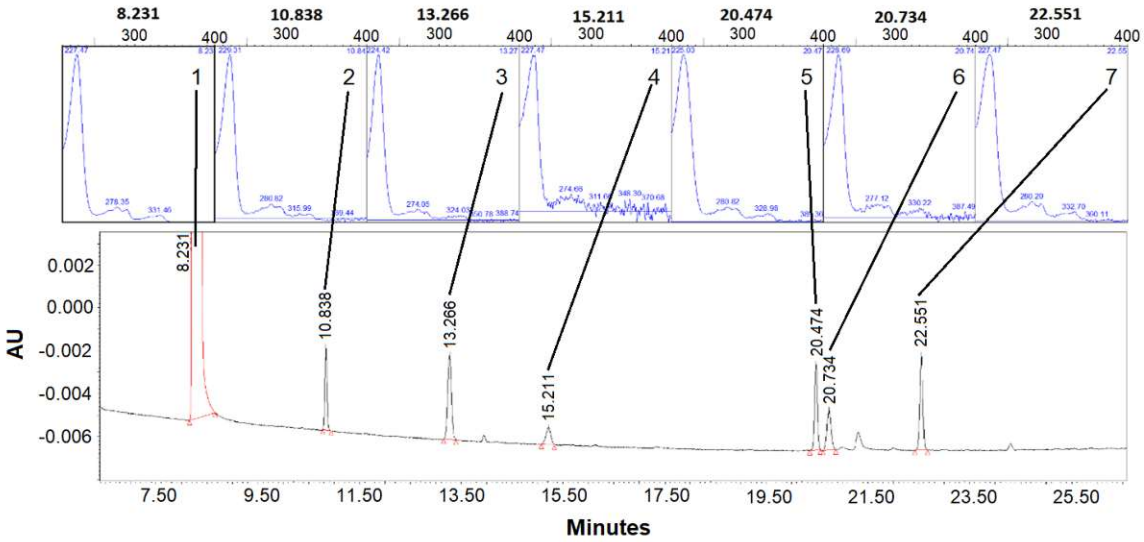

Figure 47: Optimisation run DryLab 2 at 25 °C, pH 2.06 and 1.64 % B min^{-1}.

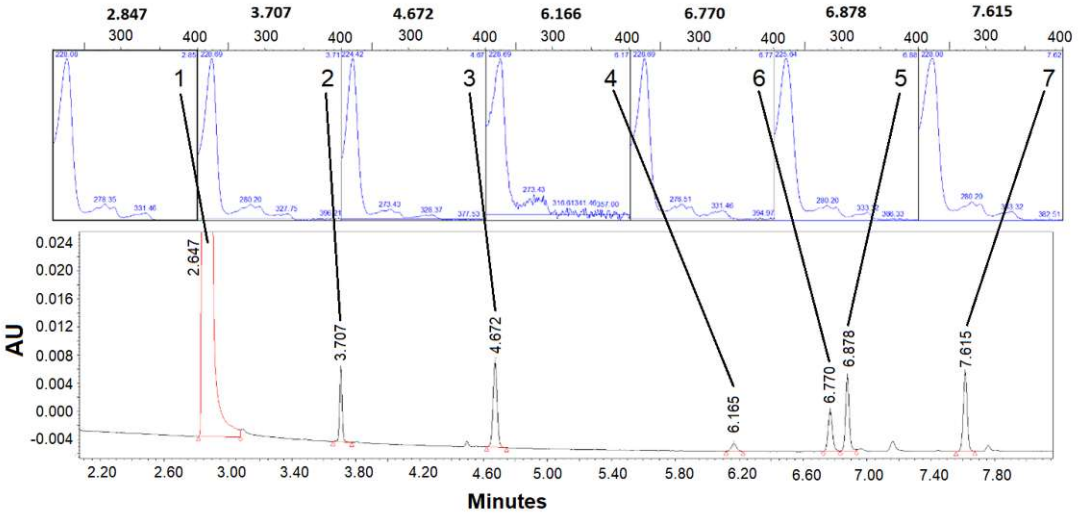

Figure 48: Optimisation run DryLab 3 at 55 °C, pH 2.06 and 6 % B min^{-1}.

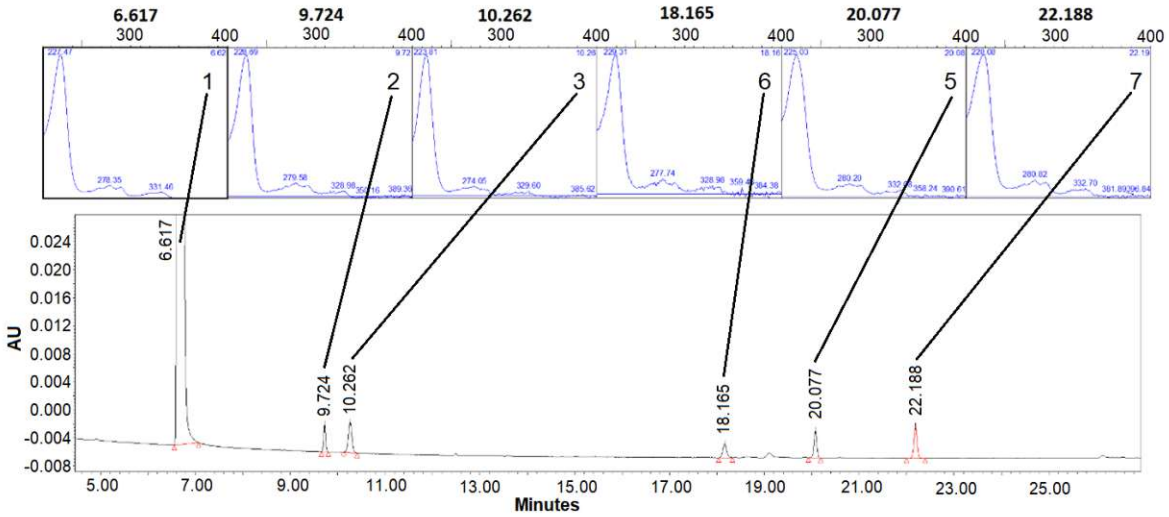

Figure 49: Optimisation run DryLab 4 at 55 °C, pH 2.06 and 1.64 % B min^{-1}.

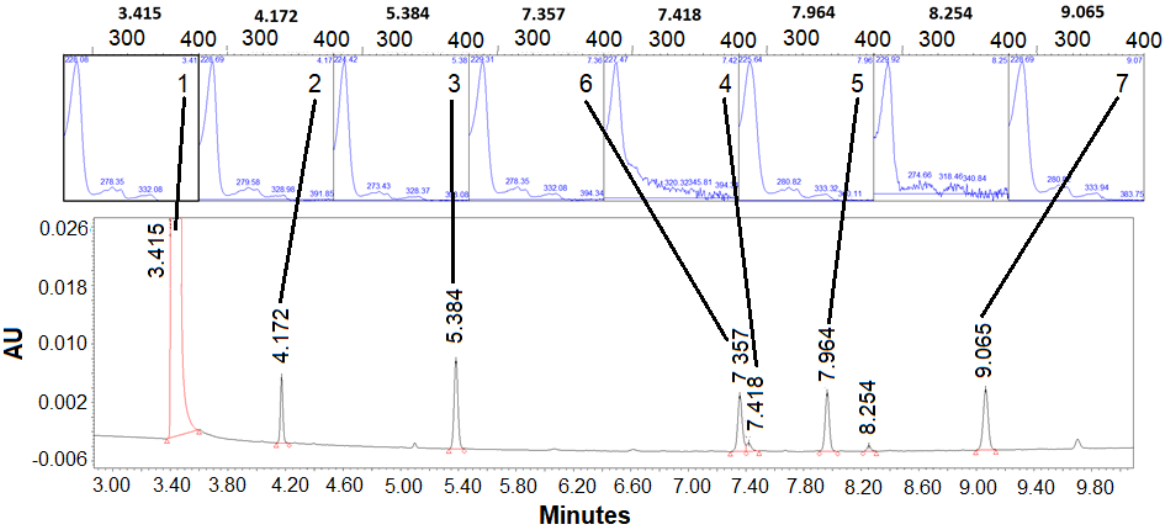

Figure 50: Optimisation run DryLab 5 at 25 °C, pH 2.4 and 6 % B min^{-1}.

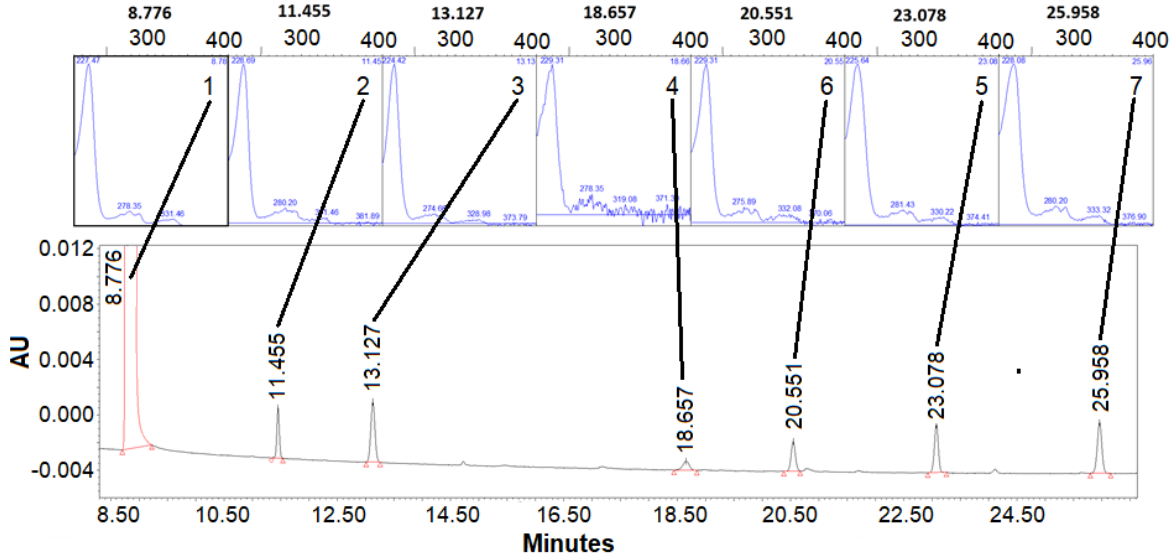

Figure 51: Optimisation run DryLab 6 at 25 °C, pH 2.4 and 1.64 % B min^{-1}.

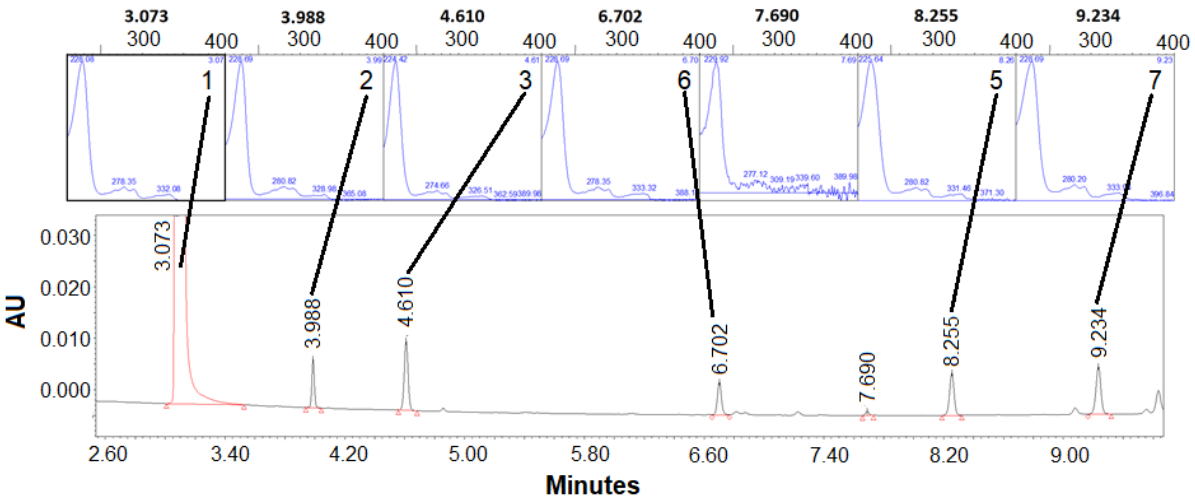

Figure 52: Optimisation run DryLab 7 at 55 °C, pH 2.4 and 6 % B min^{-1}.

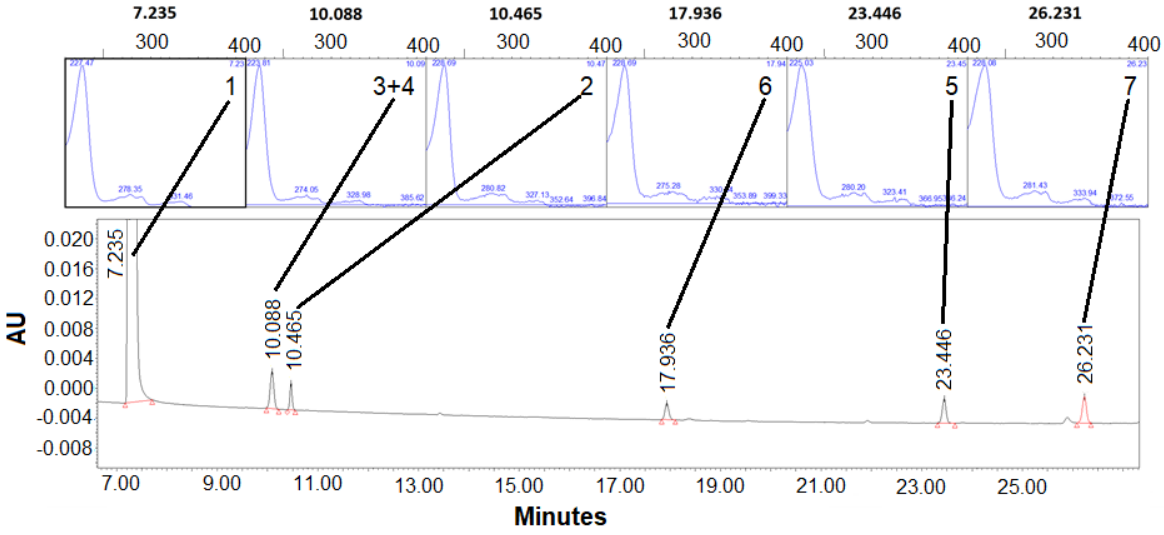

Figure 53: Optimisation run DryLab 8 at 55 °C, pH 2.4 and 1.64 % B min^{-1}.

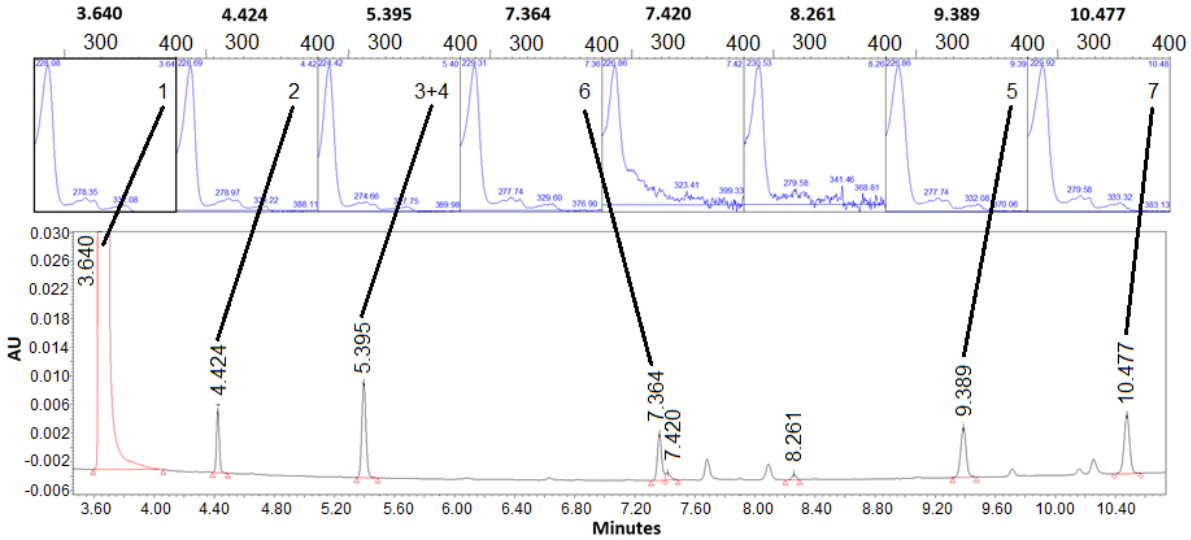

Figure 54: Optimisation run DryLab 9 at 25 °C, pH 2.8 and 6 % B min^{-1}.

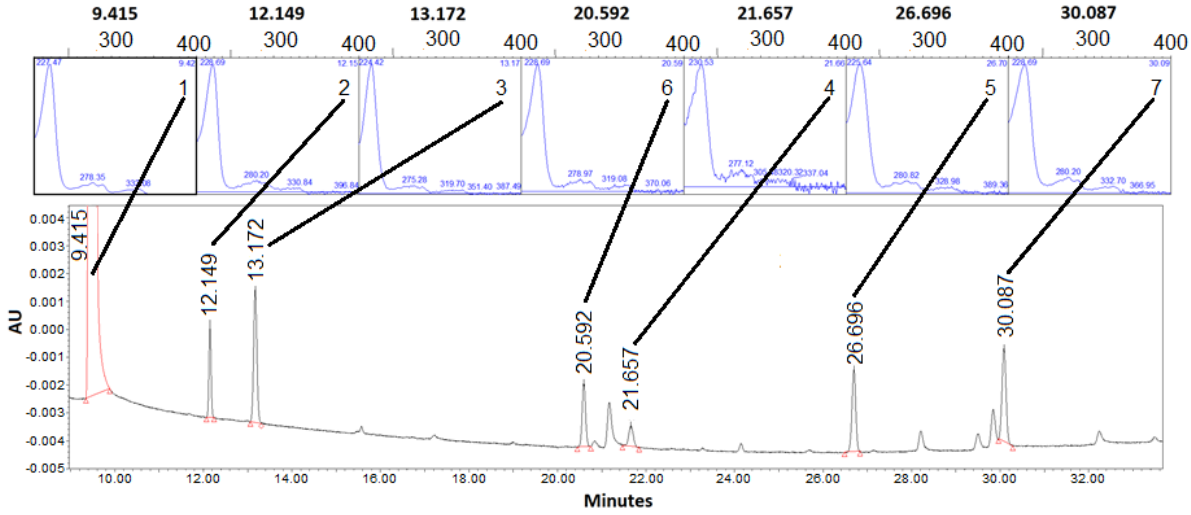

Figure 55: Optimisation run DryLab 10 at 25 °C, pH 2.8 and 1.64 % B min^{-1}.

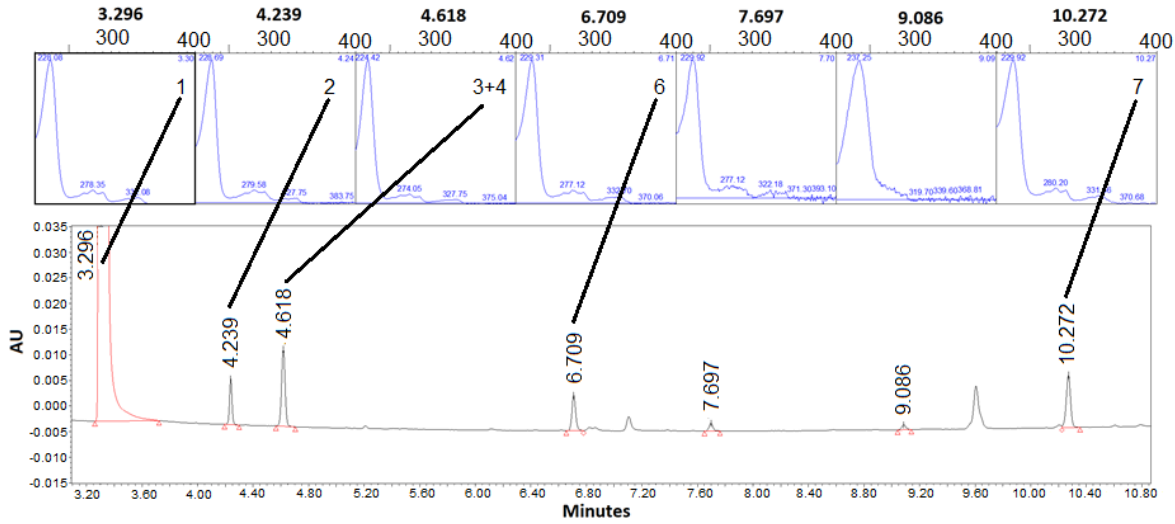

Figure 56: Optimisation run DryLab 11 at 55 °C, pH 2.8 and 6 % B min^{-1}.

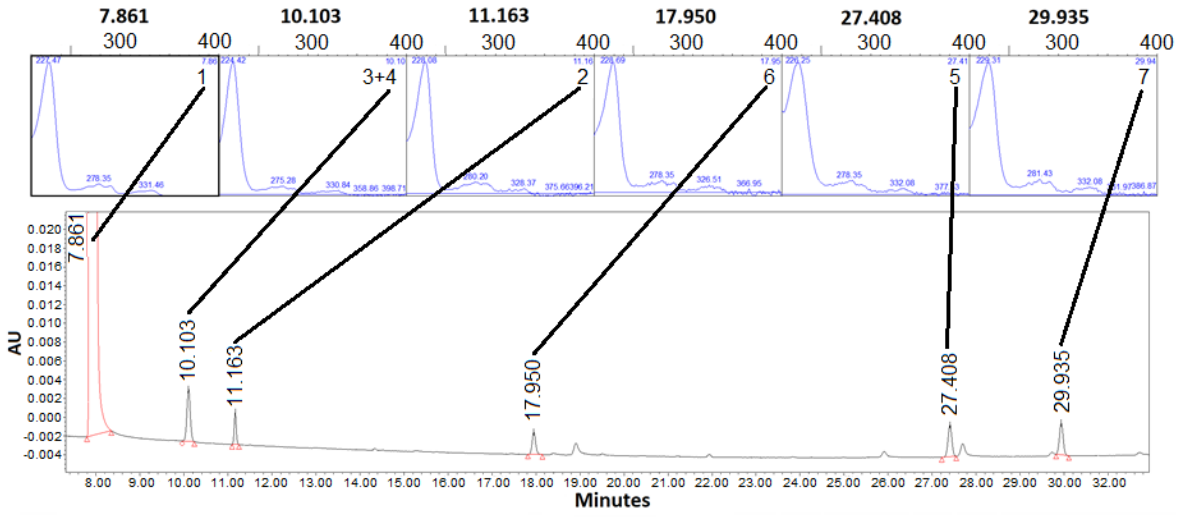

Figure 57: Optimisation run DryLab 12 at 55 °C, pH 2.8 and 1.64 % B min^{-1}.

After chosing the conditions with Drylab (see Figure 58 and Figure 59), listed in Table 19, the sample was analysed with them and the resulting chromatogram is shown in Figure 60. The peak at 5.688 in the experimental chromatogram had been omitted for the simulation as the tracking was not sufficient enough. The area of this peak differed significantly throughout the chromatograms of the optimisation runs, which interfered with the calculation of Drylab. Later it was discovered that the sample was not stable enough and that therefore, the area of this impurity differed a lot.

Table 19: Conditions for an optimal run calculated with DryLab.

Flow rate (mL min^{-1})	1	
Inj. V. (µL)	2	
Column temperature (°C)	35	
Detection Wavelength (nm)	226	
Sampling rate (points s^{-1})	20	
Run time (min)	21	
Eluent A	0.1 % H_3PO_4 in water	
Eluent B	acetonitrile	
Gradient	**t (min)**	**% Eluent B**
	Initial	10
	15	90
	18	90
	18.01	10
	21	10

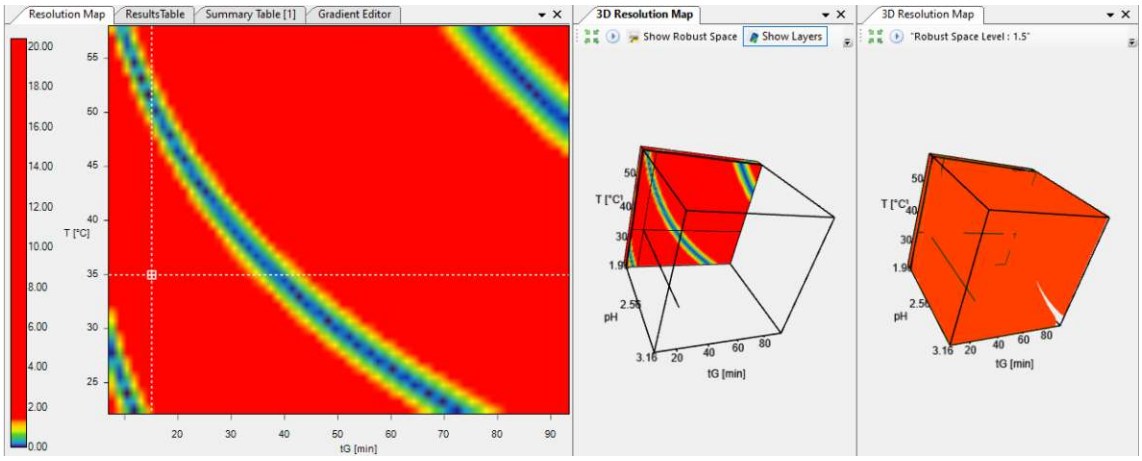

Figure 58: DryLab resolution map with red for high and blue for low resolution and yellow/green the in between values and the chosen optimal conditions shown at the intersection of the lines.

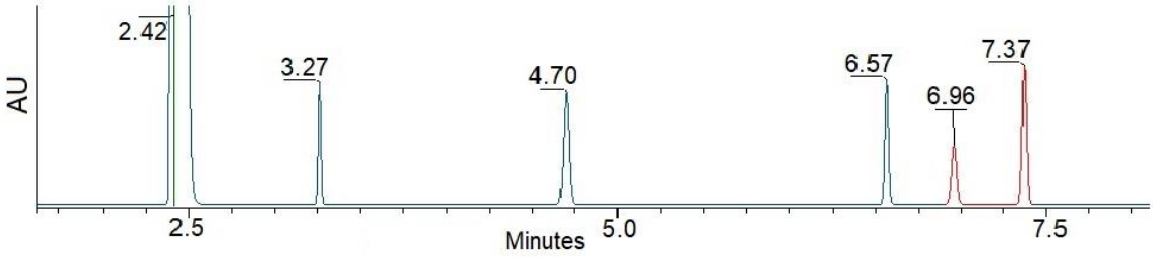

Figure 59: Simulated chromatogram of the calculated optimal conditions with DryLab.

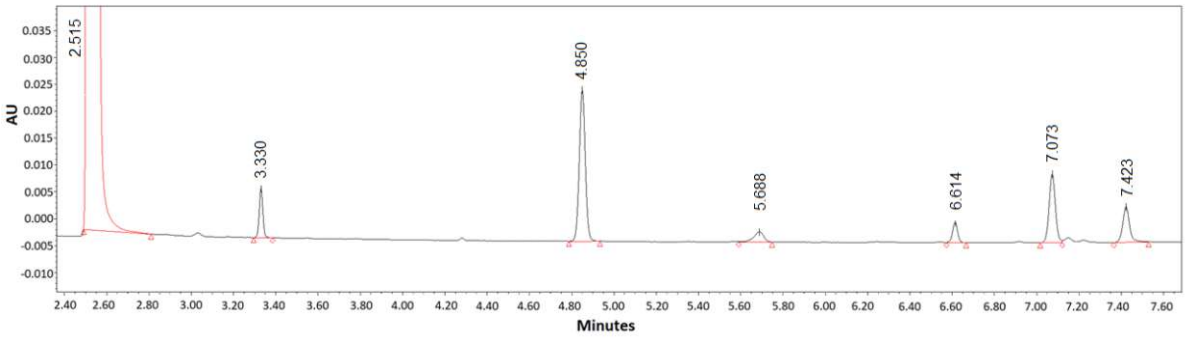

Figure 60: Experimental result of the in DryLab calculated optimal conditions on Waters Cortecs C18+.

4.10 Robustness studies

4.10.1 Initial risk assessment

For an initial risk assessment, all parameters with the most possible critical influence on the outcome of an UPLC-method were listed (see Table 20), not only for this procedure but generally. For the parameters, which are potentially critical in the examined method, ranges were defined. The bold variables are considered to have the highest or the largest influence on the results and were therefore examined further.

The ranges were defined after considering internal documents and official pharmacopeia guidelines. Since for the sample of this thesis no complex sample preparation was needed, the factors related to variability during sample preparation were not further examined.

Table 20: Possible critical parameters for HPLC methods, with the bold variables being considered as the ones, with the largest influence on the method. The crossed out are sections, which can be a big factor for other types of methods or samples, but in the example in this thesis, there was no complicated sample preparation.

Variable	**Range**
Column temperature	**± 3 °C**
Column (Batch)	
Flow rate	**± 3 %**
Injection volume	
pH - value	± 0.2
Gradient slope % B	**± 2 % B**
Buffer concentration	**± 10 %**
Detection wavelength	**± 2 nm**
Sample concentration	
Sample temperature	
Sample stability	
Sample preparation:	
Ultrasonic bath time	
Extraction time	

If there is a need to investigate the sample preparation, a separate DoE would be recommendable. The same accounts for testing the sample stability.

In the case of *1-(α-Aminobenzyl)-2-naphthol hydrochloride* a second DoE for assessing sample stability would have been appropriate, but for this thesis only the walkthrough of all steps was important, which is why the DoE for the robustness of the method parameters had a higher priority.

4.10.2 Design of Experiment

Based on the before mentioned initial risk assessment a Design of Experiment was generated. For the experimental conditions see section 3.14 and for the resulting outputs see Table 21 below.

In section 4.8 it is explained that the chromatogram was chosen even though two peaks where overlapping, due to an error at the data evaluation.

With differing the conditions, the two overlapping peaks, were sometimes separated notably. For a functioning DoE it was important to get the right peak tracking with right

peak areas. As a peak tracking of the two peaks, which sometimes seemed as one, could not be done sufficient enough and it was considered more important to take time testing the software than to rerun the whole development, the two peaks were considered as one peak for the DoE.

After measuring the experimental data and transferring it to Design Expert 11, every output needed to be analysed separately, but beforehand, two possible procedures of robustness study had to be considered.
In one case, all results are analysed and transferred to the optimisation software, even though it is clear, that results vary in a range where outputs are not critically changed. For example, when a resolution varies between 4 and 6, the variation is notable, but the resolution is never near the limit of 2, the set critical minimum criteria for the resolution.
In the second way of the robustness study, such results would not be considered in the DoE.
Both paths have advantages and disadvantages.
Considering all results means having all the data and all the influences of parameters on the outputs, if critical or not. Reducing them means influences, which may be considered mathematically significant, can be overlooked, when the excluded output remains within the ATP criteria.

The counterargument is that if not all results need to be processed through the software, the time expenditure is minimalized.
As both the ATP criteria for the resolution (see Table 21) and the %area (see Table 22) between the peaks were met with these experiments, except for the %area of peak 3, only the resolutions and the parameter influence on these were examined for representation.

The problem with assessing the %area is because of the already mentioned instability of the sample. Later throughout the experiments it became clear, that the sample is not stable over time. Therefore, the %area of the peaks does not only vary with the parameter changes but also with the stand time of the sample solution and with the made experiments, no statement about the real influence of the parameters on the %area can be made.
The data in the software was processed as explained in section 3.14.1.

Table 21: Resolution results of the DoE runs for the robustness study, with the compliance of the set ATPs as the goal.

Run	Resolution peak 1-2	Resolution peak 2-3	Resolution peak 3-4	Resolution peak 4-5	Resolution peak 5-6	Resolution peak 6-7
1	18.55	34.36	13.70	14.25	8.105	4.79
2	17.99	36.72	16.18	10.95	7.49	6.45
3	18.20	36.07	9.512	16.61	10.10	4.03
4	18.33	35.96	11.99	14.15	9.28	5.11
5	18.26	35.01	14.19	13.67	7.25	6.39
6	18.04	35.64	10.15	14.67	10.39	4.46
7	17.39	39.84	11.79	11.38	11.20	5.40
8	17.23	39.16	9.68	12.69	12.98	3.52
9	18.45	36.26	14.54	14.02	8.21	5.36
10	18.89	32.96	13.69	16.47	6.79	4.64
11	18.06	38.58	11.85	13.21	9.58	5.06
12	18.58	34.05	16.86	15.78	4.69	7.73
13	18.33	35.96	11.99	14.15	9.28	5.11

For the RSD of the areas the ATP was set as listed in Table 5, so the maximal RSD variance for the main peak was 1 %, for secondary components with a larger area of 1 % compared to the main peak was 5 %, with an area between 1 and 0.5 % it was 10 % and for areas smaller than 0.5 % compared to the main peak the RSD variance was set at 20 %.

The RSD was calculated with equation (12), the standard deviation *s* with (13) and the average $\bar{x}$ with (14).

Despite the instability of the sample, all peaks, except for peak 3 fulfil the ATP criteria. As already mentioned, it was not possible to determine the exact influence of the parameters on this output because of later discovered instability issues of the sample solution.

$$\% RSD = \frac{s}{\bar{x}} * 100 \quad (12)$$

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{(n-1)}} \quad (13)$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (14)$$

Table 22: RSD, standard deviation and average of the % areas of the peaks from the robustness study with Design Expert.

Peak	RSD	Standard deviation	Average (%)
1	0.11	0.1040	96.783
2	3.24	0.0095	0.29
3	5.56	0.0828	1.51
4	5.48	0.0086	0.16
5	6.17	0.0116	0.19
6	4.66	0.1532	0.74
7	9.85	0.0249	0.25

In Table 23 the p-values of the parameters on the outputs are given, which can be used to assess the influence of them on the resolution between peaks with the help of the flowchart in Figure 61. Only p-values lower than 0.1 show a mathematical significance, which is why they are highlighted in bold. Based on this first assessment, only the wavelength has no significant influence on the output, which was expected in this case, as the wavelength in an HPLC method should only influence the area of peaks not their positions.

The other parameters have a mathematically significant influence on the outputs. The next step in the flowcharts is to verify if the parameters are critically influencing the resolutions. For this the help of Design Expert was needed.

Table 23: P-values of the parameters influence on the resolution between peaks, with p-values lower than 0.1 showing a mathematical significance and therefore having a significant influence on the CMAs.

P-values	Gradient slope	Flow rate	Detection wavelength	Column temperature	Buffer concentration
Resolution 1-2	**0.0815**	**0.0041**	0.7153	**0.0007**	0.3117
Resolution 2-3	0.1698	**< 0.0001**	0.9616	**< 0.0001**	**0.0107**
Resolution 3	0.9776	0.5898	0.7204	**< 0.0001**	**0.0002**
Resolution 4-5	**0.0002**	**< 0.0001**	0.7711	**0.0199**	**< 0.0001**
Resolution 5-6	**0.0179**	**< 0.0001**	0.9267	**< 0.0001**	**< 0.0001**
Resolution 6-7	0.9385	0.3242	0.5478	**n.a.**	**0.0053**

The overlay plot shown in Figure 62 demonstrates how Design Expert can be used for a critical influence check. For every resolution, the criteria of minimum 2 has been set, so the plot is coloured grey in the regions where the criteria is not met and yellow if it is met.

The overlay plot axes are changed to the other parameters too and if they are all yellow as the one shown in Figure 62 it can be said that no critical influence can be seen on the outputs.

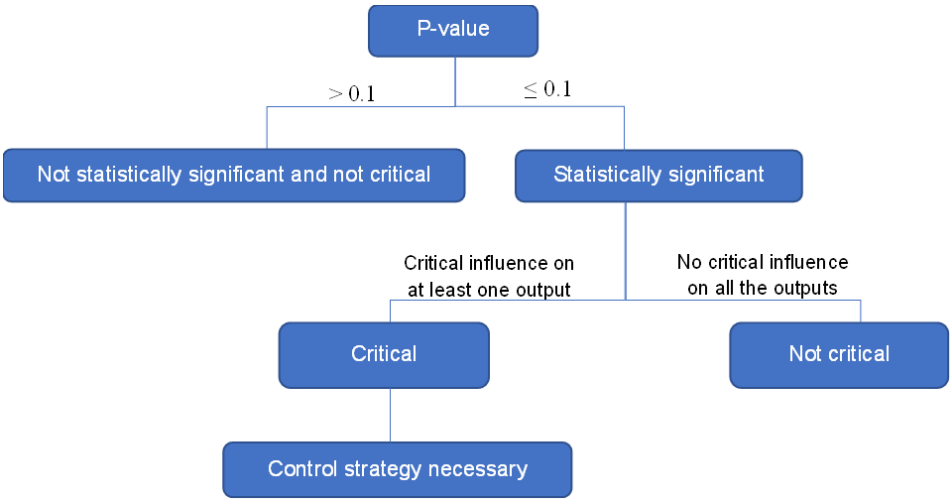

Figure 61: P-value flowchart as a decision-making aid for finding critical parameters.

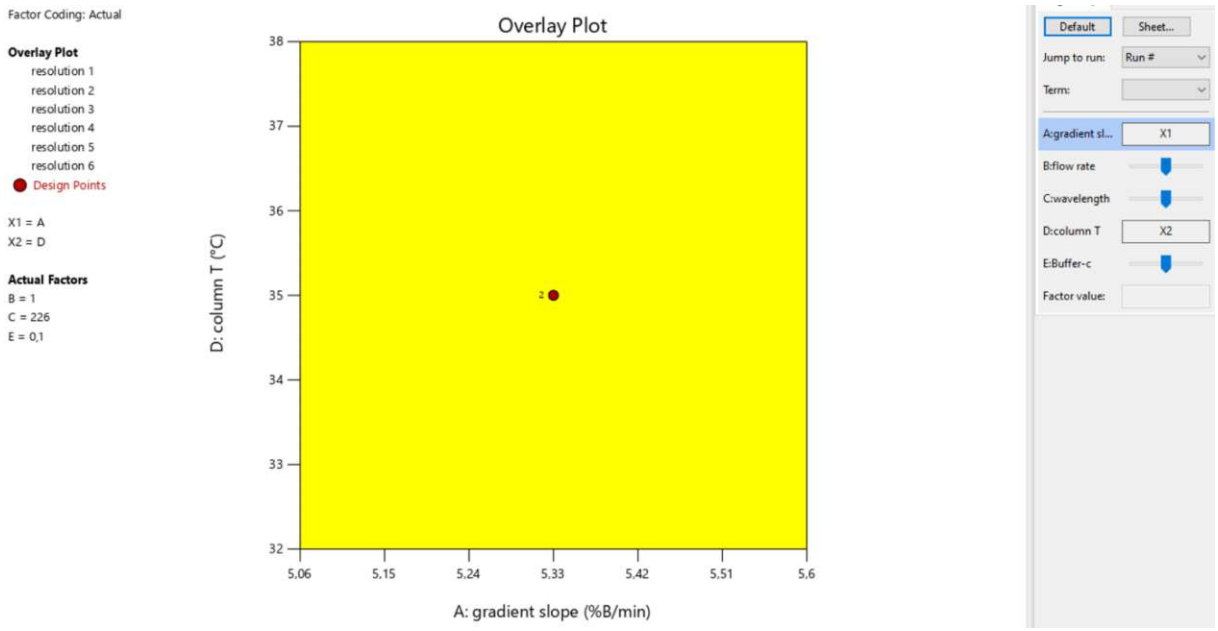

Figure 62: Overlay plot of the influence of gradient slope and temperature on the resolution, with the other parameters at their centre point.

The overlay plot can also be used to simulate the influence of the parameters outside of their measured values as is shown in Figure 63. Based on that, much higher column temperatures could get critical if the buffer concentration is simultaneously reducesd.

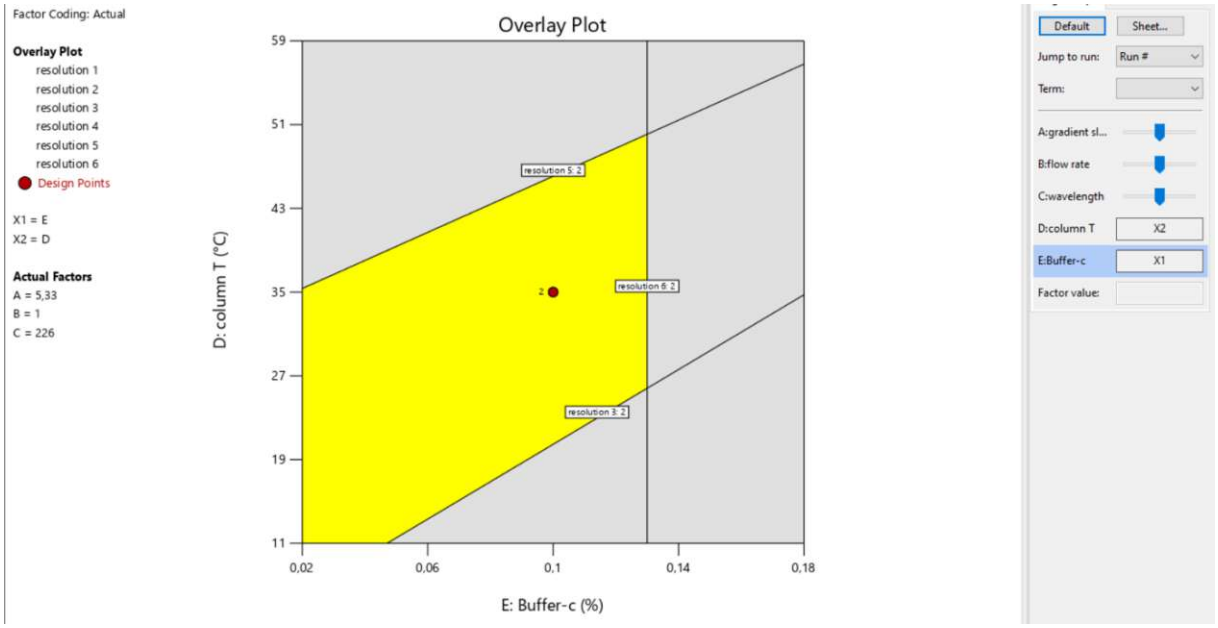

Figure 63: Overlay plot of the influence of column temperature and buffer concentration on the resolution with a range outside their measured values and the other parameters at their centre point.

4.11 Modification of the established method development

After performing a whole method development based on AQbD with the beforehand tested combinations of mobile phases, initial parameters and columns, some adjustments were made.

As the main goal was the guideline for a good method development procedure and not the perfect method itself, the modifications were not tested again for the exemplary sample but are listed here for a complete guideline of a method development.

For the final method development guideline tested in though this thesis the mobile phases are as explained in sections 0 to 4.10 with the same recommendations for a change to TFA or formic acid, when the good signal to noise relation at lower wavelengths is not needed or if the detection is changed to MS.

Adjustments were particularly made in the choice of columns. There were problems with the baseline of the Phenyl and PFP columns with the SPP phase. An example for the problem is shown in Figure 64, a chromatogram of the Thermo Scientific Accucore Phenyl-X with methanol and 0.1 % phosphoric acid as eluents.

After a thorough troubleshooting, where the system set up was scrutinised and it was also tested with a guard column, but the baseline problems were still present.

Also, the manufacturers of the columns were contacted, but their only solution was to run isocratic conditions instead of a gradient, which was no option for the method guideline.

Based on the earlier loading capacity testing, where the TPP and SPP columns performed similar, the phenyl and PFP columns with the TPP phase were chosen, because their baselines had no such problems.

That is why the final choice of the columns in the guideline for the method development procedure is as shown in the following section 4.12.

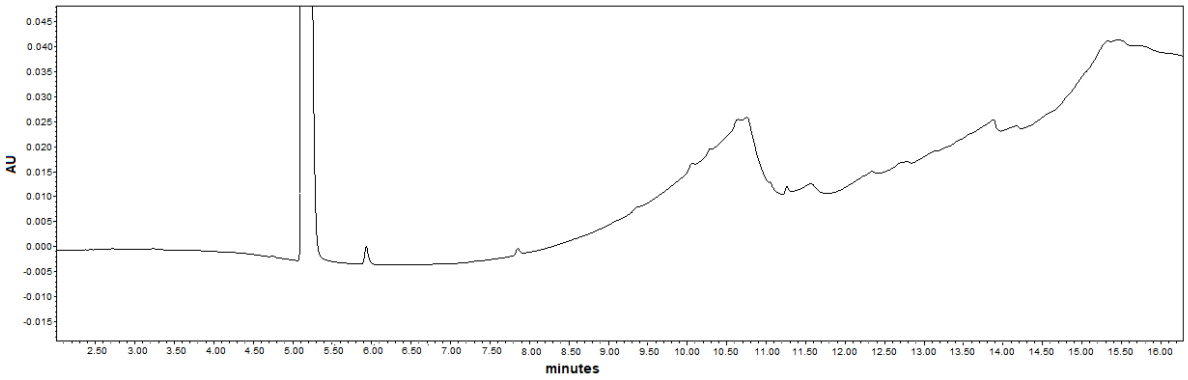

Figure 64: Chromatogram of the screening with Thermo Scientific Accucore Phenyl-X, methanol and 0.1 % phosphoric acid with an unusual baseline.

Throughout the procedure development explained in chapter 0, there are differences in Wavelength measurements and column temperature. The final values for the parameters are the range from 190-400 nm and a temperature of 40 °C.

The experiments made in the sections 3.2 to 0 are not included in the final procedure guideline, because they only have to be made once when the system set up is new or not at all, when there are given numbers for these tested parameters from the companies.

Only, if there are problems with the results these measurements can be made to check the performance of the system.

4.12 Result summary

This section summarises the final settings of the method development procedure discussed in this thesis.

It starts with the specification of the ATP and gathering know-how about the analyte. Then an initial sample preparation step needs to be established, which can be optimised later in the DoE. After an initial run, to adjust for example the concentration of the sample, the screening with the mobile phases in Table 24 and the columns listed below.

Table 24: Aqueous and organic eluents listed for the final method procedure guideline tested in this thesis

Aqueous eluents (pH-condition)	Organic eluents
0.1 % phosphoric acid in water (acid)	methanol
4 mM ammonium hydrogen phosphate (neutral)	acetonitrile
10 mM ammonium carbonate (basic)	

The final chosen columns for the method guideline developed throughout this thesis are as follows.

- Waters Cortecs C18+
- Agilent Poroshell 120 EC-C18
- Agilent Poroshell 120 Bonus RP
- Waters XSelect HSS PFP
- Waters BEH Phenyl

From the information provided in the screening a first choice for the conditions is made. Then they are optimised with DryLab and the final conditions are chosen.
In the last step a robustness study with Design Expert 11 is made around the chosen final conditions to identify critical parameters, for a better risk understanding in future routine measurements.

5 Conclusion

The analytical lifecycle defined in the (Analytical) Quality by Design is a rising theory on how to achieve a fast, precise, and robust analytical method procedure.

It is based on Quality by Design, which has already an important role in the (pharmaceutical) industry. The concept of QbD explained in ICH Q8 [2] as a systematic approach with the start of predefining objectives, using product and process understanding for a better process control and basing the knowledge on sound science and quality risk management, includes tools that help with a better development through DoE and risk assessment. This thesis includes a theoretical elaboration of mostly used definitions in QbD and their possible adaption to the analytical design. As the AQbD is still in the development phase, the definitions are not set finally. For example, in the guideline of the Analytical Life Cycle Management from USP [1] the method transfer is the last step of stage 2, whereas in other definitions it is the starting point of stage 3. Such questions are still discussed.

As explained, another important aspect is the Design of Experiment, where a lot of information can be processed more precisely but with fewer experiments than the traditional method procedure development approach.

There are already a lot of softwares for the establishment of DoE. This thesis used DryLab for finding optimal conditions and Design Expert 11 for robustness studies, but for the latter there are many more like JMP or Statease.

The type of software should depend on the department, which needs the DoE, for example the used DryLab had a focus on chromatographic methods, which therefore simulates known parameter interactions better or with less input runs than more general softwares.

But as this is a special case, most of the time for analytical methods not all interactions can be simulated and therefore, more precise DoEs have to be chosen, resulting in higher amounts of experiments. In this thesis, the theoretical elaboration of QbD together with the help of the softwares were used to gather fast and precise information for an analytical method and a guideline for a method procedure development was established.

Based on the results gathered, it can be concluded that the theory based on QbD is rising in its importance, as it is applied more and more often and also more frequently listed as a requirement. It is a good way to achieve the goal of a fast, precise, and robust analytical method procedure.

6 References

[1] USP, “PF Online,” 2020. [Online]. Available: https://online.usppf.com/usppf/document/GUID-35D7E47E-65E5-49B7-B4CC-4D96FA230821_10101_en-US?highlight=%E2%8C%A91220%E2%8C%AA%20Analytical%20Procedure%20Life%20Cycle. [Accessed 04 02 2021].

[2] Committee for Human Medicinal Products, *ICH guideline Q8 (R2) on pharmaceutical development,* 2017.

[3] Committee for Human Medicinal Products, *ICH guideline Q9 on quality risk management,* 2006.

[4] Committee for Human Medicinal Products, *ICH guideline Q10 on pharmaceutical quality system,* 2007.

[5] M. P. Gregory and e. al., “Proposed New USP General Chapter: The Analytical Procedure Lifecycle ⟨1220⟩,” 17 10 2016. [Online]. Available: https://www.uspnf.com/sites/default/files/usp_pdf/EN/USPNF/revisions/s201784.pdf. [Accessed 01 2020].

[6] R. E. Majors, “LC GC,” 01 10 2012. [Online]. Available: http://www.chromatographyonline.com/developments-hplcuhplc-column-technology?id=&pageID=1&sk=&date=. [Accessed 05 05 2020].

[7] The United States Pharmacopeial Convention, “USP,” [Online]. Available: https://apps.usp.org/app/USPNF/columnsDB.html. [Accessed 05 05 2020].

[8] L. Snyder, J. Dolan and P. Carr, “The hydrophobic-subtraction model of reversed-phase column selectivity,” *Journal of Chromatogrophy A,* p. 40, 2004.

[9] USP, “U.S. Pharmacopeial Convention,” 2020. [Online]. Available: https://apps.usp.org/app/USPNF/columnsDB.html. [Accessed 26 10 2020].

[10] K. Stavros, HPLC Made to Measure, Weinheim: WILEY-VCH Verlag GmbH & Co. KGaA, 2006.

[11] Waters, 2998 Photodiode Array Detector Operator's Guide/ Revision C, Waters Corporation, 2010.

[12] M. W. Dong, Modern HPLC for practicing Scientists, Wareham, Massachusetts: John Wiley & Sons, Inc, 2006.

[13] D. I. Molnár, “Molnár-Institute for applied chromatography,” [Online]. Available: https://molnar-institute.com/drylab/. [Accessed 26 02 2023].

[14] “Analytical Life Cycle Managment - The Coming Revolution,” *The Column,* 13 August 2018.

[15] J. Piriou, E. Bernard, M. Hertschuh and R. Olliver, “Control Strategy as the Keystone of the Product Lifecycle, from Product/Process Understanding to Continuous Process Veri□cation and Improvement,” *PHARMACEUTICAL ENGINEERING®,* January 2012.

[16] E. Kovacs, J. Ermer, P. L. McGregor, P. Nethercote, R. LoBrutto, G. P. Martin and H. Pappa, “Analytical Control Strategy,” *Stimuli to the Revision Process,* 16 Ocotber 2016.

[17] Waters, Acquility UPLC - Sample Manager - Flow Through Needle Operator's Overview and Maintenance Information, Waters Corporation, 2010-2011.

[18] Waters, Acquility UPLC H-Class - Quaternary Solvent Manager Operator's Overview and Maintenence Information, Waters Corporation, 2010.

[19] Waters, Acquility UPLC H-Class - System Guide, Waters Corporation, 2010-2016.

[20] “Advanced Materials Technology,” 2023. [Online]. Available: https://halocolumns.com/product/halo-90-a-pfp/?attribute_pa_part-size=2-7-%25c2%25b5m&attribute_pa_inner-diameter=3-0-mm&attribute_pa_column-length=100-mm&attribute_pa_format=analytical. [Accessed 23 01 2023].

[21] “Agilent Technologies, Inc.,” [Online]. Available: https://www.agilent.com/en/product/small-molecule-columns/reversed-phase-hplc-columns/infinitylab-poroshell-120#productdetails. [Accessed 23 01 2023].

[22] “Waters Corporation,” 2023. [Online]. Available: https://www.waters.com/nextgen/at/de/shop/columns/186007402-cortecs-c18-column-90a-27--m-3-mm-x-100-mm-1-pk.html. [Accessed 23 01 2023].

[23] “Thermo Fisher Scientific Inc.,” 2006-2022. [Online]. Available: https://www.thermofisher.com/order/catalog/product/27926-154630?SID=srch-srp-27926-154630. [Accessed 23 01 2023].

[24] “Phenomenex Inc.,” 2023. [Online]. Available: https://www.phenomenex.com/Error/NoPage?aspxerrorpath=/Products/Kinetex-hplc-column/Kinetex-EVO-C18. [Accessed 23 01 2023].

[25] J. DeStefano, T. Langlois and J. Kirkland, “Characteristics of Superficially-Porous Silica Particles for Fast HPLC: Some Performance Comparisons with Sub-2-µm Particles,” *Journal of Chromatographic Science, Vol. 46,* p. 7.

[26] B. Davis, L. Lundsberg and G. Cook, “PQLI Control Strategy Model and Concepts,” *Product Quality Life Cycle Implementation (PQLI) Innovations,* 17 May 2008.

[27] O. Bhusnure, S. Gholve, R. Suryawanshi, R. Sugave and J. Sangshetti, “Life cycle assessment (LCA) approach to analytical method development: a review,” *World Journal of pharmacy and pharmaceutical sciences,* 24 October 2015.

[28] M. K. Parr and A. H. Schmidt, “Life cycle management of analytical methods,” *Journal of Pharmaceutical and Biomedical Analysis,* 17 June 2017.

[29] N. Raman, U. R. Mallu and H. R. Baptu, “Analytical Quality by Design Approach to Test Method Development and Validation in Drug Substance Manufacturing,” *Hindawi Publishing Corporation, Journal of Chemistry,* 01 January 2015.

[30] S. Kromidas, HPLC Made to Mesure, Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA, 2006.

7 Appendix

Table 25: F-values of the tested TPP columns (with a length of 100 mm, an inner diameter of 2.1 mm and a pore size in the range between 1.7 and 1.9 µm) top down at pH 2.8 with a base present in the sample, with an acid present in the sample, pH 2.8 without acid or base present in the sample and at pH 6.7 with a base present in the sample.

Columns	Intersil ODS-3	Bonus RP	HSS PFP	C18 ExRS	BEH Phenyl	Sustain C18
Acquity UPLC CSH C18	45.23	253.1	60.46	21.06	22.45	14.61
	32.44	34.33	38.7	44.01	23.08	41.71
	6.65	23.66	22.74	12.53	12.21	8.21
	41.57	117.51	130.62	35.19	36.51	29.17
Inertsil ODS-3		209.22	102.96	28.52	63.94	32.81
		63.45	21.16	21.21	13.93	9.71
		26.62	21.15	6.33	6.81	3.28
		79.41	170.98	10.32	75.75	13.45
Zorbax Bonus RP			311.26	237.04	272.49	241.49
			71.92	74.41	53.66	71.37
			44.88	32.22	30.03	27.7
			247.83	88.31	152.5	91.42
HSS PFP				76.88	44.76	72.28
				26.57	27.53	26.01
				24.24	24.91	24.19
				163.32	97.45	158.47
Triart C18 ExRS					36.19	6.91
					23.24	5.1
					5.45	4.93
					67.52	6.95

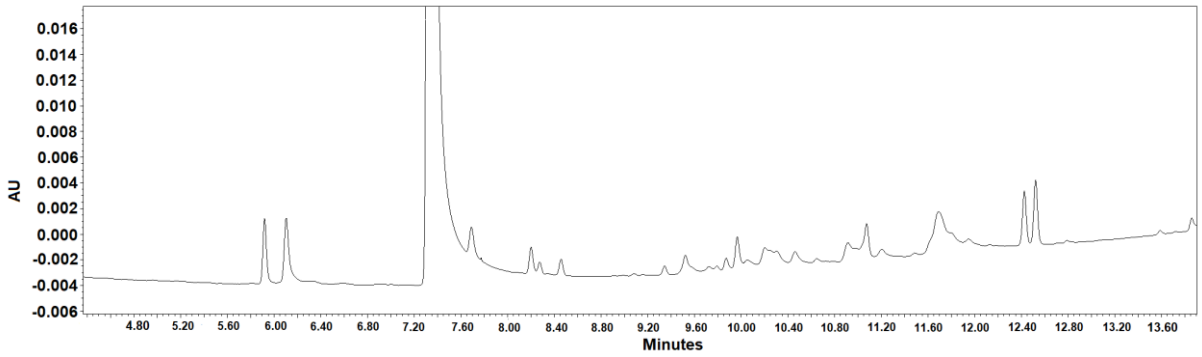

Figure 65: Chromatogram of the sample screening in section 3.12 on the column Bonus RP with a pH value of 6.7 and acetonitrile (zoomed).

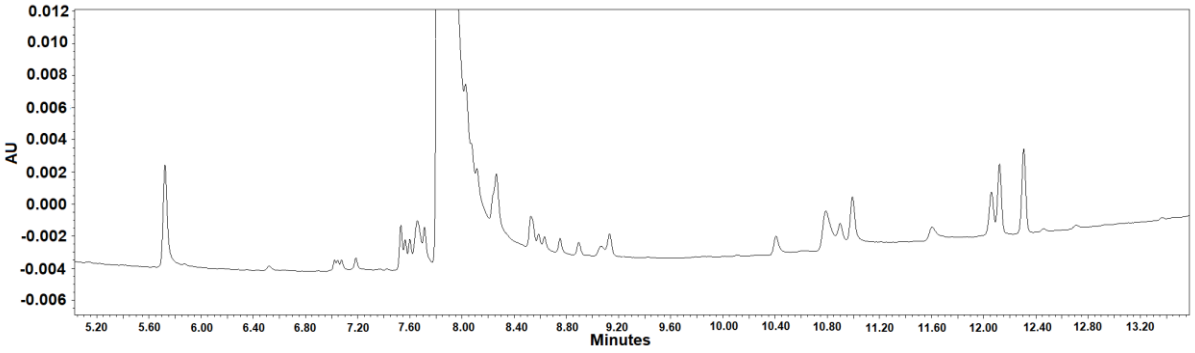

Figure 66: Chromatogram of the sample screening in section 3.12 on the column Waters Cortecs C18+ with 6.7 and acetonitrile (zoomed).

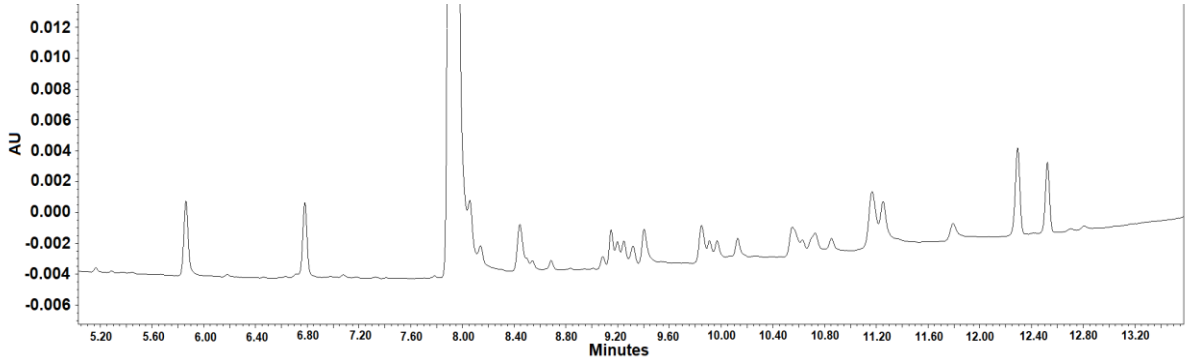

Figure 67: Chromatogram of the sample screening in section 3.12 on the column EC-C18 with 6.7 and acetonitrile (zoomed).

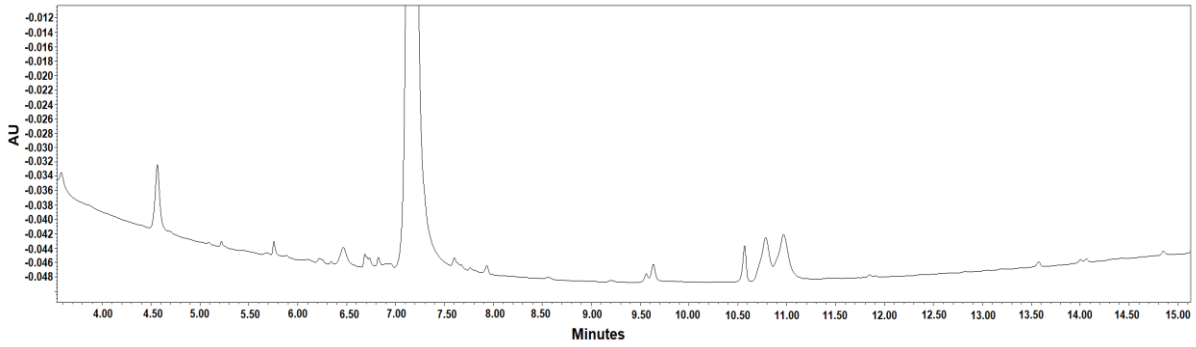

Figure 68: Chromatogram of the sample screening in section 3.12 on the column Phenomenex Kinetex C18 EVO with a pH of 10 and acetonitrile (zoomed).

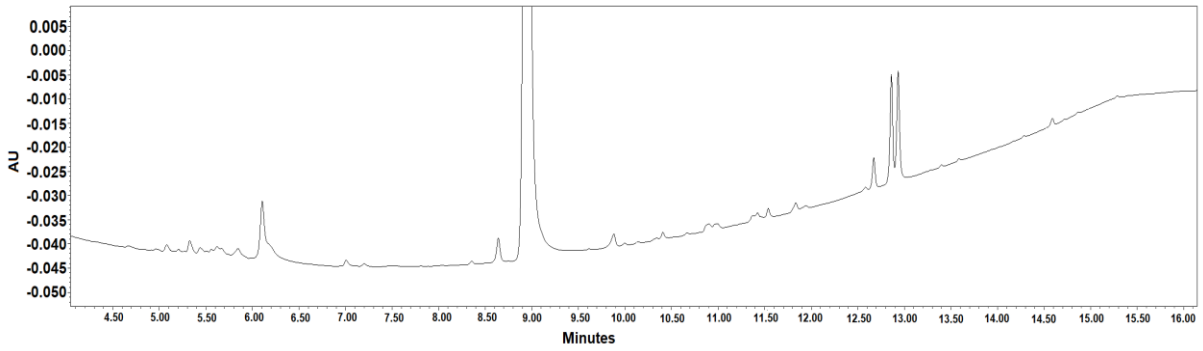

Figure 69: Chromatogram of the sample screening in section 3.12 on the column Phenomenex Kinetex C18 EVO with a pH of 10 and methanol (zoomed).

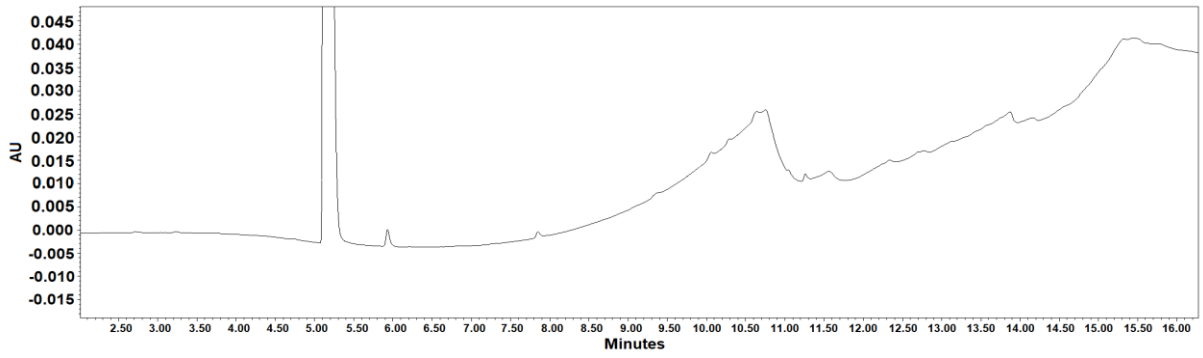

Figure 70: Chromatogram of the sample screening in section 3.12 on the column Phenyl X with a pH of 2 and methanol (zoomed).

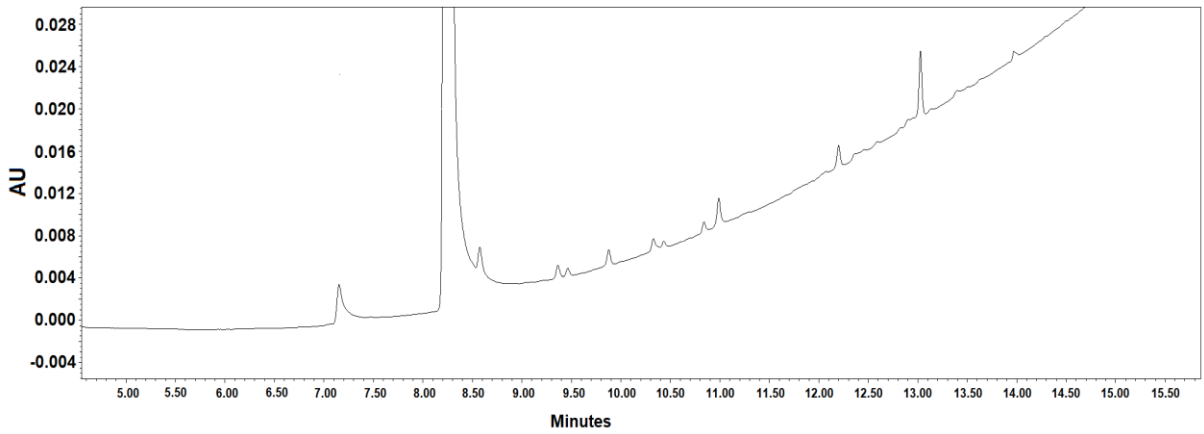

Figure 71: Chromatogram of the sample screening in section 3.12 on the column Bonus RP with a pH of 6.7 and methanol (zoomed).

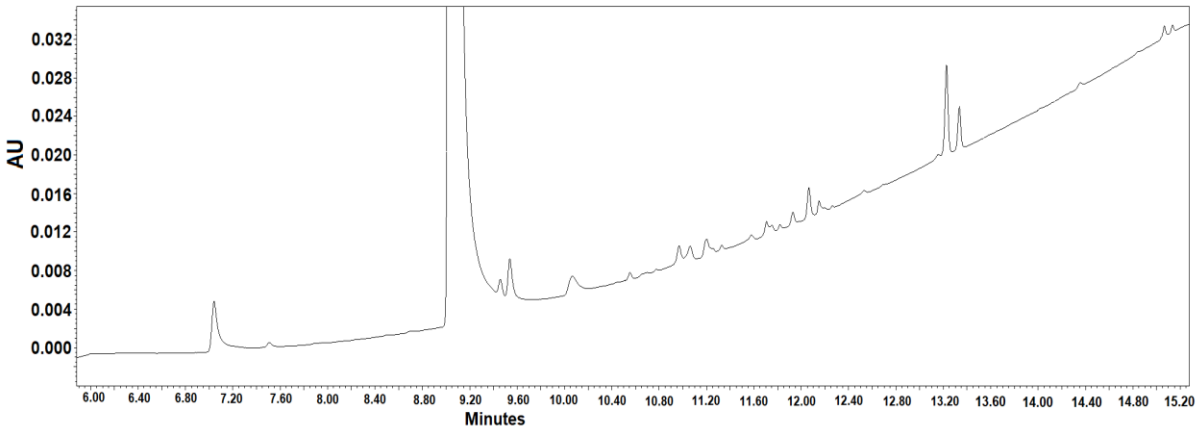

Figure 72: Chromatogram of the sample screening in section 3.12 on the column Waters Cortecs C18+ with a pH of 6.7 and methanol (zoomed).

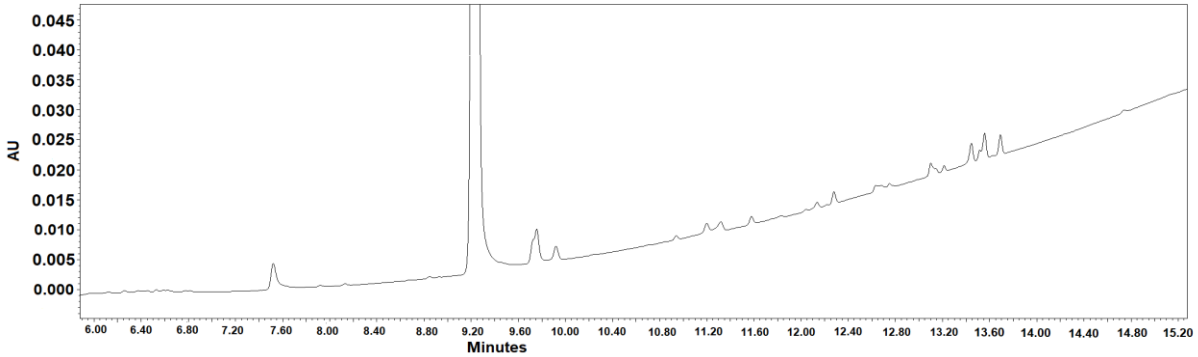

Figure 73: Chromatogram of the sample screening in section 3.12 on the column EC-C18 with a pH of 6.7 and methanol (zoomed).

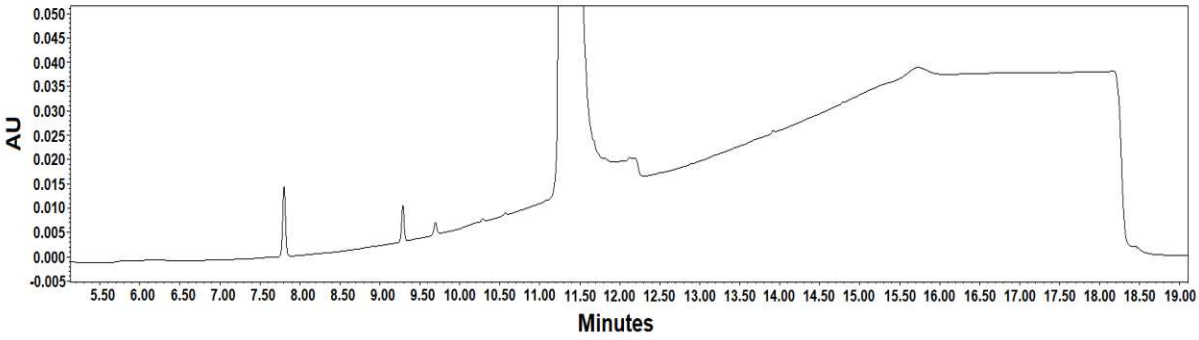

Figure 74: Chromatogram of the sample screening in section 3.12 on the column PFP with a pH of 6.7 and methanol (zoomed).

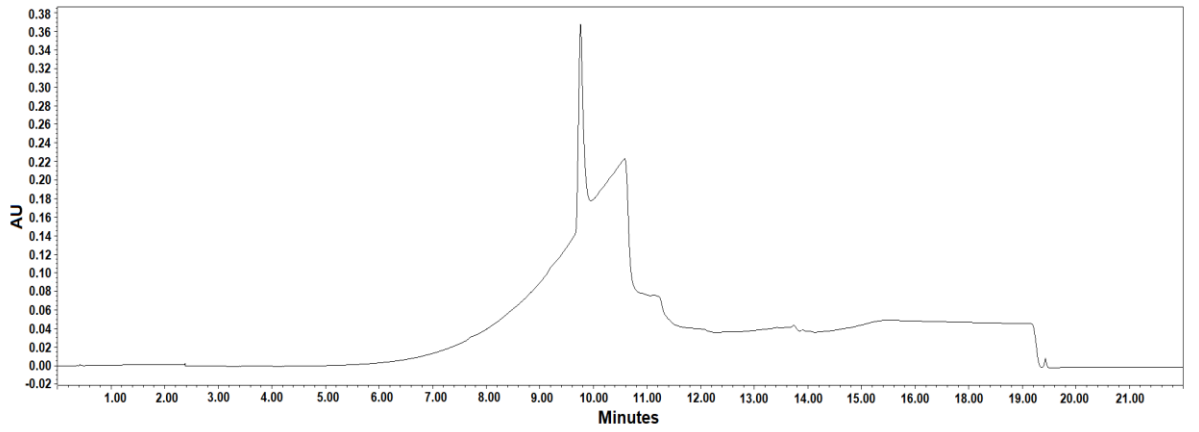

Figure 75: Chromatogram of Phenyl X with the sample screening in section 3.12 on the column with a pH of 6.7 and methanol (zoomed).

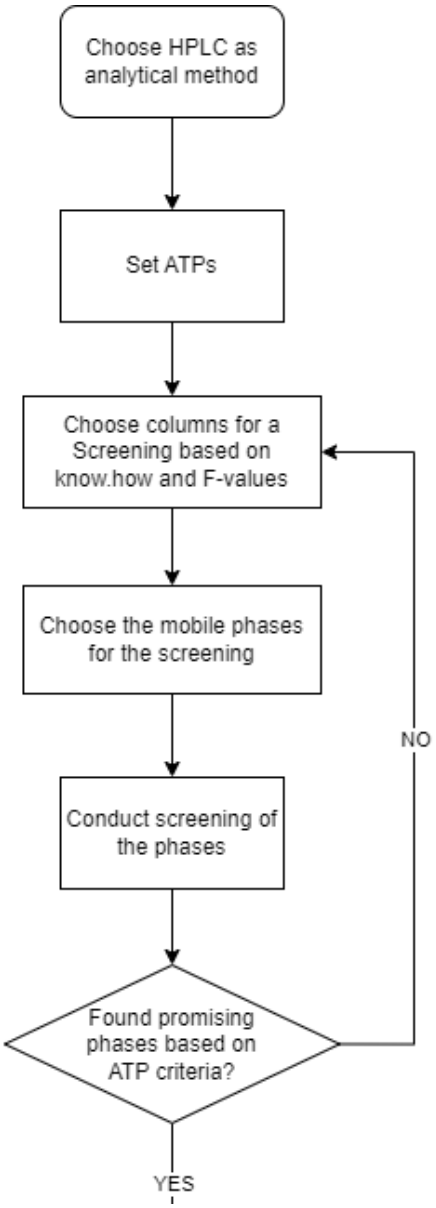

Figure 76: Flow chart as a guideline for a method development based on the Quality by Design approach part 1.

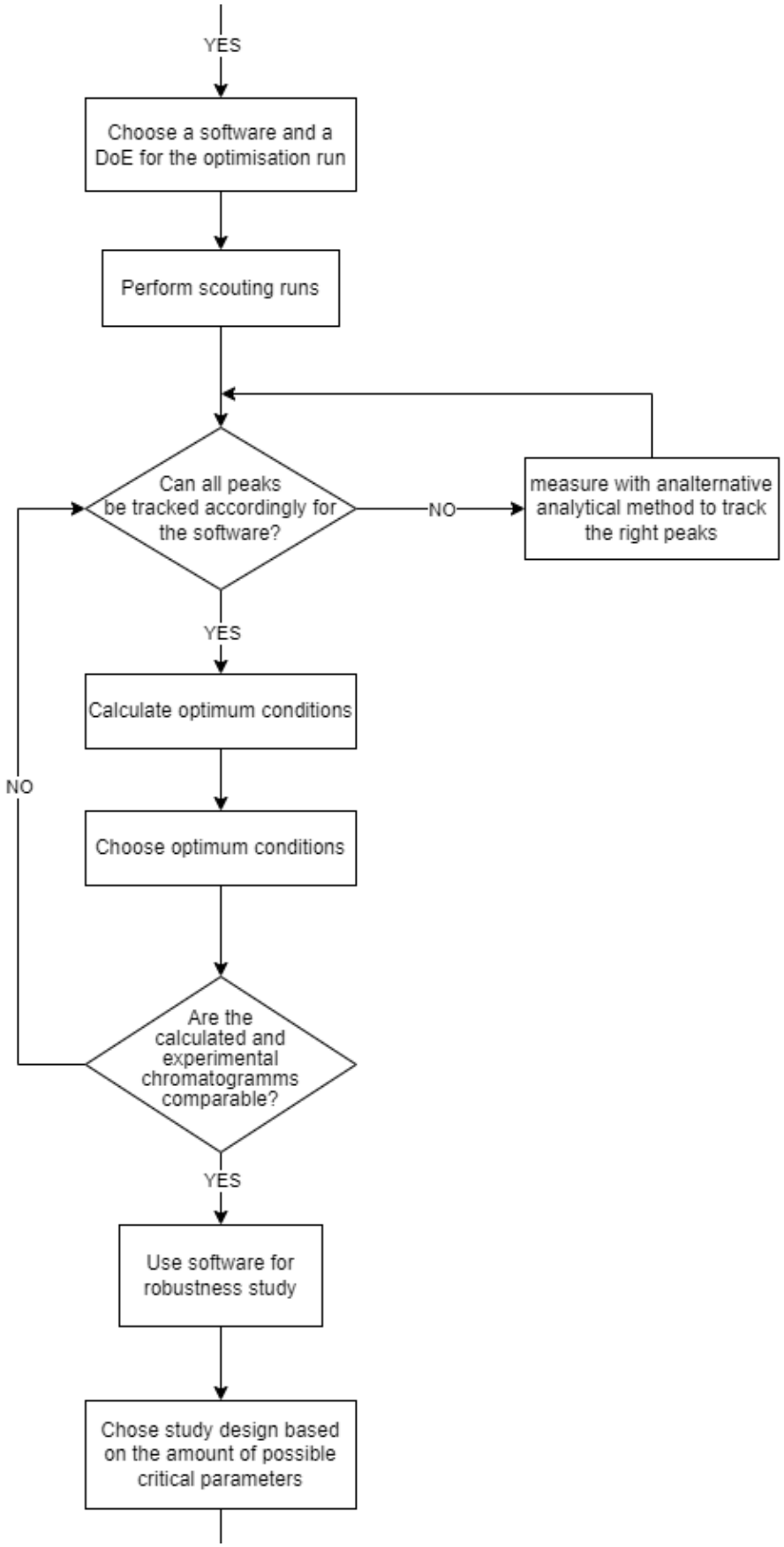

Figure 77: Flow chart as a guideline for a method development based on the Quality by Design approach part 2.

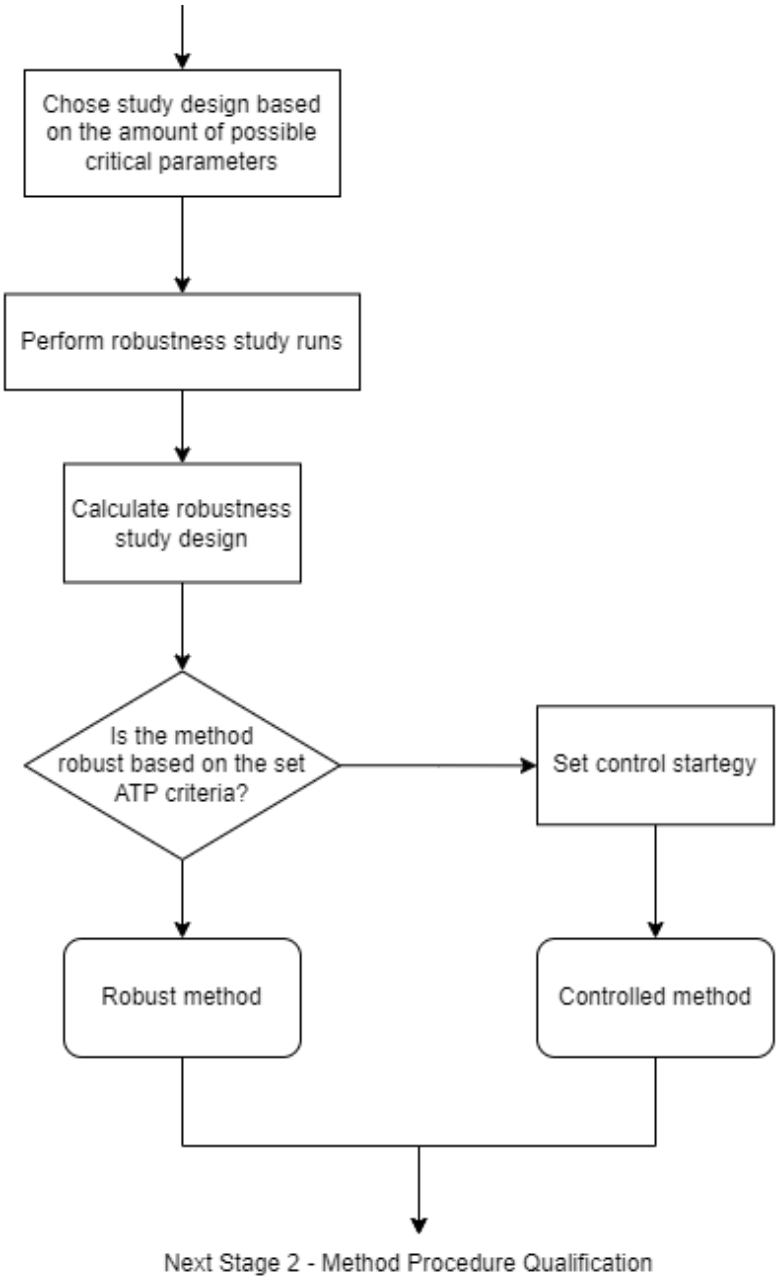

Figure 78: Flow chart as a guideline for a method development based on the Quality by Design approach part 3.