Chromatography Modelling in High Performance Liquid Chromatography Method Development

Introduction
Computer supported chromatography method development [1,2] started around 1980 as IBM released the so called IBM PC, the first "Personal Computer". The members of the project were Lloyd Snyder, John Dolan, Tom Jupille, founders of LC Resources on the US West Coast and myself, who, after returning from Csaba Horváth's Lab at Yale University founded the Institute for Applied Chromatography, located in Berlin-Kreuzberg on the 1st October 1981. The 4 of us decided to use the new technology of IBM-computers to write a program for HPLC method development. Jack Kirkland, a pioneer of HPLC at DuPont measured different properties on 1000 columns, and with Lloyd Snyder calculated the influence of the pore structure and ligand length in order to model band spreading. This was the beginning of “DryLab”, a name which Lloyd Snyder suggested for the software.

In 1988, the first iteration of the DryLab software was developed and allowed the modelling of band spreading, during optimisation of isocratic %B in DryLab I (i-isocratic) [4]. In 1989 modelling of gradient elution DryLab II (G=gradient) was developed [5]. First chromatograms for visualisation were plotted with *-characters. A few months later we were able to plot chromatograms for every change in experimental conditions. In the following years the software was further improved to isocratic multiparameter software, called “DryLab Imp”, where the user could model changes in pH, temperature, ionic strength, ternary eluent composition and ion-pair.

Optimisation of isocratic %B in DryLab I suggested for the software. The major feature of these models was their simplicity and visuality [6]. A very informative book on computer assisted method development was published in 1990 by Glajch and Snyder with 42 contributions to the theory and praxis of HPLC modelling [1] illustrating the work of leaders of the chromatographic scientific community working on separation predictions.

Some years later Sergej Galushko started his project, which he first named ‘ChromDream’ which was later renamed ‘ChromSword’. The software allowed the prediction of retention time based on a compound molecular structure which is important for those working in drug design [7,8]. To run the experiments he later introduced “AutoChromSword” software which collected runs overnight in an automated fashion.

Other companies also introduced similar software packages. Agilent developed ICOS (intelligent computer optimization software) [9]. In France “Orion” was developed by the group of Heinisch, Rocca and Tschapla [10]. In Canada Mike McBrian introduced an optimisation software for chromatography with ACDLabs (Advanced Chromatography Development) [11]. During this time programs like ‘Diamond’ and ‘PESOS’ came and went. Around 2003 the company 5-Matrix introduced “Fusion”, software which controlled Waters instruments to generate experiments and evaluate them according to statistical principles. This list is not complete and there were other software packages developed during this time, but these are beyond the scope of this article.

Theory of RPC Modelling
Retention phenomena of reversed-phase chromatography (RPC) are described in many ways by different authors. The philosophy used in DryLab is described in the “Solvophobic Theory” of Csaba Horváth, which was developed in the years 1975-1977 at Yale [12]. The fundamental concept of this theory is that retention in RPC is enforced by water, as the retarding component of the eluent. The uptake (dissolution) of nonpolar molecules in the water structure requires large amounts of energy. The retention factor k (also called the “capacity factor”) is proportional to the energy needed in this process. In the case of dibenzanthracene on a C8-phase, we find the following values for the capacity factor: k in water $k_W \approx 4000$ k in acetonitrile $k_{ACN} = 1$

Horváth and his team found that the only possible explanation for this extremely wide scale of retention times is the change in the surface tension of water altered by the
addition of acetonitrile (AN) or methanol (MeOH). The strong lipophilicity of water can easily and continuously be reduced in this way, which is what occurs in gradient elution. Thus a typical approach to method development in RPLC is to initially run a scouting gradient on a C18 column, which will typically resolve more than 95% of all compound peaks present in the sample.

Gradient elution typically starts with water or water-rich eluents. Upon injecting the sample into such a mobile phase (eluent), the water mixes with the hydrophobic sample components and forces them onto the surface of the C8 or C18 column packings. The capacity factors of organic molecules in water (Kw) are 10⁻⁶⁻¹⁰⁻⁷ times higher than in acetonitrile or methanol. By increasing the amount of the organic eluent, the retention force from water will become weaker, the surface tension of the eluent is reduced from 72 dyn/cm in water to approximately 22 dyn/cm at room temperature with a reduction in retention time occurring at the same time. This process has tremendous capabilities for separating complex mixtures in a highly reproducible manner for simultaneous qualitative and quantitative analysis.

In gradient elution, we can calculate the retention precisely for every component. Based on only two gradient runs, we can further calculate isocratic conditions and see how the k-values are reduced with increasing %B (percent organic) in the mobile phase.

The amazing ease of Reversed-Phase gradient elution is exhibited in the continuous reduction of the retention force of water by the increasing amount of the organic eluent (MeOH or AN). Fine differences in accessible solvophobic molecular surface areas, consisting of C-C, C- H and other nonpolar atomic bonds, combined with steps in the gradient, are sufficient to achieve reasonably good separations with almost any mixture in life science applications.

Modelling of Reversed-Phase separations is based on the measurement of both the retention time and the peak area [13,14]. The calculation of sample positions in the corresponding chromatograms in a Design of Experiments (DoE) enables the chromatographer to look at a small number of experiments in a virtual mode and generate a fast overview of separation choices. However by running a DoE, e.g. a 3-G-T model with 4 runs, we must realise that each chromatogram will look different. This however is the purpose of the exercise, as we want to learn how peaks move, so we can establish a model and can derive solutions for separation problems.

Experimental conditions

Column selection should be done carefully. We have a great number of RP-columns on the market. Snyder, Dolan, Carr, Engelhardt, Emeby, Tanaka and Peterson among others published excellent papers on column selectivity [15,16] including more than 500 columns and demonstrated how to select the best columns for a separation. We used a YMC C18 100Å column, 150 x 4.6 mm, 5 µm (Waters, Milford, MA, USA) with a synthetic sample mixture developed for column testing at a flow rate of 2.0 mL/min. A Shimadzu Prominence (Shimadzu Europe, Duisburg, Germany) LC with dwell volume Vd 0.4 mL and UV detection at 254 nm was used throughout the work. Modelling software was DryLab®4, v.4.0.10.15. (Molnar-Institute, Berlin, Germany). Eluent A was 0.025 M phosphate buffer at pH 2.8. Eluent B1 was acetonitrile (AN) Eluent B2 was methanol (MeOH)- and a 50:50 mix of B1: B2. Gradient times were 20 and 60 min from 5 to 95% (B1+B2) at T1: 30°C and T2: 60°C.

One and 2-dimensional models

If a chromatographer wants to understand peak movements caused by changes in experimental parameters, they must keep everything constant except one factor, one eluent B (or pH or one factor at a time, OFAT). This helps to understand how a separation might change. Initially, this may appear to be an inefficient approach to spending time, however the opposite is true, since the chromatographer understands the separation better.

Changes of other parameters can also be modelled in DryLab by calculation. The influence of the flow rate, of the column length and ID, dwell volume, gradient start and end, steps, etc. So even an OAT-model in DryLab allows the understanding of multifactorial changes. The most successful 2-dimensional model was and is still today the gradient time – temperature- or G-T-model, especially when combined with a ternary gradient elution technique [17]. The G-T-model which was used by Snyder in column characterisation [18], has lead to an extension into a 3D-resolution space, the Cube [14].

3-D-Models, the Cube

The first Cube model was demonstrated at the HPLC2009 conference in Dresden [14]. Soon afterwards, a number of papers appeared demonstrating the advantages of this new technology for industrial applications [19-23]. This new technology is especially well suited to improving the speed of older pharmacopeial methods, as shown by Schmidt, where they reduced a method's analysis times from 160 to 3 min, using DryLab and UHPLC [24].

The first step in this process is to plan a Design of Experiments (DoE) followed by the so called Peak Tracking process. The most efficient DoE is shown in Figure 1. (Figure 1).

Peak Tracking

One and 2-dimensional models

If a chromatographer wants to understand peak movements caused by changes in experimental parameters, they must keep everything constant except one factor, one eluent B (or pH or one factor at a time, OFAT). This helps to understand how a separation might change. Initially, this may appear to be an inefficient approach to spending time, however the opposite is true, since the chromatographer understands the separation better.

Changes of other parameters can also be modelled in DryLab by calculation. The influence of the flow rate, of the column length and ID, dwell volume, gradient start and end, steps, etc. So even an OAT-model in DryLab allows the understanding of multifactorial changes. The most successful 2-dimensional model was and is still today the gradient time – temperature- or G-T-model, especially when combined with a ternary gradient elution technique [17]. The G-T-model which was used by Snyder in column characterisation [18], has lead to an extension into a 3D-resolution space, the Cube [14].

3-D-Models, the Cube

The first Cube model was demonstrated at the HPLC2009 conference in Dresden [14]. Soon afterwards, a number of papers appeared demonstrating the advantages of this new technology for industrial applications [19-23]. This new technology is especially well suited to improving the speed of older pharmacopeial methods, as shown by Schmidt, where they reduced a method's analysis times from 160 to 3 min, using DryLab and UHPLC [24].

The first step in this process is to plan a Design of Experiments (DoE) followed by the so called Peak Tracking process. The most efficient DoE is shown in Figure 1. (Figure 1).

Peak Tracking

Figure 1: Design of Experiments (DoE) for the simultaneous optimisation of gradient time (tG), temperature (T) and pH of the eluents A or the ternary composition IC of the eluent B (i.e., B1+B2) in B1 with B2: AN and B3: MeOH). The pH is changed by having two aqueous eluents A1 and A2 with changing ratios. Circles represent the twelve input experiments for the 2-D model. The short gradient time tG1 is at the points 1, 5, 9, 3, 7 and 11, and the long gradient time tG2 at 2, 6, 10, 4, 8, and 12. In the present example IC was 20 min and IC2 was 60 min (5 to 95%). T1 was 30°C and T2 was 60°C. The composition of the organic eluent (B1+B2/ternary composition, IC) was varied between 0, 50 and 100% MeOH in AN.

Peak Tracking is an important step in method development, as most chromatographers using a method are afraid of unexpected changes. Therefore small variations in working conditions should be carried out to test method robustness. The question is however, “How much should we vary a parameter?”
If we change a parameter by very little, then we might not see hidden peaks. Therefore larger changes are required, e.g., two gradient times $t_{G1}$ and $t_{G2}$ with a factor 3 difference. In temperature optimization experiments we should have a difference 30-40°C and with pH, 0.6 pH units over 3 (or more) runs.

With these experiments we can create an experimental design with 4-12 runs, which is sufficient in most cases. We should learn as much as possible with the least possible number of runs.

It is widely accepted, that the so called tG-T-model is the best one to start with. It has only 4 runs and it allows simple peak tracking as shown in the following figures.

Initially the order of elution is established at the experimental points 2, 6 and 10 (see Figure 2).

A peak tracking table of a tG-T-tC (tC = ternary eluent composition) model showing different elution profiles of the same mixture of 18 compounds in fewer than 12 different conditions [14]. The peak areas in those runs have a standard deviation of ca. 2% on average and can therefore efficiently be used to track moving peaks and establish robust conditions for routine applications.

The next step is to align the 12 runs in the 3 tG-T-sheets. This is a process of looking at peak movements, peak overlaps and peak turnovers. Peak identification is based on peak areas, which represent the injected amount of the sample. Keeping this constant we get constant peak areas for a given compound in every run. Peak areas (concentration x volume = mass) are well suited to identify a peak. In peak overlaps the areas are additive. In Figure 3. we show the runs 1-2-3-4, where the organic eluent B1 is AN. Note the selectivity differences between the runs.

Then the peaks of the experiments 5-6-7-8 are aligned (Figure 4). Again there are different selectivities generated and several coeluting peak pairs observed. At the end the last sheet of runs 9-10-11-12, which is the 100% methanol-sheet, all peaks are fully tracked (Figure 5).

When peak tracking is complete, we then calculate between the 3 core sheets another 97 sheets, filling out the total space so we can simulate any chromatogram at any point in that whole space with more than 10^6 virtual chromatograms. The results are highly precise, up to 99.8% accuracy in retention.
times, which is comparable to the operational accuracies of most UHPLC instruments.

Method adjustments are much easier to implement when utilizing resolution maps, as alterations of the “set point” or “working point” in the Design Space are not considered to be changes with post-regulatory approval. This means, that changes in the Design Space (Figure 7) are possible without re-validation, allowing a much greater flexibility in the lab than in previous years.

From Figure 7 we can define several Design Spaces. The extension of the red areas (the possible Design Spaces) will give us a first idea about the robustness. We could also find suitable method parameter in methanol (front sheet of the cube in Figure 7 as well as in acetonitrile (back sheet in Figure 7). From the design space as defined in Figure 7 we can get robustness information only for the measured parameters: Gradient time, pH, and IC (%) where B1 is acetonitrile and B2 is methanol. However, as DryLab®4 is able to calculate other changes which might occur at the same time, we can calculate the influence of additional parameters like flow rate or start- and end-%B of the gradient. No additional experiments are necessary for this kind of robustness calculation. The result is shown in Figure 8.

We can see in Figure 8, on the top of the graph the selected method parameter (tG = 46 min, T = 30°C and tC = 100 % MeOH as organic eluent) with estimated possible deviations from the nominal value. The temperature is assumed to deviate from the nominal value of 30°C by not more than ±2°C, i.e. the true temperature is assumed to be in any experiment between 28 and 32°C.

On the left graph the ‘Frequency Distribution’ shows how often (N) a certain critical resolution (Rs,crit) occurs under any combination of possible, true parameter values. As can be seen from the graph, the failure rate, i.e., the number of experiments that could fall outside the required critical resolution Rs,crit = 1.5, is zero. That means that practically all experiments should fulfill the critical resolution requirement. The position of the “set point” or “working point” is of great importance, as many experiments cost enormous amount of resources. If the point is selected by trial and error, an analyst may have to change it and repeat a large number of experiments to find a new optimum. DryLab can calculate 6 experimental factors at 3 levels, i.e., P = 729 experiments in less than 1 minute!

The right graph in Figure 8 ('Regression Coefficients') describes the importance of each parameter, related to the selected deviation from the nominal value, for the critical resolution. As can be seen from the graph, temperature has the most important influence; a lower temperature gives a higher global resolution.

DryLab and the QbD movement

In 2002, the FDA instigated the development of the QbD concept which allowed more flexibility in industrial laboratories [14, 15]. DryLab demonstrated as early as 25 years ago, that systematic experimentation in HPLC is required and has contributed to the development of Quality by Design in the analytical chemistry area. It was the first software demonstrating ‘robust resolution maps’, allowing the estimation of tolerance limits for every important parameter of a separation. DryLab is therefore an important tool to help meet QbD practices.
Method transfer

Method transfer is a problem in a global economy, where products travel over borders and are used in different locations to generate the same analytical result. It is necessary to enhance this process using modelling software to ease the burden using virtual UHPLC models. This method transfer process is often instrument dependent and therefore it is important to understand how to utilize predictive software in method transfer. An example of successful method transfer using this approach is demonstrated in reference [23].

The so-called Knowledge Management Protocol, which was discussed above in a short format, is a great help in dealing with regulatory authorities. In this way methods can be developed in an inspection-safe manner.

Economic considerations of modelling in reducing waste

In a steadily growing number of publications, the value and usability of retention modelling for fast and systematic method development has been demonstrated [17-22].

During the acetonitrile shortage, it was difficult to work in the HPLC lab as acetonitrile was not readily accessible. The development of the ternary Cube was important in this situation, and demonstrated that in most cases an alternative method utilizing methanol could be implemented instead of acetonitrile [18]. We can contribute to a green chemistry by reducing waste through computer modelling and reduce our environmental impact by reducing the volume of mobile phase waste.

As we can see, there is not much difference between both methods, but the method in Figure 6 using MeOH as eluent B is more environmentally safe and is less expensive as the method in Figure 9 using AN as eluent B.

Conclusion

Methods with short analysis times can aid production of drugs faster and more economically than before, typically using UHPLC instrumentation. The use of modelling software allows the development of methods concordant with QbD criteria, increasing flexibility in routine operations. Retention and critical resolution problems can be more transparent than in the past. Method transfer is much easier using DryLab. Finally HPLC modelling is truly green as it saves time, energy and reduces waste.

References

New Control Driver for the Evaporative Light-Scattering Detector

DataApex is pleased to announce the release of a new control driver for the Evaporative Light-Scattering Detector 90LT, manufactured by SEDERE Company. This is already the 3rd Sedere detector its control and digital data acquisition has been incorporated into Clarity Chromatography SW. The driver for ELSD 90LT has been officially released with Clarity Chromatography Software version 4.0.3.

SEDERE is the leading manufacturer of Low-Temperature Evaporative Light-Scattering Detectors. SEDERE has more than 25 years of experience in the development of versatile detection systems for all chromatography Applications and their detectors are distributed worldwide.

Clarity Chromatography Software has a strong position in the chromatography data systems market. Clarity, the third generation of DataApex products, allows controlling more than 400 different instruments from the single environment and offers users very high flexibility. Clarity is highly regarded for its intuitive approach, excellent performance, cost-effectiveness and proficient technical support.

DataApex is solely focused on chromatography software development. A strong emphasis is placed on technological innovation, visionary adopion of new laboratory standards, best practices and extensive customer support. DataApex products are sold in over 80 countries around the world. Ten chromatography instrument manufacturers privately resell labeled versions of DataApex’s software.

To obtain further information about DataApex Clarity products please contact us at clarity@dataapex.com or visit www.dataapex.com.