JOURNAL OF combinatorial CHEMISTRY

Subscriber access provided by American Chemical Society

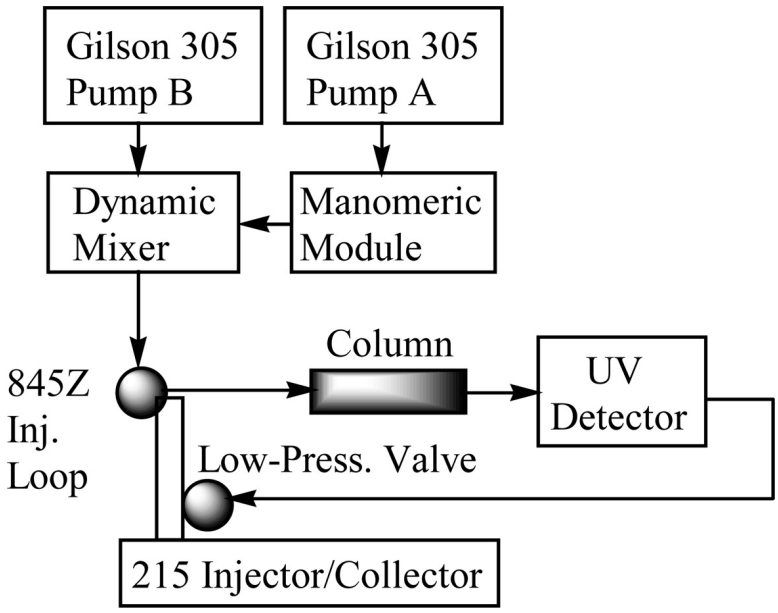
Article

Use of Design of Experiments To Optimize High-Throughput Semipreparative LC and LC/MS Methods

Derek C. Cole, Nicholas Pagano, Michael F. Kelly, and John Ellingboe

J. Comb. Chem., **2004**, 6 (1), 78-82• DOI: 10.1021/cc034013v • Publication Date (Web): 26 November 2003

Downloaded from http://pubs.acs.org on March 20, 2009



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

ACS Publications
High quality. High impact.

Journal of Combinatorial Chemistry is published by the American Chemical Society.
1155 Sixteenth Street N.W., Washington, DC 20036

Use of Design of Experiments To Optimize High-Throughput Semipreparative LC and LC/MS Methods

Derek C. Cole,* Nicholas Pagano, Michael F. Kelly, and John Ellingboe

Chemical & Screening Sciences, Wyeth Research, 401 N. Middletown Road, Pearl River, New York 10965

Received May 29, 2003

This paper describes an investigation into decreasing the run time for high-throughput semipreparative RP-HPLC methods without compromising the resolution. Experimental design was used to devise a small set of experiments in which factors, including solvent flow rate, solvent/column temperature, at-column dilution, and run time were varied systematically. The results were analyzed by means of multiple regression and partial least squares to generate a model relating the factors to the results, showing which factors are important. The model was then used to determine the optimal conditions.

Introduction

The focus of our group is lead optimization by parallel synthesis. Iterative small (20−50 member) libraries are designed to represent the chemical space which is being explored. For small diverse libraries, it is not always possible to fully optimize the synthetic steps for all the monomers; thus, the library may contain members of low purity. However, since the libraries are designed to represent chemical space, it is important that the maximum number of compounds is purified for testing. Thus, an efficient RP-HPLC method that allows the isolation of most library members at the desired purity (>85%) is required.

High-throughput semipreparative LC and LC/MS are now well-established techniques for purification of compounds produced by parallel synthesis. Most systems are run under similar conditions: fast gradient from 5 to 95% organic, 10−50 mL/min, 20 mm × 50−100 mm C18 columns, UV- or MS-directed fraction collection, 10−50-mg sample size, and 5−16-min run time.

Shorter run times have obvious advantages, that is, higher throughput, decreased solvent usage, and smaller fraction volumes, which leads to shorter evaporation times. When we attempted to decrease the run time of the semipreparative RP-HPLC method, we found the resolution decreased and was not sufficient to ensure the desired purity.

Thus, a systematic evaluation of the variable factors involved in semipreparative RP-HPLC was carried out to determine the optimal conditions. Design of experiments (DOE) using Modde 6.0 software (Umetrics, Kinnelon, NJ)[1] was used to devise a small set of experiments in which the pertinent factors, including solvent flow rate, solvent/column temperature, at-column dilution (ACD), and run time, were varied systematically. DOE ensures an organized approach in which fewer experiments are required, the experiments are maximally informative, and the influence of all the factors are taken together, giving connection between experimental results. The results were analyzed by means of multiple regression and partial least squares (PLS) to give a model relating the factors to the results. The results are viewed using bar charts and contour plots, showing which factors are important, and how they combine in influencing the results. The model may also be used to make predictions, for example, how to set the factors to achieve optimal results.

The primary role of column/solvent temperature in RP-HPLC is thought to be its effect on retention, with increasing temperature producing shorter retention times, which give poorer resolution. However, increasing temperature also leads to lower mobile phase viscosity and a reduction in the diffusion coefficient, which results in larger column plate numbers and narrower peaks, leading to improvements in resolution. Reduced mobile phase viscosity also leads to lower back-pressure, allowing higher flow rates.[2]

Chromatographic distortions and poor resolution associated with high column loading are usually caused by strong solvents, large volume injections required to load large amount of sample, or both, rather than by overloading the column packing. ACD was described by Wheat as a means to achieve greater mass loading.[3] A third pump is used to pump the sample from the injection loop to a mixing tee where it is mixed with the solvent at the initial conditions, thus diluting the strong solvent and depositing the sample at the head of the column. Blom recently described a "two-pump" configuration for running ACD in which the stronger solvent is pumped through the injector loop, pushing the sample to the mixing tee where it is diluted by the weaker solvent.[4]

A mixture of cortisone (10 mg) and reserpine (10 mg) dissolved in DMSO (0.5 mL) was used as the test mixture, since this mixture is poorly resolved under the current RP-HPLC conditions used by our group.[5] The performance of the RP-HPLC system was determined by calculating the resolution between these two components. Resolution = (t2 − t1)/(b1 + b2). t1 and t2 are the retention times of peaks

* Corresponding author. E-mail: coledc@wyeth.com.

10.1021/cc034013v CCC: $27.50 © 2004 American Chemical Society
Published on Web 11/26/2003

Table 1. Experimental Design 1 (Screening)

exp no.	temp °C	flow rate (mL/min)	ACD	total run time (min)	resolution	maximum back-pressure (psi)	av retention time (min)
1	40	40	off	8	0.64	1334	3.9
2	40	20	on	16	0.86	667	6.2
3	40	20	off	8	0.47	710	4.8
4	20	20	on	8	0.59	750	4.7
5	20	20	off	8	0.5	754	4.8
6	20	20	on	16	0.77	795	6.3
7	40	40	off	16	1.24	1200	5.2
8	40	40	on	8	0.68	1280	3.8
9	40	40	on	16	1.16	1204	5
10	20	40	off	8	0.62	1363	4
11	20	40	on	8	0.87	1200	3.9
12	30	30	off	12	0.67	1015	5.3
13	20	20	off	16	0.71	754	6.4
14	40	20	on	8	0.59	670	4.7
15	20	40	on	16	1.29	1200	5.2
16	30	30	off	12	0.64	1015	5.3
17	20	40	off	16	1.14	1435	5.2
18	40	20	off	16	0.74	710	6.4

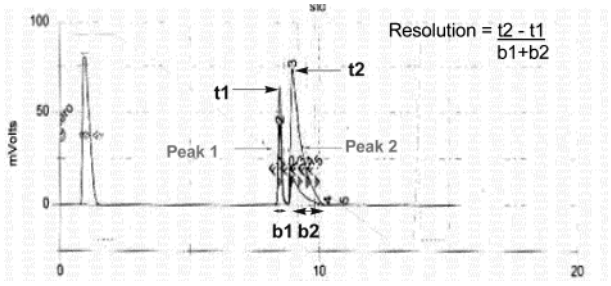


Figure 1. Chromatogram showing separation of cortisone (peak 1) and reserpine (peak 2) and t1, t2, b1, and b2.

1 and 2, and b1 and b2 are the width of the peaks at the baseline (see Figure 1).

Experimental Section

The Gilson semipreparative RP-HPLC system consists of a Gilson 215 liquid handler used as both the sample injector and fraction collector, two 305 dual solvent pumps with 50.SC pump heads (max flow, 50 mL/min), an 806 manometric module, and an 811B dynamic mixer.[6] An 845Z six-position valve mounted on top of the 215 Z arm is used as the injection valve.[7] The sample is aspirated through the needle and into the 2-mL sample loop. The valve is then switched to allow the sample to be pushed from the loop and into the column. A low-pressure valve mounted on the bottom of the 215 Z arm is used as the fraction collection valve. The system is controlled by Gilson Unipoint version 3.2 software.

In the ACD configuration, a 307 pump with a 25.SC pump head is used to pump the organic solvent (or methanol) through the sample loop to a mixing tee (0.02-in. i.d., 0.566-μL dead volume, Analytical Sales & Service part no. 66413-2), where the sample is diluted with the solvent from the main pumps. The flow rate of the ACD pump is set at 5% of the total flow, and the flow rate of the organic pump is adjusted so that the overall percent organic is unchanged.

Since the solvent flow rate under exploration is 20−40 mL/min, it is impossible to heat the column using a column heater alone; thus, the solvent was preheated by passing through a heat exchanger prior to entering the mixing tee.

Table 2. Model Summary (Screening)

	R^2	Q^2	reproducibility
resolution	0.97	0.78	0.99
max back-pressure	0.98	0.86	1
av ret time	0.99	0.98	1

The heat exchanger consists of 12 in. of 1/16-in. stainless steel tubing with an i.d. of 0.040 in. embedded in a "zigzagged" fashion inside an aluminum heat exchanger block, which is controlled by a J-Kem model 150 temperature controller (J-Kem Part No. HPLC-RC-11).[8]

Results and Discussion

The objective of the screening set of experiments was to determine which variables are the most influential and the appropriate ranges. The factors varied and the ranges were solvent temperature, 20−40 °C; flow rate, 20−40 mL/min; run time, 8−16 min; and with or without ACD.

A full factorial design with a set of 18 experiments, including two center points, was selected. The experiments were run in random order, and the calculated and measured results, that is, resolution, maximum back-pressure (measured at 95% aqueous mobile phase), and average peak retention time, are shown in Table 1.

The data shown in Table 1 were analyzed using Modde 6.0 to generate a model with linear and interaction terms. The quality of the model is summarized by the data in Table 2, where R^2 is the goodness of fit value and is a measure of how well the model fits the raw data. Q^2 is the goodness of prediction and estimates the predictive power of the model, and reproducibility is a measure of the variations of the response under the same conditions, that is, the center points. A perfect model has a value of 1 for all three parameters.

Bar charts provide an overview of which factors most influence resolution, maximum back-pressure, and average retention time. Figure 4 shows that the flow rate and run time have the biggest effect on resolution; however, ACD also increases resolution significantly. Temperature has no effect, and there is a small interaction effect between rate and time.

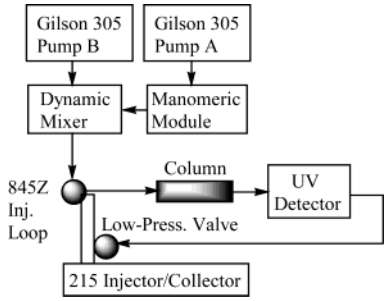


Figure 2. Semipreparative RP-HPLC flow diagram.

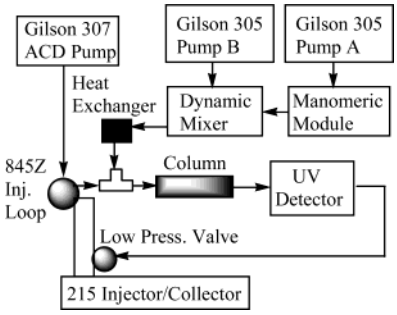


Figure 3. Semipreparative RP-HPLC flow diagram with ACD and solvent heat exchanger.

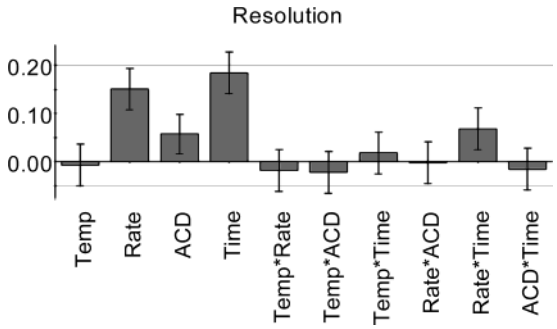


Figure 4. Factors affecting resolution. Temp = solvent/column temperature, rate = solvent flow rate, ACD = at-column dilution, and time = total run time.

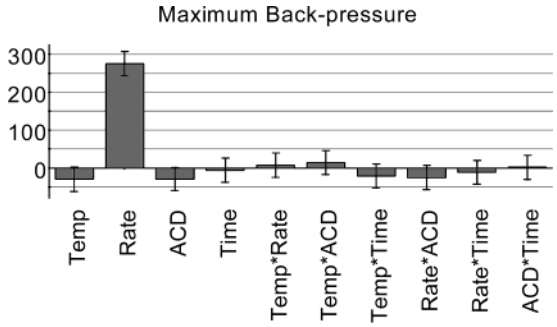


Figure 5. Factors affecting maximum back-pressure. Temp = solvent/column temperature, rate = solvent flow rate, ACD = at-column dilution, and time = total run time.

Figure 5 shows that flow rate is the predominant factor affecting pressure, with an ~280 psi increase going from 20 to 40 mL/min. There is a trend in which both increasing temperature and ACD decrease the back-pressure, with a combined effect of ~80 psi.

Increasing flow rate and decreasing run time decrease the average retention time, as shown in Figure 6.

In summary, the screening set of experiments shows that increasing the flow rate increases resolution, allowing the run time to be shortened. Increasing the solvent temperature combined with the use of ACD results in increased resolution while decreasing back-pressure. There is an interaction term between run time and flow rate which increases resolution. ACD also potentially allows increased sample loading.[3,4]

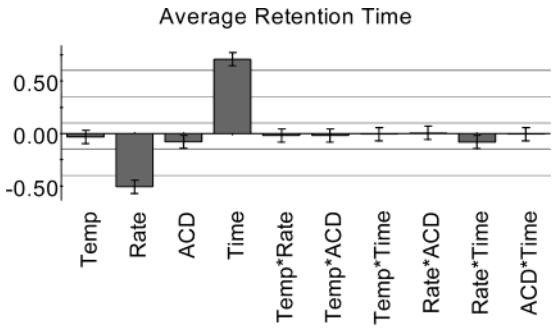


Figure 6. Factors affecting average retention time. Temp = solvent/column temperature, rate = solvent flow rate, ACD = at-column dilution, and time = total run time.

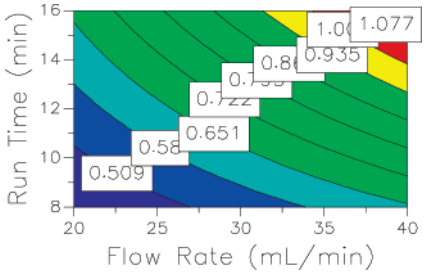


Figure 7. Contour plot showing resolution as a factor of run time vs flow rate at 20 °C with ACD off.

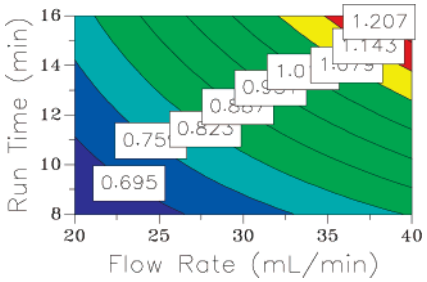


Figure 8. Contour plot showing resolution as a factor of run time vs flow rate at 20 °C with ACD on.

The results are best interpreted by viewing contour plots where the resolution is plotted in a run time vs flow rate graph (Figure 7).

The contour plots show that increases in total run time and flow rate lead to increased resolution. By comparing Figures 7 and 8, the effect of adding ACD is observed. The resolution increased from 1.077 to 1.207 for comparable 20-min runs at 40 mL/min.

Having determined the dominant variables, a response surface model (RSM) optimization design was used to determine the optimal value for these variables, or the best compromise, that is, the combination of the important factors that results in optimal operating conditions. For these experiments, the solvent temperature was set at 40 °C with ACD on, the flow rate was varied from 20 to 40 mL/min, and the run time was varied from 8 to 16 min. The randomized set of experiments, including 4 center points, and the calculated resolution are shown in Table 3.

Table 3. Experimental Design 2 (Optimization)

exp. no.	flow rate	run time	resolution
1	33.3	16	0.84
2	20	13.3	0.57
3	40	13.3	0.96
4	40	8	1.05
5	40	10.7	0.96
6	20	16	0.735
7	20	8	0.51
8	30	12	0.66
9	26.7	16	0.765
10	30	12	0.675
11	40	16	1.11
12	33.3	8	0.57
13	26.7	8	0.615
14	30	12	0.63
15	30	12	0.675
16	20	10.7	0.6

Table 4. Model Summary (Optimization)

	R^2	Q^2	reproducibility
resolution	0.92	0.77	0.98

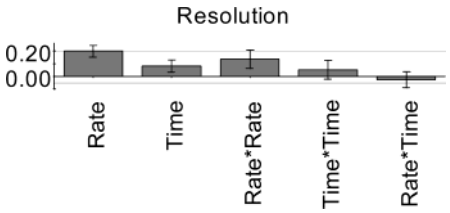


Figure 9. Factors affecting average retention time. Rate = solvent flow rate, time = total run time.

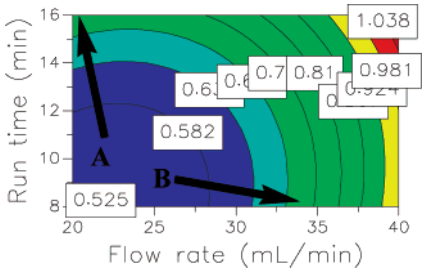


Figure 10. Contour plot showing resolution as a factor of run time vs flow rate at 40 °C with ACD on.

The data were fit to a quadratic model with the R^2, Q^2, and reproducibility values shown in Table 4, indicating an excellent model. Figure 9 shows an overview of the model, with flow rate having the most pronounced effect on resolution.

The contour plot shown in Figure 10 shows resolution as a factor of flow rate vs run time. Once again, the best resolution is achieved at longest run time and highest flow rate, with flow rate being the dominant factor. The resolution appears to be equivalent at point A, 20 mL/min for 16 min; and at point B, 34 mL/min for 8 min.

A direct experimental comparison was run using the unoptimized conditions (point A, Figure 10, 20 °C, ACD off)[5] and the optimized conditions (point B, 40 °C, ACD on).[9] The results are shown in Figures 11 and 12. Clearly, the results are much superior in Figure 12. The overall run time is halved from 16 to 8 min, and the average retention time for the peaks is shortened from 9.0 to 4.8 min, while the resolution increased from 0.54 to 0.74.

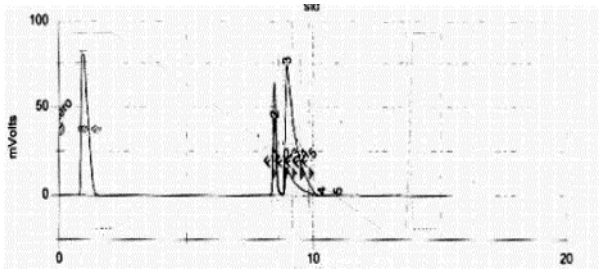


Figure 11. Unoptimized conditions: flow rate, 22.5 mL/min for 16 min (360 mL total); temperature, room temp; ACD off; 5−95% organic (acetonitrile); av retention time = 9.0 min; and resolution = 0.54.

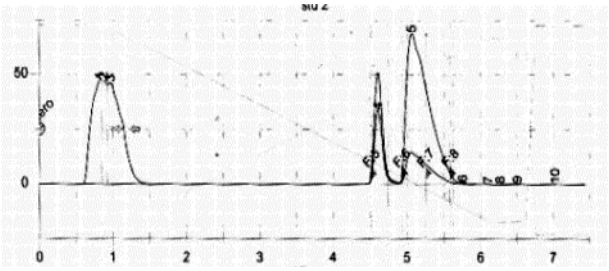


Figure 12. Optimized conditions: flow rate, 34 mL/min for 8 min (272 mL total); temperatur, 40 °C; ACD on; gradient 5−95% organic (acetonitrile); av retention time = 4.8 min; and resolution = 0.74.

Conclusions

Using DOE, we were quickly able to define and optimize the variables, which had the most important influence on the resolution of two closely eluting peaks. The optimized conditions result in a 50% reduction in run time and a 25% reduction in solvent usage while increasing the resolution. Increased resolution also leads to sharper peaks, thus, fewer fractions to combine and less solvent to evaporate. In addition, since most analytical LC/MS systems are run at elevated temperatures, running the semipreparative LC at the same temperature allows easier scaling from analytical to preparative. Increased temperature may also increase solubility, leading to fewer blockages due to sample precipitation, although no precipitation was seen with the test mixture.

Acknowledgment. The authors acknowledge Svante Wolde (Umetrics); Tom Wheat (Waters); and Joan Stevens, Paul Lefebvre, Bill Tuting, and Millie Driz (Gilson) for helpful discussions.

References and Notes

(1) www.umetrics.com.
(2) Dolan, J. W. *J. Chromatogr. A* **2002**, *965*, 195−205. See also www.selerity.com.
(3) Wheat, T. *Abstracts of Papers, Pittsburgh Conference, New Orleans, LA, March 4−9, 2001*; Pittsburgh Conference, Pittsburgh, PA, 2001; p 1948.
(4) Blom, K. F. *J. Comb. Chem.* **2002**, *4*, 295−301.
(5) Unoptimized HPLC conditions: solvent A, 0.02% TFA/water; solvent B, 0.02% TFA/acetonitrile. Gradient: time 0 min, 95% A; 2 min, 95% A; 14 min, 10% A; 15 min, 10% A; 16 min, 95% A. Flow rate, 20 mL/min. Detection, 254-nm DAD. Sample size, 20−50 mg crude dissolved in 1−2 mL DMSO/MeOH/water. Column, Waters XTerra Prep MS C18 5 μM, 19 × 50 mm. Guard column, XTerra C18 5 μM, 10 × 10 mm.

(6) www.gilson.com.
(7) Lefebvre, P. M; West, W. R. Personal communications. See also Application Note LT40004, www.gilson.com.
(8) www.j-kem.com.
(9) Optimized HPLC conditions: Solvent A, 0.02% TFA/water; solvent B, 0.02% TFA/acetonitrile. Gradient: time 0 min, 95% A; 1 min, 95% A; 7 min, 10% A; 7.5 min, 10% A; 8 min, 95% A. Flow rate, 34 mL/min. ACD, 5% of solvent A (or MeOH) pumped through sample loop. (Note: the flow rate of pump A is decreased so the overall percent of A is unchanged.) Solvent temperature, 40°C. Detection, 254-nm DAD. Sample size, 20−50 mg crude dissolved in 1−2 mL DMSO/MeOH/water. Column, Waters XTerra Prep MS C18 5 μM, 19 × 50 mm. Guard column: XTerra C18 5 μM, 10 × 10 mm.

CC034013V