CHROM. 24 540

Experimental design in the development and characterization of a high-performance liquid chromatographic method for amino acids

H. A. Hansen and C. Emborg

Department of Biotechnology, Technical University of Denmark, Block 223, DK-2800 Lyngby (Denmark)

(First received April 24th, 1992; revised manuscript received August 3rd, 1992)

ABSTRACT

In the practical establishment of an HPLC method, it is often necessary to do a considerable amount of experimentation even with methods that are not new. In this paper it is demonstrated how the structuring of a considerable part of the experiments in two-level factorial designs facilitates the localization of useful chromatographic conditions, while at the same time the system is characterized with respect to factors and interactions between factors that have an influence on the quality of the chromatograms.

INTRODUCTION

This paper deals with the very common situation in which one wishes to introduce a new analytical method in the laboratory. The literature has been thoroughly studied, the market for equipment has been surveyed, discussions with experienced colleagues have been held and in the end a decision has been made and HPLC equipment has been bought. In this work, we wished to determine amino acid concentrations in animal cell cultures in order to devise new methods of control based on the physiological demands of the cells. We decided to use ophthalaldehyde-mercaptoethanol (OPA-ME) precolumn derivatization and reversed-phase HPLC. This is a well documented method [1,2], and we expected that it might be difficult to resolve the peaks of some pairs of amino acids (glutamine-histidine, glycine-threonine, tryptophan-methionine and methionine-valine).

Many papers that deal with the development of

HPLC methods use experimental designs based on the variation of one factor at a time $[1,3-5]$. In our laboratory, we often use two-level factorial designs, because they reveal interactions between the factors [6-8]. We have examined the HPLC literature for information on the handling of factorial experiments in this common situation. However, papers are often of a more flamboyant nature [9]. We therefore think that our experience with two-level factorial designs may be of value for other workers in the field.

EXPERIMENTAL

Statistical design

The experiments were planned as full $2⁴$ factorial designs, *i.e.* experiments were performed which give all combinations of four variables each at two levels. The total number of runs was sixteen. We executed two designs. The second design was decided after the results of the first design became available. The variables (factors A, *B, C* and D) chosen for the first design are given in Table I. Factors A and *B* are further specified in Table II and factors C and *D* are further described under *Mobile phase preparation.*

Correspondence to: C. Emborg, Department of Biotechnology, Technical University of Denmark, Block 223, DK-2800 Lyngby, Denmark.

FACTORS USED IN DESIGN 1

Factors		Level		
		Low	High	
A	Gradient shape	Linear	Convex	
B	Starting concentration of eluent B	18%	22%	
\mathcal{C}	Acetate in eluent A	0 M	0.005 M	
D	Tetrahydrofuran in eluent A	0%	2%	

TABLE III FACTORS USED IN DESIGN 2

To avoid confusion we named the factors for the second design J, K, *L* and M [see Tables III and IV (factor M) and *Mobile phase preparation* (factors *J, K* and *L)].* To reduce the number of eluent reservoir shifts, the factorial design was performed in blocks in a non-random way.

Factorial design 1. For each factor combination three chromatograms were acquired. The first was a chromatogram with a hydrolysis standard containing amino acids of which fifteen give signals in the chromatograms. Then two chromatograms containing glutamine and histidine were acquired.

Factorial design 2. For each factor combination four chromatograms were acquired. The first was a chromatogram with a hydrolysis standard containing amino acids of which fifteen give signals in the chromatograms. Three other chromatograms were acquired to determine the retention times of glutamine, asparagine and tryptophan, which are not included in the hydrolysis standard, and to verify the retention times of the other components. These

three samples were constituted as follows: glutamine, tyrosine, tryptophan; asparagine, arginine, methionine; and serine, glycine, valine.

Equipment

The chromatographic system consisted of two M510 pumps and an M470 fluorescence detector (with a 5- μ l flow cell, excitation at 330 nm and emission at 418 nm) from Waters, Division of Millipore (Milford, MA, USA). A Model 7000 automatic switching valve (Rheodyne, Cotati, CA, USA) was equipped with a $20-\mu$ sample loop and driven pneumatically. Control of the system and data acquisition were performed with Baseline 810 software (Dynamic Solution, Division of Millipore) run on an NEC personal computer. The amino acid derivatives were separated on a 100×8 mm I.D. Radial-Pak Nova-Pak C₁₈ 4- μ m column (Waters). A Nova-Pak Guard-Pak C_{18} column was used to protect the analytical column.

TABLE II

CHROMATOGRAPHIC GRADIENT PROGRAMME USED IN DESIGN 1

TABLE IV

Reagents

The amino acid standard was a protein hydrolysate, Amino Acid Standard H (Pierce, Rockford, IL, USA), which does not contain asparagine, glutamine and tryptophan, and the OPA reagent used was Fluoraldehyde reagent solution (0.8 mg/ml of OPA in borate buffer containing Brij 35 and 2-mercaptoethanol) (Pierce). The OPA reagent was filtered through a 0.2 - μ m Minisart NML filter (Sartorius, Gottingen, Germany). Methanol (HPLC grade) was purchased from Labscan (Dublin, Ireland), tetrahydrofuran (LiChrosolv) from Merck (Darmstadt, Germany) and morpholine (analyticalreagent grade) from Sigma (St. Louis, MO, USA). Other chemicals (analytical-reagent grade or better) were from Merck. Aqueous solvents were based on water purified with a Milli-Q water purification system ("Milli-Q water") (Millipore, (Milford, MA, USA).

Mobile phase preparation

In two factorial designs different eluents A were used to obtain the variations designated factors C, *D, J* and K. The preparation of eluents A was based on a sodium phosphate stock solution (79.20 g of $Na₂HPO₄ \cdot 2H₂O + 1000$ ml of Milli-Q water) and a sodium acetate stock solution (54.43 g of $CH_3COONa \cdot 3H_2O + 1000$ ml of Milli-Q water).

In design 1, eluent A without acetate (factor C low level) consisted of 12 ml of sodium phosphate stock solution and 948 ml of Milli-Q water. Before addition of organic modifier, the pH was adjusted to 7.2 with phosphoric acid. Eluent A with acetate (factor C high level) was prepared from 12 ml of both sodium acetate and sodium phosphate stock solutions and 936 ml of Milli-Q water. The pH was adjusted to 7.2 with glacial acetic acid before addition of organic modifier. For factor *D* low level, 40 ml of methanol were added, and for factor *D* high level, 20 ml of methanol and 20 ml of tetrahydrofuran were added.

In design 2 eluent A consisted of 12 ml of sodium phosphate stock solution and 948 ml of Milli-Q water. Before addition of organic modifier, the pH was adjusted with phosphoric acid. For factor K low level, the pH was adjusted to 6.5, and for factor K high level, the pH was adjusted to 7.2. Then organic modifier was added. For factor *J* low level, 40 ml of methanol were added, and for factor *J* high level, 20 ml of methanol and 20 ml of tetrahydrofuran were added.

For design 1 eluent B described as low level of factor *L* in design 2 was used. In design 2 two B eluents were used, giving the high and low levels of factor *L.* Factor *L* low level was simply 350 ml of Milli-Q water and 650 ml of methanol. Factor *L* high level was 10 ml of morpholine, 4.75 ml of 85% orthophosphoric acid, 335.25 ml of Milli-Q water and 650 ml of methanol (pH 6.85 before and pH 7.54 after addition of methanol).

HPLC solvents were degassed by vacuum filtration through a $0.45~\mu m$ type HV filter (Millipore).

Derivatization

The derivatization was carried out manually in small tubes. To sample volumes of 5μ (containing 2.5 mM of each amino acid), 800 μ l OPA reagent were added. The OPA amino acid adducts are unstable compounds so the reaction procedure needs to be timed carefully. The OPA reagent was added to the sample and, after mixing, the samples were transferred into the $20-\mu l$ sample loop. After 120 s of reaction the pneumatic valve was turned for injection of the samples into the eluent stream and the chromatographic gradient programme was started.

Chromatographic conditions

The column was conditioned before analysis by first equilibrating the column with 30 ml of organic solvent, then 100 ml of 0.1% (v/v) H_3PO_4 were pumped through the column. This procedure is adapted from another HPLC method [3] in which μ Bondapack C₁₈ columns were used. It was found that flushing with phosphoric acid removed an acidsoluble component that reacted with the OPA-ME reagent and caused rapid degradation of the column. The chromatographic gradient programme used in the analysis is shown in Tables III and IV. The column was equilibrated for 5 min between each injection.

Data handling

The retention times from the chromatograms were manually transferred to a Lotus 1-2-3 spreadsheet and Yates algorithm [10] and other statistical calculations were executed with this software.

RESULTS

The array of chromatographic conditions of de-

and histidine in design 1.

sign 1 caused an almost synchronous variation in the retention times of glutamine and histidine $(7-13)$ min). The array, however, caused a small variation in the difference between the retention times of glutamine and histidine. The differences in retention times of the isolated glutamine and histidine peaks are illustrated in a block diagram (Fig. 1). There is a relatively large variation in the size of the blocks, ranging from a poor separation of glutamine and histidine $(ABCD, 0.1 \text{ min})$ to a fine separation $(C, 1)$ min). It is easily seen that the blocks in the right half of the diagram are smaller than those in the left half. This means that tetrahydrofuran in the eluent A (factor *D)* had a negative influence on the separation of glutamine and histidine. It is also seen that every other block is lower than its left neighbour. This means that a convex gradient shape (factor A) had a negative effect on the separation of the two compounds. Careful inspection of the blocks also reveals that blocks with factor B are lower than blocks without factor B. This means that the starting concentration of eluent B in the gradient had an influence on the separation. An 18% concentration gave a better separation than a 22% concentration. These statements were confirmed by the statistical analysis, which also indicated that acetate in eluent A (factor C) had no influence on the separation. The best separation in the set, however, was obtained with acetate in eluent A. On the other hand, with tetrahydrofuran in eluent A acetate had a negative effect on the separation. This suggests an interaction between acetate and tetrahydrofuran.

Fig. 1. Differences in retention times (min) between glutamine Fig. 2. Chromatogram for the hydrolysis standard under condi-
tion C in design 1.

Fig. 3. Differences in retention times (min) between glycine and threonine in design 1.

This possible *CD* interaction was visible in the statistical analysis, but below the significance level.

Fig. 2 shows the chromatogram of the hydrolysis standard obtained under the conditions that gave the best separation of glutamine and histidine (C) . Most peaks were well separated, but the pairs gly $cine$ -threonine (16 min) and methionine-valine (29 min) were close. The block diagrams for the separation of these two pairs are shown in Figs. 3 and

Fig. 4. Differences in retention times (min) between methionine and valine in design 1.

4. At first glance, Fig. 3 for the glycine-threonine separation is a mirror of the glutamine-histidine separation in Fig. 1, because tetrahydrofuran in eluent A (factor *D)* is positive for the glycine-threonine separation to the same extent as it was negative for the glutamine-histidine separation. The convex gradient shape (factor A) had a negative effect in Fig. 3 as in Fig. 1, but the starting concentration of eluent B (factor B) had no influence on the glycinethreonine separation, contrary to the case in Fig. 1. Again, the best separation was seen with acetate in eluent A, but this time together with tetrahydrofuran *(CD)*. Here also there appears to be a small *CD* interaction, but in this instance the effect is the opposite of that for glutamine–histidine case. For the separation of methionine-valine the block diagram is totally different (Fig. 4). The statistical analysis revealed no significant effects, although visual inspection of the diagram points to a possible effect of gradient shape (factor A). In this instance the convex gradient gave a better separation than the linear gradient, whereas the linear gradient was the better in the two previous examples.

Based on these experiments, we performed some experiments with lower tetrahydrofuran concentrations (1 and 0.5%) and lower levels of various gradients. We thereby found a method with 0.5% tetrahydrofuran in eluent A where sixteen of the amino acids could be quantified (Fig. 5).

Fig. 5. Chromatogram with the hydrolysis standard supplemented with asparagine, glutamine and tryptophan under the following conditions: cluents as *CD* in design 1 except for organic modifier in eluent A, which was 0.5% tetrahydrofuran and 3.5% methanol. Linear gradient programme: (min, %B) (0, 12) (18, 42.6) (20, 52) (35, 100) (37, 100) (40, 12) (45, 12); flow-rate 2.0 ml/min.

TABLE V

RETENTION TIMES AND DIFFERENCES IN RETENTION TIMES FOR METHIONINE, VALINE AND TRYPTOPHAN IN DESIGN 2

Factor	Retention time (min)			Difference in retention times (min)			
combi- nation	Trp	Met	Val	$Met-$ Trp	$Val -$ Met	$Val-$ Trp	
(1)	24.58	24.52	25.02	-0.06	0.50	0.44	
\boldsymbol{J}	24.63	24.17	24.67	-0.46	0.50	0.04	
K	24.43	24.42	24.90	-0.01	0.48	0.47	
JK	24.35	23.83	24.32	-0.52	0.49	-0.03	
L	29.15	29.53	30.35	0.38	0.82	1.20	
JL	29.07	29.07	29.87	0.00	0.80	0.80	
KL	29.27	29.68	30.50	0.41	0.82	1.23	
JKL	29.20	29.25	30.05	0.05	0.80	0.85	
M	22.27	22.10	22.68	-0.17	0.58	0.41	
JM	22.40	21.77	22.32	-0.63	0.55	-0.08	
KM	22.17	22.10	22.65	-0.07	0.55	0.48	
JKM	22.07	21.42	21.95	-0.65	0.53	-0.12	
LM.	27.48	27.83	28.75	0.35	0.92	1.27	
JLM	27.27	27.22	28.13	-0.05	0.91	0.86	
KLM	27.32	27.83	28.72	0.51	0.89	1.40	
JKLM	27.70	27.65	28.52	-0.05	0.87	0.82	

Whereas design 1 focused on optimization of the glutamine-histidine separation with a compromise with other features in the chromatograms, design 2 (Table II) focused on the separation of three amino acids, tryptophan, methionine and valine. Tryptophan is not present in the hydrolysis standard because it is degraded by acid hydrolysis. The array of design 1 did not have much influence on the separation of methionine and valine, and tryptophan was eluted close to these two peaks under the conditions hitherto tested. We thought that morpholine might cause a change in the relative affinities of the amino acids to the column. Morpholine in eluent B was therefore included as a factor together with the more traditional factors chosen. Design 2 consists of 64 chromatograms. Sixteen of these are with the hydrolysis standard. For each of the three critical amino acids there are sixteen chromatograms. In these chromatograms there are three peaks, one for the critical amino acids together with two peaks for other amino acids of the hydrolysis standard.

Table V shows the retention times obtained for the three peaks together with the differences in retention times. It can be seen that this array of chromatographic conditions caused variations in reten-

tion times for all three compounds in a fairly synchronous manner. Although methionine and valine were both more sensitive to the changes in conditions than tryptophan, there was a fair to good separation between these two compounds everywhere in the design. Morpholine in eluent B caused an increase in retention time for all compounds. This

Fig. 6. Chromatogram with the hydrolysis standard supplemented with asparagine, glutamine and tryptophan under conditions *KLM* in design 2 (Imp. $=$ impurity).

increase was larger for methionine and valine than for tryptophan and caused a change in the order of elution. Under four sets of conditions in the array *(L, KL, LM* and *KLM)* tryptophan was eluted before methionine. Fortunately, the condition for the best separation of tryptophan and methionine coincided with the best separation of methionine and valine *(KLM,* pH in eluent A, 7.2; morpholine in eluent B, 1%; and starting concentration of eluent B, 25%; no tetrahydrofuran in eluent A). As in design 1, tetrahydrofuran had no influence on the separation of methionine and valine. This separation was favoured by morpholine in eluent B and a high starting concentration of eluent B. Tetrahydrofuran had a strong negative influence on the separation of tryptophan from methionine-valine, whereas morpholine had a positive influence on this separation. Fig. 6 shows a chromatogram with the *KLM* conditions. It can be seen that the separation of tryptophan and methionine is not as good as indicated in the factorial, so further work may be needed. It is also seen that morpholine causes the baseline to slope, but with the present software this is easily dealt with. Morpholine is fluorescent, which is why the baseline slopes. Morpholine is not available as chromatographic-grade material and some impurities reacted with the OPA reagent giving peaks in the chromatogram.

Generally, morpholine had a significant influence on the retentions of many of the amino acids, on their separation and on the order of peaks, and this influence was not always positive. This is illustrated in Fig. 7, which is a block diagram of the glutamine-histidine separation. This data set may be used to illustrate how the statistical analysis may support the interpretation of the results. The block diagram in Fig. 8 shows the squares of effects obtained by the Yates analysis of the data in Fig. 7. Fig. 8 shows that the factors *J* and *L* (tetrahydrofuran and morpholine) had an effect, but that factors K and M (pH in eluent A and starting concentration of eluent B) had no effect. These conclusions could also have been drawn from a visual inspection of Fig. 7, at least regarding *J, L* and M. Based on visual inspection it is puzzling that factor *K* should have no effect. Fig. 8 solves this problem. It shows that factor *K* is involved in an interaction with factor *L*. The effect of pH in eluent A is reversed when there is morpholine in eluent B. It must

Fig. *7.* Differences in retention times (min) between glutamine and histidine in design 2.

be concluded that the use of morpholine has advantages and disadvantages. It may be useful for special separation problems, especially if better HPLC-grade material were available.

Sets of raw retention data for two compounds may be applied to the Yates algorithm directly. This is done by including "compound" in the design as a factor and treating the data as derived in a $2⁵$ facto-

Fig. 8. Squares of effects obtained by Yates analysis of the data in Fig. 7.

Fig. 9. Squares of effects obtained by Yates analysis when the glutamine and histidine retention data from design 2 is treated as a 25 factorial experiment.

rial. This is done in Fig. 9. Factor N is "compound". It can be seen that the picture from Fig. 7 is repeated in the right side of the diagram in the interactions JN, *LN* and *KLN.* Interactions involving the "compound" factor N indicate an influence on separation. The effects on the left of the diagram are much larger. This means that varying chromatographic conditions have a much larger influence on the position of the peaks, than on their mutual separation. We note here factor M (starting concentration of eluent B) and its interaction with J (tetrahydrofuran in eluent A), which apparently works in the same way on glutamine and histidine, as they are not seen in the right side of the diagram.

DISCUSSION

In work on the design and optimization of an HPLC method, there are many choices to be made. It would be an impossible task to validate all choices experimentally, so everything one does is in fact a more or less intelligent practical compromise, which is made in the frame of available time and resources. When one decides to make a 2" factorial, one must run 2" chromatograms, and one is committed to providing an array of experimental conditions. We find the effort to be worthwhile. According to textbooks, such an array should be executed

178 *H. A. Hansen and C. Emborg 1 J. Chromatogr. 626 (1992) 171-180*

in a random manner or by carefully blocking the experiments by methods also described in textbooks. Such schemes may represent a burden in an HPLC investigation with changes of eluent, compared with the execution of the array in the most practical manner. We chose to execute the arrays in the most practical manner possible because, with reliable equipment at hand, block effects should be negligible. Also, if block effects do occur, as will rarely be the case, what has happened will often be evident in the practical situation. The block effect may decrease the representability of the results, but often the information for the guidance of future work may be intact. In the worst case, some experiments must be repeated.

Another burden is the statistical treatment of the results. As seen from the Results section, almost all information for the guidance of future work can be obtained without the aid of statistical treatments. The main advantage of structured experimental designs is the feedback of structured knowledge including a knowledge of interactions between factors, and the increase in probability of finding useful experimental conditions that may serve as a platform for future work.

Software for the statistical treatments is commercially available, and it is easy to perform the manipulations of the Yates algorithm in a spreadsheet, which may be coupled to software for graphical representation of the results. The Yates algorithm, however, is also easily performed manually.

We chose to use $2⁴$ designs. The advantage of this design is, assuming the absence of three factor interactions, that it is possible to obtain a good estimate of experimental error without repetition of experiments. By virtue of its size, it gives sixteen chances of finding useful experimental conditions. In these experiments we investigated four factors in the designs. It is possible to include a fifth factor in the design by sacrificing an easy estimate of experimental error, but retaining an estimate of two-factor interactions and the sixteen chances of finding useful experimental conditions. An additional factor may increase the experimental effort in eluent changes, but for example design 2 could have included gradient shape without a substantial increase in experimental effort. The extra factor complicates the interpretation of the results without the aid of the Yates algorithm.

Fig. 10. Rearrangement of Fig. 1. Each half of the histogram illustrates a 2^{4-1} factorial.

Fig. 11. Rearrangement of Fig. 3. Each half of the histogram illustrates a 24-1 factorial.

Another possibility is the $2³$ design. It includes only one estimate of experimental error and gives only eight chances of finding useful experimental conditions. It is possible to include a fourth factor in the design by loss of the estimate of experimental error and by loss of a clear indication of two-factor interactions.

The effect of executing $2³$ instead of $2⁴$ designs may be visualized by inspection of the block diagrams in Figs. 1 and 3. Each figure includes four different $2³$ designs. It is easy to see two of these, because the left half of the diagram is one $2³$ design and the right half is another. From this one might conclude that the extra effort in a $2⁴$ design was worthwhile. The results in the two block diagrams may also be used to illustrate $2³$ designs with a fourth factor included $(2^{4-1}$ designs). In Figs. 10 and 11 the blocks are rearranged, so that each half consists of a 2^{4-1} design in Yates order. It can be seen that it is difficult to realise the effects visually, but with the help of the Yates algorithm it is clear that for the simple system in Fig. 11 without interactions, information from the confounded design is almost the same as that of the full design, whereas information on the interaction in Fig. 10 is lost in the confounded design. When we decided not to use a 2^{4-1} design, we reasoned that when we had set the system for work with four factors, the extra burden

of executing a 24 design was small compared with the overall effort. The discussion here may, however, serve to illustrate why we think that simpler designs than those considered here may be less efficient, especially if one is unlucky in the choice of experimental conditions for the first experiments. When we have found a good platform, we often use simpler designs in order to elaborate on effects and interactions that have attracted our attention in the two-level factorials.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Nordic Programme on Bioprocess Engineering under the auspices of the Nordic Fund for Technology and Industrial Development and by the Center for Process Biotechnology, Technical University of Denmark.

REFERENCES

- 1 P. Lindroth and K. Mopper, *Anal.* Chem., 51 (1979) 1667- 1674.
- 2 J. Schmidt and C. J. McClain, *J. Chromatogr.*, 419 (1987) $1 - 16$.
- 3 D. Hill, L. Burnworth, W. Skea and R. Pfeifer, *J. Liq. Chromatogr., 5* (1982) 2369-2393.
- *4* B. Persson and D. Eaker, *J. Biochem. Biophys. Methods, 21 (1990) 341-3.50.*
- *5* E. H. J. M. Jansen, R. H. van den Berg, R. Both-Miedema and L. Doorn, *J. Chromatogr.*, 553 (1991) 123-133.
- *6* P. K. Jepsen, E. Riise, K. Biedermann, P. C. R. Kristensen and C. Emborg, *Appl. Environ. Microbial., 53 (1987) 2593- 2596.*
- *7 C.* Emborg, P. K. Jepsen and K. Biedermann, *Biotechnol. Bioeng., 33 (1989) 1393-1399.*
- 8 K. Biedermann, H. Fiedler, B. S. Larsen, E. Riise, C. Emborg and P. K. Jepsen, *Appl. Environ. Microbial., 56 (1990) 1833- 1838.*
- *9 M. Righezza and J. R. Chrétien, J. Chromatogr., 556 (1991) 169-180.*
- *10 G.* E. P. Box, W. G. Hunter and J. S. Hunter, *Statistics for Experimenters -An Introduction to Design, Data Analysis, and Model Building,* Wiley, New York, 1978.