CHROMSYMP. 2245

# **Computer-assisted sample clean-up in liquid chromatography from thin-layer chromatographic data**

A. M. SIOUFFI\* and N. VOLPE

Laboratoire de Génie Chimique et Chimie Appliquée, Université Aix-Marseille 3, F-13397 Marseille Cedex *13 (France)* 

#### ABSTRACT

A method is presented where thin-layer chromatographic data with binary mixtures eluents are treated with NEMROD software in order to enhance the *AR,* between two solutes. In this procedure it is possible to selectively retain one solute or one group of solutes with a change in solvent composition. The method has been succesfully applied to derivatives of phenols in both normal and reversed-phase systems and the only requirement is the capacity factor ranking.

### INTRODUCTION

The ideal situation in liquid chromatography (LC) would be to directly inject the sample onto a chromatographic column and to obtain unambiguous separation of all solutes. Unfortunately this is not possible and in many cases sample pretreatment is required. For this purpose many strategies have been developed including liquid-liquid extraction, solid-phase extraction, column switching, etc. The topic is well documented in excellent book [l] and a recent review [2].

In many cases sample clean-up is performed either on-line or off-line with a cartridge of small dimensions. Samples are adsorbed onto the packing and a solvent is selected to retain the solutes of interest whilst the others are eluted. Conversely, appropriate choice of packing and solvent will retain the undesirable compounds and elute those of interest with the advantage of peak compression. Selection of suitable packing and solvent is usually carried out by trial and error or from knowledge of chromatographic behaviour. For example, hydrophobicity of a solute is commonly expressed as the logarithm of the partition coefficient (log  $P$ ) between 1-octanol and water. Since the logarithm of the capacity factor (log  $k'$ ) obtained from reversedphase liquid chromatography (RPLC) has been shown to have good correlation with the log *P* of several classes of chemicals, the value of k' in pure water (log  $k_w$ ) can be deduced and can be effective in the designing of sample clean-up procedures [3]. On the other hand, knowledge of the polarity of the solutes permits maximization of the interactions with the packing in order to perform selective extractions.

These procedures are very selective but some drawbacks are obvious. For ex-

ample, they may involve two immiscible solvents with the consequence of the required use of a third solvent and subsequent dilution of the solutes prior to injection in the analytical column. From the numerous data published in the literature it can be shown that extrapolation to 0% organic modifier in RPLC yields different values of  $\log k_{\rm w}$  depending on the nature of the organic modifier (methanol or acetonitrile) [4].

Two features caught our attention. (1) The performances of precolumns expressed as plate counts are not very high and can be compared to those obtained in classical thin-layer chromatography (TLC); large amounts of TLC data can be retrieved from the literature thus allowing the gathering of actual information on solute behavior. (2) From the optimization procedures advocated in high-performance liquid chromatography (HPLC) a slight change in solvent composition would dramatically change the selectivity. Thus it would be possible to enhance the selectivity by looking for the solvent composition which would produce the maximum difference in capacity factors, making it possible to separate and eliminate one solute (or one group of solutes) while keeping the others on the precolumn for a predicted time. In this paper we explore the feasibility of the concept and give some preliminary results.

# **THEORY**

We shall first make some assumptions in order to restrict the domain. We shall only consider normal-phase (NP) and RP chromatography with binary eluents. We shall not consider mixtures of three to four solvents or ion-pairing chromatography. We shall consider that the sample can be chromatographed either isocratically or with gradient elution yielding a chromatogram in which peaks are ranked from first to last. This precludes trace analysis or sample overload. The aim will be to eliminate (or to keep) the first solute. Conversely we can eliminate all the peaks with the exception of the last one.

It has been demonstrated in TLC by Soczewinski and Golkiewicz [5,6] that in NP with a silica gel packing and a binary mobile phase consisting of an apolar diluent and a polar modifier that

$$
R_M = \log [(1 - R_F)/R_F] = a_i \log X_s + b_i \tag{1}
$$

where a and *b* are characteristics of a given solute *i.* In RP the utilization of the solubility parameter concept of Schoenmakers and co-workers [7,8] resulted in the quadratic equation

$$
\log k' = a\varphi^2 + b\varphi + c \tag{2}
$$

However numerous studies have established that the empirical relationship

$$
\log k' = \log k_{\rm w} - S\varphi \tag{3}
$$

is valid within the range 20-80% of the volume of the organic modifier.

 $\varphi$  is the volume fraction of the strong solvent and  $k_{w}$  and S are constants that are characteristic of the strong solvent and solute respectively. For the sake of simplicity we shall write:

$$
k' = (1 - R_F)/R_F \tag{4}
$$

Since  $X_3$  is similar to  $\varphi$  and  $R_M = a_i \ln X_s + b_i$ .

$$
R_F = 1/[1 + \exp(a_i \ln X_s + b_i)] \tag{5}
$$

The response function selected must be based on selectivity. A resolution criterion would be meaningless since information on spot widths is rather scarce in TLC and resolution is  $R<sub>F</sub>$ -dependent with a given chromatographic system.

Separation between two components is expressed as:

 $AR_{Fij} = R_{Fi} - R_{Fj} = 1/[1 + \exp(a_i \ln X_s + b_i)] - 1/[1 + \exp(a_i \ln X_s + b_i)]$  (6)

We can select the mobile phase composition to maximize  $AR<sub>F</sub>$ . It has been shown by Nurok and Richard [9] that plots of  $AR_F$  versus mobile phase composition of a binary eluent exhibits a maximum. We shall maximize the  $AR_F$  between the first (or the last) eluted solute and the others.

The procedure is as follows: since for a given  $X_s$  there will be different  $a_i$  and  $b_i$ values we have utilized a non-fractional two-level factorial design  $2^k$ . Considering a pair of solutes we have  $a_i$ ,  $b_i$ ,  $a_j$  and  $b_j$  values, thus yielding a  $2^4$  factorial design. By fixing every a and  $b$  at + or - level, respectively, the matrix has 4 rows and 16 lines (see Table I). We are then able to construct the model matrix and determine the highest and lowest boundaries for a and *b* for every solute of the sample. The lowest values of  $a_i$  and  $b_i$  and the largest variation are determined. These yield the center of the variation domain for all a and *b,* and the steps within this variation. We chose to consider at the  $-$  level the values of a and b within the domain between the lowest boundary and the center of the domain. The  $X_s$  values are calculated which yield *A RFmax* for the *i, j* pair of solutes. As many responses as are in the model matrix are obtained, and the NEMROD software calculates the coefficient of the model by multilinear regression.

# TABLE I

	$X_1$	$X_{2}$	$X_3$	$X_4$	a,	$b_i$	$a_i$	$b_j$	$X_{n}$ (computer)	$\Delta R_{F\mathrm{max}}$
					$-1.200$	$-1.819$	$-1.490$	$-2.395$	0.53	0.07
2	$+$				$-0.750$	$-1.727$	$-1.510$	$-1.980$	0.08	0.33
3	-	$\ddot{}$			$-1.340$	$-1.243$	$-1.200$	$-1.819$	0.02	0.20
4	$+$	$\ddot{}$			$-1.080$	$-0.253$	$-1.400$	$-2.395$	0.44	0.45
5	-	-	$^{+}$		$-1.200$	$-1.819$	$-0.750$	$-1.727$	0.07	0.20
6	$^{+}$		$+$		$-0.700$	$-1.543$	$-0.798$	$-1.814$	0.63	0.05
7		$\ddot{}$	$\ddot{}$		$-1.120$	$-1.658$	$-0.960$	$-1.704$	0.10	0.09
8	$+$	$+$	$^{+}$		$-1.080$	$-0.253$	$-0.788$	$-1.814$	0.23	0.46
9	-			$\ddot{}$	$-1.270$	$-0.737$	$-1.200$	$-1.819$	0.33	0.30
10	$+$			$\ddot{}$	$-1.200$	$-1.819$	$-0.950$	$-1.727$	0.08	0.11
11	-	$\ddot{}$		$\ddot{}$	$-1.340$	$-1.243$	$-1.120$	$-1.658$	0.21	0.18
12	$+$	$+$		$\ddot{}$	$-1.080$	$-0.253$	$-1.340$	$-1.243$	0.80	0.23
13			$\ddot{}$	$\ddot{}$	$-1.400$	$-2.395$	$-0.700$	$-1.543$	0.05	0.23
14	$\ddot{}$	-	$\ddot{}$	$+$	$-1.510$	$-1.985$	$-1.080$	$-0.253$	0.61	0.36
15		$^{+}$	$+$	$\ddot{}$	$-1.080$	$-0.253$	$-1.120$	$-1.658$	0.44	0.34
16	$\ddot{}$	$+$	$+$	$\ddot{}$	$-0.960$	$-1.704$	$-0.700$	$-1.543$	0.04	0.11

EXPERIMENTAL MATRIX FOR THE NORMAL-PHASE MODEL

NEMROD is able to work with either coded (or reduced) variables or with the actual variables. The reduced variable will be written as  $a = (y - y_0)/c_y$ , where  $y_0$  is the value at the center of the domain and  $c<sub>v</sub>$  is the step of variation.

## RESULTS AND DISCUSSION

We shall consider some contaminants, especially the nitro and chloro derivatives of phenol (see Table II). From the TLC literature data we can consider either the NP or the RP mode. In Table II are the experimental  $a_i$  and  $b_i$  values from NP-TLC [lo] with silica gel as stationary phase and heptane-ethyl acetate as solvent mixture.

To check the influence of  $a_i$  and  $b_i$  values on eqn. 6 it is necessary to have a knowledge of the different levels that a and *b* can attain. To this purpose a complete  $2<sup>4</sup>$  factorial design has been constructed with yields a set of combinations  $a_i$ ,  $b_i$ ,  $a_j$  and  $b_i$ . From this factorial design, 16 pairs of compounds are characterized by their  $a_i$ ,  $b_i$ values. A simple program permits the calculation of the responses from eqn. 6. The response is the value of  $X_s$  which permits the attainment of  $AR_{Fmax}$ . Plots of  $AR_{Fmax}$ *versus* mobile phase composition are shown in Fig. 1. From Table II it is seen that variation of a is from  $-1.51$  to  $-0.70$  and b from  $-2.395$  to  $-0.253$ . From these values the center of the variation domain is easily deduced and  $+$  and  $-$  levels are attributed.

Data from Table III are treated with NEMROD software. In this procedure an empirical model is postulated of the form  $X_s = f(a_i, b_i, a_j, b_j)$ . The mathematical form is a first- or second-order polynomial. NEMROD selects the model which yields the best statistical parameters (variance, covariance, etc.). From the NEMROD computation the highest quality is obtained, nevertheless, as is obvious particularly in TLC, experiments are often required to check the validity of the model.



#### TABLE II

CHARACTERISTIC VALUES OF CHLOROPHENOLS AND NITROPHENOLS ON SILICA GEL WITH BINARY MOBILE PHASE HEPTANE-ETHYL ACETATE



Fig. 1. Plots of  $AR_{Fmax}$  versus  $X_s$ .

From the NEMROD results the best polynomial was:

$$
X_s = X_0 + X_1 a_i + X_2 b_i + X_3 a_j + X_4 b_j
$$

We must point out that it is possible to work with normalized reduced variables. In this case the comparison between the two procedures yielded the same results and the model with actual variables performed better (Table III). Selected values are as follows:

 $X_0 = 0.71835$  $X_1 = 0.61186$  $X_2 = 0.02041$  $X_3 = -0.635537$  $X_4 = 0.193513$ 

Checking the model is necessary to make sure that estimated  $X_s$  values are identical to those from curves  $AR_{Fi,j} = f(X_s)$ . We performed this checking with the "ideal" case (where  $\Delta R_F < 0.3$ ), this 0.3 value was selected from experience since the maximization of  $AR<sub>F</sub>$  is insufficient below 0.3.

#### TABLE III

Solute pair	$X$ , plot	$X_{r}$ (nrv)	$X_{n}$ (ac)	$AR_{\text{max}}$	
$2,6-NtP/2,4,6-CP$	0.174	0.100	0.100	0.42	
2.6-NtP/Penta-CP	0.226	0.180	0.200	0.34	
2,6-NtP/3,5-CP	0.154	0.080	0.080	0.40	
2-NtP/Penta-CP	0.320	0.320	0.320	0.38	
$2-NtP/2.4-CP$	0.330	0.330	0.330	0.38	
$2-NtP/3.5-CP$	0.208	0.200	0.200	0.42	
$2-NtP/2,6-NtP$		0.720	0.720	0.10	
$2,6-NtP/4-NtP$	0.560	0.480	0.540	0.10	
$4-NtP/3-NtP$	0.430	0.450	0.450	0.14	
$Ph/4$ -CP	0.645	0.410	0.410	0.02	
$3.5$ -CP/2,4-CP	0.410	0.540	0.540	0.12	

COMPARISON OF RESPONSES OBTAINED BY PLOT OF *AR,* AND MODELS FROM NOR-MALIZED REDUCED VALUES (nrv) AND FROM ACTUAL VALUES (ac)

Fig. 2 shows the comparison between the response from the model and the first derivative of the  $AR_F = f(X_s)$  function. When the first derivative = 0 it represents the maximum of that function.

Discrepancies between responses were observed: (1) when  $X_s$  is outside the range 0.2-0.8, since beyond that range the experimental data do not fit a linear regression (an example is given in the first row of Table III, as  $X_{s,model} = 0.1$  and  $X_{s,plot}$  is 0.174); (2) when the  $AR_F$  difference is less than 0.1 (for example, of the phenol and 4-chlorophenol).

# *Application to NP chromatography*

*Nitrophenols (NtPs) and chlorophenols (Cps) with silica gel as stationary phase and heptane-ethyl acetate as mobile phase.* From the mathematical model and the fixed contraints ( $AR_F > 0.1$  and  $0.2 \le X_s \le 0.8$ ) we can consider a sample clean-up on 2-NtP and 2,6-NtP. From Table IV it looks obvious that these contaminants can be separated from chlorophenols and phenol. From the  $X_s$  values a stepwise procedure can be considered according to the wishes of the analyst. We can retain the 2-NtP ( $k' = 4.42$ ) with a heptane-ethyl acetate (80:20, v/v) mobile phase and by consequence 3,5-CP and 2,4,6-CP would be eliminated  $(k' = 0.65$  for both). By a simple change to a  $(67.33, v/v)$  solvent composition elimination of pentachlorophenol, 2,4-CP and 3-CP would be obtained  $(k' =$  approximately 0.53). Of course the capacity factor of 2-NtP shifted from 4.42 to 2.57 but the gap is still large. In the third step with 54:46 ( $v/v$ ) solvent composition phenol, 4-CP and 3.4-Cp are quickly eluted  $(k' = 0.61)$ . Finally, 2-NtP is the only solute retained which can be eluted with pure ethyl acetate.

In the same way, it would be easy to separate 2,6-NtP from phenol, 3-CP, 4-CP, 3,4-CP, 2,4-CP and pentachlorophenol (Fig. 3) in a stepwise procedure.

*Nitrophenols and chlorophenols with silica gel as stationary phase and a heptanediisopropyl ether as mobile phase.* In the above procedure we could not separate 4-NtP from the others. A careful change of the polar modifier will permit us to achieve this goal. A remarkable feature of the model elaborated from NEMROD computation is that it can be used with any NP system provided that  $a_i$  and  $b_i$  for the different polar modifier are available. Nevertheless it must be pointed out that some discrepancies may occur when silica gel of different activity is used as chromatographic packing. When care is taken to normalize the experiments the mathematical model performs very well. In our calculation we used data from Matyska and Soczewinski [10] which



Fig. 2. Comparison of responses from the model and the first derivative of the  $AR<sub>F</sub> = f(X<sub>s</sub>)$  function.

TABLE IV



Solute pair	$X_{n}$ (computer)	$X_{\epsilon}$ (model)	$\Delta R_{F \text{ max}}$		
$2-NtP/Ph$	0.44	0.44	0.33		
$2-NtP/2-CP$	0.22	0.19	0.22		
$2-NtP/3-CP$	0.31	0.31	0.31		
$2-NtP/4-CP$	0.46	0.46	0.35		
$2-NtP/3,4-CP$	0.47	0.46	0.36		
2-NtP/3.5-CP	0.21	0.20	0.42		
$2-NtP/2,4-CP$	0.33	0.33	0.37		
$2-NtP/2,4,6-CP$	0.23	0.21	0.45		
2-NtP/Penta-CP	0.32	0.32	0.38		
$2.6-NtP/Ph$	0.3	0.32	0.27		
2,6-NtP/2-CP	0.16	0.07	0.42		
$2,6-NtP/3-CP$	0.22	0.18	0.35		
2.6-NtP/4-CP	0.33	0.34	0.28		
2,6-NtP/3,4-CP	0.34	0.35	0.28		
2,6-NtP/2,4-CP	0.23	0.21	0.33		
2.6-NtP/Penta-CP	0.22	0.20	0.34		

OPTIMAL MOBILE PHASE FOR SEPARATION OF CHLOROPHENOLS AND NITROPHENOLS

are very coherent. It must be remembered that the above requirements are still valid  $(i.e. 0.2 \leq X_s \leq 0.8$  and  $AR_F > 0.3$ ). Table V shows both experimental and computational values. From these data it is possible to selectively retain 4-NtP from 2,4,6-CP, 3,5-CP and 2-Cp.



Fig. 3. Selective extraction of 2,6-NtP from a mixture of chlorophenols and nitrophenols.

TABLE V

Solute pair	$a_i$	$b_i$	a,	$b_{j}$	$X_{\epsilon}$ (computer)	$X_{n}$ (model)	$\mathcal{A}R_{F\text{max}}$
$4-NtP/2,4,6-CP$	$-1.150$	$-0.253$	$-0.960$	$-1.497$	0.34	0.33	0.34
$4-NtP/4-IP$	$-1.150$	$-0.253$	$-1.200$	$-0.944$	0.67	0.59	0.17
$4-NtP/4-BrP$	$-1.150$	$-0.253$	$-1.170$	$-1.51$	0.57	0.53	0.22
$4-NtP/3-NtP$	$-1.150$	$-0.253$	$-1.210$	$-0.760$	0.78	0.63	0.12
4-NtP/Ph	$-1.150$	$-0.253$	$-1.120$	$-1.013$	0.54	0.53	0.19
$4-NtP/2-CP$	$-1.150$	$-0.253$	$-0.760$	$-1.243$	0.25	0.25	0.34
$4-NtP/3-CP$	$-1.150$	$-0.253$	$-1.050$	$-1.174$	0.44	0.45	0.24
$4-NtP/4-CP$	$-1.150$	$-0.253$	$-1.020$	$-1.105$	0.42	0.44	0.23
4-NtP/3,4-CP	$-1.150$	$-0.253$	$-1.110$	$-1.105$	0.51	0.50	0.22
$4-NtP/3,5-CP$	$-1.150$	$-0.253$	$-1.140$	$-1.704$	0.42	0.40	0.35
$4-NtP/2,4-CP$	$-1.150$	$-0.253$	$-0.980$	$-1.106$	0.39	0.42	0.24

EXTENDED MODEL FOR TREATMENT OF PARENT COMPOUNDS WITH A DIFFERENT POLAR MODIFIER

Moreover, a careful selection of the mobile phase composition permits us to obtain two peaks: one for 4-NtP and the other of the three remaining chlorophenols which have not been separated  $(k' = 2.23$  and 0.5, respectively).

In the above procedures we considered derivatives of phenol. We now attempt to separate some phenol derivatives from very different chemical species (Table VI). Conditions for application of the previous model are checked. 1-Naphthol is eluted whilst 2,6-NtP and 2-NtP are selectively retained. We note that isoquinoline can be separated from 8-methylquinoline. Phenol is eluted whilst isoquinoline is retained. Conversely it can be seen that for  $1,2-HB$ -phenol a 8% difference in the mobile phase composition is found between the computed and actual values. This may be attributed to differences in plate activity since data from different origin are gathered.

## *Application to RP chromatography*

Since selectivity observed with NP is very different from the selectivity observed in RP it would be valuable to consider the same solutes as in Table II with a reversed-

TABLE VI

EXTENDED MODEL FOR TREATMENT OF NON-PARENT COMPOUNDS WITH HEPTANE-ETHYL ACETATE



 $\text{I/O} = \text{isquinoline}; 8\text{-}MO = 8\text{-}methylquinoline}; 1\text{-}NH = 1\text{-}naphthol}; 1,2\text{-}HB = 1,2\text{-}dihylrovybenzene.$ 

# TABLE VII

### EXPERIMENTAL MATRIX FOR THE REVERSED-PHASE MODEL

For conditions see text.



# TABLE VIII

SELECTIVE EXTRACTION OF PENTA-CP OR 2,4,5-CP FROM A MIXTURE OF CHLOROPHE-NOL



phase system. The strategy for mathematical modeling is the same as above. Data were from the paper of Arai et al. **[l 11.** Table VII displays both the experiment matrix and the responses, a first-order polynomial best fits the experimental data:

$$
X_{\rm sRP} = 0.686251 + 0.064745a_i + 0.115527b_i - 0.124439a_j + 0.059049b_j
$$

We note the same features as previously observed, *i.e.* when  $AR_F \le 0.1$  there is no correlation between the response given by plot of  $AR<sub>F</sub>$  versus  $X<sub>s</sub>$  and the response of the model.

It can be seen from Table VIII that the pentachlorophenol or the 2,4,5-CP is readily separated from all others.

A solid-phase extraction with same packing as advocated by Arai *et al.* [ 1 l] will permit selective retention of the pentachlorophenol with a mobile phase composition 37:63 (v/v) and will elute 2-CP, 3-CP, 4-CP, 2,3-CP, 2,4-CP, 2,5-CP, 2,6-CP, 3,4-CP, 4-bromophenol and 3-bromophenol.

### **CONCLUSIONS**

This computer-assisted method exhibits some obvious advantages. The selection of optimal mobile phase composition for sample clean-up is effective. A large amount of TLC data is available and the model performs well with different organic modifiers in the mobile phase. The drawbacks are that secondary effects are not taken into account. We did not consider large differences in solute amounts, which will change retention according to the isotherm forms. This will be the subject of a forthcoming paper.

#### ACKNOWLEDGEMENTS

We are grateful to Dr. M. Matyska for helpful discussion. One of us (N.V.) is indebted to Parfums Christian Dior for a grant to support this work.

### **REFERENCES**

- 1 R. W. Frei and K. Zech, in R. W. Frei and K. Zech (Editors), *Selective Sample Handling and Detection in High-Performance Liquid Chromatography,* Part A, Elsevier, Amsterdam, p. 1.
- 2 R. D. McDowall, J. *Chromatogr., 492* (1989) 3.
- 3 A. Betti, G. Lodi and S. Coppi, *J. Chromatogr.*, 513 (1990) 219.
- *4* A. Robbat, Jr. and T.-Y. Liu, J. *Chromatogr., 513 (1990)* 117.
- 5 E. Soczewinski and W. Golkiewicz, J. *Chromatogr.,* 118 (1976) 91.
- 6 E. Soczewinski, J. *Chromatogr., 130* (1977) *23.*
- *7* P. J. Schoenmakers, H. A. H. Billiet and L. de Galan, J. *Chromatogr., 205* (1981) 13.
- 8 P. J. Schoenmakers, A. C. J. H. Drouen, H. A. I-f. Billiet and L. de Galan, *Chromatographia, 15* (1982) *688.*
- *9* D. Nurok and M. J. Richard, *Anal. Chem., 53* (1981) 564.
- 10 M. Matyska and E. Soczewinski, J. *Planar. Chromatogr., 3 (1990)* 144.
- 11 Y. Arai, M. Hirukawa and T. Hanai, J. *Liq. Chromatogr.,* 10 (1987) *635.*