Journal of Chromatography, 264 (1983) 1–6
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 15,818

COMPUTER ASSISTANCE IN THE SELECTION OF THE OPTIMUM COMBINATION OF SYSTEMS FOR TWO-DIMENSIONAL CHROMATOGRAPHY

MARIE-FRANCE GONNORD, FLORENCE LEVI and GEORGES GUIOCHON*
École Polytechnique, Laboratoire de Chimie Analytique Physique, Route de Saclay, 91128 Palaiseau Cédex (France)
(Received March 4th, 1983)

SUMMARY

To permit the selection of the optimum combination of chromatographic systems in two-dimensional chromatography among those available, a data bank has been created on a personal microsystem from k' or R_F values published in the literature or measured in the laboratory.

Simulations of two-dimensional chromatograms can be shown on either the system display or on a plotter. The selection of the best pair of chromatographic systems to resolve different compounds from a mixture is performed by the system after selection by the user of an optimization criterion. The pertinence of various criteria is compared.

INTRODUCTION

The capacity of a separation system can be defined as the number of compounds that it can separate from each other with a resolution of unity. In thin-layer chromatography (TLC), a peak capacity of 10 is easily achieved whereas a capacity of 30 is almost impossible to reach[1], which gives this simple technique a poor capacity compared with other chromatographic methods. On the other hand, TLC has a unique ability: two developments of the same mixture can be carried out successively along two perpendicular directions, using two different mecanisms or two solvents with different selectivities, provided that the sample is spotted in the corner of a square or rectangular plate. Two-dimensional TLC thus offers a surprisingly large spot capacity, which gives it a performance level comparable to that of modern column chromatography: it is easy to achieve a spot capacity between 100 and 250 but nearly impossible to exceed 500 (ref. 2).

The combination of this large potential spot capacity and flexibility in the choice of independent elution mechanisms explains why two-dimensional TLC has been used to carry out a large number of difficult separations and has been a useful tool in the elucidation of several biological pathways[3].

Basically, two techniques can be used: either the same chromatographic bed is developed successively with two different solvent mixtures along the two perpen-

0021-9673/83/$03.00

dicular directions, or the plate is coated with a strip of a sorbent on one edge and a large zone of a different sorbent, and again two developments are carried out with two different solvent systems. Such plates are difficult to prepare although they offer very attractive possibilities[4].

In both instances we need to optimize the combination of two systems in order to resolve the largest possible number of compounds in the sample, which often implies the use of almost all the total chromatographic area available and always requires the appropriate selection of the two different chromatographic sorbent–mobile phase systems, for each perpendicular elution, among a relatively large number of possible candidates.

TLC, together with high-performance liquid chromatography (HPLC), can supply the analyst with a large amount of retention data obtained for one-dimensional separations using a wide variety of combinations of solvents and sorbents. Indeed, TLC is often used as an auxiliary technique to predict the retention behaviour of solutes in the corresponding system used in column chromatography, with the only requirement that proper correspondence between R_F and k' be established[5].

The selection of operating conditions requires optimization of the two independent one-dimensional chromatographic systems and also of their combination. One can easily predict how long this preliminary study can take if it is carried out by trial and error.

To simplify and shorten difficult assays, we have developed an automatic computer program based on a compact data bank that we have created and stored on the disc unit of a home computer system. The aim of this paper is to present the organization of the data and the different features of the program and to describe an application of this procedure to the selection of optimum conditions for the separation of DNP-amino acids by two-dimensional TLC. In order to define an adequate separation, two different criteria have been compared based on different functions of the average distance between the various positions of the spots of the resolved compounds. The choice of the more pertinent criterion is discussed.

THEORETICAL

In TLC retention data are referenced to the migration distance L of the solvent front through the R_F value

$$R_F = z/L \tag{1}$$

where z is the migration distance of the spot. In column chromatography the column capacity factor, k', is used, where

$$k' = \frac{t_R - t_0}{t_0} \tag{2}$$

where t_R and t_0 are the retention time of the mass-centre of the solute band and the band of an unretained compound.

As long as data are collected using TLC in order to achieve two-dimensional

TLC separations or using column chromatography, the data are homogeneous and will be used without transformation. When data from both techniques must be combined or if we want to use TLC data to predict retention in column chromatography, the problem is more complex. As a first approximation we may write

$$R_F = \frac{1}{1 + k'} \tag{3}$$

but this is only an approximation, as column chromatography is carried out under steady-state conditions, while plate saturation is achieved in TLC only for small R_F values and the solvent density in the porous layer varies with length and time. It has been shown that better results are obtained if a correction factor is used[5–7], with

$$R_{F(\mathrm{TLC})} = \xi \; R_{F(\mathrm{CC})} \tag{4}$$

Usually ξ is estimated to be between 1.3 and 1.5.

It is easy to take any required value of ξ into account in our program when necessary. As we have been using TLC data to optimize two-dimensional TLC, however, it is not useful at this stage.

SYSTEM CONFIGURATION AND SOFTWARE

A Commodore CBM3032 (Commodore Business Machines, Norristown, CA, U.S.A.) was chosen as it is simple and inexpensive system. Several peripherals were necessary for this work: a CBM8050 double diskette unit (2 times 500 kilo octet) and a CBM 8023 line printer were interfaced through the IEEE-488 bus. It is also possible to connect to the same bus a graphical plotter, when necessary. Programs were written in BASIC language.

The programs have been developed with special emphasis on flexibility, and written with the intent of allowing easy modifications. The software can thus be broken down into several operations: entry from the keyboard and storage of the data to constitute a data bank, search of data in the data bank and loading in the central memory, graphical display of any required two-dimensional separation, comparison between all possible combinations of any two chromatographic systems and selection of the best combination. The software organization is modular in such a way that the user can select its own sequence procedure. The different modules are presented in Fig. 1.

Data organization

Different sets of data are stored in different files. Direct access files have been used to reduce time access to the file during the search routine. Data files were created in a format that could be read in BASIC directly. Data were stored as chains of characters..

Each chemical class (the compounds are classified according to their chemical nature: PAH, steroids or DNP-amino acids, for example) is associated with one track. On each track, each sector is devoted to the retention data obtained using one chromatographic system. Each compound is codified and each datum consists of two numerical values: the compound code and its retention value. The data concerning the same chromatographic conditions and the same chemical family are thus stored in the same sector, while data pertaining to a given set of chromatographic conditions are broken into as many sectors as there are distinct chemical families represented

Module number	*Module name*	*Purpose*
1	CREASEC	Makes the reservation sectors on a track defined for a given chemical class
2	CREASOL	Files experimental data on the given sector and completes the library
3	SEARCH	Searches through the library and the data bank the experimental data and chromatographic conditions for the required compounds and loads the results of the search procedure in the computer memory
4	AAD	Computes and plots the pattern of any two-dimensional separation. Offers different options for limited number of solvents combinations
5	OPTICAP	Performs the optimization of the choice of a pair of solvent. The choice of the criterion selected by the user is optional. If no choice, both results are displayed with experimental conditions

Fig. 1. Software modules.

in the bank.

A library file is reserved in which correspondences between chemical classes and track numbers are made. Also for each track, correspondences between sector numbers and chromatographic conditions are established together with the codification of the compound name. The library file is a sequential access file. The coordinates and entries of any new file are stored after the last information existing before, in the library file, upon creation of the new file.

When the executing program is searching for a given compound, it first goes through the library, identifies the track number associated with the compound chemical class and the compound code and then goes through the sectors of the track, identifying and loading in the central processor unit memory the data concerning the compound under study, together with the references of the chromatographic conditions.

When a series of compounds is given as an entry, if some compounds are absent from some of the sectors of the data bank, *i.e.*, there are no retention data available for them with some of the chromatographic systems investigated, priority is given to the chromatographic systems containing the larger number of compounds required by the entry. Up to 35 classes of chemical compounds and 17 chromatographic systems per class can be stored on one diskette.

Graphical display for a two-dimensional separation

Upon requirement by the user, the bank is searched for retention data of the listed compounds in the two selected systems. All data are translated into R_F values according to eqn. 3 to ensure homogeneity of presentation. A test on R_F ($0 \leqslant R_F \leqslant 1$) is performed. Thus two R_F values are assigned to each compound. They are treated as coordinates in an orthogonal graph and a map of the two-dimensional separation appears on the computer display. The position of each solute is represented as a "point" character on the display. The resolution of the CBM 3032 display being very poor, wherever a separation looks interesting it is also possible to obtain, through the graphic plotter, the plot of the map, together with all the information concerning the separation, such as chromatographic system and code number of the compounds.

The number of compounds to be treated in one run is only limited by the performances of the computer. Our programs can treat up to 40 components; this value could be extended easily. With a CBM 8096 microcomputer several hundred compounds could be treated simultaneously. An example of the plotter output is presented in Fig. 2.

Optimization of the solvent selection for a two-dimensional separation

The two-dimensional TLC separation is of no interest if the selection of the two different eluting systems is not adequate. For example, if the solvent selectivities in both directions are too similar the spots will align along the diagonal of the plate and the increase in the time of analysis is generally not worth the gain in resolution. A good separation will be obtained when the surface area of the plate over which the spots are spread is relatively large.

It is thus useful to calculate an estimate of the quality of a two-dimensional separation and, as the software developed in this work allows a fast compilation of the different possible pairs of solvents to be used, and a very rapid computation of simple parameters, to test immediately the quality of the analysis and compare those of the different 20-dimensional chromatographic system combinations.

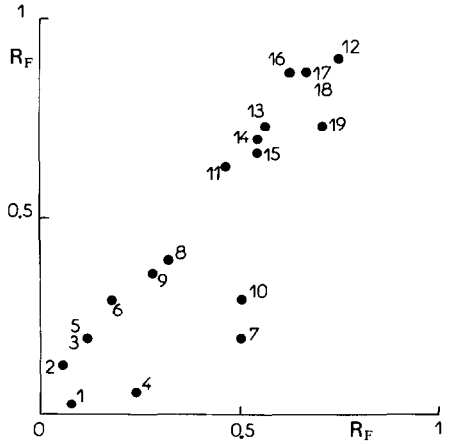
The more readily accessible measure of the quality of a separation is the distance between the different spots. Two approaches have been considered, both making use of these distances.

The first method consists in maximizing the sum of the square of all the possible distances between any pair of spots:

$$D_A = \sum_{i=1}^{k} \sum_{j=i+1}^{k} [(x_i - x_j)^2 + (y_i - y_j)^2] \quad (5)$$

The same statistical weight is given to each distance between two spots in this mode, irrespective of its value.

Another method can be considered: the sum of the inverse of the square of the distances between pairs of spots is minimized after elimination of the unresolved pairs:



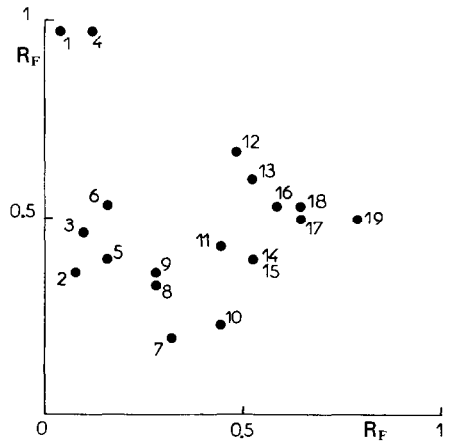


Fig. 2. TLC pattern of a two-dimensional separation of 19 DNP-amino acids after computer selection from Model D_A (eqn. 5). Solvents IV and V from ten solvents in ref. 8 were selected.

Fig. 3. TLC pattern of a two-dimensional separation of 19 DNP-amino acids after computer selection from model D_B (eqn. 6). Solvents I and IX from ten solvents in ref. were selected.

$$D_B = \sum_{i=1}^{k} \sum_{j=i=1}^{k} \frac{1}{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (6)$$

Eqn. 5 tends to give no importance to unresolved or poorly resolved spots and a pattern in which a few very excentric spots are separated from a dense spot cloud can be retained rather than a more evenly spread pattern. In contrast, eqn. 6 gives more weight to small distances and thus privileges the patterns in which the spots are more dispersed. The program also evaluates the number of unresolved pairs and can easily be made to discriminate in favour of the systems with which the larger number of spots are resolved, for example by assigning arbitrarily a distance equal to the average spot width, or even smaller, to all pairs of unresolved compounds.

RESULTS AND DISCUSSION

To exemplify the usefulness of the programs and to compare the results obtained using the two optimization methods, we have applied our procedure to a set of experiments on a series of nineteen DNP-amino acids chromatographed on polyamide layers using ten different solvent systems[8]. The published experimental data have been introduced into the data bank. Each pair of solvents has been computed and the different separations compared on the basis of either D_A or D_B. The corresponding two-dimensional separations selected as best by the program are presented in Figs. 2 and 3, respectively. It is obvious from a mere visual comparison between the two patterns than the optimization according to D_B gives a better resolved chromatogram and that the layer surface area is better covered. The spot pattern obtained with the combination of solvents 4 and 5 shows a poor difference in the selectivity of the two solvent systems as the spots are nearly all situated along the diagonal of the layer. Of course, the user has always the possibility of discarding combinations of solvent systems that are not compatible, as could be the case if the computer assistance is required to run a true two-dimensional liquid chromatograph[9].

The data base is easily modified by frequent introduction of new data. In the near future it is planned to use it in combination with the results of a new theoretical approach permitting the calculation of retention data in binary solvent mixtures from retention data in the pure solvents[10]. Then, each solvent used for one of the two perpendicular developments could be a mixture of the solvents used to acquire the data, with an optimized composition. A more powerful computer would thus be necessary.

REFERENCES

1 G. Guiochon and A. M. Siouffi, *J. Chromatogr.*, 245 (1982) 1.
2 G. Guiochon, M. F. Gonnord, A. Siouffi and M. Zakaria, *J. Chromatogr.*, 250 (1982) 1.
3 M. Zakaria, M. F. Gonnord and G. Guiochon, *J. Chromatogr.*, 271 (1983) 127.
4 A. M. Siouffi and G. Guiochon, to be published.
5 S. Roggia and G. G. Gallo, in A. Frigerio and M. McCamish (Editors), *Recent Developments in Chromatography and Electrophoresis, 10*, Elsevier, Amsterdam, 1980, p. 47.
6 A. M. Siouffi, A. Guillemonat and G. Guiochon, *J. Chromatogr.*, 137 (1977) 35.
7 L. R. Snyder, in J. C. Giddings and R. A. Keller (Editors), *Principles of Adsorption Chromatography*, Marcel Dekker, New York, 1968, p. 365.
8 K. T. Wang, J. M. K. Huang and I. S. Y. Wang, *J. Chromatogr.*, 22 (1966) 362.
9 G. Guiochon, L. A. Beaver, M. F. Gonnord, A. M. Siouffi and M. Zakaria, *Chromatographia*, 17 (1983) 121.
10 P. Jandera, H. Colin and G. Guiochon, *Chromatographia*, 16 (1982) 132.