Journal of Chromatography, 209 (1981) 377–384
Elsevier Scientific Publishing Company, Amsterdam — Printed in The Netherlands

CHROM. 13,664

STUDY OF PAPER AND THIN-LAYER CHROMATOGRAPHY OF PHENOLIC SUBSTANCES BY STATISTICAL METHODS

L. NAGELS
Rijksuniversitair Centrum Antwerpen, Laboratorium voor Algemene Scheikunde, Groenenborgerlaan 171, 2020 Antwerp (Belgium)
(Received January 19th, 1981)

SUMMARY

Classical paper chromatographic (PC) and thin-layer chromatographic (TLC) separations of phenolic compounds (including flavonoids) were classified into clusters according to their selectivities. Systems were compared two by two, by plotting the R_F values of a set of compounds in one system against the R_F values in the other system. One hundred and seventy correlation coefficients were computed and used in a hierarchical clustering method. The dendrograms obtained can be used by analysts to plan an efficient scheme for the identification or separation of phenolic substances. The potential of PC and TLC in this respect is discussed and compared with that of high-performance liquid chromatography.

INTRODUCTION

The literature contains descriptions of a wide variety of chromatographic systems for the analysis of phenolic substances of industrial, plant or animal origin. Most of these systems are based on paper chromatography (PC) and thin-layer chromatography (TLC). Gas chromatographic (GC) data also exist, and high-performance liquid chromatography (HPLC) is being used increasingly[1].

A qualitative comparison of HPLC with PC and TLC seemed appropriate for several reasons. There is no doubt that the quantitative analysis of phenolic substances can be carried out more accurately with HPLC than with PC and TLC systems. However, HPLC may not be competitive with PC and TLC in qualitative analysis (identification). PC and TLC offer the possibility of two-dimensional elution and selective detection, and simultaneous analysis on different systems can also easily be achieved. These factors can increase the information provided by the technique[2]. In the PC and TLC of phenolic substances, different chromatographic systems can be obtained by combining a variety of mobile phases with a single stationary phase. This will increase the information obtained only if changing the mobile phase also results in a significant change in selectivity, and this is only vaguely known for most systems at present (for an explanation of the selectivity of a chromatographic system, the reader is referred to ref. 3). The aim of this paper is to compare the selectivities that

0021-9673/81/0000–0000/$02.50

are obtained in the PC and TLC of phenolic compounds by varying either the mobile phase or the stationary phase. We therefore used a well defined set of published classical chromatographic systems. They are classified into "clusters" with analogous selectivities by means of statistical methods.

RESULTS AND DISCUSSION

Some classical papers on the PC and TLC of phenolic substances were chosen arbitrarily from the literature[4–7]. Two of them were concerned with PC[4,7], another with cellulose TLC[6] and the last[5] with TLC on mixed adsorbents. We compared the selectivities of the different systems used by one author, and when possible we also compared the selectivities of systems used by different authors.

Comparisons were made by plotting the R_F values of a set of phenolic compounds in one system against the R_F values in another system. Correlation (or absence of correlation) between the two was expressed by the correlation coefficient (ϱ). In our opinion, this coefficient is a good parameter of what we call qualitatively the "selectivity difference" between two chromatographic systems. These coefficients were used in a hierarchical clustering method as described by Massart and De Clercq[8], and the results were plotted in a "dendrogram".

Table I indicates which TLC and PC systems were compared in this way. Dendrograms showing the correlation of the systems of each publication separately are shown in Fig. 1. Fig. 2 shows correlations calculated from the overlap of data from two publications. These two figures were obtained from the algorithmic treatment of 119 correlation coefficients (119 different pairs of chromatographic systems were investigated using a computer program). Therefore, Figs. 1 and 2 represent the compression of a large amount of chromatographic data relating to phenolic compounds. They can be interpreted by chromatographers without requiring a thorough knowledge of statistics.

When there was a small spread of R_F values in a certain system, there was no point in looking for a correlation with another system. Therefore, we needed a measure of the spread of R_F values over the chromatograms. We divided the R_F scale into four equal parts and calculated for each system the number of R_F values in each part. We then compared this distribution with a hypothetical one in which the set of R_F values was distributed equally over the four parts (null hypothesis). As a measure of divergence from this null hypothesis we used the χ^2 value. These values are given in Table I. They were generally better (*i.e.*, smaller) in some publications (Jangaard, Van Sumere *et al.*) than in others (Reio, Jay *et al.*). As a measure of the eluting power of a chromatographic solvent system, the mean R_F value ($\overline{R_F}$) was calculated (see Table I). Based on the χ^2 values, we rejected two systems from Reio (RA and RF) for further comparison. Although they were not omitted from the dendrogram in Fig. 1, they were not used further in this work.

The resulting systems from this author could be divided into two selectivity groups, *i.e.*, [RB, RC, RD] and [RE] (see Fig. 1). Especially RB and RC showed a high correlation, and RE was the only system with completely different selectivity (see also the "scatter diagrams" in Fig. 3). Systems RE and RD had the lowest χ^2 value and when used in conjunction they should yield the largest amount of information for this set of phenolic compounds. One might argue that the subdivision of Reio's

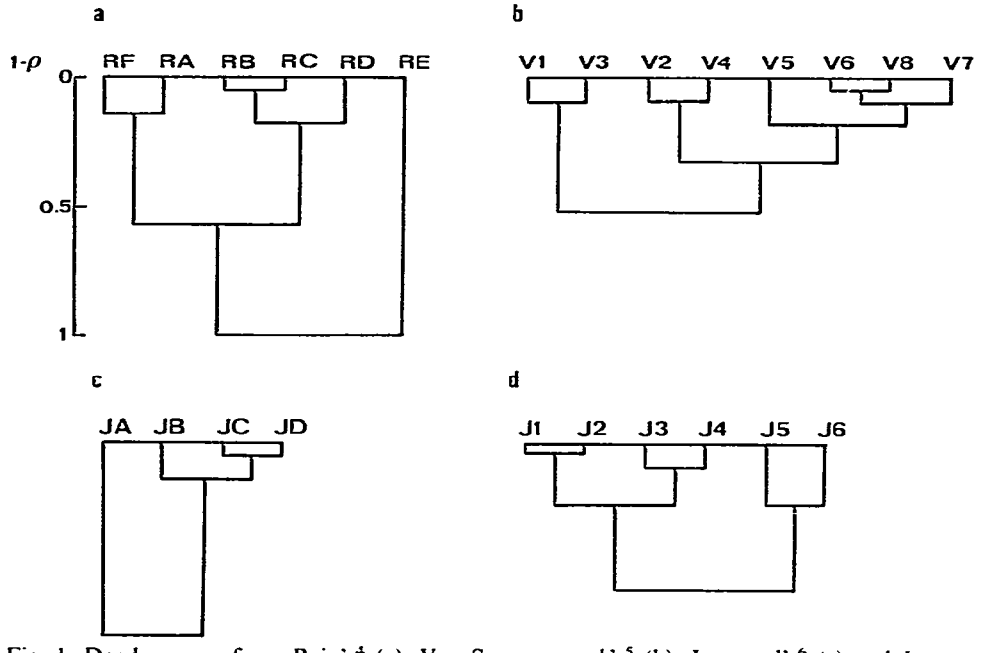

Fig. 1. Dendrograms from Reio's[4] (a), Van Sumere *et al.*'s[5] (b), Jangaard's[6] (c) and Jay *et al.*'s[7] (d) chromatographic systems (see also Table I for more information).

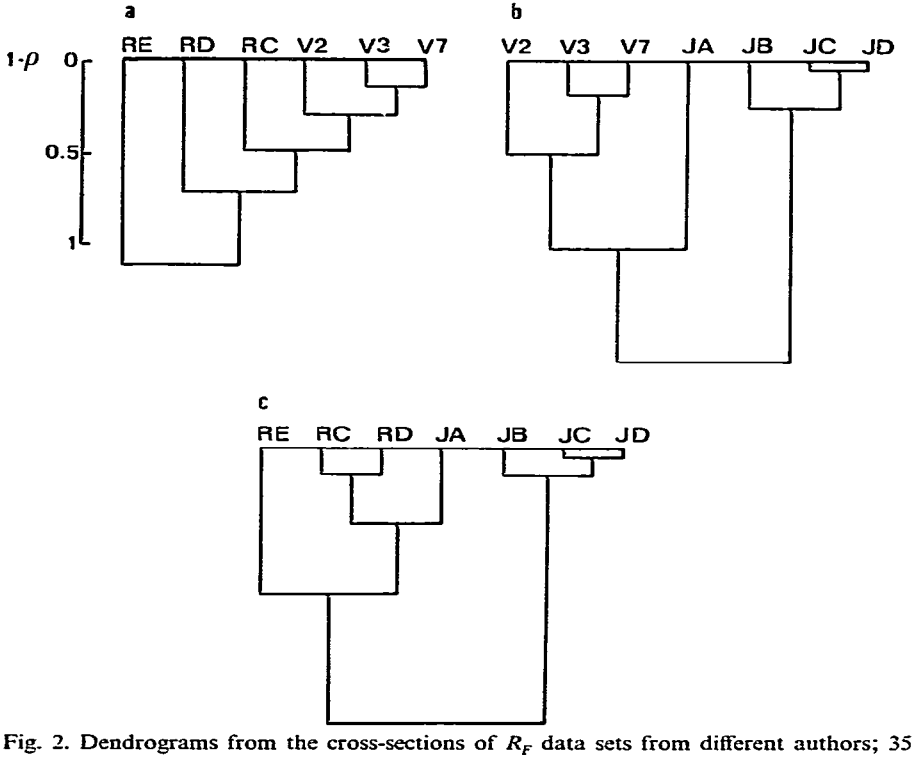

Fig. 2. Dendrograms from the cross-sections of R_F data sets from different authors; 35 R_F data were obtained from the cross-section Reio × Van Sumere *et al.* (a), 32 R_F data from Van Sumere *et al.* × Jangaard (b) and 51 R_F data from Reio × Jangaard (c).

TABLE I

PUBLISHED PC AND TLC SYSTEMS USED IN THIS STUDY

Reference	*Procedure*	*Systems*	*Mobile phase*	χ^2	$\bar{R}_F$
Reio[4]	228 phenolic compounds: PC	RF	Ethyl methyl ketone–acetone–formic acid–water	463	0.86
		RE	Water–ethyl methyl ketone–diethylamine	62	0.55
		RA	Methyl isobutyl ketone–formic acid–water	476	0.82
		RB	Chloroform–methanol–formic acid–water	151	0.68
		RC	Benzene–ethyl methyl ketone–formic acid–water	214	0.65
		RD	Benzene–formic acid–water	48	0.49
Van Sumere *et al.*[5]	81 phenolic compounds: TLC on silica gel (V1, V5); on steamed silica gel (V2, V6); on silica gel–cellulose (V3, V7); and on steamed silica gel–cellulose (V4, V8)	A = V1	Toluene–ethyl formate–formic acid	41.5	0.42
		B* = V2	Toluene–ethyl formate–formic acid	1.61	0.51
		C = V3	Toluene–ethyl formate–formic acid	29.8	0.47
		D* = V4	Toluene–ethyl formate–formic acid	5.90	0.49
		A = V5	Chloroform–acetic acid–water	10.9	0.43
		B* = V6	Chloroform–acetic acid–water	9.61	0.49
		C = V7	Chloroform–acetic acid–water	11.2	0.53
		D* = V8	Chloroform–acetic acid–water	22	0.55
Jangaard[6]	69 phenolic compounds, flavonoids and coumarins: cellulose TLC	JA [§]	Isopropanol–NH_3–water	18.8	0.38
		JB	20% potassium chloride	25.3	0.33
		JC [§]	2% formic acid	6.68	0.42
		JD	10% acetic acid	17.0	0.51
Jay *et al.*[7]	86 flavonoids: PC (J1, J2, J3, J4) TLC on polyamide (J5, J6)	J1	Acetic acid–water	27.9	0.63
		J2	Acetic acid–water–HCl	79.7	0.76
		J3	*n*-Butanol–acetic acid–water	149	0.83
		J4	*tert.*-Butanol–acetic acid–water	109	0.77
		J5	Benzene–methyl ethyl ketone–methanol	4.80	0.56
		J6	Benzene–petrol–methyl ethyl ketone–methanol	92.3	0.21

[§] In Jangaard's paper[6], the headings JA and JC were erroneously exchanged.

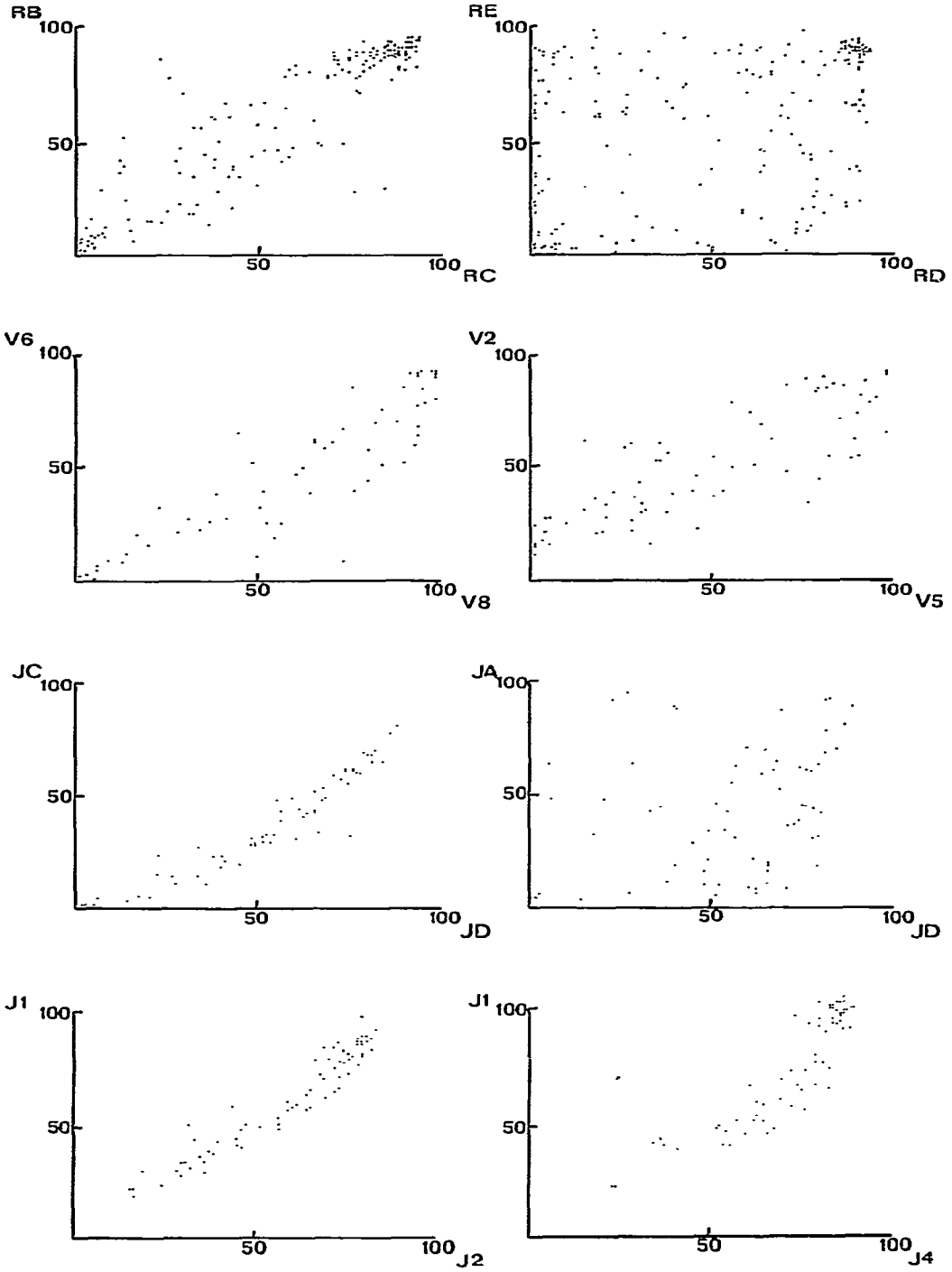

Fig. 3. Scatter diagrams obtained by plotting $R_F \times 100$ values from two different chromatographic systems. The left-hand scatter diagrams show highly correlated systems, and on the right-hand side less correlated systems from the same authors are shown.

systems into two groups, one group using an acidic solvent [RB, RC, RD] and the other using a basic solvent [RE], could have been induced by the fact that the subgroup of organic acids (77 R_F data) behaved completely differently from the subgroup of "neutral" phenolics (137 R_F data). We calculated the dendrograms for both subgroups, and only qualitative differences from the total group were noted: the difference between [RB, RC, RD] and [RE] was greater for the "acidic" subgroup than for the "neutral" subgroup. However, both subgroups yielded a dendrogram that was qualitatively comparable to that shown in Fig. 1 for the total group.

Jangaard's chromatographic systems could also be divided into two subgroups (see dendrogram B), *i.e.*, [JB, JC, JD] and [JA]. Surprisingly, three chemically completely different mobile phases fell into one subgroup. Two similar mobile phases (2% formic acid and 10% acetic acid) effected an uncorrelated elution behaviour. Jangaard suggested the use of the eluent pairs JA–JC, JB–JC, JD–JC and JD–JB in two-dimensional chromatography. This was completely in contradiction to the dendrogram that we calculated from his results.

On this basis we should rather suggest the combinations JA–JB, JA–JC and JA–JD for this purpose. Closer inspection of Jangaard's data revealed that when JA and JC were exchanged, Jangaard's results correlated with our findings. This meant that the mobile phase in the JA system was isopropanol–ammonia–water, whereas it was 2% formic acid in the JC system. Hence, in ref. 6, the column headings JA and JC were probably erroneously exchanged. Using a set of 66 phenolic compounds available in our laboratory, we tested this possibility. We chromatographed these compounds on MN300 + F_{254} cellulose TLC plates from Macherey, Nagel & Co. (Düren, G.F.R.) (the same type as used by Jangaard) with 2% formic acid and 10% acetic acid as mobile phases. The correlation between these two systems was clear (ϱ = 0.83). Even when an HPTLC stationary phase such as cellulose HPTLC + F_{254} (Merck, Darmstadt, G.F.R.), was used, these two eluents effected the same elution behaviour (ϱ = 0.86). Division of Jangaard's set of R_F values into "acidic" (40 R_F data) and "neutral" (39 R_F data) subgroups yielded qualitatively analogous dendrograms.

Van Sumere *et al.*'s systems could clearly be divided into several pairs, V1–V3, V2–V4, V6–V8 and V5–V7 (see Table I). These pairs were composed of two systems of steamed or non-steamed plates (see Fig. 3). This shows that the use of mixed layers had little or no effect on selectivity, whereas treatment of the plates with steam alters the selectivity to a certain extent. About the same variation in selectivity is obtained by changing the mobile phase from TEF to CAW: subgroups [V1, V2, V3, V4] and [V5, V6, V7, V8] were formed.

The dendrogram obtained with the acidic compounds subgroup (40 R_F data) was analoguous to that from the total group which is shown in Fig. 1. The neutral compounds subgroup (39 R_F data) behaved slightly different.

Cross-sections of chromatographic data sets from the different authors were also compared. We determined which phenolic compounds were found in two such collections, and these new, smaller data sets were used in the multiple correlation search. It was possible to do this for the systems Reio × Van Sumere *et al.*, Van Sumere *et al.* × Jangaard and Reio × Jangaard. Jay *et al.*'s data could not be used in this search, as they investigated compounds of a different nature (flavonoids), so that there was only a small overlap with the other authors' data.

Dendrograms of correlations between systems from the different publications are shown in Fig. 2. From each publication we selected three or four systems belonging to different selectivity groups. In one such group the choice was made on the basis of a favourable χ^2 value (best spreading of R_F values). The dendrogram from Van Sumere *et al.*'s data and Reio's data (Fig. 1) shows that generally there was little or no correlation between the two, except for Reio's system C, which correlated moderately well with the three systems of Van Sumere *et al.* This indicated an analoguous separation mechanism on RC and V2, V3, V7. From the RXJ dendrogram it was clear that JA and RC + RD had comparable selectivities, although the correlation was only moderate. It can be seen that all of these systems with a cellulose stationary phase could be divided into four groups with completely different selectivities, *i.e.*, [RE], [RC, RD], [JA] and [JB, JC, JD] (see Fig. 2, dendrogram c).

In accordance with the theories of Massart and De Clercq[8], it would be advantageous to use combinations of systems from these four groups when one is faced with the problem of product identification.

The VSX J comparison is also illustrated in a dendrogram in Fig. 2. The systems of Van Sumere *et al.* and of Jangaard were completely uncorrelated. Product identification in these systems is most likely to be successful when one uses a combination of systems chosen from the three groups [V2, V3, V7], [JA] and [JB, JC, JD].

As Jay *et al.*'s publication was concerned with a completely different set of phenolic compounds, namely the flavonoids, it is discussed separately. Using the dendrogram derived from Jay *et al.*'s system (Fig. 1), we could make subdivisions into three groups: [J1, J2], [J3, J4] and [J5, J6]. J1 and J2 had practically identical selectivities. The difference between the two eluents was their eluting power (J2 > J1; compare the $\bar{R}_F$ values in Table I). Systems J3 and J4 were completely comparable on the basis of selectivity and eluting power. There is also little difference between the two groups J1, J2 and J3, J4 (Fig. 3). Switching from PC to polyamide TLC changed the selectivity, but only to a limited extent.

The best chromatographic system for this set of flavonoids was undoubtedly J5 (lowest χ^2 and good $\bar{R}_F$ value). Maximum information on the nature of an unknown flavonoid should be obtained by chromatography on J1 and J5 when only two of the authors' systems were chosen. J3, J4 and J6 have χ^2 values too high for this set of flavonoids.

CONCLUSION

We detected only two sufficiently different selectivities in Reio's PC systems and in Jangaard's TLC separations. Jay *et al.*'s PC separations also split into two selectivity groups. They obtained two additional selectivities by using polyamide as a stationary phase. A study of Van Sumere *et al.*'s data set showed that replacing simple layers (silica) by mixed layers (cellulose/silica) did not change the selectivity. The water content of the layers had a great effect on the selectivity. In fact, one can select three systems from Van Sumere *et al.* with sufficiently different selectivities that have the same stationary phase. We found an almost complete lack of correlation between systems from the different authors used in our cross-correlation study.

The dendrograms in Figs. 1 and 2 can be used in the chromatography of phenolic compounds for the following purposes:

(1) to decide which combination of systems will yield the most information about the nature of an unknown phenolic compound (i.e., qualitative analysis); the use of systems with comparable selectivities can be avoided as they do not provide new information;

(2) to choose an ideal pair of mobile phases for two-dimensional chromatography;

(3) when newly developed stationary phases such as HPTLC cellulose plates are used; compilation of R_F data in these systems could be restricted to a set of selected mobile phases that yield sufficiently different selectivities.

The PC and TLC of phenolic compounds offer only two selectivities when one stationary phase is used. Changes of mobile phases are needed only in order to obtain a better spread of the compounds over the chromatogram (minimizing χ^2). Up to three different systems can be obtained with one stationary phase when the water content of the latter is changed (*e.g.*, Van Sumere *et al.*'s data). A change of stationary phase results in a complete change in selectivity (only one exception is noted, namely that substituting cellulose HPTLC for cellulose TLC only slightly affects the selectivity). This is what makes TLC advantageeous over HPLC in qualitative analysis. Indeed, with the former technique one can easily perform analyses on different systems at the same time. Only a few practical systems are known in the HPLC analysis of phenolic compounds. If the latter technique is to have the same capabilities as PC and TLC for the qualitative analysis of these compounds, other systems will have to be found.

ACKNOWLEDGEMENTS

We thank Mr. W. van Dongen for assistance and useful discussions, and Miss L. Van Roy for the TLC separations.

REFERENCES

1 L. Nagels, W. van Dongen, J. de Brucker and H. de Pooter, *J. Chromatogr.*, 187 (1980) 181–187.
2 D. L. Massart, *J. Chromatogr.*, 79 (1973) 157–163.
3 B. L. Karger, L. R. Snyder and Cs. Horváth, *An Introduction to Separation Science*, Wiley, New York, 1973.
4 L. Reio, *J. Chromatogr.*, 1 (1958) 338–373.
5 C. F. van Sumere, G. Wolf, H. Teuchy and J. Kint, *J. Chromatogr.*, 20 (1965) 48–60.
6 N. O. Jangaard, *J. Chromatogr.*, 50 (1970) 146–148.
7 M. Jay, J. F. Gonnet, E. Wollenweber and B. Voirin, *Phytochemistry*, 14 (1975) 1605–1612.
8 D. L. Massart and H. De Clercq, *Anal. Chem.*, 46 (1974) 1988–1992.