# Distribution of $R_f$ Values and Combination of Systems in Thin-Layer Chromatography

## H. De Clercq, E. Blockeel, E. Defrise-Gussenhoven,[1] and D. L. Massart

*Farmaceutisch Instituut, Vrije Universiteit Brussel, Paardenstraat 67, B-1640 Sint-Genesius-Rode, Belgium*

Recently, we have discussed the problem of making a rational selection of a restricted set from a large number of available chromatographic systems for the separation of a particular group of substances. A clustering technique, called numerical taxonomy, was proposed for this purpose and applied to TLC (*1*) and GLC (*2, 3*). A question which remains unanswered, is how many chromatographic systems must be chosen. For GLC, this question was answered by Dupuis and Dijkstra (*4*), who have shown that information theory allows one to calculate the amount of information obtained, if one knows the standard deviation of the retention index distribution (which was shown to be Gaussian for the GLC columns) and the correlation coefficients between the columns. The amount of information can then be compared with the amount of information needed.

Connors (*5*) has addressed the question of how many TLC and/or PC (paper chromatographic) systems should be used in seeking evidence of identity. By system is meant here, the combination of any particular stationary phase with a developing solvent. This author considers the distribution of the $R_f$ values of a set of substances over $10^m$ cells in a separation space defined by $m$ supposedly independent systems. For example, when $m = 2$, the $R_{fx}$–$R_{fy}$ diagram is divided into $10^2$ cells, each with dimensions $(0.1)^2$ (in $R_f$ units) and two substances are considered to be separated when they are found in different cells. If one counts the number of cells containing 0, 1, 2, 3, ... substances then, according to Connors, a Poisson distribution given by

$$p(k) = \frac{e^{-\lambda}\lambda^k}{k} \qquad (1)$$

is obtained. $p(k)$ = the probability that $k$ points will be found in a cell. $\lambda = T/N$, where $T$ = total number of points and $N$ = total number of cells.

The author gives experimental evidence for four such cases and concludes that the agreement between theory and experiment is rather good. The Poisson distribution is thought to occur in all cases, where the $m$ systems are independent. Although this is not said explicitly, this should also be the case for $m = 1$.

These observations and proposals have been submitted to a critical examination by us. Although a $\chi^2$ test cannot be carried out for the examples cited by Connors, one can conclude that the experimental distributions in the examples of this table agree well with the proposed Poisson distribution (the $\chi^2$-test cannot be carried out because the number of degrees of freedom is found equal to 0 in all cases: the number of degrees of freedom is obtained by subtracting 2 from the total number of classes, of which the theoretical frequency exceeds 8). However, when this proposition is investigated more thoroughly, it appears that this agreement occurs only in a very limited number of cases. A Poisson distribution is obtained when there is a random distribution of points, or, as Connors puts it, if the probability of occupancy of a cell by a point is independent of the location of the cell. This is rarely the case, as many practic-

ing chromatographers know, and there are two main reasons for this. The first is that in the set of substances, pairs or larger groups of very strongly related substances often occur. These are nearly always to be found in the same cell, so that there can no longer be a question of randomness in distribution. Such pairs occur in nearly all sets. In the set of 26 water-soluble food dyes, which served as an example in the article in which the selection procedure with numerical taxonomy was introduced (*1*), Orange GGN and Sunset Yellow occurred in the same cell in all 10 systems.

The second and more important reason is that in most systems not all $R_f$ values are equiprobable. To show this, the distributions were determined for all the $R_f$ values in the different TLC systems given in chapters by Maček (*6*) about synthetic drugs (1317 $R_f$ values), by Luckner (*7*) about drugs extracted from plants (692 $R_f$ values), by Sunshine (*8*) for toxicologically important drugs (1518 $R_f$ values), and by Moffat (*9*) for PC and TLC of basic drugs (800 $R_f$ values) (Figure 1).

To avoid confusion between the distribution of $R_f$ values given in Figure 1 and the Poisson distribution of the cells according to their occupancy, we shall in this article call the former R-distribution and the latter P-distribution.

The data used cannot be considered a biased set of data since use has been made of all the $R_f$ values of a complete chapter or, for Moffat's set, of a very large data set.

Neither can it be argued that inefficient TLC systems or systems which do not constitute a good choice for the sepa-

**Table I. Number of Cells Containing $k$ Points (Data from Ref. 10)**

| $k$ | System | | | | | | | Theory,[a] % |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 0 | 5 | 3 | 5 | 2 | 5 | 2 | | 0.1 |
| 1 | 1 | 1 | 1 | | 2 | 1 | | 0.7 |
| 2 | 1 | | | 1 | | | 2 | 2.3 |
| 3 | | 2 | 3 | 2 | 1 | 5 | 1 | 5.4 |
| 4 | 1 | 3 | | | | 3 | 4 | 9.3 |
| 5 | 1 | 2 | | 3 | 1 | 3 | 3 | 13.0 |
| 6 | 2 | 1 | 2 | 3 | 2 | 1 | 3 | 15.0 |
| 7 | | 1 | | 1 | 2 | | 3 | 14.9 |
| 8 | 2 | 1 | | 1 | | 1 | 1 | 12.9 |
| 9 | 2 | | 2 | 1 | | | | 10.0 |
| 10 | | 1 | | 2 | 1 | | | 6.9 |
| 11 | 1 | 1 | 1 | 1 | 1 | | | 4.4 |
| 12 | 1 | | | | | | | 2.5 |
| 13 | | 1 | 3 | 2 | 1 | | | 1.4 |
| 14 | | | | | 1 | | 1 | 0.7 |
| 15 | | 1 | | 1 | 1 | 1 | 1 | 0.3 |
| 16 | | | | | | | 1 | 0.1 |
| 17 | 1 | | 2 | 1 | 1 | 1 | | 0.1 |
| 18 | 1 | 2 | | | | 1 | | 0.0 |
| 19 | | 1 | | | | | | |
| 23 | 1 | | | | 1 | | | |
| 32 | | | | | | 1 | | |
| $\hat{\sigma}^2$ | 44.6 | 36.9 | 38.3 | 20.1 | 45.7 | 61.3 | 16.2 | |

[a] Poisson distribution for 139 points in 20 classes.
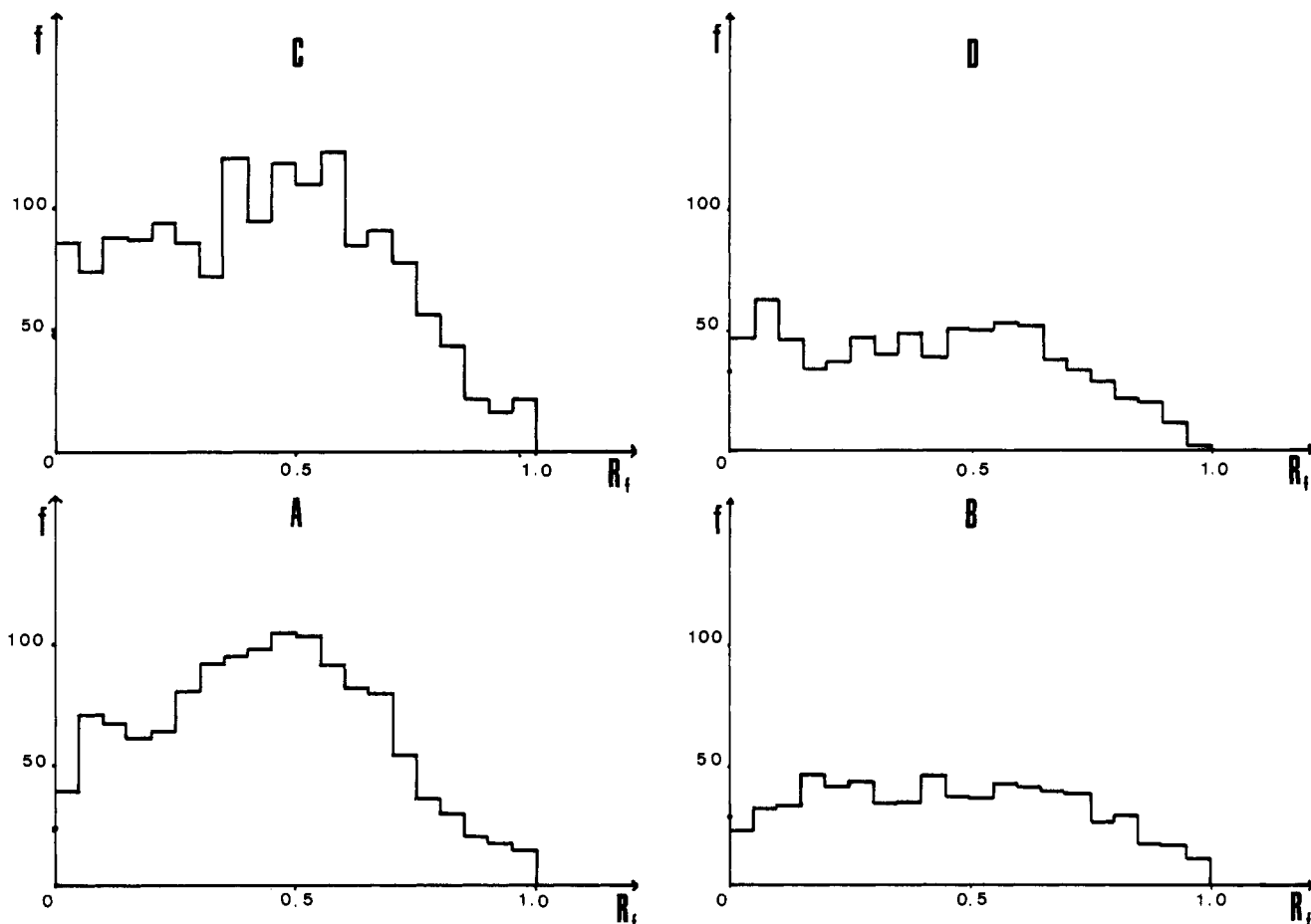
**Figure 1.** Distributions of TLC $R_f$ values for synthetic drugs (A) Ref. 6; drugs extracted from plants (B) Ref. 7; toxicologically important drugs (C) Ref. 8; basic drugs (D) Ref. 9

ration of the substances concerned were studied: the systems from which the $R_f$ values originate are generally accepted as among the best available. It can be seen that in each case, $R_f$ values of more than 0.75 are distinctly less probable than $R_f$ values around 0.5. The same is true to a lesser degree for $R_f$ values between 0.01 and 0.3. On the contrary, values of 0 are more frequent than many other values.

It should be realized that the R-distributions of Figure 1 represent mean situations and are not illustrations of individual systems. In fact, whereas in a given system generally the same type of distribution is found, different systems can have dissimilar R-distributions. These can go from more or less rectangular, to pseudo-normal (most values around 0.5), U-shaped (many low and high values), or skewed with most values in the first half of the plates or in the last half. One has to conclude that not only is there no equiprobability of $R_f$ values and therefore also no Poisson P-distribution, but also there is no other generally valid R- or P-distribution which would allow calculations about the number of systems that are sufficient for a specified degree of separation.

To demonstrate that the P-distribution is generally no Poisson distribution, we have investigated more particularly some of the few available systems where sufficient data are present, namely the 7 systems proposed for 139 therapeutically significant organic bases by Sunshine (10). In Table I, the data have been classified according to the number of cells containing $k$ points. For these large sets, a smaller dimension (0.05 $R_f$ units) is used than the one proposed by Connors (0.10). This has no influence, however, on the conclusions.

In each system, a Poisson distribution with $\lambda$ = 139/20 should be found (except for system 7 where $\lambda$ should be 130/20). This cannot be tested with a $\chi^2$ test because the number of occurrences is too small. From a visual comparison, with the theoretical frequencies in Table I, one observes however that the frequency of cells with 0 points or with an occupancy of 10 or more seems much higher than theory would permit. A test which can be carried out is the computation of the variance which for a Poisson distribution should be equal to the mean (6.95 or 6.50, resp.). It is quite clear from Table I that this is not the case.

Although we do not doubt that in some cases a Poisson P-distribution can be found, the same observations which have been made for the data sets cited, can in fact be made also for many of the 2-dimensional systems used as illustrations by Connors. A cursory investigation shows that 20 out of the 35 systems investigated by this author have frequency of $R_f$ values <0.2 or >0.8 which is much too low. This is not surprising since Connors states "that (it is desirable that) $R_f$ values fall in the range 0.2–0.8 and not near 0.0 or 1". However, in that case, the proposed calculations should not use $N$ = 100 cells but rather $N$ = 36.

Since all systems, where $R_f$ values of <0.2 or >0.8 are frequent, seem to have been excluded and since Connors' calculations are erroneous for the other systems (the value of $N$ used is incorrect), it is obvious that the proposed Poisson distribution can only occur by chance in a few cases. Connors' conclusions about the number of systems to be used as a function of the number of substances in the set are therefore not generally valid. They do provide a tool to calculate the number of systems required for identification under certain assumptions, namely that the Poisson distri-

bution describes all systems employed. As explained in the present article, this is rarely the case. Since all the phenomena which are discussed here lead to a lesser degree of separation, his conclusions may be considered as giving the lower limit of systems necessary to carry out the identification, but not the number that is sufficient in practice. This limit is however of entirely academic interest.

It also follows from these considerations, that unlike GLC it is not possible to calculate a priori the number of TLC and/or PC systems necessary to obtain the separation of a given set of substances. Instead one has to determine what are the best combinations of 2, 3, ... systems and to evaluate the results obtained. An objective way to do this, is to select the best combinations, using numerical taxonomy, and to evaluate the results using information theory (1) or the discriminating power concept (9).

Finally, it should be added that our criticisms concern only the possibility of applying Poisson distributions to describe the results of TLC or PC separation procedures in practice. Connors has made the very useful point that uncorrelated systems are better than correlated systems for identification purposes. The $R_{fx}$, $R_{fy}$ plots introduced by him, constitute a practical visual guide to identifying the former. Numerical taxonomy and other classification procedures allow this achievement in a more formal way.

## LITERATURE CITED

(1) D. L. Massart and H. De Clercq, Anal. Chem., **46**, 1988 (1974).
(2) D. L. Massart, P. Lenders, and M. Lauwereys, J. Chromatogr. Sci., **12**, 617 (1974).
(3) A. Eskes, P. F. Dupuis, A. Dijkstra, H. De Clercq, and D. L. Massart, Anal. Chem., **47**, 2168 (1975).
(4) P. F. Dupuis and A. Dijkstra, Anal. Chem., **47**, 379 (1975).
(5) K. A. Connors, Anal. Chem., **46**, 53 (1974).
(6) K. Macek, in "Pharmaceutical Applications of Thin-Layer and Paper Chromatography", K. Macek, Ed., Elsevier Publishing Co., Amsterdam, New York (1972).
(7) M. Luckner, Ref. 6.
(8) I. Sunshine, "Handbook of Analytical Toxicology", Chemical Rubber Co., Cleveland, Ohio, 1969.
(9) A. C. Moffat and K. W. Smalldon, J. Chromatogr., **90**, 9 (1974).
(10) I. Sunshine, W. Fike, and H. Landesman, J. Forensic Sci., **11**, 428 (1966).

# Effects of Metal Loading to Increase Absorptivity in Laser Pyrolysis–Gas Chromatography

## Nicholas E. Vanderborgh,[1] Ray L. Hanson, and Candace Brower

The University of New Mexico, Albuquerque, N.M. 87131

Pulsed lasers have been well characterized as intense and rapid thermal sources for pyrolysis–gas chromatography (1). Impinging a focused laser beam onto a sample contained in a quartz tube attached on-line to a gas chromatograph permits a rapid and convenient method for sampling high-molecular-weight solids. This method shows good promise for the characterization of a variety of different sample types (2).

Intense laser pulses couple into solid samples to produce significant thermal degradation. However, milder conditions of pyrolysis, using energies in the one joule/pulse range are not always absorbed efficiently and, consequently, little sample degradation results. Several methods have been described to circumvent this problem. Work has been described in which carbon has been introduced into the sample (3, 4). In this way, absorbing centers are introduced and the thermal energy is then transferred into the sample. Other workers have described converting samples into absorbing derivatives (5). Organic acids which did not pyrolyze were converted to copper salts which could be efficiently degraded. Samples have also been deposited on solid absorbing materials (6). Laser energy deposited into the solid is then transferred into the sample. Each of these techniques has limitations.

Carbon loading on organic compounds clearly changes the product distribution pattern (4, 7). Loading carbon into a sample changes the energetics of the plume quenching process to lead to a more carbon-rich distribution. Much information about the sample is lost using this procedure. Formation of derivatives is a complex chemical process involving considerable sampling time and some chemical systems (polymers, for example) are not easily

coupled to absorbing centers. Impinging the laser energy onto a solid absorbing center spreads and lessens the pulse intensity. This readily follows from heat capacity considerations. Thus, work in these laboratories has centered upon finding loading materials which have the experimental simplicity of carbon but do not change product distributions. This led to a series of experiments where powdered metals were intimately mixed into organic compounds for laser pyrolysis–gas chromatography (LPGC). We report here information that suggests that powdered nickel is especially convenient for this task.

## EXPERIMENTAL

The laser used for this study was a TRG 104-A pulsed ruby system. The laser was powered by a Xenon flash lamp with 840 joules of input energy. Measured output energy was approximately 3 joule/pulse. The laser energy (694.3 nm) was focused using a 21-cm quartz lens onto the sample contained in a 6-mm quartz tube attached directly to the input port of the GC used for product separation and analysis.

The gas chromatograph (Varian 940) was equipped with a 15-cm precolumn (Poropak S, 100–120 mesh) in series with a 1.8-m analytical column (Poropak Q, 100–120 mesh). Work was done isothermally at 150 °C (using helium as the carrier at 25 ml/minute). The results were obtained with the FID coupled to a CSI-208 digital integrator.

Two compounds were selected for pyrolysis: anthracene and cholesterol. Nickel powder, 100–200 mesh, was ground with the sample to intimately mix the sample with the absorbing center. Various concentrations of nickel were prepared by transferring samples into a stainless steel die and pressing at 20,000 psi. Samples were then removed from the press and transferred to the pyrolysis chamber. Helium flow was established past the sample and the laser fired.

## RESULTS AND DISCUSSION

The results of laser-induced pyrolysis with two different hydrocarbon compounds are described here. One of these,

---

[1] Present address, Los Alamos Scientific Laboratory, Los Alamos, N.M. 87545.