

This article was downloaded by:

On: 24 January 2011

Access details: *Access Details: Free Access*

Publisher *Taylor & Francis*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Liquid Chromatography & Related Technologies

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713597273>

A Study of an Empirical Equation for the Evaluation of Separation Efficiency in Chromatography

Darko Kantoci^a

^a Laboratory of Chemical Endocrinology, Loma Linda University School of Medicine, Loma Linda, CA

To cite this Article Kantoci, Darko(1996) 'A Study of an Empirical Equation for the Evaluation of Separation Efficiency in Chromatography', *Journal of Liquid Chromatography & Related Technologies*, 19: 6, 853 — 866

To link to this Article: DOI: 10.1080/10826079608001917

URL: <http://dx.doi.org/10.1080/10826079608001917>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

A STUDY OF AN EMPIRICAL EQUATION FOR THE EVALUATION OF SEPARATION EFFICIENCY IN CHROMATOGRAPHY

Darko Kantoci

Loma Linda University
School of Medicine
Laboratory of Chemical Endocrinology,
Loma Linda, CA 92350

ABSTRACT

An empirical equation has been developed to evaluate chromatographic separation efficiency. This equation can be expressed as the product of a "separation" term, a "capacity" term and an "alignment between peaks" term. It can be shown that this equation is insensitive to other column or separation parameters. Its sensitivity depends only upon the resolution between peaks. This in turn allows the linking of the equation with any column or separation parameter during the optimization process as a response function. The utility of this equation has been extensively tested with the aid of computer-simulated overlapping peaks. This method is applicable to any separation technique that is based upon separation of compounds due to differences in partitioning between the stationary and mobile phases (column chromatography, HPLC, counter-current distribution, etc.). This equation can also be used to evaluate the quantity of each separated or overlapped compound giving a true picture of separation efficiency.

INTRODUCTION

Due to today's increasingly stringent requirements for pure pharmaceutical, agricultural and industrial chemicals, an increased burden has been placed upon the analytical chemist. As such, one must not only choose the best separation technique (column chromatography, HPLC, counter-current chromatography, molecular exclusion, etc.), but also optimize all parameters in order to achieve the maximum resolution of the main constituent of a mixture. In order to obtain a suitable chromatographic solvent, the novice must rely on solvent systems reported in literature for the separation of similar types of compounds and use trial-and-error techniques to arrive at a suitable solvent composition. An experienced analytical chemist may select better solvent combinations, but still may not have optimized conditions so to achieve the best resolution. Regardless of whether a novice or an expert separation chemist is confronted with a separation problem, the question inevitably arises as to what criteria one uses in assessing separation efficiency.

The literature contains a number of more or less complicated equations¹⁻⁵ attempting to describe column efficiency in terms of resolution between peaks and involving one or more separation parameters such as N (theoretic plate number), R_s (column resolution), α (column selectivity) and k' (capacity factor) among others. Said⁶ first developed an equation for separation of a pair of compounds (Said's recovery index) for Gaussian peaks that is not restricted to elution chromatography systems. Unfortunately this method is useful only for binary systems that afford a Gaussian distribution of peaks. This is not adequate for most real world separations. What is needed is a mathematical expression by which to evaluate the efficiency of the separation of any number of components with any peak shape, using any separation process. This expression should require only experimental data and empirical coefficients. The pioneering work in this area had been done by Rony,⁷⁻⁹ Giddings¹⁰ and De Clerk.¹¹

Numerous functions describing the chromatographic response function (CRF) have been derived and some tested by Wegscheider.⁵ Most of these methods are based on the assumption that, if the composition of a mixture is known (which is often not the case), the only problem remaining is to optimize the parameters of the equation describing the resolution of the mixture. A number of different approaches to finding an adequate solution to, and refinement of, these equations have also been described.¹² Still, all of these equations suffer from an interdependence of the separation parameters themselves, and thus prohibit generalization of a separation response function.

A problem often encountered by the separation chemist during the analysis of a complex mixture is the determination of the number and quantity of each compound in the sample. Still pronounced complete peak overlap is common during the routine HPLC separation and spectral analysis of extracts obtained from natural products. With the advent of today's UV/visible liquid chromatography diode-array detectors, and lineshape simulation programs, reliable estimates of the number of compounds contained within overlapping peaks, their retention times and relative abundance, can be at least estimated. Such information can subsequently be introduced into the equation described below, which assesses the separation efficiency for the system used as a numerical value between 0 and 1. When linked with other separation parameters (solvent composition, column dimensions, flow rate, etc.), it can provide the chromatographer with a method of optimizing separation leading to improved resolution of the overlapping peaks.

EXPERIMENTAL

Programs were written for the calculation of separation efficiency using GWBASIC v.3.22., (Microsoft corp.). Overlap calculations and testing of the equations was done on an IBM PS/2 and Macintosh Quadra 650 computer. Polynomial curve fitting was done using Cricket Graph v.1.3 and all drawings were generated with either Microsoft Excel, v5.0 (Microsoft corp.) or SuperPaint v.2.0 (Silicon Beach Software) software on a Macintosh computer.

THEORY

The three factors should be considered for the evaluation of the separation efficiency: peak overlap, quantity of the material under the peak and peak apex distribution.

The first factor, peak overlap or separation factor, evaluates how well peaks are separated from each other, thus evaluates their mutual overlap (E1).

The second factor evaluates the quantity of pure and overlapped compound(s) under the peak or part of it (E2). This factor measures mass distribution during separation. For instance we might have a severe overlap, but the mass of the compound(s) under the overlap is small compared to the total quantity of the separated material. Therefore the contribution to the separation efficiency will be small, and vice versa.

The ideal separation has all peaks evenly distributed throughout the chromatogram. The third factor (A) evaluates the peak distribution compared to ideal, and is designated as a alignment between peaks. The impact of this factor to the product of E1 and E2 is significant in cases where some peaks are grouped in a chromatogram and others are far apart. This indicates a poor solvent or column selection, wrong gradient profile, etc.

The separation efficiency (E) can be then expressed as the product of a "separation" term (E1), a "capacity" term (E2) and an "alignment between peaks" term (A), *i.e.*:

$$E = E1 * E2 * A * 100 \quad (1)$$

E1 is a "separation" factor expressed as the number of fractions containing only one compound ($T_{(1)}$) in total number of fractions ($T_{(n)}$) containing "n" compounds. C is the highest number of compounds present in any fraction and b is an empirical coefficient which will be explained later. The separation is then defined as:

$$E1 = \frac{\left(\sum T_{(1)}\right)^b}{\sum_{n=1}^c \left[n * \left(\sum T_{(n)}\right)^b\right]} \quad (2)$$

E2 is the "capacity" factor which evaluates the quantity of material in each fraction ($W_{(n)}$) containing the same number of compounds (n). W_t represents total quantity of material present. The E2 is actually a percentage of a compound in a mixture divided by the number of compounds (n) present in a fraction. a is an empirical coefficient which will be explained later, therefore E2 is defined as:

$$E2 = \sum_{n=1}^c \frac{W_{(n)}}{n^a * W_t} \quad (3)$$

The quantity may be expressed as the area under a curve or the mass of the fractions as long as it is consistent through the calculations.

Combining E1 and E2 and multiplication by 100 (for conversion to percentage), then rearranging defines E' which represents the separation efficiency based on the number of compounds under a peak with its relative amount. The exponents "a" and "b" represents empirical coefficients. Their values will be determined later.

$$E' = \frac{100 * \left(\sum T_{(i)} \right)^b * \sum_{n=1}^C \frac{W_{(n)}}{n^a * W_t}}{\sum_{n=1}^C \left[n * \left(\sum T_{(n)} \right)^b \right]} \quad (4)$$

The alignment factor¹³ (A) which keeps track of the separation between peaks and tries to separate them as far as possible, may be expressed as:

$$A = \frac{\prod_{i=1}^{P-1} \left[Rt_{(i+1)} - Rt_{(i)} \right]}{\left[\frac{Rt_{(max)} - Rt_{(min)}}{P - 1} \right]^{(P-1)}} \quad (5)$$

Where $Rt_{(i)}$ is the retention time for an individual peak or fraction at its apex. $Rt_{(max)}$ is the peak with maximal retention time or fraction containing last compound, $Rt_{(min)}$ is the peak with shortest retention time or fraction number containing the first compound, and P is number of peaks.

Thus the final equation for E, with alignment incorporated becomes $E = E' * A$, i.e.

$$E = \frac{100 * \left(\sum T_{(1)} \right)^b * \left[\sum_{n=1}^C \frac{W_{(n)}}{n^a * W_t} \right] * \left\{ \prod_{i=1}^{P-1} [Rt_{(i+1)} - Rt_{(i)}] \right\}}{\left\{ \sum_{n=1}^C \left[n * \left(\sum T_{(n)} \right)^b \right] \right\} * \left[\frac{Rt_{(max)} - Rt_{(min)}}{P - 1} \right]^{(P-1)}} \quad (6)$$

The only necessary data required for input into equations (4) or (6) are: the number of peaks (P) and their positions [Rt_(i)], the total number of compounds (C) contained under each peak (T_(n)), and the weight of each fraction W_(n) (or the area under the peaks).

As will be demonstrated, the function is nondimensional, and as such is insensitive to the shape of the peak, type of separation, column size, flow, and quantity of material separated. Subsequently, E takes into account the peak width and could be affected by k' and the selectivity, α. Equation (6) does not require approximations that are necessary for Purnell's¹⁴ or Knox's¹⁵ methods. Nor does it suffer from column parameter factors or other methods of separation. Thus, equation (6) leads to a general solution of the optimization problem. The functions E, E', E1, E2 are fitted in the range between 0 and 1 (0 represents no separation and 1 base line resolution between two or more peaks) near or on baseline.

RESULTS AND DISCUSSION

To test equations, hypothetical test mixtures were generated by computer simulation using a BASIC program that contained all the necessary peaks and separation between peak parameters.

Example of Estimation of Exponent "b" for E1:

A test matrix was generated which represents the number of compounds (1 and n) in fractions (1-12) for 13 separations. This matrix represents 13 separations as follows: the separation 1 represents 12 fractions of peak each

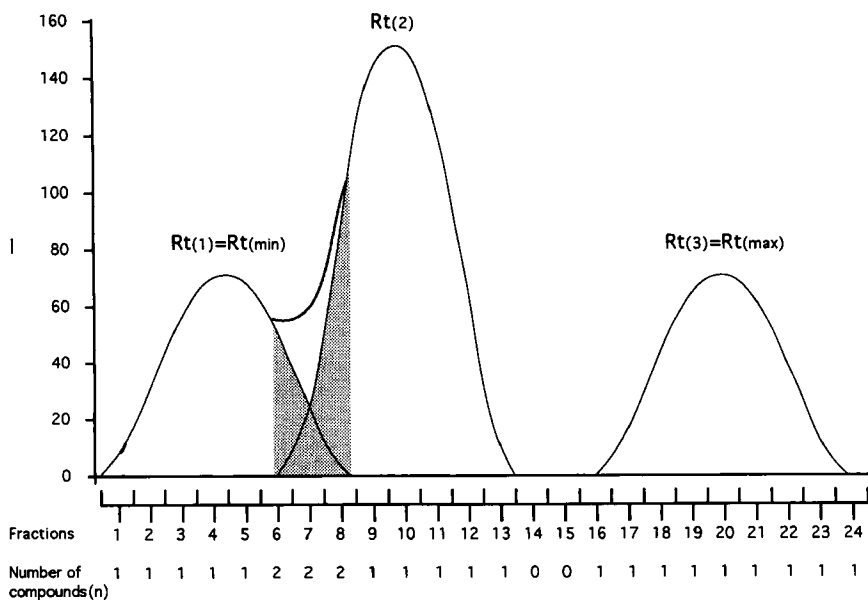


Figure 1. Partial overlap of two peaks

Table 1

Validity of Equation (1)

Separation #	Fractions											
	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	n
3	1	1	1	1	1	1	1	1	1	1	n	n
4	1	1	1	1	1	1	1	1	1	n	n	n
5	1	1	1	1	1	1	1	1	n	n	n	n
6	1	1	1	1	1	1	1	n	n	n	n	n
7	1	1	1	1	1	1	n	n	n	n	n	n
8	1	1	1	1	1	n	n	n	n	n	n	n
9	1	1	1	1	n	n	n	n	n	n	n	n
10	1	1	1	n	n	n	n	n	n	n	n	n
11	1	1	n	n	n	n	n	n	n	n	n	n
12	1	n	n	n	n	n	n	n	n	n	n	n
13	n	n	n	n	n	n	n	n	n	n	n	n

contains one compound; in separation 2, the fractions 1 to 11 contains one compound and fraction 12 contain n compounds, in separation 3 the fractions 11 and 12 contain n compounds; and so on. The last separation (13) is a single peak that contains n compounds (total overlap).

To prove validity of equation (1), substituting the number of compounds as 1 (separation 1) [$n=1$, $T_{(1)}=12$, $T_{(n)}=T_{(1)}=12$] in equation (1) for any value of b, $E1=1$; if the number of compounds under the peak is ∞ [$n=\infty$, $T_{(1)}=0$, $T_{(n)}=\infty$] for any value of b, $E1=0$ or in our example for separation 13, if $n=2$ then $T_{(1)}=0$, $T_{(2)}=13$ for any value of b, $E1=0$ (Table 1).

The equation E1 actually counts fractions (time increments) with only one compound and compare them with the number of fractions with two or more compounds. This might be illustrated as in Figure 1.

The number of fractions with 1 compound ($n=1$) is 19 [$T_{(1)}=19$]

The number of fractions with 2 compounds ($n=2$) is 3 [$T_{(2)}=3$]

$C=2$ (maximal number of compounds in single fraction)

$$\text{In this case } E1 = \frac{19^b}{1 \cdot 19^b + 2 \cdot 3^b}$$

if $b=1.5$ then $E1=0.89$ (89%) (disregarding E2 and A)

in more complex example where three compounds overlap (Figure 2) we have:

The number of fractions with 1 compound ($n=1$) is 13 [$T_{(1)}=13$]

The number of fractions with 2 compounds ($n=2$) is 6 [$T_{(2)}=6$]

The number of fractions with 3 compounds ($n=3$) is 3 [$T_{(3)}=3$]

$C=3$ (maximal number of compounds in single fraction)

$$\text{in this case } E1 = \frac{13^b}{1 \cdot 13^b + 2 \cdot 6^b + 3 \cdot 3^b}$$

if $b=1.5$ then $E1=0.51$ (51%) (disregarding E2 and A)

In order to determine the best value of exponent "b", the exponent "b" was varied over a range between 0.5 to 2.0, and the number of compounds (n) was taken as 2 to simplify presentation. The calculation can be expanded to any number of compounds.

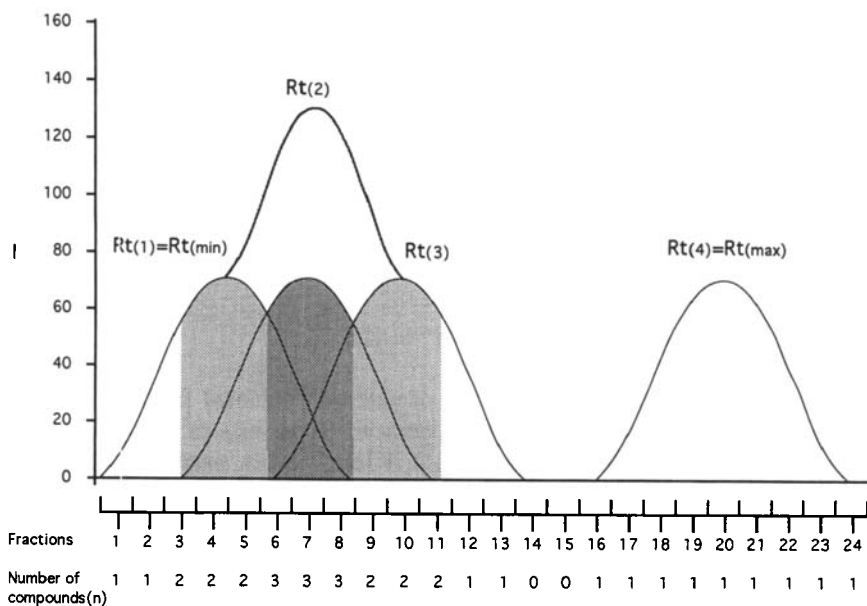


Figure 2. Three peak overlap.

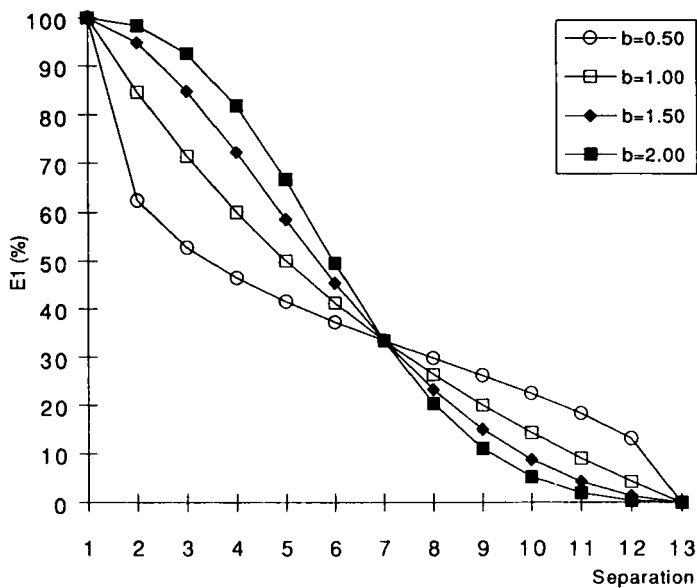


Figure 3. Plot of data for E1 with differing exponent "b"; n=2

By examination of the Figure 3, "b" lies between 1 and 2. For overlap of 11/12, E1=84.6% (b=1), resembling big drop of 15.4%. With the same value of b, for the overlap of 11/12, the E1=4.35%. For the value of b=2 and overlap of 11/12, the E1=98.4% that is close to 1, and for overlap of 11/12, the E1=0.41%, close to 0. The exponent "b" of 1.5 was chosen as algebraic mean to describe presented set of data for separation efficiency. This choice of "b" exhibits a slow drop in the beginning, a faster drop in the middle and again a slower drop at the end, thus resembling sigmoidal shape.

Example of Estimation of Exponent "a" in E2:

To test equation E2, an initial set of data was obtained from a randomly chosen real peak from a real HPLC separation. From this data, a polynomial equation was determined by non weighted least square curve fitting which described the peak. The factor x^5 is so small and does not have significant impact on fit.

The polynomial equation for the peak was determined to be:

$$y = 2.6279 - 3.3561 x + 1.2608 x^2 - 6.3247 \cdot 10^{-2} x^3 + 8.5469 \cdot 10^{-4} x^4$$

Applying this polynomial equation, new hypothetical sets of data could be generated which simulated conditions ranging from total overlap of two peaks to complete resolution of the peaks. This data could then be used in subsequent calculations of various E2 values for each set of conditions.

The details of this procedure are as follows:

The peak was divided into 12 sections to simplify calculations and the area under each section of the peak determined by integration of polynomial:

$$P = \int_{x_1}^{x_2} (2.6279 - 3.3561x + 1.2608x^2 - 6.3247 \cdot 10^{-2} x^3 + 8.5469 \cdot 10^{-4} x^4) dx$$

For estimation of the exponent "a" in E2, conditions involving complete overlap to complete resolution of two peaks were performed in stages. First, the peak was overlapped with one identical sized peak and E2 calculated. Then the second peak was shifted one place to the right and E2 again calculated. This procedure was repeated until both peaks were fully resolved (13 steps).

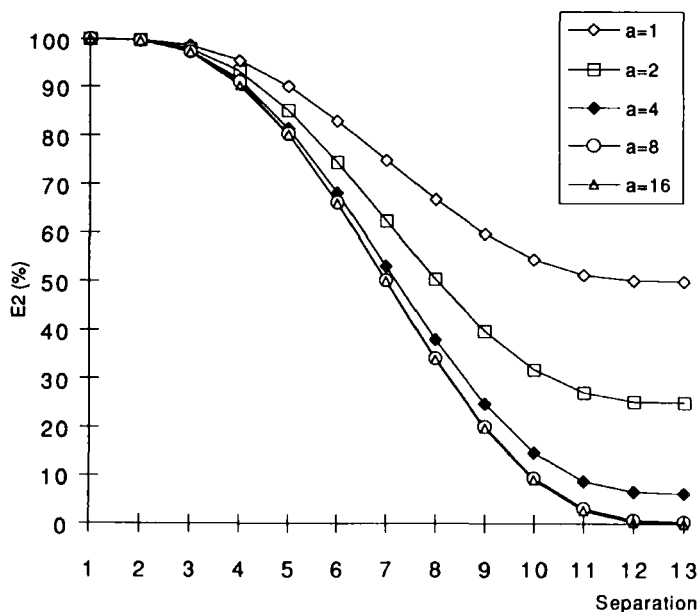


Figure 4. Plot of data for E2 with differing exponents "a".

The calculation of E2 was done for different exponents of "a" ($n=2$) and results are shown in Figure 4.

For equation (3) the possible value for exponent "a" could be 8 when used by itself, because for total overlapping of peaks if $a=8$, then E2 is close to zero. We can see that $E2 \rightarrow 0$ (total overlap), when $a \rightarrow \infty$. With the value of 16 for "a" the E2 is 10^{-5} .

The function was further tested using computer generated peaks by assessing separation between two peaks with greatly differing areas (ratios varying between 1:1 to 1:32). Assuming that both peaks have the same width, the average drop is 1.87% for 50% overlap to 0.07% for 9% overlap in the range of 1:1 to 1:32 of the relative area ratio. From the results obtained apparently response is not affected with size of peaks, but only with their mutual overlapping.

The Testing of Equation E' (4):

Equation (4) was calculated using $b=1.5$ with differing exponents of "a".

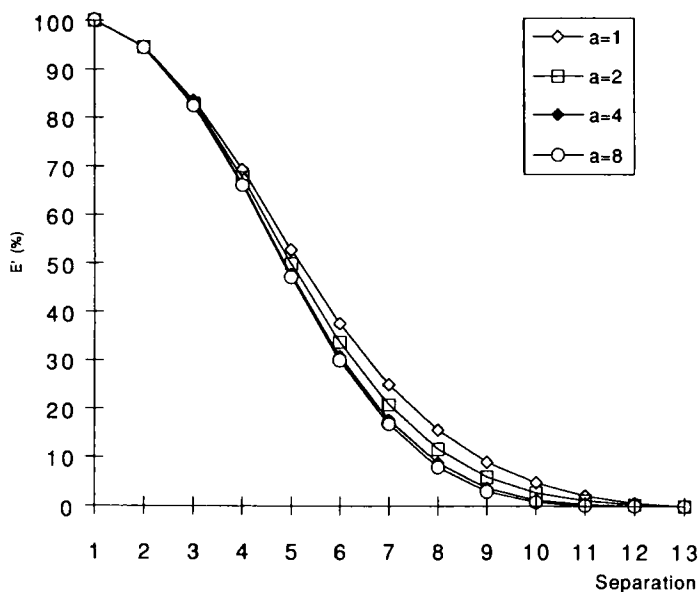


Figure 5. Plot of data for E' , $b=1.5$ with differing exponents of "a".

Table 2

Description of Symbols

E, E_1, E_2	separation of efficiency
A	alignment
$T_{(1)}$	number of fractions which contain only one compound (regardless of R_f)
n	number of compounds in each fraction
$T_{(n)}$	number of fractions (test tubes or ΔR_t etc.) with "n" compounds
C	total number of compounds
W_t	total weight or area under peaks of all fractions after separation
$W_{(n)}$	weight of each fraction or group of fractions (or area) with same n
a, b	empirical coefficients
$R_{t(i)}$	retention time for individual peak
$R_{t(max)}$	peak with maximal retention time
$R_{t(min)}$	peak with minimal retention time
P	number of peaks

From Figure 5, it is evident that equation (4) is not very sensitive to the value of exponent "a".

An important characteristic of E2 and consequently E' and E is that it is not necessary to know the ratio of compounds in overlapping parts of a peak, but only the total quantity of this part of a peak, and of course, the number of compounds contained in the overlapped peak. If we do not know the number of compounds under overlaying peaks, an estimation will serve well. Later during optimization other compounds may emerge from the overlapped peaks. Misjudged numbers of compounds under a peak causes only a small error.

CONCLUSION

Equation (4) and, consequently, equation (6), is independent of conventional chromatography column parameters (column width or length, stationary phase composition, mobile phase composition, elution time, flow, etc.). This independence allows one to link this equation to the any optimization method (e.g. Simplex¹⁶) as a response function in which one or more parameters are optimized at the time (length of column, quantity of sample, composition of mobile phase, etc.). An account can be taken of all peaks present or only peaks of interest (for example, the best resolution of valuable material from impurities) during optimization process. The results can, however, be generalized to a case involving "n" compounds. The only prerequisite parameters for the calculation of separation efficiency necessary to enter in equation are $R_{t(i)}$, $T_{(n)}$, C, P, and $W_{(n)}$. The values for factors a and b from above analysis are $a=2$, $b=1.5$.

ACKNOWLEDGMENT

The author wish to express appreciation to Dr. Delbert Herald, Arizona State University and Dr. Kane Denike for assisting in the preparation of the manuscript and comments.

REFERENCES

1. P. J. Schoenmakers, J. K. Strasters, A. Bartha, *J.Chromatogr.*, **458**, 355-370 (1988).
2. G. Eppert, *Z. Chem.*, **26(9)**, 325-329 (1986).

3. D. W. Osten, B. R. Kowalski, *Anal. Chem.*, **56**, 991-995 (1984).
4. H. Vink, *J. Chromatogr.*, **69**, 237-242 (1972).
5. W. Wegscheider, E. P. Lankmayr, M. Otto, *Anal. Chim. Acta*, **150**, 87-103 (1983).
6. A. S. Said, *J. Gas. Chrom.*, **2**, 60 (1964).
7. P. R. Rony, *Separation Science*, **3(3)**, 239-248 (1968).
8. P. R. Rony, *Separation Science*, **3(4)**, 357-374 (1968).
9. P. R. Rony, *Separation Science*, **3(5)**, 425-453 (1968).
10. J. C. Giddings, *Separation Science and Technology*, **13(1)**, 3-24 (1978).
11. K. De Clerk, T. S. Buys, *Separation Science*, **7(4)**, 371-387 (1972).
12. R. P. W. Scott, (Ed.), **Gas Chromatography**, Butterworth, London, 1960, p. 429.
13. D. Kantoci, G. R. Pettit, and Z. Cichacz, *J. Liquid Chrom.*, **14(6)**, 1149-1160 (1991).
14. J. H. Purnell, *J. Chem. Soc.*, 1268 (1960).
15. J. H. Knox, *J. Chem. Soc.*, 433 (1961).
16. C. L. Shavers, M. L. Parsons, S. N. J. Deming, *J. Chem. Educ.*, **56**, 307 (1979).

Received July 12, 1995

Accepted September 27, 1995

Manuscript 3927