CHROMSYMP. 1483

COMPUTER-ASSISTED RETENTION PREDICTION SYSTEM FOR OLIGONUCLEOTIDES IN GRADIENT ANION-EXCHANGE CHROMATOGRAPHY

YOSHINOBU BABA* and MIYUKI FUKUDA

Department of Chemistry, Faculty of Education, Oita University, Dannoharu 700, Oita 870-11 (Japan)

and

NORIMASA YOZA

Department of Chemistry, Faculty of Science, Kyushu University, Hakozaki, Higashiku, Fukuoka 812 (Japan)

SUMMARY

A technique is presented for the retention prediction of single-stranded homo-oligonucleotides under gradient elution conditions in anion-exchange chromatography. The prediction system is based on the theory proposed by Jandera and Churaček. In the present system, the theory was modified by the extrapolation method, utilizing the linear relationship between the log of the capacity factor and the number of nucleotides. The modified theory allowed the prediction of retention times for high-molecular-weight oligonucleotides, which could not be calculated from the original theory. Oligoadenylate was used as a standard oligonucleotide to demonstrate the accuracy of the prediction system. By use of this system, the retention times of oligoadenylates up to 25–35 bases in chain length were predicted within 8% errors under both binary-linear and binary-convex gradient shapes.

INTRODUCTION

In the different steps of cloning and sequencing of nucleic acids, high-performance liquid chromatography (HPLC) is particularly useful for the separation of fragments of DNA and RNA[1–18]. To find suitable separation conditions of these fragments by trial-and-error experiments is a time-consuming and troublesome task. Although computer-assisted retention prediction has been recognized as a powerful tool for optimizing elution conditions[19–31] in isocratic and gradient HPLC, no successful attempts have been reported on a prediction system for fragments of DNA and RNA.

We have developed a computer-assisted retention prediction system for inorganic polyphosphates in isocratic[30] and gradient[29,31] anion-exchange chromatography. The prediction system for inorganic polyphosphates has been successfully applied to the optimization of isocratic[27] and gradient[28] elution conditions and has provided a substantial saving in the time required for optimization over the conventional (trial-and-error) method.

Oligonucleotides as well as inorganic polyphosphates are anionic species, containing negatively charged phosphate groups. In anion-exchange chromatography, the behaviour of oligonucleotides is expected to be similar to that of inorganic polyphosphates, because the separation is primarily based on the ionic interactions of positively charged groups on the anion exchanger and negatively charged phosphate groups on the solutes.

In the present paper, the theory and procedure of computer-assisted retention prediction under gradient elution conditions are described for the oligomeric series of oligonucleotides, the simplest fragments of DNA or RNA. Oligoadenylate, A_n, in which n is the number of nucleotides, was chosen as a standard oligonucleotide. Fundamental chromatographic parameters necessary for the prediction system were obtained by isocratic elution chromatography. The capacity factor, k', of individual oligoadenylates containing less than eight nucleotides were measured under various isocratic elution conditions: eluent concentrations, C, and column temperature. The log k' values of oligoadenylates were found to be linearly related to the number of nucleotides, n. The k' values for oligoadenylates with n values > 7, which are difficult to obtain, were estimated by extrapolation of the linear relationship between log k' and n. The constants a and b were calculated from the relationship between k' and C. The retention times under gradient elution conditions were predicted by means of the present system, in which fundamental parameters were stored. The predicted retention times in gradient anion-exchange chromatography were compared with the observed ones with binary-linear and binary-convex gradient shapes.

EXPERIMENTAL

Chemicals

Unless otherwise stated, guaranteed reagents from Wako (Osaka, Japan) were used without further purification. Polyadenylate, sodium salt [poly(A)] was obtained from Yamasa Shoyu (Chiba, Japan). Oligoadenylate fragments from poly(A) were prepared by alkaline hydrolysis of poly(A) according to the literature[5]. Poly(A), digested in alkaline solution, gave a series of oligoadenylates, A_n, with varying chain lengths. Each A_n contains 2'- and 3'-terminal phosphates.

HPLC equipment

A TRI-ROTAR II HPLC system (Jasco, Tokyo, Japan) was used. The separations were performed on a commercially available column[10] (Shim-pack WAX-1, a weakly basic anion exchanger, 50 mm × 4 mm I.D.; Shimadzu, Kyoto, Japan). The Shim-pack WAX-1 is packed with a spherical silica support, having a particle diameter of 3 μm and a pore size of 100 Å and tertiary amino groups chemically bonded to the silica support. The column dead-time, t_0, was measured by injecting water and found to be 0.55 min. The separation column was surrounded by a jacket containing circulating water at constant temperature within ± 0.1°C. The sample solution (100 μl) was injected into a separation column and chromatographed at a flow-rate of 1.0 ml/min. Oligoadenylate fragments, eluted from the column, were monitored at 260 nm by means of an UVIDEC-100 spectrophotometric detector (Jasco).

Eluent

Fundamental chromatographic parameters were measured in isocratic elution with phosphate buffer solutions (pH 6.8), containing appropriate concentrations of phosphate and 20% acetonitrile.

The binary gradient elution technique was used for oligoadenylate separations. Buffer A was 0.01 M phosphate buffer (pH 6.8) containing 20% acetonitrile and buffer B was 0.3 M phosphate buffer (pH 6.8) containing 20% acetonitrile.

Calculation

Retention times and band widths were calculated with a personal computer PC-9801 (NEC, Tokyo, Japan). The program was written in BASIC. The full listings are available on request from the authors.

RESULTS AND DISCUSSION

Procedure for prediction of retention times and band widths

The first step in the procedure for prediction of retention times of oligonucleotides under gradient elution conditions is to determine constants a and b characteristic of each solute. These constants are calculated from the relatonship between the eluent concentratons, C, and the capacity factors, k', measured under isocratic elution conditions[20,27]:

$$k' = aC^{-b} \tag{1}$$

In gradient ion-exchange chromatography, the solute retenton times, t_R (min), for oligonucleotides were calculated from eqn. 2 by means of constants a and b[20-22,29,31].

$$t_R = (1/u)\{(1/B')[(xb + 1)B' \, at_0 u + C_i^{(xb + 1)/x}]^{1/(xb + 1)} - C_i^{1/x}/B'\} + t_0 \tag{2}$$

where u is the flow-rate (ml/min) and t_0 is the column dead-time (min); C_i, B' and x are adjustable parameters for the gradient profile.

The following form of the gradient-profile function was chosen so as to be applicable to a great variety of gradient shapes

$$C = (C_i^{1/x} + Bt)^x$$
$$B = (C_f^{1/x} - C_i^{1/x})/t_f \tag{3}$$

where C is the eluent concentration (M) at time t (min), C_i is the initial eluent concentration (M) at the beginning of the gradient elution ($t = 0$ min) and C_f is the final one at the end of gradient elution ($t = t_f$); $B' = B/u$. The parameter x characterizes the shape of the gradient profile: at $x = 1$, the change in eluent concentration is linear, while the concentration gradient is convex for $x < 1$ and concave for $x > 1$. Gradient profiles with the parameters $x = 0.58$ and $x = 1$ were used in the present study, as shown in Fig. 1.

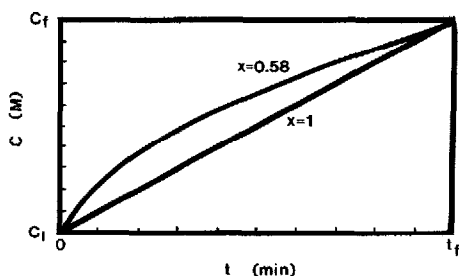Band widths of oligonucleotides should be estimated to calculate chromato-

Fig. 1. Gradient profiles used in the present study. For details see text.

graphic resoluton in addition to solute retention time. Band widths, w ($=4\sigma$ in time units) are calculated as follows

$$w = 4t_R/N_g^{1/2} \tag{4}$$

where N_g is the plate number in gradient elution chromatography.

Fig. 2 illustrates the plots of observed band widths against observed retention times in gradient anion-exchange chromatography of oligoadenylates at column temperatures of 40 (Fig. 5), 50 an 60°C. It was surprising that the column temperature did not contribute to a decrease in band widths, as shown in Fig. 2.

The calculated band widths (solid line in Fig. 2) were in good agreement with the observed ones, when N_g was 5000. As a result, in the present system, band widths at all column temperatures were calculated from eqn. 4, using this value of N_g.

The simple expression (eqn. 4) is a first approximation for the estimation of band widths in gradient elution chromatography. Other N_g values should be chosen for the estimation of band widths in other chromatographic systems, because the value of 5000 would be applicable for a specific chromatographic system.
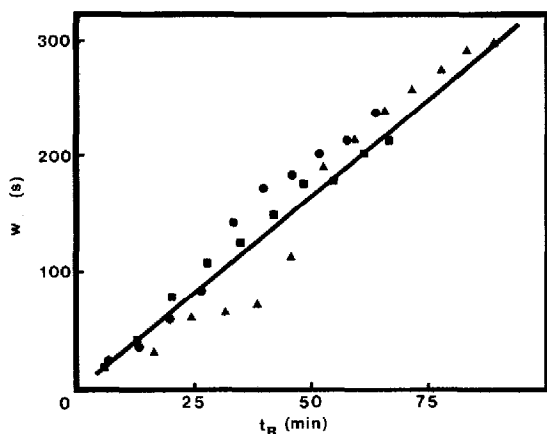

Fig. 2. Relationship between the band width, w, and retention time, t_R, at various temperatures: ● $= 40$ (Fig. 5); ▲ $= 50$; ■ $= 60°C$ in gradient elution chromatography. The solid line represents the band width calculated from eqn. 4 using an N_g value of 5000.
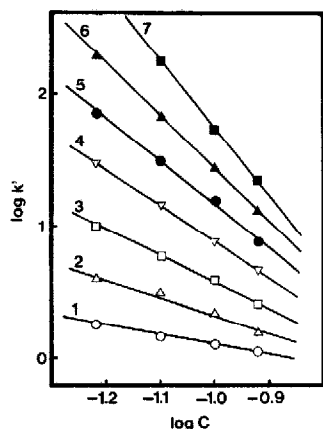
Fig. 3. Plots of log k' vs. log C at pH 6.8 and a column temperature of 40°C. Eluent: phosphate buffer solution (pH 6.8) containing 20% acetonitrile. Numbers indicate the number of nucleotides in the oligoadenylates.

Determination of constants a and b in isocratic elution

As described in the above section, constants a and b must be determined prior to the prediction of retention times. As shown in Fig. 3, the plot of log k' vs. log C gave straight lines with correlation coefficients between 0.985 and 0.999 at a column temperature of 40°C. Constants a and b for individual solutes of purified low-molecular-weight nucleotides (from A_1 to A_7) were determined from the slope and intercept of the straight line in Fig. 3 and are compiled in Table I.

In previous papers, we demonstrated that the column temperature was an indispensable factor in improving the chromatographic resolution of inorganic polyphosphates in both isocratic[27] and gradient[28] anion-exchange chromatography. However, it scarcely affected the chromatographic behaviour of oligonucleotides. Consequently, constants a and b in Table I were used for the prediction of retention times at all column temperatures.

Isocratic elution experiments to determine k' values were performed for individual oligonucleotides up to seven bases in chain length, as shown in Fig. 3. When individual nucleotides were difficult to obtain, such as the purified higher oligo-

TABLE I

VALUES OF CONSTANTS a AND b IN EQN. 1 AT A COLUMN TEMPERATURE OF 40°C

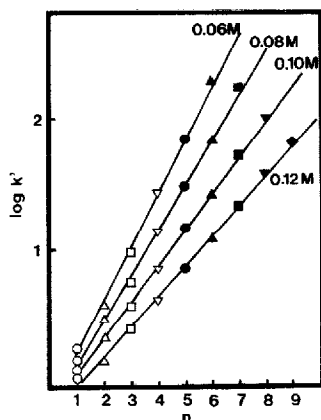| Solute	b	a
A_1	0.737	$2.32 \cdot 10^{-1}$
A_2	1.37	$8.95 \cdot 10^{-2}$
A_3	1.92	$4.50 \cdot 10^{-2}$
A_4	2.71	$1.44 \cdot 10^{-2}$
A_5	3.25	$8.09 \cdot 10^{-3}$
A_6	3.90	$3.41 \cdot 10^{-3}$
A_7	5.01	$5.38 \cdot 10^{-4}$

Fig. 4. Plots of log k' vs. the number of nucleotides of oligonucleotides, n, at various eluent concentrations (molar concentrations of phosphate buffer, pH 6.8). Column temperature: 40°C. ○ = Mono-; △ = di-; □ = tri-; ▽ = tetra-; ● = penta-; ▲ = hexa-; ■ = hepta-; ▼ = octa-; ◆ = nonaadenylate.

nucleotides or polynucleotides, an extrapolation method was employed to estimate k' values.

In isocratic elution reversed-phase chromatography, the log of the capacity factor for the members of an oligomeric series was related linearly to the polymerization number of the series for various organic compounds[32–37]. The relationship can be written as

$$\log k' = cn + d \tag{5}$$

where c and d are constant and n is the polymerization number. We found a similar relatonship in isocratic anion-exchange chromatography for oligomeric series of inorganic polyphosphates[29] and applied the extrapolation of the relationship to the retention prediction of higher polyphosphates.

The log k' values for the oligomeric series of oligoadenylates were found to increase linearly with increasing number of nucleotides, n, in isocratic anion-exchange chromatography (correlation coefficient = 0.999) at each eluent concentration (Fig. 4). Constants c and d of eqn. 5 were determined by linear regression analysis and are listed in Table II.

TABLE II

VALUES OF c AND d IN EQN. 5 FOR VARIOUS PHOSPHATE BUFFER CONCENTRATIONS

Buffer concentration (M)	c	d
0.06	0.409	−0.189
0.08	0.343	−0.209
0.10	0.278	−0.212
0.12	0.229	−0.252

TABLE III

ESTIMATED LOG a AND b VALUES AT A COLUMN TEMPERATURE OF 40°C

Solute	b	$\log a$	Solute	b	$\log a$	
A_8	5.10	-3.00	A_{22}	13.4	-7.35	
A_9	5.61	-3.33	A_{23}	14.0	-7.86	
A_{10}	6.21	-3.65	A_{24}	14.6	-8.18	
A_{11}	6.81	-3.97	A_{25}	15.2	-8.51	
A_{12}	7.41	-4.30	A_{26}	15.8	-8.83	
A_{13}	8.02	-4.62	A_{27}	16.5	-9.15	
A_{14}	8.62	-4.94	A_{28}	17.1	-9.48	
A_{15}	9.22	-5.27	A_{29}	17.7	-9.80	
A_{16}	9.82	-5.59	A_{30}	18.3	-10.12	
A_{17}	10.4	-5.92	A_{31}	19.0	-10.96	
A_{18}	11.0	-6.24	A_{32}	19.6	-11.30	
A_{19}	11.6	-6.56	A_{33}	20.2	-11.65	
A_{20}	12.2	-6.89	A_{34}	20.8	-11.99	
A_{21}	12.8	-7.21	A_{35}	21.4	-12.33	

The capacity factors of oligoadenylates containing more than seven nucleotides at various eluent concentrations were estimated by the extrapolation of the relationships. As a result, constants a and b for higher oligonucleotides were calculated from eqn. 1 and compiled in Tabe III.

Constants a and b listed in Tables I and III were stored in the program, which is described in a previous paper[31]. After the constants for solutes are obtained and stored in a computer system, the prediction of the retention time for each solute in gradient anion-exchange chromatography becomes possible.

TABLE IV

OBSERVED (Obs.) AND CALCULATED (Calc.) RETENTION TIMES (min) UNDER GRADIENT ELUTION CONDITIONS

$x = 1$; $C_i = 0.01$ M; $C_f = 0.3$ M; $t_f = 128$ min; $T = 40$°C.

Solute	Obs.	Calc.	Error(%)	Solute	Obs.	Calc.	Error(%)	
A_1	4.72	4.05	-14	A_{14}	76.3	79.9	4.7	
A_2	11.6	11.6	0	A_{15}	79.4	82.3	3.7	
A_3	19.8	20.3	2.5	A_{16}	82.2	84.4	2.7	
A_4	28.2	30.9	9.6	A_{17}	84.9	86.4	1.8	
A_5	35.5	40.2	13	A_{18}	87.5	88.1	0.69	
A_6	42.1	45.9	9.0	A_{19}	89.8	89.7	-0.11	
A_7	47.9	52.3	9.2	A_{20}	92.0	91.2	-0.87	
A_8	53.3	57.9	8.6	A_{21}	94.1	92.6	-1.6	
A_9	57.2	62.8	9.8	A_{22}	96.1	93.8	-2.3	
A_{10}	61.7	67.1	8.8	A_{23}	97.4	94.9	-2.6	
A_{11}	65.8	70.9	7.8	A_{24}	99.2	96.0	-3.2	
A_{12}	69.5	74.3	6.9	A_{25}	101	97.0	-4.0	
A_{13}	73.0	77.2	5.8					

TABLE V

OBSERVED (Obs.) AND CALCULATED (Calc.) RETENTION TIMES (min)

$x = 0.58$; $C_i = 0.01$ M; $C_f = 0.3$ M; $t_f = 240$ min; $T = 40°C$.

Solute	Obs.	Calc.	Error(%)	Solute	Obs.	Calc.	Error(%)
A_1	5.04	3.74	−26	A_{19}	122	116	−4.9
A_2	11.7	8.70	−26	A_{20}	127	120	−5.5
A_3	17.8	14.8	−17	A_{21}	132	123	−6.8
A_4	23.7	23.7	0	A_{22}	136	126	−7.4
A_5	30.7	33.3	8.5	A_{23}	140	130	−7.1
A_6	38.0	39.8	4.7	A_{24}	145	132	−9.0
A_7	45.0	48.1	6.9	A_{25}	149	135	−9.4
A_8	52.0	56.0	7.7	A_{26}	153	137	−10
A_9	58.8	63.8	8.0	A_{27}	158	140	−11
A_{10}	65.1	70.6	8.4	A_{28}	162	142	−12
A_{11}	72.0	77.2	7.2	A_{29}	165	144	−13
A_{12}	78.9	83.4	5.7	A_{30}	169	146	−14
A_{13}	86.0	89.2	3.7	A_{31}	172	147	−15
A_{14}	92.9	94.6	1.8	A_{32}	175	149	−15
A_{15}	99.6	99.6	0	A_{33}	179	151	−16
A_{16}	106	104	−1.9	A_{34}	182	152	−16
A_{17}	112	109	−2.7	A_{35}	185	154	−17
A_{18}	117	113	−3.4				

Prediction of retention times for oligonucleotides

To predict retention times, the following parameters are input into the prediction system: column dead-time (min); flow-rate (ml/min); N_g value; adjustable parameters for the gradient profile (x, C_i, C_f and t_f).

The retention times for oligoadenylates were predicted and compared with the observed t_R values to test the performance of the present system under gradient elution conditions. Tables IV and V list the results with binary-linear ($x = 1$) and binary-convex ($x = 0.58$) gradient shapes (Fig. 1), respectively. The 60 observed t_R values were predicted with an average error of 7.7%.

To discuss the accuracy of the prediction in more detail, the errors of prediction and the range of n values are 6.4% ($n < 25$) and 13.9% ($n \geqslant 25$). Predictions reported in gradient reversed-phase chromatography[21,25,26,38,39] were achieved in the range of errors from 3 to 8%. In comparison with the reported errors, our result for lower oligoadenylates was satisfactory but that for higher oligoadenylates was slightly poor. The reason is that the constants a and b used in the predictions were extrapolated from the data for the first seven oligoadenylates, while the t_R values reported in reversed-phase chromatography were predicted without extrapolation methods. As a result, we conclude that the extrapolation of the relationship between $\log k'$ and n is successful in the estimation of capacity factors for higher oligoadenylates under gradient elution conditions.

As an example, the predicted chromatogram displayed on the monitor screen was compared with a chromatogram observed under the gradient elution conditions in Table V, as shown in Fig. 5. In the predicted chromatogram, the peaks were given gaussian profiles by use of the predicted retention times (Table V) and band widths
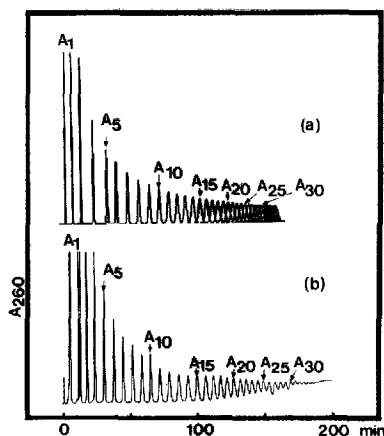
Fig. 5. Predicted (a) and observed (b) chromatograms for a polyadenylate partial hydrolyzate. (a) Gradient elution conditions as in Table V. (b) Column: Shim-pack WAX-1 (weak anion exchanger, 50 mm × 4 mm I.D.). Flow-rate: 1.0 ml/min. Buffer: A, 0.01 M phosphate (pH 6.8) containing 20% acetonitrile; B, 0.3 M phosphate (pH 6.8) containing 20% acetonitrile. Gradient elution conditions as in Table V. Column temperature: 40°C.

calculated by eqn. 4. The area of each peak was assumed to be the same. The characteristics of each chromatogram are similar with respect to the retention times, band widths and resolution.

When the present prediction system is applied to the other separation variables including stationary phases, mobile phases and equipment, constants a and b in eqn. 1 will be obtained with a minimum number of experiments (at least three sets of experiments under isocratic elution conditions).

ACKNOWLEDGEMENT

REFERENCES

1 J. B. Crowther and R. A. Hartwick, in P. R. Brown (Editor), *HPLC in Nucleic Acid Research*, Marcel Dekker, New York, 1984, Ch. 11, pp. 195–213.
2 L. W. McLaughlin and R. Bischoff, *J. Chromatogr.*, 418 (1987) 51.
3 R. Hecker and D. Riesner, *J. Chromatogr.*, 418 (1987) 97.
4 R. D. Wells, S. C. Hardies, G. T. Horn, B. Klein, J. E. Larson, S. K. Neuendorf, N. Panayotatos, R. K. Patient and E. Selsing, *Methods Enzymol.*, 65 (1980) 327.
5 J. D. Pearson and F. E. Regnier, *J. Chromatogr.*, 255 (1983) 137.
6 M. Colpan, D. Riesner, *J. Chromatogr.*, 296 (1984) 339.
7 R. R. Drager and F. E. Regnier, *Anal. Biochem.*, 145 (1985) 47.
8 R. R. Drager and F. E. Regnier, *J. Chromatogr.*, 359 (1986) 147.
9 P. A. D. Edwardson, C. R. Lowe and T. Atkinson, *J. Chromatogr.*, 368 (1986) 363.
10 T. Ueda and Y. Ishida, *J. Chromatogr.*, 386 (1987) 273.
11 J. B. Crowther, S. D. Fazio and R. A. Hartwick, *J. Chromatogr.*, 282 (1983) 619.

12 R. Bischoff and L. W. McLaughlin, *J. Chromatogr.*, 296 (1984) 329.
13 R. Bischoff and L. W. McLaughlin, *Anal. Biochem.*, 151 (1985) 526.
14 E. Selsing, J. E. Larson and R. D. Wells, *Anal. Biochem.*, 99 (1979) 213.
15 R. Lohrmann and L. E. Orgel, *J. Mol. Biol.*, 142 (1980) 555.
16 H. Sawai, *Nucleic Acids Res. Symp. Ser.*, No. 15 (1984) 105.
17 G. D. McFarland and P. N. Borer, *Nucleic Acids Res.*, 7 (1979) 1067.
18 H.-J. Fritz, R. Belagaje, E. L. Brown, R. H. Fritz, R. A. Jones, R. G. Lees and H. G. Khorana, *Biochemistry*, 17 (1978) 1259.
19 G. D'Agostino, L. Castagnetta, F. Mitchell and M. J. O'Hare, *J. Chromatogr.*, 338 (1985) 1.
20 P. Jandera and J. Churáček, *J. Chromatogr.*, 91 (1974) 223.
21 P. Jandera and J. Churáček, *Adv. Chromatogr. (N.Y.)*, 19 (1981) 125.
22 P. Jandera and J. Churáček, *Gradient Elution in Column Liquid Chromatography*, Elsevier, Amsterdam, 1985.
23 L. R. Snyder, J. W. Dolan and J. R. Gant, *J. Chromatogr.*, 165 (1979) 3.
24 K. Jinno and K. Kawasaki, *J. Chromatogr.*, 316 (1984) 1.
25 S. A. Tomellini, R. A. Hartwick and H. B. Woodruff, *Anal. Chem.*, 57 (1985) 811.
26 K. Jinno and M. Kuwajima, *Chromatograhia*, 21 (1986) 541.
27 Y. Baba, N. Yoza and S. Ohashi, *J. Chromatogr.*, 348 (1985) 27.
28 Y. Baba, N. Yoza and S. Ohashi, *J. Chromatogr.*, 350 (1985) 119.
29 Y. Baba, N. Yoza and S. Ohashi, *J. Chromatogr.*, 350 (1985) 461.
30 Y. Baba, *J. Assoc. Pers. Comput. Chem.*, 7, No. 1 (1985) 119.
31 Y. Baba, *J. Assoc. Pers. Comput. Chem.*, 7, No. 3 (1985) 41.
32 H. Colin and G. Guiochon, *J. Chromatogr. Sci.*, 18 (1980) 54.
33 P. Dufek and E. Smolková, *J. Chromatogr.*, 257 (1983) 247.
34 P. Dufek, *J. Chromatogr.*, 281 (1983) 49.
35 A. Tchapla, H. Colin and G. Guiochon, *Anal. Chem.*, 56 (1984) 621.
36 P. Jandera, *J. Chromatogr.*, 314 (1984) 13.
37 P. Jandera, *Chromatographia*, 19 (1984) 101.
38 P. Jandera, J. Churáček and L. Svoboda, *J. Chromatogr.*, 174 (1979) 35.
39 S. A. Tomellini, S.-H. Hsu and R. A. Hartwick, *Anal Chem.*, 58 (1986) 904.