JOURNAL OF
CHROMATOGRAPHY A
INCLUDING ELECTROPHORESIS AND OTHER SEPARATION METHODS

ELSEVIER

Scope. The *Journal of Chromatography A* publishes papers on all aspects of **chromatography, electrophoresis** and related methods. Contributions consist mainly of research papers dealing with chromatographic theory, instrumental developments and their applications. In the *Symposium volumes*, which are under separate editorship, proceedings of symposia on chromatography, electrophoresis and related methods are published. *Journal of Chromatography B: Biomedical Applications* —This journal, which is under separate editorship, deals with the following aspects: developments in and applications of chromatographic and electrophoretic techniques related to clinical diagnosis or alterations during medical treatment; screening and profiling of body fluids or tissues related to the analysis of active substances and to metabolic disorders; drug level monitoring and pharmacokinetic studies; clinical toxicology; forensic medicine; veterinary medicine; occupational medicine; results from basic medical research with direct consequences in clinical practice.

Submission of Papers. The preferred medium of submission is on disk with accompanying manuscript (see *Electronic manuscripts* in the Instructions to Authors, which can be obtained from the publisher, Elsevier Science B.V., P.O. Box 330, 1000 AH Amsterdam, Netherlands). Manuscripts (in English; *four* copies are required) should be submitted to: Editorial Office of *Journal of Chromatography A*, P.O. Box 681, 1000 AR Amsterdam, Netherlands, Telefax (+31-20) 485 2304, or to: The Editor of *Journal of Chromatography B: Biomedical Applications*, P.O. Box 681, 1000 AR Amsterdam, Netherlands. Review articles are invited or proposed in writing to the Editors who welcome suggestions for subjects. An outline of the proposed review should first be forwarded to the Editors for preliminary discussion prior to preparation. Submission of an article is understood to imply that the article is original and unpublished and is not being considered for publication elsewhere. For copyright regulations, see below.

Publication information. *Journal of Chromatography A* (ISSN 0021-9673): for 1995 Vols. 683–714 are scheduled for publication. *Journal of Chromatography B: Biomedical Applications* (ISSN 0378-4347): for 1995 Vols. 663–674 are scheduled for publication. Subscription prices for *Journal of Chromatography A*, *Journal of Chromatography B: Biomedical Applications* or a combined subscription are available upon request from the publisher. Subscriptions are accepted on a prepaid basis only and are entered on a calendar year basis. Issues are sent by surface mail except to the following countries where air delivery via SAL is ensured: Argentina, Australia, Brazil, Canada, China, Hong Kong, India, Israel, Japan, Malaysia, Mexico, New Zealand, Pakistan, Singapore, South Africa, South Korea, Taiwan, Thailand, USA. For all other countries airmail rates are available upon request. Claims for missing issues must be made within six months of our publication (mailing) date. Please address all your requests regarding orders and subscription queries to: Elsevier Science B.V., Journal Department, P.O. Box 211, 1000 AE Amsterdam, Netherlands. Tel.: (+31-20) 485 3642; Fax: (+31-20) 485 3598. Customers in the USA and Canada wishing information on this and other Elsevier journals, please contact Journal Information Center, Elsevier Science Inc., 655 Avenue of the Americas, New York, NY 10010, USA, Tel. (+1-212) 633 3750, Telefax (+1-212) 633 3764.

Abstracts/Contents Lists published in Analytical Abstracts, Biochemical Abstracts, Biological Abstracts, Chemical Abstracts, Chemical Titles, Chromatography Abstracts, Current Awareness in Biological Sciences (CABS), Current Contents/Life Sciences, Current Contents/Physical, Chemical & Earth Sciences, Deep-Sea Research/Part B: Oceanographic Literature Review, Excerpta Medica, Index Medicus, Mass Spectrometry Bulletin, PASCAL-CNRS, Referativnyi Zhurnal, Research Alert and Science Citation Index.

US Mailing Notice. *Journal of Chromatography A* (ISSN 0021-9673) is published weekly (total 52 issues) by Elsevier Science B.V., (Sara Burgerhartstraat 25, P.O. Box 211, 1000 AE Amsterdam, Netherlands). Annual subscription price in the USA US$ 5389.00 (US$ price valid in North, Central and South America only) including air speed delivery. Second class postage paid at Jamaica, NY 11431. **USA POSTMASTERS**: Send address changes to *Journal of Chromatography A*, Publications Expediting, Inc., 200 Meacham Avenue, Elmont, NY 11003. Airfreight and mailing in the USA by Publications Expediting.

See inside back cover for Publication Schedule, Information for Authors and information on Advertisements.

CONTENTS

REGULAR PAPERS

Contents (continued)

SHORT COMMUNICATIONS

Column Liquid Chromatography

JOURNAL OF
CHROMATOGRAPHY A

INCLUDING ELECTROPHORESIS AND OTHER SEPARATION METHODS

ELSEVIER

Amsterdam – Lausanne – New York – Oxford – Shannon – Tokyo

ห้องสมุดกรมวิ... ...เกษตรบริการ

- 7 ก.ย. 2538

Retention mechanism and implications for selectivity for a group of dihydropyridines in ionic micellar liquid chromatography

J.M. Saz, M.L. Marina*

Department of Analytical Chemistry, Faculty of Sciences, University of Alcalá de Henares, Ctra. Madrid–Barcelona Km 33.6, 28871 Alcalá de Henares (Madrid), Spain

Abstract

The retention behaviour of a group of dihydropyridines in micellar liquid chromatography was studied using sodium dodecyl sulphate and hexadecyltrimethylammonium bromide as surfactants in the mobile phase containing 5% of n-butanol and a C_{18} column. When the surfactant concentration in the mobile phase is increased, a tendency to change from a three partition equilibria mechanism to direct transfer of solutes from micelles to the stationary phase is observed for both surfactants. This progressive change in the retention mechanism is explained through the large micellar phase–water partition coefficients of these compounds and the increase produced in the fraction of solute molecules in the micellar phase due to the increase in the volume of this phase originating from the increase in surfactant concentration. As a result, the selectivity coefficients show a tendency to match the ratio of the stationary phase to micellar phase partition coefficients of these compounds, constituting further proof of the progressive change in the retention mechanism when the surfactant concentration is increased.

1. Introduction

The use of micellar systems as mobile phases in high-performance liquid chromatography (HPLC) has given rise to micellar liquid chromatography (MLC) [1]. This confers on HPLC the unique properties of micelles, allowing the analysis of ionic solutes [2,3] and of polar and nonpolar solutes simultaneously [4] and improving the selectivity through variations of the surfactant nature and concentration [5–7].

Micelles constitute a different phase or pseudo-phase in the bulk mobile phase and retention

is determined by the solute affinity toward the aqueous mobile phase, the stationary phase and the micellar mobile phase. A three partition equilibria mechanism has been proposed to explain retention in MLC [8]. Each solute undergoes partitioning between the stationary phase, the aqueous phase and the micellar pseudo-phase. Two equations that arise from this mechanism have been developed to predict retention in MLC [8,9]:

$$\frac{V_s}{(V_e - V_m)} = \frac{v(P_{mw} - 1)}{P_{sw}} \cdot C_M + \frac{1}{P_{sw}} \tag{1}$$

$$\frac{1}{k'} = \frac{V_m K_2}{V_s K_1 [L]} \cdot C_M + \frac{V_m}{V_s K_1 [L]} \tag{2}$$

* Corresponding author.

where k' is the capacity factor of the solute, K_1 and K_2 the solute–stationary phase and –micelle association constants, C_M the micellized surfactant concentration (total surfactant concentration minus the critical micellar concentration, cmc), [L] the stationary phase concentration, V_s, V_m, V_e and v the stationary and mobile phase volume, the eluting volume of the solute and the molar volume of the surfactant, respectively, and P_{mw} and P_{sw} the micelle–water and stationary phase–water partition coefficients of the solute, respectively.

Borgerding et al. [10] have proposed a limit theory for those compounds whose affinity towards the micellar phase is large enough to experience a direct transfer from this phase to the stationary phase. These solutes only undergo partitioning between the micellar pseudo-phase and the stationary phase, as the amount of solute in the aqueous phase is almost negligible. The capacity factor (k') for these compounds is defined by

$$k' = \frac{V_s}{V_m} \cdot \frac{P_{sm}}{vC_M} \qquad (3)$$

The implications of this direct transfer for chromatographic parameters such as selectivity have not been considered in the data published previously.

In this study, new evidence of the direct transfer theory is presented to help interpret the chromatographic behaviour of a group of dihydropyridines (DHPs) in an MLC system in which sodium dodecyl sulphate (SDS) and hexadecyltrimethylammonium bromide (CTAB) are used as surfactants in the mobile phase. The results of the application of this theory to selectivity and some supporting data are presented.

2. Experimental

The HPLC system consisted of a Model 510 pump, U6K injector, Model 481 variable-wavelength UV–Vis detector and Model 740 data module (all from Waters). The analytical columns (150 mm × 3.9 mm I.D.) were packed by a

column packing company (Tracer, Barcelona, Spain) using a commercially available C_{18} packing material (e.g., 5-μm Spherisorb ODS-2). The column was water-jacketed and temperature controlled (30 ± 1°C) by a circulating bath.

All reagents (e.g., SDS, CTAB, n-butanol and sodium phosphate buffer) were purchased from Merck. DHPs were synthesized in the Department of Organic Chemistry, University of Alcalá de Henares, Spain. Fig. 1 groups the structures of the DHPs studied and their identification numbers used through the paper.

Mobile phases were prepared by dissolving in HPLC-grade water the appropriate amounts of surfactant, n-butanol and sodium phosphate buffer in an ultrasonic bath followed by filtration through a 0.45-μm filter. Mobile phases were degassed in the ultrasonic bath prior to their utilization. Stock solutions of test solutes were prepared in the mobile phase itself and injected directly (20 μl) into the chromatographic system. The solute concentration was arbitrarily adjusted to permit detection. Solutes were injected in triplicate for every mobile phase condition used and the mean value of the retention times was used for calculations. The void volume was determined from the first deviation of the baseline and the stationary phase volume was taken as the difference between the total volume of the column and the void volume.

3. Results and discussion

3.1. Retention mechanism

The predicted linear relationship between $1/k'$ or $V_s/(V_e - V_m)$ versus C_M according to Eqs. 1 and 2 has been demonstrated in literature for many different types of solutes, surfactants and stationary phases [10–12]. Further, good agreement has been found between the partition or binding constants obtained from these equations and those determined by alternative methods [9,13].

As predicted by theory, a linear relationship between $1/k'$ and C_M is also shown in Fig. 2a–d for the studied DHPs when SDS or CTAB is
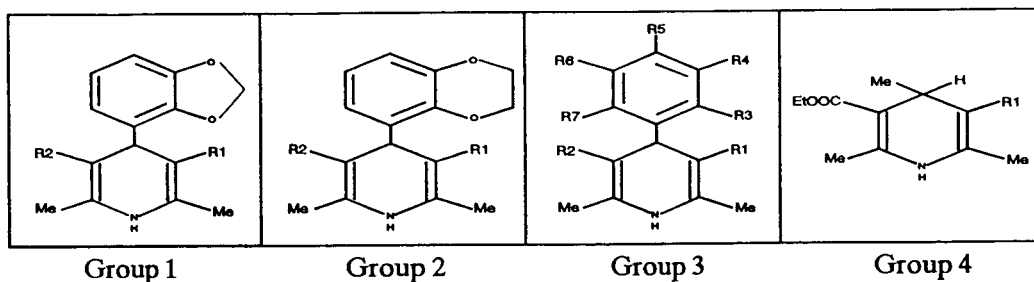
Group 1	-R$_1$	-R$_2$
1	-COOCH$_2$-(1)	-COOCH$_3$
2	-COOCH$_3$	-COOCH$_3$
3	-COOCH$_2$CH$_2$OCH$_3$	-COOCH$_2$CH$_2$OCH$_3$
5	-COOCH$_2$CH$_3$	-COOCH$_2$CH$_3$
6	-COOCH$_2$CH$_2$OCH$_3$	-COOIsp
7	-COOIsp	-COOCH$_2$CH$_3$
12	-COOCH$_2$CH$_3$	-COOCH$_3$
14	-COOIsp	-COOIsp
17	-COOIsp	-COOCH$_3$
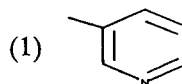
Group 2	-R$_1$	-R$_2$
4	-COOCH$_2$CH$_3$	-COOCH$_3$

Group 3	-R$_1$	-R$_2$	-R$_3$	-R$_4$	-R$_5$	-R$_6$	-R$_7$
8	-COOCH$_2$CH$_3$	-COOCH$_3$	-H	-NO$_2$	-H	-H	-H
9	-COOCH$_2$CH$_3$	-COOCH$_3$	-Cl	-Cl	-H	-H	-H
10	-COOCH$_2$CH$_3$	-COOCH$_3$	-OCH$_2$OCH$_3$	-H	-H	-H	-H
11	-COOCH$_2$CH$_3$	-COOCH$_3$	-H	-OCH$_3$	-H	-OCH$_3$	-H
13	-COOCH$_2$CH$_3$	-COOCH$_3$	-H	-H	-H	-H	-H
15	-COOCH$_2$CH$_3$	-COOCH$_3$	-H	-OCH$_3$	-OCH$_3$	-H	-H
16	-COOCH$_2$CH$_3$	-COOCH$_3$	-OCH$_3$	-OCH$_3$	-H	-H	-H
18	-COOCH$_2$CH$_3$	-COOCH$_3$	-H	-OCH$_3$	-OH	-OCH$_3$	-H
19	-COOCH$_2$CH$_3$	-COOCH$_3$	-H	-OCH$_3$	-OCH$_3$	-OCH$_3$	-H
20	-COOCH$_2$CH$_3$	-COOCH$_2$CH$_3$	-Cl	-H	-H	-H	-Cl
21	-COOCH$_2$CH$_3$	-COOCH$_2$CH$_3$	-Cl	-H	-H	-H	-H

Group 4	-R$_1$
22	-COOCH$_3$
23	-COOCH$_2$CH$_2$-(2)
24	-COOCH$_2$CH$_3$
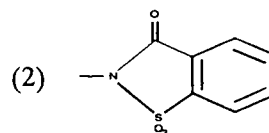
Fig. 1. Structures of the DPHs studied and the number assigned to each.
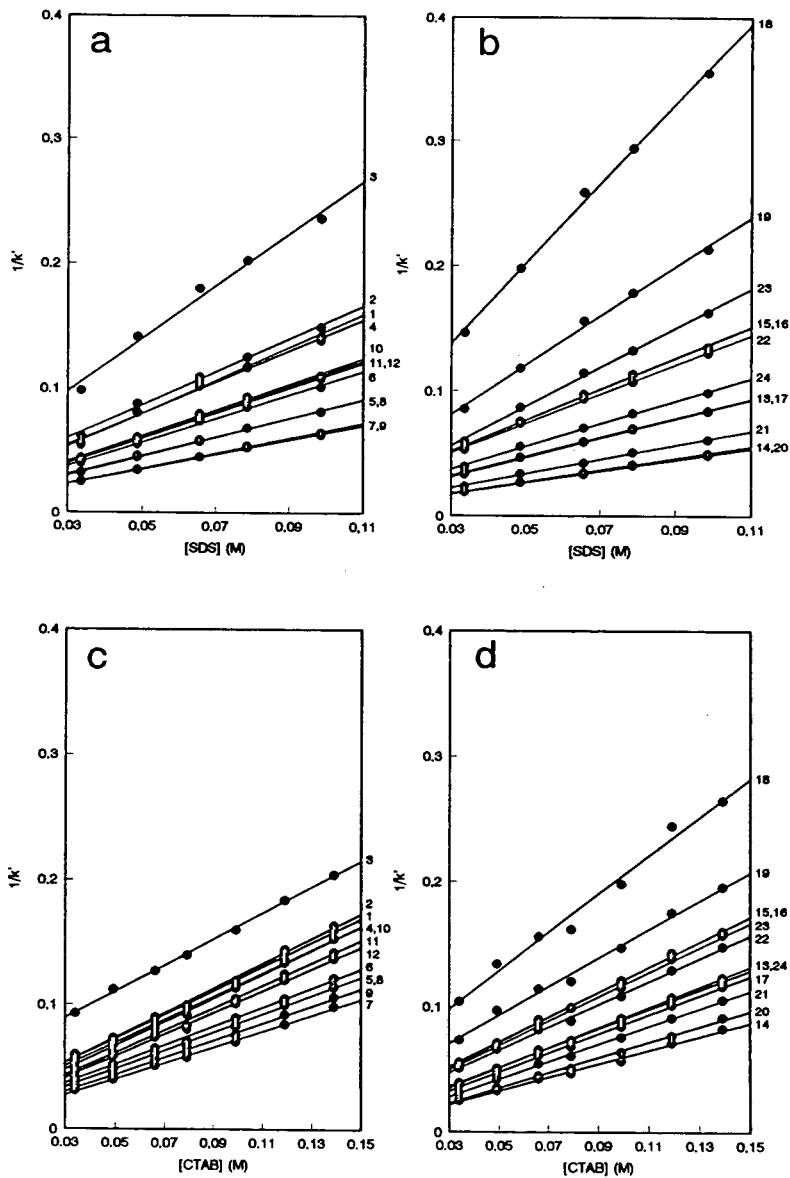
Fig. 2. Inverse of capacity factor ($1/k'$) versus micellized surfactant concentration (C_M) in the mobile phase (total surfactant concentration minus its critical micellar concentration) for each DHP. The composition of mobile phases was 0.01 M sodium phosphate buffer (pH 6.7)–5% (v/v) n-butanol with a total surfactant concentration ranging from 0.035 to 0.1 M for SDS and 0.035 to 0.14 M for CTAB.

present in the mobile phase above their cmc. A small amount of an organic additive (e.g., n-butanol) and phosphate buffer of pH 6.7 were

added to the mobile phase to decrease the retention of solutes and to control the pH, as these types of compound present a large re-

tention and are quoted in the literature as alkaloid compounds [14]. A similar linear relationship also exists for $V_s/(V_e - V_m)$ versus C_M (data not shown). The parameters of these straight lines are given in Tables 1 and 2 and were used to determine K_2, P_{sw} and P_{mw} using Eqs. 1 and 2. The stationary phase–micellar phase partition coefficients (P_{sm}) were obtained from the ratio P_{sw}/P_{mw}.

The relative error for the solute–micelle association constants shown in Tables 1 and 2 was determined statistically. The means ± standard deviations were 26 ± 7% for SDS and 16 ± 7% for CTAB. In general terms, we observed that the error increased when the intercept of straight lines in Fig. 2a–d decreased. This is primarily due to the fact that the K_2 constants are determined from the slope/intercept ratio of these

straight lines and, therefore, small intercept values produce large statistical errors in the determination of solute–micelle association constants.

In general terms, the K_2 constants and the related P_{mw} coefficients shown in Tables 1 and 2 and in Fig. 3 are larger for SDS than CTAB. SDS and CTAB possess a long hydrocarbon chain with twelve and sixteen carbons, respectively. This means that hydrophobic interactions between DHPs and SDS or CTAB should be higher, or at least similar (depending on the molecular size of the solute), in magnitude for CTAB than SDS. If DHPs bind to SDS and CTAB micelles only through hydrophobic forces, then the solute–micelle association parameters (e.g., K_2 and P_{mw}) should also be higher for CTAB or at least similar for both. As

Table 1
Parameters of straight lines obtained when retention factors of Eqs. 1 and 2 using SDS are plotted versus the micellized surfactant concentration and the equilibrium constants calculated from these parameters

DHP	Eq. 1				Eq. 2						
	Slope	Intercept ($\times 10^2$)	r	K_2	Slope	Intercept ($\times 10^2$)	r	K_2	P_{sw}	P_{mw}	P_{sm}
1	1.32	1.37	0.9937	96.16	1.64	1.70	0.9936	96.01	58.64	391.24	0.15
2	1.32	2.06	0.9971	63.95	1.63	2.56	0.9971	63.97	39.10	261.04	0.15
3	2.10	3.44	0.9932	61.07	2.61	4.26	0.9932	61.22	23.46	249.86	0.09
4	1.25	1.74	0.9965	71.63	1.54	2.16	0.9964	71.50	46.25	291.65	0.16
5	0.75	0.86	0.9983	86.83	1.01	0.64	0.9989	158.73	156.84	646.24	0.24
6	0.95	0.90	0.9979	105.37	1.18	1.13	0.9978	104.29	88.63	424.94	0.21
7	0.59	0.56	0.9988	104.02	0.73	0.69	0.9988	105.26	144.48	428.89	0.34
8	0.76	0.69	0.9991	110.52	0.95	0.85	0.9990	111.10	117.08	452.63	0.26
9	0.61	0.49	0.9990	125.27	0.76	0.61	0.9990	123.43	163.24	502.75	0.32
10	1.03	1.03	0.9984	99.85	1.28	1.28	0.9985	99.91	78.11	407.14	0.19
11	1.00	0.98	0.9993	102.37	1.24	1.22	0.9994	102.12	82.19	416.12	0.20
12	1.00	1.11	0.9979	90.00	1.24	1.37	0.9980	90.83	73.01	370.23	0.20
13	0.77	0.90	0.9993	85.73	0.95	1.11	0.9993	85.62	89.76	349.05	0.26
14	0.45	0.40	0.9990	112.68	0.56	0.50	0.9991	112.37	200.65	457.79	0.44
15	1.25	1.42	0.9986	87.79	1.54	1.76	0.9985	87.64	56.74	357.26	0.16
16	1.25	1.42	0.9985	87.53	1.55	1.77	0.9985	87.38	56.50	356.20	0.16
17	0.78	0.75	0.9991	103.90	0.97	0.92	0.9990	104.81	108.21	427.06	0.25
18	3.21	4.18	0.9988	76.81	3.98	5.19	0.9988	76.80	19.31	313.20	0.06
19	1.97	2.19	0.9980	90.01	2.45	2.73	0.9980	89.75	36.65	365.84	0.10
20	0.47	0.37	0.9990	125.71	0.58	0.47	0.9990	124.06	212.74	505.31	0.42
21	0.57	0.53	0.9984	106.02	0.70	0.65	0.9984	107.99	153.05	439.98	0.35
22	1.18	1.47	0.9985	80.56	1.46	1.82	0.9985	80.33	54.82	327.54	0.17
23	1.57	0.89	0.9988	176.66	1.95	1.10	0.9988	177.54	91.00	722.71	0.13
24	0.91	0.99	0.9987	92.31	1.13	1.23	0.9986	91.61	81.03	373.40	0.22

Table 2
Parameters of straight lines obtained when retention factors of Eqs. 1 and 2 using CTAB are plotted versus the micellized surfactant concentration and the equilibrium constants calculated from these parameters.

DHP	Eq. 1				Eq. 2						
	Slope	Intercept $(\times 10^2)$	r	K_2	Slope	Intercept $(\times 10^2)$	r	K_2	P_{sw}	P_{mw}	P_{sm}
1	0.97	2.49	0.9997	38.85	1.20	3.09	0.9997	38.89	32.28	106.22	0.30
2	0.99	2.49	0.9997	39.88	1.23	3.09	0.9997	39.87	33.53	114.02	0.29
3	1.05	5.81	0.9983	18.03	1.30	7.20	0.9992	18.00	13.89	50.50	0.28
4	0.93	2.38	0.9989	38.90	1.15	2.96	0.9989	38.81	35.16	112.13	0.31
5	0.73	1.17	0.9992	62.61	0.91	1.46	0.9992	62.40	68.38	171.85	0.40
6	0.76	1.46	0.9998	52.00	0.94	1.82	0.9998	51.94	57.64	150.86	0.38
7	0.64	0.78	0.9989	82.66	0.80	0.97	0.9990	82.44	103.25	227.26	0.45
8	0.73	1.24	0.9995	59.30	0.91	1.53	0.9995	59.22	69.34	174.57	0.40
9	0.69	0.92	0.9986	74.43	0.85	1.15	0.9986	74.29	87.36	206.05	0.42
10	0.95	1.95	0.9990	48.68	1.18	2.42	0.9990	48.74	41.32	134.65	0.31
11	0.89	1.74	0.9998	51.38	1.11	2.16	0.9998	51.28	48.83	150.23	0.33
12	0.87	1.59	0.9999	54.70	1.08	1.96	0.9999	54.79	53.67	160.23	0.33
13	0.81	1.11	0.9999	73.13	1.00	1.38	0.9999	72.95	76.66	212.89	0.36
14	0.55	0.59	0.9999	92.13	0.68	0.74	0.9999	92.45	149.25	278.03	0.54
15	1.01	2.11	0.9996	47.68	1.25	2.62	0.9996	47.62	38.23	132.04	0.29
16	0.99	2.25	0.9991	44.12	1.23	2.79	0.9991	44.17	35.88	122.51	0.29
17	0.76	1.05	0.9991	72.33	0.95	1.13	0.9988	84.39	88.08	231.74	0.38
18	1.54	5.20	0.9920	29.55	1.90	6.46	0.9920	29.45	15.49	81.99	0.19
19	1.16	3.50	0.9990	33.06	1.44	4.33	0.9990	33.18	22.22	87.43	0.25
20	0.62	0.40	0.9988	153.64	0.76	0.50	0.9988	153.31	200.11	421.80	0.47
21	0.71	0.66	0.9992	107.54	0.88	0.82	0.9992	107.51	122.91	297.29	0.41
22	0.91	2.02	0.9994	45.37	1.13	2.50	0.9994	45.28	41.34	130.15	0.32
23	0.97	2.19	0.9993	44.41	1.20	2.72	0.9993	44.34	39.90	133.86	0.30
24	0.78	1.26	0.9994	61.99	0.96	1.56	0.9994	61.85	64.18	171.19	0.37

these parameters are much larger for SDS than CTAB, it is most likely that differences in the intensity of hydrophilic interactions between DHPs and the ionic surface of SDS and CTAB
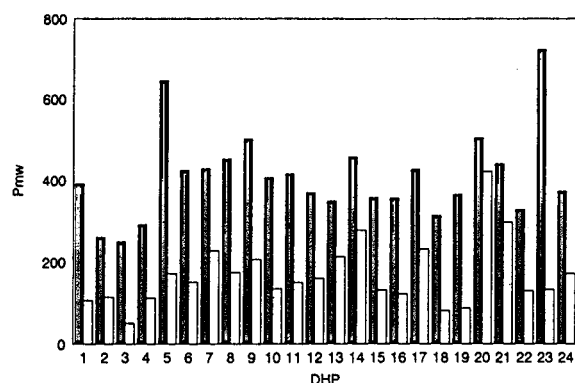
Fig. 3. P_{mw} coefficients of DHPs for either SDS (■) and CTAB (□).

micelles could be responsible for the higher affinity of DHPs to SDS than CTAB. This result is in accordance with the reported alkaline properties of this class of compounds [14], which could indicate the possibility that these DHPs will be positively charged under the experimental conditions used. This idea is also supported by the fact that SDS micelles, owing to their anionic surface, produce an increase in the local concentration of protons on their surface by a factor of ca. 100 [15], which can lead to an increase in the pK_a of a given compound of between 0.5 and 3.0 units [16]. For other compounds such as benzene and naphthalene derivatives under similar experimental conditions, larger solute–micelle association constants with CTAB have been reported [12]. This may be due to the diverse molecular structures of these derivatives and DHPs. In fact, benzene and naphthalene deriva-

tives have π-electrons which can interact with the positive surface of CTAB micelles.

As shown in Tables 1 and 2 and Fig. 3, the solute–micelle interaction is much larger for SDS than for CTAB. However, the P_{mw} coefficients with CTAB are also large. This could be mainly due to (i) the high hydrophobicity of these compounds (e.g., log P_{ow} ranges from 2.8 to 5.0 for the DHPs studied in this work) and (ii) the π-electrons which almost all of these DHPs possess in their phenyl ring which could interact with the positive surface of CTAB micelles.

P_{mw} coefficients for either SDS and CTAB, with the exception of three cases, possess values larger than 100 (Tables 1 and 2 and Fig. 3). This indicates the possibility that DHPs will be present in the mobile phase mainly in the micellar pseudo-phase, that is, their chromatographic behaviour could be explained by the direct transfer theory proposed by Borgerding et al. [10]. To demonstrate this idea, the experimental k' values and those obtained by this theory (Eq. 3) need to be the same. We calculated for every DHP and chromatographic condition used the k' values expected by Eq. 3. Fig. 4 shows a representative group of data for DHPs with small, intermediate and large P_{mw} coefficients, and the general pattern is similar in all instances. As shown by Fig. 4, when the surfactant concentration is increased the k' values of the direct transfer theory tend to match the experimental values. Therefore, there is a change from a three partition equilibria mechanism to direct transfer of DHPs from the micellar pseudo-phase to the stationary phase when the surfactant concentration in the bulk mobile phase is increased. This could be explained through the increase in the micellar pseudo-phase volume which, owing to the large P_{mw} coefficients of these compounds (Tables 1 and 2 and Fig. 3), causes an increase in the micellar–water solute molecules fraction. Table 3 shows an estimate of the ratio of molecules between the micellar and aqueous phases for the compounds shown in Fig. 4. These data were obtained from the P_{mw} coefficients and the micellar and aqueous phase volumes.

Fig. 5 shows the DHP P_{sw} coefficients when either SDS or CTAB is present in the mobile phase. The P_{sw} coefficients are slightly higher for SDS than for CTAB. This difference may be evidence of the surfactant adsorption on the stationary phase, at least for one of the two surfactants, otherwise the P_{sw} coefficients of these compounds should be independent of the nature of the surfactant used in the mobile phase. Assuming surfactant adsorption on the stationary phase, as the affinity of DHPs for SDS micelles is higher than for CTAB micelles (Fig. 3), the affinity towards the SDS-modified stationary phase is expected also to be higher than for CTAB, as demonstrated by Fig. 5. The magnitude and difference between the P_{sw} coefficients for both surfactants (Fig. 5) are much lower than for the P_{mw} coefficients (Fig. 3). Therefore, it is likely that the participation of the surfactant molecules adsorbed in the solute retention will be small and then DHPs are retained primarily through interaction with the C_{18} chains.

A linear relationship between P_{mw} and P_{sw} coefficients is clearly shown in Fig. 6 for CTAB but not for SDS. This result could indicate interaction of the solute with the micellar and the surfactant-modified stationary phase similar in nature for CTAB and different for SDS. Assuming that surfactant adsorbed-molecules on the stationary phase for either SDS or CTAB probably participate to a small extent in the retention mechanism (Fig. 5) and solute–stationary phase interaction occurs primarily through the C_{18} chains, the nature of the solute–micelle interaction could be mainly controlled by hydrophobic forces for CTAB. The worse correlation between P_{mw} and P_{sw} coefficients for SDS could be explained by considering the above-mentioned participation of hydrophilic interactions on the DHP–micelle association complex.

3.2. Selectivity coefficients

The selectivity coefficients were calculated for each pair of compounds. Fig. 7 shows a representative group of selected data among those compounds with small, intermediate and large P_{mw} coefficients. The general pattern observed in almost all instances is that the selectivity co-
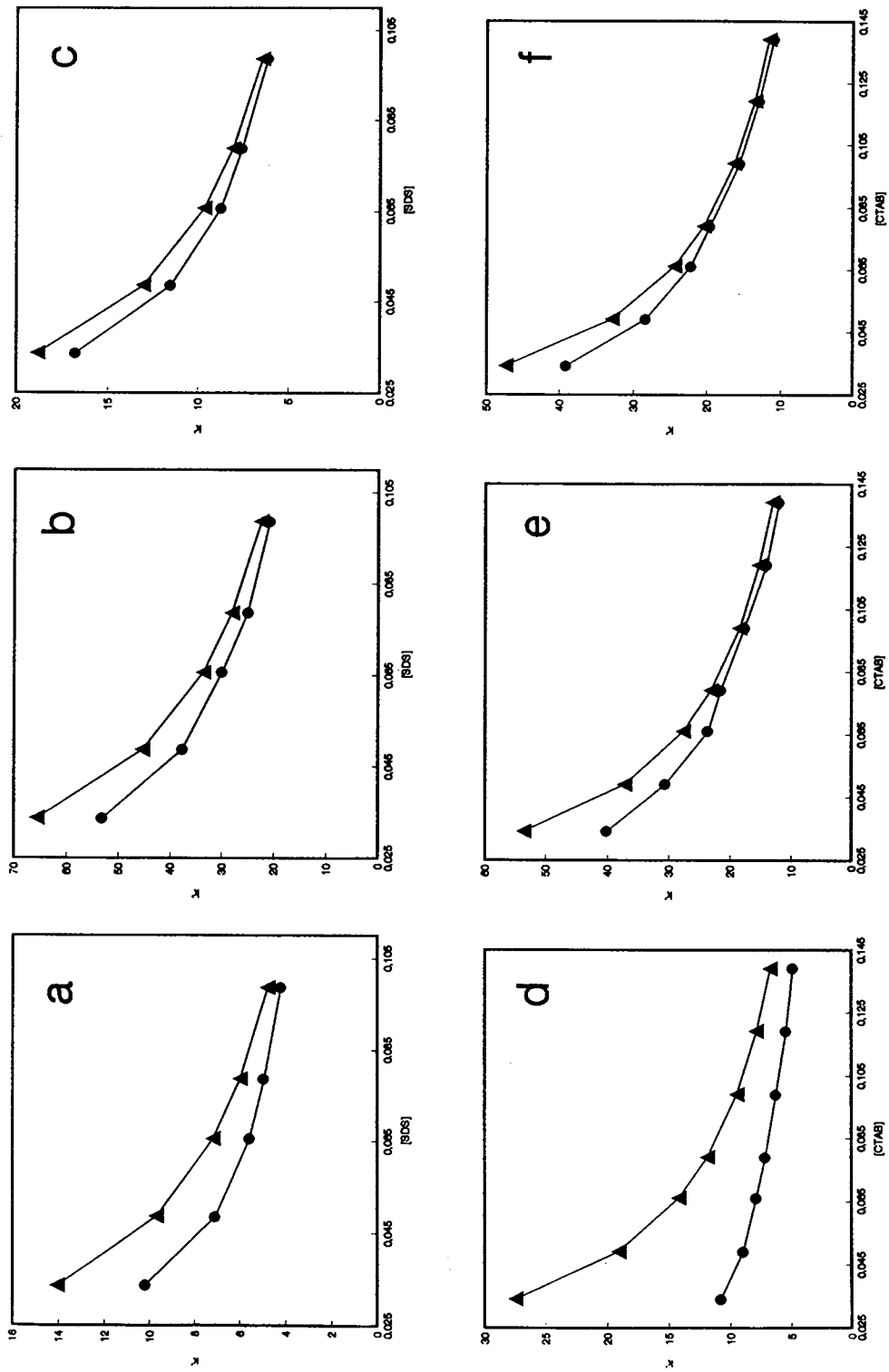
Fig. 4. ($\bullet$) Experimental and ($\blacktriangle$) theoretical capacity factors (k') of three representative groups of DHPs with small, intermediate and large P_{mw} coefficients as a function of surfactant concentration (M) in the mobile phase for (a–c) SDS and (d–f) CTAB. The compounds represented are (a and d) 3, (b and e) 14, (c) 23 and (f) 20.

Table 3
Micellar–water ratio of solute molecules for conditions used in Fig. 4.

		DHP 3	DHP 14	DHP 23	DHP 20
[SDS] (M)	0.035	2.1	3.8	6.0	
	0.050	3.0	5.5	8.8	
	0.067	4.1	7.5	11.9	
	0.080	4.9	9.0	14.3	
	0.100	6.2	11.4	18.0	
[CTAB] (M)	0.035	0.6	3.5		5.3
	0.050	0.9	5.1		7.7
	0.067	1.2	6.8		10.4
	0.080	1.5	8.2		12.5
	0.100	1.9	10.4		15.8
	0.120	2.3	12.6		19.1
	0.140	2.7	14.8		22.5

efficients decrease, tending to the P_{sm} coefficient ratio of solutes, when the surfactant concentration in the mobile phase increases. There are only a few cases for which the tendency is like that shown in Fig. 7a and d: the selectivity coefficients increase when the surfactant concentration increases but also tending to the P_{sm} coefficient ratio of solutes. These results can be explained by Eq. 3 based on the direct transfer of the solute from the micellar phase to the stationary phase postulated above when the
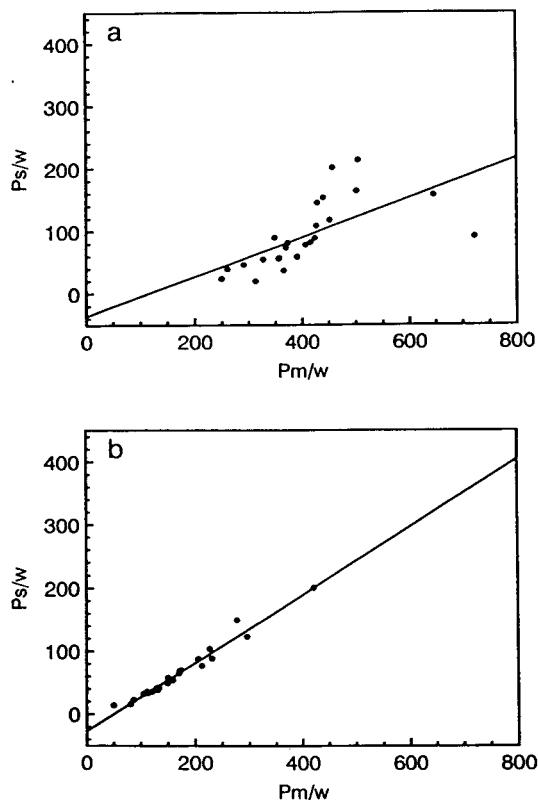
Fig. 6. Correlation between P_{sw} and P_{mw} coefficients of DHPs for (a) SDS and (b) CTAB.
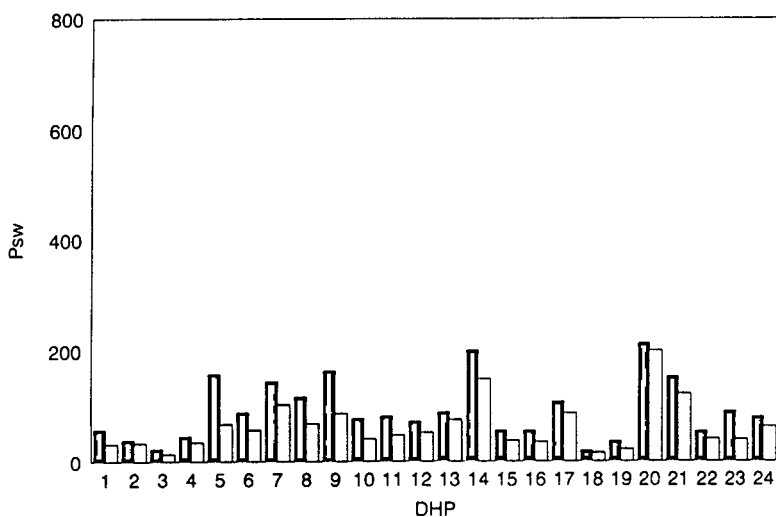
Fig. 5. P_{sw} coefficients of DHPs when SDS (■) or CTAB (□) is used as surfactant modifier in the mobile phase.
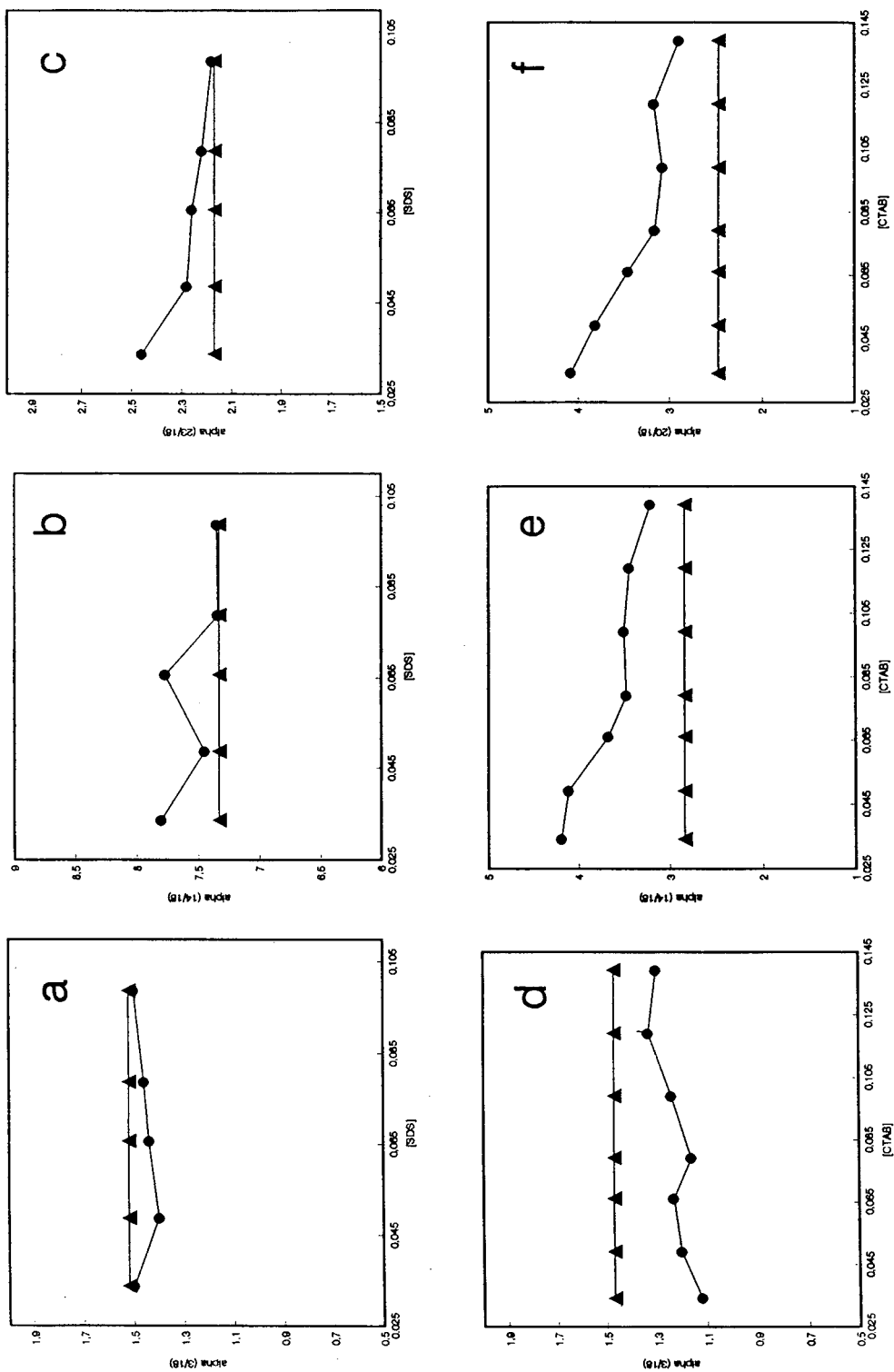
Fig. 7. Selectivity coefficients (●) for some representative pairs of DHPs with small, intermediate and large P_{mw} coefficients versus surfactant concentration (M) in the mobile phase for SDS and CTAB. Horizontal straight lines (▲) represent the ratio of P_{sm} coefficients for each pair for SDS and CTAB.

concentration of surfactant in the mobile phase is high. If Eq. 3 is valid for these compounds when the surfactant concentration is high, then the selectivity should tend to the ratio of P_{sm} coefficients, as demonstrated in Fig. 7.

Developing a mathematical expression for these compounds derived from the three partition equilibria theory, it is possible to explain the tendency of selectivity coefficients to the P_{sm} coefficient ratio when the concentration of surfactant increases:

$$k' = \frac{q_s}{q_{aq} + q_M}$$
$$= \frac{V_s[XL]}{V_m(1 - vC_M)[X] + V_m vC_M[XM]} \quad (4)$$

where q_s, q_{aq} and q_M are the amounts of solute in the stationary, aqueous and micellar phases, respectively, and [XL], [X] and [XM] are the concentrations of the solute in the stationary, aqueous and micellar phases, respectively. Dividing by [XM]:

$$k' = \frac{V_s P_{sm}}{V_m(1 - vC_M)(1/P_{mw}) + V_m vC_M}$$
$$= \frac{V_s}{V_m} \cdot \frac{P_{sm}}{vC_M(1 - 1/P_{mw}) + 1/P_{mw}} \quad (5)$$

The selectivity coefficient for a given pair of compounds (e.g., compounds **1** and **2**) will be defined by

$$\alpha = \frac{P_{sm1}[vC_M(1 - 1/P_{mw2}) + 1/P_{mw2}]}{P_{sm2}[vC_M(1 - 1/P_{mw1}) + 1/P_{mw1}]} \quad (6)$$

Eq. 6 shows that if $P_{mw1} > P_{mw2}$ then α will decrease when the surfactant concentration (e.g., C_M) is increased, otherwise α will decrease. This explains the cases for which α increases with C_M. As the P_{mw} coefficients for these compounds are close to or larger than 100 for either SDS or CTAB (Tables 1 and 2), it is possible to consider $(1 - 1/P_{mw}) \approx 1$. Substituting in Eq. 6:

$$\alpha = \frac{P_{sm1}}{P_{sm2}} \cdot \frac{vC_M + 1/P_{mw2}}{vC_M + 1/P_{mw1}} \quad (7)$$

When C_M is increased, vC_M could be large enough to make $(vC_M + 1/P_{mw}) \approx vC_M$, that is

$$\alpha = \frac{P_{sm1}}{P_{sm2}} \quad (8)$$

Therefore, as expected by theory, the results obtained in this work demonstrate that when the surfactant concentration in the mobile phase is increased, the selectivity coefficients tend to a constant value which corresponds to the P_{sm} coefficient ratio of the solutes. This tendency is due to a change in the retention mechanism from a three partition equilibria mechanism to a direct transfer mechanism when the concentration of surfactant in the mobile phase is increased. The general pattern observed in the literature for selectivity versus surfactant concentration in the mobile phase is a decrease in selectivity when the surfactant concentration increases [5,7]. Other workers have observed an increase in selectivity versus surfactant concentration in the mobile phase [7]. However, in all instances, selectivity tends to a limiting value which can be explained by the above-mentioned change to a direct transfer mechanism in the solute retention when the surfactant concentration in the mobile phase increases.

The selectivity coefficients calculated in this work are higher for SDS than for CTAB for every surfactant concentration studied. This result is in good agreement with the hypothesis that solutes can be located in micelles in different microenvironments of different polarity. The selectivity will be larger if the difference in polarity between the mobile phase and stationary phase environment occupied by solutes is also larger [17]. As discussed previously, Fig. 6 could be showing a similar nature for solute–micelle and –stationary phase interaction for CTAB, but not for SDS, which may be the reason for the higher selectivity of SDS.

Finally, the selectivity results are similar to those obtained previously for benzene and naphthalene derivatives under the same chromatographic conditions, for which selectivity is also higher for SDS than for CTAB and improves when the surfactant concentration in the mobile phase decreases [12].

4. Conclusions

A direct transfer from micelles to the stationary phase is presented to explain the retention behaviour and selectivity coefficients for a group of DHPs in MLC when surfactant concentration in the mobile phase increases.

A primarily hydrophobic interaction of these DHPs with CTAB and a hydrophilic interaction with SDS is suggested according to the correlation between P_{mw} and P_{sw} coefficients. The selectivity results are also in good agreement with this hypothesis, as DHPs show larger selectivity coefficients for SDS than for CTAB, which may be due to some differences in the polarity of the microenvironmental of DHPs in SDS and CTAB micelles.

Acknowledgements

References

[1] D.W. Armstrong and S.J. Henry, *J. Liq. Chromatogr.*, 3 (1980) 657.
[2] F.G.P. Mullins and G.F. Kirkbright, *Analyst*, 109 (1984) 1217.
[3] G.F. Kirkbright and F.G.P. Mullins, *Analyst*, 109 (1984) 493.
[4] P. Yarmchuk, R. Weinberger, R.F. Hirsch and L.J. Cline Love, *Anal. Chem.*, 54 (1982) 2233.
[5] M.A. García, S. Vera, M. Bombín and M.L. Marina, *J. Chromatogr.*, 646 (1993) 297.
[6] J.K. Strasters, E.D. Breyer, A.H. Rodgers and M.G. Khaledi, *J. Chromatogr.*, 511 (1990) 17.
[7] M.G. Khaledi, J.K. Strasters, A.H. Rodgers and E.D. Breyer, *Anal. Chem.*, 62 (1990) 130.
[8] D.W. Armstrong and F. Nome, *Anal. Chem.*, 53 (1981) 1662.
[9] M. Arunyanart and L.J. Cline Love, *Anal. Chem.*, 56 (1984) 1557.
[10] M.F. Borgerding, F.H. Quina, W.L. Hinze, J. Bowermaster and H.M. McNair, *Anal. Chem.*, 60 (1988) 2520.
[11] M.L. Marina, S. Vera and A.R. Rodríguez, *Chromatographia*, 28 (1989) 379.
[12] M.A. García, S. Vera and M.L. Marina, *Chromatographia*, 32 (1991) 148.
[13] E. Pramauro and E. Pelizzetti, *Anal. Chim. Acta*, 154 (1983) 153.
[14] P. Mageney, R. Gosmini, S. Raussou, M. Commercon and A. Alexakis, *J. Org. Chem.* 59 (1994) 1877.
[15] I.A. Pastre and M.G. Neumann, *J. Photochem. Photobiol. A*, 79 (1994) 1.
[16] G.S. Hartley, *Trans. Faraday Soc.*, 30 (1934) 444.
[17] M.G. Khaledi, *Anal. Chem.*, 60 (1988) 876.

Recycling in preparative liquid chromatography

Frédéric Charton[a,b,c,1], Michel Bailly[a], Georges Guiochon[b,c,*]

[a]*Laboratoire des Sciences du Génie Chimique, 1 Rue Grandville, BP 451, 54001 Nancy Cedex, France*
[b]*Department of Chemistry, University of Tennessee, Knoxville, TN 37996-1600, USA*
[c]*Division of Analytical Chemistry, Oak Ridge National Laboratory, Oak Ridge, TN 37881-6120, USA*

Abstract

A comparison is made between the performance achieved in preparative chromatography when using elution, the most usual operating mode of this technique, and the different implementations of recycling. This work is based on the experimental results obtained in the separation of the enantiomers of ketoprofen on a cellulose-based stationary phase, using the various chromatographic modes. It uses also the modeling of non-linear chromatography to compare the different operating modes considered. Theoretical and experimental results are presented and compared. In some cases, it is possible to achieve simultaneously both an increase in the production rate and a decrease in the amount of eluent needed.

1. Introduction

High-performance preparative liquid chromatography is becoming an increasingly popular industrial separation process [1]. It uses fine particle stationary phases, typically in the 10–30 μm range, thus ensuring faster mass transfers than in classical adsorption processes. Although high column efficiency is generally achieved, the production rate is limited by the relatively low capacity of the stationary phases, so the preparative chromatographic process remains expensive and requires careful optimization in order to minimize production costs. In particular, it has long been recognized that preparative columns must be operated under strongly overloaded conditions, in a concentration range where adsorption isotherms are no longer linear.

A number of modifications of the conventional elution mode have been suggested to enhance further the performance of the process [2–4]. Some, such as simulated moving bed [4], are complex and require dedicated equipment. Others are simple and can be implemented with minor changes of the equipment used in elution. Bailly and Tondeur [5–7] proposed different implementations of recycling, taking advantage of non-linear effects arising in chromatography at high concentrations. Using the ideal model of chromatography, they studied the application of these different modes to the separation of binary mixtures by ion-exchange chromatography and showed the potential advantages. Extension of this work to other modes of chromatography is straightforward. However, they did not consider

---

* Corresponding author. Address for correspondence: Department of Chemistry, University of Tennessee, Knoxville, TN 37996-1600, USA.
[1] Present address: Separex, 54240 Champigneulles, France.

the influence of a finite column efficiency, and there are no studies comparing the performance of the various implementations of recycling. It must be pointed out at this stage that these methods are not related to the conventional "peak shaving" technique, developed by analytical chromatographers. In this method the eluent is continuously recycled from the exit back to the column inlet as long as its composition does not meet one of the criteria for fraction collection. In the present case, the eluent containing the material to be recycled is collected, stored until the end of the cycle and injected at the beginning of the new cycle, before the complement of fresh feed or mixed with it. A steady state is achieved in which the amount of material in the column, the chromatogram and the production are all the same for each cycle, whereas in recycling with peak shaving an injection of pure fresh feed is done only every so many cycles, after the amount of material has been reduced to none by shaving at the end of each cycle [3].

We have undertaken a comparative investigation of the eluent consumption, the production rate and the recovery yield at a given purity of conventional elution and of the recycling procedures [5–7]. This study takes into account the finite rate of mass transfer kinetics in the column. It involves numerical calculations and experimental determinations, all made for the same separation problem. As a case in point, we chose the separation of the enantiomers of ketoprofen. To perform this separation we used a chiral stationary phase (CSP) prepared by adsorption of cellulose tris(4-methylbenzoate) on a macroporous silica support [8], and a mixture of n-hexane, 2-propanol and acetic acid as the mobile phase. Among CSPs, adsorbed cellulose esters are characterized by a high loading capacity and a fast rate of mass transfer kinetics, and hence a high column efficiency [9].

This separation was chosen because enantiomer purification is a problem of current concern in the pharmaceutical industry. Further, these binary separations permit simple calculations while giving relevant results of practical importance [10].

2. Theory

2.1. Description of the chromatographic modes studied

In the elution mode, the simultaneous achievement of a large recovery yield and a high product purity requires a significant degree of resolution between the two bands at the column outlet. This limits seriously the amount of feed which can be injected in each cycle. If the feed amount is increased beyond this limit to obtain a higher production rate, a mixed zone of growing importance appears in the chromatogram. To achieve both the high recovery and the purity required, this fraction must be recycled. Different implementations are possible.

Besides the classical overloaded elution mode (Fig. 1), we have studied three different recycling procedures, the recycling of an intermediate, mixed zone (Fig. 2), the recycling of the dilute tail of the second band (Fig. 3) and a combination of these two methods (Fig. 4). Recycling must be combined with sample injection. Fresh sample may be added to the recycled fraction at the end of every cycle, or at the end of every second cycle, or never, the next sample being injected only when all the amount injected has been purified. This last approach has been discussed previously [3]. In this work, we assume that some amount of fresh sample is added to the recyclate at the beginning of each new cycle. In this case, operation under cyclic steady state conditions is required in order to achieve the potential performance of the method.

Elution chromatography
Elution (Fig. 1) is still the most usual operating mode in preparative chromatography. The
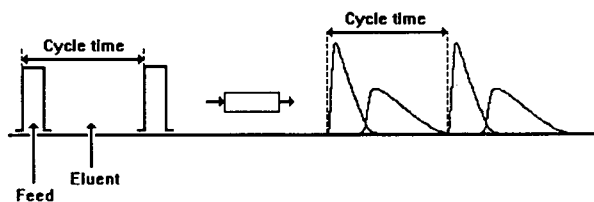
Fig. 1. Schematic representation of elution chromatography.

column is swept by a stream of eluent at constant flow-rate. A given amount of feed is injected periodically at its inlet. The cycle time or period between two successive injections is chosen in order to avoid delay between successive chromatograms. In the separation of a binary mixture, the goal is the production of one component with a minimum degree of purity. Switching valves at appropriate times begins and stops the collection of the useful component. These "cut times" are chosen in order to achieve the required purity.

The performance of the process can be evaluated using one of several criteria: the production rate, the amount of stationary phase immobilized, the amount of solvent needed for the purification of a unit amount of component and the concentration of the collected product. Our purpose is to study operating modes that could successfully replace elution chromatography. In this work, we deal with two operating modes which have already been introduced in the literature [5,7], and with an original combination of them. In all instances, the implementation requires a single column.

Recycling of an intermediate cut

In "recycling with mixing" (Fig. 2A), the non-separated, intermediate zone of the chromatogram is collected during any given cycle, lumped
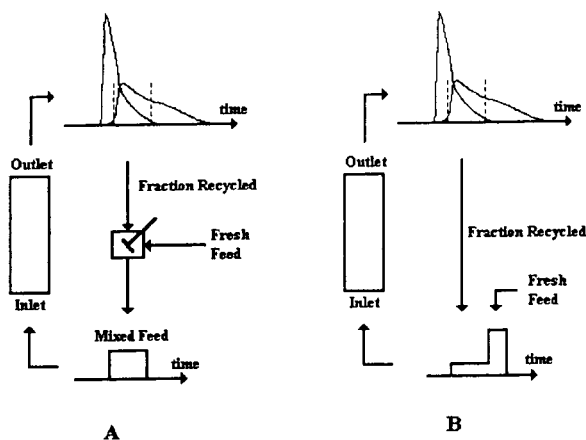
into one fraction and mixed with a given amount of the fresh feed solution. The resulting mixed feed (whose composition differs from that of the actual or fresh feed) is injected into the column as a pulse at the beginning of the next cycle. The cut times are determined in order to satisfy the product purity requirements.

Assuming an infinitely efficient column and a separation factor independent of the concentrations (e.g., with a competitive Langmuir isotherm), Bailly and Tondeur [5] have shown that this procedure could save eluent and give more concentrated products than the elution mode. They derived equations permitting the calculation of the optimum operating parameters of this process.

Because recycled fraction and fresh feed have different compositions, their mixing results in a loss of the separation work performed by the column. "Segmented recycling" (Fig. 2B) is similar to "recycling with mixing", but avoids this loss of separation. The mixed zone of the chromatogram is collected and lumped into one fraction, as in the recycling with mixing mode, but in the next cycle the recycled fraction is pumped into the column just before the new injection of fresh feed is performed [7].

Recycling of the second peak tail

In most instances, equilibrium isotherms are convex upwards, so each band exhibits a steep front, or shock layer, and a diffuse rear boundary [10]. Thus, the peak of the more retained component has a long tail, and the purified fraction containing this component is dilute. The phenomenon may still be aggravated by the consequence of the tag-along effect [10], which spreads the last component over a wide retention range. Rather than collect the highly dilute part of the fraction, it may be more advantageous to recycle it. The corresponding procedure is described in Fig. 3. The injection is made as in "segmented recycling", the recycled fraction being pumped into the column just ahead of the new fresh feed sample. Note that, because the tail is collected, stored and injected ahead of the fresh feed, the tail will not grow longer progressively but will stabilize after a few cycles.

Fig. 2. Schematic representations of (A) "recycling with mixing" and (B) "segmented recycling".
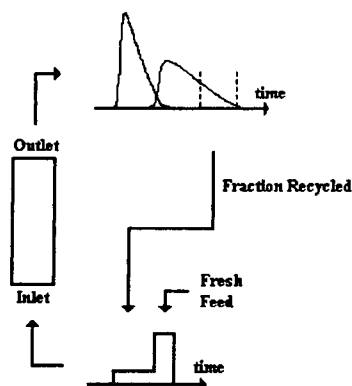
Fig. 3. Schematic representation of "recycling of the second peak tail".
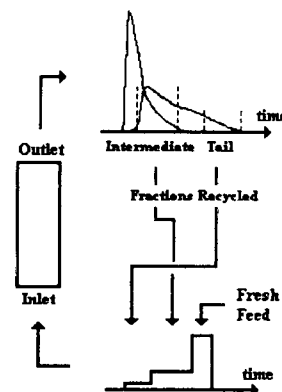
Fig. 4. Schematic representation of "combination of recyclings".

This is what distinguishes this method from the conventional "peak shaving" method [3]. The cut point for the recycling of this tail depends on the optimization criterion.

Combination of recycling modes

Because of the dispersive rear front, the production rate of the more retained component of a mixture is always lower than that of the less retained component in elution, and the collected fractions are less concentrated. The operating modes based on the sole recycling of the intermediate, mixed zone suffer from the same drawback. Hence, it seems attractive to combine the two recycling modes discussed above, as shown in Fig. 4. The injection is made in three steps. The recycled fraction containing the dilute end of the more retained component band is injected first, followed by the mixed, recycled fraction and, finally, a new amount of fresh feed.

2.2. Economic criteria and adjustable variables

Economic criteria were used to optimize the experimental conditions of the chromatographic procedures and compare their performances. The relationships between the cost components of the process and its technical parameters are investigated first.

Economic criteria

Optimizing an industrial process comes down to minimizing its production cost. Following the cost analyses given by Nicoud and Colin [1] and by Felinger and Guiochon [11], the components of the production cost of a chromatographic separation can be arranged into three main groups of contributions, the operating costs, OC, the cost of the crude material lost, CC, and the fixed costs, FC, with

$$PC = CC + OC + FC \qquad (1)$$

CC is proportional to the recovery yield, Y_i, of the desired component i:

$$Y_i = \frac{\text{mass of compound of interest recovered}}{\text{mass of compound of interest injected}}$$

$$(2)$$

To permit a meaningful comparison, the experimental conditions will be chosen so that very close recovery yields be reached with the various operating schemes. In this case, the cost contribution of the loss of crude material has no effect on the comparison of the various procedures.

OC includes mainly the cost of the solvent required, the energy spent and the packing material [1,2,11]. Regeneration of the eluent for its recycling is essential for the economics of the chromatographic process. Its cost, including the cost of the energy required, added to the cost of

fresh solvent, and of waste management, constitutes the essential of the solvent cost and the main part of OC in most instances [1]. Although often minor [1], the cost of the packing material must also be taken into account. Labor costs are sometimes included in OC [1], which is correct if the duration of a production campaign is inversely proportional to the production rate. Because labor costs are more controlled by safety regulations and possible union contracts than by the actual production rate of a given unit, they are part of the fixed costs for continuous productions [11].

The unit operating costs, or separation cost per unit amount produced, can be reasonably estimated using the following three parameters:

(i) the volume of eluent needed to produce a unit amount of the purified product;

(ii) the specific amount of stationary phase used, or volume of stationary phase immobilized in a column of unit cross-sectional area; and

(iii) the productivity or specific production rate of the desired component, i.e., the mass of this compound produced per unit time and unit column cross-section.

FC represents fixed costs linked to the equipment investments. It includes the capital cost and, at least in the case of short batch production, the labor costs. The additional fixed costs of recycling include only the purchase of a few additional valves and the addition to the control program of the few lines of code needed to operate them, so we can neglect it. As all the operating procedures considered here are performed using virtually the same equipment and the same column, the fixed costs are essentially the same for all the operating modes examined.

Hence the comparison between the cost of the different operation modes will be based on a comparison of the maximum production rates possible and of the volume of eluent required.

Adjustable variables

The performance of the chromatographic separation process, and hence the production cost, depend on a number of parameters that can be adjusted prior to a new campaign. These parameters are:

(i) the concentration of the fresh feed injected, C_F; since in the present case we study the purification of either enantiomers from their racemic mixture, $C_{1F} = C_{2F}$;

(ii) V_F, the volume of fresh feed injected at the beginning of each cycle;

(iii) u_0, the mobile phase velocity;

(iv) L, the column length; because of the availability of dynamic axial compression columns, the column length can be considered as adjustable, at least within a certain range;

(v) for all the recycling modes, the composition and the volume of the recycled fractions are important intermediate parameters. They depend on the previous parameters and on the required values of the yield and purity, but are conveniently treated as variables.

Because the aim of this study is a comparison between the performance of the various operating modes, not their individual optimization, we do not need to optimize separately L and u_0. For each efficiency, the injection conditions (fresh feed and recycling) were optimized. This is done as followed. The composition of the fresh feed and the column efficiency are given, in addition to the purity and recovery yield required for the component of interest. The parameters that must be optimized are the volume of fresh feed injected and the characteristics of the recycled fraction. The optimum values of these parameters can be calculated easily in the case of an infinitely efficient column, using the analytical equations derived by Bailly and Tondeur [5–7] for the ideal model. Starting with this initial guess, numerical solutions of the equilibrium-dispersive model [10] permit the rapid determination of the optimum conditions for any column efficiency. The number of combinations of L and u_0 that permit one to obtain a certain value of N_P with a given column is infinite, but it has been shown that, for a given value of N_P, the production rate does not depend much on the individual choice of L and u_0 [11]. Further, the values of these parameters do not influence the ratios of production rates and of eluent consumptions in elution and in one of the recycling modes. Another advantage of this choice is that, for a given column, the amount of stationary

phase immobilized is the same for all the operating modes, and this criterion disappears from the comparison.

2.3. Modeling of non-linear high-performance liquid chromatography

In high-performance liquid chromatography, the columns have a finite but high efficiency. This fact makes the use of the equilibrium-dispersive model of chromatography particularly attractive [10]. In this model, the differential mass balance equation for one component in a chromatographic column is written as

$$\frac{\partial C_i}{\partial t} + \frac{1 - \varepsilon_T}{\varepsilon_T} \frac{\partial q_i}{\partial t} + u_0 \cdot \frac{\partial C_i}{\partial x} = D_{i,\mathrm{ap}} \cdot \frac{\partial^2 C_i}{\partial x^2} \qquad (3)$$

where C and q are the concentrations in the mobile and the stationary phases at equilibrium, respectively, ε_T is the total porosity of the bed, and is derived from the retention time, t_0, of a non-retained component whose propagation velocity is u_0, t_0 is also called the column holdup time and $D_{i,\mathrm{ap}}$ is the apparent dispersion coefficient.

In the equilibrium-dispersive model, constant equilibrium between the mobile and stationary phases is assumed, and q is related to C by the equilibrium isotherm. The finite column efficiency due to axial dispersion and to the finite rate of the mass transfer kinetics is taken into account by the use of the lumped apparent dispersion coefficient, D_{ap}. We assume D_{ap} to be constant and equal to its value under linear conditions. Thus, D_{ap} is related to the column length, L, and to the number of theoretical plates, N_P, by the equation

$$D_{\mathrm{ap}} = \frac{u_0 L}{2 N_P} \qquad (4)$$

For a multi-component system, we must write as many Eqs. 3 as there are components. In this case, however, the different components of the mixture compete for access to the adsorption sites on the stationary phase. Thus, the amount q_i of component i adsorbed at equilibrium depends not only on C_i, but also on the concentrations of all the other components in the mobile phase, through the competitive adsorption isotherm:

$$q_i = q_i(C_1, C_2, \ldots) \qquad (5)$$

The determination of the competitive adsorption isotherms of the system components is the keystone of our modeling effort. Different approaches and equations to evaluate competitive isotherms, combined with the use of the equilibrium-dispersive model, have already been used successfully to calculate band profiles [10–15].

Finally, the solution of the system of partial differential equations of the problem requires proper initial and boundary conditions. The initial conditions are identical for all the operating schemes studied: the column is initially empty. As the column efficiency is usually high, the boundary conditions can be simplified by neglecting the dispersion effect [3,10,11], so they are reduced to the injection conditions at the column inlet.

Only numerical solutions of the equilibrium-dispersive model of chromatography can be calculated. The algorithms available have been reviewed recently [10,16]. For most practical applications, the forward–backward finite difference method proposed by Rouchon et al. [16] is the fastest and most efficient algorithm.

3. Experimental

3.1. Equipment

The experiments were performed using an HP1090 liquid chromatograph (Hewlett-Packard, Palo Alto, CA, USA), equipped with a multi-solvent delivery system, an automatic sample injector with a 250-μl loop, a diode-array UV detector and a computer data acquisition system. Acquired data were downloaded to one of the computers at the University of Tennessee Computer Center. Also, a Gilson (Middleton, WI, USA) Model 203 fraction collector was used to complement the HP system.

3.2. Materials

Column

A 25 cm × 0.46 cm I.D. Chiralcel OJ column (Daicel, Tokyo, Japan) was used (average particle size 10 μm). This stationary phase was cellulose tris(4-methylbenzoate) adsorbed on macroporous silica [8]. The total column porosity ($\varepsilon_T = 0.674$) was derived from the retention time of the solvent peak. The use of this column was restricted by two independent constraints imposed by the manufacturer, allowing a maximum flow-rate of 1.5 ml/min and a maximum inlet pressure of 50 bar.

Mobile phase and chemicals

In all chromatographic experiments, the mobile phase was *n*-hexane–2-propanol–acetic acid (90:10:0.5, v/v/v). Hexane and 2-propanol were purchased from Burdick and Jackson (Muskegon, MI, USA) and acetic acid from Mallinckrodt (Paris, KY, USA). The *S*-(+)-enantiomer (purity 99%) and a racemic mixture (purity >99%) of ketoprofen were obtained from Rhône-Poulenc (Centre de Recherches de Vitry, Vitry, France). All these compounds were used as received.

3.3. Procedures

All the experiments were performed at 30°C. We determined the adsorption data for the pure *S*-(+)-enantiomer and the competitive adsorption data for different mixtures including the racemic mixture. The column efficiency was also measured at different mobile phase flow-rates. Very small sample amounts (2 μg) were injected. The column efficiency was derived from the width of the peaks at half-height. It was almost the same for the two components.

Determination of adsorption data

The single-component adsorption data were determined using the classical frontal analysis technique [17]. The experiments were performed at a flow-rate of 0.8 ml/min. Data were acquired in two concentration ranges. Twenty data points were measured between 0.2 and 4.3 g/l with detection at 365 nm and ten data points between 0.04 and 0.4 g/l with detection at 295 nm. At these two wavelengths, the detector response was found to be linear in the corresponding concentration range.

Competitive adsorption data were measured using the binary frontal analysis method [18]. In all experiments, the initial concentration of the enantiomers in the column was zero. Hence the concentration of the more retained component in the intermediate plateau was also zero and there was no need to analyze the corresponding eluate. This procedure greatly simplifies the experiments and improves the accuracy of the results. The concentration of the pure less retained component on this intermediate plateau was derived from the calibration graph at 370 nm, at which wavelength the detector response was linear over the whole concentration range studied.

The retention volumes and the corresponding concentrations were inserted into the equation given by Jacobson et al. [18] to determine the amounts adsorbed at equilibrium. The measurements of adsorption data could be made for values of the total concentration up to ca. 5.0 g/l. At higher concentrations, the intermediate plateau, where the less retained component was alone, disappeared.

Determination of elution profiles

Because the HP1090 chromatograph is not equipped with a large enough sample loop, the injections were made by programming the multisolvent delivery system. The individual elution profiles in the mixed band region were determined by collecting fractions of the eluent and reinjecting them on to the same column under analytical conditions. The time of collection for each fraction was chosen according to the flow-rate so that the fraction volume was ca. 80 μl. As the UV spectrum is achiral, the relative concentration of the two enantiomers in a fraction is equal to the peak-area ratio. This permits the rapid determination of the individual concentration profiles.

4. Results and discussion

4.1. Modeling of the separation of the enantiomers of ketoprofen on a cellulose-based stationary phase

We need to find a suitable model for the competitive isotherm data and a proper correlation for the dependence of the column efficiency on the mobile phase flow-rate. These models must then be validated.

Modeling of the competitive equilibrium isotherm

Typically, one of the two enantiomers was not available as a pure compound. Hence it was not possible to determine the competitive isotherm of the system following the classical procedure [12–15,19]. Only the pure S-(+)-enantiomer, which is also the more retained, and the racemic mixture were available, allowing the determination of the single-component adsorption data for the S-(+)-enantiomer, and the multi-component adsorption data for 1:1 and 1:3 [$R(-)$/$S(+)$] mixtures of R-(−)- and S-(+)-enantiomers. These experimental data were fitted together to obtain the best possible equilibrium isotherm.

The Scatchard plot of q/C vs. q of the experimental adsorption data for the S-(+)-enantiomer is not linear, which suggests that the classical Langmuir model cannot account for these data. Previous studies [12,13,15,19] have shown that on the surface of enantioselective phases two types of sites co-exist, enantioselective and non-selective sites. The former adsorb more strongly, but have a lower saturation capacity than the latter. A two-site adsorption model was successfully adopted, as for other enantiomers on the same stationary phase [19].

Accordingly, we fitted the experimental data to a bi-Langmuir isotherm. The first term represents the contribution of the chiral recognition mechanism and the second term accounts for the non-selective interactions. Previous, independent investigations of the chiral recognition process on cellulose-based CSPs have also shown the existence of two types of retention mechanism [20,21].

In compliance with the assumption that no chiral selectivity is involved in the second term of a bi-Langmuir isotherm, the coefficients A and B (Table 1) must be the same for both enantiomers, which reduces to six the total number of parameters of the two isotherms. We do not know at this stage the nature of the chiral retention mechanism, and thermodynamics cannot identify it. However, by analogy with previous results, we further assume that the column saturation capacities of the two enantiomers on the enantioselective sites are equal. Hence, $a_{(-)}/b_{(-)} = a_{(+)}/b_{(+)}$ (Table 1), which further reduces to five the number of parameters and ensures the thermodynamic consistency of the competitive isotherm model by satisfying the Gibbs–Duhem equation [22]. In spite of all these satisfactory properties, however, the isotherm model must be considered as empirical.

Five different sets of experimental data, one set for the pure S-(+)-enantiomer, one set each with the 1:1 mixture and the 1:3 mixture, for both enantiomers were handled together. We used a non-linear regression program based on a simplex algorithm to minimize the following objective function [19]:

Table 1
Isotherm parameters

Type of sites	Isomer	$a_{(*)}$ or A	$b_{(*)}$ or B (l/g)	q_s (g/l)
Selective	R-(−)-	4.40	0.202	21.7
Selective	S-(+)-	6.79	0.312	21.7
Non-selective	R-(−)- and S-(+)-	2.89	0.019	152.1

No. of parameters = 5. Model of competitive isotherm:

$$q_{(*)} = \frac{a_{(*)}C_{(*)}}{1 + b_{(-)}C_{(-)} + b_{(+)}C_{(+)}} + \frac{AC_{(*)}}{1 + B[C_{(-)} + C_{(+)}]}$$

where $* = -$ or $+$. The coefficients $a_{(+)}$, $a_{(-)}$, $b_{(+)}$, and $b_{(-)}$ are the coefficients of the Langmuir adsorption isotherm of the (+) and (−) isomers, respectively, on the chiral selective site; the coefficients A and B are the coefficients of the Langmuir adsorption isotherm of either isomer on the nonselective sites.

$$\delta = \delta_{(-)} + \delta_{(+)} = \sqrt{\frac{1}{N_{(-)}} \sum_1^{N_{(-)}} \left(\frac{q_{\exp_i} - q_{\mathrm{cal}_i}}{q_{\exp_i}} \right)^2_{(-)}}$$

$$+ \sqrt{\frac{1}{N_{(+)}} \sum_1^{N_{(+)}} \left(\frac{q_{\exp_j} - q_{\mathrm{cal}_j}}{q_{\exp_j}} \right)^2_{(-)}}$$

$$(6)$$

where $q_{\exp}$ and q_{cal} are the experimental amount adsorbed and the value derived from the isotherm (Table 1), respectively, both functions of $C_{(-)}$ and $C_{(+)}$. The S-$(+)$-enantiomer adsorption data were only involved in $\delta_{(+)}$ and used with $C_{(-)} = 0$. The best values obtained for the isotherm parameters are summarized in Table 1. A comparison of experimental and calculated adsorption data is made in Fig. 5, demonstrating an excellent fit of the model to the experimental data. The average difference between the calculated and experimental adsorption data was less than 1% for both components. The initial slopes of the isotherm were in very good agreement with the retention times derived from chromatograms obtained under linear conditions, when very small amounts were injected.

Efficiency of the column

The column efficiency was measured for different values of the mobile phase flow velocity, between 0.03 and 0.20 cm/s, corresponding to flow-rates between 0.2 and 1.4 ml/min for a 0.46 cm I.D. column. Very small sample amounts (2 μg) were injected. The column efficiency was derived from the width of the peaks at half-height. Fig. 6 shows a plot of the experimental values of the height equivalent to a theoretical plate (HETP) (symbols) versus the flow velocity, u_0. The range of flow-rates studied was limited because of the constraint of a maximum of 1.5 ml/min given by the column manufacturer, independent of the maximum inlet pressure equal to 50 bar [with ΔP (bar) $= 8.37\,L$ (cm) u_0 (cm/s) $\leq$ 50 bar].

The column efficiency was almost the same for the two components. The difference observed at the highest velocity is lower than the experimental inaccuracy. In the velocity domain studied, a
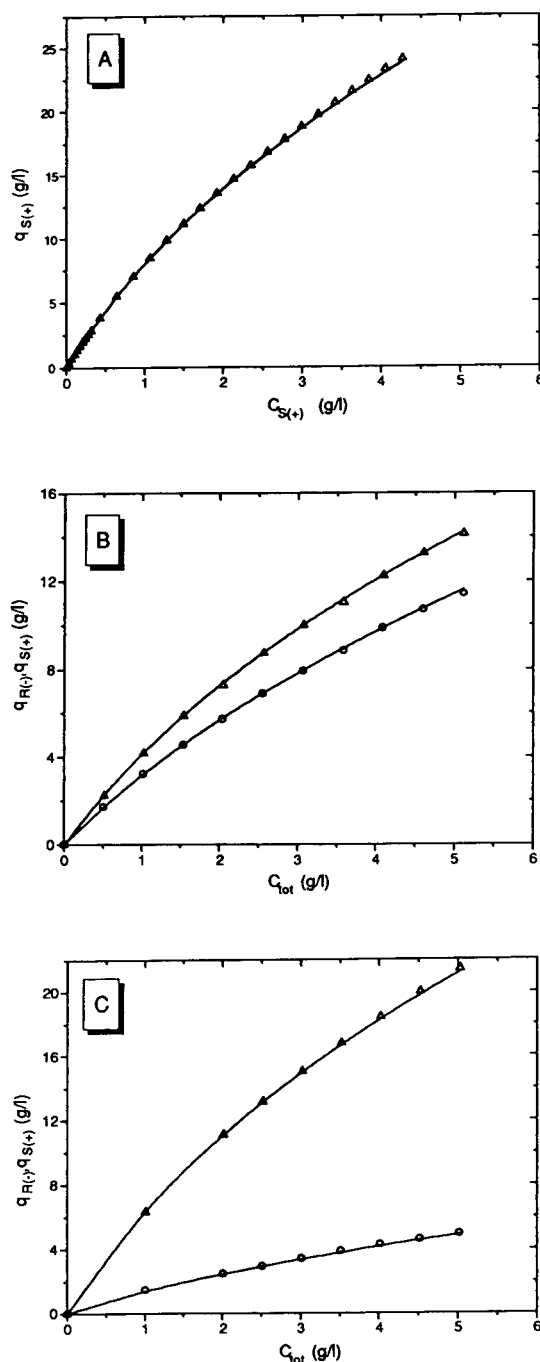
Fig. 5. Experimental adsorption data (symbols) for the ketoprofen on a Chiralcel OJ stationary phase at 30°C. $\bigcirc = R$-$(-)$-enantiomer; $\triangle = S$-$(+)$-enantiomer. (A) S-$(+)$-alone; (B) $C_{(-)} = C_{(+)}$; (C) $C_{(+)} \approx 3.5 C_{(-)}$.
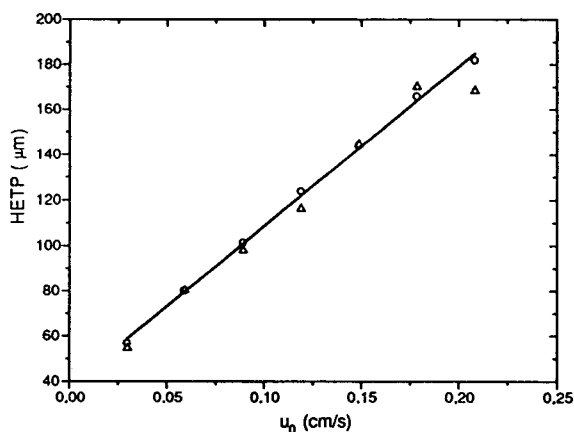
Fig. 6. Experimental [$\bigcirc$ = R-(-)- and $\triangle$ = S-(+)-enantiomer] and calculated (line) according to Eq. 8 plate heights as a function of the velocity u_0

simple linear relationship could be used to relate the HETP and u_0:

$$\text{HETP}(\mu\text{m}) = 707.4\, u_0 \text{ (cm/s)} + 37.6 \qquad (7)$$

Similar behavior has been previously observed [23]. Horváth and Lin [24] suggested this representation at the relatively (depending on the stationary phase) high mobile phase velocities usually selected in preparative chromatography.

Validation of the model

To check the validity of the band profiles calculations, we compare in Fig. 7 the calculated and experimental individual band profiles obtained in two different cases. The calculated profiles were obtained using the equilibrium-dispersive model [10], the competitive isotherm (Table 1), and the correlation between column plate height and mobile phase velocity (Eq. 7, Fig. 6). The experimental individual band profiles of the two enantiomers were obtained by analysis of collected fractions.

The comparison was done for large samples of two different binary mixtures, having relative compositions 1:1 and 1:3. The total amount injected was about 9 mg for the 1:1 mixture and 7 mg for the 1:3 mixture. These amounts correspond to the high degree of column overload typical of non-linear chromatography. The injec-

tion concentration exceeded only slightly the range within which the experimental adsorption data had been determined. Two different flow-rates were used for the two experiments, 0.6 ml/min for the 1:1 mixture, corresponding to an efficiency of 2500 theoretical plates, and 1.2 ml/min for the 1:3 mixture, corresponding to 1500 plates for our column.

As shown in Fig. 7, there is an excellent agreement in both instances between the experimental and the calculated band profiles, demonstrating the validity of the equilibrium-dispersive model, as reported previously [10–16,19]. In both instances, the first component band is slightly shorter than predicted, by approximately 10%. With the 1:3 mixture (Fig. 7B), we observe a small retention time difference of ca. 2% between the calculated and measured profiles. This difference could be explained by a corresponding error on the set flow-rate.

The agreement observed in Fig. 7 between experimental and calculated profiles justifies the use of the equilibrium-dispersive model and the bi-Langmuir isotherm model in all further calculations required by the comparison between recycling procedures.

4.2. Comparison of the various operating modes

We are interested in the production from the racemic mixture of one or other of the two enantiomers, with a purity of 99%. Because of the fundamentally unsymmetrical behavior of the interactions between the two component bands in non-linear chromatography [10], the results of the comparison differ strongly depending on whether the desired component is the less or the more retained component. These two cases are studied successively in the next two sections.

As mentioned in the Theoretical section, the variables involved in the comparison are the injection concentrations and volumes and the column efficiency N_p. For a given value of N_p, we have calculated the optimum injection conditions for the various operating modes and evaluated the performances achieved (prod-
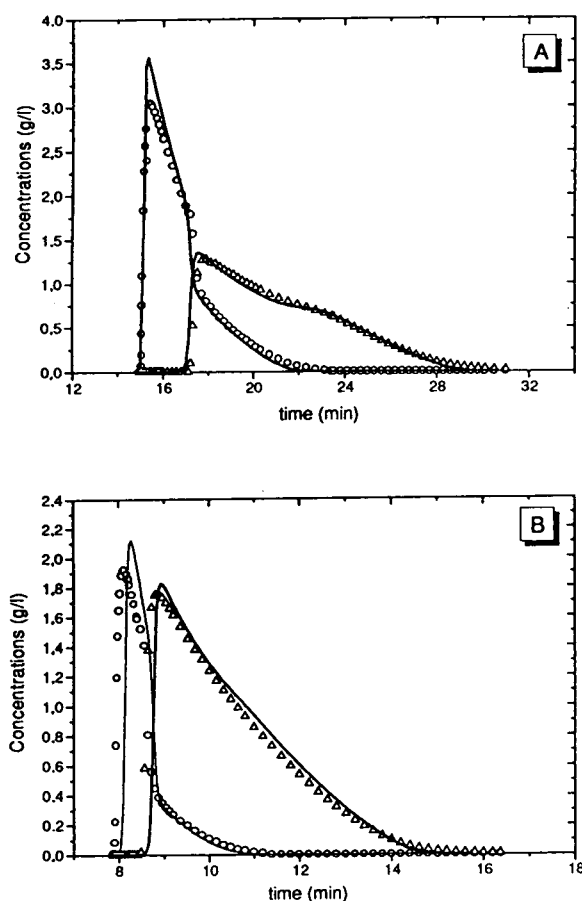
Fig. 7. Comparison between experimental [○ = R-(−)- and △ = S-(+)-enantiomer] and calculated (lines) individual profiles for an injection of a binary mixture. (A) $C_{(-)} = C_{(+)} = 3.72$ g/l; volume injected, 1.2 ml; flow-rate, 0.6 ml/min; column efficiency, 2500 plates. (B) $C_{(-)} = 1.49$ g/l; $C_{(+)} = 4.74$ g/l; volume injected, 1.2 ml; flow-rate, 1.2 ml/min; column efficiency, 1500 plates.

uctivity, eluent consumption). These calculations were repeated for several values of the number of theoretical plates.

The total concentration of the fresh feed was set equal to 6 g/l ($C_{1F} = C_{2F} = 3$ g/l). Although this injection concentration is not optimum, it was chosen for experimental, practical purposes because it is easy to achieve. Higher injection concentrations are not useful. Numerical calculations showed that no improvements of the economic parameters was obtained when the total injection concentration of the fresh feed ex-

ceeded 12 g/l. Moreover, this choice of C_{1F} and C_{2F} was observed to have no influence on the conclusions of the comparison.

First case: production of the less retained component

Figs. 8 and 9 illustrate the performance of the recycling operating modes compared with that of simple elution (for the absolute performance of the elution mode and of combination of recyclings, see below). They show plots of the production rate (Figs. 8A and 9A) and the

Fig. 8. Comparative performances of the operating modes studied with respect to elution chromatography for various column efficiencies. The less retained component is the product of interest. Minimum recovery yield sought = 99% (purity = 99%). ○ = "Recycling of the second peak tail"; □ = "segmented recycling"; △ = "combinations of recyclings".

Fig. 9. Same as Fig. 9, but minimum recovery yield sought = 90%.

model is very useful because it gives a good understanding and descriptive tool, and it permits a rapid estimation of approximate values of the optimum conditions, allowing a rapid choice of the proper experimental conditions for band calculations or for actual experiments.

Recycling of the second peak tail. As shown in Figs. 8 and 9, the performance of the mode "recycling of the second peak tail" relative to that of elution is almost independent of the column efficiency. A nearly constant but significant eluent saving, between 13 and 20%, is achieved by recycling the highly dilute fraction, but at the cost of a few percent decrease in production rate. With this recycling mode, the degree of separation reached at the column outlet is of the same order as in elution chromatography, but slightly lower. To obtain the same recovery yield, the amount of fresh feed injected must be slightly smaller. This causes a decrease in the production rate. It is not surprising that recycling the tail of the second band has a minor effect on the production rate and solvent consumption for the first band.

"Segmented recycling" and "recycling with mixing". Different results are obtained in "segmented recycling". At high efficiencies, this mode also offers a compromise between a loss of productivity and a decrease of the eluent consumption. The terms of the compromise are very close to those of "recycling of the tail" of the second peak.

By contrast, at low column efficiency, both an increase in the production rate and a decrease in the eluent consumption can be achieved simultaneously with "segmented recycling" and with "recycling with mixing". This is a very attractive proposition. The importance of the improvement increases with decreasing column efficiency. For example, with an efficiency of 500 theoretical plates and a recovery yield of 90%, the production rate can be increased by 50% and the eluent consumption halved. For a recovery yield of 99%, the gain in solvent consumption is close to 70% while the production rate is multiplied by 2.7, a considerable gain. We note that the rela-

eluent consumption (Figs. 8B and 9B) versus the column efficiency. Figs. 8 and 9 correspond to two different values of the required recovery yield, 99% and 90%, respectively. The injection conditions selected are optimum for the corresponding recovery yield constraint.

In the elution mode, the volume of fresh feed injected is the only variable there is to optimize for each column efficiency. In practice, this volume is the one for which the recovery yield required is just achieved. For the other operating schemes, the determination of the optimum injection conditions is more complex. A precise analysis is required for all of them. The ideal

tive retention of the two enantiomers at infinite dilution (Table 1) is 9.68/7.29 = 1.33. With this relative retention and the value of the retention factor (3.1), a resolution of 1.05 is achieved with a 500-plate column.

Fig. 10 explains the relative behavior of the "segmented recycling" and elution modes of operation. It shows plots of the optimum amount of fresh feed injected at the beginning of each cycle versus the column plate number, N_P, relative to the optimum amount which would be injected on an infinitely efficient column, for the two operating modes. Fig. 10A and B correspond to recovery yields of 99% and 90% (always with a purity constraint of 99%), respectively. For an infinitely efficient column, the optimum amount of fresh feed injected for each cycle would be the same with both modes. At high but finite efficiencies, the optimum amounts of fresh feed injected are almost identical. The decrease in production rate observed at high efficiency in "segmented recycling" is due to a longer cycle time. The eluent saving stems from an increase by a factor of 2–3 of the concentration of the less retained component in the eluent. This increase comes from the enhancement at higher concentrations of the non-linear effects involved in preparative chromatography. In this case the results are similar to those predicted by the ideal model.
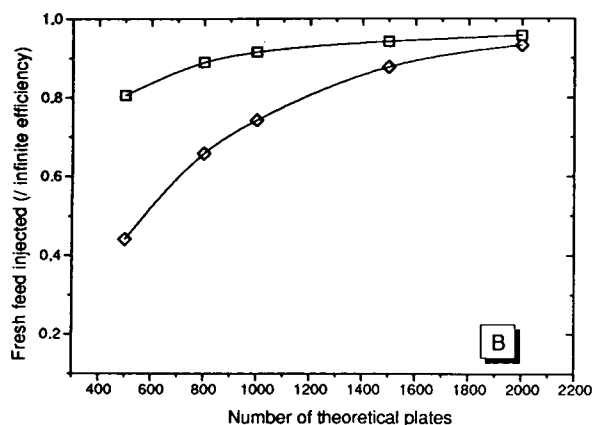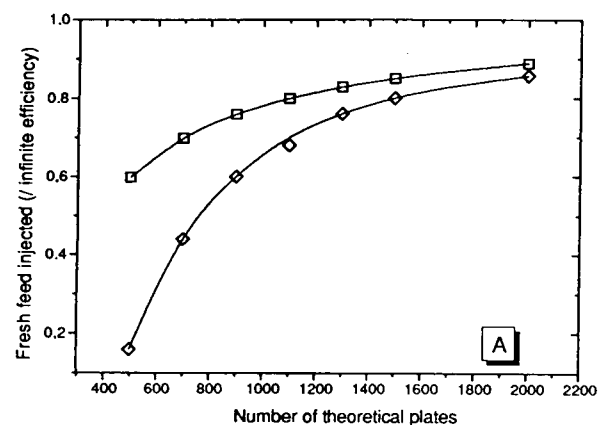
At low column efficiencies, the optimum amount of feed injected in the "segmented recycling" mode is much higher. The negative influence of the poor column efficiency (i.e., of the kinetic and hydrodynamic effects) is much stronger in the elution mode than in recycling. Elution performance is more strongly affected than recycling performance at small numbers of theoretical plates. This is especially true when the required degree of separation is high, i.e., for a required recovery yield of 99%.

The comparative performance of the "recycling with mixing" mode (Fig. 2A) is not given in Figs. 8 and 9, for the sake of clarity. It is too similar to that of "segmented recycling", albeit slightly inferior.

Fig. 10. Amount of fresh feed injected at each cycle for various column efficiencies compared with the amount injected to reach the same purity and recovery yield on a column of infinite efficiency. The less retained component is the product of interest (purity = 99%). Recovery yield = (A) 99% and (B) 90%. ◇ = Elution chromatography; □ = "segmented-recycling".

"Combination of recyclings". This combination of operating mode associates positively the performance achieved with the two simple modes. Its main interest appears at high column efficiencies, where the compromise between productivity and eluent consumption that it affords is the most attractive. For instance, with a 2000-plate column and a recovery yield of 99%, there is a 15% decrease in the production rate, but also a 40% saving in the eluent consumption. At low column efficiencies, the results achieved are

close to those given by the mode of "segmented recycling".

Conclusion. It appears that, for the production of the first component, the operating modes of "segmented recycling" and "combination of recyclings" offer better performance than elution, with often a higher production rate and always a lower solvent consumption. These modes are less sensitive to the spreading effects of a finite column efficiency than the conventional elution mode. Accordingly, we can expect their optimum operating conditions to be found at lower values of the column efficiency. At best, both an increase in production rate and a decrease in eluent consumption will be achieved. At worst, significant eluent savings will have to be paid for by a moderate decrease of the production rate. When the eluent consumption is the major contribution to the production cost, recycling could be justified even at the expense of a significant decrease in the production rate.

We compare in Fig. 11 the plots of the absolute values of the production rate or productivity and the eluent consumption obtained in elution and in "combination of recyclings" as a function of the column efficiency. The plots in Fig. 11
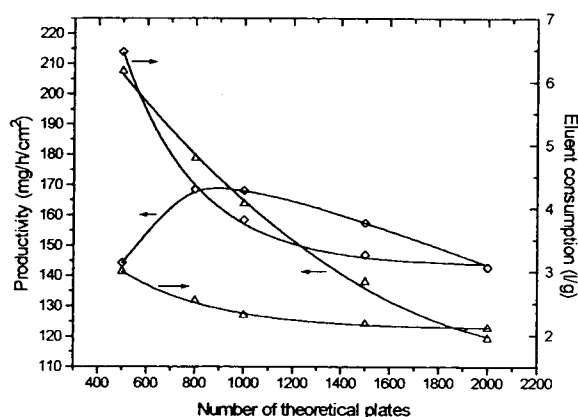
Fig. 11. Dependence of the production rate or productivity (mg/h·cm^2) for the less retained isomer and of the eluent consumption (l/g) on the column efficiency (◇) in the conventional elution mode and (△) in "combination of recyclings". Minimum recovery yield = 90%; purity = 99%.

were obtained for a recovery yield of 90% and a purity of 99%. The column length, the mobile phase velocity and the amount of feed injected were optimized separately for each value of the column efficiency, as explained above, using Eq. 7 and the condition $\Delta P \leq 50$ bar to derive the mobile phase velocity. These plots demonstrate that the production rate of elution is maximum for an efficiency between 850 and 900 theoretical plates and that the solvent consumption decreases steadily with increasing efficiency, with minimum changes to be expected beyond 2000 plates. These results are in excellent agreement with those derived by Felinger and Guiochon [11]. By contrast, the production rate with "combination of recycling" increases with decreasing efficiency and would exhibit a maximum only below 500 theoretical plates. This is a result of the relative insensitivity of the production rate on the hydrodynamics and kinetic effects, as noted above. As in elution, the solvent consumption in "combination of recycling" decreases with increasing column efficiency but it is much lower, especially at low efficiencies. Fig. 11 permits further direct comparisons between the performance of elution and "combination of recycling". For example, for $N_P = 800$, a number close to that for which the production rate is maximum in elution, the combination of recyclings offers nearly the same production rate but a 40% saving in eluent consumption. For $N_P = 500$, this combination permits a production rate 25% larger than elution did at $N_P = 800$ plates and at the same time a 30% decrease in eluent consumption. This example illustrates the attractive potential of the "combination of recyclings" mode in preparative chromatography.

The substitution of conventional elution by a recycling procedure, although admittedly more complex to implement, is especially attractive when, for one reason or another, it is difficult to obtain an efficient column (see Figs. 9 and 10). In such a case, the performance of the chromatographic process is improved considerably. This is the case for certain chiral phases, for example for microcrystalline cellulose triacetate, which is the most widely used phase for the industrial separation of enantiomers by preparative chro-

matography, and with which it is difficult to prepare columns having an efficiency in excess of ca. 1000 plates per metre [13,25]. In this case, the recycling operating modes presented here could be very attractive.

Finally, the selectivity of the system studied here, ca. 1.33, is relatively high. For systems exhibiting a lower selectivity, similar conclusions are expected. Much higher column efficiencies would be required, but the advantage resulting from the enhancement of the displacement effect when the sample size is increased, as allowed by recycling procedures, would permit the same separation to be obtained at efficiencies that are low compared with those which the elution mode would require. Further work is required to reach a definitive conclusion on this point.

Second case: production of the more retained component

If the product of interest is the more retained component, it has been shown that the production rate of a purified product does not increase significantly with increasing sample size beyond the stage where the two bands begin to interfere significantly and the whole amount of second component injected cannot be collected [26]. For an infinitely efficient column, the optimum sample size is that for which the recovery yield begins to drop below 100%. The result is similar for a real column. There is an optimum amount injected, which gives both the maximum production rate and the minimum eluent consumption [27,28]. The corresponding recovery yield of the more retained component depends on the column efficiency. It tends towards 100% with increasing column efficiency.

Fig. 12 compares the performances of "segmented recycling" and "combination of recyclings" with that of the elution mode. The optimum amount injected for elution was determined for each column efficiency. For the other two operating modes, the recovery yield was very close to 99%. Fig. 12 shows that in this case, the simple elution mode is the most favorable approach. "Segmented recycling" and "combination of recyclings" have performances that are nearly independent of the column ef-
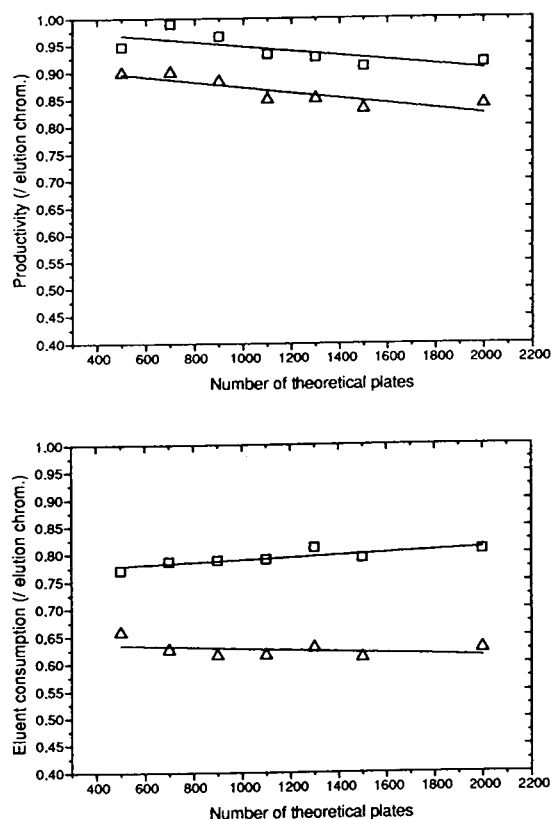
Fig. 12. Comparative performances of the operating modes studied to elution chromatography for various column efficiencies. The more retained component is the product of interest. □ = "Segmented recycling"; △ = "combination of recyclings".

ficiency. "Segmented recycling" causes a 5–10% decrease in production rate, but also a 20–25% eluent saving. "Combination of recyclings" gives a 10–15% decrease in production rate, but a 35–40% eluent saving. The compromise offered by "combination of recyclings" may be attractive for industrial separations, when the eluent consumption is the major part of the production costs.

It must be pointed out that a considerable improvement of the separation performance would also be observed at low efficiencies if a high minimum recovery yield was required. The situation would be the same as in Fig. 9 when the less retained component is the product of interest.

4.3. Experimental illustration

Experiments were conducted with the chromatographic system and equipment described above to investigate the conclusions of the theoretical study and the validity of the models used. The column efficiency was 2000 theoretical plates, for a length of 25 cm. The flow-rate was set at 0.8 ml/min, corresponding to a velocity $u = 0.119$ cm/s. Our objective was the production of either enantiomers with a purity and a recovery yield of 99%. As explained above, the total concentration of the fresh feed was chosen equal to 6 g/l, so the injection volumes would be large enough to permit the injection of accurately known volumes using the solvent delivery system.

The various modes of recycling operate under cyclic steady-state conditions. A dynamic equilibrium must be reached before measurements can be carried out. If a production run is started with the injection of pure feed, the chromatogram, the production per cycle at a given purity and the composition of the recycled fraction will drift toward their steady-state values but it may take a significant number of cycles before this dynamic equilibrium is reached. This process can be accelerated by a judicious choice of the start-up conditions, i.e., by mixing the fresh feed with a solution having a composition close to that of the recycled fraction. An educated guess is required, which is easy if previous purifications have been made with the same feed. For a new separation, the easiest method consists in calculating, then preparing the solutions corresponding to the composition of the recycled fractions of the cyclic steady state [5]. This preparation is easily achieved using the solvent delivery system of our chromatograph.

Figs. 13–16 compare the experimental (symbols) and the calculated (lines) chromatograms obtained during the first cycles of the various operating schemes. For these first cycles, synthetic solutions were prepared according to the numerical prediction of the corresponding cyclic steady state with a 2000-plate column. For "recycling with mixing" (Fig. 14), a synthetic solution with the calculated composition of the
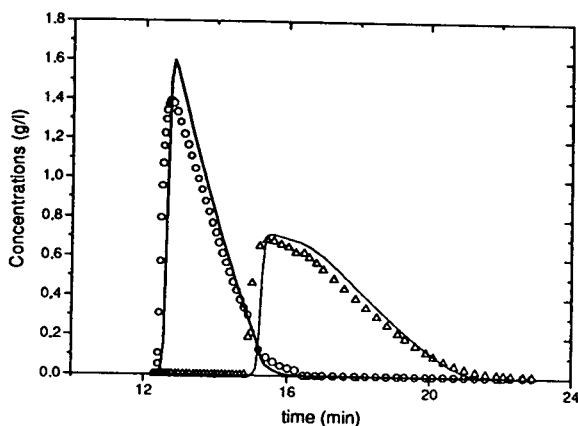
Fig. 13. Comparison between experimental [$\bigcirc$ = R-(−)- and $\triangle$ = S-(+)-enantiomer] and calculated (lines) concentration profiles for elution chromatography. Flow-rate, 0.8 ml/min (2000 plates); $C_{(-)} = C_{(+)} = 3.0$ g/l; volume injected, 0.62 ml.

mixed feed was injected as a rectangular plug. In the case of "segmented recycling", the first cycle (Fig. 15A) was repeated several times and the mixed zone of the chromatogram was collected according to the cut times given by numerical simulations. The resulting solution was used for a second cycle (Fig. 15B). For "recycling of the second peak tail" (Fig. 16), we made several
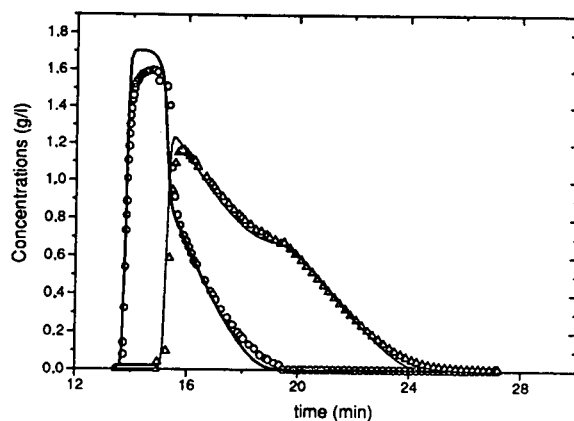
Fig. 14. Comparison between experimental [$\bigcirc$ = R-(−)- and $\triangle$ = S-(+)-enantiomer] and calculated (lines) concentration profiles of the first cycle of "recycling with mixing". Mixed feed: $C_{(-)} = 0.99$ g/l; $C_{(+)} = 1.38$ g/l; volume injected, 3.06 ml.

Fig. 16. Comparison between experimental [$\bigcirc = R$-$(-)$- and $\triangle = S$-$(+)$-enantiomer] and calculated (lines) concentration profiles of the first cycle of "recycling of the second peak tail". Fresh feed: $C_{(-)} = C_{(+)} = 3.0$ g/l; volume injected, 0.62 ml. Recycled fraction: $C_{(+)} = 78$ mg/l; volume injected, 2.24 ml.
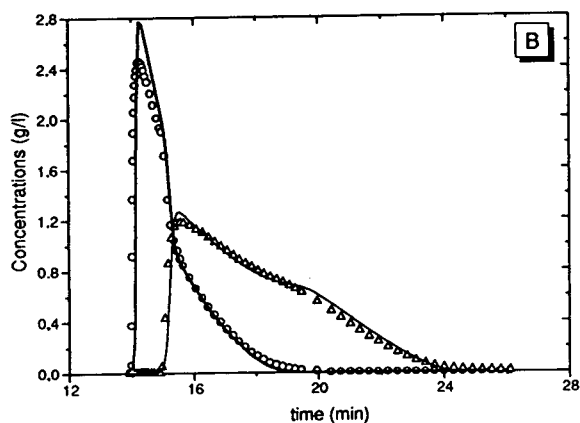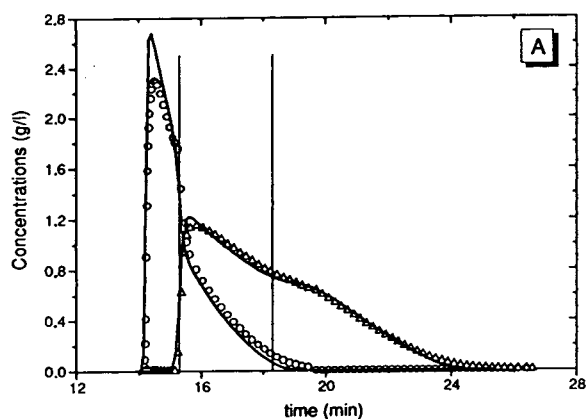
Fig. 15. Comparison between experimental [$\bigcirc = R$-$(-)$- and $\triangle = S$-$(+)$-enantiomer] and calculated (lines) concentration profiles of (A) the first and (B) the second cycles of "segmented recycling". Fresh feed: $C_{(-)} = C_{(+)} = 3.02$ g/l; volume injected, 0.62 ml. Recycled fractions: $C_{(-)} = 0.490$ g/l; $C_{(+)} = 0.975$ g/l; volume injected, 2.40 ml (first cycle); and $C_{(-)} = 0.541$ g/l; $C_{(+)} = 1.028$ g/l; volume injected, 2.40 ml (second cycle).

injections of fresh feed and collected the rear part of the elution profiles. This solution was used to generate the volume of recycled fraction of the more retained component which was needed.

A very good agreement is generally observed between the experimental and calculated band profiles shown in these figures, as also seen in Fig. 7. The model used is able to predict with excellent qualitative agreement the results of actual series of experiments. The purity and recovery yield of the products were derived from the experimental profiles, as well as the values of the economic criteria. However, the first conclusion from these determinations was that the required values of the purity and the recovery yield were not exactly achieved in the experiments (Table 2). In spite of the excellent agreement between calculated and measured band profiles (Figs. 7 and 13–16), the model does not predict accurately enough the performance of the operating modes. Errors of the same order of magnitude have been reported by Jacobson et al. [15,29].

One of the main reasons for this error is that the model does not predict correctly the tail of the less retained component band in the mixed zone of the chromatogram. Some empirical adjustment (amounts injected, cut times) remains necessary. This error originates from inaccuracies in the competitive isotherm model, which does not predict exactly the magnitude of the displacement effect (e.g., Fig. 14). The errors made in the determination of the competitive isotherm are in part due to the unavailability of the less retained compound, in part owing to the use of a simple model.

Table 2
Purity and recovery yield of the collected fractions

Mode	P_1 (%)	Y_1 (%)	P_2 (%)	Y_2 (%)
Elution	98.1	95.6	96.9	93.3
"Recycling with mixing"	98.1	99.2	97.4	~100
"Segmented recycling"	97.0	97.0	97.0	96.7
"Rectckubg of the second peak tail"	98.1	95.6	96.9	93.3

The aim was a purity (P) of 99% and a recovery yield (Y) of 99%.

5. Conclusions

Although it is accurate enough to give excellent results in qualitative comparisons between calculated and experimental band profiles, the combination of techniques used for the modeling of competitive equilibrium isotherms and for the calculation of band profiles in non-linear chromatography does not predict exactly the recovery yield or the degree of purity achieved. Nevertheless, there is a semi-quantitative agreement, and important conclusions can be derived from calculated profiles regarding the optimization of experimental parameters.

The theoretical investigation made of several operating modes based on recycling of intermediate fractions shows that they might often be more economically attractive than conventional elution chromatography. Almost always, the eluent consumption is reduced significantly. In a number of cases, the production rate of the first component of a binary mixture is markedly increased. In the case studied, the separation required is easy, the column efficiency needed is low and the use of recycling brings dramatic improvements in both the production rate, which increases, and the solvent consumption, which decreases. The advantage of recycling remains probably as important when the separation carried out is difficult.

The major part of the separation cost of the processes using high-performance preparative liquid chromatography, at least when organic solvents are used, is related to the solvent costs. Accordingly, the operating modes studied here are of great interest because of their potential for significant eluent savings.

Symbols

a, b	Isotherm coefficients, enantioselective sites (B l/g)
A, B	Isotherm coefficients, non-enantioselective sites (B: l/g)
C	Concentrations in the mobile phase (g/l)
C_{1F}, C_{2F}	Concentrations of the fresh feed (g/l)
D_{ap}	Apparent dispersion coefficient
L	Column length
N_P	Column efficiency
q	Concentrations in the adsorbed phase
t	Time
t_0	Column dead time
u_0	Propagation velocity of a non-adsorbed component
x	Abscissa along the column

z	Reduced abscissa
ε_T	Total porosity of the bed
θ_i	Recovery yield of component i

References

[1] R.M. Nicoud and H. Colin, *LC·GC*, 8, No. 1 (1990) 22.

[2] H. Colin, in M. Perrut (Editor), *Proceedings of PREP'92, Nancy, France, April 1992*, Société Française de Chimie, Paris, 1992, p. 259.

[3] A. Seidel-Morgestern and G. Guiochon, *AIChE J.*, 39 (1993) 809.

[4] B. Balannec and G. Hotier, *Rev. Inst. Fr. Pet.*, 46 (1991) 803.

[5] M. Bailly and D. Tondeur, *Chem. Eng. Sci.*, 37 (1982) 1199.

[6] M. Bailly and D. Tondeur, *Chem. Eng. Sci.*, 36 (1981) 455.

[7] M. Bailly and D. Tondeur, *Chem. Eng. Process.*, 18 (1984) 293.

[8] Y. Okamoto, M. Kawashima and K. Hatada, *J. Am. Chem. Soc.*, 106 (1984) 5357.

[9] Y. Okamoto, M. Kawashima, R. Aburatani, K. Hatada, T. Nishiyama and M. Masuda, *Chem. Lett.*, (1986) 1237.

[10] S. Golshan-Shirazi and G. Guiochon, in F. Dondi and G. Guiochon (Editors), *Theoretical Advancement in Chromatography and Related Separation Techniques (NATO ASI Series)*, Kluwer, Dordrecht, 1992, p. 35.

[11] A. Felinger and G. Guiochon, *AIChE J.*, 40 (1994).

[12] S. Jacobson, S. Golshan-Shirazi and G. Guiochon, *AIChE J.*, 37 (1991) 836.

[13] A. Seidel-Morgestern and G. Guiochon, *Chem. Eng. Sci.*, 48 (1993) 2787.

[14] A.M. Katti, M. Czok and G. Guiochon, *J. Chromatogr.*, 556 (1991) 205.

[15] S. Jacobson, A. Felinger and G. Guiochon, *Biotechnol. Prog.*, 8 (1993) 533.

[16] P. Rouchon, M. Schonauer, P. Valentin and G. Guiochon, *Sep. Sci. Technol.*, 22 (1987) 1793.

[17] D.H. James and C.S.G. Phillips, *J. Chem. Soc.*, (1954) 1066.

[18] J.M. Jacobson, J.H. Frenz and C. Horváth, *Ind. Eng. Chem. Res.*, 26 (1987) 43.

[19] F. Charton, S. Jacobson and G. Guiochon, *J. Chromatogr.*, 630 (1993) 21.

[20] I.W. Wainer, R. Stiffin and T. Shibata, *J. Chromatogr.*, 411 (1987) 139.

[21] Y. Fukui, A. Ichida, T. Shibata and M. Kyozo, *J. Chromatogr.*, (1990) 85.

[22] M.D. LeVan and T. Vermeulen, *J. Phys. Chem.*, 85 (1981) 3247.

[23] R. Endele, I. Halász and K. Unger, *J. Chromatogr.*, 99 (1974) 377.

[24] C. Horváth and H.J. Lin, *J. Chromatogr.*, 149 (1978) 43.

[25] E. Francotte and A. Junker-Buchheit, *J. Chromatogr.*, 576 (1992) 576.

[26] S. Golshan-Shirazi and G. Guiochon, *Anal. Chem.*, 61 (1989) 1276.

[27] A. Katti and G. Guiochon, *Anal. Chem.*, 61 (1989) 982.

[28] S. Golshan-Shirazi and G. Guiochon, *Anal. Chem.*, 61 (1989) 1368.

[29] S. Jacobson, A. Felinger and G. Guiochon, *Biotechnol. Bioeng.*, 40 (1993) 1210.

Synthesis and spectrometric characterization of a true diol bonded phase

Joseph J. Pesek*, Maria T. Matyska

Department of Chemistry, San Jose State University, 1 Washington Square, San Jose, CA 95192, USA

Abstract

The syntheses of two types of diol bonded phases using a silanization/hydrosilation pathway are described. One product is similar to the phase produced by organosilanization but without the problem of potential side products and utilizes allyl glycidyl ether as the bonding olefin to the intermediate hydride. The second product is a "true" diol phase without any heteroatoms from the surface to the diol moiety and involves reacting 7-octene-1,2-diol with hydride silica. Both products can be obtained in good yield and the bonded phases can be characterized spectrometrically by diffuse reflectance IR Fourier transform and cross-polarization magic angle spinning NMR spectrometry. The only complication to this reaction scheme is bonding of the olefin diol to the silica surface via one of the alcohol groups (esterification). It appears that this is a problem when the original silica contains a significant number of unassociated silanol groups.

1. Introduction

The use of epoxy and/or diol phases in HPLC has been established for many years [1,2]. The normal synthetic route involves bonding of 3-glycidyloxypropyltrimethoxysilane to silica followed by hydrolysis of the epoxide ring to form the diol [2]. The epoxide is a useful intermediate for the production of certain ion-exchange [3] or affinity phases [4–7]. In earlier studies it was suggested that bonding of the silane to silica under slightly acidic conditions (pH 5–6) was necessary to preserve the oxirane ring [3,4,8]. However the presence of the reactive epoxide species leads to complications in the bonding reaction that can often lead to a variety of products instead of the desired epoxy or diol

materials. A recent study has shown that temperature, pH and solvent all have a considerable effect on the presence and nature of species other than the intermediate epoxide or the diol bonded phases [9]. Among the species identified on the silica surface were various glycol ethers, polyaddition products, and cyclic oligosilanes. Therefore, it is not surprising that diol phases from different sources, and perhaps even different batches, display varying degrees of hydrophilic and hydrophobic properties. In order to solve the problem of variability in the diol product phase, another approach to synthesizing this material is needed.

The reactive glycidyloxypropyl and trimethoxy species on the bonding moiety complicate the process which leads to the formation of more than one product. A possible solution to this problem is to replace the organosilanization

* Corresponding author.

process by a silanization/hydrosilation reaction scheme [10]. This approach eliminates the organosilanization reagent and immediately removes the possibility of reactions involving the trimethoxy species. The key to this synthetic scheme is the formation of a hydride intermediate on the silica surface. This can be accomplished by the reaction of silica in the presence of the hydrolysis product of triethoxysilane (TES) under carefully controlled conditions. The reaction can be described as follows:

$$-Si-OH \ + \ (HO)_3-Si-H \ \rightarrow \ -Si-O-Si-H$$
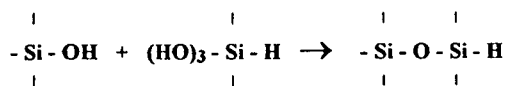
With the correct TES:silica ratio and the appropriate concentration of acid catalyst, a monolayer of hydride can be formed on the surface. The net effect is to essentially replace all of the silanols with hydrides on the surface. While 100% efficiency is not possible due to incomplete condensation of TES on the surface, the small size of hydride in comparison to other organic moieties makes this process much more efficient than typical organosilanization reactions [10]. Indeed, not all silanols can be removed from silica because some are "buried" in micropores that would be inaccessible to the silanization reagent (hydrolysis product of TES). However, in theory very few accessible silanols should be present on the silica surface when compared to bonded phases produced by organosilanization.

Once the hydride intermediate has been formed, then there are two options for forming a diol bonded material. The first is a synthetic path similar to the one used for preparation of epoxy and diol phases by the conventional method. It involves bonding allyl glycidyl ether (AGE) to the hydride intermediate as follows:

$$-Si-H \ + \ CH_2=CH-CH_2-O-CH_2-CH \overset{O}{\underset{\diagup \diagdown}{-}} CH_2 \ \overset{cat.}{\rightarrow}$$

$$-Si-CH_2-CH_2-CH_2-O-CH_2-CH \overset{O}{\underset{\diagup \diagdown}{-}} CH_2$$
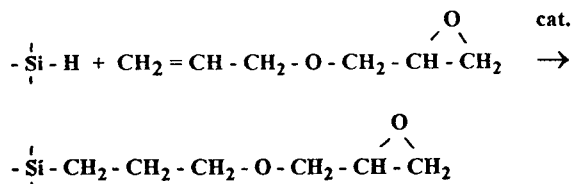
After bonding of the oxirane compound, then formation of the diol can be accomplished by opening of the epoxide ring in acid solution:

$$-Si-CH_2-CH_2-CH_2-O-CH_2-CH \overset{O}{\underset{\diagup \diagdown}{-}} CH_2 \ \overset{H_2O}{\underset{H^+}{\rightarrow}}$$

$$-Si-CH_2-CH_2-CH_2-O-CH_2-\overset{OH}{\underset{|}{CH}}-CH_2-OH$$
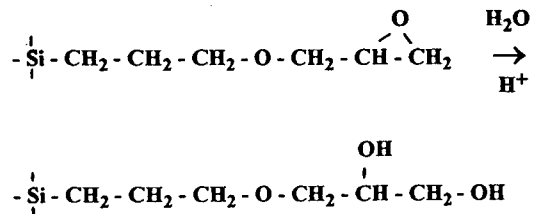
While this general procedure is similar to that of the organosilanization reaction, it eliminates those side reactions which involve the trimethoxy reactive groups on the organosilane. However, care must still be exercised when bonding the olefin since too high of a concentration of this species can result in polymerization rather than monomeric bonding. Even if the bonding scheme works perfectly, the final product is not a true diol since the ether linkage is still present leading to a slightly higher hydrophilicity. However, this reaction scheme still preserves the option of utilizing the epoxide moiety as a binding site in such applications as affinity chromatography.

The possibility of forming a true diol bonded phase can be realized by reacting the hydride intermediate with 7-octene-1,2-diol (7-OD).

$$-Si-H \ + \ CH_2=CH-(CH_2)_4-\overset{OH}{\underset{|}{CH}}-CH_2-OH \ \overset{cat.}{\rightarrow}$$

$$-Si-(CH_2)_6-\overset{OH}{\underset{|}{CH}}-CH_2-OH$$
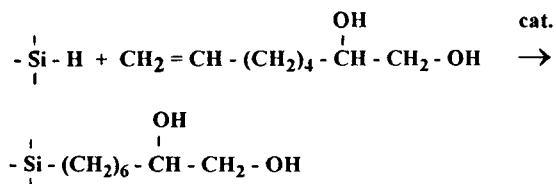
In this procedure the olefin is presumably the most reactive site and hydrosilation is the preferred pathway leading to the formation of the diol bonded phase. The absence of other functional groups or heteroatoms on the above product results in a "true" diol bonded phase. Some caution must be exercised in this reaction scheme since bonding via the terminal hydroxy group is possible leading to a silicon–oxygen–carbon linkage at the surface [11] and an olefin at the opposite end of the organic moiety.

$$-\overset{|}{\underset{|}{Si}} - H \; + \; CH_2 = CH - (CH_2)_4 - \overset{\overset{\displaystyle OH}{|}}{CH} - CH_2 - OH \quad \xrightarrow{\text{cat.}}$$

$$-\overset{|}{\underset{|}{Si}} - O - CH_2 - \overset{\overset{\displaystyle OH}{|}}{CH} - (CH_2)_4 - CH = CH_2$$
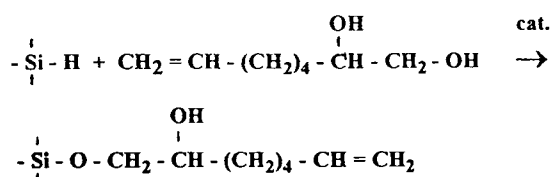
Since such competitive reactions have not been studied before on silica, careful spectrometric characterization of the hydrosilation product must be undertaken to determine if esterification contributes significantly to the population of bonded organic species.

2. Experimental

Solid-state NMR spectra, cross-polarization (CP) with magic angle spinning (MAS) were obtained on a Bruker MSL 300 spectrometer. Samples of 200–300 mg in a double bearing ZrO_2 rotor were measured at a spinning rate of 5 kHz. For ^{13}C spectra the contact time was 5 ms with a repetition rate of 5 s. The chemical shifts were referenced to tetramethylsilane by using the carbonyl signal of glycine as a standard. For ^{29}Si, CP-MAS-NMR spectra were recorded using a contact time of 5 ms and a pulse repetition rate of 5 s with poly(hydrido)siloxane as the standard.

Diffuse reflectance IR Fourier transform (DRIFT) spectra were obtained with a Perkin-Elmer Model 1800 Fourier transform (FT) IR spectrometer equipped with a deuterated triglycine sulfate (DTGS) detector. Spectra were taken in the 4000–450 cm^{-1} region with a nominal resolution of 2 cm^{-1}. Two hundred sample

scans were acquired against pure KBr as a reference. Spectra were normalized to 100% transmittance.

Carbon analysis was carried out by a conventional combustion method with a Perkin-Elmer Model 240C elemental analyzer equipped with a Model 56 recorder. Samples of 5–15 mg were analyzed.

2.1. Materials

The silica gels used are listed in Table 1. The samples were dried in a vacuum oven at 110°C for 48 h before the silanization. AGE and 7-OD (both Aldrich) were used to prepare the bonded silicas. A 10 mM hexachloroplatinic acid (3.75% as Pt, Aldrich) solution in 2-propanol was used as the catalyst for the hydrosilation reaction. p-Dioxane (J.T. Baker) was dried by contact with calcium hydride (Sigma) for several days and then distilled before use. TES (Huls America) was used as received. Distilled water was purified in the laboratory using a Millipore Model Milli-Q deionization system.

The concentration of surface-bonded groups was obtained from carbon percentages determined from elemental analysis according to the equation of Berendsen and De Galan [12]. The specific surface area used in the calculation was the value given by the supplier. See Table 2.

Hydride silica intermediates were prepared by a previously reported silanization method [10]. The hydride-modified supports were formed by reacting TES (45 ml of 0.5 M TES/dioxane solution) with the silica substrates (5.00 g of each silica) in the presence of water, an acid catalyst

Table 1
Physical characteristics of bare silica gels

No.	Type	Porosity			Supplier
		d_p	S_{BET}	D	
1	Vydac 101 TP, lot 900 201	6.58	106.5	380	Separations Group (Hesperia, CA, USA)
2	Davisil, grade 633	32–74	480.0	60	Aldrich (Milwaukee, WI, USA)

d_p = Particle diameter (μm); S_{BET} = specific surface area (m^2/g); D = pore diameter (Å). All data from supplier.

Table 2
Carbon content and surface coverages (α_R) of diol silica samples

No.	Sample	Corrected (% C)[a]	α_R (μmol/m^2)
1	Vydac-diol (AGE)[b]	4.16	5.87
2	Vydac-diol (AGE)	2.18	2.96
3	Davisil-diol (AGE)	7.24	2.41
4	Vydac-diol (7-OD), grey	3.35	3.45
5	Davisil-diol (7OD), grey	10.10	2.58
6	Vydac-diol (7OD), white	3.54	3.65
7	Davisil-diol (7OD), white	9.77	2.48
8	Vydac-diol (7OD), white, large batch	4.08	4.25

[a] % C after subtracting the amount before hydrosilation reaction.
[b] Possible polymerization of AGE.

(7 ml of 2.3 M HCl solution) and dioxane (100 ml). Under those experimental conditions an SiH monolayer was covalently bonded to the silica surface. The final product was washed consecutively with 50-ml portions of water–tetrahydrofuran (THF) (20:80), THF and diethyl ether (twice with each solvent), dried at room temperature and then in a vacuum oven at 110°C for at least 6 h.

The chemically bonded phases were synthesized by a modification of previous methods [13,14]. The solution of olefin (2.65 ml) in toluene (40.82 ml) and catalyst (1.53 ml of 10 mM H$_2$PtCl$_6$) was heated to about 60–70°C while being magnetically stirred for 1 h, then the sample of hydride modified silica (2.60 g) was slowly added. The reaction was allowed to proceed for 96 h at 90 ± 1°C. After this time period the solvent was removed and the solid washed with 30-ml portions of toluene (three times) followed by similar washings with dichloromethane and diethyl ether. The product was dried at room temperature first for 1 h and then under vacuum at 110°C overnight. The chemical surface modification with AGE was carried out in standard glassware as described above but under nitrogen. The hydrolysis of AGE to form the diol consisted of placing the epoxy material in a stirred 0.05 M HCl solution at 65 ± 1°C for 4 h.

3. Results and discussion

In order to properly interpret the spectral results of the product phases, it was necessary to first characterize the hydride intermediate. A surprising result was obtained in routine carbon analysis when both the bare silica and the hydride were used for background measurements in the determination of the ligand concentration on the product phases. The bare Davisil did not have a measurable amount of carbon while the hydride intermediate had 0.4% C after silanization with TES. The bare Vydac silica and the hydride contained 0.3 and 1.0% carbon respectively. The presence of the carbon containing material was confirmed by DRIFT spectra in all cases when peaks in the 3000–2800 cm^{-1} range were observed. In the case of Vydac material, the ^{13}C CP-MAS-NMR spectrum (Fig. 1A) of the bare silica displays a prominent peak at about 50 ppm (most likely a methoxy group) while the spectrum of the hydride material (Fig. 1B) shows peaks at 17 and 60 ppm (also detectable in the bare silica spectrum). These latter chemical shifts correspond to the methyl and methylene resonances of the ethoxy moiety. The presence of the ethoxy species on the surface can be accounted for by either incomplete hydrolysis of TES or by subsequent bonding of the ethanol released during TES hydrolysis to the silica surface. What was even more surprising was that these peaks persisted even after extensive washing under aqueous conditions since the ethoxy moiety is bonded to the silica surface via a labile Si–O–C linkage. Even though some carbon is detected by elemental analysis and DRIFT for the Davisil hydride, no peaks were observed in the ^{13}C CP-MAS-NMR spectrum which indicates that residual ethoxy species are not a significant factor on this material.

Each of the two proposed reaction schemes for producing the diol on hydride-modified silica were tested and studied spectrometrically. The following sections summarize the DRIFT and CP-MAS-NMR results for the two types of diol.

3.1. Diol via allyl glycidyl ether

The differences between the final material
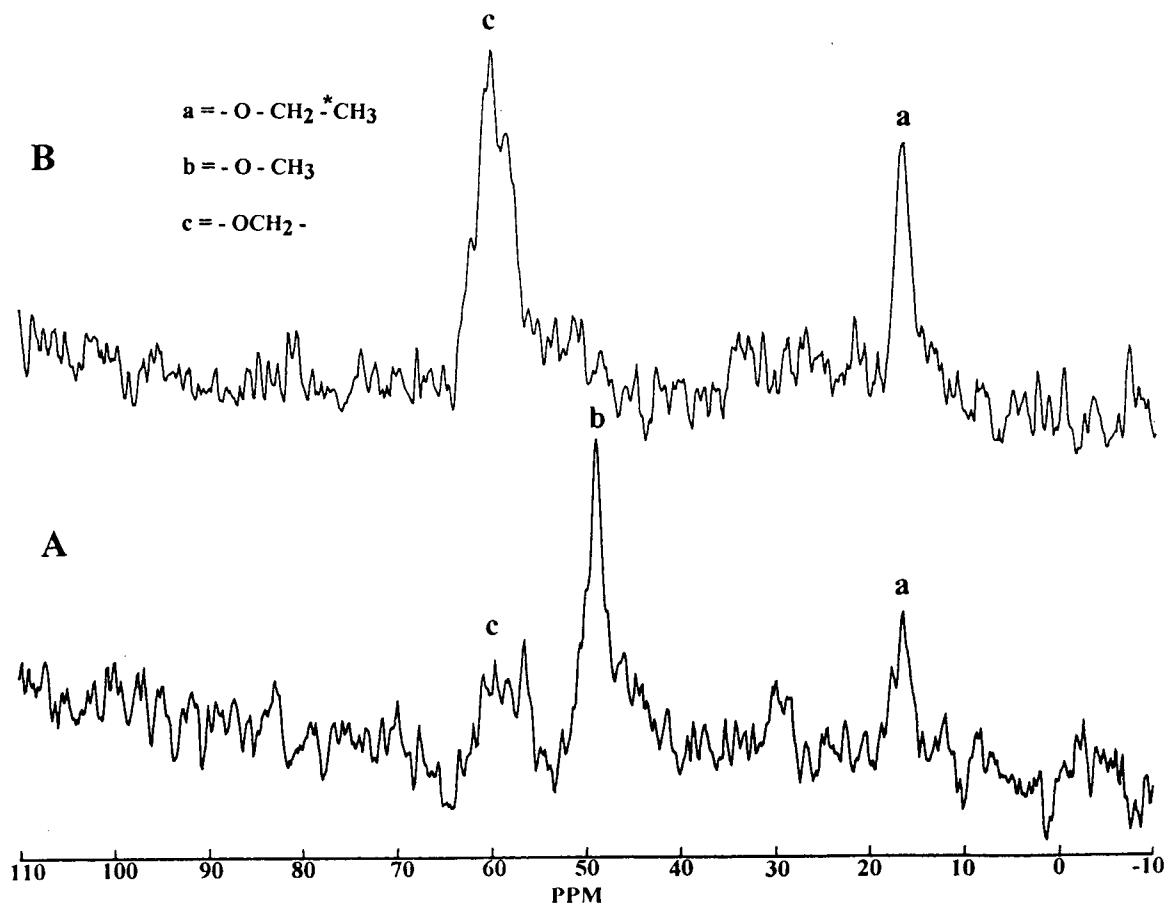
Fig. 1. ^{13}C CP-MAS-NMR spectra of bare (A) Vydac silica and the same material after reaction with the hydrolysis product of triethoxysilane (B).

produced by the silanization/hydrosilation and organosilanization processes revolve mainly around the number of possible bonded moieties. As enumerated above, the latter process can lead to several different species in varying proportions on the silica surface depending on the reaction conditions [9]. However, the hydride-based material generally leads to only a single bonded material if the concentration of AGE is kept low. Then polymerization of this molecule will be limited with only the epoxy bonded phase on the surface.

Fig. 2 shows the DRIFT spectrum of the hydride intermediate on Davisil silica (spectrum A). The characteristic Si–H stretching frequency is readily apparent at 2250 cm^{-1}. Upon reaction with AGE, the Si–H peak diminishes in intensity and strong C–H stretching bands appear in the 3000–2800 cm^{-1} range as shown in Fig. 2 (spectrum B). These results indicate successful formation of at least the epoxide bonded phase. Evaluation of the extent of the hydrolysis process cannot be made from these spectra. A similar analysis can be made of the DRIFT spectra for the same reactions on Vydac silica. It is also not possible to ascertain conclusively from the DRIFT spectra whether any side reactions or polymerization have taken place during the bonding process.

The hydrosilation process results in a direct
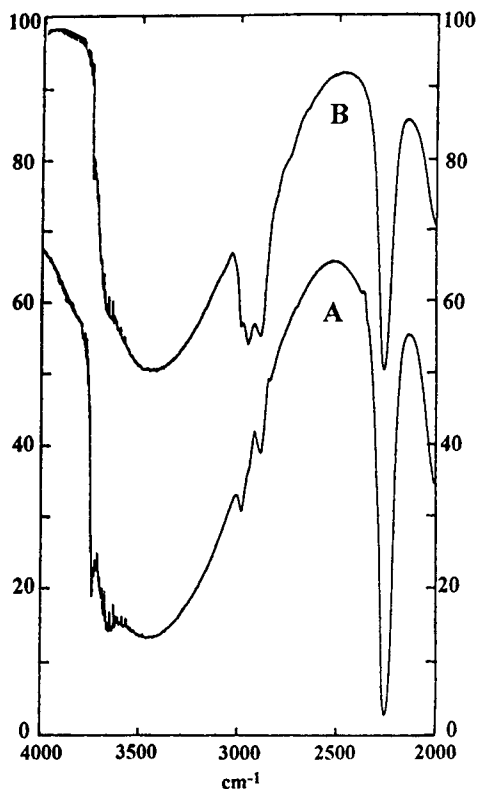
Fig. 2. Partial DRIFT spectra of hydride intermediate on Davisil (A) and the product obtained from the reaction with AGE (B).

silicon–carbon bond at the surface. This can be easily detected in the ^{29}Si CP-MAS-NMR spectrum. Fig. 3 shows the ^{29}Si NMR spectra of the AGE product on both Davisil (A) and Vydac (B) silicas. In both spectra the peak for the silicon atom directly bonded to carbon can be seen at −65 ppm. The intensity of this peak is low because there are relatively few protons close-by to provide the cross-polarization necessary for increased sensitivity. In addition to the appearance of the new peak at −65 ppm, a decrease in the Si–H peak intensity at −85 ppm is also observed. This is expected since success of the hydrosilation reaction converts hydride sites into bonded organic moieties.

Fig. 4 shows the ^{13}C CP-MAS-NMR spectrum of the AGE moiety bonded on Vydac silica and subsequently hydrolyzed by HCl. This spectrum can be used to confirm both the success of the

bonding reaction as well as the hydrolysis from the epoxide to the diol. The peaks for the ethoxy group identified on the hydride can be seen at 16 and 60 ppm. The peak at 7 ppm represents the carbon directly bonded to the silica surface. The peak at 22 ppm is due to the methylene group next to the surface bonded carbon. The peaks at 71 ppm and 65 ppm (shoulder) are due to the remainder of the carbon atoms in the molecule. Each of these carbons is bonded to an oxygen atom and accounts for the downfield chemical shift. If the hydrolysis reaction is not complete a peak at 45 ppm indicates the presence of the epoxide bonded material [8,15,16]. Absence of this peak indicates success of the acid hydrolysis and the presence of the resulting diol.

3.2. Diol via 7-octene-1,2-diol

The results above indicate the ease of bonding the AGE moiety to either a wide-pore (Vydac) or a narrow-pore (Davisil) silica as well as the success of hydrolyzing the epoxide ring to the diol product. However, the final bonded moiety is not a true diol since it is composed of two three-carbon chains connected via an ether linkage. Because of the versatility of the hydrosilation reaction in the presence of the hydride surface, it is possible to make a true diol phase, i.e. with no heteroatoms between the silica and the 1,2-diol moiety, using a terminal olefin with two OH groups on the last two carbons at the opposite end of the molecule. To test this possibility, 7-OD was chosen as the olefin to be bonded. This molecule possess a hydrocarbon chain which is long enough so that the diol functionality is significantly removed from the surface with a hydrophobic buffer in-between.

Fig. 5 shows the DRIFT spectra of products from the reaction of 7-OD on both Davisil hydride (A) and Vydac hydride (B). In both cases, a comparison with the starting hydride shows a decrease in intensity of the Si–H band at 2255 cm^{-1} and a concomitant appearance of C–H stretching bands between 2800–3000 cm^{-1}. These results conclusively prove that bonding of a diol-containing olefin can be readily accomplished on a hydride-modified silica surface.
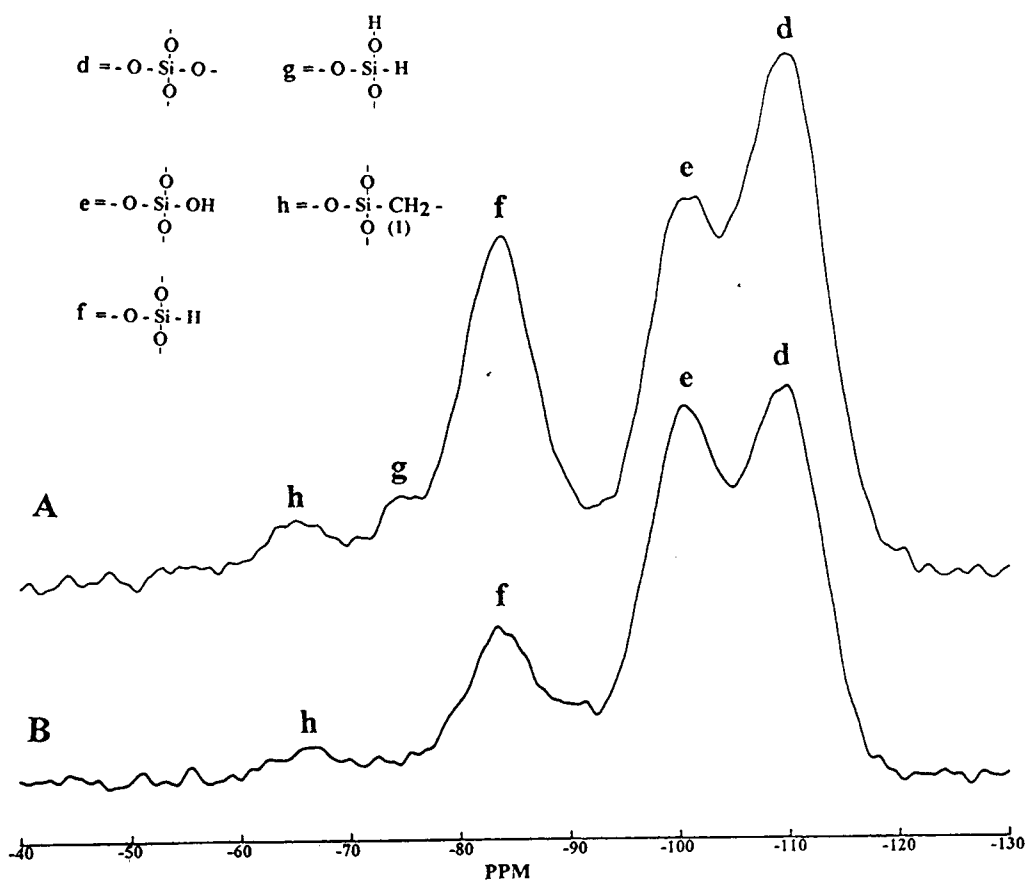
Fig. 3. ^{29}Si CP-MAS-NMR spectra of the products from the reaction of the hydride intermediate and AGE on Davisil (A) and Vydac (B) silicas.

However, closer examination of the C–H stretching region in the DRIFT spectra (Fig. 6) of the Davisil diol shows an additional band above 3000 cm^{-1}. Its location at 3080 cm^{-1} corresponds to the expected position of an allylic functional group. While its intensity is low, it does indicate that at least a small amount of the olefin is bonded via one of the OH groups. Since most of the silanols on the bare silica have been replaced by hydrides, it is not surprising that only a small amount of the olefin is bonded via a Si–O–C linkage. Nevertheless, this result proves that there is a competition between the hydrosilation and the silanol/alcohol condensation reactions. However, the same effect is not observed on the Vydac silica. This may be due to the fact the bare Vydac silica contains relatively few

isolated silanols or impurities on the surface in comparison to Davisil silica. Therefore, the silanols which remain after TES silanization on Vydac are probably still the less-reactive associated Si–OH's.

Further confirmation of the bonding process involving hydrosilation with 7-OD can be acquired from CP-MAS-NMR spectra. Fig. 7 shows the ^{29}Si CP-MAS-NMR spectra of both the Vydac (A) and Davisil (B) products. In each case there is a decrease in the intensity of the peak at −85 ppm (O_3–Si–H) and the appearance of a new peak at −65 ppm (O_3–Si–C–). A similar conclusion about the intensity of the peak at −65 ppm to that for the AGE product can be made for 7-OD. That is its intensity is low because there are relatively few protons nearby.
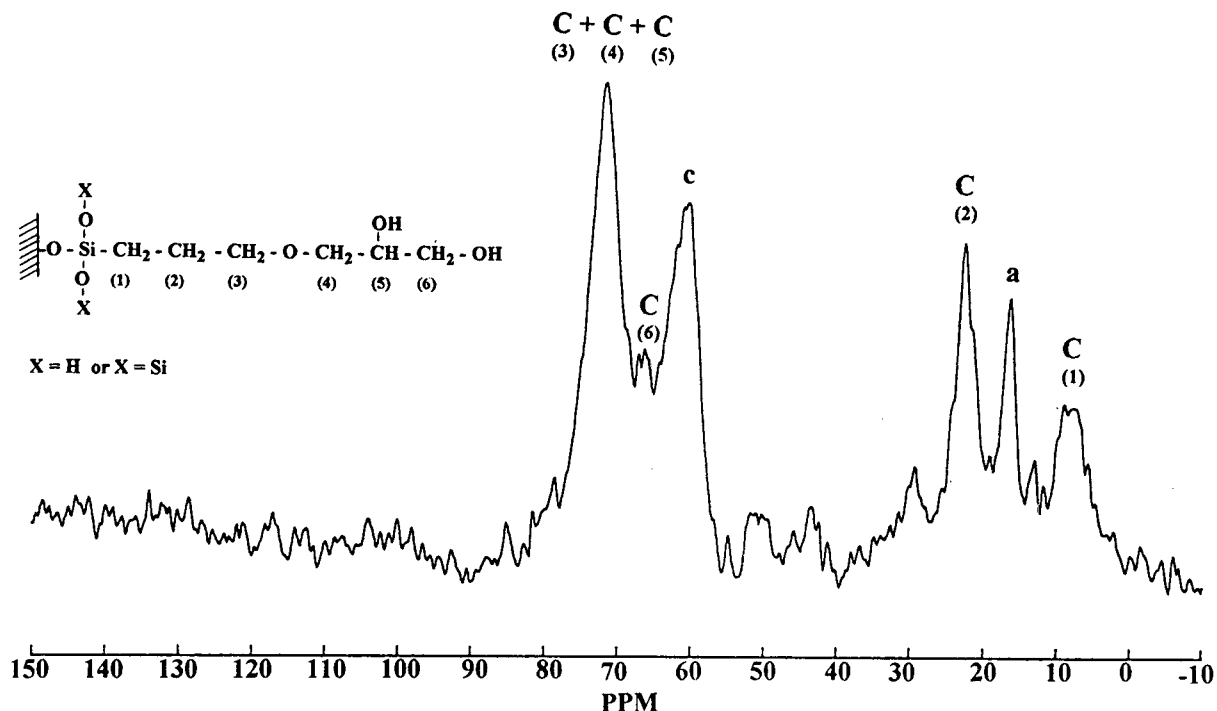
Fig. 4. ^{13}C CP-MAS-NMR spectrum of the product from the reaction of the hydride intermediate on Vydac silica and AGE after hydrolysis.

However, its appearance in the spectrum proves the success of the hydrosilation reaction.

Fig. 8 shows the ^{13}C CP-MAS-NMR spectrum for the same two products. A peak for the Si–C linkage can be seen at 12 ppm in the Davisil spectrum (A). The remainder of the methylene groups appear between 22–32 ppm. Two peaks for the–CH_2–OH and–CH(OH) groups appear at 67 and 72 ppm, respectively. In fact, with the shoulder on the 32 ppm peak, resonances for each of the eight carbons in the molecule can be identified. The Davisil spectrum contains two extra peaks at 124 and 132 ppm. These peaks are in the olefinic region of the spectrum and give further support to the conclusion that some of the 7-OD bonds via one of the alcohol groups to form an esterification product. In some cases two additional peaks can be observed at 113 and 136 ppm suggesting that perhaps bonding can occur at either alcohol group. Spectrum B in Fig. 8 is the spectrum which results from bonding 7-OD to Vydac hydride. The same peaks enumerated
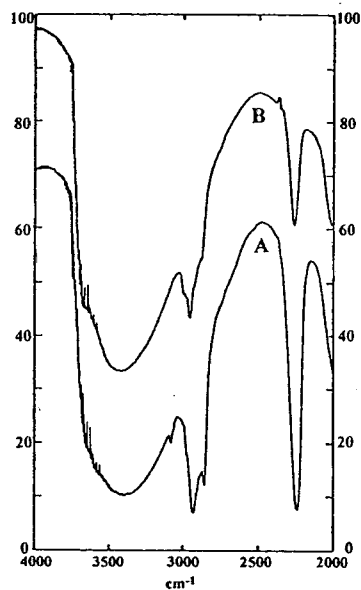
Fig. 5. Partial DRIFT spectra of the products from the reaction of the hydride intermediate and 7-OD on Davisil (A) and Vydac (B) silicas.
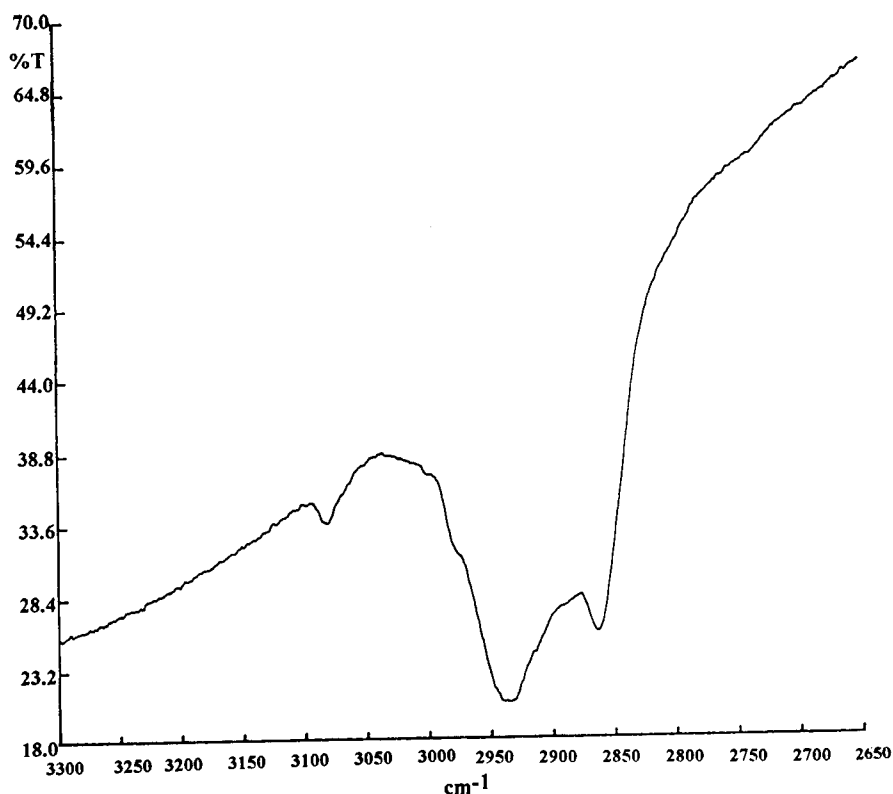
Fig. 6. Partial DRIFT spectrum of C–H stretching region of the product from the reaction of the hydride intermediate on Davisil and 7-OD. T = Transmittance.

above which document the success of the hydrosilation reaction can be identified in this spectrum as well. In addition, the two peaks for the ethoxy moiety can be seen at 16 and 60 ppm. However, no peaks are seen in the olefinic portion of the spectrum. This confirms the results obtained by DRIFT for the two materials which indicate some bonding through one of the alcohols for Davisil but not for Vydac.

3.3. Catalyst reduction

A known complication of the hydrosilation reaction for the bonding of olefins to hydride silica is the possible reduction of the catalyst to the elemental state and resulting deposition on the surface [13]. At high concentrations of catalyst (100 mM), the surface of Vydac and Davisil silica became gray for both the AGE and the 7-OD reactions. Fig. 9 shows the ^{29}Si CP-MAS-NMR spectrum of the 7-OD reaction product on Vydac hydride in the presence of high catalyst concentration. A large decrease in the Si–H peak at −85 ppm can be seen. The material has a very noticeable grey color which lends further support to the assumption that some reduction of Pt(IV) to Pt(0) has taken place. Absolute confirmation of this result was determined from the electron spectrometry for chemical analysis (ESCA) spectrum of this and other gray products which were obtained from AGE and 7-OD reactions at high catalyst concentration. Quantitative measurements of Pt on the surface ranged from 0.01 to 0.12 atom%. Therefore, all subsequent syntheses were done with 10 mM of catalyst and in each case the products obtained were white with no detectable Pt on the surface by ESCA.
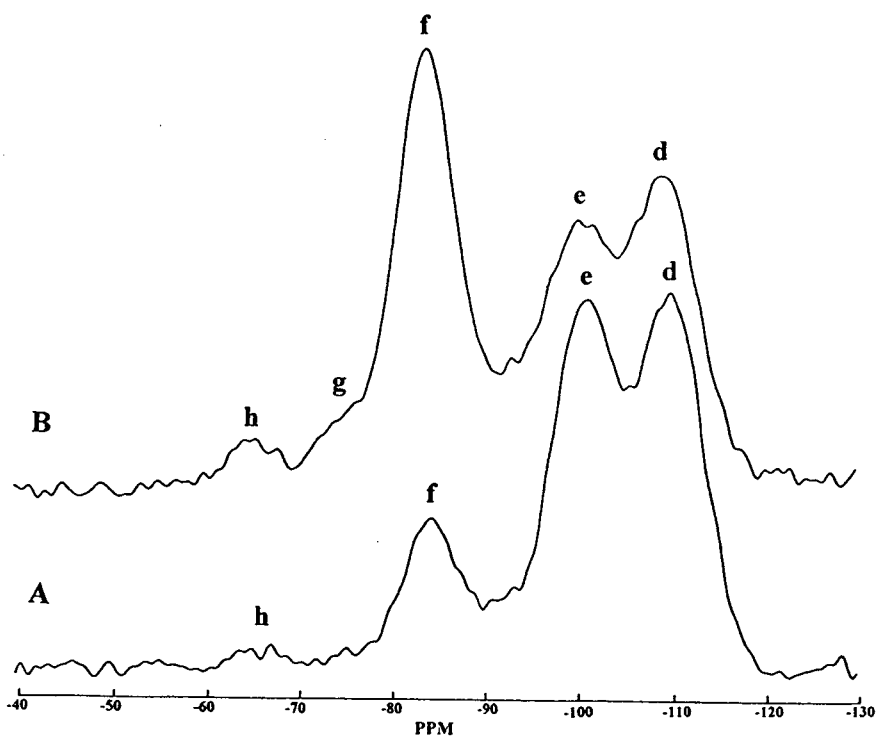
Fig. 7. ^{29}Si CP-MAS-NMR spectra of the products from the reaction of the hydride intermediate and 7-OD on Vydac (A) and Davisil (B) silicas.
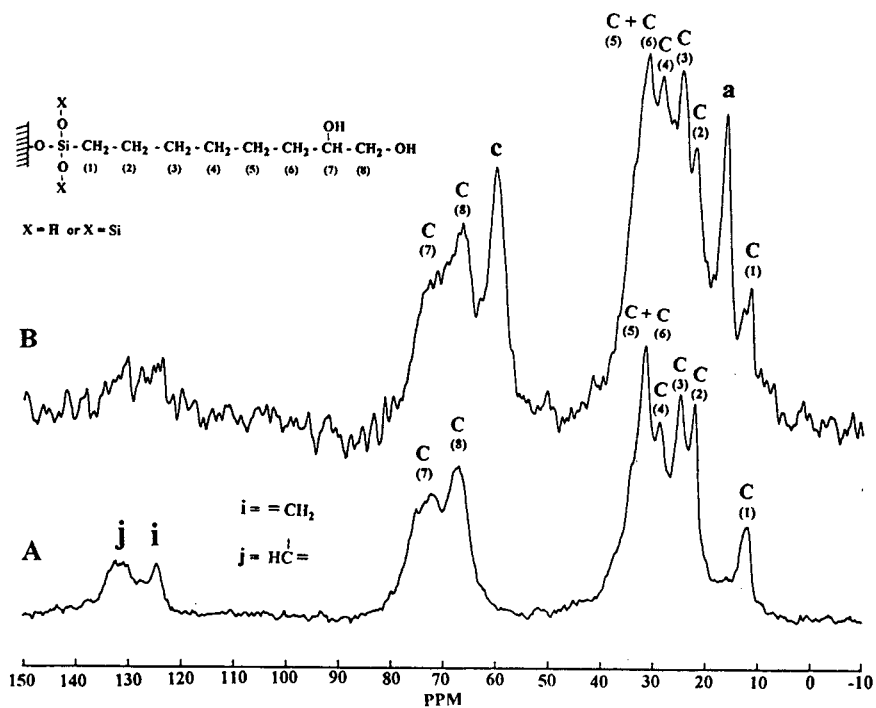
Fig. 8. ^{13}C CP-MAS-NMR spectra of the products from the reaction of the hydride intermediate and 7-OD on Davisil (A) and Vydac (B) silicas.
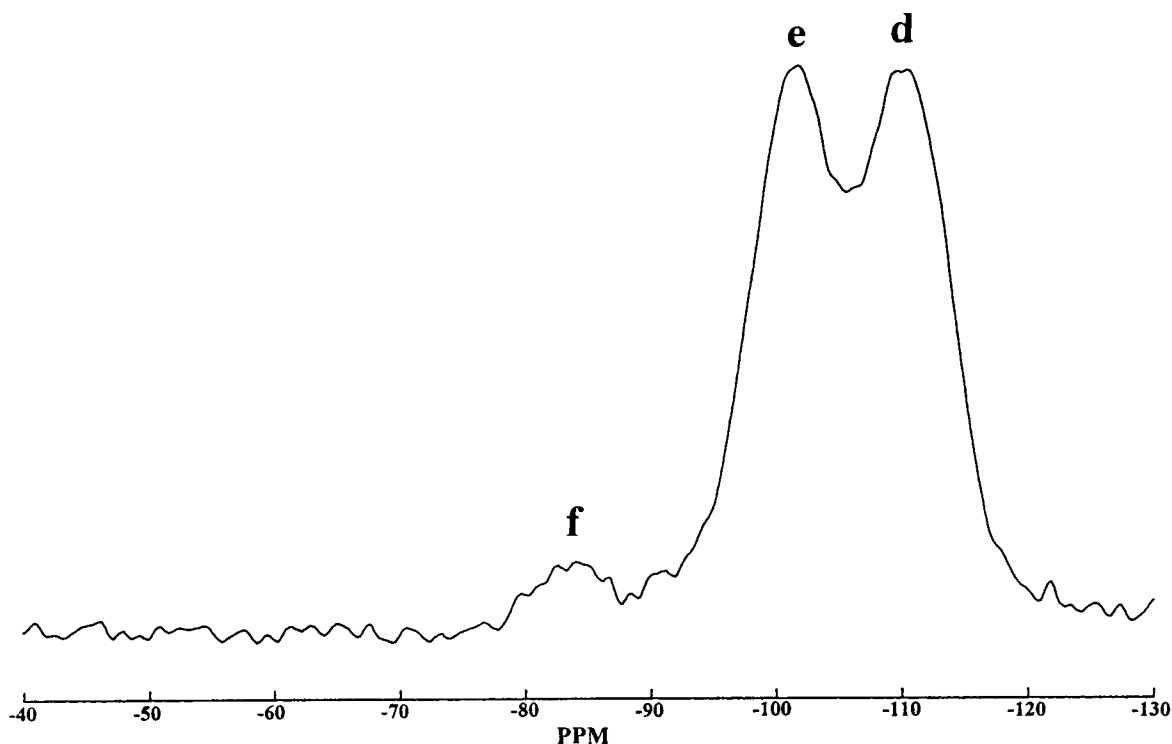
Fig. 9. ^{29}Si CP-MAS-NMR spectrum of the product from the reaction of the hydride intermediate and 7-OD on Vydac silica in the presence of high catalyst concentration.

In conclusion, the above synthetic pathways provide two options for synthesizing potentially uniform and reproducible bonded diol phases for chromatography. Both proceed through the hydride intermediate and via hydrosilation under transition metal catalysis to the final product. Each method is successful if the catalyst concentration is not too high to prevent reduction and deposition of the reduced metal on the surface. The two products are slightly different in that the AGE material has an ether linkage between two three-carbon chains while the 7-OD phase contains only methylene groups until the final two carbons which have the alcohol groups. In the latter reaction, the degree of hydrophobicity vs. hydrophilicity could be controlled by varying the length of the hydrocarbon chain in the original olefin. The only complication of the 7-OD reaction appears to be the potential competition from bonding through the alcohol groups (esterification). However, it appears that if the original silica contains very few unassociated (free) silanols then this reaction is minimized, i.e. undetectable by either DRIFT or ^{13}C CP-MAS-NMR analysis.

Acknowledgements

References

[1] F.R. Regnier and R. Noel, *J. Chromatogr. Sci.*, 14 (1976) 316.

[2] D.E. Schmidt, Jr., R.W. Giese, D. Conron and B.L. Karger, *Anal. Chem.*, 52 (1980) 177.

[3] S.H. Chang, K.M. Gooding and F.E. Regnier, *J. Chromatogr.*, 120 (1976) 321.

[4] R.R. Walters, *J. Chromatogr.*, 249 (1982) 19.

[5] T. Ohta, T. Inoue and Y. Fukumoto, *Chromatographia*, 30 (1990) 410.

[6] D.J. Phillips, B. Bell-Alden, M. Cava, E.R. Grover, W.H. Mandeville, R. Mastico, W. Sawlivich, G. Vella and A. Weston, *J. Chromatogr.*, 536 (1991) 95.

[7] T. Ohta, K. Ishimura and S. Takitani, *Chromatographia*, 33 (1992) 113.

[8] G.R. Bogart, D.E. Leyden, T.M. Wade, W. Schafer and P.W. Carr, *J. Chromatogr.*, 483 (1989) 209.

[9] B. Porsch, *J. Chromatogr.*, 653 (1993) 1.

[10] C. Chu, E. Jonsson, M. Auvinen, J.J. Pesek and J.E. Sandoval, *Anal. Chem.*, 65 (1993) 808.

[11] K.K. Unger, *Porous Silica*, Elsevier, Amsterdam, 1979.

[12] G. Berendsen and L. De Galan, *J. Liq. Chromatogr.*, 1 (1978) 561.

[13] J.E. Sandoval and J.J. Pesek, *Anal. Chem.*, 63 (1991) 2634.

[14] J.J. Pesek, J.E. Sandoval and M. Su, *J. Chromatogr.*, 630 (1993) 95.

[15] E. Bayer, K. Albert, J. Reiners, M. Nieder and D. Muller, *J. Chromatogr.*, 264 (1983) 197.

[16] A. Tuel, H. Hommel, A.P. Legrand, M.F. Gonnard, E. Mincsovics and A.M. Siouffi, *J. Chim. Phys.*, 89 (1992) 477.

ELSEVIER

Resolution of racemic drugs on a new chiral column based on silica-immobilized cellobiohydrolase
Characterization of the basic properties of the column

Jörgen Hermansson*, Anders Grahn

ChromTech AB, Box 6056, S-129 06 Hägersten, Sweden

Abstract

The basic properties of a new chiral column based on silica-immobilized cellobiohydrolase 1 (CBH 1) were studied using 32 different drugs and endogenous compounds as model compounds. All the compounds were resolved with high separation factors and high resolution. The stability of the CBH 1 column was studied by pumping two different mobile phases through the column, one of pH 6 and the other of pH 7. A 34-l volume of mobile phase, corresponding to about 38 000 column volumes, was pumped through the column. The enantio-selectivity, the resolution, the separation efficiency and the retention of the test compounds were unchanged after the passage of 38 000 column volumes of mobile phase. The most important tool for affecting the enantioselectivity, resolution and retention is the pH of the mobile phase. Uncharged modifiers such as acetonitrile and 1- and 2-propanol were also used. Unique effects were obtained for a series of drugs and endogenous compounds, related to epinephrine, by increasing the 2-propanol concentration in the mobile phase from 0 to 20%. The capacity factor for the most retained enantiomer could be selectively increased by increasing the 2-propanol concentration in the mobile phase, resulting in a substantial improvement in the enantioselectivity. For one compound, octopamine, the capacity factor of the most retained enantiomer increased by almost 100%, resulting in a very large increase in the separation factor from 1.9 to 3.3. The resolution increased from 4.8 to 7.8. The relationship between the molecular structure and the enantioselectivity was also studied. The data indicate that the steric bulk on a basic nitrogen influences the enantioselectivity to a large extent, in addition to the degree of ring substitution. Further, the Chiral-CBH column seems to be very well suited for the resolution of preferentially basic drugs of different character.

1. Introduction

Proteins of different types are an important group of compounds in nature. Enzymes transform a substrate into a product, drugs and endogenous compounds are transported in the body as complexes with proteins such as albumin

---

* Corresponding author.

and α_1-acid glycoprotein (AGP). The proteins are asymmetric molecules which means that they can bind a low-molecular-mass ligand, such as a drug, stereoselectively. This property of proteins has been utilized for the chromatographic resolution of enantiomers by immobilization of the protein on a matrix. One of the most frequently used chiral stationary phases (CSPs) today is a column based on the protein AGP. The extreme-

ly broad applicability of the AGP column has been documented in a large number of publications (e.g., [1–6]). This broad applicability of AGP can be assigned to the unique primary, secondary and tertiary structure that nature has built into the protein.

Other protein phases with a narrower applicability are bovine serum albumin and the human serum albumin columns [7,8]. The latest addition to this category of CSPs is the ovomucoid column [9] and a column based on cellobiohydrolase 1 [10–13].

Cellobiohydrolase 1 (CBH 1) is a protein with very interesting properties concerning the ability to discriminate between enantiomers. CBH 1 is very well suited for use as a chiral selector in liquid chromatographic columns owing to the very high stability of the enzyme. The reason for this high stability might be that nature designed the enzyme to work extracellularly. CBH 1 has a molecular mass of 64 000. It consists of 497 amino acid residues, stabilized by 12 disulphide bridges. The carbohydrate content is about 6%. Recently the three-dimensional structure of the active site of CBH 1 has been elucidated by X-ray crystallography [14]. The binding site is a tunnel with the dimensions $4 \times 7 \times 40$ Å. There are seven acidic amino acid residues, four tryptophan residues and also tyrosine, serine, threonine, arginine and histidine lining the tunnel. This gives the prerequisites for obtaining stereoselective binding of a broad range of chiral solutes.

In this paper, a new chiral column, Chiral-CBH, based on cellobiohydrolase, is described. The properties of the Chiral-CBH column differ significantly from those of the CBH 1 column referred to above, owing to large differences in the surface chemistry of the base silica and large differences in the immobilization technique.

2. Experimental

2.1. Apparatus

The chromatographic system consisted of an LKB Model 2150 HPLC pump (Pharmacia,

Uppsala, Sweden), a Kontron (Eching/Munich, Germany) HPLC 360 autosampler equipped with a 20-μl loop and a Spectra 100 variable-wavelength UV detector (Spectra-Physics, San Jose, CA, USA). The experimental data were collected and analysed on a Kontron 450 MT2 data system, which also controlled the autosampler. pH was measured with an Orion SA520 pH meter connected to a Ross Model 8102 combination electrode. The column used was a commercially available Chiral-CBH (ChromTech, Hägersten, Sweden) with the dimensions 100×4.0 mm I.D. (5-μm particles). A Chiral-CBH guard column (ChromTech) was coupled in front of the analytical column. The dimensions of the guard column were 10×3.0 mm.

2.2. Chemicals

2-Propanol and acetonitrile, of HPLC grade, were obtained from Lab-Scan (Dublin, Ireland). The drug compounds were obtained from different companies. The salts used for the preparation of mobile phases were of analytical-reagent grade.

2.3. Chromatographic conditions

The experiments were performed at room temperature. The flow-rate was 0.9 ml/min and the UV detector was operated at 210 nm. The sample concentrations were in the range 0.02–0.05 mg/ml. Mobile phases containing phosphate buffers were prepared from sodium dihydrogenphosphate and the pH was adjusted with sodium hydroxide. Disodium EDTA was added before adjusting the pH. Thereafter an appropriate amount of the organic modifier was added, followed by water to the final volume. Mobile phases containing acetate buffers were prepared from sodium acetate and the pH was adjusted with acetic acid. Disodium EDTA was added before adjusting the pH. Thereafter an appropriate amount of the organic modifier was added, followed by water to the final volume. The hold-up volume (V_m) was measured by injection of distilled water.

3. Results and discussion

3.1. Binding of solutes to CBH 1

CBH 1 is a very stable enzyme with a molecular mass of 64 000 and an isoelectric point of 3.9. This means that the CBH 1 molecule has a net negative charge in the pH range normally used in chromatographic experiments (4–7). In the chromatography of basic compounds in this pH range, ionic binding between the solute and the protein is an important type of interaction, in order to retain basic solutes on this column. Ionic binding has, in addition to ion-pair adsorption (ion pairs are neutral species and are bound by hydrophobic interaction and hydrogen bonding), been demonstrated to be very important for the retention of basic drugs on an AGP column [1,2,6]. From the three-dimensional structure of CBH 1 it can be seen that the binding site is a tunnel with the dimensions 4 × 7 × 40 Å [14]. The binding site contains seven acidic amino acid residues and aspartic and glutamic acid [14]. This means that it seems very likely that the ion-exchange mechanism is an important retention mechanism on the CBH 1 column. Further, the binding site (the tunnel) also contains four tryptophan residues and also other amino acid residues such as tyrosine, serine, threonine, arginine and histidine, which means that both hydrogen bonding and hydrophobic interaction are very important for the retention of the solutes. From the above it follows that the enantioselectivity and the retention of the solutes can be affected by the pH of the mobile phase and also the nature and concentration of an uncharged organic modifier such as 1- or 2-propanol or acetonitrile. pH is a very important factor since it affects both the degree of charge of the solute and the chiral bonding properties of the protein. An uncharged modifier can also strongly influence the stereoselective binding of the solutes.

3.2. Retention, enantioselectivity and resolution versus the concentration of uncharged organic modifier

Unique effects were observed on the retention, the enantioselectivity and the resolution for a series of epinephrine analogues on increasing the 2-propanol concentration, as demonstrated in Table 1. As can be seen, the capacity factor of the most retained octopamine enantiomer is almost doubled by increasing the 2-propanol concentration from 0 to 20%, whereas the capacity factor for the first-eluted enantiomer is almost unchanged. This means that the separation factor and the resolution increase substantially on increasing the 2-propanol concentration. The separation factor of octopamine increases from 1.90 to 3.32 and the resolution increases from 4.8 to 7.8. Similar effects, but not so dramatic, were also observed for the analogues norepinephrine, normethanephrine, phenylethanolamine and epinephrine, as demonstrated in Table 1. To our knowledge, such effects have never been observed previously on any other chiral column. From Table 1 it can be seen that the enantioselectivity for all epinephrine analogues increases on increasing the 2-propanol concentration. The largest increase in the capacity factors was observed for compounds containing a primary amino group, whereas the capacity factors for metanephrine and epinephrine, containing a secondary amino group, demonstrated a deviatiant behaviour. The capacity factors for both enantiomers of metanephrine decreased slightly and a very small increase in the capacity factor was observed for the last-eluted enantiomer of epinephrine. Replacing 2-propanol with 5% acetonitrile in the same buffer resulted in lower enantioselectivity and lower retention for all compounds.

The unique behaviour of increasing retention with increasing 2-propanol concentration cannot easily be explained. A reasonable explanation could be that the addition of 2-propanol induces a reversible change of the secondary structure, which means that the binding site changes character and exposes other amino acid residues in the binding site (the tunnel), giving a higher affinity of the last-eluted enantiomer. As the retention of the least retained enantiomer is only slightly affected, it is reasonable to assume that the enantiomers are retained at different locations, i.e., different subsites, in the 40 Å long

Table 1
Effects of uncharged organic modifiers on k', α and R_s for epinephrine analogues

Modifier concentration	Octopamine				Metanephrine				Normetanephrine				Epinephrine				Norepinephrine				Phenylethanolamine			
	k'_1	k'_2	α	R_s	k'_1	k'_2	α	R_s	k'_1	k'_2	α	R_s	k'_1	k'_2	α	R_s	k'_1	k'_2	α	R_s	k'_1	k'_2	α	R_s
None	3.04	5.79	1.90	4.83	2.83	3.30	1.17	0.91	3.68	6.30	1.72	3.67	2.25	2.97	1.32	1.62	3.60	6.29	1.74	3.75	–	–	–	–
0.5% 2-PrOH	3.01	6.04	2.01	5.23	2.68	3.23	1.20	1.08	3.57	6.43	1.80	3.98	2.19	2.98	1.36	1.83	3.53	6.41	1.82	3.96	2.25	2.96	1.31	1.95
5% 2-PrOH	3.03	7.57	2.50	6.81	2.40	3.02	1.26	1.43	3.28	6.68	2.04	4.82	2.08	3.08	1.48	2.26	3.40	7.19	2.12	4.99	2.18	3.06	1.40	2.44
10% 2-PrOH	3.02	8.61	2.85	7.52	2.20	2.82	1.29	1.57	3.06	6.84	2.23	5.20	2.04	3.12	1.53	2.48	3.27	7.60	2.32	5.51	2.19	3.19	1.46	2.64
20% 2-PrOH	3.31	11.0	3.32	7.80	2.19	2.83	1.29	1.52	3.03	7.42	2.45	4.87	2.27	3.37	1.49	2.21	3.43	8.40	2.45	5.30	2.51	3.74	1.49	2.77
5% CH$_3$CN	2.35	4.27	1.82	3.48	1.95	2.24	1.15	0.83	2.60	3.93	1.51	2.19	1.76	2.34	1.33	1.74	2.68	4.37	1.63	3.07	1.83	2.27	1.24	1.56

Column: Chiral-CBH, 100×4.0 mm I.D. Mobile phase: 10 mM sodium phosphate buffer (pH 6.0)–50 μM disodium EDTA, having different concentrations of modifiers.
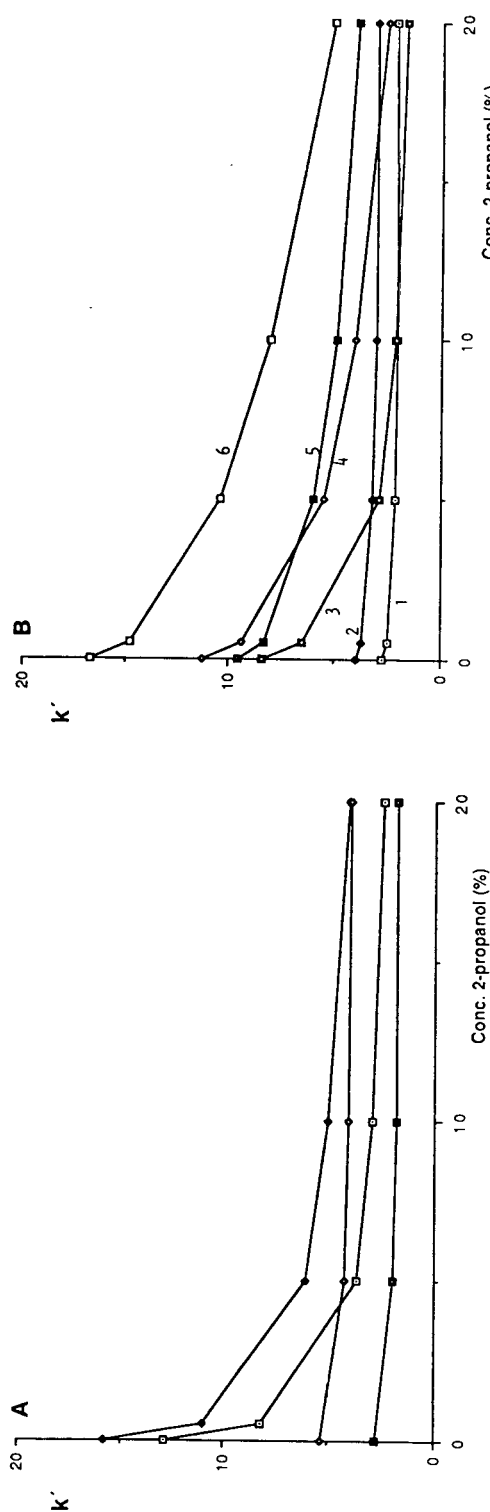
Fig. 1. Capacity factors, k', versus 2-propanol concentration. Column, Chiral-CBH (100×4.0 mm I.D.); mobile phases, 10 mM sodium phosphate buffer (pH6.0)–50 μM disodium EDTA containing 0–20% 2-propanol. (A) Samples: □ = k'_1 talinolol; ■ = k'_2 talinolol; ◆ = k'_1 atenolol; ◇ = k'_2 atenolol. (B) Samples: 1 = k'_1 cathinone; 2 = k'_2 cathinone; 3 = k'_1 laudanosine; 4 = k'_2 laudanosine; 5 = k'_1 dobutamine; 6 = k'_2 dobutamine.

binding site (the tunnel) of CBH 1. The decrease in the capacity factor observed for metanephrine with increasing modifier concentration might be the result of competition for hydrogen bonding and/or hydrophobic interaction.

When the unique properties reported above were observed, we were also interested in establishing whether that kind of effect could be observed for other classes of compounds with other characteristics. Five different basic compounds of different character were included in the study: the amino ketone cathinone, two amino alcohols, the two well known β-blockers atenolol and talinolol, the secondary amine dobutamine and the tertiary amine laudanosine. The structures are shown in Table 3.

Fig. 1A and B demonstrate the influence of the 2-propanol concentration on the retention of the five model compounds. As can be seen in Fig. 1A, the capacity factors of the enantiomers of atenolol are affected to a very limited extent by increasing the concentration of 2-propanol to 20%, i.e., a behavior similar to that observed for metanephrine. The retention of the talinolol enantiomers is strongly reduced. From Fig. 1B it can be seen that the amino ketone cathinone demonstrates a behaviour similar to that of

atenolol. The capacity factor of cathinone is almost constant in the concentration range of 2-propanol used in the study. This behaviour may indicate that hydrophobic interaction is not very important in the retention process for these solutes. The capacity factors of dobutamine and laudanosine are reduced to a relatively large extent. It is very interesting that the capacity factor of the most retained enantiomer of dobutamine decreases more rapidly than that of the least retained enantiomer. For laudanosine the reversed situation occurred.

The relative influence of a modifier on the retention of the enantiomers of a solute is very important for the enantioselectivity, as is clearly demonstrated in Fig. 2. As can be seen, the enantioselectivity of atenolol, talinolol and laudanosine increases strongly at 2-propanol concentrations up to 10%. At higher concentrations the enantioselectivity is almost unaffected for atenolol and talinolol. Similar effects have been observed previously on the AGP column for carboxylic acids such as tiaprofenic acid and some weak acids such as methylphenobarbital and some barbituric acid derivatives [2–4], and also on the CBH 1 column for some basic drugs [11]. The separation factor for dobutamine is
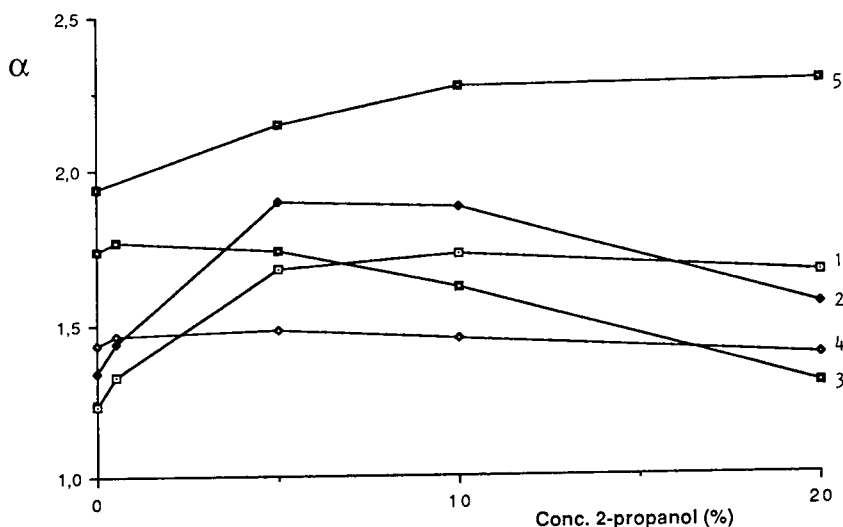
Fig. 2. Effect on the separation factor, α, of increasing the 2-propanol concentration. Conditions as in Fig. 1. Samples: 1 = talinolol; 2 = laudanosine; 3 = dobutamine; 4 = cathinone; 5 = atenolol.
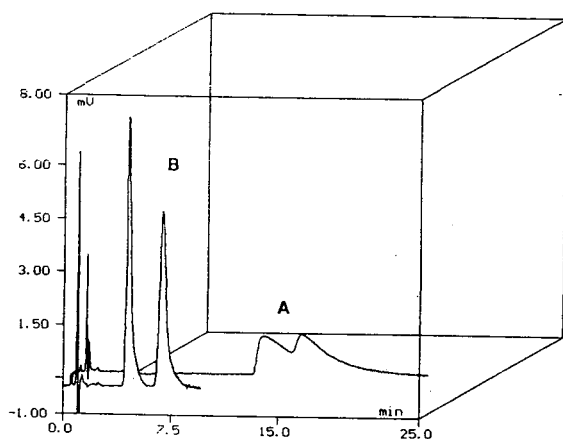
Fig. 3. Improvement of the chromatographic performance by addition of 2-propanol to the mobile phase. Column, Chiral-CBH (100 × 4.0 mm I.D.); sample, talinolol; mobile phase, (A) 10 mM sodium phosphate buffer (pH 6.0)–50 μM disodium EDTA and (B) 5% 2-propanol in 10 mM sodium phosphate buffer (pH 6.0)–50 μM disodium EDTA.

reduced on increasing the modifier concentration, whereas the separation factor for cathinone remains almost constant. The resolution increases for all compounds with increasing 2-propanol concentration up to 5%. For two of the compounds, laudanosine and atenolol, the res-

olution continues to increase up to a 2-propanol concentration of 10% in the mobile phase.

The reason for the very strong increase in the resolution on increasing the 2-propanol concentration is that the enantioselectivity increases with increasing modifier concentration. However, the separation efficiency and the peak symmetry are also highly affected by the 2-propanol concentration, as can be seen in Fig. 3. Chromatography of talinolol using a phosphate buffer of pH 6 without 2-propanol results in a low separation efficiency, low enantioselectivity and only partial resolution. However, addition of 5% 2-propanol gives baseline resolution, as can be seen from Fig. 3. This is the result of higher enantioselectivity and better chromatographic performance when 2-propanol was added to the mobile phase.

The reason for the better chromatographic performance obtained for talinolol with 2-propanol in the mobile phase might be due to masking of a high affinity site by 2-propanol. This is supported by the fact that the retention is drastically reduced in the concentration range where the improvement of the chromatographic performance is obtained. It is very interesting to compare the behaviour of the two analogues

Table 2
Chromatographic properties of structural analogues of epinephrine

HO—CHCH2NH—R4

R1, R3, R2 (structure)
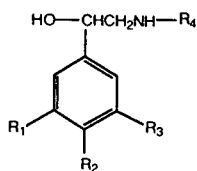
Solute	R_1	R_2	R_3	R_4	k'_1	α	R_s
Phenylethanolamine	H	H	H	H	2.18	1.40	2.44
Octopamine	H	OH	H	H	3.03	2.50	6.81
Norepinephrine	H	OH	OH	H	3.40	2.12	4.99
Epinephrine	H	OH	OH	CH₃	2.08	1.48	2.26
Normetanephrine	H	OH	CH₃O	H	3.28	2.04	4.82
Metanephrine	H	OH	CH₃O	CH₃	2.40	1.26	1.43
Isoprenaline	H	OH	OH	CH(CH₃)₂	1.80	1.14	0.75
Metaproterenol	OH	H	OH	CH(CH₃)₂	2.80	1.00	–
Terbutaline	OH	H	OH	C(CH₃)₃	1.61	1.18	0.86

Column: Chiral-CBH, 100 × 4.0 mm I.D. Mobile phase: 5% 2-propanol in 10 mM sodium phosphate buffer (pH 6.0)–50 μM disodium EDTA.

talinolol and atenolol with increasing 2-propanol concentration in the mobile phase (see Fig. 1A). The retention of atenolol is affected to a very limited extent by the 2-propanol concentration in the range 0–20% compared with the effect obtained for talinolol. This may indicate that they are bound in a different way, or to different subsites in the 40 Å long binding site.

3.3. Influence of solute structure on retention and enantioselectivity

A series of analogues of epinephrine were utilized as model compounds in order to study the influence of the molecular structure on the enantioselectivity and the retention. The compounds were chromatographed using a mobile phase of 5% 2-propanol in 10 mM sodium phosphate buffer (pH 6.0) containing 50 μM disodium EDTA. The results are summarized in Table 2. Phenylethanolamine, octopamine and norepinephrine have a primary amino group in the side-chain but they differ in the ring substitution. Phenylethanolamine has an unsubstituted ring, octopamine has a hydroxy group in the *para* position and norepinephrine has two hydroxy groups in *meta* and *para* positions to the side-chain. The retention increases in the order phenylethanolamine $(k'_1 = 2.18) <$ octopamine $(k'_1 = 3.03) <$ norepinephrine $(k'_1 = 3.40)$, i.e., the retention increases with increasing number of hydroxy groups in the aromatic ring. This indicates that the hydroxy groups are involved in hydrogen bonding with hydrogen bonding groups in the binding site. The highest enantioselectivity of the above three compounds is obtained for octopamine with an α value of 2.5 and a resolution, R_s, of 6.8.

Comparison of norepinephrine with epinephrine, with a methyl group on the basic nitrogen, demonstrates that the introduction of the methyl group strongly influences both the retention and the enantioselectivity. The separation factor is reduced from 2.12 to 1.48. The retention is also reduced by about 40%. The same effects were observed for the pair normetanephrine and metanephrine. Higher enantioselectivity and retention were observed for normetanephrine,

which has a primary amino group in the side-chain. This might be due to the higher affinity in the ionic binding because the positively charged nitrogen atom will come closer to the anionic amino acid residue in the binding site. It can be seen that an even larger substituent, i.e., an isopropyl or an isobutyl group, instead of a methyl group on the basic nitrogen gives a further decrease in both the enantioselectivity and the retention. However, these observations are valid for this series of compounds and it is probably too early to draw any general conclusions concerning the relationship between the molecular structure and the enantioselectivity. More data have to be collected. For example, it has been observed for alprenolol and its de-sisopropyl analogue that the dealkylated analogue gave a lower enantioselectivity and higher retention [11]. However, as can be seen from Table 2, most of the compounds can be resolved with very high R_s values.

By chromatography of a series of analogues with different steric bulk on a basic nitrogen on the AGP column, it was observed that the enantioselectivity for many series of compounds increases with increasing steric bulk on the basic nitrogen atom [3,15].

3.4. Applicability of the Chiral-CBH column

The applicability of the Chiral-CBH column was tested by injecting a large number of drug compounds of different character. The molecular structures, the capacity factors of the least retained enantiomers, the separation factors (α) and the resolution (R_s) are summarized in Table 3. As can be seen, high enantioselectivity and very high resolution were obtained. Fig. 4A–D show chromatograms of four different types of drugs; atenolol, cimetidine sulphoxide, norepinephrine and toliprolol, resolved on a CBH 1 column. As can be seen, a very good chromatographic performance was obtained. With the present knowledge, our data indicate that the CBH 1 column is very useful for the resolution of a wide variety of basic drug compounds. In addition, among the basic drugs the column demonstrated excellent properties for amino
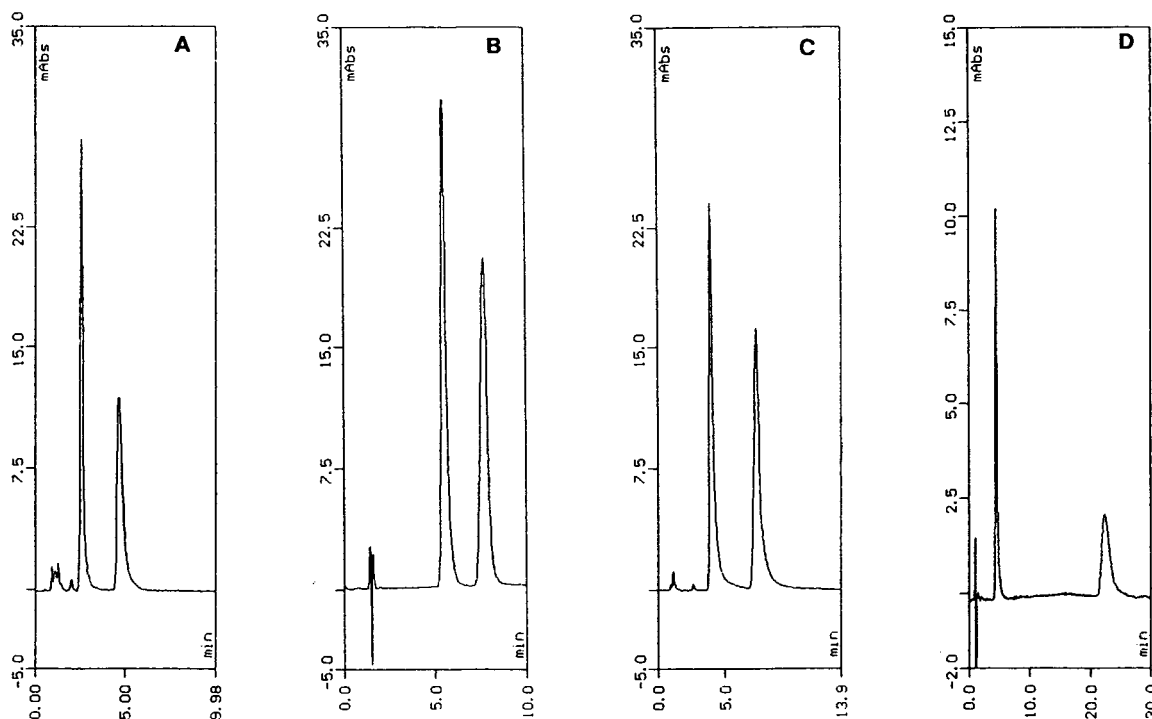
Fig. 4. Resolution of enantiomers on the Chiral-CBH column. (A) Sample, atenolol; column, Chiral-CBH (100 × 4.0 mm I.D.); mobile phase, 5% 2-propanol in 10 mM sodium phosphate buffer (pH 6.0)–50 μM disodium EDTA. (B) Sample, cimetidine sulphoxide; column, Chiral-CBH (150 × 4.0 mm I.D.); mobile phase, 10 mM sodium phosphate buffer (pH 6.0)–50 μM disodium EDTA. (C) Sample, norepinephrine; conditions as in (A). (D) Sample, toliprolol; conditions as in (A).

alcohols, which also was demonstrated on the first-generation CBH 1 column [10–13]. All the β-blockers that were tested were easily resolved, as can be seen from Table 3.

3.5. Stability of the CBH column

Stability studies were performed using a 100 × 4.0 mm I.D. column with a guard column (10 × 3.0 mm I.D.) in front of the analytical column. Two different mobile phases were used: 5% 2-propanol in 10 mM phosphate buffer (pH 6.0) containing 50 μM disodium EDTA and 5% 2-propanol in 10 mM phosphate buffer (pH 7.0) containing 50 μM disodium EDTA. The reason for choosing these mobile phases for the stability tests was that most of the separations that have been performed on the CBH column so far have been performed in the pH range 5–7. EDTA was added to the mobile phase in order to

complex the metal ions present. The mobile phase was pumped at a flow-rate of 0.9 ml/min, to obtain the same linear velocity as on a column having an I.D. of 4.6 mm and using a flow-rate of 1.0 ml/min.

In total 34 l of mobile phase were pumped through the column at pH 6.0, which corresponds to about 38 000 column volumes. It also corresponds to the continuous use of the column for 8 h per day, 5 days per week for 16 weeks. The guard column was exchanged after every 4 l of mobile phase (corresponding to about 4500 column volumes) had passed through the column and thereafter the three test compounds, cathinone, metoprolol and octopamine, were injected.

Fig. 5A shows the capacity factors of the most retained enantiomers of the three test compounds and Fig. 5B shows the resolution, R_s, versus the volume of mobile phase pumped

Table 3
Examples of baseline-resolved drugs and endogenous compounds

Substance	Formula	k_1'	α	R_s	Mobile phase[a]
Acebutolol		3.77	4.16	7.33	1
Atenolol		1.69	2.26	4.19	2
Betaxolol		3.82	3.63	6.45	1
Bisoprolol		3.21	3.95	7.82	2
Carbuterol		2.92	1.66	4.06	4
Cathinone		2.36	1.54	3.29	2
Cimetidine sulphoxide[b]		3.04	1.54	3.28	3
Dobutamine		5.62	1.78	3.41	2
Dropropizine		4.16	1.49	2.85	4
Epanolol		9.52	2.22	3.64	5

Table 3 (*continued*)
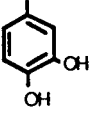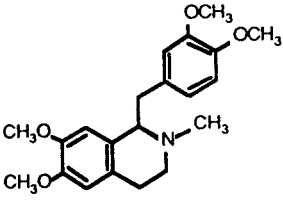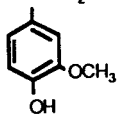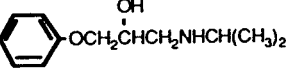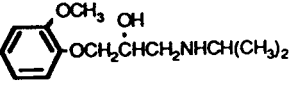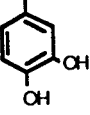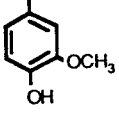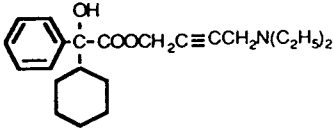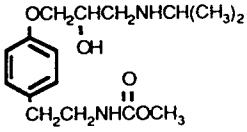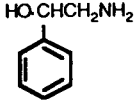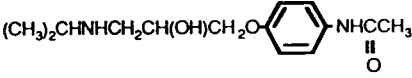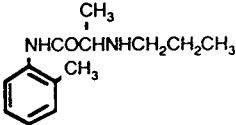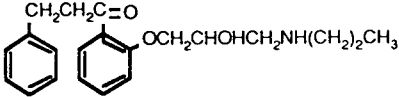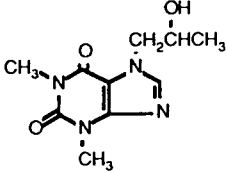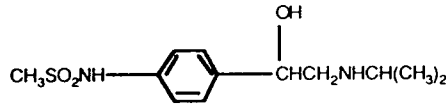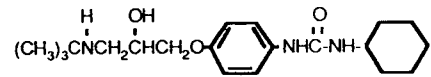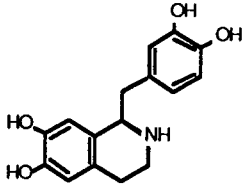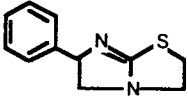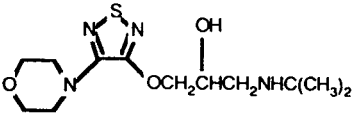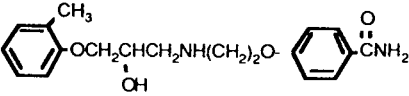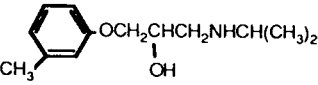
Substance	Formula	k'_1	α	R_s	Mobile phase[a]
Epinephrine	HO.CHCH$_2$NHCH$_3$	4.26	1.69	4.00	6
Laudanosine		2.82	2.00	3.00	2
Metanephrine	HO.CHCH$_2$NHCH$_3$	8.52	1.38	2.65	4
Metoprolol	CH$_3$OCH$_2$CH$_2$—OCH$_2$CHCH$_2$NHCH(CH$_3$)$_2$	3.59	3.16	7.15	2
Moprolol	OCH$_2$CHCH$_2$NHCH(CH$_3$)$_2$	4.72	1.91	4.21	2
Norepinephrine	HO.CHCH$_2$NH$_2$	3.40	2.12	4.99	2
Normetanephrine	HO.CHCH$_2$NH$_2$	3.28	2.04	4.82	2
Octopamine	HO.CHCH$_2$NH$_2$	3.03	2.50	6.81	2

Table 3 (*continued*)

Substance	Formula	k_1'	α	R_s	Mobile phase[a]
Oxybutynine		5.20	2.08	3.42	2
Pamatolol		4.01	2.31	5.34	2
Phenylethanolamine		4.62	1.54	3.26	6
Practolol		8.66	1.36	2.51	7
Prilocaine		3.16	1.48	2.90	4
Propafenone		7.06	2.04	4.10	1
Proxyphylline		1.22	1.63	2.47	4
Sotalol[b]		6.66	1.26	2.52	4
Talinolol		3.12	1.77	3.14	2

Table 3 (continued)

Substance	Formula	k_1'	α	R_s	Mobile phase[a]
Tetrahydropapaveroline		3.17	1.76	3.45	8
Tetramisole		2.45	1.58	3.29	9
Timolol		1.55	4.10	5.34	9
Tolamolol		3.46	2.07	3.68	5
Toliprolol		3.95	6.05	10.8	2

Column: Chiral-CBH, 100 × 4.0 mm I.D.

[a] Mobile phases: 1 = 5% 2-propanol in 10 mM sodium acetate buffer (pH 5.5)–50 μM disodium EDTA; 2 = 5% 2-propanol in 10 mM sodium phosphate buffer (pH 6.0)–50 μM disodium EDTA; 3 = 10 mM sodium phosphate buffer (pH 6.0)–50 μM disodium EDTA; 4 = 5% 2-propanol in 10 mM sodium phosphate buffer (pH 7.0)–50 μM disodium EDTA; 5 = 5% 2-propanol in 10 mM sodium acetate buffer (pH 5.0)–50 μM disodium EDTA; 6 = 5% 2-propanol in 10 mM sodium phosphate buffer (pH 6.5)–50 μM disodium EDTA; 7 = 5% acetonitrile in 10 mM sodium phosphate buffer (pH 7.0)–50 μM disodium EDTA; 8 = 5% acetonitrile in 10 mM sodium acetate buffer (pH 5.5)–50 μM disodium EDTA; 9 = 10% 2-propanol in 10 mM sodium phosphate buffer (pH 6.0)–50 μM disodium EDTA.

[b] Sotalol and cimetidine sulphoxide were resolved on a column with the dimensions 150 × 4.0 mm I.D.

through the column. As can be seen, both the retention and the resolution are almost constant. No deterioration of the column properties was observed. Most likely the column could be run with an additional 38 000 column volumes or more without a negative influence on the chromatographic properties of the analytical column. Fig. 6 shows two chromatograms of cathinone, obtained before the stability studies and after 38 000 column volumes of mobile phase had been pumped through the column.

As was mentioned above, an additional stability study was also performed using a higher pH of the buffer (7.0). The difference in this study was that the guard column was exchanged after the passage of 7–8 l of mobile phase. Cathinone and prilocaine were used as test compounds. Fig. 7A and B show the capacity factors of the most
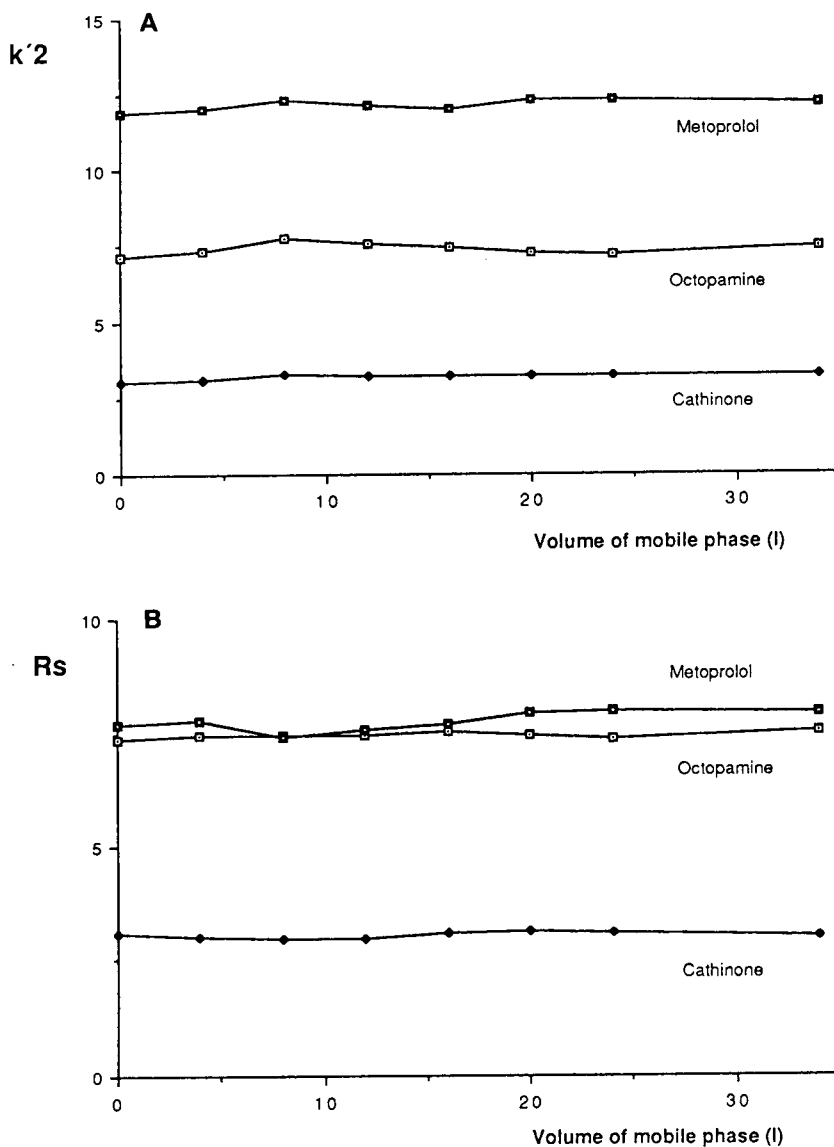
Fig. 5. (A) Capacity factors of the most retained enantiomer versus the volume of mobile phase pumped through the column. Column, Chiral-CBH (100×4.0 mm I.D.) with a Chiral-CBH guard column (10×3.0 mm I.D.); mobile phase, 5% 2-propanol in 10 mM sodium phosphate buffer (pH 6.0)–50 μM disodium EDTA. (B) Resolution versus the volume of mobile phase pumped through the column. Conditions as in (A).

retained enantiomer and the resolution, R_s, versus the volume of mobile phase pumped through the column. In total 22 l, corresponding to about 24 500 column volumes of mobile phase, were pumped in this stability study. As can be seen, the retention and the resolution of

the basic test compounds were almost constant and no negative influence on the chromatographic performance was observed. Therefore, it is reasonable to assume that there is no difference between the stability of the column at pH 6 and 7. The guard column is used in order to trap
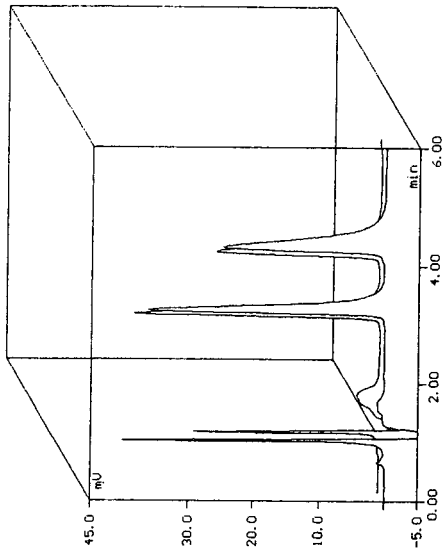
Fig. 6. Chromatograms of cathinone obtained before and after (front chromatogram) the passage of 34 l of mobile phase. Conditions as in Fig. 5A.
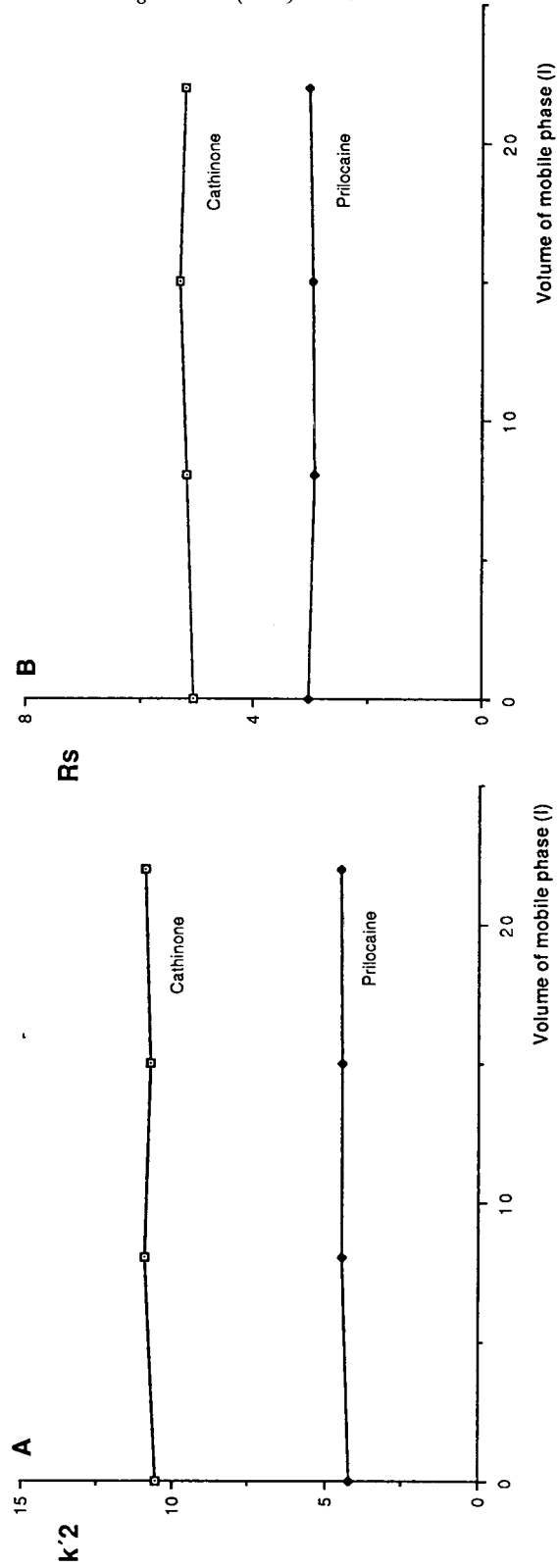
Fig. 7. (A) Capacity factors of the most retained enantiomer versus the volume of mobile phase pumped through the column. Column, Chiral-CBH (100×4.0 mm I.D.) with a Chiral-CBH guard column (10×3.0 mm I.D.); mobile phase, 5% 2-propanol in 10 mM sodium phosphate buffer (pH 7.0)–50 μM disodium EDTA. (B) Resolution versus the volume of mobile phase pumped through the column. Conditions as in (A).

compounds and metal ions that are bound to the column with high affinity.

On exchanging the guard column regularly the analytical column can be used for very long periods of time without a reduction of its chromatographic properties, as has been demonstrated in the stability studies. In the study performed at pH 6 the guard column was exchanged after 4 l of mobile phase. However, at pH 7 the guard column was exchanged after about 7–8 l of mobile phase. The stability studies demonstrated that the guard column should be replaced after about 7–8 l of mobile phase irrespective of the pH of the mobile phase. When 7–8 l of mobile phase had been pumped through the column, the resolution was decreased to some extent, which is the result of a small decrease in the separation efficiency caused by the guard column. However, the analytical column was totally unaffected by the large volumes of mobile phase that were pumped through the column, as is demonstrated in Figs. 5–7.

From this experiment, it can be concluded that the guard column traps something from the mobile phase that affects the migration of the sample on the guard column. Obviously, it must be present in the mobile phase at a very low concentration and it also has a very high affinity to the CBH 1 guard column. This conclusion is supported by the fact that the chromatographic properties of the analytical column are totally unaffected. As was mentioned above, it can be hypothesized that it is metal ions that are trapped on the guard column.

References

[1] J. Hermansson, *J. Chromatogr.*, 269 (1983) 71.
[2] J. Hermansson and G. Schill, in P.A. Brown and R.A. Hartwick (Editors), *High Performance Liquid Chromatography (Monographs on Analytical Chemistry Series)*, Wiley-Interscience, New York, 1988, pp. 337–374.
[3] J. Hermansson, *Trends Anal. Chem.*, 8 (1989) 251.
[4] J. Hermansson and I. Hermansson, *J. Chromatogr.*, 666 (1994) 181.
[5] I. Fitos, J. Visy, M. Simonyi and J. Hermansson, *J. Chromatogr.*, 609 (1992) 163.
[6] J. Hermansson and A. Grahn, *J. Chromatogr.*, in press.
[7] S. Allenmark, B. Bomgren and H. Borén, *J. Chromatogr.*, 264 (1983) 63.
[8] E. Domenici, C. Bertucci, P. Salvadori, G. Felix, S. Motellier and I. Wainer, *Chromatographia*, 29 (1990) 170.
[9] T. Miwa, T. Mayakawa, M. Kayano and Y. Miyake, *J. Chromatogr.*, 408 (1987) 316.
[10] P. Erlandsson, I. Marle, L. Hansson, R. Isaksson, C. Pettersson and G. Pettersson, *J. Am. Chem. Soc.*, 112 (1990) 4573.
[11] I. Marle, P. Erlandsson, L. Hansson, R. Isaksson, C. Pettersson, G. Pettersson, *J. Chromatogr.*, 586 (1991) 233.
[12] I. Marle, S. Jönsson, R. Isaksson, C. Pettersson, *J. Chromatogr.*, 648 (1993) 333.
[13] R. Isaksson, C. Pettersson, G. Pettersson, S. Jönsson, I. Marle, J. Stålberg and J. Hermansson, *Trends Anal. Chem.*, in press.
[14] C. Divne, J. Ståhlberg, T. Reinikainen, L. Ruohonen, G. Pettersson, J.K.C. Knowles, T. Teeri and A. Jones, *Science*, in press.
[15] U. Norinder and J. Hermansson, *Chirality*, 5 (1991) 422.

New ligands for boronate affinity chromatography

Xiao-Chuan Liu*, William H. Scouten

Biotechnology Center, Department of Chemistry and Biochemistry, Utah State University, Logan, UT 84322-4700, USA

Abstract

A new type of boronate affinity ligand was synthesized which contained an internal coordinate bond between a carbonyl oxygen and a boron atom. This coordination makes the boron atom tetrahedral, which is favorable for boronate esterification with cis-diols. The ligand formed a catechol ester in 2-(N-cyclohexylamino)ethanesulfonic acid, Tris and phosphate buffers, respectively, at 0.5, 0.6 and 1.1 pH units lower than did phenylboronic acid. When the ligand was coupled to an agarose gel, the new matrix esterified with catechol at pH 7.5. When the ligand was coupled to a cellulose gel, the new matrix esterified with catechol at pH 7.0. In comparison, the commercial m-aminophenylboronic acid–agarose does not form an ester with catechol below pH 8.0.

1. Introduction

Boronate affinity chromatography was first used to separate nucleic acid components and carbohydrates by Weith et al. in 1970 [1]. Since then, the specificity of boronate has been used to separate cis-diol-containing compounds, including catechols, nucleic acids, glycoproteins and carbohydrates [2]. In recent years it has been applied to separate γ-glutamyltransferase in patients with hepatocellular carcinoma [3]; fast boronate affinity chromatography of glycosylated hemoglobin [4]; purify human platelet glycocalicin [5]; determine glycosylated albumin concentration [6]; determine 5-S-cysteinyldopa in human urine [7]; isolate 3,4-dihydroxyphenylalanine-containing proteins [8]; measure glycosylated hemoglobin in dried blood [9]; and isolate 2-hydroxycarboxylic acids [10]. Boronate affinity chromatography has also been used in

* Corresponding author.

clinical studies. For example, immobilized phenylboronate has been used to measure the level of glycosylated hemoglobin in the diagnosis of diabetes [11–13] and glycemia [14]. Measuring levels of hypoxanthine, uridine and inosine using boronate chromatography has been useful in the diagnosis of several pathological disorders [15,16]. Thus boronate affinity chromatography has been employed for a wide variety of applications in both basic biochemistry and clinical chemistry.

The earliest, and most widely used, boronate ligand is 3-aminophenylboronic acid (3aPBA or mPBA). It can be coupled to a solid support through the anilino group, and the m-amino substitution lowers its pK_a [17]. In all applications of 3aPBA, the pH must be basic, i.e. pH > 8. The pK_a of 3aPBA is 8.8 [18], so the pH should be as high as reasonably possible for optimum binding. However, in many cases the analytes lose their biological activities at such high pH values, which is the major limitation to

expanded use of boronate affinity chromatography [19]. Attempts to lower the pK_a of boronate ligands have involved ligands such as p-bromophenyl boronate [17,20], p-(ω-aminoethyl) phenylboronate [21], and p-vinylbenzene boronate [22] and, more recently, the introduction of strong electron-withdrawing groups on the phenyl ring. For example, Soundararajan et al. [19] put a (N-methyl) carboxyamido group on the phenyl ring, while Singhal et al. [23] put a nitro group on the phenylboronate. Aliphatic boronate ligands with a five-membered stable complex formation have also been employed [24], and proposed [25]. In spite of these efforts, the goal of obtaining a ligand which can form a complex with cis-diols at neutral pH has not been attained [26].

The pH is a crucial factor in the reaction between boronates and cis-diols. Phenylboronic acid does not complex with D-glucose below pH 6, but is fully complexed above pH 9 [27]. It is not known if other factors determine the pH at which boronates react with cis-diols. The pK_a of the boronate may be the critical factor, but the optimal pH for esterification is not the same as the pK_a of the ligand. For example, the maximal esterification of adenosine occurred at pH 7.8, 7.3 and 8.2 during incubation with phenylboronic acid (PBA, $pK_a = 8.9$), p-nitrophenylboronic acid (4nPBA, $pK_a = 7.0$), and p-methylphenylboronic acid (4mPBA, $pK_a \approx 9$), respectively [28]. Maximal esterification with PBA occurs at 1.1 pH unit below its pK_a, whereas maximal esterification with 4nPBA occurs at 0.3 pH unit above its pK_a. This is similar to esterification between β-methylribofuranoside (mRib) and 3aPBA ($pK_a = 8.8$) and 4nPBA [23]. Esterification occurs 1.8 pH units below the pK_a of 3aPBA and 0.4 pH unit below the pK_a of 4nPBA. Thus the pK_a of a ligand is not consistent with the pH at which esterification occurs, which indicates that lowering the pK_a of a ligand may not be the essential factor needed to lower the optimal pH for esterification.

A tetrahedral boronate may be the favorable configuration for forming esters with cis-diol compounds, possibly because a tetrahedral boronate is less strained. We propose that the concentration of tetrahedral conformation, rather than pK_a, is the key factor for esterification. This is supported by the observation that in the esterification of ethylene glycol with 3-acetamidophenylboronic acid (trigonal boronate) and 3,6-diacetamidophenylboronic acid (tetrahedral boronate) in dimethyl sulfoxide (DMSO), the formation constant of the latter is 115 times higher than the former [29]. Further evidence comes from the observation by Wulff et al. [30], who showed that the rate of esterification of PBA with α-propylene glycol was several orders greater in the presence of piperidine, probably because the piperidine nitrogen donated a pair of electrons to the boron atom to form a coordinate nitrogen–boron bond. The resulting boronate was tetrahedral, as shown by ^{11}B NMR. In contrast, although triethylamine has a pK similar to that of piperidine, it did not accelerate the rate of esterification. Therefore, basicity is not the key here. One possible explanation is that steric hindrance associated with triethylamine prevents the formation of a tetrahedral boronate. If so, a boronate affinity adsorbent with the structure shown in Fig. 1 would be a better chromatography matrix at lower pH values than conventional boronate affinity adsorbents.

To determine this, we synthesized and crystallized the model compounds catechol [2-(diethylamino)carbonyl,4-methyl]phenylboronate (I) and catechol [2-(diisopropylamino)carbonyl]phenylboronate (II) (Fig. 2). X-Ray crystallography demonstrated that there was indeed an internal coordination bond between carbonyl oxygen and boron, and that the boron was tetrahedral, which was also supported by ^{11}B NMR [31]. Subsequently, we synthesized the boronate affinity ligand, catechol [2-(diethyla-
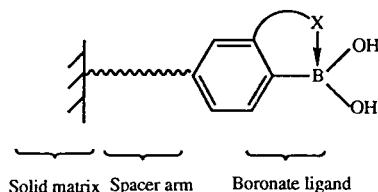
Fig. 1. The proposed internally coordinated boronate affinity matrix.
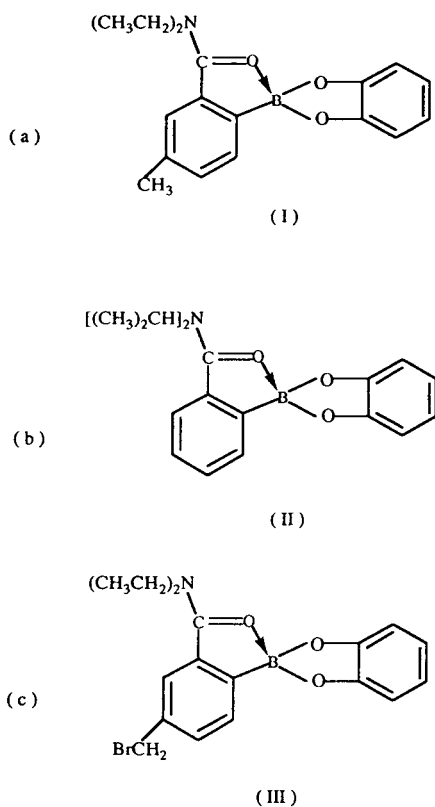
Fig. 2. The boronate affinity ligands synthesized: (a) catechol [2-(diethylamino)carbonyl,4-methyl]phenylboronate (I); (b) catechol [2-(diisopropylamino)carbonyl]phenylboronate (II); (c) catechol [2-(diethylamino)carbonyl,4-bromomethyl]phenylboronate (III).

mino)carbonyl, 4-bromomethyl]phenyl boronate (III) (Fig. 2), which was coupled to agarose and cellulose and used in boronate affinity chromatography.

2. Experimental

2.1. Materials

sec.-Butyllithium (^{sec}BuLi), N,N,N′,N′-tetramethylethylenediamine (TMEDA), tetrahydrofuran (THF), dimethylformamide (DMF), 2,5-dimethoxybenzyl alcohol, trimethyl borate, carbon tetrachloride, bromine, PBA and catechol were purchased from Aldrich (Milwaukee, WI,

USA). N,N-Diethyl-m-tolumide and N,N-diisopropylbenzamide were purchased from Lancaster (Windham, NH, USA). Cysteamine agarose, sulfhydryl cellulose, m-aminophenylboronic acid–agarose (3aPBA-agarose), dithiothreitol (DTT), sorbitol, 2-(N-cyclohexylamino)ethanesulfonic acid (CHES), Tris and 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) were purchased from Sigma (St. Louis, MO, USA). Azomethine H and 5,5′-dithio-bis-(2-nitrobenzoic acid) (DTNB, Ellman's reagent), were products of Pierce (Rockford, IL, USA). All other reagents were of analytical grade. NMR was performed on a Varian XL300 NMR spectrometer; elemental analysis was done by Atlantic Microlab (Norcross, GA, USA); mass spectrometry was performed by Mass Spectrometry Services, Montana State University; chromatography columns were from Sepracor (Marlborough, MA, USA); liquid chromatography was performed on a Bio-Rad Econo System (Hercules, CA, USA).

2.2. Methods

Synthesis of ligands (Fig. 3)

Synthesis of I *and* II

A modification of the method described by Beak and Brown [32] was used to synthesize I and II. All glassware and syringes were oven-dried. The reactions were carried out under purified nitrogen. THF was freshly distilled from potassium metal under nitrogen. TMEDA was freshly distilled from calcium hydride under nitrogen. ^{sec}BuLi was standardized prior to use by titration using 2,5-dimethoxybenzyl alcohol as an indicator [33]. The procedure used to synthesize I is described below.

TMEDA (2 ml, 13.0 mmol) in 30 ml dry THF at −78°C was added dropwise to 12 ml ^{sec}BuLi (13.0 mmol). After 10 min, 2.1 g of N,N-diethyl-m-tolumide (11.0 mmol) in 20 ml dry THF were added dropwise. The mixture was then stirred for 1 h at −78°C. Trimethyl borate, 8.0 ml (70.4 mmol), was added rapidly and the mixture was then stirred for an additional 20 h during which it was allowed to warm to room temperature. The
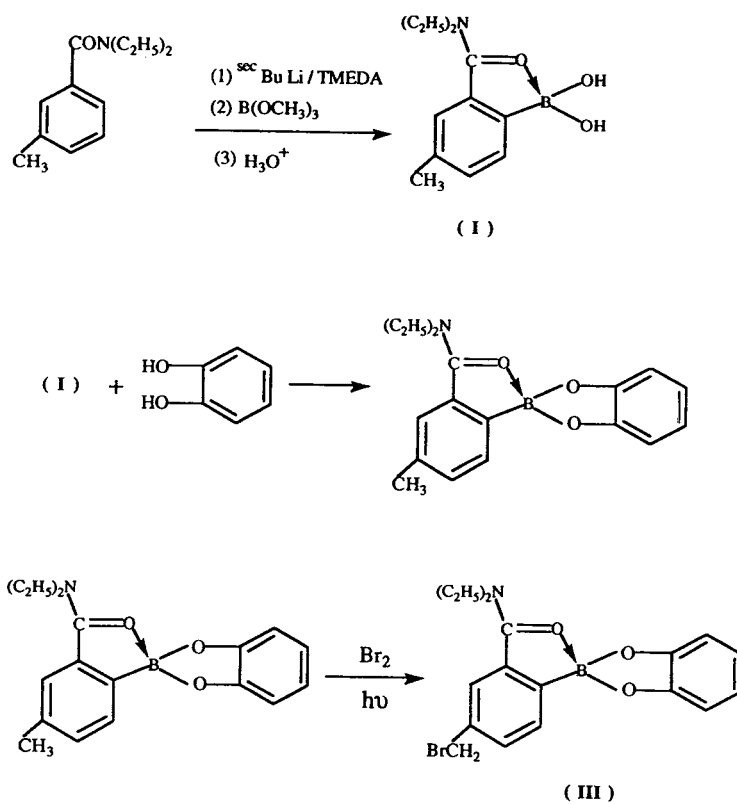
Fig. 3. Synthesis of III.

reaction mixture was then poured into a separating funnel containing 30 ml saturated aqueous NH_4Cl (pH 5.7) and 30 ml diethyl ether. The water phase was extracted twice and the combined ether phase was dried over $MgSO_4$. Evaporation of solvent in vacuo gave 2.46 g (10.5 mmol, 95.4%) [2-(diethylamino)carbonyl,4-methyl]phenylboronic acid.

Using a Dean–Stark water separator, the azeotrope of water and benzene was distilled from a solution of 2.1 g (8.9 mmol) of [2-(diethylamino)carbonyl,4-methyl]phenylboronic acid in 100 ml benzene. The reaction mixture cooled to 70°C and 1.0 g of catechol (9.1 mmol) was added. Distillation then continued for 3 h. Evaporation of the solvent in vacuo yielded 2.91 g (89.1%) of crude product, I. Recrystallization from toluene gave white needles. m.p. 173–175°C. ^{1}H NMR (300 MHz, C^2HCl_3) δ (ppm) 7.66 (d, 1H), 7.50 (s, 1H), 7.44 (d, 1H), 6.88 (m,

2H), 6.78 (m, 2H), 3.95 (q, 2H), 3.70 (q, 2H), 2.45 (s, 3H), 1.50 (t, 3H), 1.30 (t, 3H). ^{11}B NMR δ (ppm) 13.5 ppm, single peak. IR (KBr), ν 1620 cm^{-1}, 806 cm^{-1}, 740 cm^{-1}, 701 cm^{-1}. Mass spectrum (m/z, ion): 295, 223, 195, 167, 105. Analysis: calculated for $C_{18}H_{20}O_3BN \cdot 0.5C_7H_8$: C, 72.69; H, 6.81; N, 3.94; found: C, 72.16; H, 6.68; N, 3.91.

II was synthesized according to the same procedure, using N,N-diisopropylbenzamide in place of N,N-diethyl-m-tolumide m.p. 179–180°C. ^{1}H NMR (300 MHz, C^2HCl_3) δ (ppm) 7.78 (d, 1H), 7.68 (d, 1H), 7.60 (t, 1H), 7.42 (t, 1H), 6.80 (m, 4H), 5.05 (m, 1H), 3.80 (m, 1H), 1.46 (q, 12H). ^{11}B NMR δ (ppm) 13.5 ppm, single peak. IR (KBr), ν 1620 cm^{-1}, 746 cm^{-1}. Mass spectrum (m/z, ion): 323, 280, 223, 195, 167, 136, 105. Analysis: calculated for $C_{19}H_{22}O_3BN \cdot C_6H_6O_2$: C, 69.30; H, 6.51; N, 3.23; found: C, 69.21; H, 6.57; N, 3.13.

Synthesis of ligand **III** [34].

II (0.828 g, 2.68 mmol) in 100 ml dry carbon tetrachloride in a 200-ml flask was equipped with a Claisen head, a condenser and a receiver. Carbon tetrachloride (20 ml) was distilled from the mixture to insure the absence of water in the solvent. The distillation apparatus was replaced quickly by a dropping funnel containing 4.30 g of bromine (26.8 mmol) in 15 ml of dry carbon tetrachloride. Bromine was added dropwise with illumination from two 150-W tungsten bulbs held 3 in. (1 in. = 2.54 cm) from the flask. The mixture was stirred vigorously and maintained at a gentle reflux until the color due to the presence of bromine in reaction flask disappeared. Addition of bromine ceased when the color no longer disappeared after 5 min. Evaporation of the solvent yielded 0.94 g of a gray cream powder, which was used for coupling without further purification. ^{1}H NMR showed that the benzylmethyl group was brominated. ^{11}B NMR showed a peak at 13.5 ppm.

Interaction with catechol in solution

The reaction of catechol [2-(diisopropyla-mino)carbonyl]phenylboronic acid (DICP) and PBA with catechol (1:2 ratio) in CHES, Tris and phosphate buffers at various pH values was studied by ^{11}B NMR on a Varian XL300 NMR spectrometer. Samples contained 10% ^{2}H$_2$O to provide an NMR locking signal.

^{11}B NMR spectroscopy

^{11}B NMR spectra were recorded at 25°C on a Varian XL300 NMR spectrometer at 96.25 MHz, using boron trifluoride diethyl etherate as the external standard. Boronate derivatives and catechol were dissolved in various buffers containing 10% ^{2}H$_2$O at different pH values. ^{11}B NMR chemical shifts were measured above the boron background present in the NMR tubes, and the ^{11}B NMR peaks appeared as sharp peaks above the bell-shaped background.

Coupling of the boronate ligand [35,36] (Fig. 4)

Coupling of the boronate ligand on agarose

Cysteamine agarose (5 ml) was washed with excess anhydrous DMF to remove water, then transferred to a 100-ml flask containing 0.1 g **III**, 0.05 g anhydrous K$_2$CO$_3$ and 50 ml DMF. The mixture was stirred for 4 h at room temperature
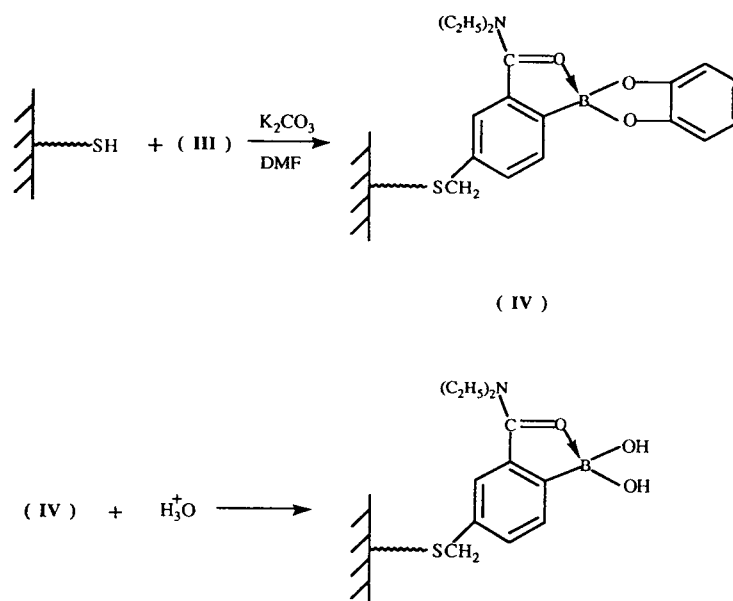
Fig. 4. Immobilization of the boronate affinity ligand **III** on a solid matrix.

under nitrogen, after which the gel was washed with 50 ml DMF, then 50 ml 50 mM HEPES, pH 8.5.

Coupling of the boronate ligand on cellulose

Sulfhydryl cellulose (1 g) was added to a 100-ml flask containing 15 ml anhydrous DMF and 0.415 g DTT. The mixture was stirred for 1 h at room temperature under nitrogen. The gel was washed thoroughly using DMF, and then transferred to a 50-ml flask containing 0.55 g **III** and 0.05 g anhydrous K_2CO_3. The mixture was stirred for 4 h at room temperature under nitrogen. After reaction, the gel was washed with DMF, then 50 mM HEPES, pH 8.5.

During the coupling reactions, the sulfhydryl concentration was monitored by Ellman's assay [37], and the boron concentration was monitored by the azomethine H method [38].

Affinity chromatography

A 1-ml column (Sepracor) was packed with 1 ml of gel. The column was connected to an automated liquid chromatographic system (Bio-Rad). It was first washed with 1% acetic acid to hydrolyze catechol boronate to boronic acid, after which it was equilibrated with 50 mM HEPES. Catechol samples were dissolved in the equilibration buffer. After application of the sample, the column was washed with equilibration buffer until the absorbency returned to the baseline. Elution was accomplished using 1% acetic acid or 0.2 M sorbitol in the equilibration buffer. In the case of elution using sorbitol, the column was regenerated by treatment with 1% acetic acid, followed by re-equilibration with the washing buffer.

3. Results and discussion

3.1. DICP and PBA interact with catechol in solution

Although DICP can not be coupled to a solid matrix, we used it as a model to study boronate interaction with *cis*-diols in solution. The results were in good agreement with those obtained

using immobilized boronate affinity columns, as demonstrated by Singhal et al. [23]. Phenylboronic acid was used as the standard since its structure is closely related to that of commercial boronate affinity matrices.

In CHES buffers (Fig. 5), catechol boronate ester between phenylboronic acid and catechol formed at pH 9.0, as shown by ^{11}B NMR, whereas catechol boronate ester between [2-(di-isopropylamino)carbonyl]phenylboronic acid and catechol occurred at pH 7.9. Free phenylboronic acid has a chemical shift of 29.6 ppm on ^{11}B NMR, which indicates a trigonal boron. The chemical shift of catechol phenylboronates were approximately 10 ppm (9.6 ppm for DICP ester and 10.4 ppm for PBA ester in this case) on ^{11}B NMR, which indicates a tetrahedral boron. Free
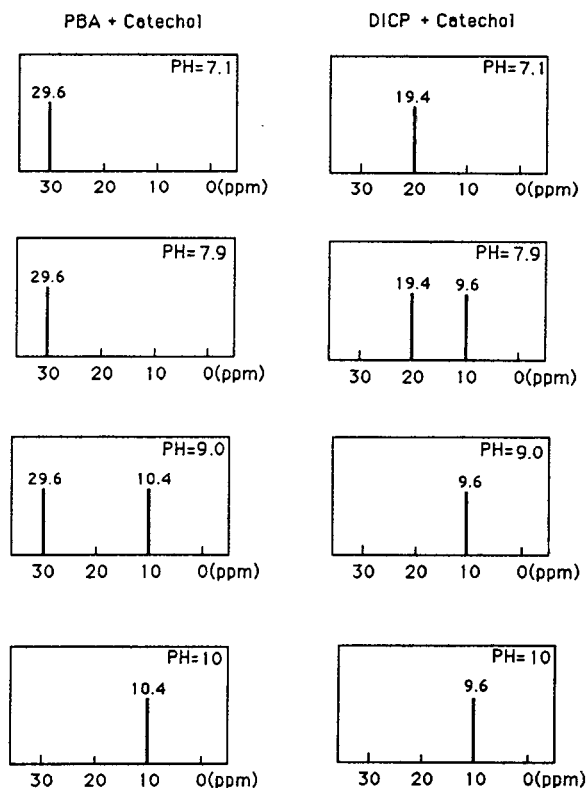
Fig. 5. Interaction between boronate affinity ligands and catechol. The reaction of PBA and DICP with catechol (1:2 ratio) in CHES buffer at various pH values. Numbers shown are chemical shift on ^{11}B NMR. Samples contained 10% ^2H_2O to provide an NMR locking signal.

[2-(diisopropylamino)carbonyl]phenylboronic acid had a chemical shift of 19.4 ppm on ^{11}B NMR, suggesting that the boron was partially tetrahedral. At pH values higher than the "critical pH" (where the catechol esters start to form), all boronate groups eventually became esterified with catechol, and demonstrated only one ^{11}B NMR peak at about 10 ppm. At pH values lower than the "critical pH", no esterification occurred, and only a peak for the free boronate ligand (29.6 ppm for PBA and 19.4 ppm for DICP) was seen. As shown in Fig. 5, [2-(diisopropylamino)carbonyl]phenylboronic acid esterified with catechol at 1.1 pH unit lower than did phenylboronic acid.

In Tris buffers (Fig. 6), DICP began to form an ester bond with catechol at pH 7.0 compared

to 7.6 for PBA, which is 0.6 pH unit lower. In phosphate buffers (Fig. 7), esterification between [2-(diisopropylamino)carbonyl]phenyl boronic acid and catechol occurred at pH 6.6, while esterification between phenylboronic acid and catechol formed at pH 7.1, which is 0.5 pH unit lower than for PBA. It is known that buffers have various effects on boronate affinity chromatography [39].

3.2. Affinity chromatography on [2-(diethylamino)carbonyl,4-bromomethyl]phenylboronate–agarose (DECBP-agarose)

The chromatography of catechol on DECBP-agarose gel is shown in Fig. 8a and b. The
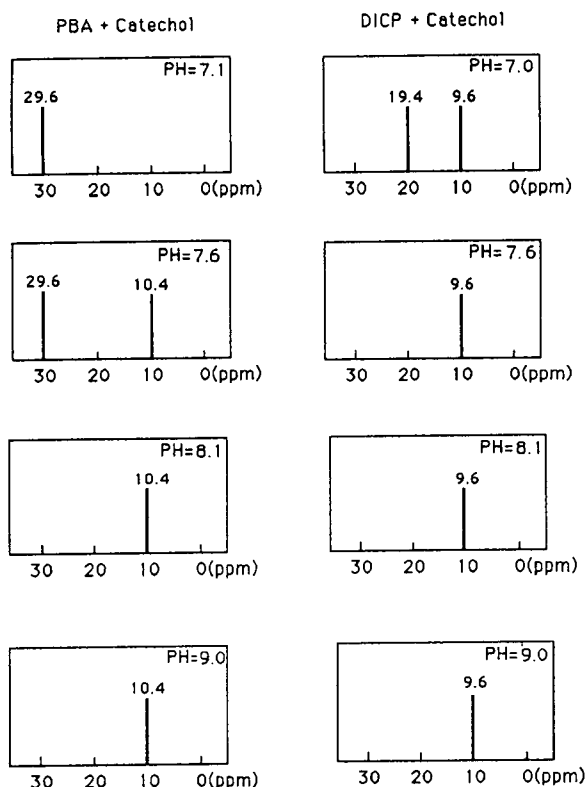
Fig. 6. Interaction between boronate affinity ligands and catechol. PBA and DICP were reacted with catechol (1:2 ratio) in Tris buffers at various pH values. Numbers shown are chemical shift on ^{11}B NMR. Samples contained 10% ^{2}H$_{2}$O to provide NMR locking signal.
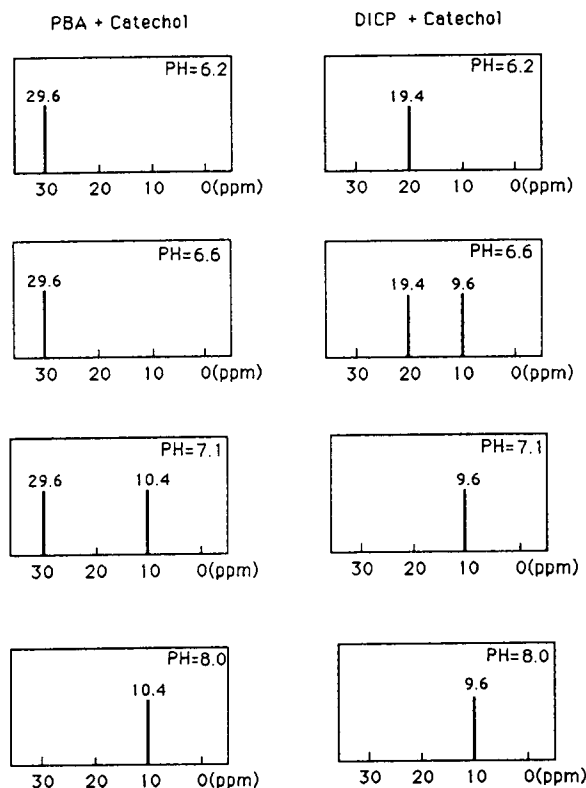
Fig. 7. Interaction between boronate affinity ligands and catechol. PBA and DICP were reacted with catechol (1:2 ratio) in phosphate buffers at various pH values. Numbers shown are chemical shift on ^{11}B NMR. Samples contained 10% ^{2}H$_{2}$O to provide NMR locking signal.
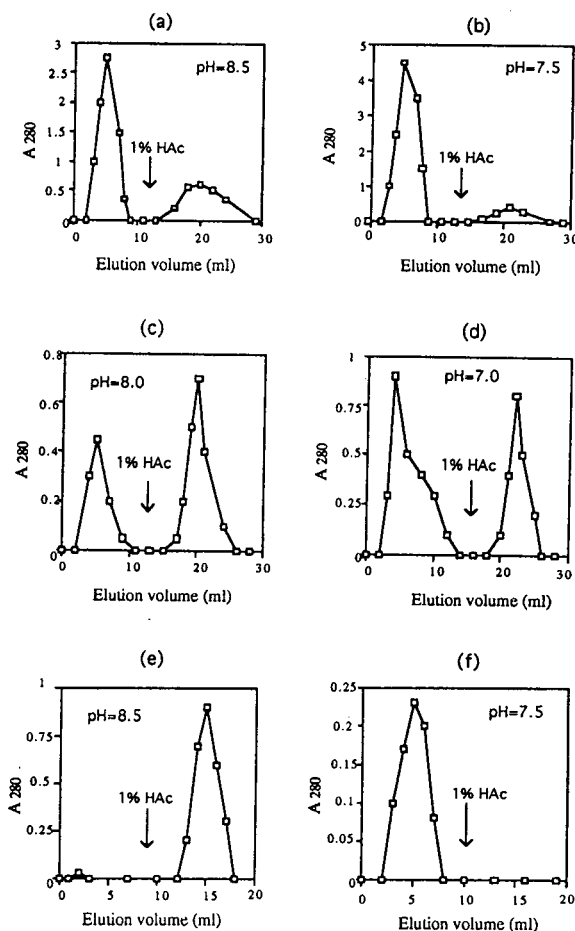
Fig. 8. Chromatography of catechol on DECBP-agarose gel, DECBP-cellulose gel and mPBA-agarose gel. (a) Column: 1 ml DECBP-agarose gel (0.14 μmol boronate); washing buffer: 50 mM HEPES, pH 8.5; sample: 0.25 μmol catechol in washing buffer; elution: 1% acetic acid; flow-rate: 2.0 ml/min. (b) Column: 1 ml DECBP-agarose gel (0.14 μmol boronate); washing buffer: 50 mM HEPES, pH 7.5; sample: 0.33 μmol catechol in washing buffer; elution: 1% acetic acid; flow-rate: 2.0 ml/min. (c) Column: 1 ml DECBP-cellulose gel (0.66 μmol boronate); washing buffer: 50 mM HEPES, pH 8.0; sample: 1.0 μmol catechol in washing buffer; elution: 1% acetic acid; flow-rate: 1.0 ml/min. (d) Column: 1 ml DECBP-cellulose gel (0.66 μmol boronate); washing buffer: 50 mM HEPES, pH 7.0; sample: 1.0 μmol catechol in washing buffer; elution: 1% acetic acid; flow-rate: 1.0 ml/min. (e) Column: 1 ml mPBA-agarose gel (14 μmol boronate); washing buffer: 50 mM HEPES, pH 8.5; sample: 2.5 μmol catechol in washing buffer; elution: 1% acetic acid; flow-rate: 2.0 ml/min. (f) Column: 1 ml mPBA-agarose gel (14 μmol boronate); washing buffer: 50 mM HEPES, pH 7.5; sample: 0.6 μmol catechol in washing buffer; elution: 1% acetic acid; flow-rate: 2.0 ml/min.

column was overloaded with an excess of catechol. After unbound catechol was eluted with washing buffer, 1% acetic acid was used to elute catechol bound to the gel. At pH 8.5, catechol was easily bound to the gel, and subsequently eluted with 1% acetic acid. At pH 7.5, catechol binding also occurred although the elution peak was somewhat smaller than that at pH 8.5.

3.3. Affinity chromatography on [2-(diethyl-amino)carbonyl,4-bromomethyl]phenyl-boronate–cellulose (DECBP-cellulose)

The chromatography of catechol on DECBP-cellulose gel, as shown in Fig. 8c and d, followed the same procedures as for DECBP-agarose. At pH 8.0 and 7.0, catechol bound to the gel, and gave a sharp peak upon elution by 1% acetic acid. Elution using 0.2 M sorbitol in the washing buffer gave similar results. Chromatography at pH 6.0 showed that some binding occurred, although the amount of binding was considerably less than at pH 7.0.

3.4. Affinity chromatography on m-amino-phenylboronic acid–agarose (mPBA-agarose)

These experiments were performed as a comparison to commercial mPBA-agarose gel. As shown in Fig. 8e and f, at pH 8.5, 50 mM HEPES, catechol was bound to the gel, and gave a sharp elution peak. However, at pH 7.5, as was expected, no catechol was bound to the gel. Results from additional chromatography showed no binding occurred until the pH reached 8.0.

These investigations demonstrated that the new internally coordinated boronate ligands can esterify with catechol at neutral pH values, both in solution and when bound to the boronate affinity columns. In solution, DICP esterified with catechol at 0.5, 0.6 and 1.1 pH unit lower than PBA did in CHES, Tris and phosphate buffers, respectively. When III was coupled to agarose gel, the new matrix esterified with catechol at pH 7.5. When III was coupled to cellulose gel, the new matrix esterifies with catechol

at pH 7.1. In comparison, the commercial mPBA-agarose did not esterify with catechol below pH 8.0. In addition, the adsorbent might also reduce secondary ionic and charge transfer interactions in boronate affinity chromatography [39], since no boron anion is formed and the boron no longer has an empty orbital for charge transfer. Therefore, utilizing this type of ligand might increase the overall application of boronate affinity chromatography.

Acknowledgements

References

[1] H.L. Weith, J.L. Wiebers and P.T. Gilham, *Biochemistry*, 9 (1970) 4396.
[2] W.H. Scouten, *Solid Phase Biochemistry*, Wiley, New York, (1983) p.149.
[3] T. Yamamoto, Y. Amuro, Y. Matsuda, H. Nakaoka, S. Shimomura, T. Hada and K. Higashino, *Am. J. Gastroenterol.*, 86 (1991) 495.
[4] S. Hjertén and J. Li, *J. Chromatogr.*, 500 (1990) 543.
[5] R. DeCristofaro, R. Landolfi, B. Bizzi and M. Castagnola, *J. Chromatogr.*, 426 (1988) 376.
[6] R.N. Jahnson and J.R. Baker, *Clin. Chem. (Winston-Salem, N.C.)*, 34 (1988) 1456.
[7] C. Hansson, B. Kaagedal and M. Kaellgerg, *J. Chromatogr.*, 420 (1987) 146.
[8] C.J. Hawkins, M.F. Lavin, D.L. Parry and I.L. Ross, *Anal. Biochem.*, 159 (1986) 187.
[9] J. Eross, D. Kreutzmann, C. Crowell and M. Silink, *Clin. Chem. (Winston-Salem, N.C.)*, 32 (1986) 2222.
[10] S. Higa and S. Kishimoto, *Anal. Biochem.*, 154 (1986) 71.
[11] D.C. Klenk, G.T. Hermanson, R.I. Krohn, E.K. Fujimoto, A.K. Malia, P.K. Smith, J.D. England, H.M. Wiedmeyer, R.R. Little and D.E. Goldstein, *Clin. Chem. (Winston-Salem, N.C.)*, 28 (1982) 2088.
[12] B.J. Gould, P.M. Hall and G.H. Cook, *Clin. Chim. Acta*, 125 (1982) 41.
[13] R. Kluckiger, T. Woodtli and W. Berger, *Diabetes*, 33 (1984) 73.
[14] H.F. Bunn, *Am. J. Med.*, 70 (1981) 325.
[15] K. Nakano, K. Shindo, K. Yaraka and H. Yamamoto, *J. Chromatogr.*, 332 (1985) 21.
[16] K. Nakano, K. Shindo, K. Yaraka and H. Yamamoto, *J. Chromatogr.*, 332 (1985) 127.
[17] A. Yurkevich, I. Kolodkina, E. Ivanova and E. Pichuzhkina, *Carbohydr. Res.*, 43 (1975) 215.
[18] H.-G. Boit, *Beilstein*, 16(3), Springer, Berlin, 1974, p. 1284.
[19] S. Soundararajan, M. Badawi, C.M. Kohlrust and J. Hageman, *Anal. Biochem.*, 178 (1989) 125.
[20] E. Ivanova, S. Panchenko, I. Kolodkina and A. Yurkevich, *Z. Obshcheikhimii*, 45 (1975) 208.
[21] V. Akparov and V. Stepanov, *J. Chromatogr.*, 155 (1978) 329.
[22] C. Elliger, B. Chan and W. Stanley, *J. Chromatogr.*, 104 (1975) 57.
[23] R.P. Singhal, B. Ramamurthy, N. Govindraj and Y. Sarwar, *J. Chromatogr.*, 543 (1991) 17.
[24] V. Adamek, X. Liu, Y. Zhang, K. Adamkova and W.H. Scouten, *J. Chromatogr.*, 625 (1992) 91.
[25] R.P. Singhal and S.S.M. DeSilva, *Adv. Chromatogr.*, 31 (1992) 307.
[26] J.R. Mazzeo and I.S. Krull, *Biochromatography*, 4 (1989) 124.
[27] S.A. Barker, A.K. Chopra, B.W. Hatt and P.J. Somers, *Carbohydr. Res.*, 26 (1973) 33.
[28] R.J. Ferrier, *Adv. Carbohydr. Chem. Biochem.*, 35 (1978) 31.
[29] S.X. Cai and J.F.W. Keana, *Bioconjugate Chem.*, 2 (1991) 317.
[30] G. Wulff, R. Dederichs, R. Grotstollen and C. Jupe, in T.C.J. Gribnau, J. Visser and R.J.F. Nivard (Editors), *Affinity Chromatography and Related Techniques —Theoretical Aspects/Industrial and Biomedical Applications (Proceedings of the 4th International Symposium, Veldhoven, 22–26 June 1981)*, Elsevier, Amsterdam, 1982, p. 207.
[31] X. Liu, J.L. Hubbard and W.H. Scouten, *J. Organomet. Chem.*, (1994) submitted for publication.
[32] P. Beak and R.A. Brown, *J. Org. Chem.*, 47 (1982) 34.
[33] M.R. Winkle, J.M. Lansinger and R.C. Ronald, *Chem. Commun.*, (1980) 87.
[34] P. Tschampel and H.R. Snyder, *J. Org. Chem.*, 29 (1964) 2168.
[35] J.M. Khurana and P.K. Sahoo, *Synthetic Commun.*, (1992) 1691.
[36] Y.S. Or, R.F. Clark and J.R. Luly, *J. Org. Chem.*, 56 (1991) 3146.
[37] G.L. Ellman, *Arch. Biochem. Biophys.*, 82 (1959) 70.
[38] T.P. Gaines and G.A. Mitchell, *Commun. Soil Sci. Plant Anal.*, 10 (1979) 1099
[39] S. Fulton, *Boronate Ligands in Biochemical Separations*, Amicon Corp., Danvers, MA, 1981, p. 6.

Alignment of chromatographic profiles for principal component analysis: a prerequisite for fingerprinting methods☆

Gunnar Malmquist[1], Rolf Danielsson[*]

Institute of Chemistry, Department of Analytical Chemistry, Uppsala University, P.O. Box 531, S-751 21 Uppsala, Sweden

Abstract

In fingerprinting methods, small differences between chromatograms with rather complex appearance have to be detected. Pattern recognition methods based on principal component analysis (PCA) could be a useful tool, but for chromatographic profiles as input data a severe problem is the great impact of chromatographic variations compared with true variations in sample composition. The problem has been analysed in terms of parameter variations for exponentially modified Gaussian peaks, and a procedure has been developed to align a sample chromatogram towards a target chromatogram in order to compensate for (i) small shifts in retention time (not due to different sample components), (ii) common variations in peak area (not due to sample composition) and (iii) variations in level and slope of the baseline. The effects of the alignment procedure on the PCA is demonstrated for a set of chromatographic profiles intended for peptide mapping.

1. Introduction

Several important analytical techniques used for biological samples rely on a comparison of chromatographic profiles. Examples of application areas are food and beverage analysis [1], DNA fingerprinting [2], pyrolysis–GC [3] and peptide mapping [4,5]. Sometimes the original sample is analysed for its components, whereas in other instances the sample is fragmented before the analysis. The analysis is performed by a separation of the components or fragments, e.g., by electrophoresis or chromatography, and the resulting profile is used as a pattern or fingerprint for the sample. In many instances, the evaluation consists of a direct visual comparison with a reference sample, in order to detect profile differences.

An important characteristic of these fingerprinting methods, is that the significant information may be contained in the presence or absence of certain components or fragments. Small differences between samples that overall are similar can, for instance, be detected by the altered retention time for a certain fragment containing the modified site. This differentiates these methods from techniques in which the determination of certain components in the samples provide the information. Another aspect is that it is seldom necessary, or even possible, to identify and quantify all peaks in the profile. The

---

overall appearance of the fingerprint is instead used for discrimination purposes, making these methods qualitative rather than quantitative.

Peptide mapping is a fingerprinting method frequently used for quality control in the biotechnological production of recombinant DNA-derived proteins [6]. It is necessary to establish that the amino acid sequence is the same for different production batches. The protein is fragmented, by chemical cleavage or enzymatic digestion, and the resulting peptide fragments are separated, usually by reversed-phase liquid chromatography (RPLC). This technique will be used as an illustrative example in the present paper, while a more detailed discussion of the experimental aspects will be given in the accompanying paper [7].

The chromatographic separation is associated with several sources of variation that may have a large impact on the overall pattern. Variations in the mobile phase composition, gradient reproducibility, temperature variations and column variability lead to shifts in the chromatographic pattern, making the evaluation more difficult. If several chromatograms are to be compared, these matters should be considered. On the other hand, if two sequential chromatograms, run with the same mobile phase preparation and on the same column, are compared, these chromatographic variations might not be very severe.

The result of a manual comparison of chromatographic profiles may depend on the individuals performing the comparison. Multivariate methods that can cope with the variations in the digestion of reference samples make the evaluation less subjective. The idea is to gather a whole set of reference chromatograms that represents the normal variations in the chromatographic profile caused by the experimental conditions. Principal component analysis (PCA) [8] can be used to identify the main variation sources and to highlight the peaks where the variations are reflected. New test samples can subsequently be classified by multivariate classification methods.

The multivariate data analysis is conducted on a data set where the objects, the reference chromatograms in this case, are described by a number of variables. Some of the different approaches for conversion of chromatographic traces to variable values that have been suggested will be briefly discussed here.

The most intuitive approach is to represent the chromatograms by integration reports, i.e., the retention time and peak area for all detected peaks. The peak areas for all, or selected, peaks may constitute the variable values, provided a correct assignment of the peaks can be made between the chromatograms. An important prerequisite for multivariate analysis is that the variations should be expressed as different levels of the variables, not as shifts between variables [9]. Incorrect peak assignments imply that the quantity of a certain solute will be contained in different variables between the objects, thus reducing the quality of the data set.

The peak assignment is usually based entirely on retention time matching, unless a specific detector, e.g., a mass spectrometer, can identify the peaks. The assignment process is simplified if the retention times of the peaks are synchronized between the chromatograms. Several methods for peak synchronization, where the retention times of the detected peaks are adjusted using reference peaks that can be identified in all chromatograms, have been developed for multivariate analysis in fingerprinting contexts. These methods are conceptually related to the use of retention indices in GC. The retention time matching is only qualitative, as the retention times are used for identification purposes only. Crawford and Hellmuth [10] used two internal standard peaks to make minor linear adjustments of the retention times. More elaborate methods using multiple reference peaks have also been presented, e.g., by Mayfield and Bertsch [11], Pino et al. [12], and Parrish et al. [13].

Multivariate analysis of chromatograms represented by peak areas for a number of peaks is of limited value in peptide mapping, where the purpose is to detect modifications of the amino acid sequence. All peaks correspond to potential modification sites, and a single amino acid substitution may lead to a large change in retention.

G. Malmquist, R. Danielsson / J. Chromatogr. A 687 (1994) 71–88

73

This not only complicates the peak assignment, but implies that no relevant assignment can be made for the modified fragment as it actually constitutes a new peak. No variable will be present in the data set to contain its peak area.

The variable values can instead be defined by window summation, where the chromatogram is divided into a number of consecutive retention time segments. Each segment corresponds to one variable in the data set, and the variable value is calculated as the sum of the total signal within the window. This approach has been used, for instance, by Headley and Hardy [1] for detection of contaminants in whiskey based on GC profiles. Recently, Armanino et al. [14] presented a similar method for extraction of information regarding air pollution from GC profiles. In the latter instance, the window summation was preceded by retention alignment. A serious drawback of the window summation approach is the inherent loss of resolution. The effect of the summation is essentially a decrease in the sampling frequency, where each variable value corresponds to an averaged signal within the window. This means that variations in the size of a small peak may not be detected if it occurs in the same window as a large peak.

Fingerprinting methods in general, and peptide mapping in particular, put strong demands on the resolution and peak capacity, thus enhancing the previously mentioned drawback of window summation. A more adequate approach would be to use the entire chromatographic profile, i.e., the digitalized detector signal, directly. The data set will then have one variable for each data point in the chromatogram. Such large data sets were previously considered impractical, forcing variable reductions prior to the multivariate analysis, but the advent of powerful personal computers allows the direct analysis of large data sets. The present paper deals with some special aspects of chromatographic profiles in multivariate analysis. A preprocessing procedure is presented that facilitates the characterization of a set of reference chromatograms by PCA. An application of this procedure concerning multivariate classification of peptide

mapping chromatograms is discussed in the accompanying paper [7].

2. Chromatographic profiles in PCA

It is important to realize that PCA is sensitive to all variations in the data set, i.e., not only those emanating from differences between the samples but also to variations caused by the chromatographic process. The latter aspect will be treated in some detail, illustrating the influence of chromatographic variations in PCA.

A commonly used model for chromatographic peaks is the exponentially modified Gaussian function, EMG [15]. With this function the response $y(t)$ is characterized by four parameters:

$$y(t) = \frac{A}{\tau} \exp\left[\frac{1}{2}\left(\frac{\sigma}{\tau}\right)^2 - \left(\frac{t - t_r}{\tau}\right)\right] \int_{-\infty}^{z\sqrt{2}} \exp(-x^2)\,dx$$

(1)

where A is the peak area, t_r and σ are the retention time and width (standard deviation) of a unit area Gaussian peak and τ is the time constant of the modifying exponential decay (tailing). The upper limit of integration is given by $z = (t - t_r)/\sigma - \sigma/\tau$.

2.1. Linear approximation for peak difference

Small shifts in the EMG parameters give rise to variations in the resulting peaks. In order to compare two similar peaks we can look at the difference Δy, which can be approximated by

$$\Delta y = \frac{\partial y}{\partial A} \cdot \Delta A + \frac{\partial y}{\partial t_r} \cdot \Delta t_r + \frac{\partial y}{\partial \sigma} \cdot \Delta \sigma + \frac{\partial y}{\partial \tau} \cdot \Delta \tau \quad (2)$$

i.e., a linear combination of the parameter shifts with the corresponding partial derivatives as coefficients. These partial derivatives with respect to the parameters are derived in the Appendix. Although the EMG function is rather complex, the results can be related to the origi-

nal peak in a simple way. Except for scaling constants they are

$$\frac{\partial y}{\partial A} = y(t) \tag{3}$$

(i.e., the EMG function itself);

$$\frac{\partial y}{\partial t_r} = \frac{dy}{dt} \tag{4}$$

(i.e., the time derivative);

$$\frac{\partial y}{\partial \sigma} = \frac{d^2 y}{dt^2} \tag{5}$$

(i.e., the second time derivative);

$$\frac{\partial y}{\partial \tau} = \frac{dy}{dt} * e^{-t/\tau} \tag{6}$$

(i.e., the time derivative, exponentially modified by convolution once more).

The first two partial derivatives can easily be perceived. The area parameter is just a scaling constant, and a change in retention time appears as a shift along the time axis.

The linear approximation holds for each point in time, i.e., the response difference $\Delta y(t)$ can be regarded as a linear combination of the partial derivatives (as functions of time) with the parameter shifts as coefficients. It is valid only for small deviations in the parameters, except for the scaling parameter A, where a linear relation holds for all values. In Fig. 1 the validity of the
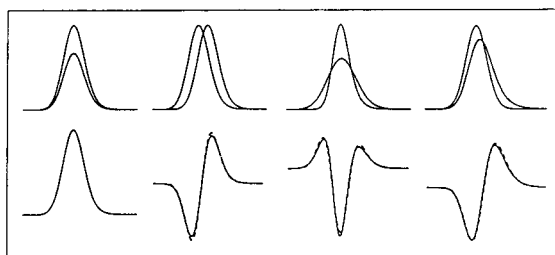
Fig. 1. Top: EMG peaks with deviations in A, t_r, σ, and τ, respectively (from left to right). Bottom: the resulting differences (solid lines) and the linear approximations using partial derivatives (dashed lines). The derivatives were obtained for $A = 1$ (arbitrary units), $t_r = 40\%$, $\sigma = 8\%$ and $\tau = 5\%$ compared with the length of the time scale. The parameter deviations were $\Delta A = \pm 0.2$, $\Delta t_r = \pm 4\%$, $\Delta \sigma = \pm 2.5\%$ and $\Delta \tau = \pm 2.5\%$.

approximation is demonstrated for certain values of parameter deviation. These values are chosen to represent a practical limit for the linear approximation. The area deviation is merely chosen to give a realistic picture of all parameter variations.

2.2. PCA of the EMG parameter shifts

To characterize the variations for a set of similar chromatographic peaks (e.g., a set of single-peak chromatograms), we can apply PCA [8]. With this technique, the chromatograms are described as linear combinations of deviations $p_1(t), p_2(t), \ldots$ from the mean chromatogram $y_{\text{mean}}(t)$. For chromatogram i one obtains

$$y_i(t) = y_{\text{mean}}(t) + s_{i1}p_1(t) + s_{i2}p_2(t) + \cdots \tag{7}$$

Although written as time functions, the chromatogram $y_i(t)$ and the deviations $p_1(t)$, $p_2(t), \ldots$ are represented as sampled values at t_j, for $j = 1, 2, 3, \ldots$ In the context of PCA, the sampled values of the chromatograms are assigned to separate variables, one for each point in time. The results of PCA are the coefficients $s_{i1}, s_{i2}, \ldots$, one set for each chromatogram, and the sampled values $p_1(t_j), p_2(t_j), \ldots$ The latter are referred to as loadings (loading vectors), deviation patterns that are common for the whole set of chromatograms. The loadings are constructed one by one, starting with $p_1(t_j)$, in such a way that the current approximations are as close to $y_i(t)$ as possible in the sense of least squares (all chromatograms included). The individual chromatograms are then represented by the coefficients $s_{i1}, s_{i2}, \ldots$, which are called the scores for chromatogram i.

Using PCA, the variations around the mean chromatogram are thus described by a number of components with contributions $s_{i1}p_1(t), s_{i2}p_2(t)$, etc. The first component accounts for as much variation as possible in the original set of chromatograms, the second plays the same role for the remaining residuals, and so on. This is reflected in a decreasing series of measures for "explained variance", the sum of which approaches 100% when the number of components

equals the number of independent sources of variation. Owing to the effect of noise and non-linearities, the optimum number of components must be determined by some more elaborate procedure, usually by cross-validation [16].

2.3. PCA for single peak variations

If the variations arise from shifts in only one of the EMG parameters, we will obtain one main component from the PCA. The loadings for this component, i.e., the shape of the variation, then corresponds to the partial derivative for that parameter (since the deviation from the average curve can be approximated by Eq. 2 for each curve). To verify this, a series of five EMG peaks were generated for each parameter (A, t_r, σ and τ), with parameter values evenly distributed within the same range as before (cf., Fig. 1). Except for a normalizing constant, the loadings obtained with PCA were almost identical with the partial derivatives (cf., Fig. 1), and the scores were related to the parameter deviations in a linear way. This is demonstrated by the correlation coefficients listed in Table 1, where also the amount of explained variance is given. The deviation from unity (100%) indicates the influence of non-linear relationships for the parameter in question.

When there are variations in more than one of the parameters, there will be additional principal components. In Fig. 2 a design is shown for deviations in area and also in retention time,

Table 1
Results from PCA of chromatograms with variations in one parameter

Parameter	Correlation coefficient		Explained variance (%)
	a^a	b^b	
ΔA	1.0000	1.0000	100
Δt_r	0.9993	0.9999	97.4
$\Delta \sigma$	0.9965	0.9946	98.4
$\Delta \tau$	0.9978	0.9974	98.4

[a] Correlation between the PC1 loadings and the partial derivative.
[b] Correlation between the PC1 scores and the parameter values.
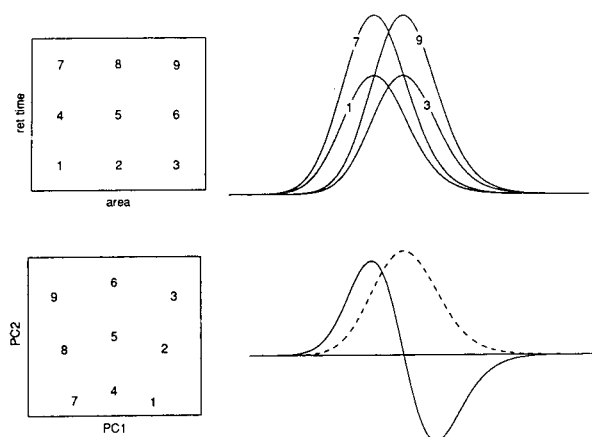
Fig. 2. Top: design of variations in A and t_r and the EMG peaks for the corner points, ($\pm \Delta A$, $\pm \Delta t_r$). Bottom: score plot for PC1 and PC2 (left) and the loadings for PC1 (solid line) and PC2 (dashed line).

together with the resulting curves corresponding to the corners of the square design. Two principal components are obtained, which together explain 98.6% of the variance. The PCA results are also depicted in Fig. 2. The two-dimensional score plot reflects the design pattern, with the deviations in retention time along the first principal component (PC1) and those in area along the second (PC2). In the two loading plots, the peak derivative shape connected with Δt_r dominates PC1 and the peak shape for ΔA dominates PC2. Thus the shifts in retention time had the greatest influence on the peaks, and the first component alone could explain 64.4% of the variations. The distortion of the design pattern shows the influence of non-linearities and interactions between the parameters.

A similar design for deviations in σ and τ, and the resulting curves for the corner points, are shown in Fig. 3 together with the results of PCA. Now the design seems to be tilted in the score plot, and the two loading plots are actually linear combinations of the two partial derivatives. Thus PCA does not separate the true sources of variations, the principal components are constructed as linearly independent combinations (orthogonal loading vectors). In the preceding case the two partial derivatives were actually
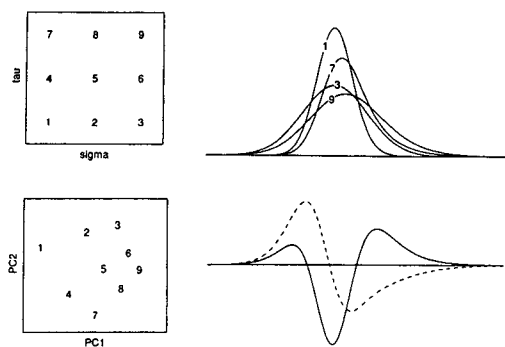
Fig. 3. Top: design of variations in σ and τ and the EMG peaks for the corner points, ($\pm\Delta\sigma$, $\pm\Delta\tau$). Bottom: score plot for PC1 and PC2 (left) and the loadings for PC1 (solid line) and PC2 (dashed line).

Fig. 4. Top: design of area variations for the two peaks (A_1 = 1, ΔA_1 = ±0.2, A_2 = 0.8, ΔA_2 = ±0.2) and the EMG peaks for the corner points. Bottom: score plot for PC1 and PC2 (left) and the loadings for PC1 (solid line) and PC2 (dashed line).

almost linearly independent (orthogonal functions), which explains the successful separation. For the deviations in σ and τ the first component, PC1, accounts for 74.7% of the variations, and the second, PC2, for 22.6%. Again, non-linearities and interactions are responsible for the distortion of the regular design pattern.

It should be noted that the main features of the parameter design patterns were revealed by PCA without any evaluation of the parameter values from the curves. Such values must be obtained from the statistical moments, with much problems related to the baseline, or by non-linear regression, which for the EMG function is not a trivial task.

2.4. PCA for multiple peak variations

With more than one peak in the chromatogram, there are more parameters with possible deviations to be accounted for. A simple example will show some features of PCA in the case of multiple peaks. In Fig. 4 the design for area deviations in two peaks is shown, together with the curves corresponding to the corners. The results from PCA, i.e., the scores and the loadings, are also shown in Fig. 4. Again, the design pattern is tilted in the score plot and the diagonal direction of PC1 maximizes the variations.

When the area for two peaks varies independently, the loading plots for PC1 and PC2 will

reflect the sum and difference, respectively, of the single peak variation pattern. In a real case, however, the area variations are often coupled. Variations in the injected amount of sample have the same influence on both peaks, corresponding to one principal component with loadings in accordance with the sum (cf., PC1 in Fig. 4). On the other hand, another common situation is variations in the area distribution between peaks. When this is the case, the loadings for the corresponding principal component will be the difference (cf., PC2 in Fig. 4).

3. Alignment of chromatographic profiles

When PCA is applied to a set of chromatographic profiles with multiple peaks, several types of variation sources may be encountered. Usually we are looking for variations between samples, and the variations induced by the chromatographic procedure are then a severe complication. The maximum variation principle in PCA implies mixing of these sources within the principal components, and the variations of interest may be difficult to discern. In general the main source of variation is small shifts in retention time, systematic for all peaks as well as random for individual peaks, caused by, for example, variations in flow-rate, mobile phase composition, or gradient slope. Furthermore, the

overall variations in the chromatographic signal due to the injected amount of sample, detector sensitivity, etc., are reflected in the first principal components. The true sample variations may be small in comparison, and hence difficult to distinguish in the results of PCA. To detect differences between the samples that are independent of the chromatographic variations, these variations have to be reduced. Otherwise, the retention shifts and differences in the injected amount may conceal significant information regarding the samples.

The unwanted variations can be minimized by increased reproducibility in the chromatographic procedure, but multivariate analysis with chromatographic profiles as input data is sensitive to even minute variations. This requires a post-chromatographic alignment of the profiles, much in the same spirit as the retention time matching methods described in the Introduction. However, it must be stressed that chromatographic profiles as input data requires a quantitative alignment of the time scale in contrast to the more qualitative matching of peaks in the former instance.

As early as 1979, Reiner et al. [17] suggested a method to compensate for retention shifts. Each data point in the profile was adjusted towards a reference chromatogram according to a "time-warping" function. The aligned chromatograms were used for visual comparisons only, without any use of multivariate methods. Another solution has recently been proposed by Andersson and Hämäläinen [18], commercially available as the software ChromPro [19]. The entire profile is adjusted towards a selected target chromatogram, using two parameters corresponding to linear displacement of the profile and compression/expansion of the time scale, respectively. The parameter values are determined by non-linear regression using a simplex procedure, maximizing the correlation within two selected retention windows. The aligned chromatogram is calculated by linear interpolation between the two windows, preferably situated in the start and the end of the chromatogram. This approach is useful for chromatograms with relatively broad peaks. In some instances, the chromatographic

profile has to be divided into several segments, each aligned individually [18]. A similar procedure has been used by Wathelet and Marlier [20] for alignment of migration distances in electrophoresis. In the work of Armanino et al. [14], the chromatograms are synchronized using multiple peaks prior to the window summation. Unfortunately, no details were given about the synchronizing algorithm.

The high peak capacity and separation power necessary for peptide mapping fingerprints often require segmented gradients in order to obtain adequate resolution within a reasonable analysis time [21]. Linear expansion or compression of the time scale will not be sufficient in these cases. The large number of peaks and the sensitivity of the retention towards the mobile phase composition further emphasize that the necessary alignment function may be non-linear. This requires a more elaborate method with individual alignment of peaks throughout the whole chromatogram. A method for such alignment for NMR profiles has been used by Vogels et al. [22]. In this instance, individual groups of lines in the two spectra are matched and brought to the same resonance position. In chromatography, however, it is seldom possible to match all peaks individually. There is a need for some interpolation of the alignment between the selected peaks.

Normalization to constant area is a common procedure used to compensate for the different amounts of injected sample. This method implies so-called closure of the data set, i.e., if one peak increases the size of other peaks must decrease [23]. This may lead to artificial correlations in the data set and thus degrade the quality of the data. In the proposed method, a selective normalization is made, by considering only selected peaks in the calculation of the normalization factor.

In this work, a combined procedure was developed to reduce the chromatographic variations. The idea is to align the sample chromatograms towards a target chromatogram in order to compensate for (i) small shifts in retention time (not due to different sample components), (ii) common variations in peak area (not due to

sample composition) and (iii) variations in level and slope of the baseline.

The alignment procedure consists of several steps, which will be illustrated with an example concerning peptide mapping. From a set of chromatograms obtained as described under Experimental, one representative chromatogram was selected as the target chromatogram. The sample chromatogram was arbitrarily chosen from the others. Each chromatogram is represented by 4900 data points, the peak width corresponding to about 30 points. Later the effects on PCA for the total set, when all chromatograms are aligned to the selected target, will be discussed.

3.1. Comparative chromatogram plot

To facilitate the alignment procedure a descriptive plot of the chromatograms is utilized. The sample chromatogram (the one to adjust) is plotted versus the target chromatogram in a comparative chromatogram plot (CCP). The features of this plot is demonstrated by a simplified example. Two simulated chromatograms (Fig. 5, left), with peaks differing in selected ways, illustrate the corresponding effects on the CCP (Fig. 5, right). Especially small retention
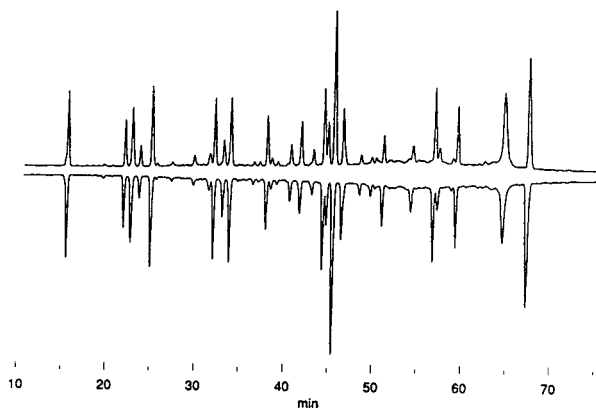
Fig. 6. Mirror plot of (top) the sample and (bottom) the target chromatograms.

shifts, the main obstacle when applying PCA to chromatograms, are clearly revealed as "loops". Although well indicated by a curved line, deviations in band broadening (σ) are not accounted for in the alignment procedure to be described.

For our real example the two chromatograms are shown in the traditional "mirror plot" (Fig. 6) and also the comparative chromatogram plot (Fig. 7). The dominant feature in the CCP is the wide loops, indicating that the retention time
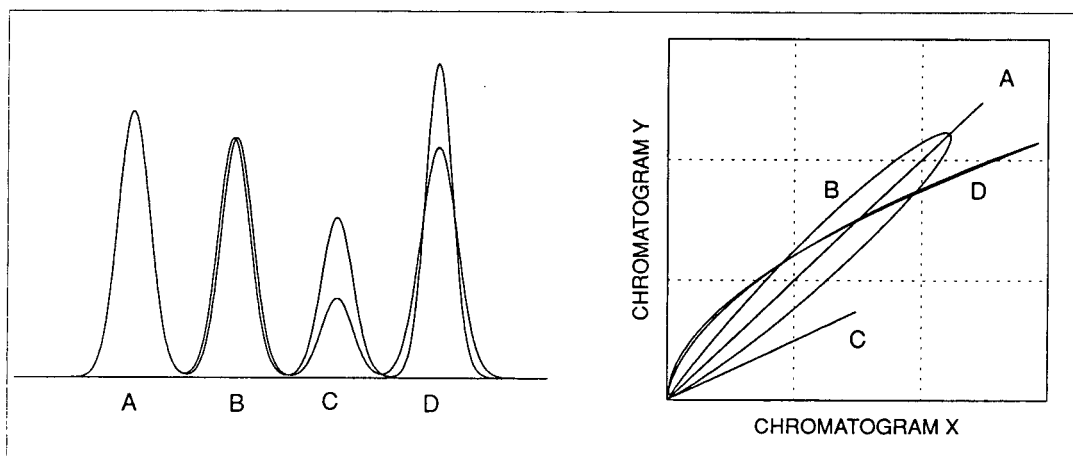
Fig. 5. Left: two chromatograms with four EMG peaks: A = no parameter shifts; B = shift in t_r; C = shift in A; D = shift in σ. Right: the CCP of the two chromatograms.
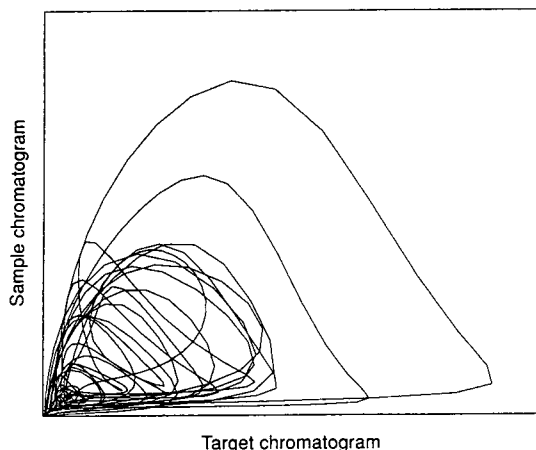
Fig. 7. The CCP for the untreated sample chromatogram.

scale is shifted between the two chromatograms (cf. peak B in Fig. 5).

3.2. Retention alignment

The first part of the alignment procedure is to align the retention for peaks that are assumed to correspond to the same sample component. For the two chromatograms $y_{sample}(t)$ and $y_{target}(t)$ the cross-correlation function (ccf) is defined as

$$ccf(\Delta t) = \int y_{sample}(t - \Delta t) y_{target}(t) \, dt \qquad (8)$$

which for equidistant data points is evaluated as the sum

$$ccf(\Delta t) = \sum y_{sample}(t_i - \Delta t) y_{target}(t_i) \qquad (9)$$

This function is calculated for discrete values $\Delta t = 0, \pm 1, \pm 2, \ldots$, where the integers refer to the time displacement expressed as number of sampling intervals.

In the first step, the cross-correlation function is calculated for all data points in the two chromatograms, and a maximum is obtained for some Δt. This integer time shift value is used as an overall time shift for a coarse precorrection of the test chromatogram by renumbering the data points. This corrects for any large shift of the chromatogram along the time axis.

In the second step, a limited number of peaks, those with the largest peak heights, are selected from the target chromatogram. If the number is not too high, the selected peaks correspond to main components and should appear in the sample chromatogram also, as we are dealing with sets of apparently similar chromatograms. For each of these peaks, a section of the target chromatogram is taken around the peak centre. The length of this section is typically a few peak widths, and related to the largest time shift allowed. The cross-correlation function for each section and the corresponding section of the sample chromatogram is calculated, and from its maximum the time correction for each selected peak is obtained. The position of the ccf maximum indicates how many data points (i.e., sampling intervals) the section of the sample chromatogram must be shifted to match that of the target chromatogram as close as possible.

The individual time shifts, which are valid at the centre of the selected peaks, are used to construct a time displacement function $\delta t(t)$ for the sample chromatogram. The time shift for all points between these fix points is calculated by linear interpolation as shown in Fig. 8, where the fix points are indicated by circles. By visual inspection of the time displacement function, any mis-matches for the selected peaks are detected, e.g., as obvious outliers in the curve, and can be removed.

For the sample chromatogram a corrected time scale, $t' = t + \delta t(t)$, is now applied. The sample chromatogram is evaluated at the points
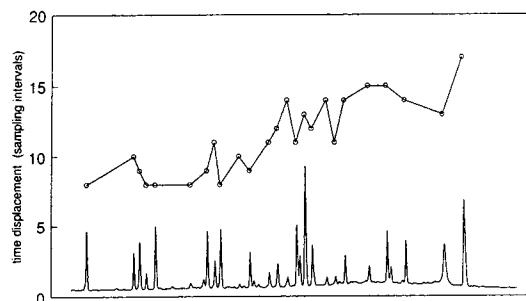
Fig. 8. The time displacement function for the coarse time alignment. The target chromatogram is shown for identification of selected peaks.

in time where the target chromatogram is sampled

$$y'_{sample}(t) = y_{sample}(t') = y_{sample}[t + \delta t(t)] \qquad (10)$$

When the corrected time value lies between two adjacent points in the sample chromatogram, $y'_{sample}(t)$ is calculated from the adjacent points by linear interpolation. The CCP for the sample chromatogram at this stage is shown in Fig. 9.

In the third step, a larger number of peaks in the target chromatogram are involved. For all peaks identified as local maxima, the cross-correlation function with the same section of the pre-aligned sample chromatogram is calculated. This time a smaller number of shifts (sampling intervals) are allowed, corresponding to about half the peak width. The maximum value of the *ccf* for each peak is used together with the peak height in the target chromatogram to sort the peaks, with highest preference for large peaks that correlates well with the corresponding portion of the sample chromatogram.

So far the time shifts have been obtained as integer values of the sampling interval. In the first two steps the sample chromatogram was actually recalculated (coarse alignment of selected main peaks). The result of the third step was a list of integer time shift values for a larger

number of peaks, but so far no adjustment was made.

Because even small variations in retention times have a great influence on the principal components, a fourth fine-tuning step is desirable. A certain number of peaks are taken from the sorted list, and for each one the *ccf* is calculated for five points around the maximum found in the previous step. To these values a polynomial is fitted, and if the maximum of the polynomial lies within the five points, the peak is a candidate for matching. Again a piece-wise linear time displacement function $\delta't(t)$ is constructed from the positions of the *ccf* maxima (now non-integer values), the validity of which should be checked by visual inspection as before.

The final alignment of the sample chromatogram according to

$$y''_{sample}(t) = y'_{sample}[t + \delta't(t)] \qquad (11)$$

is obtained in the same way as before. The comparative chromatogram plot after time alignment is displayed in Fig. 10, which should be compared with that for the untreated data (Fig. 7).

3.3. Response corrections

In Fig. 10, the CCP so far, a dashed line indicates the trace for identical peaks, i.e., with

Fig. 9. The CCP after the coarse time alignment.
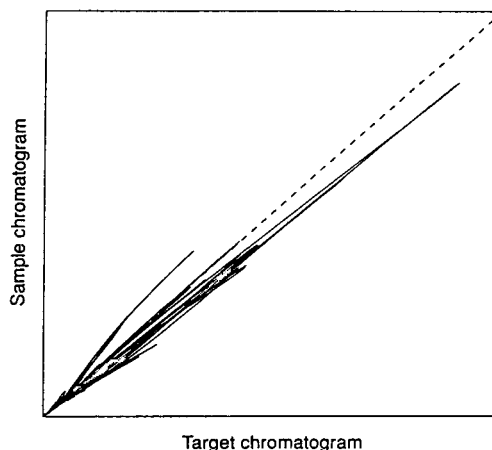
Fig. 10. The CCP after the final time alignment. The dashed line is the trace for identical peaks.

the same chromatographic response at all sampling events. It is seen that the majority of points fall on one side of this line. Obviously there is a need for a common correction factor for all peaks, but some smaller peaks do show deviations from the main trace. These true differences in peak area must be considered when the normalization is performed. The proposed method applies a least-squares fit, which also takes into account possible differences in the baseline characteristics (level and slope). The linear regression model is

$$y''_{sample}(t) = a + bt + cy_{target}(t) \tag{12}$$

The regression is obtained iteratively, initially using all data points. Then all points too far from the ideal trace are excluded from the regression, according to the condition

$$|y''_{sample}(t) - a - bt - cy_{target}(t)| < k + ly_{target}(t) \tag{13}$$

The constants k and l define a wedge-shaped strip around the ideal diagonal trace, and points outside this strip are rejected. The regression is repeated without these points, leading to new values for a, b and c and a new test for exclusion. Finally, the set of excluded points is constant, which ends the regression procedure.
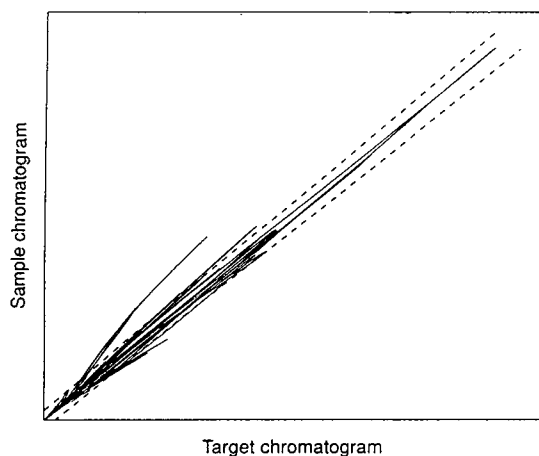
Fig. 11. The CCP after time alignment and response correction. The dashed lines indicate conditions for data points used in regression.
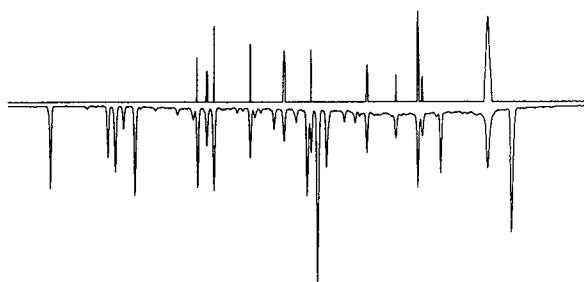
Fig. 12. Peaks with true area differences in the sample chromatogram (top) compared to the target chromatogram (bottom).

The final correction of the sample chromatogram is then

$$y'''_{sample}(t) = [y''_{sample}(t) - a - bt]/c \tag{14}$$

and the resulting comparative chromatogram plot is shown in Fig. 11, where also the lines of rejection are indicated. The segments that fall outside these lines are regarded as peaks with a true area difference, and a chromatogram where only such peaks are shown is a tool for pinpointing the sample differences (Fig. 12). A similar method for response correction, multiplicative scattering correction (MSC), has been used to pretreat near-infrared reflectance spectra prior to multivariate calibration [24].

4. Experimental

4.1. Tryptic digests

The tryptic digests of equine cytochrome c (Sigma, St. Louis, MO, USA) were prepared according to the procedure described by Renlund et al. [25], with the exception that the concentration of trypsin (Sigma) was decreased to 0.2 μg/μl [26]. The procedure was also scaled up five times, by increasing the volumes in all steps. Four replicated preparations of the reagents for denaturation and cysteine reduction, desalting buffer and the trypsin solution were used.

4.2. Chromatographic procedure

The tryptic digests were injected by a CMA 200/240 refrigerated (4°C) autosampler (CMA Microdialysis, Stockholm, Sweden), and separated on a SuperPac Pep-S C_2/C_{18} (5 μm, 100 Å) column (250 × 4 mm I.D.) using a precolumn (10 × 4 mm I.D.) packed with the same material. The separations were performed with a Model 2249 gradient pump with detection at 215 nm by a Model 2141 variable-wavelength monitor. The chromatographic system was controlled by HPLCmanager software, also used to store the chromatograms prior to the multivariate analysis. All chromatographic columns and instruments were from Pharmacia Biotech (Uppsala, Sweden), except where indicated.

The separations were performed by gradient elution (flow-rate 1 ml/min). The mobile phases were consistently prepared by weighing instead of volumetric measurements. The aqueous phase (A) consisted of 50 mM phosphate buffer (pH 2.5), prepared by mixing fixed amounts of stock solutions of phosphoric acid and sodium dihydrogenphosphate (both from Merck, Darmstadt, Germany). The organic phase (B) consisted of acetonitrile–A (80:20). The acetonitrile was of gradient grade (Merck). The mobile phases were degassed by sparging with helium for 5 min (A) and 10 min (B). The samples (125 μl) were eluted with a linear gradient from 0 to 60% B in 96 min, corresponding to a gradient slope of 0.5% acetonitrile/ml. All calculations were implemented in the programming environment ASYST (Macmillan Software, New York, USA).

5. Results and discussion

To test the performance of the proposed alignment and selective normalization procedure, 27 replicated digests of equine cytochrome c were prepared as described under Experimental. The digestions were divided into two sets, consisting of fourteen and thirteen samples, respectively. The protein samples in each set were simultaneously digested in the same ther-

mostated digestion block. Within each set, digestions were made with two separate preparations of all reagents, e.g., denaturing agent, cysteine reducing agent and trypsin solution. After the complete digestion, the digests were stored at 4°C and chromatographed twice, each replicate performed with freshly prepared mobile phases. The time between the digestion and the last injection were in all instances less than 4 days (this storage time can be accepted, as shown by Dougherty et al. [27]). The training set for PCA consists of 54 chromatograms, as described in Table 2.

The training set was pretreated by the procedure presented above. A suitable target chromatogram can be chosen by arbitrarily selecting one of the training set chromatograms with intermediate retention for the majority of the peaks. Another approach is to perform PCA on the raw, non-aligned, chromatograms and select the chromatogram that correlates best with the first principal component [20]. For this data set, chromatogram No. 27 (see Table 2) was selected, as it fulfilled both these criteria.

The proposed alignment and selective normalization procedure was tested by performing PCA on the data set in different phases of the procedure. The three versions of the data set correspond to (i) raw data, (ii) retention aligned chromatograms and (iii) the final data set where the selective normalization and baseline correction is included. It is important to realize that the pretreatment of the chromatograms has reduced the total amount of variation by removing the retention shifts and differences in the injected amount. The total sum of squares is calculated by

$$SS_{\text{tot}} = \sum_{i=1}^{N} \sum_{j=1}^{P} [y_i(t_j) - \bar{y}(t_j)]^2 \qquad (15)$$

where N and P are the number of objects and data points (variables), respectively. The average chromatogram is denoted by $\bar{y}(t_j)$. Comparison of the total sum of squares between the three phases of the pretreatment (see Table 3) shows that the variations have been reduced to about 1.6% of the initial amount.

Table 2
Design of the training set

Digest	Amount of protein (mg)	Chromatogram	Digestion set	Reagent preparation	Mobile phase preparation
1	0.62	1/15	1	a	A/C
2	0.56	2/16	1	b	A/C
3	0.64	3/17	1	a	A/C
4	0.48	4/18	1	b	A/C
5	0.65	5/19	1	a	A/C
6	0.53	6/20	1	b	A/C
7	0.48	7/21	1	a	A/C
8	0.55	8/22	1	b	B/D
9	0.47	9/23	1	a	B/D
10	0.53	10/24	1	b	B/D
11	0.45	11/25	1	a	B/D
12	0.59	12/26	1	b	B/D
13	0.60	13/27[a]	1	a	B/D
14	0.55	14/28	1	b	B/D
15	0.47	29/42	2	c	E/G
16	0.51	30/43	2	d	E/G
17	0.45	31/44	2	c	E/G
18	0.49	32/45	2	d	E/G
19	0.44	33/46	2	c	E/G
20	0.61	34/47	2	d	E/G
21	0.52	35/48	2	d	F/H
22	0.61	36/49	2	c	F/H
23	0.55	37/50	2	d	F/H
24	0.57	38/51	2	c	F/H
25	0.43	39/52	2	d	F/H
26	0.53	40/53	2	c	F/H
27	0.51	41/54	2	d	F/H

Each digest was chromatographed twice with replicated mobile phase preparations.
[a] Indicates the selected target chromatogram.

Table 3
Results from PCA on the data set in different phases of the pretreatment

Data set	SS_{tot} [a]	Rank[b]	Explained variance (% of SS_{tot})						Total[c]
			PC1	PC2	PC3	PC4	PC5	PC6	
Raw data	1186	6	41.2	24.1	16.0	7.6	3.1	2.0	94.0
Aligned data	64.7	4	84.8	9.9	1.6	1.4			97.7
Final data	19.2	5	67.8	10.9	6.8	5.5	2.4		93.4

[a] Total sum of squares in the data set, calculated by Eq. 15.
[b] Number of significant principal components according to cross-validation.
[c] Cumulative explained variance with the significant components.

5.1. Characterization by principal component analysis

The data set consists of 54 objects (chromatograms), each described by 4900 variables (sampling interval 0.8 s). Before the PCA, the (54 × 4900)-dimensional matrix was mean centred, i.e., for each individual variable the mean over all chromatograms was subtracted from each chromatogram.

The number of significant principal components is not very important if the main purpose of the PCA is to characterize the data set, looking for patterns and groupings between the chromatograms. The principal components are calculated so that the amount of variance that is explained by the individual components decreases for each additional component (see Table 3). This means that the main information regarding the variations in the data set is found in the first few components. In this case it is sufficient to calculate an arbitrarily chosen number of principal components and then examine the score plots for the first components. Nevertheless, cross-validation [16] can be used to determine the number of components corresponding to the model with the best predictive ability.

To assess the influence of the chromatographic variations in the data set, PCA was first performed on the raw, non-aligned chromatograms (see Table 3). The loading plot for the first principal component, explaining about 41% of the initial variation, is shown in Fig. 13. From
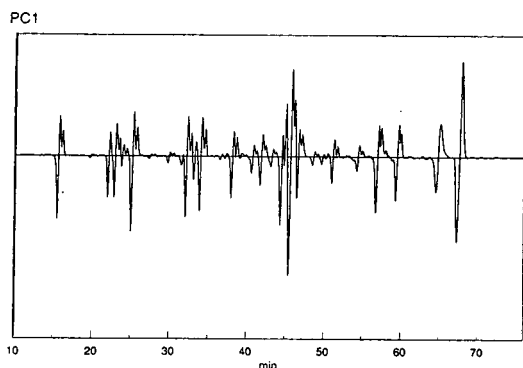
the complicated pattern, and the many regions resembling the first time derivative, it is obvious that the retention shifts influence the data set to a very high degree. The score plot of the two first principal components for the raw data (Fig. 14) shows a strong clustering according to the mobile phase preparations. The replicated chromatograms of the same digest are in most instances very far from each other. The pattern can be interpreted as a general trend, going from the first chromatograms in the upper right part of the score plot via the middle left part to the last chromatograms in the lower right part. This indicates that the underlying factor might be related to time and not to random deviations in the mobile phase composition. One possible explanation is a gradual degradation of the performance of the chromatographic column, e.g., by the acidic mobile phase. Such prominent variations caused by the chromatographic process will certainly obscure the interesting, sample-dependent, variations in the data set. The use of a multivariate classification method for detection of abnormal samples will probably not be successful with non-aligned chromatograms.

In the next phase, PCA was performed on the retention aligned chromatograms, prior to the selective normalization and baseline adjustment (see Table 3). The loadings for PC1, explaining about 84% (data not shown), are all positive and the overall pattern is very similar to PC0, i.e., the average chromatogram. The similarity can be quantified by calculation of the correlation between the variable averages and the loadings for PC1. The good correlation, $r = 0.96$, indicates that the dominant source of variation is related to the injected amount of sample. The correlation between the scores in PC1 and the initial amount of protein (see Table 2), is lower, $r = 0.88$. This suggests that the injected amount is influenced also by other factors, e.g., the recovery in the sample pretreatment or the digestion efficiency. The differences in the injected amount will prevent the characterization of variations between samples.

Finally, after retention alignment, selective normalization and baseline adjustment, the data set was characterized by PCA (see Table 3). If

Fig. 13. PC1 loadings with the raw data set.
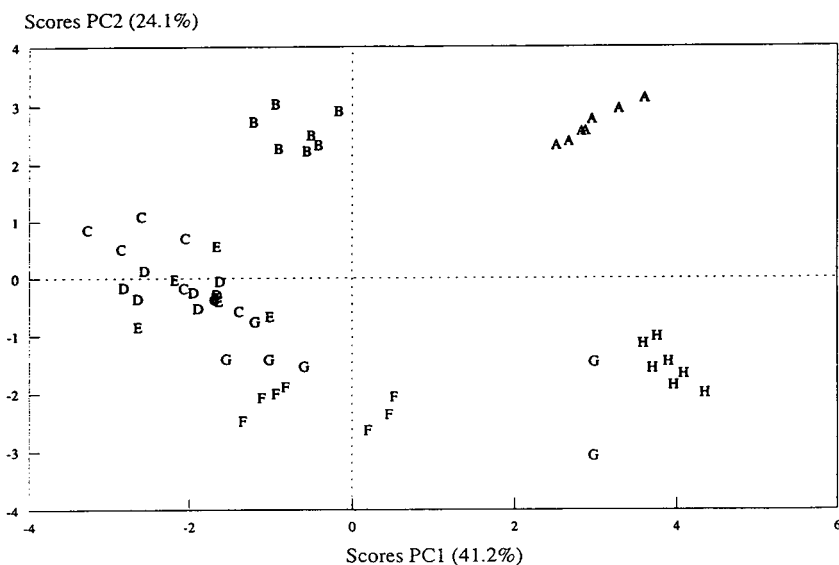
Fig. 14. Score plot of the two first principal components calculated for the raw data. The letters A–H designate the mobile phase preparations (see Table 2).

the scores for the two first principal components (together explaining about 79% of all variations in the final data set) are plotted against each other (see Fig. 15), an interesting pattern is revealed. The objects (chromatograms) are separated according to the preparation of the reagents, with the preparation labelled "a" (see Table 2), situated in the upper part of the score plot, i.e. with high scores in the second component. In the lower part of the score plot, a less
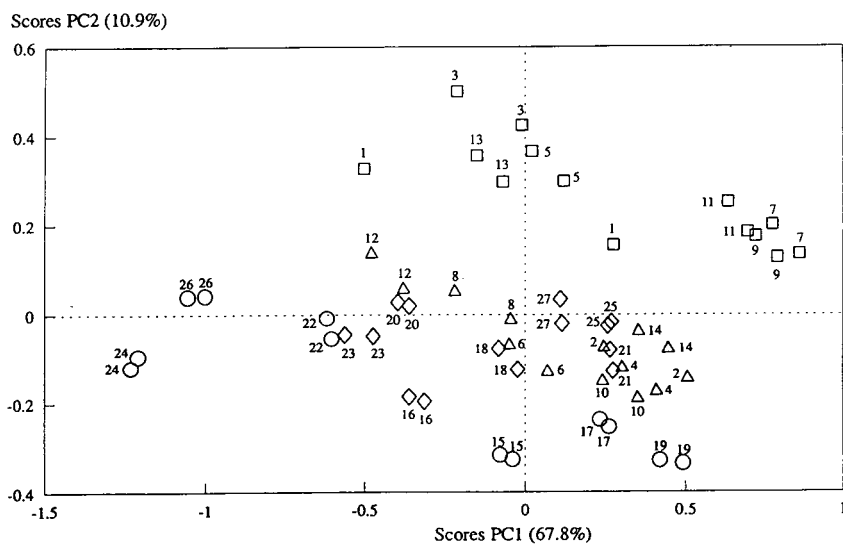
Fig. 15. Score plot for the two first principal components calculated for the final data set. The symbols denote the replicated reagent preparations (see Table 2): □ = a; △ = b; ○ = c; ◇ = d. The numbers indicate the individual digests (see Table 2).

pronounced separation is observed for the preparation labelled "c". The two preparations "b" and "d" are not separated and situated in the middle of the score plot.

The different digests (see Table 2) are also indicated in Fig. 15. The two replicated chromatograms of the same digest are generally close to each other in the score plot, showing that the chromatographic variations are to a large extent removed from the data set. One exception can be seen for digest 1, where the two chromatograms are fairly far apart in the score plot. This deviation can be mainly attributed to the observed large difference in the peak, probably corresponding to undigested cytochrome c, as discussed below.

No systematic information that can be easily interpreted is found in the scores for the higher principal components (data not shown), but these components are nevertheless important to characterize the data set fully.

The loading plot for the first principal component, explaining about 68% of all variations, is shown in Fig. 16. The component is dominated by negative loadings for the broad peak eluted at approximately 64 min (cf., Fig. 6). This peak is believed to be connected with the undigested cytochrome c, which means that the most important variation in the final data set is the degree of digestion. It is also natural that this variation is preserved after the selective normalization procedure, as the amount of undigested

protein is negatively correlated with the amount of digested protein, i.e., the majority of peaks. The first component is not correlated with the average chromatogram (by loadings) or the initial amount of protein (by scores).

The second component, responsible for the separation between reagent preparations, has a more complicated pattern (Fig. 17). Some parts of the loading plot, e.g., the first peak at 16 min and the last peak at 68 min, show an anomalous pattern. This indicates that there are remaining variations in peak shape.

There are some types of chromatographic variations that cannot be removed by the current method. Unfortunately, there is to our knowledge no other procedure that would be successful in the following cases either. In the presence of overlapping peaks, it is not possible to correct for variations in the overlap between the peaks. This could probably be achieved by fitting an appropriate peak shape model, e.g., Gaussian or exponentially modified Gaussian, to each peak in the chromatogram. These fitted peaks could then be aligned and the resulting chromatogram calculated. This is not an attractive solution, however, owing to the large number of peaks present in most fingerprint chromatograms. The proposed strategy does not compensate for variations in peak width and tailing. Such variations could be caused by column degradation, which is a serious problem in all fingerprinting methods, regardless of the type of evaluation and interpretation that is used [12].
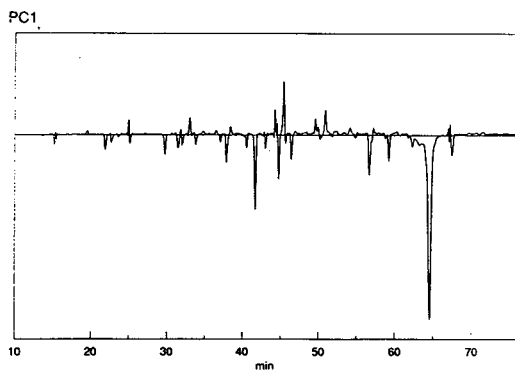
Fig. 16. PC1 loadings for the final data set.
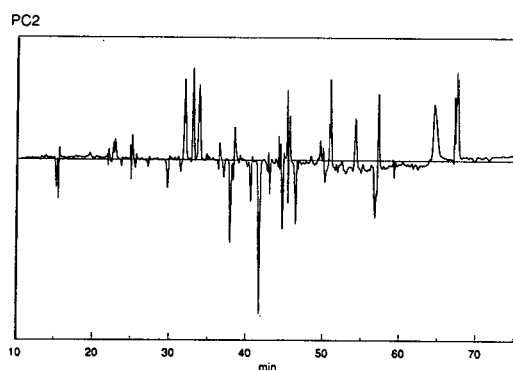
Fig. 17. PC2 loadings for the final data set.

The proposed alignment procedure is nevertheless capable of reducing the chromatographic variations in the current data set to an acceptable level. This is indicated in the score plot (Fig. 15) by the generally small deviations for the replicated chromatograms of the same digest. The characterization by PCA could describe differences between the digests that were unrelated to variations in peak overlap and column degradation.

6. Conclusions

By using the proposed retention alignment procedure and the selective normalization of peak heights, it is possible to perform principal component analysis on complex chromatographic data sets. When there are many peaks in the chromatograms, and there is a possibility that peaks can disappear and new peaks appear, it is beneficial to use the whole chromatographic profile for the analysis. The proposed method can be made fully automatic and allows the processing of numerous chromatographic profiles in a data set.

For the peptide mapping data set used in this study it was possible to detect differences between the digests that would have been obscured by the chromatographic variations if proper alignment had not been performed.

The proposed pretreatment method is also applicable to other situations where complex chromatographic data sets are treated by multivariate data analysis, e.g., pyrolysis–GC [3]. In the accompanying paper [7], multivariate classification of tryptic digests is suggested as an objective evaluation method for peptide mapping.

Appendix

Despite the complex appearance of the EMG function in the time domain (Eq. 1), the calculations can readily be performed in the frequency domain. The Fourier transform of the EMG function is

$$\tilde{y}(\omega) = A \exp[-(\omega\sigma')^2] \exp(-j\omega t_r')/(1 + j\omega\tau') \tag{A1}$$

The primed versions of the time related parameters t_r, σ and τ are scaled with the constant $2\pi/T$, where T is the duration of the time function.

The equation can be interpreted according to three frequency-dependent factors:

(i) $\exp[-(\omega\sigma')^2]$, a Gaussian peak of width σ with unit area around $t = 0$;

(ii) $\exp(-j\omega t_r')$, shifting this peak along the time axis by t_r;

(iii) $1/(1 + j\omega\tau')$, convolution of the shifted peak with an exponential decay with time constant τ (tailing).

Finally this unit area peak is multiplied with the area parameter A.

The transform can easily be calculated for $\omega = 0, 1, 2, \ldots, N$, and the time function is then obtained for $2N$ equally spaced points in time by the inverse fast Fourier transform (IFFT). All simulated chromatograms with EMG peaks in this work were obtained in this way.

Moreover, the separation of the parameters in the four factors facilitates the calculation of the partial derivatives:

$$\frac{\partial \tilde{y}}{\partial A} = (1/A)\tilde{y}(\omega) \tag{A2}$$

$$\frac{\partial \tilde{y}}{\partial t_r'} = (-j\omega)\tilde{y}(\omega) \tag{A3}$$

$$\frac{\partial \tilde{y}}{\partial \sigma'} = -2\sigma'(\omega^2)\tilde{y}(\omega) = 2\sigma'(j\omega)(j\omega)\tilde{y}(\omega) \tag{A4}$$

$$\frac{\partial \tilde{y}}{\partial \tau'} = [(-j\omega)/(1 + j\omega\tau')]\tilde{y}(\omega) \tag{A5}$$

By including the scaling constant $2\pi/T$, the partial derivatives with respect to the original, unprimed, parameters are obtained. They can be

transformed to the time domain by IFFT. However, the shape of the partial derivatives in the time domain can be predicted without numerical calculations. Apart from constants, i.e., factors not containing ω, the derivatives are the original EMG function, possibly multiplied with the factors $j\omega$ and $1/(1 + j\omega\tau)$. Multiplication with these factors in the frequency domain corresponds to the time operations d/dt (time differentiation) and convolution with $\exp(-t/\tau)$, respectively. Thus partial differentiation corresponds to the following shape modifications of the EMG time function:

$\partial/\partial A$ unmodified;

$\partial/\partial t_r$ time differentiation;

$\partial/\partial\sigma$ time differentiation twice;

$\partial/\partial\tau$ time differentiation and exponential convolution.

References

[1] L.M. Headley and J.K. Hardy, *J. Food Sci.*, 57 (1992) 980.

[2] J.S.C. Smith and O.S. Smith, *Adv. Agron.*, 47 (1992) 85.

[3] J.A. Pino, J.E. McMurry, P.C. Jurs, B.K. Lavine and A.M. Harper, *Anal. Chem.*, 57 (1985) 295.

[4] F.E. Regnier, *LC·GC*, 5 (1987) 392.

[5] F.E. Regnier, *LC·GC*, 5 (1987) 472.

[6] W.S. Hancock, *LC·GC*, 5 No. 4 (1992) 30.

[7] G. Malmquist, *J. Chromatogr.* 687 (1994) 89.

[8] S. Wold, K. Esbensen and P. Geladi, *Chemometr. Intell. Lab. Syst.*, 2 (1987) 37.

[9] S. Wold, C. Albano, W.J. Dunn, III, U. Edlund, K. Esbensen, P. Geladi, S. Hellberg, E. Johansson, W. Lindberg and M. Sjöström, in B.R. Kowalski (Editor), *Chemometrics: Mathematics and Statistics in Chemistry*, Reidel, Dordrecht, 1984, p. 17.

[10] N.R. Crawford and W.W. Hellmuth, *Fuel*, 69 (1990) 443.

[11] H.T. Mayfield and W. Bertsch, *Comput. Appl. Lab.*, 1 (1983) 130.

[12] J.A. Pino, J.E. McMurry, P.C. Jurs, B.K. Lavine and A.M. Harper, *Anal. Chem.*, 57 (1985) 295.

[13] M.E. Parrish, B.W. Good, F.S. Hsu, F.W. Hatch, D.M. Ennis, D.R. Douglas, J.H. Shelton and D.C. Watson, *Anal. Chem.*, 53 (1981) 826.

[14] C. Armanino, M. Forina, L. Bonfanti and M. Maspero, *Anal. Chim. Acta*, 284 (1993) 73.

[15] J.P. Foley and M.S. Jeansonne, *J. Chromatogr. Sci.*, 29 (1991) 258.

[16] S. Wold, *Technometrics*, 20 (1978) 397.

[17] E. Reiner, L.E. Abbey, T.F. Moran, P. Papamichalis and R.W. Schafer, *Biomed. Mass Spectrom.*, 6 (1979) 491.

[18] R. Andersson and M.D. Hämäläinen, *Chemometr. Intell. Lab. Syst.*, 22 (1994) 49.

[19] *ChromPro*, BioTriMark, Björkkulla, Funbo, Uppsala, Sweden.

[20] B. Wathelet and M. Marlier, *Chemometr. Intell. Lab. Syst.*, 4 (1988) 327.

[21] R.C. Chloupek, W.S. Hancock and L.R. Snyder, *J. Chromatogr.*, 594 (1992) 65.

[22] J.T.W.E. Vogels, A.C. Tas, F. van den Berg and J. van der Greef, *Chemometr. Intell. Lab. Syst.*, 21 (1993) 249.

[23] E. Johansson, S. Wold and K. Sjödin, *Anal. Chem.*, 56 (1984) 1685.

[24] P. Geladi, D. MacDougall and H. Martens, *Appl. Spectrosc.*, 39 (1985) 491.

[25] S. Renlund, I.-M. Klintrot, M. Nunn, J.L. Schrimsher, C. Wernstedt and U. Hellman, *J. Chromatogr.*, 512 (1990) 325.

[26] S. Renlund, personal communication, 1992.

[27] J.J. Dougherty, Jr., L.M. Snyder, R.L. Sinclair and R.H. Robins, *Anal. Biochem.*, 190 (1990) 7.

Multivariate evaluation of peptide mapping using the entire chromatographic profile [☆]

Gunnar Malmquist[1]

Institute of Chemistry, Department of Analytical Chemistry, Uppsala University, P.O. Box 531, S-751 21 Uppsala, Sweden

Abstract

Peptide mapping is an important analytical technique for quality control of rDNA-derived proteins. The evaluation in peptide mapping is complicated by variations in the digestion and the chromatographic separation. The variation sources in peptide mapping are briefly reviewed. A multivariate evaluation method that can account for the digestion variations is presented. The method utilizes the entire chromatographic profile as input data, eliminating the need for peak size and retention time determinations. The influence of chromatographic variations should be reduced by proper pretreatment of the chromatograms, in order to allow classification of protein samples. The method is intended to facilitate the evaluation in peptide mapping and is capable of handling numerous chromatograms in a data set.

1. Introduction

Many therapeutically important peptides and proteins, e.g., insulin and growth hormone, are currently produced by the recombinant DNA (rDNA) technique [1]. This sophisticated technology is based on insertion of foreign genetic material, coding for the substance of interest, into a host cell. The host cells will then produce the desired substance in addition to their natural production. Biotechnological production of pharmaceuticals requires very rigorous quality con-

trol owing to the risk of undesirable protein modifications and the numerous possibilities for contamination of the product. The integrity of the amino acid sequence of the protein has to be confirmed for each production batch. An introduction to the rDNA technique and the analytical aspects of quality control in biotechnology was given by Garnick et al. [2].

Peptide mapping is an indispensable analytical method in biotechnology for quality control of rDNA-derived proteins [2]. Peptide mapping consists of fragmentation of the protein by enzymatic digestion or chemical cleavage, with subsequent separation of the fragments. The fragmentation is in most instances performed by enzymatic digestion with trypsin [3], while the separation is usually performed by gradient elution reversed-phase liquid chromatography

---

☆ Parts of this material were previously presented at *Analysis of Peptides, Stockholm, 2–4 June, 1993,* and *Analysdagarna, Lund, 14–18 June 1993.*
[1] Present address: Pharmacia Biotech, R&D, S-751 82 Uppsala, Sweden.

(RPLC) [4]. Digestion with trypsin gives specific cleavage of the protein at the C-terminal side of arginine and lysine residues, resulting in a large number of fairly short fragments with average size 7–12 residues [3]. This requires a high peak capacity in the chromatographic separation and often necessitates the use of segmented gradients. The high resolving power obtained with shallow gradients in RPLC [5] is one of the main reasons for the popularity of RPLC in peptide mapping.

The chromatogram can be regarded as a fingerprint of the protein, where the overall appearance is used to assess the integrity of the amino acid sequence. Modification of one amino acid will alter the properties of one fragment, which may be detected as a change in retention for that fragment. For instance, substitution of a single amino acid in tissue-type plasminogen activator ($M_r \approx 64\,000$) leads to a significant change in the retention of one fragment [2]. The evaluation of peptide mapping is traditionally performed by visual comparison of the sample chromatogram with a reference chromatogram of a digested protein with the correct sequence.

Both mutations in the DNA sequence and translation errors may lead to incorporation of erroneous amino acids in the protein. Further variants can be formed by post-translational modifications of the protein, mainly by degradation processes. Proteolytic or chemical cleavage of the protein, oxidation of methionine residues and deamidation of asparagine residues all give rise to protein variants that may be detected by peptide mapping [1].

Other applications of the technique include the characterization of naturally occurring protein variants [6] and the identification of animal species [7]. One possible application is the detection of contaminants in the sample, but this approach has not yet been much employed, owing to the difficulty to detect small peaks in a complex chromatogram containing 20–150 peaks [8].

There are several sources of variations associated with the enzymatic digestion of protein samples, that may lead to variations in the fingerprint even for identical samples. For trypsin digestion, such deviations from the normal cleavage pattern have been reported to arise from chymotryptic cleavages, partial digestion and incomplete digestion [3,9]. The final fingerprint may also be influenced by incomplete cysteine reduction or alkylation [10]. Deamidation of asparagine or glutamine residues can be induced after the digestion, e.g. by improper storage of the digests in a non-refrigerated autosampler [10].

The commercial trypsin preparations are usually treated with L-1-tosylamide-2-phenylethyl chloromethyl ketone (TPCK) in order to reduce the chymotryptic activity in the preparation. Chymotryptic cleavages are nevertheless frequently observed, owing to the minute amounts of chymotrypsin that may remain despite the TPCK treatment [11]. Non-specific cleavage fragments can occasionally be caused by trypsin cleavages at less favourable sites, e.g., adjacent to proline residues [12]. The cleavage pattern may vary between trypsin preparations obtained from different manufacturers, and even between batches from the same vendor [11].

Chloupek et al. [9] showed an interesting example of variations in the amount of non-specific cleavage fragments. The additional peaks may reduce the possibility to detect contaminants in the sample, and should therefore be kept at a minimum. The amount of non-specific cleavages could unfortunately not be reduced by changes in the digestion conditions, e.g., buffer type, temperature and reaction time. Purification of TPCK-treated trypsin by RPLC did not reduce the chymotryptic activity either.

Partial digestion will take place if the protein contains a series of two to four adjacent basic amino acids, all potential cleavage sites. When cleavage at a random position within this series has occurred, trypsin will not cleave the terminal amino acids. This will lead to the formation of varying amounts of overlapping fragments, differing in the first and last positions [3]. Variations in the yield of the digestion leads to different amounts of the resulting fragments, and a variable amount of undigested protein remaining in

the sample. Incomplete digestion may also lead to formation of partially digested fragments.

Other important sources of variation in peptide mapping are connected with the chromatographic separation. Preparation of fresh mobile phases will inevitably introduce small differences in the pH and possibly in the amount of organic modifier. The retention of peptides in RPLC is very sensitive to the composition of the mobile phase [13]. Changes in the amount of mobile phase additives, e.g. ion-pairing agents, may influence the peak retention [14]. The reproducible generation of shallow gradients is difficult even with modern LC instrumentation [15]. This may lead to slight retention shifts, especially at the beginning of the gradient [2]. Temperature variations [14], the gradual degradation of column performance [11,16] and column to column differences [17] are additional possible sources of variations in the profile. Dong and Tran [4,8] have provided recommendations for reproducible chromatographic separations of tryptic digests.

The possible deviations from the expected cleavage pattern, together with the retention variations caused by the chromatographic process, implies that peptide mapping is a very demanding analytical technique. Development and validation of a successful peptide mapping method require great effort, where the expertise of both biochemists and analytical chemists is necessary. Despite these problems, peptide mapping is the most important technique for assessment of the amino acid sequence integrity in proteins.

The visual comparison of peptide mapping chromatograms is complicated by variations in the digestion and the chromatographic separation. The evaluation will be more or less subjective and requires great experience. A more unbiased evaluation can be made by multivariate pattern recognition methods capable of handling the experimental variations. Pattern recognition [18] is a category of chemometric methods suited for the characterization of complex data sets. A multivariate evaluation method for peptide mapping is proposed in this paper, where test sam-ples are classified by SIMCA [19], a multivariate classification method based on principal component analysis (PCA).

2. Multivariate evaluation of peptide mapping

A set of reference chromatograms, obtained for digests of samples with the correct sequence, is accumulated. The data set should cover the normal variations encountered in both the digestion and the chromatographic separation. The chromatograms are represented by the entire profile, i.e., the digitalized detector signal where each sampled data point corresponds to one variable in the data set. This is advantageous for evaluation of peptide mapping, as discussed in the accompanying paper [20]. This data set can be characterized by PCA [21], expressing the main variations in the data set. Chromatograms of test sample digests can subsequently be classified by the pattern recognition method SIMCA [19]. A brief introduction to SIMCA is given below to facilitate the discussion on the evaluation method.

Multivariate analysis of chromatographic profiles requires that the chromatographic variations are reduced. The possible differences between the samples will be obscured by the chromatographic variations, unless proper pretreatment is performed. In the accompanying paper [20], this was illustrated by simulated data and a peptide mapping data set. A method for pretreatment of chromatographic profiles, intended to remove slight retention shifts caused by the chromatographic process, was developed. Compensation for variations in the injected amount was made by a selective normalization procedure that also allowed correction for baseline differences.

2.1. Multivariate classification with SIMCA

SIMCA, an acronym for soft independent modelling of class analogy, is a multivariate classification method based on PCA [19]. The classification is based on a model of the similarities between the known members of a class in

a training set. The training set in this particular instance is composed of the accumulated reference maps, all of which belong to the same class. Each class is described by a few principal components calculated for the members of the class in the training set only. Objects, i.e., chromatograms, that belong to a class will be situated close to the hyperplane spanned by the principal components. A tolerance region is established for the class, in order to quantify the similarity between a test sample and the members of the class. The tolerance region defines the boundaries for the residuals and the normal range of scores along the principal components. This is illustrated in Fig. 1, where a hypothetical class described by two principal components is depicted. Outliers and non-members will be situated outside the tolerance region.

Classification of tryptic digests based on their chromatographic profiles constitutes a special case of SIMCA, where only the acceptable reference samples in the training set form a proper class of similar objects. Deviating test samples can be different in many disparate ways, leading to the formation of an asymmetric class
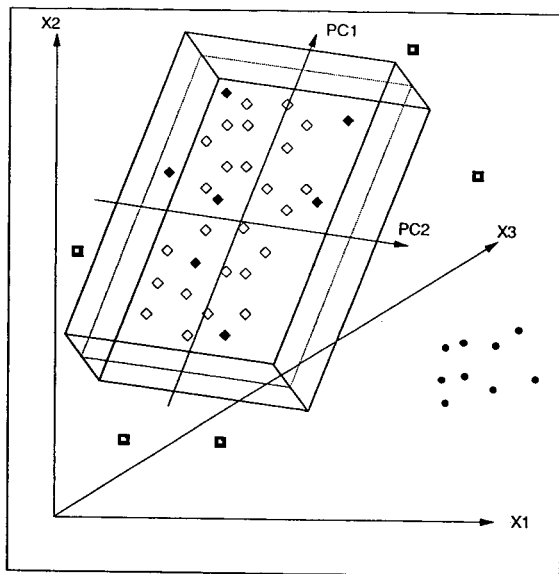
[22]. SIMCA can cope with this situation by assigning test objects to the acceptable class only if they fall within the tolerance region. All objects falling outside the region are referred to the asymmetric class of unacceptable samples.

The classification of a test sample by SIMCA is performed with an F-test according to Wold et al. [22] and Sharaf et al. [23]:

$$F = \frac{s_i^2}{s_0^2} \qquad (1)$$

where s_i^2 is the class distance for test object i, composed of the residual variance and the deviation in scores, and s_0^2 is the residual variance of the class members in the training set. The F-value is compared with a critical F-value where the degrees of freedom can be calculated from the number of variables, M, the number of objects in the training set, N, and the number of principal components used to describe the class, A. If the F-value for the test sample exceeds the critical F-value, the sample is classified as an outlier or non-member of the class.

The critical F-value can be chosen with $(M - A)$ and $(N - A - 1)(M - A)$ degrees of freedom in the numerator and the denominator, respectively [23]. In data sets with many variables this will lead to a very low critical F-value that will reject many acceptable samples. Gemperline et al. [24] addressed this problem and proposed another method to calculate the degrees of freedom for classification of test samples. When a single test sample is classified, the appropriate degrees of freedom should be one in the numerator and $(N - A - 1)$ in the denominator. The critical F-value is calculated for one-tailed tests both at the 0.10 and 0.05 significance levels. If the F-value for the test sample exceeds the critical value at the 0.05 level it is rejected, i.e. classified as unacceptable. The sample will be assigned as an outlier if the F-value is between the 0.10 and 0.05 levels. Test samples with F-values below the 0.10 level are regarded as acceptable. This approach was used in this paper for classification of the tryptic maps, together with a graphical presentation of classification results, as discussed below.

Fig. 1. SIMCA class model and tolerance region. Unfilled symbols denote objects in the training set and filled symbols refer to test set objects.

3. Experimental

3.1. Tryptic digests

The tryptic digests of equine and bovine cyto-chrome c (Sigma, St. Louis, MO, USA) were prepared according to the procedure described by Renlund et al. [25], with the exception that the concentration of trypsin (Sigma) was decreased to 0.2 μg/μl [26]. The procedure was also scaled up fivefold, by increasing the volumes in all steps. Four replicate preparations of the reagents for denaturation and cysteine reduction, desalting buffer and the trypsin solution were used.

3.2. Chromatographic procedure

The tryptic digests were injected by a CMA 200/240 refrigerated (4°C) autosampler (CMA Microdialysis, Stockholm, Sweden), and separated on a SuperPac Pep-S C_2/C_{18} (5 μm, 100 Å) column (250 × 4 mm I.D.) using a precolumn (10 × 4 mm I.D.) packed with the same material. The separations were performed with a Model 2249 gradient pump with detection at 215 nm with a Model 2141 variable-wavelength monitor. The chromatographic system was controlled by HPLCmanager software, also used to store the chromatograms prior to the multivariate analysis. All chromatographic columns and instruments were from Pharmacia Biotech (Uppsala, Sweden) except where indicated.

The separations were performed by gradient elution (flow-rate 1 ml/min). The mobile phases were consistently prepared by weighing instead of volumetric measurements. The aqueous phase (A) consisted of 50 mM phosphate buffer (pH 2.5), prepared by mixing fixed amounts of stock solutions of phosphoric acid and sodium dihydrogenphosphate (both from Merck, Darmstadt, Germany). The organic phase (B) consisted of acetonitrile–A (80:20). The acetonitrile was of gradient grade (Merck). The mobile phases were degassed by sparging with helium for 5 min (A) and 10 min (B). The samples (125 μl) were eluted with a linear gradient from 0 to 60% B in

96 min, corresponding to a gradient slope of 0.5% acetonitrile/ml.

All calculations were implemented in the programming environment ASYST (Macmillan Software, New York, USA).

4. Results and discussion

The chromatographic separation in this paper was not optimized with respect to the resolution of the fragments. The composition of the mobile phase and the shape of the gradient were merely chosen to give acceptable resolution within a reasonable analysis time. The proposed evaluation method is intended to facilitate the interpretation of the peptide mapping results, possibly without extreme requirements regarding the resolution. However, special precautions were taken to minimize the variations in the mobile phase composition. The critical aspect of the mobile phase composition in this context is reproducibility, not accuracy. It is not so important whether the pH of the aqueous phase is 2.50 or 2.55, as long as it is consistent throughout the data set. The pH of the aqueous buffer in the mobile phase is often established by titration of the acid with a base until the desired pH is achieved. Higher precision in the pH may be obtained by instead mixing stock solutions of the acid and the corresponding salt. The mobile phase preparation in this study was entirely based on weighing instead of volumetric measurements, in order to increase the reproducibility.

4.1. Description of the data sets

The training set of reference tryptic maps consisted of 50 objects (chromatograms), each described by 4900 variables, i.e., sampled data points 0.8 s apart. The training set originates from the set of 54 reference chromatograms (27 digests, each chromatographed twice) that had been characterized by PCA [20]. Two digests (corresponding to chromatograms 9, 23, 29 and 42 in the original set) were excluded from the training set.

The evaluation method was tested by simulated spiking, where one of the excluded chromatograms (29) was used as a template. Test chromatograms were obtained by addition of one Gaussian peak to the original chromatogram. The peak width of the added peak was set approximately equal to the width of the original peaks, i.e., with $\sigma = 0.1$ ml. Twenty-seven distinct peaks in the template were selected for the simulated spiking (see Fig. 2). Three positions for each selected template peak, corresponding to complete co-elution and shoulder peaks on the leading and trailing edge of the template peak, were independently used as the retention volume of the added Gaussian peak. The theoretical resolution between the template peak and the added shoulder peak was 0.5. Twelve baseline positions were also selected to estimate the level of detection for well resolved peaks. The height of the added peak was in all instances set at 3, 5, 7 and 10% of the largest peak in the template (hereafter refered to as 3–10% FS, respectively). The use of 91 peak positions and

four peak heights at each position led to a test set consisting of 364 chromatograms, each containing one added Gaussian peak. The chromatograms were pretreated according to the previously described procedure [20] and subsequently classified by SIMCA.

4.2. Multivariate classification

The training set had been characterized by PCA [20], and cross-validation [27] indicated that five principal components was optimum for the description of the training set. Inclusion of additional components in the class model may improve the classifications by SIMCA, however. Nevertheless, it is very important not to include too many components in the class model, in order to avoid bad classification results for new samples. Gemperline et al. [24] suggested that the number of components could be determined by the optimum classification accuracy. Two types of classification errors are possible which are conceptually related to the Type 1 and Type
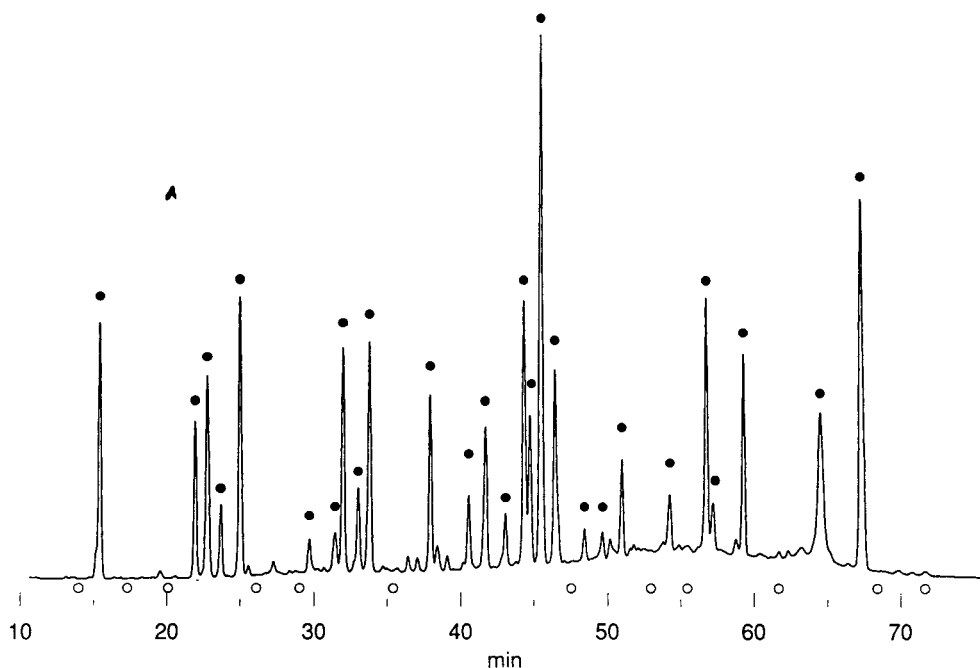
Fig. 2. Simulated spiking. A Gaussian peak is added at one position in each test set chromatogram. ● = Peak positions; ○ = baseline positions. Chromatographic conditions as in Experimental.

2 errors in significance testing that are discussed in textbooks on statistics, e.g., by Miller and Miller [28]. Error of Type 1 refers to rejection of a true null hypothesis. In the present case, this corresponds to rejection of reference chromatograms, i.e., F-values above the critical F-value at the 0.05 significance level. Inclusion of additional principal components in the class model will increase the risk of Type 1 errors as the residual variance of the class members in the training set will decrease (cf., Eq. 1). Error of Type 2, on the other hand, refers to a failure to reject a false null hypothesis, corresponding to acceptance of deviant chromatograms (F-value below the critical F-value at the 0.10 significance level). The risk for Type 2 errors will be reduced by increasing the number of principal components in the class model.

An extended cross-validation, or jack-knife, procedure was carried out to assess the classification accuracy, and to find the optimum balance between the risks for Type 1 and Type 2 errors. Independent reference chromatograms were obtained by successively excluding two acceptable digests, i.e., four chromatograms, from the training set (for this procedure, the original set of 54 reference chromatograms was used). The principal components were calculated for the remaining objects, and the excluded objects were classified according to the classification rules outlined above. This was repeated until all chromatograms in the training set had been excluded once. Fig. 3 shows the percentage of Type 1 errors from the extended cross-validation as a function of the number of components in the class model. Class models with five or six components resulted in 100% acceptance of the reference chromatograms, i.e., no Type 1 errors. No reference chromatograms were rejected if up to nine components were used in the class model; however, three chromatograms were classified as outliers.

A test set of deviant chromatograms were obtained by the simulated spiking, where one Gaussian peak had been added to the template. Fig. 4 shows the percentage of Type 2 errors for test set chromatograms with different peak heights for the Gaussian peak added as complete
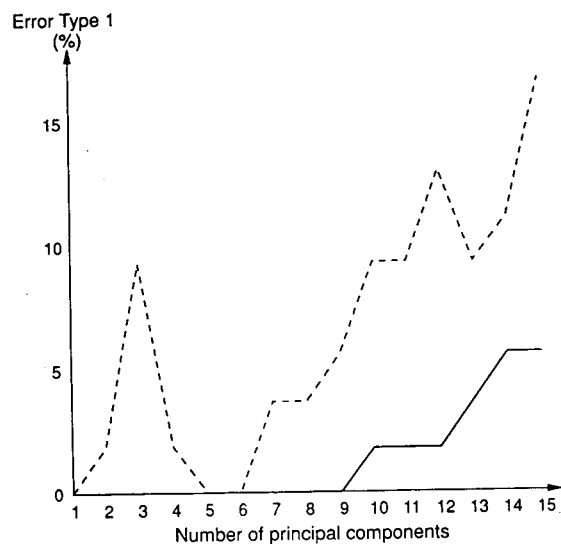
Fig. 3. Percentage of Type 1 errors, i.e., erroneous rejection of reference chromatograms. The dashed line indicates the percentage of the chromatograms that were either rejected or classified as outliers. The solid line refers to rejected reference chromatograms.
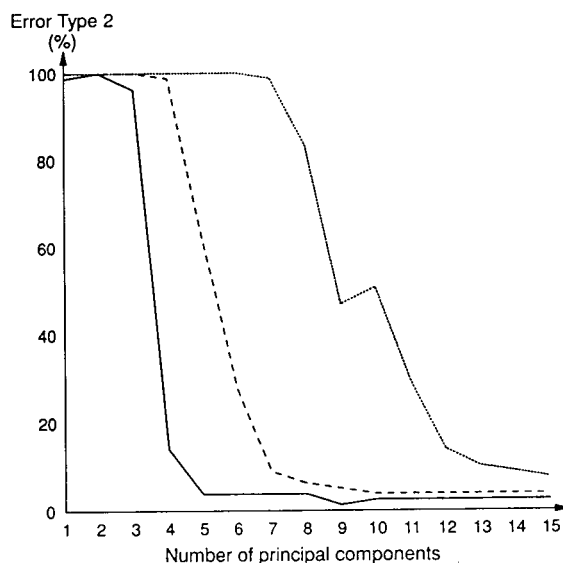
Fig. 4. Percentage of Type 2 errors, i.e., erroneous acceptance of spiked chromatograms in the test set. The solid line refers to simulated spiking with a peak height corresponding to 10% FS. The dashed and dotted lines refer to peak heights of 7 and 5% FS, respectively.

co-elution or as a shoulder on an existing peak. Chromatograms with a Gaussian peak corresponding to 10% of the largest peak in the template chromatogram are in most instances detected as deviant already with five components in the class model, while additional components are necessary to detect the smaller peaks. A class model with nine components will detect the majority of the test chromatograms where a 7% FS peak was added and about half of the chromatograms with a 5% FS peak.

Combination of these results indicates that a class model with nine principal components provides the best balance between the two types of classification errors, and is therefore considered optimum for classification purposes.

All 364 objects in the test set were classified, using nine principal components in the class model. The results are summarized in Table 1.

Table 1
Classification by SIMCA with nine principal components in the class model

Type of object[a]	Classification results[b]		
	Rejected[c]	Outliers	Accepted
Reference	0	5.6 (3)	94.4 (51)
Non-resolved, 10% FS	96.2 (76)	2.5 (2)	1.3 (1)
Non-resolved, 7% FS	77.2 (61)	17.7 (14)	5.1 (4)
Non-resolved, 5% FS	0	53.2 (42)	46.8 (37)
Non-resolved, 3% FS	0	0	100
Baseline, 10% FS	100	0	0
Baseline, 7% FS	100	0	0
Baseline, 5% FS	0	100	0
Baseline, 3% FS	0	0	100

[a] Reference refers to the 54 training set samples where the classifications were made by the extended cross-validation procedure. Non-resolved refers to chromatograms where the Gaussian peak was added as co-eluting or shoulder. Baseline refers to simulated spiking at baseline positions.
[b] The classification results are expressed in percent and the corresponding number of objects are indicated in parentheses where appropriate.
[c] Rejected refers to objects with F-values larger than the critical F-value at the 0.05 significance level. Outliers are objects with F-values between the critical F-value at the 0.05 and 0.10 significance levels. Objects with F-values below the critical F-value at the 0.05 significance level are accepted.

The peak-height level necessary for detection in case of co-elution or low resolution is about 7% relative to the largest peak in the chromatogram (see Fig. 5). Co-elution with a peak that has a large variation in the training set will reduce the possibility for detection. The majority of the objects on the 7% FS level that were erroneously accepted were spiked at the broad peak eluted at approximately 64 min. This peak exposed the largest variation in the training set, as revealed by its dominant role in the loadings on the first principal component [20]. New peaks at the 5% FS level may be detected if they are baseline resolved, although only as outliers. The detection limit can thus be improved by optimization of the chromatographic separation. Segmented gradients could be utilized to increase the resolution between the fragments in some parts of the chromatogram. This may increase the probability that a new peak will be well resolved from the original peaks.

The method was also tested with mixtures of bovine and equine cytochrome c. The two protein forms are phylogenetically related [29], but differ in three amino acid positions out of 104 [30]. The pure forms of the proteins are easily distinguishable by visual comparison of tryptic maps (see the two upper chromatograms in Fig. 6). The ability of the method to detect amino acid sequence modifications was assessed by adding a small amount of bovine cytochrome c (1–25%) as an impurity to the equine protein. The bottom chromatogram in Fig. 6 shows a peptide map of a sample where about 7% bovine protein had been added. This sample was classified as unacceptable by the proposed method, as were samples with larger amounts of the impurity. Smaller amounts could not be distinguished from the normal variability of the peptide mapping method.

4.3. Practical considerations

A serious problem with all chromatographic fingerprinting methods is the influence of the gradual degradation of column performance. This has been observed to affect the outcome of multivariate characterization methods in
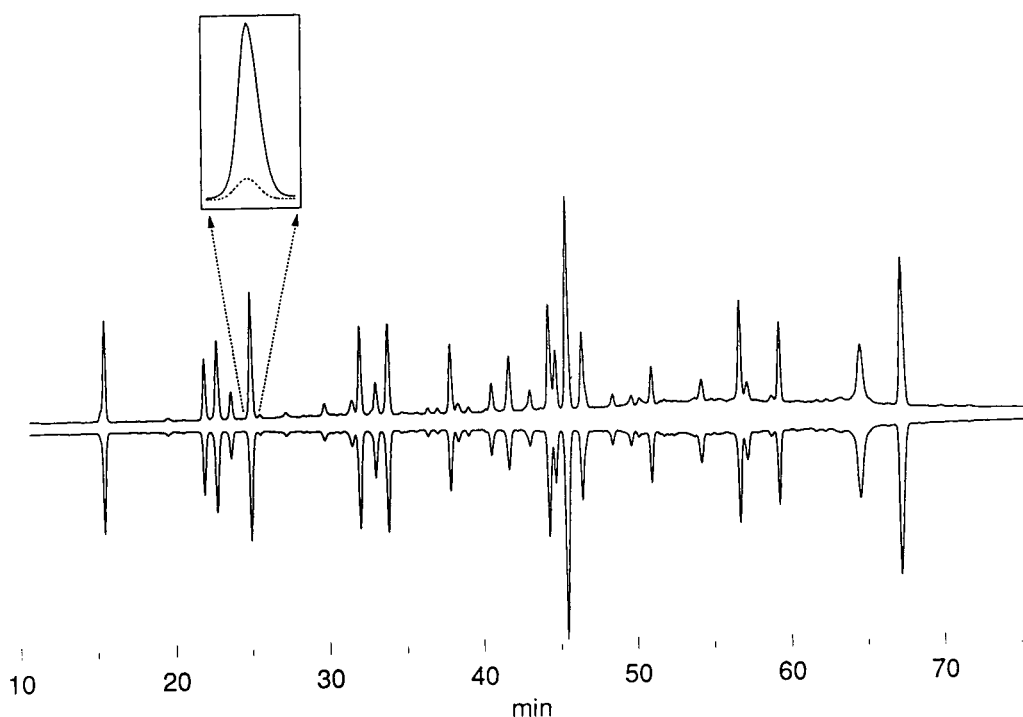
Fig. 5. Top: simulated spiking with addition of a Gaussian peak corresponding to 7% of the largest peak in the chromatogram. Bottom: reference chromatogram. Both chromatograms are shown after the pretreatment to facilitate the comparison. The inset shows the position and size of the added peak. Chromatographic conditions as in Experimental.

pyrolysis–GC [31] and classifications in GC [32]. The effect of stationary phase degradation on predictions in experimental design has recently been demonstrated [33]. The normal chromatographic pattern in peptide mapping is also influenced by column degradation [11,16]. Proper experimental precautions should be taken to maximize the stability of the column, e.g. by using high-purity mobile phase additives and regular cleaning of the column to remove any adsorbed contaminants. The acidic mobile phases commonly used in peptide mapping may cause column degradation introduced by cleavage of the bonded groups or end caps, exposing the silanol groups [8]. The most common mobile phase system in peptide mapping is based on trifluoroacetic acid (TFA) added both to the aqueous and organic components of the mobile phase. The volatility of the TFA system is a distinct advantage, facilitating mobile phase re-

moval [4]. The aqueous part of the mobile phase in this study consisted of a phosphate buffer (pH 2.5), known to give different selectivity for tryptic fragments compared with TFA-based mobile phases [8,12]. The phosphate buffer is less acidic than TFA-based mobile phases (typical pH 2.0), which might reduce the column degradation, and thus possibly give better long-term reproducibility. It may also be worthwhile to explore the possibilities of polymer-based or zirconium-based reversed-phase columns, which have been claimed to have better pH stability than silica columns [4,34].

The traditional approach in peptide mapping is to produce a reference map together with each new test sample map, and make a visual comparison of the two chromatograms. The multivariate evaluation method is not intended to replace the visual inspection of the chromatograms, but to produce a less subjective criterion
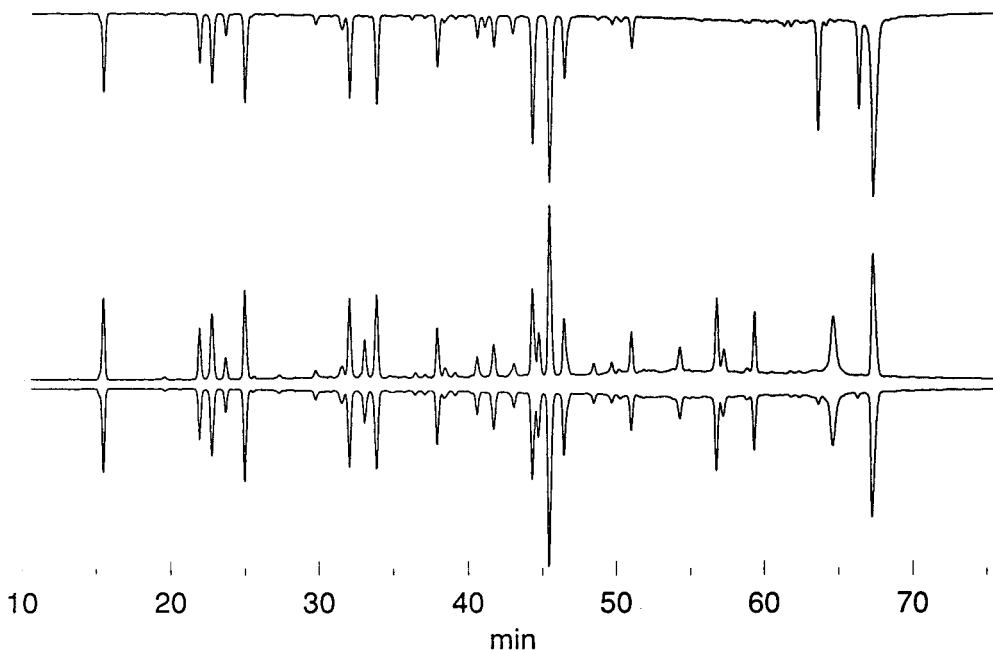
Fig. 6. Chromatograms of (top) a tryptic digest of bovine cytochrome c, (middle) equine cytochrome c and (bottom) a mixture in which 7% of bovine cytochrome c was added to the equine protein. All chromatograms are shown after the pretreatment to facilitate the comparison. Chromatographic conditions as in Experimental.

for classification. Each test sample should be digested and the fragments separated with the same mobile phase preparation as a new reference digest.

A system suitability test for the training set can be provided by a graphical evaluation procedure, where the F-values of the training set are plotted together with the test sample F-values. This is illustrated in Fig. 7, where some selected test samples are shown. The two critical F-values corresponding to outliers and rejected samples are indicated, thus allowing a graphical classification of the test samples. The graphical representation is used as a control chart where the F-value of the test sample is inserted together with the F-value of the corresponding reference chromatogram. The training set is valid as long as the new reference chromatograms are classified as acceptable. This is indicated by F-values for the new reference chromatograms below the critical F-value at the 0.10 significance level.

Visual inspection of the chromatograms in the training set reveals that the column had deteriorated throughout the study. This was also con-
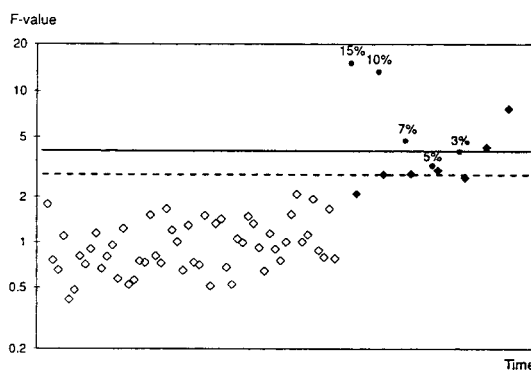
Fig. 7. System suitability test for the training set as a function of time. The solid line indicates the limit for rejection and the dashed line shows the limit for outliers. Note the logarithmic scale on the ordinate. ◇ = Training set objects; ◆ = reference chromatograms in the test set; ○ = chromatograms with bovine cytochrome c added.

firmed by the characterization of the raw data made previously [20]. At the end of the study, the system suitability test indicated that the training set no longer is valid for classifications of test samples (see Fig. 7). One possible solution is the approach presented by Headley and Hardy [32]. They used a dynamic training set, where a set of new reference chromatograms was included in the training set after validation. The oldest chromatograms in the training set were simultaneously removed in order to keep the size of the set constant. A new class model has to be calculated by PCA each time the content of the training set is modified. The training set could be used for a prolonged period of time in their case, despite the observed changes in the experimental conditions, e.g., column degradation.

The visual detection of impurities or modified fragments is, of course, easier if the retention times for these fragments are known. The proposed classification method, on the other hand, treats the entire chromatographic profile and will detect additional peaks without prior knowledge of their position. An indication of the position of the deviant peaks can be obtained by inspection of the residual vector for the suspect chromatogram. Large residuals are expected for regions that deviate from the normal peak pattern in the training set.

General detection limits in peptide mapping are difficult to express, as the peak size necessary for detection is dependent on the elution position of the modified fragment. The general experience of Chloupek et al. [9] is that new, well resolved peaks must be greater than 5% (mol/mol) and co-eluting peaks greater than 15% for visual detection. Instrumental detection of 8% of a spiked contaminant eluted with baseline resolution has been reported [35]. Dougherty et al. [16] used an extensive characterization of the variability in the amount of individual fragments to achieve detection of spiked impurities in rDNA-derived somatotropin in 2–4.5% levels for some specific fragments [16]. The purpose of the method presented in this paper is not primarily to decrease the detection limit, but to facilitate the interpretation of the peptide maps. The automatic processing of numerous chromatograms in a data set, and the unbiased evaluation, are the main benefits.

5. Conclusions

The proposed multivariate evaluation method for fingerprinting techniques is one approach towards the full exploitation of the information contained in complex chromatograms. Numerous chromatograms in a data set can be automatically processed. The method is based on the entire chromatographic profile, thus eliminating the determination of retention time and peak area for the large number of peaks commonly encountered in peptide mapping. The differences between the samples are highlighted by reduction of the chromatographic variations in the data. The use of a training set to describe the normal variations in the cleavage pattern will improve the possibilities for detection of amino acid sequence modifications and contaminants in the protein sample. Multivariate classifications of protein samples according to their peptide maps are less subjective than the traditional visual inspection of two chromatograms.

This approach may also be useful in areas other than peptide mapping, as long as the profiles are relatively similar. In other cases where the profiles are very different, classification will be fairly straightforward from a direct visual comparison [7]. DNA fingerprinting [36] and pyrolysis–GC [37] are examples of other fingerprinting techniques that could be facilitated by the proposed method.

thanked for providing insight into the peptide mapping technique.

References

[1] W.S. Hancock, LC · GC Int., 5, No. 4 (1992) 30.
[2] R.L. Garnick, N.J. Solli and P.A. Papa, Anal. Chem., 60 (1988) 2546.
[3] F.E. Regnier, LC · GC, 5 (1987) 392.
[4] M.W. Dong and A.D. Tran, J. Chromatogr., 499 (1990) 125.
[5] L.R. Snyder, in C. Horváth (Editor), High Performance Liquid Chromatography: Advances and Perspectives, Vol. 1, Academic Press, New York, 1980, p. 208.
[6] G.A. Ross, P. Lorkin and D. Perret, J. Chromatogr., 636 (1993) 69.
[7] H. Rehbein, Electrophoresis, 13 (1992) 805.
[8] M.W. Dong, Adv. Chromatogr., 32 (1992) 21.
[9] R.C. Chloupek, J.E. Battersby and W.S. Hancock, in C.T. Mant and R.S. Hodges (Editors), HPLC of Peptides and Proteins: Separation, Analysis, and Conformation, CRC Press, Boca Raton, FL, 1991, p. 825.
[10] S. Borman, Anal. Chem., 59 (1987) 969A.
[11] K.L. Stone and K.R. Williams, in D.H. Schlesinger (Editor), Macromolecular Sequencing and Synthesis, Alan R. Liss, New York, 1988, Ch. 2, p. 7.
[12] R.C. Chloupek, R.J. Harris, C.K. Leonard, R.G. Keck, R.G. Keyt, M.W. Spellman, A.J.S. Jones and W.S. Hancock, J. Chromatogr., 463 (1989) 375.
[13] N. Lundell, J. Chromatogr., 639 (1993) 97.
[14] M. Herold, D.N. Heiger and R. Grimm, Am Lab., August (1993) 20J.
[15] E.R. Hoff, LC · GC, 2, No. 6, (1989) 28.
[16] J.J. Dougherty, Jr., L.M. Snyder, R.L. Sinclair and R.H. Robins, Anal. Biochem., 190 (1990) 7.
[17] R.C. Chloupek, W.S. Hancock and L.R. Snyder, J. Chromatogr., 594 (1992) 65.
[18] R.G. Brereton (Editor), Multivariate Pattern Recognition in Chemometrics, Elsevier, Amsterdam, 1992.
[19] S. Wold, C. Albano, W.J. Dunn, III, U. Edlund, K. Esbensen, P. Geladi, S. Hellberg, E. Johansson, W. Londberg and M. Sjöström, in B.R. Kowalski (Editor), Chemometrics: Mathematics and Statistics in Chemistry, Reidel, Dordrecht, 1984, p. 17.
[20] G. Malmquist and R. Danielsson, J. Chromatogr., 687 (1994) 71.
[21] S. Wold, K. Esbensen and P. Geladi, Chemometr. Intell. Lab. Syst., 2 (1987) 37.
[22] S. Wold, C. Albano, W.J. Dunn, III, K. Esbensen, S. Hellberg, E. Johansson and M. Sjöström, in H. Martens and H. Russwurm, Jr. (Editors), Food Research and Data Analysis, Applied Science, London, 1983, p. 147.
[23] M.A. Sharaf, D.L. Illman and B.R. Kowalski, Chemometrics, Wiley, New York, 1986, Ch. 6, p. 179.
[24] P.J. Gemperline, L.D. Webber and F.O. Cox, Anal. Chem., 61 (1989) 138.
[25] S. Renlund, I.-M. Klintrot, M. Nunn, J.L. Schrimsher, C. Wernstedt and U. Hellman, J. Chromatogr., 512 (1990) 325.
[26] S. Renlund, personal communication, 1992.
[27] S. Wold, Technometrics, 20 (1978) 397.
[28] J.C. Miller and J.N. Miller, Statistics for Analytical Chemistry, Ellis Horwood, Chichester, 3rd ed., 1993, Ch. 3, p. 75.
[29] M.A. Sharaf, B.R. Kowalski and B. Weinstein, Z. Naturforsch. Teil C, 35 (1980) 508.
[30] Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Georgetown University Medical Center, Georgetown, Washington, DC, 1972.
[31] B.K. Lavine, Chemometr. Intell. Lab. Syst., 15 (1992) 219.
[32] L.M. Headley and J.K. Hardy, J. Food Sci., 57 (1992) 980.
[33] B. Bourguignon and D.L. Massart, Anal. Chim. Acta, 282 (1993) 33.
[34] H.-J. Wirth, K.-O. Eriksson, P. Holt, M.I. Aguilar and M.T.W. Hearn, J. Chromatogr., 646 (1993) 129.
[35] D.G. Bursryn, T. Copmann, M. Dinowitz, R. Garnick, A. Losikoff, A. Lubiniecki, M.S. Rubino and M. Wiebe, Biopharm, 4 (1991) 22.
[36] J.S.C. Smith and O.S. Smith, Adv. Agron., 47 (1992) 85.
[37] J.A. Pino, J.E. McMurry, P.C. Jurs, B.K. Lavine and A.M. Harper, Anal. Chem., 57 (1985) 295.

Correlation of structure and retention behaviour in reversed-phase high-performance liquid chromatography
I. Leucine-enkephalin-related glycoconjugates

Lidija Varga-Defterdarović,* Štefica Horvat, Mare Skurić, Jaroslav Horvat

Ruđer Bošković Institute, Department of Organic Chemistry and Biochemistry, P.O. Box 1016, Bijenička c. 54, 41001 Zagreb, Croatia

Abstract

The chromatographic behaviour of leucine-enkephalin-related glycoconjugates with an ester-, ether- and amide-type of linkage was investigated by reversed-phase high-performance liquid chromatography using trifluoroacetic acid as ion-pairing agent and methanol as modifier of the aqueous phase. The results show that the position and the type of sugar–peptide linkage, the type of sugar moiety introduced and the degree of carbohydrate protection contribute to the overall retention of the glycopeptides studied.

1. Introduction

In recent years, reversed-phase high-performance liquid chromatography (RP-HPLC) has been widely used for the separation of naturally occurring glycoproteins [1,2] and in the final purification of synthetic glycopeptides [3–7]. The carbohydrate chain of the glycoproteins often carries a highly specific biological recognition structure and extensive studies have been performed on the chromatographic analysis of these oligosaccharide systems [8–10]. However, only a few reports have appeared [11–13] dealing with the relationships between the structure and RP-HPLC retention characteristics of natural or synthetic glycopeptides.

Since the discovery of the enkephalins [14]

(Tyr–Gly–Gly–Phe–Leu/Met), many endogenous opioid peptides have been detected in mammals. This important group of peptides produce a wide range of central and peripheral effects, which, in addition to spinal and supraspinal analgesia, include tolerance and physical dependence, respiratory depression, euphoria and other behavioural effects, effects on gastrointestinal motility and cardiovascular and immune functions [15].

We have shown that the introduction of the sugar moiety at the fifth position of the enkephalins significantly influenced the biological activity profile of the parent peptide. Interestingly, both the type of linkage and the kind of sugar moiety introduced were of major consequence with regard to opioid receptor selectivity [16–20]. In addition, leucine-enkephalin (Tyr–Gly–Gly–Phe–Leu) glycoconjugates exhibited antiviral activity against HIV-1 which was significantly

---

* Corresponding author.

higher than the activity of the parent peptide itself [21].

In continuation of our investigations on the influence of incorporated sugar moieties on different aspects of the behaviour and activity of leucine-enkephalin, in this paper we give the correlation between the incorporated sugar moieties and the chromatographic behaviour, using RP-HPLC of leucine-enkephalin-related glycoconjugates 1–12 (Fig. 1). The study was carried out with both gradient and isocratic elution with methanol, using trifluoroacetic acid as ion-pair-
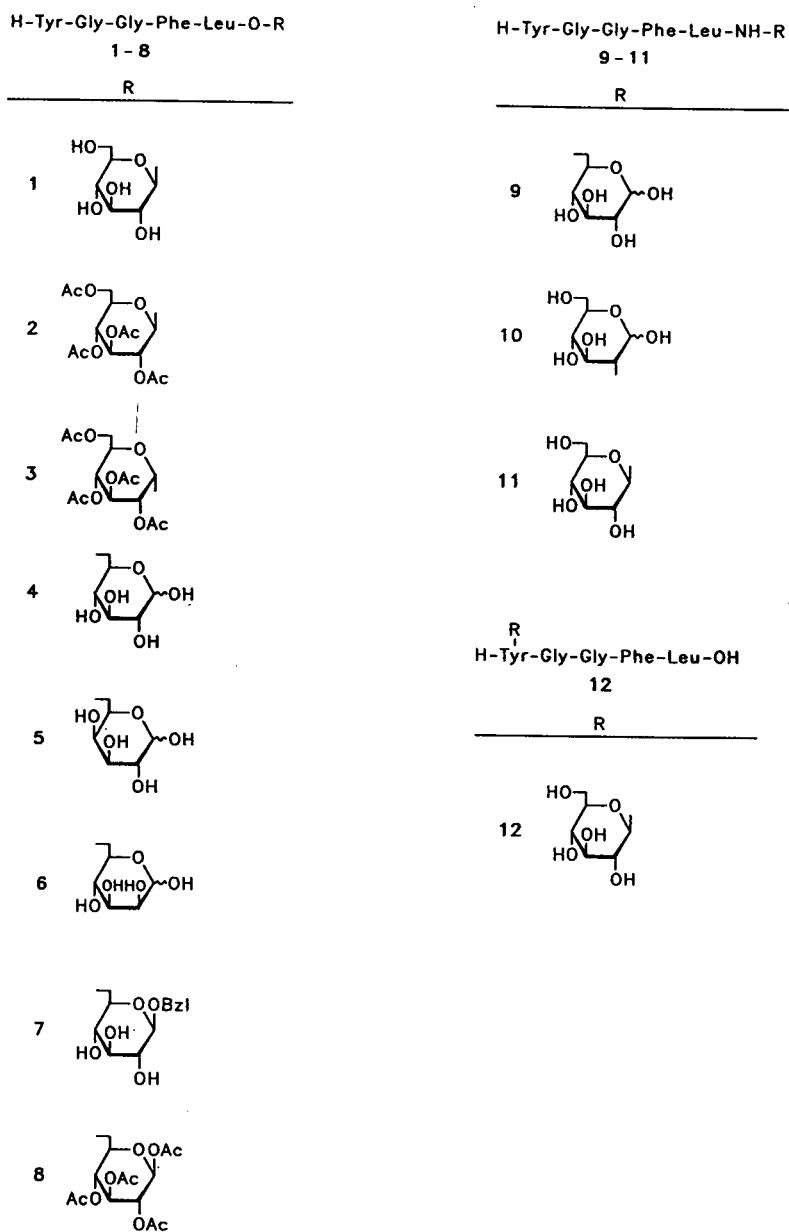
Fig. 1. Structures of leucine-enkephalin-related glycoconjugates with ester (**1–8**), amide (**9–11**) and ether (**12**) types of linkage.

ing agent. The effects of carbohydrates on the retention of the parent peptide were determined as a function of the type of sugar moiety, the degree of sugar moiety protection and the type and position of the sugar–peptide linkage.

2. Experimental

2.1. Column

An n-octadecyl Si 100 analytical column (250 × 4.6 mm I.D., 5 μm) (Serva) was used. The dead volume was determined to be 3.58 ml by injection of 5 μg of uracil. The column was operated at a flow-rate of 0.5 ml/min. All measurements were made at 25°C.

2.2. Instrumentation

The HPLC equipment consisted of Varian Model 9010 liquid chromatograph equipped with a Rheodyne Model 7125 injector, a Varian Model 4400 integrator and a Varian Model 9050 variable-wavelength UV–Vis detector. UV detection was performed at 280 nm.

2.3. Mobile phase

Solvent A was 0.1% trifluoroacetic acid (TFA) in methanol–water (40:60) (pH 2.50) and solvent B was 0.1% TFA in methanol–water (75:25) (pH 2.82). Samples were eluted in the linear gradient mode with methanol (100% A → 50% A + 50% B; 40.0% → 57.5% methanol and 100% A → 25% A + 75% B; 40.0% → 66.25% methanol) in 0.1% TFA over a 30-min time period, and also in the isocratic mode with different concentrations of methanol.

For analytical HPLC the samples were dissolved in solvent A at 0.5 mg/ml and 100 μl of the solution were injected.

2.4. Chemicals

Methanol was of HPLC grade (Aldrich, Milwaukee, WI, USA) and trifluoroacetic acid was of spectroscopic grade (Uvasol; Merck, Darm-

stadt, Germany). Leucine-enkephalin ([Leu[5]]E) and leucine-enkephalinamide ([Leu[5]]E-NH$_2$) were purchased from Sigma (St. Louis, MO, USA). Leucine-enkephalin methyl ester ([Leu[5]]E-OMe) was prepared as described previously [22]. Compounds 1–12 were synthesized as described [16–20]. HPLC analysis of some glycoconjugates indicated the presence of two peaks (found to be diastereomers) which were separated by repetitive injections (100 μl, concentration of 15 mg/ml in solvent A) under the conditions given in Section 3. The structure and homogeneity of glycopeptides were confirmed by microanalysis (C, H, N), NMR spectroscopy using a Varian Gemini 300 instrument and RP-HPLC.

3. Results and discussion

The results obtained with linear gradient elution for leucine-enkephalin-related glycoconjugates with ester (1–8), amide (9–11) and ether (12) types of linkage, and also for some of their D-Leu[5] isomers [two isomeric products were obtained during the synthesis of 4–7, while in others only small amounts (1–3 and 8) or no (9–12) of racemized products were detected] are shown in Table 1.

In general, incorporation of an unprotected carbohydrate, owing to the increased hydrophilicity of the molecule, decreased the retention time of the parent unmodified leucine-enkephalin. Differences in the retention times indicate that the type of unprotected sugar moiety and the type and position of the linkage affect the HPLC retention behaviour of the glycoconjugates studied. Compounds 1 and 4 in which the peptide is linked through the ester bond to either C-1 or C-6 of the identical carbohydrate moiety (D-glucose) indicate a smaller hydrophilicity of the 1-O-glycoconjugate 1, reflected in a stronger retention, than that of the 6-O-derivative 4. The investigation of the chromatographic behaviour of 4–6 having identical type and position of the sugar–peptide linkage but different monosaccharide moieties (D-

Table 1
Retention data for leucine-enkephalin (Tyr–Gly–Gly–Phe–Leu) and related glycoconjugates

Compound	Sugar	Type of linkage	Position of linkage	Retention time (min)	
				L-Isomer	D-Isomer
1	Glc	Ester	1	11.60	
2	β-GlcAc$_4$	Ester	1	24.49	
3	α-GlcAc$_4$	Ester	1	24.49	
4	Glc	Ester	6	10.60	16.61[a]
5	Gal	Ester	6	11.53	18.19[a]
6	Man	Ester	6	10.60	16.47[a]
7	GlcBzl	Ester	6	18.11	22.05[a]
8	GlcAc$_4$	Ester	6	22.77	
9	6-NH$_2$-Glc	Amide	6	10.81	
10	2-NH$_2$-Glc	Amide	2	11.03	
11	1-NH$_2$-Glc	Amide	1	11.67	
12	Glc	Ether	1	11.10	
[Leu^5]E				16.32	23.84[b]
[Leu^5]E-OMe				19.12	22.70[c]
[Leu^5]E-NH$_2$				12.82	

RP-HPLC conditions: linear gradient of methanol in 0.1% TFA (compounds **2** and **3**, 40.0 to 66.25% methanol over a 30-min time period; all aother compounds, 40.0 to 57.5% methanol over a 30-min time period); flow-rate, 0.5 ml/min; load, 50 μg per 100 μl of solvent A; temperature, 25°C; UV detection at 280 nm.

[a] Leucine-enkephalin-related glycoconjugates containing D-Leu at the fifth position of the peptide backbone.
[b] [D-Leu^5]E.
[c] [D-Phe^4,Leu^5]E-OMe.

glucose, D-galactose, D-mannose) revealed that with gradient elution the retention remained unaffected with configurational change at C-2 (**4** and **6**). In contrast, introduction of D-galactose into the leucine-enkephalin molecule (**5**) decreased the hydrophilicity of the overall molecule, resulting in increased retention on the column. Concerning the effect of the type of C-1 sugar–peptide linkage, we found no significant difference in the retention behaviours of ester **1** and amide **11**. The ether derivative **12** showed increased hydrophilicity, probably owing to the presence of a free carboxyl group in the molecule. Comparison of the chromatographic behaviours of the C-6 glycoconjugates **4** and **9** revealed that an amide bond contributes to a slightly stronger retention on a reversed-phase column. Different chromatographic behaviour of glycoconjugates with the amide type of linkage at C-6, C-2 and C-1 positions of the sugar molecule (**9**, **10** and **11**, respectively) was also observed. Thus, the 1-NH-glycopeptide **11** binds

to the column more strongly than **9** and **10**, and this is reflected in the decreasing order **11** > **10** > **9**. Amidation of the C-terminal carboxyl group of leucine-enkephalin also decreases the retention of the parent peptide compound but to a lesser extent than in **9–11**.

Inversion of the amino acid configuration at the Phe^4 or Leu^5 position in leucine-enkephalin in addition to the glycoconjugates studied leads to a considerably stronger retention on a reversed-phase column and the retention times of negative (L-D) isomers have consistently been found to be greater than those for positive (L-L) isomers (Table 1).

As expected, the incorporation of partially or fully protected monosaccharides into leucine-enkephalin resulted in increased hydrophobicity of the overall molecule and thus in stronger retention on the column (Table 1). Among the fully acetylated glycopeptides studied (**2**, **3** and **8**) we observed the same elution order as with the corresponding unprotected glycopeptides **1**
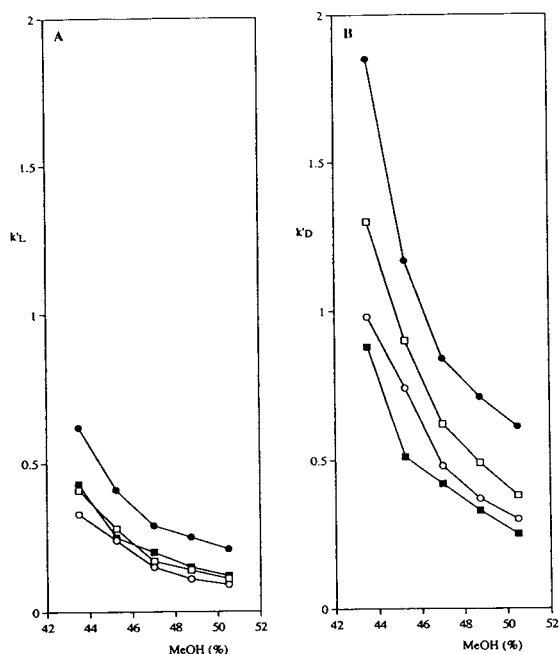
Fig. 2. Effect of methanol concentration on (A) capacity factors (k'_L) of glycoconjugates 4–6 and (B) capacity factors (k'_D) of D-Leu^5-isomers of compounds 4–6. RP-HPLC conditions: isocratic elution at different concentrations of methanol in 0.1% trifluoroacetic acid; flow-rate, 0.5 ml/min; load, 50 μg per 100 μl of solvent A; temperature, 25°C; UV detection at 280 nm. ○ = 4; □ = 5; ◼ = 6; ● = [Leu^5]E.

Fig. 3. Effect of methanol concentration on (A) separation factor (α) and (B) resolution (R_s) of some epimeric leucine-enkephalin-glycoconjugates. RP-HPLC conditions as in Fig. 2. ○ = 4; □ = 5; ◼ = 6; ◆ = 7; ● = [Leu^5]E.

and 4. Accordingly, 1-O-conjugates 2 and 3 were retained considerably more strongly on the column than the 6-O-derivative 8. Concerning the influence of the anomeric form on retention, under the gradient conditions employed there was no difference in the chromatographic behaviour of 2 (β-anomer) and 3 (α-anomer). The same observation was made with isocratic elution (50.5% methanol in 0.1% trifluoroacetic acid). However, under the same chromatographic conditions, the epimer of 2, containing at the C-terminal peptide backbone D-Leu residue, was retained more strongly than the corresponding epimer of 3.

We investigated the chromatographic behaviour of 1–12 at different methanol concentrations. A remarkable change in the retention was observed when the methanol concentration was changed from 40% to 50%. With increasing

percentage of methanol in the mobile phase, the retention of all the compounds studied decreased, following the same elution pattern as under gradient conditions, except for 4–6. As can be seen from the k'_L versus methanol concentration plot in Fig. 2A, the retention decreased in the order (except at 45.25% methanol) *manno* (6) >*galacto* (5) > *gluco* (4). Interestingly, the elution orders were the same under gradient and isocratic conditions for the corresponding isomers of conjugates 4–6 having D-leucine in the peptide part of the molecule (Fig. 2B).

The effect of different methanol concentrations in the mobile phase on the separation factor α of epimeric glycoconjugates and also the parent epimeric peptides is presented in Fig. 3A. As shown, for all the compounds examined the composition of the mobile phase has no or little influence on the α values.

Resolutions (R_s) of the studied diastereomers of leucine-enkephalin-related glycoconjugates were inferior to those of epimeric leucine-enkephalin (Fig. 3B).

In conclusion, the investigation of the chromatographic behaviour of leucine-enkephalin-related glycopeptides 1–12 demonstrated that the

position and type of sugar–peptide linkage, the structure of the sugar moiety and the degree of the sugar moiety protection, influence their retention on a reversed-phase column. Further, the feasibility of RP-HPLC for separating epimeric glycopeptides has been demonstrated.

Abbreviations used for monosaccharides

Gal	D-galactopyranose
Glc	D-glucopyranose
GlcAc$_4$	1,2,3,4-tetra-O-acetyl-β-D-glucopyranose
α-GlcAc$_4$	2,3,4,6-tetra-O-acetyl-α-D-glucopyranose
β-GlcAc$_4$	2,3,4,6-tetra-O-acetyl-β-D-glucopyranose
GlcBzl	benzyl β-D-glucopyranoside
Man	D-mannopyranose
1-NH$_2$-Glc	β-D-glucopyranosylamine
2-NH$_2$-Glc	2-amino-2-deoxy-D-glucopyranose
6-NH$_2$-Glc	6-amino-6-deoxy-D-glucopyranose

Acknowledgements

References

[1] H.P.J. Bennett, N.G. Seidah, S. Benjannet, S. Solomon and M. Chretien, Int. J. Pept. Protein Res., 27 (1986) 306.

[2] K.-H. Strube, F. Lottspeich and R. Geyer, Eur. J. Biochem., 184 (1989) 119.

[3] J.L. Torres, P. Clapes, I. Haro, G. Valencia, F. Reig and J.M. Garcia-Anton, Chromatographia, 25 (1988) 891.

[4] J.L. Torres, I. Haro, G. Valencia, F. Reig and J.M. Garica-Anton, Experientia, 45 (1989) 574.

[5] F. Filira, L. Biondi, F. Cavaggion, B. Scolaro and R. Rocchi, Int. J. Pept. Protein Res., 36 (1990) 86.

[6] J.L. Torres, E. Bardaji and G. Valencia, Methods Neurosci., 6 (1991) 35.

[7] J.L. Torres, F. Reig, G. Valencia, R.E. Rodriguez and J.M. Garcia-Anton, Int. J. Pept. Protein Res., 31 (1988) 474.

[8] K. Kakehi, S. Suzuki, S. Honda and Y.C. Lee, Anal. Biochem., 199 (1991) 256.

[9] R. Mögele, B. Pabel and R. Galensa, J. Chromatogr., 591 (1992) 165.

[10] T. Akiyama, J. Chromatogr., 588 (1991) 53.

[11] H. Morehead, P. McKay and R. Wetzel, Anal. Biochem., 126 (1982) 29.

[12] M. Hollosi, E. Kollat, I. Laczko, K.F. Medzihradszky, J. Thurin and L. Otvos, Jr., Tetrahedron Lett., 32 (1991) 1531.

[13] L. Otvos, Jr., L. Urge and J. Thurin, J. Chromatogr., 599 (1992) 43.

[14] J. Hughes, T.W. Smith, H.W. Kosterlitz, L.A. Fothergill, B.A. Morgan and R.H. Morris, Nature, 258 (1975) 577.

[15] G.A. Olson, R.D. Olson and A.J. Kastin, Peptides, 12 (1991) 1407.

[16] L. Varga, Š. Horvat, C. Lemieux and P.W. Schiller, Int. J. Pept. Protein Res., 30 (1987) 371.

[17] J. Horvat, Š. Horvat, C. Lemieux and P.W. Schiller, Int. J. Pept. Protein Res., 31 (1988) 499.

[18] L. Varga-Defterdarović, Š. Horvat, N.N. Chung and P.W. Schiller, Int. J. Pept. Protein Res., 39 (1992) 12.

[19] Š. Horvat, J. Horvat, L. Varga-Defterdarović, K. Pavelić, N.N. Chung and P.W. Schiller, Int. J. Pept. Protein Res., 41 (1993) 399.

[20] M. Skurić, J. Horvat, Š. Horvat, N.N. Chung and P.W. Schiller, Int. J. Pept. Protein Res., 43 (1994) 402.

[21] Š. Horvat, L. Varga, J. Horvat, A. Pfützner, H. Suhartono and H. Rübsamen-Waigmann, Helv. Chim. Acta, 74 (1991) 951.

[22] N.S. Agarwal, V.J. Hruby, R. Katz, W. Klee and M. Nirenberg, Biochem. Biophys. Res. Commun., 76 (1977) 129.

Correlation of structure and retention behaviour in reversed-phase high-performance liquid chromatography
II. Methionine-enkephalin-related glycoconjugates

Lidija Varga-Defterdarović*, Štefica Horvat, Mare Skurić, Jaroslav Horvat

Ruđer Bošković Institute, Department of Organic Chemistry and Biochemistry, P.O. Box 1016, Bijenička c. 54, 41001 Zagreb, Croatia

Abstract

Reversed-phase high-performance liquid chromatographic elution data for methionine-enkephalin-related glycoconjugates were analysed as a function of the identity and position of the sugar–peptide linkage. It was shown that binding to the column could be correlated with the degree of sugar moiety protection. Replacement of either the phenylalanine or methionine residue in the peptide backbone of the glycoconjugates with its D-enantiomer leads to a considerably stronger retention on a reversed-phase column. The dependence of retention times on the methanol concentration in the mobile phase suggested that, under the conditions studied, there are different retention mechanisms for glycopeptides containing unprotected sugar moieties in the molecule.

1. Introduction

The results of a previous study [1] showing that glycation significantly influenced the retention characteristics of the endogenous opioid pentapeptide leucine-enkephalin [2] (Tyr–Gly–Gly–Phe–Leu) led us to investigate the chromatographic behaviour of the structurally similar opioid pentapeptide methionine-enkephalin [2] (Tyr–Gly–Gly–Phe–Met) and related glycopeptides under reversed-phase HPLC conditions. This seemed important as we observed that incorporation of the identical carbohydrate moieties into either the leucine- or methionine-enkephalin molecule produced different effects on the biological activity of the parent peptide

compounds [3,4]. Therefore, the objectives of this work were to determine the retention parameters of some C-terminally glycated methionine-enkephalin derivatives under various conditions and to find correlations between the retention properties and physico-chemical parameters of the investigated glycopeptides.

2. Experimental

2.1. Instrumentation

The chromatographic system consisted of a Varian Model 9010 liquid chromatograph equipped with a Rheodyne Model 7125 injector, Varian Model 4400 integrator and Varian Model 9050 variable-wavelength UV–Vis detector. The

---

* Corresponding author.

detection wavelength was set at 280 nm and the flow-rate was 0.5 ml/min. A Serva *n*-octadecyl Si 100 column (250 × 4.6 mm I.D., 5 μm) was employed for both analytical and semi-preparative separations. The column dead volume (3.58 ml) was determined from the retention of uracil. The RP-HPLC conditions are given in the tables and figures. For analytical HPLC the samples were dissolved in 0.1% trifluoroacetic acid (TFA) in methanol–water (40:60) at a 0.5 mg/ml concentration and 100 μl of the solution were injected.

2.2. Chemicals

Methanol was of HPLC grade (Aldrich, Milwaukee, WI, USA) and trifluoroacetic acid was of spectroscopic grade (Uvasol; Merck, Darmstadt, Germany). Methionine-enkephalin ([Met[5]]E), leucine-enkephalin ([Leu[5]]E) and [D-Ala[2], Met[5]]enkephalin were purchased from Sigma (St. Louis, MO, USA).

The methionine-enkephalin-related glycoconjugates 1–8 (Fig. 1) and leucine-enkephalin glycoconjugates 9–11 were synthesized as described previously [3–5]. HPLC analysis of each glycoconjugate indicated the presence of two peaks (found to be diastereomers) which were

separated by repetitive injections (100 μl, concentration of 15 mg/ml in 40% methanol–0.1% TFA) under the conditions given in Section 3. The structures and homogeneity of the glycopeptides were confirmed by microanalysis, NMR spectroscopy using a Varian Gemini 300 instrument and RP-HPLC.

3. Results and discussion

A series of glycoconjugates (1–8, Fig. 1) in which opioid peptide, methionine-enkephalin or related analogue, [D-Ala[2], Met[5]]-enkephalin (Tyr–D-Ala–Gly–Phe–Met), have been linked through an ester bond either to the 6-OH or 1-OH of various D-glycopyranose moieties were synthesized and analysed by RP-HPLC using trifluoroacetic acid as ion-pairing agent and methanol as modifier of the aqueous phase.

Depending on the synthetic conditions, two isomeric products were obtained in each instance due to the racemization at either the C-terminal amino acid residue (Met) or the penultimate residue (Phe). The diastereomers were separated by RP-HPLC (the only exceptions being glycoconjugates 6 and 8, for which only a small amount of racemized product was produced

H–Tyr–X–Gly–Phe–Met–O–R

1–8
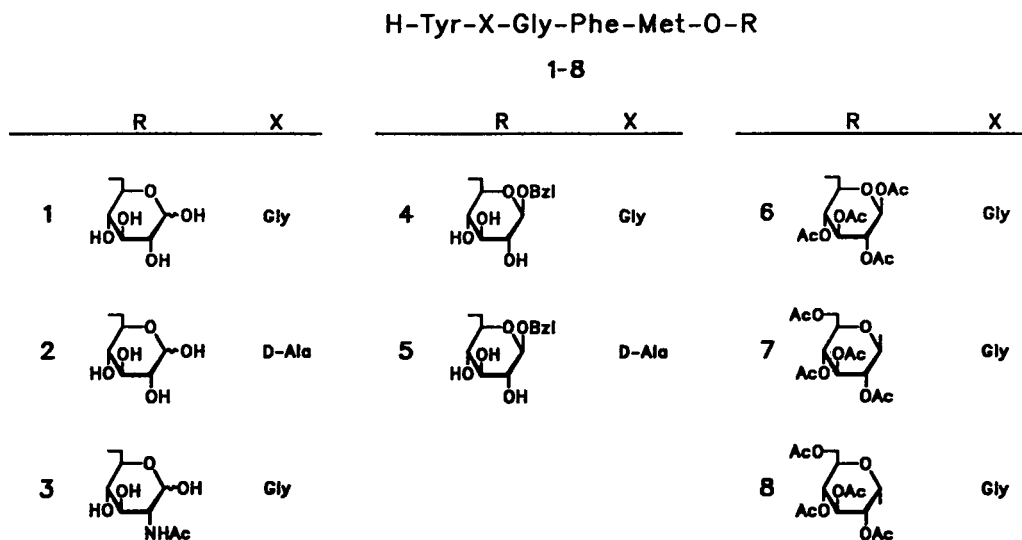
Fig. 1. Structures of methionine-enkephalin-related glycoconjugates **1–8**.

Table 1
Retention data for methionine-enkephalin (Tyr–Gly–Gly–Phe–Met) and related glycoconjugates

Compound	Sugar	Type of linkage	Position of linkage	Retention time (min)	
				L-Isomer	D-Isomer
1	Glc	Ester	6	9.45	13.39[a]
2	Glc	Ester	6	9.95	13.39[a]
3	GlcNAc	Ester	6	9.74	13.60[a]
4	GlcBzl	Ester	6	20.33	25.06[a]
5	GlcBzl	Ester	6	20.98	25.78[a]
6	GlcAc$_4$	Ester	6	20.55	
7	β-GlcAc$_4$	Ester	1	26.71	29.79[b]
8	α-GlcAc$_4$	Ester	1	27.57	
[Met^5]E				11.53	16.61[c]
[D-Ala^2, Met^5]E				11.74	

RP-HPLC conditions: linear gradient of methanol (40 to 57.5%) in 0.1% aqueous TFA over a 30-min time period; flow-rate, 0.5 ml/min; load, 50 μg per 100 μl of 0.1% TFA in methanol–water (40:60); temperature, 25°C; UV detection at 280 nm.
[a] Methionine-enkephalin-related glycoconjugate having D-Met at the fifth position of the peptide part of the molecule.
[b] Methionine-enkephalin-related glycoconjugate having D-Phe at the fourth position of the peptide part of the molecule.
[c] Tyr–Gly–Gly–Phe–D-Met.

during synthesis) and their chromatographic behaviour was examined using gradient (Table 1) and isocratic elution at different concentrations of methanol.

The results in Table 1 demonstrate that incorporation of a single sugar moiety caused either a decrease or an increase in the elution time of non-glycated parent peptides depending on the structure of the carbohydrate moiety introduced. The retention times of glycopeptides 1–3 containing unprotected sugar moieties, compared with methionine-enkephalin or its D-Ala^2-analogue, were shortened owing to the incorporation of the hydrophilic carbohydrate [1,6,7], whereas increased retention times were observed when partially (4 and 5) or fully protected monosaccharides (glycoconjugates 6–8) were introduced into the enkephalin molecule. Fig. 2 illustrates the gradient elution pattern of methionine-enkephalin and of some related glycoconjugates.

Interestingly, although it might be expected that amino acid sequence would have an effect on opioid peptide retention, replacement of the Gly^2 residue in methionine-enkephalin with D-Ala did not lead to a significant increase in the hydrophobic interaction in RP-HPLC in either

the peptide Tyr–D-Ala–Gly–Phe–Met or in the corresponding glycoconjugates 2 and 5 (Table 1). In contrast, inversion of the amino acid configuration at the Met^5 or Phe^4 position caused a dramatic increase in the retention times of all the compounds studied (Table 1). This is in accordance with the elution order of epimers of leucine-enkephalin-related glycoconjugates previously observed in RP-HPLC, with L–D isomers having longer retention times than the corresponding L–L isomers [1]. It could be concluded that replacement of the Met or Phe residue by the corresponding D-enantiomer results in a decreased molecular polarity and much larger hydrophobic contact area. According to the data in Table 1, under the RP-HPLC conditions employed, excellent column selectivity and resolution were obtained for all the diastereomeric pairs studied.

In order to find an explanation for the different biological activities of structurally related leucine- and methionine-enkephalin glycoconjugates [3,4], we compared changes in the retention time of the parent peptide, with gradient elution, caused by glycation with the same carbohydrate moiety. We observed (Table 2) that the introduction of the free D-glucose moiety into
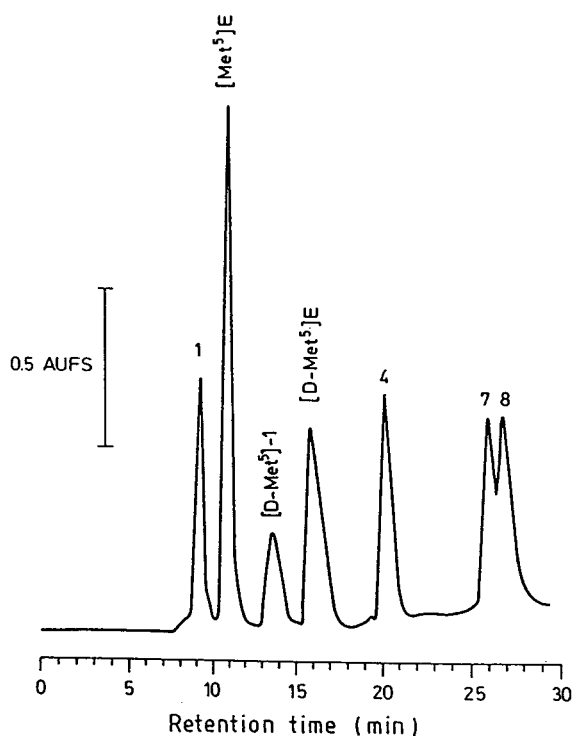
Fig. 2. Gradient elution profile of methionine-enkephalin and related glycoconjugates. Solutes: 6-O-(Tyr–Gly–Gly–Phe–Met)-D-glucopyranose (1); Tyr–Gly–Gly–Phe–Met ([Met⁵]E); 6-O-(Tyr–Gly–Gly–Phe–D-Met)-D-glucopyranose ([D-Met⁵]-1); Tyr–Gly–Gly–Phe–D-Met ([D-Met⁵]E); benzyl 6-O-(Tyr–Gly–Gly–Phe–Met)-β-D-glucopyranoside (4); 2,3,4,6-tetra-O-acetyl-1-O-(Tyr–Gly–Gly–Phe–Met)-β-D-glucopyranose (7); 2,3,4,6-tetra-O-acetyl-1-O-(Tyr–Gly–Gly–Phe–Met)-α-D-glucopyranose (8). RP-HPLC conditions as in Table 1.

methionine-enkephalin (1) decreases the retention time less (compared with the parent opioid pentapeptide) than incorporation of the same sugar moiety into leucine-enkephalin (9). A similar trend was observed for their isomers in which Met or Leu residue in the peptide backbone was replaced by the corresponding D-enantiomer ([D-Met⁵]-1 and [D-Leu⁵]-9, respectively). However, just the opposite trend was observed when methionine- and leucine-enkephalin were modified by incorporation of either partially protected (benzyl β-D-glucopyranoside) (4 and 10, respectively) or fully protected (1,2,3,4-tetra-O-acetyl-β-D-glucopyranose) (6 and 11, respec-

tively) carbohydrates (Table 2). The sugar moieties mentioned increased the retention times of the parent unmodified enkephalins, but now the methionine-enkephalin molecule was influenced much more than leucine-enkephalin. Moreover, D-isomers of glycoconjugates 4 and 10, having Tyr–Gly–Gly–Phe–D-Met and Tyr–Gly–Gly–Phe–D-Leu as the peptide moiety, showed the largest difference in retention times, Δt being +8.4 min for [D-Met⁵]-4 and −1.8 min for [D-Leu⁵]-10. This confirmed the suggestion that the general hydrophobicity of unfolded peptides [8] and the contribution of the free or derivatized carbohydrates [9] on the retention time in RP-HPLC are not the only determining factors of the elution pattern of glycopeptides [10]. We assume that a chiral incorporated sugar moiety caused different conformational changes in the parent opioid peptides, influencing the overall hydrophobicity of the related glycopeptides and consequently produced also a different effect on the biological activity of the parent peptide molecule.

Fig. 3A and B illustrate the dependence of the chromatographic behaviour of 1–8 on the methanol concentration in the mobile phase obtained in a series of isocratic experiments. A remarkable change in the retention of glycopeptides 1–8 was observed when methanol concentration was changed from 40% to 50%, as can be seen from the log k'_L versus methanol concentration plots in Fig. 3A and B. Unlike partially or fully protected glycoconjugates 4–8, for which the plots of log k'_L versus methanol concentration were almost linear (Fig. 3B), the plots for 1–3, having an unprotected monosaccharide moiety in the molecule, were non-linear under the conditions studied (Fig. 3A). If we assume that the requirements for linearity of the log k'_L versus methanol concentration for the hyprophobic (solvophobic) retention mechanism on an ODS ligand were fulfilled [11–13] then the chromatographic behaviour of methionine-enkephalin-related glycoconjugates 1–3, having the 1-OH position of the sugar moiety free is affected by processes other than solvophobic. It is important to mention that in these experiments the pH change was negligible (from pH 2.62 to 2.70),

Table 2
Comparison of the observed differences in the retention times of parent peptides after incorporation of the identical monosaccharide moieties into methionine- and leucine-enkephalin

Compound	Parent peptide	Carbohydrate moiety	Observed difference in retention time, Δt (min)
1	Tyr–Gly–Gly–Phe–Met	Glc	−2.08
9[a]	Tyr–Gly–Gly–Phe–Leu	Glc	−5.72
[D-Met⁵]-1	Tyr–Gly–Gly–Phe–D-Met	Glc	−3.22
[D-Leu⁵]-9	Tyr–Gly–Gly–Phe–D-Leu	Glc	−7.23
4	Tyr–Gly–Gly–Phe–Met	GlcBzl	+8.80
10[a]	Tyr–Gly–Gly–Phe–Leu	GlcBzl	+1.79
[D-Met⁵]-4	Tyr–Gly–Gly–Phe–D-Met	GlcBzl	+8.45
[D-Leu⁵]-10	Tyr–Gly–Gly–Phe–D-Leu	GlcBzl	−1.79
6	Tyr–Gly–Gly–Phe–Met	GlcAc₄	+9.02
11[a]	Tyr–Gly–Gly–Phe–Leu	GlcAc₄	+6.45

Based on the retention times data for methionine-enkephalin and related glycoconjugates presented in Table 1 and on the data for leucine-enkephalin and related glycoconjugates presented in Table 1 in Part I [1]. Difference in retention time calculated by comparison with the retention of the parent peptides.

[a] Compounds 9–11 are leucine-enkephalin-related glycoconjugates; 9 = 6-O-(Tyr–Gly–Gly–Phe–Leu)-D-glucopyranose [3]; 10 = benzyl 6-O-(Tyr–Gly–Gly–Phe–Leu)-β-D-glucopyranoside [3]; 11 = 1,2,3,4-tetra-O-acetyl-6-O-(Tyr–Gly–Gly–Phe–Leu)-β-D-glucopyranose [3]. For structures of compounds 9–11, see Part I [1].

hence the effect of pH on the retention time of 1–8 can be neglected.

4. Conclusions

In this study we undertook a detailed analysis of the retention times of methionine-enkephalin-related glycoconjugates and their optical isomers as a function of the identity and position of the sugar–peptide linkage. We found that the major determinant of the elution time in RP-HPLC is the degree of the sugar moiety protection. Whereas unprotected sugars decreased the retention time, partially or fully protected mono-saccharides dramatically increased the retention time of the parent peptide compound.

All pairs of diastereomers of the glycoconjugates studied can be well separated on the reversed-phase column, without using any chiral column. The retention times obtained for L-isomers were smaller than those for D-isomers, reflecting the increased hydrophobicity of the glycoconjugates in which either the phenylalanine or methionine residue was replaced by the corresponding D-enantiomer.

It has been demonstrated that glycation of the two closely related opioid pentapeptides methionine- and leucine-enkephalin, with identical monosaccharides, produced different effects on the retention times of the parent peptides.

The investigation of the dependence of the capacity factors on solvent composition suggested that the retention mechanisms of methionine-enkephalin-related glycoconjugates containing unprotected sugars in the molecule are influenced by the free hydroxyl group at C-1 of the carbohydrate moiety.

Abbreviations used for monosaccharides

Glc	D-glucopyranose
GlcAc₄	1,2,3,4-tetra-O-acetyl-β-D-glucopyranose
α-GlcAc₄	2,3,4,6-tetra-O-acetyl-α-D-glucopyranose
β-GlcAc₄	2,3,4,6-tetra-O-acetyl-β-D-glucopyranose
GlcBzl	benzyl β-D-glucopyranoside
GlcNAc	2-acetamido-2-deoxy-D-glucopyranose
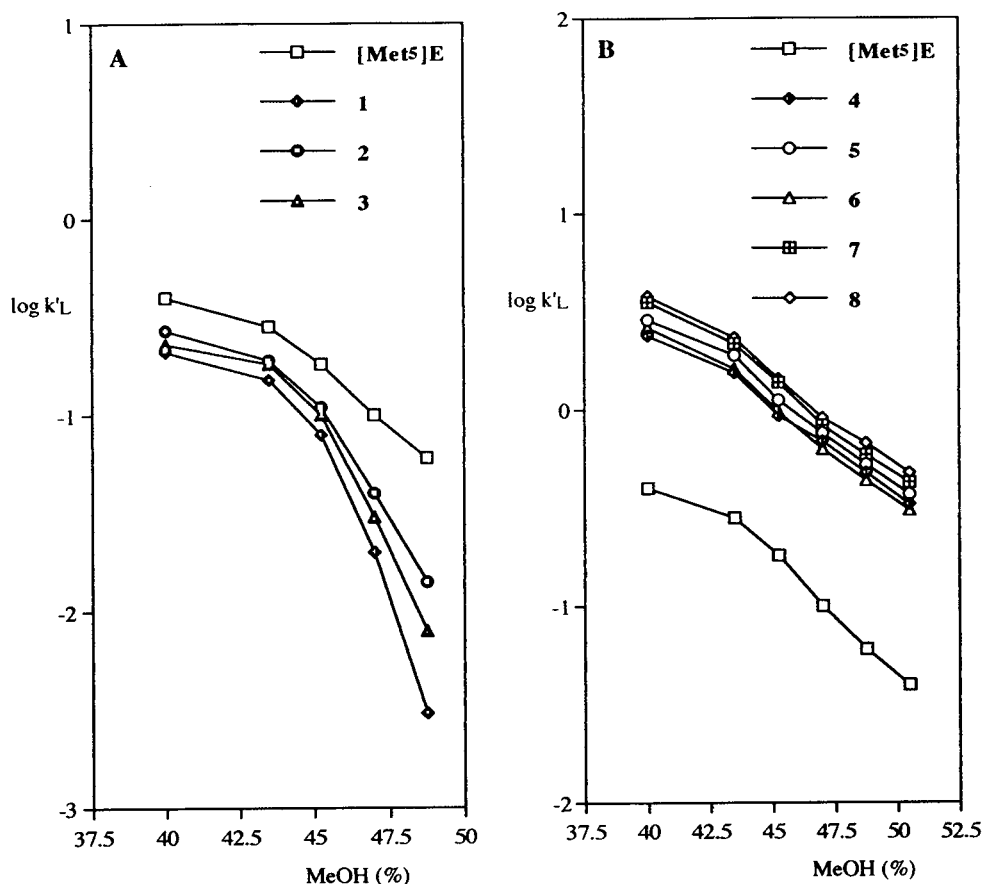
Fig. 3. Log k'_L values resulting from the chromatography of (A) glycoconjugates **1–3** and (B) glycoconjugates **4–8** with isocratic elution with concentrations of methanol in 0.1% aqueous TFA in the mobile phase of 40.00, 43.50, 45.25, 47.00, 48.75 and 50.50%. Other RP-HPLC conditions as in Table 1.

References

[1] L. Varga-Defterdarović, Š. Horvat, M. Skurić and J. Horvat, *J. Chromatogr.*, 687 (1994) 101.
[2] J. Hughes, T.W. Smith, H.W. Kosterlitz, L.A. Fothergill, B.A. Morgan and R.H. Morris, *Nature*, 258 (1975) 577.
[3] J. Horvat, Š. Horvat, C. Lemieux and P.W. Schiller, *Int. J. Pept. Protein Res.*, 31 (1988) 499.
[4] Š. Horvat, J. Horvat, L. Varga-Defterdarović, K. Pavelić, N.N. Chung and P.W. Schiller, *Int. J. Pept. Protein Res.*, 41 (1993) 399.

[5] M. Skurić, J. Horvat, Š. Horvat, N.N. Chung and P.W. Schiller, *Int. J. Pept. Protein Res.*, 43 (1994) 402.
[6] C.A. Browne, H.P.J. Bennett and S. Solomon, *Anal. Biochem.*, 124 (1982) 201.
[7] L. Gorbics, L. Urge, E. Otvos-Papp and L. Otvos, Jr., *J. Chromatogr.*, 637 (1993) 43.
[8] C.T. Mant, N.E. Zhou and R.S. Hodges, *J. Chromatogr.*, 476 (1989) 363.
[9] Y.C. Lee, B.I. Lee, N. Tomiya and N. Takahashi, *Anal. Biochem.*, 188 (1990) 259.
[10] L. Otvos, Jr., L. Urge and J. Thurin, *J. Chromator.*, 599 (1992) 43.
[11] Cs. Horvath, W. Melander and I. Molnar, *Anal. Chem.*, 49 (1977) 142.
[12] Ch. Yi, J.L. Fashing and P.R. Brown, *J. Chromatogr.*, 339 (1985) 75.
[13] A. Kaibara, C. Hohda, N. Hirahata, M. Hirose and T. Nakagawa, *Chromatographia*, 29 (1990) 275.

Determination of non-protein amino acids and toxins in *Lathyrus* by high-performance liquid chromatography with precolumn phenyl isothiocyanate derivatization

Jehangir K. Khan, Yu-Haey Kuo, Naod Kebede, Fernand Lambein*

Laboratory of Physiological Chemistry, Faculty of Medicine, University of Ghent, K.L. Ledeganckstraat 35, B-9000 Ghent, Belgium

Abstract

A simple procedure for the precolumn derivatization of toxic and non-toxic non-protein amino acids occurring in the legume crop *Lathyrus sativus* and other *Lathyrus* species with phenyl isothiocyanate and an HPLC method for the separation of the derivatives in the nano- and picomole range are reported. The results are compared with those of conventional ion-exchange chromatography for the separation of non-protein amino acids with postcolumn ninhydrin reaction. The relative standard deviations for the two methods are compared.

1. Introduction

Automated amino acid analysis was pioneered by Moore and Stein in the early 1950s [1]. They developed an amino acid analyser that automatically coupled the amino acid separation on anion-exchange columns with quantification based on the ninhydrin reaction. This has been the standard method for amino acid analysis since then, with the main variations being in the choice of buffers: sodium citrate-based buffers for the rapid determination of protein amino acids and lithium citrate buffers for the high-resolution analysis of physiological fluids or the determination of non-protein amino acids. Amino acid determination by reversed-phase HPLC after precolumn derivatization with Edman's reagent, phenyl isothiocyanate (PITC), is

now a well established method [2,3]. The chromatographic separation of the amino acid derivatives is a relatively fast procedure, and applications in medical research for the rapid determination of some selected amino acids in brain, plasma and urine have been reported [4,5]. However, for the identification of unknown amino acids, more information is available on the separation of unusual amino acids by ion-exchange methods than by the more recent HPLC methods.

Lathyrus sativus (khesari in India and Bangladesh, guaya in Ethiopia, san li dow in China, pois carré in France) is a popular drought-tolerant crop and foodstuff for several hundred million people in drought-prone areas of Africa and Asia. The overconsumption of *L. sativus* seed can cause an upper motor neurone degenerative disease known as neurolathyrism. The major neurotoxin present in the seeds and seedlings has

* Corresponding author.

been determined as 3-N-oxalyl-L-2,3-diamino-propanoic acid (β-ODAP, sometimes referred to as BOAA or β-oxalylaminoalanine) [6]. Recently, a chromatographic method to determine the neurotoxin β-ODAP and to differentiate it from the non-toxic α-isomer has been developed [7]. Higher plants produce a great diversity of non-protein amino acids, the physiological and ecological importance of which is poorly understood. The genus *Lathyrus* is especially rich in unusual amino acids, some of which contain a heterocyclic ring [8]; several have toxic properties [9]. The neurotoxin β-ODAP and homoarginine are the major amino acids in the seeds of *Lathyrus sativus*, while in the seedlings homoserine appears as a major free non-protein amino acid along with some heterocyclic amino acids derived from the isoxazolin-5-one ring [10]. BIA or β-(isoxazolin-5-on-2-yl)alanine (compound I) is the precursor for the neurotoxin β-ODAP [11–14]. The higher homologue of BIA is ACI or 2-(3-amino-3-carboxypropyl)isoxazolin-5-one (compound VI), which is found only in the genus *Lathyrus*. When ACI is given to young chicks along with food or by intraperitoneal injection, the symptoms of neurotoxicity develop [15]. ACI can be hydrolysed or photolysed by UV radiation with the formation of the neurotoxic compound 2,4-diaminobutanoic acid (DABA) [16] found in *Lathyrus sylvestris*. Another *Lathyrus* metabolite, CEI or 2-cyanoethylisoxazolin-5-one (compound VIII), can be hydrolysed, photolysed by UV radiation or metabolized with the formation of β-aminopropionitrile (BAPN), the osteolathyrogen from *L. odoratus* [16]. These toxic substances in the seedlings of *L. sativus* indicate that, while the seeds of *L. sativus* are mainly neurolathyritic (β-ODAP), the seedlings also contain an osteolathyritic toxin compound VIII in addition to a second neurolathyrogen compound VI [10,17]. Lathyrine or β-(2-aminopyrimidin-4-yl)alanine, which is metabolically linked to homoarginine, is another heterocyclic amino acid that is specific to *Lathyrus* species [18].

In order to determine more correctly the overall toxicity of *Lathyrus* species, we have developed an HPLC method that can simultaneously determine these various metabolites, and which is more sensitive and much faster than the conventional ninhydrin method.

2. Experimental

2.1. Materials

Phenyl isothiocyanate (99%) and triethylamine (99+%) were purchased from Aldrich, lithium citrate, citric acid and HCl from Merck, lithium chloride, lithium hydroxide and EDTA from Sigma and ammonium acetate from UCB. Acetonitrile and methanol were of HPLC grade. HPLC-grade water was obtained with an Elgastat UHQPS deionizing system. β-ODAP was purified from plant extracts [12] and also obtained by chemical synthesis [19]. The isoxazolinone compounds I, VI and VIII and lathyrine were purified from appropriate *Lathyrus* species. Standard amino acids were obtained from Pierce and Sigma.

2.2. Chromatographic systems

A Sykam S432 amino acid analyser with a postcolumn ninhydrin reactor was used for conventional amino acid analysis. The instrument was connected with an on-line UV detector (Linear UV-106), followed by the postcolumn ninhydrin reactor and then by detection at 570 nm. Two Chromatopac C-R6A (Shimadzu) integrators were used for data acquisition of the signals at 254 nm before postcolumn derivatization and 570 nm after postcolumn derivatization with ninhydrin. For the Sykam amino acid analyser an amino acid column (150 × 4 mm I.D.) packed with sulphonated polystyrene–divinylbenzene (LCA KO4), a strong cation exchanger, in the lithium form was used. This set-up and the buffers used [17] can be considered as a newer version of the system used previously for the identification and determination of UV-absorbing non-protein amino acids [20].

For the determination of precolumn-derivatized amino acids a Waters–Millipore Model 625 LC system, equipped with a column oven (Wa-

ters) to optimize the separation conditions was used. This HPLC system was connected with a Waters Model 991 photodiode-array detector with a scanning range from 200 to 800 nm to detect compounds of interest which did not react with PITC, and allowing identification of some peaks from the absorption spectrum. Data acquisition and integration for the Waters system were effected with Millennium 2010 chromatography manager software. For reversed-phase HPLC an Alltima C_{18} column (250×4.6 mm I.D.; 5-μm particle size) from Alltech was used. A guard column cartridge (Alltima C_{18}, 5 μm) was directly connected to the column.

2.3. Precolumn derivatization procedure

The derivatization reagents were freshly prepared every day by mixing methanol–water–triethylamine–PITC (7:1:1:1, v/v). Volumes of 100 μl of standard amino acids or samples were dried under vacuum. The residue was dissolved in 40 μl of the coupling buffer methanol–water–triethylamine (2:2:1, v/v/v), immediately dried under vacuum and then mixed with 60 μl of PITC derivatization reagent and allowed to stand at room temperature for 20 min. Excess reagent was then removed in vacuo. Before analysis the PITC derivatives were dissolved in 1 ml of buffer A, centrifuged for 10 min at 47 000 g, and filtered through a 0.45-μm Millipore filter.

2.4. Preparation of standard solutions

The standard amino acid solution for the ion-exchange amino acid analyser contained 0.5 mM of each amino acid in lithium buffer (pH 2.2) (injection buffer). A 40-μl volume was used per injection, containing 20 nmol of each amino acid. For the PITC derivatization, a stock solution, containing 6.25 mM of each amino acid in water, was used. After PITC derivatization and dissolution in buffer A, 20 μl were used per injection, containing 12.5 nmol of most amino acids, except aspartic acid, tyrosine and tryptophan (5 nmol each) and compound VIII (0.08 nmol). A standard mixture of ten non-protein amino acids occurring in *Lathyrus*

species, i.e., β-ODAP, compounds I, VI and VIII, Hse, Har, GABA, DAPRO (L-2,3-diaminopropanoic acid), BAPN and DABA was derivatized to evaluate the reproducibility and linearity of the method (Table 1).

2.5. Mobile phases

Solvent A consisted of 0.1 M ammonium acetate and was prepared freshly every other day. Solvent B was 0.1 M ammonium acetate in acetonitrile–methanol–water (46:10:44, v/v/v). Both buffers were adjusted to pH 6.5 with glacial acetic acid, filtered through a 0.22-μm membrane filter and degassed by purging with helium. The column temperature was optimized at 43°C. The buffers for the amino acid analyser were prepared as described previously [17].

3. Results and discussion

Standard non-protein amino acids were derivatized using the procedure described for precolumn derivatization. The separation of 30 compounds is shown in Fig. 1. Most of the amino acids in the standard were clearly resolved. Only phenylalanine and DABA were not separated under the conditions used and co-eluted at 45.3 min in peak 27. GABA and threonine, which co-eluted in peak 13, were only separated with a new column under optimum conditions. An unnatural amino acid, DL-allylglycine, was included as an internal standard, eluting at 34.5 min and not interfering with any of the compounds used.

The same standard solution was also analysed with the conventional amino acid analyser using postcolumn ninhydrin reaction and detection at 570 nm. The relative standard deviations for peak areas and retention times of the PITC amino acids are compared with the data for the ninhydrin reaction of ten amino acids in Table 1. The regression coefficients were calculated for twelve derivatizations of amino acid standards, ranging in amount from 100 pmol to 15 nmol, and they indicate a good linearity in this range.

All the selected non-protein amino acids,
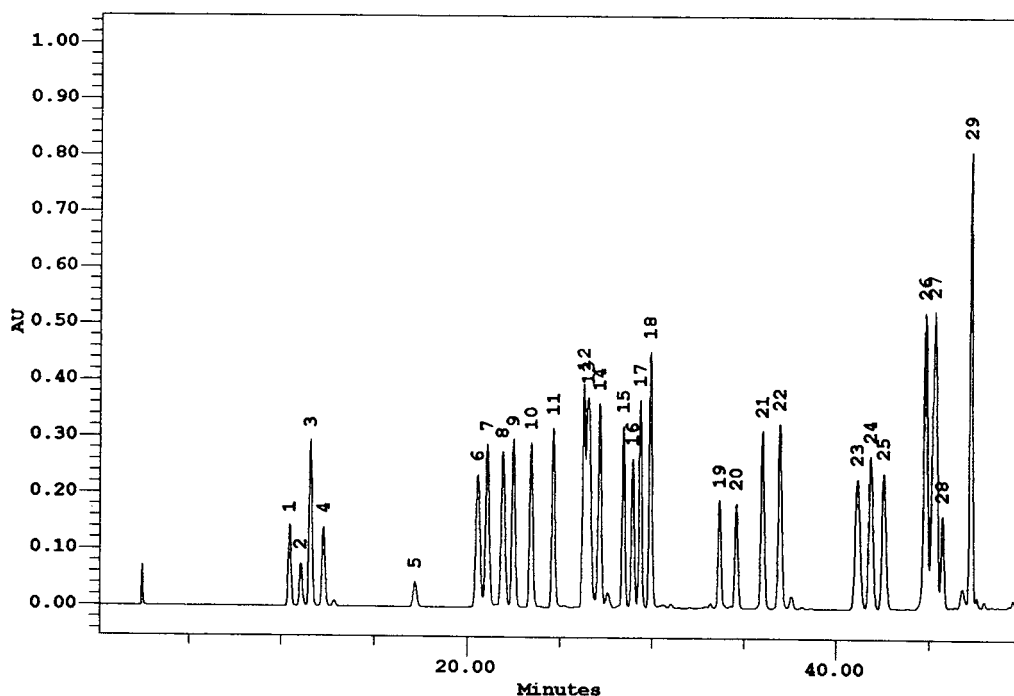
Fig. 1. Chromatogram of an amino acid standard mixture derivatized with PITC. Column temperature, 43°C. Peaks: 1 = Asp; 2 = compound VIII; 3 = β-ODAP; 4 = Glu; 5 = α-aminoadipic acid; 6 = Ser; 7 = Asn; 8 = Gly; 9 = Gln; 10 = Hse; 11 = His; 12 = Arg; 13 = Thr + GABA; 14 = Ala; 15 = Har; 16 = compound I; 17 = lathyrine; 18 = compound VI; 19 = Tyr; 20 = allylglycine (internal standard); 21 = Val; 22 = Met; 23 = BAPN; 24 = Ile; 25 = Leu; 26 = DAPRO; 27 = DABA + Phe; 28 = Trp; 29 = Lys.

Table 1
Relative standard deviations for peak areas and retention times (RT) of PITC amino acids compared with ninhydrin reaction

Amino acid	R.S.D. (%) for ninhydrin peak area ($n = 8$)	R.S.D. (%) for PITC peak area ($n = 10$)	R.S.D. (%) for ninhydrin RT (min) ($n = 8$)	R.S.D. (%) for PITC RT (min) ($n = 10$)	Regression coefficients for PITC ($n = 12$)[a]
β-ODAP	6.72	1.72	0.40	1.28	0.99826[b]
Compound I	9.35	2.07	1.85	0.73	0.92606
Homoserine	11.59	2.50	0.72	1.13	0.98782
Compound VI	3.68	1.99	0.54	0.63	0.98632
GABA	12.85	2.08	0.17	1.22	0.99824
DAPRO	3.87	3.01	0.19	0.33	0.99885
DABA	4.18	2.38	0.199	0.17	0.99112
BAPN	–	2.67	–	0.45	0.99953
Compound VIII[c]	–	3.66	–	1.24	0.98035
Homoarginine	11.29	2.60	0.37	0.56	0.97600

[a] Regression coefficients are from twelve derivatizations between 100 pmol and 15 nmol, except where indicated.
[b] Linear regression between 20 pmol and 3 nmol.
[c] Compound VIII does not react with PITC but can be detected at 254 nm.

Table 2
Gradient program for PITC amino acid analysis

Time (min)	Solvent A (%)	Solvent B (%)
0	100	0
15	90	10
30	60	40
40	50	50
50	0	100
55	0	100
57	100	0
65	100	0

Higher temperatures decreased the retention times and altered some peak resolutions; optimum resolution for this method was achieved at 43°C.

With the conventional amino acid analyser the amino acids in *Lathyrus sativus* seedlings were separated on the cation-exchange resin column (Li^+ form) using stepwise elution with five lithium buffers of increasing pH until the fourth buffer, and then by increasing the ionic strength. Selectivity effects in this system are primarily due to changes in the ionization of the amino acids and the general order of elution is acidic < neutral < basic amino acids. Li^+ buffers show a better selectivity than Na^+ buffers [21].

On the amino acid analyser compound VI elutes from the column at pH 3.30 and can be easily detected because of its absorbance at 254 nm before reacting with ninhydrin and at 570 nm after reacting with ninhydrin. Compound VIII is slightly retained and elutes at pH 2.75. Compound I, homoserine and homoarginine elute at

except the isoxazolinone derivative compound VIII, reacted with PITC under alkaline conditions to yield PITC derivatives. The reaction is rapid and quantitative and can take place at room temperature. The gradient programme used for the chromatographic separation of the selected non-protein amino acids, toxins and some protein amino acids is given in Table 2.

Fig. 2. Free amino acids and toxins in 3-day-old seedlings of *Lathyrus sativus*, without cotyledons, after PITC derivatization. Peak numbers as in Fig. 1.
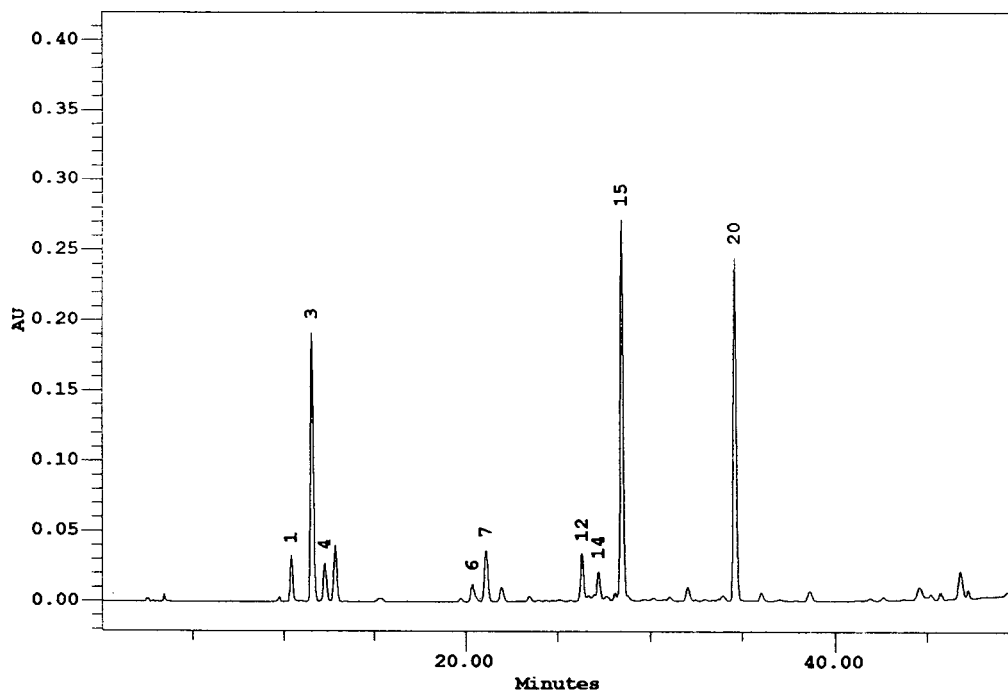
Fig. 3. Free amino acids in dry seeds of *Lathyrus sativus* after PITC derivatization. Peak numbers as in Fig. 1.

pH 2.75, 3.10 and 3.30, respectively. Homoarginine has a retention time of 195 min on the amino acid analyser with the buffer system used for high resolution, whereas on a reversed-phase column with PITC derivatization the retention time is 28.4 min. Compound VIII has a retention time of 11 min on the reversed-phase column. Although compound VIII does not react with PITC, it can be easily detected at 254 nm owing to the isoxazolinone ring. Compound I–PITC, VI–PITC and the osteotoxic β-aminopropionitrile (BAPN–PITC) have retention times of 28.9, 29.9 and 41.1 min, respectively. In addition to the shortened time of analysis, the relative standard deviation being 2–3% (2.47 ± 0.56%) for PITC and 4–13% (7.94 ± 3.80%) for the ninhydrin method, the reproducibility of quantification is superior to the ion-exchange method followed by ninhydrin reaction. The reproducibilities of the retention times are similar (Table 1).

Compound VIII is present in the seedlings and, when fed to vertebrates, it can be metabolized to BAPN [15]. Skeletal lesions that have been reported in neurolathyrism patients have recently been linked to the habit of consuming the young shoots of *Lathyrus sativus* and the presence of compound VIII [22,23].

Recent advances in the understanding of neuronal degeneration caused by overactivation of excitatory amino acid receptors has raised intriguing questions regarding additive or synergistic actions of compounds present in *Lathyrus sativus*. This excitotoxic process is perhaps not confined to the action of β-ODAP on neurones but may also extend to the toxic action of α-amino adipic acid on astrocytes [24].

Homoarginine in the seeds of *Lathyrus sativus* (up to 0.7%) [10] has been considered as a positive factor because it can be converted into the essential amino acid lysine by the mammalian liver. However, it is also a precursor of nitric oxide (NO), and as NO mediates glutamate neurotoxicity [25], a new direction for

research on and understanding of neuronal damage on consumption of the seeds of *Lathyrus sativus* is opened up.

The method reported here for the rapid determination of all the metabolites that may play a role in this toxicity should be useful for determining the overall toxicity of the seeds and seedlings (Figs. 2 and 3) of the many varieties of *Lathyrus sativus* and related *Lathyrus* species available, and of culinary preparations made from *Lathyrus*.

Acknowledgements

References

[1] S. Moore and W.H. Stein, *J. Biol. Chem.*, 192 (1951) 663.
[2] R.L. Heinrikson and S.C. Meredith, *Anal. Biochem.*, 136 (1984) 65.
[3] B.A. Bidlingmeyer, S.A. Cohen and T.L. Tarvin, *J. Chromatogr.*, 336 (1984) 93.
[4] S. Gunawan, N.Y. Walton and D.M. Treiman, *J. Chromatogr.*, 503 (1990) 177.
[5] R.A. Sherwood, A.C. Titheradge and D.A. Richards, *J. Chromatogr.*, 528 (1990) 293.
[6] S.L.N. Rao, P.R. Adiga and P.S. Sarma, *Biochemistry*, 3 (1964) 432.
[7] J.K. Khan, N. Kebede, Y.H. Kuo, F. Lambein and A. De Bruyn, *Anal. Biochem.*, 208 (1993) 237.
[8] F. Lambein, Y.-H. Kuo and A. De Bruyn, *Phytochem. (Life Sci. Adv.)*, 11 (1992) 145.
[9] F. Lambein, Y.-H. Kuo, G. Ongena, F. Ikegami and I. Murakoshi, in K. Takai (Editor), *Frontiers and New Horizons in Amino Acid Research*, Elsevier, Amsterdam, 1992, p. 99.
[10] F. Lambein, J.K. Khan and Y.H. Kuo, *Planta Med.*, 58 (1992) 380.
[11] F. Lambein, G. Ongena and Y.H. Kuo, *Phytochemistry*, 29 (1990) 3793.
[12] Y.H. Kuo and F. Lambein, *Phytochemistry*, 30 (1991) 3241.
[13] Y.H. Kuo, J.K. Khan and F. Lambein, *Phytochemistry*, 35 (1994) 911.
[14] Y.H. Kuo, F. Lambein, L.C. Mellor, R.M. Adlington and J.E. Baldwin, *Phytochemistry*, in press.
[15] F. Lambein and B. De Vos, *Arch. Int. Physiol. Biochim.*, 89 (1981) 66.
[16] A. De Bruyn, G. Verhegge and F. Lambein, *Planta Med.*, 58 (1992) 159.
[17] F. Lambein, J.K. Khan, Y.H. Kuo, C.G. Campbell and C.J. Briggs, *Nat. Toxins*, 1 (1993) 246.
[18] E.G. Brown and N.F. Al-Baldawi, *Biochem. J.*, 164 (1977) 589.
[19] F.L. Harrison, P.B. Nunn and R.R. Hill, *Phytochemistry*, 16 (1977) 1211.
[20] Y.H. Kuo, F. Lambein, F. Ikegami and R. Van Parijs, *Plant Physiol.*, 70 (1982) 1283.
[21] R.L. Cunico and T. Schlabach, *J. Chromatogr.*, 266 (1983) 461.
[22] D.F. Cohn and M. Streifler, *Arch. Suisses Neurol. Neuroch. Psychiat.*, 128 (1981) 151.
[23] A. Haque, M. Hossain, F. Lambein and E.A. Bell, *Nat. Toxins*, in press.
[24] R.J. Bridges, C. Hatalski, S.N. Shim and P.B. Nunn, *Brain Res.*, 561 (1991) 262.
[25] V.L. Dawson, T.M. Dawson, E.D. London, D.S. Bredt and S.M. Snyder, *Proc. Natl. Acad. Sci. U.S.A.*, (1991) 6368.

Gas chromatographic analysis of acid gases and single/mixed alkanolamines

Pushkar Shahi, Yun-feng Hu, Amit Chakma*

Department of Chemical and Petroleum Engineering, University of Calgary, 2500 University Drive, N.W., Calgary, AB T2N 1N4, Canada

First received 1 March 1994; revised manuscript received 3 August 1994

Abstract

A gas chromatographic (GC) technique is presented for the analysis of acid gases, alkanolamines and their degradation products in solution. The GC technique uses a Tenax GC column with a thermal conductivity detector. This simple and reliable GC technique can determine acid gases, alkanolamines and water in the solvent by a single sample injection and requires less time as compared to other existing analytical methods. The acid gases used in this study are CO_2 and H_2S and the alkanolamines used are monoethanolamine, methyldiethanolamine, 2-amino-2-methyl-1-propanol, diethanolamine and their degradation samples. This method has been compared to conventional techniques and the advantages of the GC technique have been demonstrated.

1. Introduction

Natural gas produced from gas fields contain varying amounts of CO_2 and H_2S. These acid gases must be removed from the natural gas prior to its transportation and subsequent use. Amine-based solvents are commonly used for their separation. In the amine process for natural gas sweetening, acid gases are absorbed into the solvent mostly by chemical reactions. The absorbed acid gases are then stripped off the rich solvent solutions in a stripping column and re-used. These solvents are also used for the simultaneous absorption of two or more acidic gases or for the selective removal of one over the other gases present in a mixture. Recently, tailor-made solvents consisting of amine blends are also being used since they offer energy savings and flexibility of operation over conventional processes based on single amines.

One of the key parameters in the design and operation of the absorber and regenerator in a gas plant is the so-called *acid gas loading*, which is the amount of acid gas that can be absorbed per unit amount of the solvent usually expressed in terms of mol of acid gas absorbed/mol of the amine solvent. Once this variable has been set, it immediately dictates the solution circulation rate in the plant and hence the economics of plant operation. This acid gas loading measurement indicates the extent of absorption or desorption taking place. Industrial amine units for the separation of acid gases from gaseous mixtures can be better designed and more efficiently

---

* Corresponding author.

operated if acid gas loadings (mol acid gas/mol amine) for different process streams can be accurately measured. For the proper operation of an ethanolamine gas plant, it is desirable to operate it close to the design parameters, among which the acid gas loading is of primary importance. Failure to do so not only results in an excessive use of regeneration energy but may also result in an unsatisfactory performance of the plant and give rise to any of the following problems: the produced gas may be off-specification, operation at higher acid gas loadings may cause corrosion of the process equipment, higher acid gas loading combined with high temperatures may cause the solution to degrade forming undesirable and non-regenerable products. Build-up of degradation products in the solvent directly causes a reduction in the absorption capacity of the solution. The degradation compounds so formed may also contribute to other problems such as corrosion [1], foaming [2] and fouling [3]. Therefore, it is important to closely monitor process variables, particularly acid gas loading, in order to ensure trouble-free and efficient operation of the plant. It is possible to save millions of dollars in energy, solvent and corrosion caused by analyzing for the acid gas content of the process streams [4].

There are various methods available for the separate analysis of acid gases, amines and their degradation product concentration in solution. However, these existing methods are either very tedious or inaccurate in some specific cases. Moreover, none of these methods can determine all species present in the sample simultaneously by a single analysis. Only recently analytical equipment capable of determining CO_2 and H_2S concentrations in amine solutions based on a combination of UV and IR spectroscopy has been developed [4].

In this paper a gas chromatographic (GC) technique is presented, which is capable of performing the analysis of acid gases, amines and their degradation products by a single sample injection into the chromatograph. This results in a considerable savings in time and also ensures that the sample quality is unperturbed.

2. Analytical techniques

2.1. CO_2 and H_2S loading

Some of the commonly used techniques for the determination of CO_2 loading in amines are (i) volumetric method well suited to routine analysis, (ii) quantitative precipitation of the dissolved gas as metal salt with simultaneous formation of an acid and (iii) titrimetric methods [5]. More often than not, when the sample contains very low concentration of dissolved CO_2, none of these methods is satisfactory. H_2S loading may be determined by a wet chemistry method as well [6].

2.2. Amines and their degradation products

For the analysis of amines and their degradation products in particular, a number of analytical methods have been described by Choy [7]. These methods include wet chemical techniques, IR and UV spectroscopy, paper and thin-layer chromatography and GC. The total amine concentration in the solution can be determined by a simple acid–base titration using a suitable indicator. However, if there is a mixture of two or more amines, individual analysis becomes a problem unless their titration end points are reasonably far apart. The selection of a column for amine analysis by GC involves a number of problems. Brydia and Persinger [8] described a chromatographic technique for the analysis of ethanolamines. Excessive peak tailing took place while employing direct GC techniques so they investigated derivatization with trifluoroacetic anhydride prior to chromatographic separation. Derivatization makes the amine more volatile, less polar, and hence, more amenable to GC analysis [8–10]. Trifluoroacetic anhydride, however, reacted with water and the resulting acid again caused tailing problems. This problem was taken care of by Piekos et al. [9], by converting the alkanolamines into trimethylsilyl derivatives. This silylation process yielded fairly stable compounds which were more easily separated and identified by GC. But a major limitation of this

method was that water concentrations of only up to 5% could be tolerated provided that the silylating agent was used in excess.

Since the water content of industrial amine solution typically ranges from 50 to 90%, the method suggested by Piekos et al. [9] was not directly applicable. Choy and Meisen [10] modified the technique by stripping water from the amine solution using air. Removal of water, however, resulted in the precipitation of some of the degradation products. The dried sample was then dissolved in dimethylformamide and the resulting mixture silylated. It was then separated using a Chromosorb column and flame ionization detection.

The method employed by Choy and Meisen [10], although reliable, was too time consuming for wide industrial applications because of the extensive sample preparation. Also, incomplete silylation of certain compounds posed a problem in the reliability of the technique. Saha et al. [11], used a Tenax GC column to separate alkanolamines. This column successfully separated a mixture of monoethanolamine (MEA), diethanolamine (DEA) and triethanolamine (TEA) within about 8 min. A Perkin-Elmer Model 900 gas chromatograph with a flame ionization detector was used by them. Tenax GC is a porous polymer based on 2,6-diphenyl-*p*-phenylene oxide, which has a weakly interacting surface and can be used at temperatures upto 450°C [12]. Operating the column at so high temperatures posed the probable problem of thermal degradation [13]. However, Saha et al. [11], showed that alkanolamines did not undergo rapid thermal degradation up to 375°C. As a result of the short residence time of the solvent in the column, degradation was considered insignificant [13]. Kennard and Meisen [14] presented a relatively simple technique using a Tenax GC column but the technique was limited to amines and their degradation products. Dawodu and Meisen [15] have reported the use of various other columns exclusively for the detection of amines and their degradation compounds. These include Tenax TA column (packed column), Supelcowax 10 (capillary col-

umn), DB-Wax column (capillary column) and HP-17 (capillary column). They showed that the Supelcowax 10, a capillary column lined with polyethylene glycol, performed better than a packed Tenax GC column for the analysis of fresh and partially degraded alkanolamine solutions.

2.3. Combined analysis of gases and amines

A GC technique was devised by Wisniewski [16], which could analyze for amines on one column and the acid gases and water on another column. Even that could not be accomplished in a single sample injection. While analyzing for the acid gases and water, the column had to be pretreated with HCl. This acid scrubbed the amine from the sample and allowed the acid gases and water to elute for analysis. However, saturation of the column with the amine made reproducibility difficult. Ethanolamines being very reactive compounds with a polar hydroxyl and amine group have a strong adsorption affinity for siliceous column supports. Also, the success of a stationary liquid support was improbable due to the strong physicochemical interactions and slow diffusion of the amines through the associated liquids [13]. Robbins and Bullin [13] reported the development of a GC-based technique for the determination of acid gases and hydrocarbons in gas mixtures and also the loading of acid gases in alkanolamine samples. A Tenax GC column was initially used by them which apparently proved incapable of producing a separation between the light gaseous components. To separate the light gaseous components, a simple column-switching device with two columns in series was employed. The first column containing Tenax GC was designed to give good separation between the amine and light gaseous components. The second column containing Poropak Q was designed to separate the light gaseous components. Poropak Q is an ethylvinylbenzene–divinylbenzene copolymer that is cross-linked and can withstand temperatures upto 250°C before column degradation occurs. Upon sample injection the sample vapor-

ized and the light gaseous components flowed rapidly through the two columns. The amine with its interacting groups travelled much slower. The valve and flow arrangement protected the Poropak Q column from irreversible adsorption or deactivation by the amine. However, apart from the fact that they did not report an extensive comparison with other techniques to confirm their results, the reported technique has a few limitations. Normally, in a single GC system, the oven chamber is common to both the columns. As a result, temperature in the oven cannot be exceeded above 250°C, since it would adversely affect the Poropak Q column. All alkanolamines do not completely elute before this temperature. As a result, the full extent of operability of the Tenax GC column which can function upto 450°C cannot be utilized. There was no mention of the analysis or detection of degradation compounds. Even if there was an attempt to analyze these compounds it would be limited by the temperature restriction. It is expected that there would be some passage of amine with time into the Poropak Q column thus causing irreversible adsorption of amine and thus deactivation of the column. Kim and Sartori [17] have also reported the use of Tenax columns in combination with other columns for the analysis of degraded amine samples. Accordingly two analyses were carried out for each sample; one on the Tenax GC column to determine the high-molecular-mass compounds and the other on an SE-30/Chrom G HP column for the lighter compounds.

3. Experimental

Analysis of samples in the laboratory was performed by the volumetric method, the wet chemistry technique and the GC technique and a comparative study was carried out.

3.1. Volumetric method

This method involves reacting a known amount of sample with a given concentration of sulphuric acid solution in a closed vessel. The volume of CO_2 evolved is measured and is converted to its mol equivalent at the existing temperature and pressure conditions [5].

The analysis of CO_2 and other acid gases dissolved in bases, by acidification of the solution and measurement of the volume of evolved gas, necessitates corrections for temperature, pressure, static head, solubility etc., which were carried out accordingly. The buffer solution used in the manometer was either one of the mixtures enlisted below, each of which was tested for CO_2 absorption and was found not to absorb any amount whatsoever.

Buffer mixture I: Prepared by dissolving 100 g of sodium sulphate in 500 ml of water and adding 20 ml of concentrated H_2SO_4. Ten drops of methyl red are then added to the final solution.

Buffer mixture II: This buffer is prepared by dissolving 100 g of sodium chloride in 350 ml of water and adding 1 g of sodium hydrogencarbonate. A 2-ml volume of methyl orange indicator is finally added to the resulting solution.

This method, however, suffered from the following disadvantages:

(i) Very small changes in the CO_2 concentration were indistinguishable, as it did not reflect in the volume of CO_2 evolved on acidification.

(ii) Repeatability was found to be poor.

(iii) There is always a residual amount of CO_2 left in the sample after reacting with acid. This can be confirmed by a GC analysis, albeit at the risk of letting H_2SO_4 enter the chromatographic column. Introducing H_2SO_4 into the column causes column damage and is therefore avoided. This residual CO_2 is not obtained even by mild heating of the acidified solution [18]. Though it is a very small amount it becomes appreciable when CO_2 loadings are low. The residual CO_2 is not a constant volume which can be determined once and added to the CO_2 volume finally obtained. It varies with the concentration of the amine as well as its loading.

(iv) If the released gas is to be swept by inert gas followed by absorption into a solvent, then the time consumption increases and a single analysis might take over 1 h.

3.2. Wet chemistry method

For the determination of CO_2 concentrations in the amine by this method, the amine sample is mixed with an excess of standard base and heated to boiling. Since the amine–acid gas complex is thermally unstable, the acid gas is converted into an ionic species and is precipitated as an appropriate metal salt. The filtrate is titrated with a standard acid to determine the concentration of uncarbonated amine in the sample. Bromocresol green, cresol red and phenolphthalein are commonly used to indicate the end points [13,21–23]. The total amine in the solution is usually determined by titrating the liquid sample with a standard acid in the presence of an indicator. Bromophenol blue, methyl orange and methyl red are used to indicate the end point in this titration [19,22,24]. The CO_2 content of the solution is calculated by the difference between the total amine and the uncarbonated amine values assuming a 1:1 stoichiometry between the CO_2 and the amine [5]. Hikita et al. [22] confirmed this stoichiometric ratio for tertiary amine, TEA.

Several problems are typically encountered in the preparation of carbonated amine samples under pressure. Significant flashing usually takes place in most cases where the pressure is released on rich amine solutions. This method does not give good results unless sufficient time is allowed for complete precipitation. Other problems include the loss of some of the precipitate while washing, filtering and drying. The method fails totally when the loading of samples is very low (less than 0.06). Thus, in short, the following problems can be highlighted for this technique: (i) the whole procedure is very tedious and time consuming; a single sample analysis could take well above an hour, (ii) this method, in general, seems to be applicable only at high loadings of CO_2.

The wet chemistry technique used in the laboratory comprised the $BaCl_2$ precipitation technique. This method involved preparing a solution of 0.1 M $BaCl_2$ in water which had previously been heated and bubbled with nitrogen. At low pressure (close to atmospheric), the amine sample was directly added to excess of this solution to form barium carbonate precipitate. The precipitate so obtained was then filtered using Whatman 42 or Whatman 5 filter paper. All along the filtration process, the sample was kept covered to disallow any contact with air. After this, the filter paper along with the precipitate were washed with distilled water until the filtrate reached a pH of 5–6. The precipitate along with the filter paper was then dissolved in water until a pH of 4.0 was reached. It was then titrated against 0.1 M HCl to determine the CO_2 content. At high pressure, sample was directly withdrawn into a caustic solution to fix the CO_2 present in the carbonate form. But care has to be taken not to collect amine sample in excess of caustic. The NaOH amount should be just 2–3 times the amount of sample. A shortcut method in the precipitation technique is the gravimetric method whereby the precipitate is washed, dried and directly weighed to give the amount of CO_2 present by stoichiometric calculations from the precipitation reaction.

3.3. GC technique

The GC technique developed involved a number of trials with various operating conditions of the chromatograph and use of a suitable column. For the analysis of a sample consisting of CO_2, H_2S, water and amines together, separation could not be obtained in a single column. A combination of two columns in series, more or less on the lines of Robbins and Bullin [13] had to be resorted to. A combination of Tenax GC and Haysep Q columns was tried with a valve-switch arrangement (Fig. 1). Initially, all the
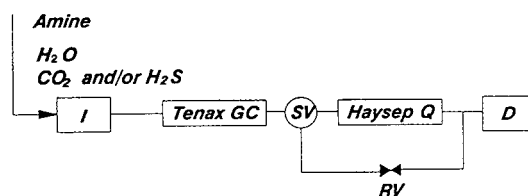

Fig. 1. GC dual-column configuration showing valve-switch arrangement. I = Injection port; SV = switch valve; RV = restriction valve; D = detector.

lighter components passed through the Haysep Q column until the H_2S peak was obtained. Then the valve was switched, so that only amine went through. Haysep Q can normally only withstand temperatures upto 250°C, after which polymer degradation occurs. However, for samples containing only CO_2, excellent detection of the gas, amine and its degradation products could be obtained by a single injection of the sample with a single Tenax GC column.

Equipment used

A Hewlett-Packard Series 5890 A gas chromatograph was used for analysis. It employs a 6 ft. × 1/8 in. (1 ft. = 30.48 cm; 1 in. = 2.54 cm) column packed with 80–100 mesh Tenax GC (175–147 μm). Tenax GC is a porous polymer based on 2,6-diphenyl-*p*-phenylene oxide, which has a weakly interacting surface and can be used at temperatures upto 450°C [12]. The Haysep Q packed column used was 8 ft. in length. Thermal conductivity detection was used. No sample preparation was required as the column was unaffected by the presence of water. Temperature programming was done with an initial temperature of 35°C, initial time of 1 min, oven maximum temperature of 300°C, final temperature of 280°C, injector port temperature of 280°C, detector temperature of 300°C and the rate of rise in temperature was 30°C/min. Helium was used as the carrier gas with a fixed flow-rate of 30 ml/min.

Procedure

Carbonated samples of different amines and their mixtures were prepared and injected into the GC column at the injection port temperature of 280°C. A precision syringe (Model 7102 KN, Hamilton) fitted with a Channey adapter and a 2-μl needle (Model 7102 RN, Hamilton) were used for sample injection. The injected sample size was 0.5 μl. This immediately vaporized the sample in the injection port. Dissolved and chemically combined CO_2 elutes from the column first, followed by water and amine. Degradation products in the sample give distinct peaks and are easily identifiable. Each sample was

injected at least three times and the peak areas of the components were averaged.

Equipment performance and maintenance

The septum at the injection port was changed after 20–25 injections to prevent any leakage. The column was conditioned after every 150 injections for 8–10 h. Calibration curves allowed determination of concentration of each species.

Calibration

Apart from the use of a suitable column and appropriate temperature programming, the calibration curves hold the key to the success of this technique. Calibration curves were prepared for individual and mixed amines as well as for their CO_2 content. The GC column was calibrated by using two different schemes. For standard samples prepared under equilibrium conditions, the peak area obtained was calibrated against literature values for solution loading. On the other hand, for solutions which were carbonated for a short time (less than 30 min) and therefore not under equilibrium, the column was calibrated using the volumetric technique. This was done by taking at least five readings using the volumetric technique and then averaging them to obtain the solution loading. These calibration curves were separate for CO_2 in the gaseous phase and that in the liquid (amine) phase. The chromatographic area of CO_2 present in amine solution did not correspond to that present in gaseous phase. This is an important criterion which should be taken into account while preparing calibration curves.

4. Results and discussion

Carbonated samples of amines and their mixtures which had attained equilibrium with CO_2 were analyzed by the different techniques and the results were compared with the average solubility values of CO_2 existing in literature. The exact amine concentration for single amine solvents was determined by titration. Both qualitative and quantitative analysis of various samples of carbonated amines and their mixtures

could be accurately carried out using the GC technique developed.

4.1. Elution order

For carbonated amine samples, the elution order was according to the sequence: air/nitrogen, CO_2, water, amine and degradation products. For samples containing H_2S, the elution order was: air/nitrogen, water, H_2S, amine and degradation products.

4.2. Separation with dual columns

Figs. 2 and 3 show the chromatograms obtained in the case of employing a combination of Tenax GC and Haysep Q columns in series with the valve-switching arrangements (Fig. 1). As can be seen in Fig. 2, separate and sharp peaks

Fig. 3. Chromatogram using dual column showing air, water, H_2S and DEA peaks. $H_2S = 3.4$ min.

were obtained for a carbonated DEA sample. In the case of a DEA sample containing H_2S, a very wide peak was obtained for H_2S and there seemed to be an interference with the water peak. While employing dual columns, care has to be taken to switch valves at the right time to avoid any amine to enter the Haysep Q column, thereby causing irreversible adsorption or deactivation by the amine.

4.3. Separation with Tenax GC column

Analysis of CO_2 in single amines

Figs. 4 and 5 show chromatograms of CO_2-loaded MEA and MDEA samples, respectively. The air, CO_2, water and amine peaks were distinct. The air peak was due to the air entrained in the sample from the injection syringe. The CO_2 loadings of the MEA and MDEA
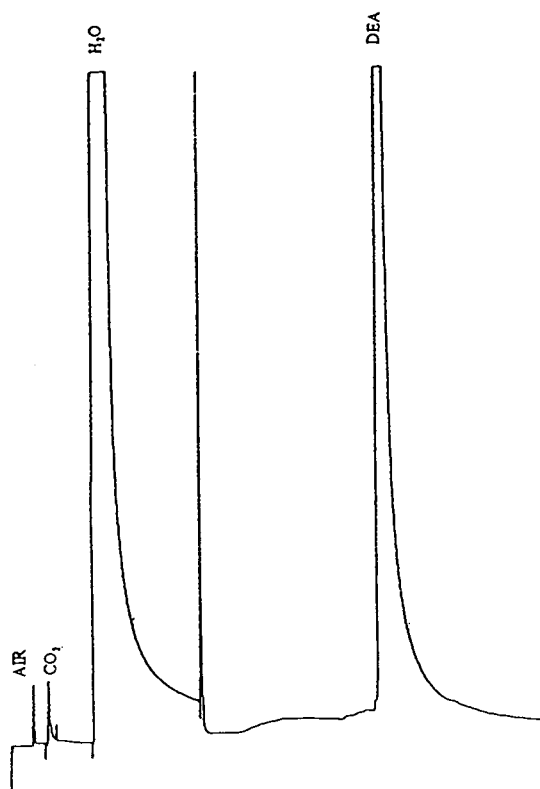

Fig. 2. Chromatogram using dual column showing air, CO_2, water and DEA peaks. Retention times (in min): air = 0.233; $CO_2 = 0.451$; water = 2.517; DEA = 9.6.
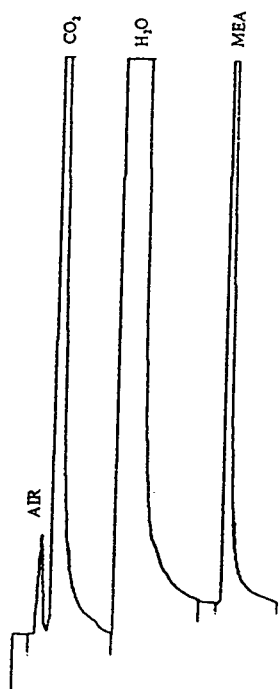
Fig. 4. Chromatogram using Tenax GC column showing air, CO_2, water and MEA peaks. MEA = 6.31 min.

Fig. 5. Chromatogram using Tenax GC column showing air, CO_2, water and MDEA peaks. MDEA = 8.147 min.

samples are 0.46 and 0.02, respectively. As can be seen such low acid gas loading levels can be adequately detected and quantified by the GC technique.

Analysis of CO_2 in amine mixtures

Fig. 6 shows a chromatogram of a CO_2-loaded sample of an amine blend consisting of 5% (w/w) MEA and 45% (w/w) MDEA. The separation of the MEA and MDEA peaks is also very sharp. It should be noted that the technique does not give an indication of the loading of CO_2 for individual amines. It provides an overall loading for the mixture.

Fig. 7 shows a carbonated blended amine sample along with degradation products. This chromatogram is a classic example of the capability of this GC technique to analyze as many amines as are present with the associated degradation products and the dissolved CO_2. Industrial samples could be tested from time to time

for the detection of amine degradation products to maintain the quality of amine being used.

Analysis of CO_2 in partially degraded amine solutions

Fig. 8 shows a chromatogram of a partially degraded DEA sample containing a number of degradation compounds. The results are reproducible even at very low concentrations of CO_2. The GC technique developed provides the concentration of all the species from a single sample injection. The complete analysis takes less than 10 min when no degradation products are present. In the presence of high-boiling degradation products, the analysis can be completed within 30 min. This translates into a considerable savings in time when compared to conventional techniques.
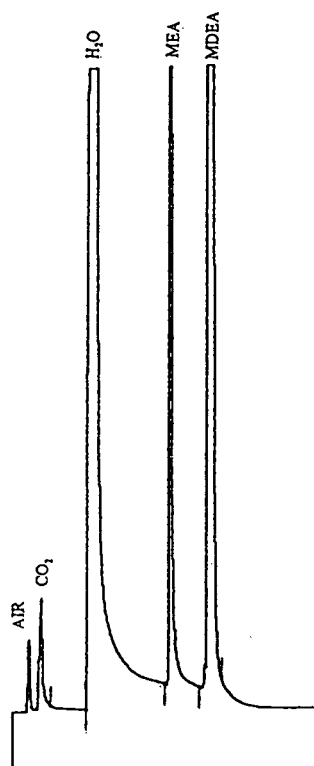
Fig. 6. Chromatogram using Tenax GC column showing air, CO_2, water, MEA and MDEA peaks.
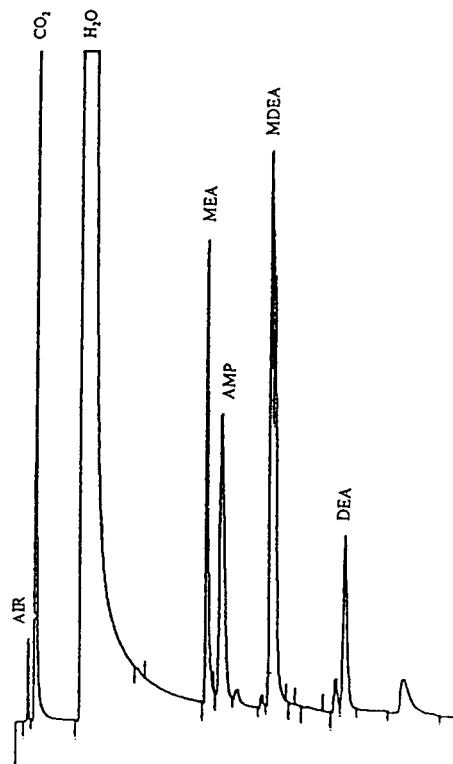
Fig. 7. Chromatogram using Tenax GC column showing air, CO_2, water, MEA, 2-amino-2-methyl-1-propanol (AMP), MDEA and DEA peaks. AMP = 6.85 min.

Sensitivity

Most of the available techniques are satisfactory for high loading of amines. The GC technique developed here, however, can detect CO_2 in highly loaded samples as well as samples of amines which have been merely exposed to air. Loadings as low as 0.003 were detected by this technique. Fig. 9 shows the detection of CO_2 in an MDEA solution which was just exposed to air (loading 0.003).

Comparison of the GC technique with conventional methods

A comparison of different techniques has been presented in Table 1. Each sample was analyzed at least three times. As can be seen from the table, reproducibility for the volumetric technique was within 16% for the equilibrated samples. However, for the non-equilibrated 50%

MDEA samples, the reproducibility varied significantly. The $BaCl_2$ method could not determine CO_2 in the non-equilibrated sample. For equilibrium samples, it consistently gave lower loading values compared to the literature values of solubility. Table 2 shows the standard deviations for various techniques from the mean loading values obtained and from the actual literature values. The superiority of the GC technique can be easily seen from the low values of its standard deviation.

5. Conclusions

(1) A GC technique based on a single column has been developed to analyze gas-treating

Table 1
Comparison of various techniques for CO_2 analysis in amines

Amine (%, w/w)	Sample	Equilibrium solubility	CO_2 loading (mol CO_2/mol amine)		
			GC method	Volumetric method	BaCl$_2$ method
MEA	1	0.735[a]	0.725	0.82	0.56
(15%)	2		0.73	0.78	0.64
	3		0.732	0.93	0.7
MDEA	1	0.83[b]	0.82	0.67	0.8
(50%)	2		0.822	0.75	0.73
	3		0.83	0.76	0.69
5% MEA +	1	0.6[c]	0.6	0.7	0.42
45% MDEA	2		0.6	0.65	0.5
	3		0.6	0.68	0.51
MDEA	1	0.06[d]	0.058	0.03	Traces
(50%)	2		0.059	0.1	
	3		0.0625	0.1	

MDEA = Methyldiethanolamine; MEA = monoethanolamine.
[a] From Ref. [20].
[b] From Ref. [19].
[c] Calculated from solubility model developed for mixed amines.
[d] Non-equilibrium sample.

Table 2
Standard deviation of loading values obtained by various techniques

Amine (%, w/w)	Standard deviation					
	From mean loading			From literature value		
	GC method	Volumetric method	BaCl$_2$ method	GC method	Volumetric method	BaCl$_2$ method
MEA (15%)	$2.5 \cdot 10^{-3}$	0.052	0.049	$5.78 \cdot 10^{-3}$	0.108	0.101
MDEA (50%)	$5.4 \cdot 10^{-3}$	0.035	0.039	$6.4 \cdot 10^{-3}$	0.096	0.087
5% MEA + 45% MDEA	0	0.019	0.035	0	0.068	0.087
MDEA (50%)	$1.67 \cdot 10^{-3}$	0.028	–	$1.67 \cdot 10^{-3}$	0.032	0.112

Deviations calculated for values reported in Table 1.

amine solvents and dissolved acid gases in solution.

(2) No sample preparation is required for analysis.

(3) The method introduces specificity, repeatability, reliability and accuracy.

(4) CO_2 loadings as low as 0.003 can be detected accurately.
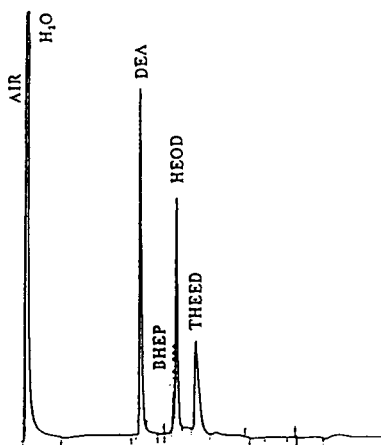
Fig. 8. Chromatogram using Tenax GC column showing air, CO_2, DEA and degradation products N,N-bis(hydroxy-ethyl)piperazine (BHEP), N,N,N-tris(hydroxyethyl)ethyl-enediamine (THEED) and 3-(hydroxyethyl)-2-oxazolidone (HEOD). BHEP = 12.1 min; HEOD = 12.6 min; THEED = 15.3 min.

(5) H_2S can also be detected using the single column when CO_2 is not present.

(6) For simultaneous determination of H_2S and CO_2 in a sample, dual columns have to be employed with a valve-switching arrangement.

Acknowledgements
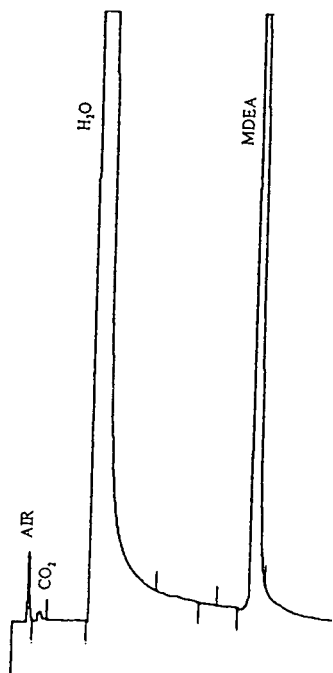
Fig. 9. Chromatogram using an extremely low loading of CO_2 in MDEA (loading = 0.003).

References

[1] A. Chakma and A. Meisen, *Ind. Eng. Chem. Prod. Res. Dev.*, 25 (1986) 627.
[2] C.R. Pauley, R. Hashemi and S. Caothien, *Oil Gas J.*, 87 (1989) 67.
[3] A. Chakma and A. Meisen, *Can. J. Chem. Eng.*, 65 (1987) 264.
[4] D. Ballard, E. Baughman and K.G. Schick, presented at the *42nd Lawrence Reid Gas Condensation Conference, Norman, OK, 1992.*
[5] R.H. Weiland and O. Trass, *Anal. Chem.*, 41 (1969) 1709.
[6] M.H. Li and K.P. Shen, *J. Chem. Eng. Data*, 38 (1993) 105.
[7] E.T. Choy, *Masters Thesis*, University of British Columbia, Vancouver, 1978.
[8] L.E. Brydia and H.E. Persinger, *Anal. Chem.*, 39 (1967) 1318.
[9] R. Piekos, K. Kobyiczyk and J. Grzybowski, *Anal. Chem.*, 47 (1975) 1157.
[10] E.T. Choy and A. Meisen, *J. Chromatogr.*, 187 (1980) 145.
[11] N.C. Saha, S.K. Jain and R.K. Dua, *Chromatographia*, 10 (1977) 368.
[12] R. van Wijk, *J. Chromatogr. Sci.*, 8 (1970) 418.
[13] G.D. Robbins and J.A. Bullin, *Energy Prog.*, 4 (1984) 229.
[14] M.L. Kennard and A. Meisen, *J. Chromatogr.*, 267 (1983) 373.
[15] O.F. Dawodu and A. Meisen, *J. Chromatogr.*, 629 (1993) 297.
[16] D.F. Wisniewski, presented at the *Gas Conditioning Conference, Norman, OK, 28–29 March 1961.*
[17] C.J. Kim and G. Sartori, *Int. J. Chem. Kinet.*, 16 (1984) 1257.
[18] R.W. Swick, D.L. Buchanan and A. Nakao, *Anal. Chem.*, 24 (1952) 2000.

[19] F.Y. Jou, A.E. Mather and F.D. Otto, *Ind. Eng. Chem. Proc. Des. Dev.*, 21 (1982) 539.

[20] J.H. Jones, H.R. Froning and E.E. Claytor, Jr., *J. Chem. Eng. Data*, 4 (1959) 82.

[21] P.W. Coldrey and I.J. Harris, *Can. J. Chem. Eng.*, 54 (1976) 566.

[22] H. Hikita, S. Asai, H. Ishikawa and M. Honda, *Chem. Eng. J.*, 13 (1977) 7.

[23] M.B. Jensen, E. Jorgensen and C. Faurholt, *Acta Chem. Scand.*, 8 (1954) 1137.

[24] *Gas Conditioning Fact Book*, Dow Chemical, Midland, MI, 1962.

Application of solid-phase microextraction and gas chromatography with electron-capture and mass spectrometric detection for the determination of hexachlorocyclohexanes in soil solutions

Peter Popp[a],[*], Karsten Kalbitz[b], Gudrun Oppermann[a]

[a]Department of Analytical Chemistry, Centre for Environmental Research Leipzig/Halle GmbH, Permoserstrasse 15, 04318 Leipzig, Germany
[b]Department of Soil Sciences, Centre for Environmental Research Leipzig/Halle GmbH, Hallesche Strasse 44, 06246 Bad Lauchstaedt, Germany

Abstract

Solid-phase microextraction (SPME) is a method for the extraction of organic compounds from aqueous samples. The analytes are extracted into a stationary phase placed on a fused-silica fibre and are thermally desorbed in the injector of a gas chromatograph. The connection of GC with electron-capture (ECD) and mass spectrometric (MS) detection with the SPME method makes it possible to determine low concentrations of organochlorine compounds in aqueous solutions. With hexachlorocyclohexanes (HCHs) detection limits between 5 ng/l (for α- and γ-HCH with the combination of SPME and GC–ECD) and 80 ng/l (for β-HCH with the combination of SPME and GC–MS) were calculated. The SPME–GC method was used to investigate the mobility of HCHs in wetland soils near Bitterfeld. The results of this study show the high mobility of β-HCH despite the low water solubility and the long persistence of β-HCH in soils. The proportion of β-HCH in the total HCH concentration is higher in soil solutions (80–90%) than in soils.

1. Introduction

Hexachlorocyclohexanes (HCHs), mainly β-HCH, are widespread in the riverine area of the river Mulde near Bitterfeld, so it is necessary to investigate the mobility of HCHs in the wetland soils to assess the danger for ground water, plants, animals and humans. The large number of samples and the urgency of this problem require a rapid and inexpensive method for the extraction of HCHs from aqueous samples, in particular from soil solutions. Such a method was found using solid-phase microextraction (SPME). First described by Pawliszyn and co-workers [1–6], SPME is a solvent-free, rapid and inexpensive method for the extraction of organic compounds from aqueous samples. The SPME technique is based on chemically modified fused-silica fibres fixed inside a syringe. The fibre with the immobilized organic film is exposed to aqueous samples and organic compounds are extracted from the water into the silicon phase by

* Corresponding author.

diffusion processes. After this procedure the microextractor is directly inserted into the split–splitless or on-column injector of a gas chromatograph.

First applications of this method to the determination of contaminants in water have shown that SPME is a practical alternative to other commonly used extraction techniques (headspace, purge-and-trap, liquid–liquid extraction).

The main objective of this work was the optimization of the SPME procedure for the measurement of HCHs and the application of the optimized procedure coupled with GC–ECD and GC–MS for the determination of these compounds in environmental samples, especially for the determination of the concentration distributions of α-, β- and γ-HCH in soil solutions from different sites.

2. Theoretical

The amount of an analyte absorbed by the fibre at equilibrium (infinite volume assumed) is linearly dependent on the concentration of the analyte in the aqueous phase according to

$$n_s = KV_sC_{aq}$$

were n_s is the number of moles of the analyte absorbed by the stationary phase, K is the distribution constant of a compound between the stationary and the aqueous phase, V_s is the volume of the stationary phase and C_{aq} is the initial concentration of the analyte in the aqueous phase.

Louch et al. [4] showed that in a finite volume V_{aq} the amount of the analyte in the stationary phase is given by

$$n_s = \frac{KV_sV_{aq}C_{aq}}{KV_s + V_{aq}}$$

This means that a linear relationship between the concentration of analytes in aqueous samples and the response of a GC detector is to be expected if the absorption conditions in the sample vial and the desorption conditions in the injection port of the GC are reproducible. Louch et al. [4] also developed a mathematical model for the dynamics of the absorption process. Assuming that the dynamics of extraction is a diffusion-determined process and solving Fick's second law of diffusion, they calculated time profiles for perfectly stirred samples of infinite volume.

The diffusion of analytes into the fibre coating from an unstirred solution of finite volume was also calculated. It was shown that the time necessary to reach equilibrium for a perfectly agitated sample is relatively short. Without mixing, the equilibrium time increases considerably.

3. Experimental

3.1. Site description and extraction of soil solution

The sites Bobbau, Keller and Spittel are located in the floodplain of the stream Spittelgraben. This stream flows into the river Mulde and was used for several decades as a waste water channel of the chemical industry. During flood events the Spittelgraben covers the wetlands with its highly contaminated water. The sites Spittel and Keller are situated only a few metres from the river bank. They have often been flooded by the Spittelgraben and the sandy soils of this riverine area are very polluted with heavy metals and organic contaminants [7,8]. The site Bobbau is further away from the edges of the Spittelgraben and the sandy soil shows low contamination. The loamy soil in Greppin is located in the floodplain of the river Mulde and is highly polluted with heavy metals and organic contaminants [7,8].

Four small lysimeters (five lysimeters for the Spittel site) with a surface soil area of 400 cm^2 and a depth of 25 cm were used for the extraction of soil solution. The lysimeters were obtained as undisturbed monoliths. Some soil properties are presented in Table 1. The experiments, including the simulation of several rainfall events, were carried out under greenhouse conditions. Irrigation of the lysimeters was ef-

Table 1
Soil properties

Site	pH	Clay + fine silt (%)	Carbon (%)	CEC (cmol$_c$ kg^{-1})[a]
Bobbau	6.1	6	2.7	10.1
Keller	4.4	13	24	32.6
Spittel	3.3	8	7.6	33.1
Greppin	6.4	22	10.5	25.3

[a] Cation exchange capacity, measured in cmol$_{charge}$ per kg of soil.

fected with distilled water. After completed percolation through the soil monoliths the percolates were collected in glass bottles, filtered through 0.45-μm filters and stored at $-18°C$ until analysed. Altogether 24 percolates per site (34 for the Spittel site) were analysed for HCHs.

3.2. SPME procedure and chromatographic conditions

The studies were carried out with an SPME device from Supelco (100-μm polydimethylsiloxane solid-phase microextraction fibre assembly).

The SPME procedure is very simple. First the fibre is withdrawn into the needle of the syringe and the syringe is used to penetrate the septum of the sample vial. Then the fibre is inserted in the aqueous phase. When equilibrium is reached, the fibre is again withdrawn into the needle and the syringe needle is removed from the vial. The last step is the thermal desorption of the analytes in the injector of the gas chromatograph.

The gas chromatograph used was a Chrompack CP 9000 device with ECD. For the experiments a 25 m × 0.32 mm I.D. Ultra 1 capillary column (Hewlett-Packard) with a 0.52-μm film thickness was used. The carrier gas and the make-up gas were nitrogen. A split–splitless injector was used in the splitless mode and maintained at 200°C. after the optimization experiments a 2.0-min desorption time was chosen. The column temperature programme was as follows: initial temperature 60°C (held for 3 min), increased at 10°C/min to 250°C and held at the final temperature for 20 min. The detector temperature was 250°C.

For the identification of the HCHs some GC–MS measurements were performed. A combina-

tion of an HP 5890II gas chromatograph and a mass spectrometer (Hewlett-Packard) was used. A capillary column of 0.25 mm I.D. and a 0.25-μm film thickness were employed. The carrier gas was helium 6.0 (99.9999% pure). A split–splitless injector in the splitless mode was used and the temperature programme was the same as described for the GC–ECD measurements. The transfer line and mass spectrometer were held at 280°C. The mass spectrometer worked in the single-ion monitoring (SIM) mode.

For calibration and optimization of the SPME, doubly distilled water was spiked with a mixture of α-, β-, γ- and δ-HCH dissolved in methanol.

4. Results and discussion

4.1. Optimization of SPME

The theoretical considerations show the necessity to optimize carefully the mixing process, the exposure time of the fibre in an aqueous sample and the desorption time (exposure time of the fibre in the GC injection port). Optimization also includes the determination of the linear dynamic range of the SPME procedure in connection with the chromatographic detector used.

The first step is to examine the time required for the HCHs to reach equilibrium depending on the turning speed of the stirrer. Fig. 1 shows the peak area of the ECD signal versus the exposure time for lindane. Without stirring (A) the equilibrium is reached after a time $\gg$ 60 min. Using a magnetic stirrer at 250 turns/min (B) the equilibrium time is reached within 40–60 min. An increase in the stirring speed up to 1000 turns/min (C) ensures that the aqueous sample is
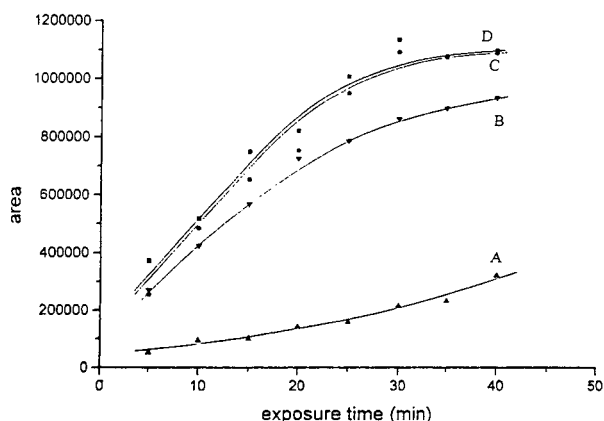
Fig. 1. Exposure time profiles for γ-HCH as a function of the stirring conditions. Concentration of γ-HCH, 0.6 ng/ml water. (A) Without stirring; (B) stirring speed 250 turns/min; (C) stirring speed 1000 turns/min, central position; (D) stirring speed 1000 turns/min, non-central position.
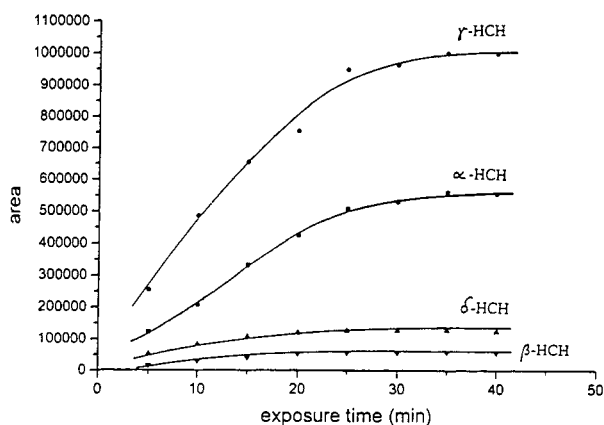
Fig. 2. Exposure time profiles for α-, β-, γ- and δ-HCH under optimum stirring conditions. Concentration of α-, β- and δ-HCH, 0.3 ng/ml water; concentration of γ-HCH, 0.6 ng/ml water.

nearly perfectly agitated and equilibrium is reached within 20–30 min. In these cases the fibre was placed in the middle of the 4-ml vial used and therefore also in the middle of the rotation paraboloid formed by magnetic stirring. In a further case (D) the fibre was exposed at a distance 6 mm from the centre. That means the fibre was placed in a region of violent agitation of the sample. The equilibrium time was also 20–30 min and the results in experiments C and D were nearly identical. Because of the better reproducibility, the central position (C) was chosen.

Fig. 2 shows the equilibrium time profiles for α-, β-, γ- and δ-HCH under optimum conditions (stirring speed 1000 turns/min). Because the exposure time profiles under constant conditions (speed of the stirrer) are well reproducible, it is possible to choose exposure times lower than the equilibrium time. The only disadvantage is the decrease in sensitivity. For the measurement of a large number of soil solutions with relatively high HCH concentrations an exposure time of 10 min and a optimum stirring speed of 1000 turns/min were chosen.

The next step is to ensure that the exposure time of the fibre within the GC injector is long enough to desorb the compounds completely from the silicon phase. Fig. 3 shows desorption

time profiles of the HCHs for an injector temperature of 200°C. The desorption is nearly completely after 0.5 min and a desorption time of 2 min was chosen for the soil solution measurements.

Depending on the partition coefficient of an analyte, the sample can be significantly depleted in a single extraction. Fig. 4 shows the decrease in the concentration of analytes in a 4-ml vial as a function of the number of extractions. Because of this decrease in concentration, for each 4-ml sample only one extraction was performed.
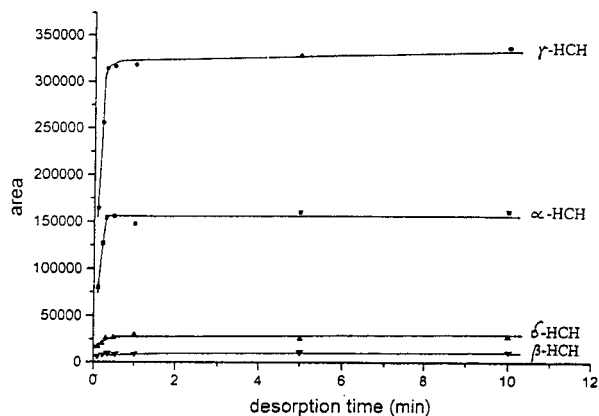
Fig. 3. Desorption time profiles for α-, β-, γ- and δ-HCH. Concentration of α-, β- and δ-HCH, 0.3 ng/ml water; concentration of γ-HCH, 0.6 ng/ml water.
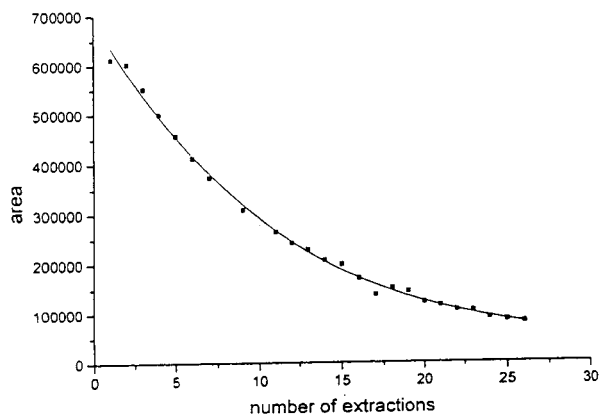
Fig. 4. Dependence of the ECD signal for γ-HCH on the number of extractions. Initial γ-HCH concentration, 0.6 ng/ml water.

After establishing the exposure time and the desorption time, the linear dynamic range of the detectors coupled with the SPME procedure was investigated. In Fig. 5 the ECD signal versus the concentration of the HCHs in doubly distilled water is shown. For α-, γ- and δ-HCH the linear dynamic range exceeds more than three orders of magnitude, and for β-HCH the linear range extends to more than 300.

MS was applied in the SIM mode with the characteristic ions at *m/z* 109, 183, and 219. The linearity of the SPME–GC–MS procedure (Fig. 6) is between two and three orders of magnitude.
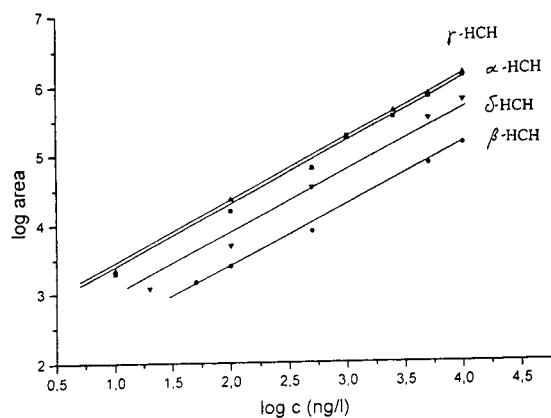
Fig. 5. Linearity of the SPME–ECD method for α-, β-, γ- and δ-HCH.
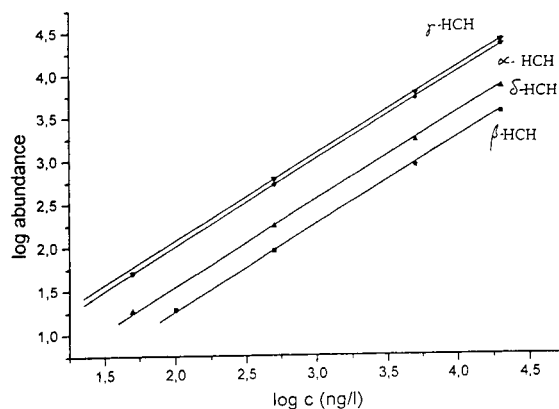
Fig. 6. Linearity of the SPME–MS method for α-, β-, γ- and δ-HCH.

A peak with a signal-to-noise ratio of 3 was defined as the detection limit. Under these conditions the detection limits in Table 2 were calculated.

4.2. Analysis of soil solutions

After optimization, the SPME–GC–ECD procedure was used for the determination of α-, β- and γ-HCH in soil solution. GC–MS combination was used when the identification of the HCHs was difficult. Fig. 7 shows the ECD results for a Spittel site sample with a dominant β-HCH peak (4.28 μg/l), an α-HCH concentration of 0.12 μg/l and a γ-HCH concentration of 0.03 μg/l. The peaks following that of γ-HCH are caused by other chlorinated compounds which are superimposed on the δ-HCH peak, but the chromatograms obtained using MS in the SIM mode showed that the δ-HCH concentrations in the samples are neglectible. Fig. 8 shows

Table 2
Detection limits with the SPME–GC combinations

Compound	SPME–GC–ECD (ng/l)	SPME–GC–MS (ng/l)
α-HCH	5	12
β-HCH	32	80
γ-HCH	5	13
δ-HCH	12	40

Fig. 7. Chromatogram of a soil solution obtained using ECD. Numbers at peaks are retention times in min.

that none of the three peaks with retention times of 18.62, 18.77 and 18.89 min corresponds to δ-HCH, because neither the retention time (18.84 min) nor the ion intensity ratio relate to this compound.

The concentration distributions of α-, β- and γ-HCH in Bobbau, Keller, Spittel and Greppin samples (see Fig. 9) were determined on 106 samples of the soil solutions. To ensure that the external calibration (addition of HCHs to doubly distilled water) is transferable to the soil solutions, some of the contaminated samples were

Fig. 8. Chromatogram of a soil solution obtained using MS. Time in min.

Fig. 9. Concentration of HCHs in soil solutions from different sites.

spiked with known α-, β- and γ-HCH concentrations.

The increase in the peak areas was comparable to the peak areas obtained after adding the HCHs to doubly distilled water. This experiment demonstrated that we could assume that the soil solution matrix in this case would not affect the calibration and that we could transfer the calibration graphs obtained for clean water.

Fig. 9 demonstrates the dependence of the HCH concentration on the distance of the site from the river bank and consequently on the frequency of flood events in the riverine area of the Spittelgraben with high HCH concentrations in the Keller and Spittel and low concentrations in the Bobbau sites. Nevertheless, the β-HCH concentration in the probably low-contaminated soil from the Bobbau site exceeds the drinking water limit for German drinking water supplies (0.1 μg/l) for chlorinated pesticides. At the Greppin site, which is located in the floodplain of the river Mulde, the HCH concentrations in the percolates are less than those at the Keller and Spittel sites. The dominance of β-HCH within the HCHs and its high concentration and also the low concentrations of γ-HCH found in the percolates cann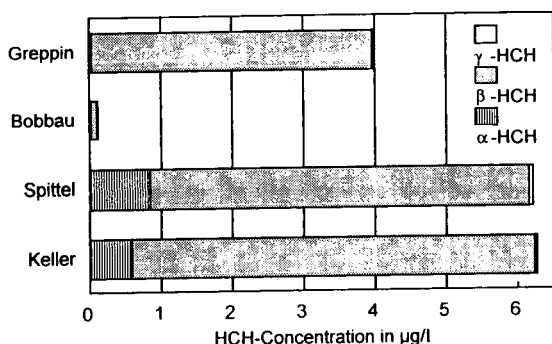ot be explained by the solubility of the HCHs. γ-HCH is the most soluble isomer and β-HCH is only slightly soluble in water [9]. However, the degradation of α- and γ-HCH is faster than the degradation of β-HCH [10,11]. Probably relative enrichment of β-HCH took place in the soils in these riverine areas.

Additionally, an influence of the dissolved organic matter in the soil solution on the β-HCH concentration is assumed. The dissolved organic matter is able to increase the solubility of hydrophobic compounds and therefore its concentration in the soil solution [12,13]. Further, α- and β-HCH are by-products of lindane production, so substantial soil contamination with γ-HCH cannot be expected. The differences in the β-HCH concentration in the percolates between the Spittel and Keller sites on the one hand and the Greppin site on the other may also result from the influence of soil properties on the mobility of β-HCH. At the Greppin site the soil texture is heavier and the pH higher than at the Spittel and Keller sites, but the Spittel and Keller soils contain more organic matter than the Greppin soil. The high mobility and long persistence of β-HCH become clear on comparing the proportions of β-HCH in the total HCH concentration between the soil and the soil solution. Borsdorf et al. [14] determined for some soils in the floodplain of the river Mulde a proportion of β-HCH in the total HCH concentration varying between 52% and 72%. This proportion as determined for the soil solutions of the four sites discussed in this paper varies between 80% and 99%. This is an indication of the highest mobility and the longest persistence of β-HCH among the other HCHs, although β-HCH shows the lowest water solubility. Additional experiments are necessary to elucidate the influence of the soil and the components of the soil solution on the mobility and persistence of β-HCH.

References

[1] D.W. Potter and J. Pawliszyn, *J. Chromatogr.*, 625 (1992) 247–255.

[2] C.L. Arthur, D.W. Potter, K.D. Buchholz, S. Motlagh and J. Pawliszyn, *LG · GC*, 10 (1992) 656–661.

[3] C.L. Arthur, K. Pratt, S. Motlagh, J. Pawliszyn and R.P. Belardi, *J. High Resolut. Chromatogr.*, 15 (1992) 741–744.

[4] D. Louch, S. Motlagh and J. Pawliszyn, *Anal. Chem.*, 64 (1992) 1187–1199.

[5] C.L. Arthur, L.M. Killam, K.D. Buchholz, J. Pawliszyn and J.R. Berg, *Anal. Chem.*, 65 (1992) 1960–1966.

[6] K.D. Buchholz and J. Pawliszyn, *Environ. Sci. Technol.*, 27 (1993) 2844–2848.

[7] M. Lauer, T. Heymann and C. Schneider, *Schadstoffe Umwelt*, 10 (1992) 163–170.

[8] R.W. Scholz, N. Nothbaum, T.W. May, R. Brockmann, H. Bode, K.-H. Deubel and U. Hippe, *Schadstoffe Umwelt*, 10 (1992) 171–179.

[9] D. Eichler, *Materialien zur DFG-Veranstaltung 28./ 29.11.79 und 6.3.1980*, Deutsche Forschungsgemeinschaft, Weinheim, 1979–80, pp. 14–17.

[10] G. Jagnow, K. Haider and P.C. Ellwardt, *Arch. Microbiol.*, 115 (1977) 285–292.

[11] K. Haider, *Z. Naturforsch.*, 34 (1979) 1066–1069.

[12] L.L. Henry and I.H. Suffet, *Adv. Chem. Ser.*, 219 (1989) 159–171.

[13] B.J. Eadie, N.R. Morehead, J.V. Klump and P.F. Landrum, *J. Great Lakes Res.*, 18 (1992) 91–97.

[14] H. Borsdorf, C. Opp and J. Stach, *Chem. Tech. (Leipzig)*, 6 (1993) 467–474.

ELSEVIER

Determination of carbohydrates, sugar acids and alditols by capillary electrophoresis and electrochemical detection at a copper electrode

Jiannong Ye[1], Richard P. Baldwin*

Department of Chemistry, University of Louisville, Louisville, KY 40292, USA

Abstract

Capillary electrophoresis combined with electrochemical detection at copper electrodes has been shown to provide a simple and sensitive method for the direct analysis of samples containing a wide range of carbohydrate compounds including simple sugars, sugar acids and alditols. In this approach, both the separation and the detection required the use of a strongly alkaline medium whose hydroxide content could be varied to optimize conveniently the migration times and resolution obtained. Detection consisted of a direct oxidation that required no derivatization and yielded detection limits at or below the fmol level for most of the carbohydrate species.

1. Introduction

Because carbohydrates are the most abundant class of organic compounds in nature and are found universally distributed among plants, animals and microorganisms, the development of analytical methods for their determination is an increasingly important research area. However, despite extensive efforts, carbohydrate analysis still presents a challenge, especially for complex, natural samples. First, resolution of the many closely related carbohydrate compounds possible in a sample poses a problem that is difficult, if not impossible, for most applicable separation techniques. To date, high-performance liquid chromatography (HPLC) has most commonly

been employed for this purpose. Second, carbohydrate compounds generally do not possess chromophores which absorb appreciably at accessible UV–visible wavelengths and therefore are not readily detected by conventional spectrophotometric methods unless rigorous and time-consuming derivatization is first carried out. Until recently, electrochemical detection (ED) played only a very limited role in carbohydrate detection. This is due to the fact that, although carbohydrates can be oxidized by numerous chemical reagents, they are not normally electroactive at carbon electrodes, which are the most commonly used working electrodes in ED. Over the past decade, oxidation at platinum and gold electrodes has gained some popularity for the detection of underivatized carbohydrates [1,2]. However, because of electrode fouling caused by adsorption of the sugar oxidation products, the applied potential for these elec-

---

* Corresponding author.
[1] Present address: Department of Chemistry, East China Normal University, Shanghai 200062, China.

trodes must be continuously pulsed for stable, long-term response to be possible.

Very recently, two developments have occurred which have provided attractive new tools for carbohydrate determination. First, capillary electrophoresis (CE) has become firmly established as a separation technique which, compared to HPLC, offers extremely high separation efficiency [3]. For example, for a 50 cm long capillary tube, the number of theoretical plates typically reaches 100 000 or more, which is about 10 times higher than that for a typical 25 cm long HPLC column. As a result, much higher resolution and much greater separating power for carbohydrate samples is potentially available with CE. Second, investigations using Cu electrodes [4–10] have shown that carbohydrates can be oxidized facilely at these surfaces at constant applied potential both in bulk solution and in HPLC detection schemes. This permits less expensive, less complicated, but still highly sensitive detection of these species. Moreover, it has been found that amperometric detection at Cu electrodes can be employed to detect not only carbohydrates themselves but also many related compounds including alditols and aldonic, aldaric and uronic acids [4]. Therefore, CE's high efficiency coupled with convenient, sensitive and dependable detection at Cu electrodes would seem to represent an ideal analytical tool for the determination of carbohydrates and their derivatives. Although the Cu electrode detection systems always require strongly alkaline conditions —which drastically limits the stationary phase options when used as HPLC detectors— high pH levels are perfectly compatible with the fused-silica capillaries used in CE.

To this point, only a few reports describing the use of CE–ED with metallic electrodes for the separation and determination of carbohydrate compounds have been published. In these studies, both pulsed amperometric detection at Au electrodes [11] and constant potential detection at Cu electrodes [12,13] have been successfully employed for sensing purposes following the CE separation. However, the analytes included consisted primarily of simple mono- and disaccharides, with little attention paid to im-

portant derivatives such as alditols and acidic sugars. In view of the varying charges exhibited by all these compounds under the high pH conditions required for oxidation at the metallic electrodes, CE–ED appeared to present a logical and attractive analysis approach for samples containing not only simple sugars but also many of their important derivatives. In an attempt to realize this possibility and thereby expand the analytical tools available for carbohydrate-related compounds in complex samples, we report here the capabilities of CE–ED with Cu electrodes for the determination of alditols and aldonic, uronic and aldaric acids as well as ordinary sugars. Such analyses, which often present problems when the separation is carried out by liquid chromatography [6,14–17], were in fact able to be accomplished in an extremely straightforward manner by a simple CE–ED procedure.

2. Experimental

2.1. Reagents

Stock solutions of all sugars, alditols and sugar acids (purchased from Sigma and Aldrich) were prepared fresh daily in deionized water. Just prior to use, samples to be injected were prepared from these stock solutions by dilution to the desired concentration with the NaOH solution used as separation electrolyte. Experiments involving enzymatic oxidation of glucose were conducted by adding 1.0 mg of glucose oxidase (Type X, purchased from Sigma) to a stirred pH 5.3 phosphate buffer solution (200 ml, room temperature) containing 1.0 mM β-D-glucose and 0.50 mM glucitol (which was used as an internal reference).

2.2. Apparatus

All CE experiments were performed on a laboratory-built instrument with a 30 kV high-voltage power supply (Model 30B, Hipotronics, Millerton, NY, USA) and an 80 cm length of 25 μm I.D. × 360 μm O.D. fused-silica capillary

(Polymicro Technologies, Phoenix, AZ, USA). In order to protect the operator from accidental exposure to high voltages, the entire capillary, the electrolyte reservoirs, and all electrodes were enclosed in a Plexiglass box equipped with a safety switch wired to shut down the power supply whenever the box was opened. In addition, the outlet end of the capillary was always maintained at ground. A 1-in.-diameter (1 in. = 2.54 cm) plastic vial served both as the cathode compartment of the CE instrument and as the electrochemical cell for the CE detection. Before insertion into the vial through a small slot cut into its side, the outlet and of the capillary was made as flat as possible by use of a fiber cleaver (Newport, Irvine, CA, USA). The electrophoresis medium was always just an NaOH solution; NaOH concentrations used ranged from 25 to 250 mM. To minimize the effect of CO_2 pickup from the atmosphere, NaOH solutions were replaced daily. Sample injection was carried out by electromigration by immersing the inlet of the capillary in the sample solution and applying a high voltage for a suitable time period.

The design and performance of the ED system employed have been described previously [13]. The specific working electrode used was a 100-μm-diameter Cu magnet wire (Newark Electronics, Chicago, IL, USA) whose side areas were covered with a non-conductive coating. The procedure for construction of these working electrodes was detailed previously [13]. The working electrode was arranged in a wall-jet configuration in which the Cu wire was inserted into the cathode compartment of the CE instrument and then positioned up against the capillary outlet with the help of an Oriel (Stratford, CT, USA) Model 14901 micropositioner. In this configuration, the CE effluent impinged directly onto the disk-shaped Cu electrode and then flowed radially outward across its surface. Also placed into this compartment were an Ag/AgCl (3 M NaCl) reference electrode and a platinum wire counter electrode. Control of the applied potential and measurement of the resulting current were carried out with a Bioanalytical Systems (West Lafayette, IN, USA) Model LC-4B amperometric detector.

3. Results and discussion

3.1. Nature of the CE separation

The electrochemical behavior of carbohydrates at metallic Cu electrodes in strongly alkaline solution has been reported previously [4,5,18,19]. The range of carbohydrate compounds that can be oxidized usefully at the Cu electrode has been shown to include mono- and oligosaccharides, both reducing and non-reducing, and many related compounds such as alditols, aldonic acids, uronic acids and aldaric acids, with the primary molecular feature needed for oxidative response being the presence of multiple aliphatic hydroxyl groups [19]. Of course, if CE–ED is to be effective for the identification and quantitation of any of these compounds in authentic samples, the CE system employed must be effective in separating them from one another and from other oxidizable species present in the sample matrix. Fortunately, CE's extremely high separation power can usually meet this requirement. In the work presented here, this was accomplished with a strongly basic separation medium (25–250 mM NaOH) which served both to maintain the activity of the Cu electrode for ED detection and to maintain all the analytes in anionic form for the differential migration required for CE separation.

The electropherograms shown in Fig. 1 for sample mixtures containing glucose and galactose as well as their respective alditol and acidic sugar derivatives illustrate the nature of the separation that can be achieved by means of CE. In both cases, the order of elution observed (alditol, aldose, aldonic acid, uronic acid, aldaric acid) was exactly that expected on the basis of the compounds' pK_a values and resulting ionic character in the 50 mM NaOH electrophoresis medium employed. Because the polarity of the voltage applied across the capillary was set so that the detector end was negative compared to the injection end, cationic species migrated toward, and anions away from, the capillary exit. Thus, it was expected that the nearly uncharged alditols (pK_a values ca. 13.6) would elute first
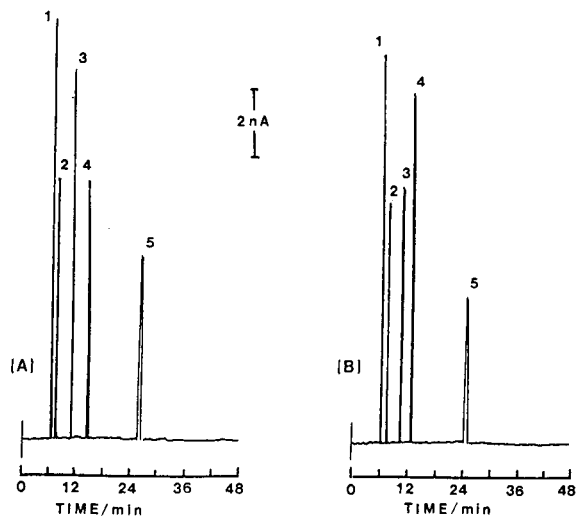
Fig. 1. Electropherograms of glucose (A) and galactose (B) families. Labeled peaks in (A) correspond to glucitol (1, 250 μM), glucose (2, 250 μM), gluconic acid (3, 500 μM), glucuronic acid (4, 1 mM) and glucaric acid (5, 2 mM). Labeled peaks in (B) correspond to galactitol (1, 250 μM), galactose (2, 250 μM), galactonic acid (3, 500 μM), galacturonic acid (4, 1 mM) and galactaric acid (5, 2 mM). Electrophoresis medium: 50 mM NaOH; working electrode: 100-μm Cu disk at + 0.60 V vs. Ag/AgCl; separation voltage: 30 kV; injection by electromigration (30 kV for 3 s).

and the di-anionic aldaric acids should take the longest time. The specific migration times seen for the glucose family as well as the average charges calculated for these pH conditions are shown in Table 1.

An interesting aspect of the CE separations in

Fig. 1 is that, because of the high efficiency of the CE approach and the different charge character of the various carbohydrate derivatives, the different classes of compounds were relatively easy to resolve from one another. This is not always the case with HPLC-based methods [6,14–17]. A more challenging analytical task involves the resolution of individual members of each class from one another —e.g., the determination of one aldonic acid in the presence of another. Significantly, it appears that the CE–ED approach presents very useful capabilities for this application as well. Shown in Fig. 2, for instance, is a single electropherogram obtained for a sample mixture containing all ten glucose and galactose species from Fig. 1. Exactly the same CE conditions were employed for this separation as earlier except that the separation voltage was decreased from 30 to 15 kV in order to facilitate the resolution of glucitol and galactitol. Apart from these two isomers whose pK_a values differ by only about 0.1 [20], the other sample components were comparatively easy to resolve completely. Of course, the decrease in CE potential had the expected effect of lengthening the time required for the separation.

An alternative approach, which offers the possibility of fine-tuning the separation further, consists of adjusting the pH of the electrophoresis medium so as to adjust the net charge on the very weakly acidic hydroxyl groups. Increasing the pH can be of special utility, for example, for alditols and simple sugars to increase the

Table 1
Migration times and estimated charges for glucose derivatives

	Glucitol	Glucose	Gluconic acid	Glucuronic acid	Glucaric acid
Migration time (min)	6.6	8.1	10.8	12.9	26.4
pK_a [a]	13.6	12.3	3.86	–[c]	3.77, 6.08
Molecular mass (g/mol)	182.2	180.2	196.2	194.1	212.2
Charge[b]	−0.11	−0.65	−1.0	−1.65[c]	−2.0

[a] From Refs. [20–22].
[b] Charges were calculated from the compounds' pK_a values listed above for the pH 12.7 electrophoresis medium.
[c] The pK_a values for glucuronic acid are not available; calculation is based on the assumption that, because of the structural similarities, these values are similar to those of gluconic acid (3.8) and glucose (12.3).
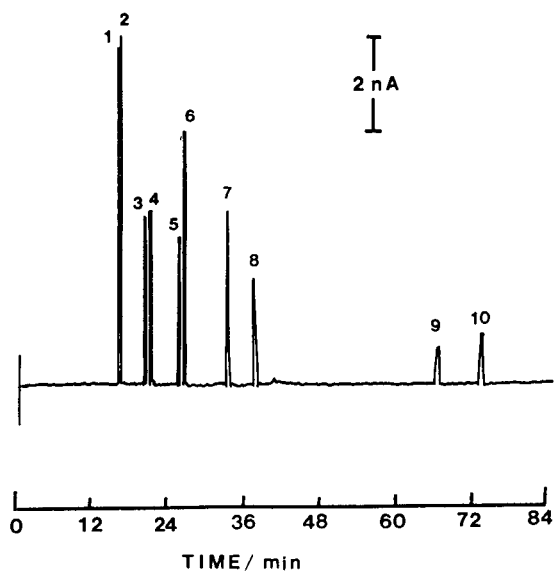
J. Ye, R.P. Baldwin / J. Chromatogr. A 687 (1994) 141–148

145

Fig. 2. Electropherogram of a mixture of glucose and galactose families. Labeled peaks correspond to galactitol (1, 125 μM), glucitol (2, 125 μM), galactose (3, 125 μM), glucose (4, 125 μM), galactonic acid (5, 250 μM), gluconic acid (6, 250 μM), galacturonic acid (7, 500 μM), glucuronic acid (8, 500 μM), galactaric acid (9, 1 mM) and glucaric acid (10, 1 mM). Electrophoresis medium: 50 mM NaOH; working electrode: 100-μm Cu disk at +0.60 V vs. Ag/AgCl; separation voltage: 15 kV; injection by electromigration (15 kV for 5 s).

Fig. 3. Electropherogram of alditol mixture. Labeled peaks correspond to ethylene glycol (1, 1 mM), glycerol (2, 500 μM), inositol (3, 160 μM), erythritol (4, 250 μM), galactitol (5, 250 μM), adonitol (6, 250 μM), glucitol (7, 250 μM) and mannitol (8, 250 μM). Electrophoresis media: (a) 25 mM NaOH, (b) 0.10 M NaOH, (c) 0.20 M NaOH and (d) 0.25 M NaOH. Working electrode: 100-μm Cu disk at +0.60 V vs. Ag/AgCl; separation voltage: 15 kV; injection by electromigration (15 kV for 3 s).

migration times and improve the separations obtained for these compounds. This is illustrated clearly in Fig. 3 which shows the electropherograms obtained for a mixture of eight different alditols at four different NaOH concentrations, all other CE conditions remaining the same. While no separation at all occurred when the electrophoresis medium was 25 mM NaOH (Fig. 3a), the separation improved as the OH⁻ level was increased, with complete resolution of the mixture achieved at NaOH concentrations of 250 mM and above (Fig. 3d). Because the pK_a values for alditols are typically 13–14, highly basic conditions are required for significant dissociation and differential electrophoretic migration to occur. Otherwise, the alditols are nearly neutral in charge and, due to the electroosmotic flow, emerge from the capillary at practically the same time. Of course, because of their greater

anionic character under the high NaOH conditions of Fig. 3d, simple sugars and the acidic sugars remained on the capillary much longer than the alditols and would therefore present no problem for the alditol analysis.

3.2. Analytical applications

The analytical results for a test mixture containing all ten members of the glucose and

Table 2
Analytical performance of Cu electrode CE–ED for glucose and galactose compounds

Peak	Compound	Detection limit (fmol)	Linear range	No. of theoretical plates	R.S.D. (%)[a]
1	Galactitol	0.5	$0.5\ \mu M$–$1\ mM$	145 000	4.9
2	Glucitol	0.5	$0.5\ \mu M$–$1\ mM$	147 000	5.8
3	Galactose	1	$1\ \mu M$–$1\ mM$	154 000	5.5
4	Glucose	1	$1\ \mu M$–$1\ mM$	166 000	5.2
5	Galactonic acid	2	$2\ \mu M$–$1\ mM$	151 000	8.2
6	Gluconic acid	1.5	$1.5\ \mu M$–$1\ mM$	155 000	6.0
7	Galacturonic acid	3	$3\ \mu M$–$2\ mM$	150 000	7.7
8	Glucuronic acid	4	$4\ \mu M$–$2\ mM$	180 000	8.9
9	Galactaric acid	8	$8\ \mu M$–$4\ mM$	71 000	7.0
10	Glucaric acid	6	$6\ \mu M$–$4\ mM$	81 000	7.8

CE–ED conditions as in Fig. 2.
[a] These numbers represent the relative standard deviations for four identical injections.

galactose families are summarized in Table 2, and analogous results for a sample containing only alditols are shown in Table 3. For most of these compounds, peak heights varied linearly over a roughly three-order-of-magnitude range with detection limits (signal-to-noise ratio 3) near or below the fmol level. The reproducibility and stability of the CE–ED system were quite reasonable in practice. Repeated injections of the same analyte typically exhibited relative standard deviations in the 5–7% range —which was essentially the reproducibility of the electrokinetic injection process employed here. The Cu electrode surfaces were generally used for periods of a month or more without replacement or any special treatment. Over the course of two weeks of continuous use in the CE–ED apparatus, only a gradual decrease in electrode response (less than 10%) was observed. Finally, separation efficiencies were generally greater than 150 000 theoretical plates; the exceptions to this were galactaric and glucaric acids which had very long migration times under the particular CE conditions in effect.

These results are noteworthy in one additional respect. The ED system employed here was based on a wall-jet configuration [13] which allows the use of much larger electrodes than has normally been the case for CE–ED. This makes electrode construction and alignment a compara-

Table 3
Analytical performance of Cu electrode CE–ED for alditols

Peak	Alditol	Detection limit (fmol)	Linear range	No. of theoretical plates	R.S.D. (%)[a]
1	Ethylene glycol	19.2	$16\ \mu M$–$4\ mM$	148 000	4.6
2	Glycerol	2.4	$2\ \mu M$–$2\ mM$	166 000	10.2
3	Inositol	0.38	$0.32\ \mu M$–$0.64\ mM$	158 000	5.5
4	Erythritol	0.6	$0.5\ \mu M$–$1\ mM$	164 000	6.7
5	Galactitol	0.6	$0.1\ \mu M$–$1\ mM$	190 000	5.0
6	Adonitol	0.6	$0.5\ \mu M$–$1\ mM$	209 000	6.5
7	Glucitol	3	$0.5\ \mu M$–$1\ mM$	215 000	6.3
8	Mannitol	4	$0.5\ \mu M$–$1mM$	180 000	5.9

CE–ED conditions as in Fig. 3d.
[a] These numbers represent the relative standard deviations for six identical injections.

tively simple process and greatly enhances the stability and reproducibility of the detector response. As a result, the CE–ED method described is not difficult to implement technically. This characteristic should make the approach attractive for many applications involving the analysis of carbohydrates and related compounds in a wide variety of complex natural samples. Two specific examples are illustrated below.

Fig. 4 shows a typical electropherogram obtained at a Cu electrode for a commercial apple juice (Carolina Gold, 100% natural). The only sample pretreatment prior to injection was dilution by a factor of 300 with the running electrolyte solution. The electropherogram exhibits four well-separated peaks whose migration times match those of glucitol, sucrose, glucose and fructose standards. Based on the calibration curves obtained for glucitol and glucose, the contents of these two compounds in the original apple juice were estimated to be 1.55 ± 0.05 and 12.36 ± 0.62 mg/ml, respectively. The separation obtained here was much more efficient than that reported by HPLC [14] where the same components were observed for apple juice but the resulting peaks were not able to be resolved completely.
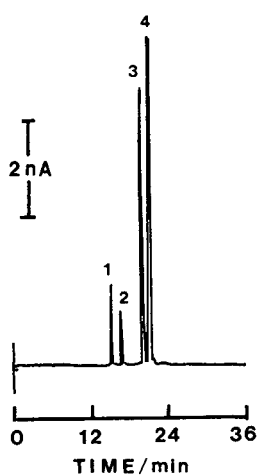
It is well known that glucose oxidase (GOx)

catalyzes the oxidation of β-D-glucose to form gluconic acid and hydrogen peroxide. Because both glucose and gluconic acid peaks can be monitored very effectively by the Cu electrode CE–ED approach, the enzymatic process can thus be followed conveniently. Fig. 5 shows three different stages of such a GOx oxidation. In Fig. 5a, the enzyme had not yet been added to the system which consisted of 200 ml of pH 5.3 phosphate buffer containing 0.5 mM glucitol (used as an internal standard) and 1.0 mM glucose. Therefore, as expected, both glucitol (peak 1, 6.6 min) and glucose (peak 2, 8.2 min) were readily detected; and no other peaks were seen. Fig. 5b shows the electropherogram obtained for a sample after GOx had been added to the system for 5 min. As expected, the glucitol peak remained unchanged, while that for glucose decreased. In the meantime, two new peaks (peaks 3 and 4), resulting from the formation of the gluconic acid and hydrogen peroxide products of the glucose oxidation, appeared at 10.9 and 24.0 min, respectively. After 25 min, the resulting electropherogram (Fig. 5c) shows the glucitol peak to remain unchanged, that for glucose to become even smaller, and those for gluconic acid and hydrogen peroxide to be larger. After about 90 min, when the glucose peak had vanished completely, the enzymatic oxidation was obviously complete. Such simultaneous peak-height monitoring of glucose as well

Fig. 4. Electropherogram of apple juice. Peaks: 1 = glucitol; 2 = sucrose; 3 = glucose; 4 = fructose. CE–ED conditions as in Fig. 2.
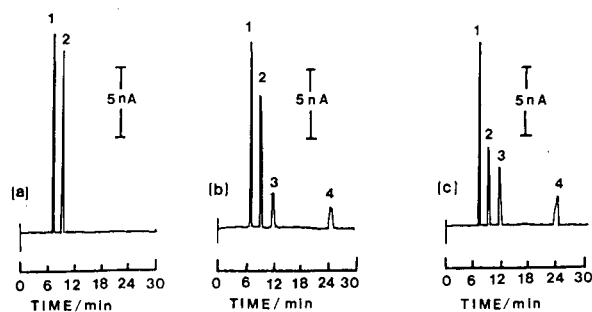

Fig. 5. Electropherograms obtained during different stages of enzymatic oxidation of D-glucose. Peaks: 1 = glucitol; 2 = glucose; 3 = gluconic acid; 4 = hydrogen peroxide. Electropherograms a, b and c correspond to samples taken 0, 5 and 25 min after adding glucose oxidase, respectively. CE–ED conditions as in Fig. 1.

as its enzymatic oxidation products could provide useful information in studying the kinetics of GOx-catalyzed reactions and, more importantly, in process monitoring and control applications. Extension of the approach to other enzyme systems involving carbohydrate metabolism should be straightforward.

Acknowledgement

References

[1] S. Hughes and D.C. Johnson, *J. Agric. Food Chem.*, 149 (1983) 1.
[2] D.C. Johnson and W.R. LaCourse, *Anal. Chem.*, 62 (1990) 589A.
[3] R. Weinberger, *Practical Capillary Electrophoresis*, Academic Press, San Diego, CA, 1993.
[4] S.V. Prabhu and R.P. Baldwin, *Anal. Chem.*, 61 (1989) 2258.
[5] P. Luo, S.V. Prabhu and R.P. Baldwin, *Anal. Chem.*, 62 (1990) 752.
[6] S.V. Prabhu and R.P. Baldwin, *J. Chromatogr.*, 503 (1990) 227.
[7] J.M. Zadeii, J. Marioli and T. Kuwana, *Anal. Chem.*, 63 (1991) 649.
[8] Y. Xie and C.O. Huber, *Anal. Chem.*, 63 (1991) 1714.
[9] S. Mannino, M. Rossi and S. Ratti, *Electroanalysis*, 3 (1991) 711.
[10] P. Luo, M.Z. Luo and R.P. Baldwin, *J. Chem. Educ.*, 70 (1993) 679.
[11] T.J. O'Shea, S.M. Lunte and W.R. LaCourse, *Anal. Chem.*, 65 (1993) 948.
[12] L.A. Colon, R. Dadoo and R.N. Zare, *Anal. Chem.*, 65 (1993) 476.
[13] J. Ye and R.P. Baldwin, *Anal. Chem.*, 65 (1993) 3525.
[14] R. Schwarzenbach, *J. Chromatogr.*, 140 (1977) 304.
[15] S. Honda, M. Takahashi, S. Shimada, K. Kakehi and S. Ganno, *Anal. Biochem.*, 128 (1983) 429.
[16] S. Honda, *Anal. Biochem.*, 140 (1984) 1.
[17] A.M. Tolbert and R.P. Baldwin, *Electroanalysis*, 1 (1989) 389.
[18] J.M. Marioli and T. Kuwana, *Electrochim. Acta*, 37 (1992) 1187.
[19] M.Z. Luo and R.P. Baldwin, *J. Electroanal. Chem.*, in press.
[20] J.A. Dean (Editor), *Handbook of Organic Chemistry*, McGraw-Hill, New York, 1987, Section 8, pp. 22–57.
[21] J.A. Dean (Editor), *Lange's Handbook of Chemistry*, McGraw-Hill, New York, 10th ed., 1979, Ch. 5, pp. 17–41.
[22] M. Windholz (Editor), *The Merck Index*, 9th ed., Merck & Co., Rahway, NJ, 1976.

ELSEVIER

Separation of chlorophyll-c_1 and -c_2 by micellar electrokinetic capillary chromatography

Koichi Saitoh*, Hisashi Kato, Norio Teramae

Department of Chemistry, Faculty of Science, Tohoku University, Sendai, 980-77, Japan

Abstract

Successful separation of chlorophyll c_1 (Chl-c_1), chlorophyll c_2 (Chl-c_2) and their demetallated forms (pheoporphyrins) was performed by the electrokinetic chromatographic mode of capillary zone electrophoresis in a running solution containing micelles of sodium dodecyl sulfate and dimethylformamide. Chl-c_1 migrated with higher velocity than Chl-c_2, and their migration times were shorter than those of the corresponding demetallated forms. Complete resolution was obtained for these compounds within 20 min under the conditions of migration distance (i.e., effective capillary length) 50 cm and electric field 429 V/cm. Application to the separation of chlorophylls in a typical brown seaweed was demonstrated.

1. Introduction

Chlorophylls are typical instances of naturally occurring metal (magnesium) chelate complexes that are well known as photosynthetic pigments. Different chlorophylls possessing a chlorin or porphyrin macrocyclic structure are found in accordance with the species of plants. For example, chlorin-type compounds, such as chlorophyll *a* (Chl-*a*) and chlorophyll *b* (Chl-*b*), are found in green algae (*Chlorophyta*) and higher plants, and porphyrin-type compounds, such as Chl-c_1 and Chl-c_2, in brown algae (*Phaeophyta*).

The determination of chlorophylls and their degradation products in natural samples from the sea, rivers, lakes and other sources gives considerable information about the biological activity in different environments. Traditional analytical methods for chlorophylls are based on spectrophotometry or fluorimetry, and high-performance liquid chromatography (HPLC) is today a more promising approach. HPLC is fairly effective for separation of Chl-*a* and Chl-*b* [1–3], but few such successful separations of Chl-c_1 and Chl-c_2 have been reported so far. Exceptional instances of the separation of Chl-c_1 and Chl-c_2 were developed by means of thin-layer chromatography [4] and conventional column liquid chromatography [5] by Jeffrey using a specially prepared polyethylene powder, and by HPLC using a commercially available octadecyl-bonded vinyl alcohol copolymer in this laboratory [6]. However, methods with a much higher resolving ability need to be developed for the accurate measurement of these chlorophylls and related compounds in real samples.

There is increasing interest in the high resolving ability of capillary zone electrophoresis (CZE) and also its expanded mode, micellar

* Corresponding author.

electrokinetic   capillary   chromatography
(MECC) [7]. The applicability of MECC to the
separation of metal chelates was previously con-
firmed for uncharged bis- and trisacetylacetonato
complexes of di- and trivalent metal ions, respec-
tively [8,9], and bioporphyrins, such as protopor-
phyrin and haematoporphyrin, in the free acid
forms and the complexed forms with copper(II)
and zinc(II) [10].

This paper describes the feasibility of applying
MECC to the separation of chlorophylls, such as
Chl-c_1 and Chl-c_2, and related compounds, in
particular. A solution containing sodium dodecyl
sulfate (SDS) micelles and dimethylformamide
(DMF) is applied as the running solution. The
effects of the composition of the running solution
on the migration velocity and resolution of these
chlorophylls are detailed.

2. Experimental

2.1. Chlorophyll-c

Algal pigments were extracted with acetone
from fresh brown seaweed Undaria pinnatifida
obtained from Onagawa Bay, Miyagi, Japan.
The acetone extract was shaken with both light
petroleum and a saturated aqueous solution of
sodium chloride to remove Chl-a and carotenes.
Chlorophyll c was transferred from the resulting
water-rich phase to ethyl acetate, and then
purified on an octadecyl-bonded silica gel col-
umn with methanol–water (90:10, v/v). The
coloured fraction of the eluate was shaken with
both ethyl acetate and a saturated aqueous
solution of sodium chloride, and the ethyl ace-
tate phase was passed through a cellulose column
with pure ethyl acetate. A mixture of Chl-c_1 and
Chl-c_2 (see Fig. 1) was obtained after removal of
the solvent from the coloured fraction of the
eluate. The mixture was resolved into Chl-c_1 and
Chl-c_2 by reversed-phase HPLC as detailed
previously [6].

Pheoporphyrin c_1 (Pheo-c_1) and pheopor-
phyrin c_2 (Pheo-c_2) were prepared by demetalla-
tion of Chl-c_1 and Chl-c_2, respectively, using 0.1
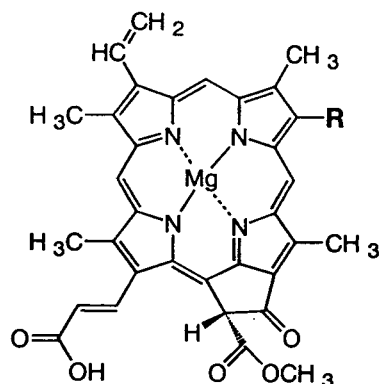M hydrochloric acid.

Fig. 1. Structure of Chl-c_1 (R = ethyl) and Chl-c_2 (R =
vinyl).

2.2. Electrophoresis

A Jasco (Tokyo, Japan) Model CE-800 CZE
system was used with a fused-silica capillary (70
cm × 50 μm I.D. × 375 μm O.D.) (Scientific
Glass Engineering, Ringwood, Australia). The
running solution filling the capillary was pre-
pared to contain a certain amount of SDS and
DMF in a solution adjusted to pH 11 with 20
mM N-cyclohexyl-3-aminopropanesulfonic acid
(CAPS) and 17 mM sodium hydroxide.

A methanol solution of chlorophyll(s) was
diluted with the same composition as the running
solution, and then was introduced into the posi-
tive end of the capillary by siphoning (10 cm, 5
s). Electrophoresis was run with an applied
voltage of 30 kV (electric field 429 V/cm) and at
25°C. Detection was performed by on-column
measurement of UV absorption at 430 nm
through a 50 μm × 0.75 mm slit located 20 cm
from the grounded end of the capillary.

3. Results

A simple buffer solution adjusted to pH 11
with CAPS and sodium hydroxide was not ap-
plicable to the medium (or running solution) for

the electrophoresis owing to the low solubility of the substances to be studied in the solution. The solubility should be enhanced by the addition of appropriate surfactant micelles and/or organic solvent to the buffer solution. In this work, SDS and DMF were taken as the former and the latter additives, respectively. SDS is popular as a surfactant in MECC. DMF was chosen from consideration of its high dissolving capability for chlorophylls, low viscosity, high dielectric constant and high miscibility with water. The solubilities of the chlorophylls and their demetallated forms of interest were sufficient for the following experiments in both solutions containing SDS micelles and DMF. DMF-containing micellar solutions were effective for performing successful MECC separations of porphyrins [10,11].

3.1. Formation of micelles of SDS

The formation of micelles of SDS in a DMF-containing solution was examined by conductimetry [10]. The critical micelle concentration (CMC) of SDS was estimated from the break of the proportionality between the specific conductivity change and the total concentration of SDS in the solution. The CMC values in different compositions of DMF-containing solution are given in Table 1.

Table 1
Critical micelle concentration (CMC) of SDS in DMF-containing solutions[a] at 25°C

DMF (%, v/v)	CMC (mM)
0	4.4
9.1	6.4
16.7	8.0
23.1	10.0
28.6	7.7
33.3	5.1
37.5	3.8

[a] Mixtures of DMF with an aqueous solution of 20 mM CAPS and 17 mM sodium hydroxide (pH 11).

3.2. Effects of SDS and DMF on electrophoretic migration

The migration behaviour of the chlorophyll compounds was studied in buffer solutions (pH 11) containing SDS and DMF at concentrations up to about 50 mM and 47% (v/v), respectively. The migration time of methanol was considered to be equal to the migration time of the electroosmotic flow of the solution (t_0).

When the running solution was a simple micellar solution of SDS (concentration 25 mM), all chlorophyll compounds, including Chl-c_1, Chl-c_2, Pheo-c_1 and Pheo-c_2, moved in the capillary and were detected at a migration time (t_S) identical with that of uncharged and hydrophobic compounds, such as Oil Yellow OB. The t_S value obtained for these compounds was larger than t_0. These results implied that all these compounds were completely solubilized in the negatively charged SDS micelles dispersed in the solution. Accordingly, the compounds could not be resolved on the basis of differential partitioning between the micelle pseudo-phase and bulk solution phase in the micellar solution.

The migration velocity of each chlorophyll compound varied with the addition of DMF to the running solution containing SDS micelles. When the running solution was a mixture of buffer solution (pH 11) and DMF containing 25 mM SDS, the t_S values of Chl-c_1, Chl-c_2, Pheo-c_1 and Pheo-c_2 increased with increasing proportion of DMF in the mixture, as shown in Fig. 2. A significant resolution was obtained for both the Chl-c_1–Chl-c_2 and the Pheo-c_1–Pheo-c_2 pairs when the DMF concentration was >20%. The t_0 value also increased with increasing DMF content. The electroosmotic velocity of a liquid in a capillary is, in general, directly proportional to the zeta potential on the capillary wall, the dielectric constant of the liquid and the electric field strength, but inversely proportional to the viscosity of the liquid. Therefore, the observed decrease in the velocity of the electroosmotic flow was probably caused by the diminution of the dielectric constant and/or increase in the viscosity of the solution due to the addition of
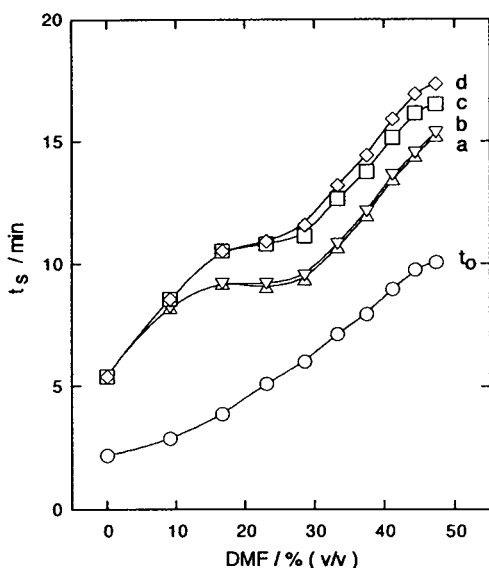
Fig. 2. Dependence of the migration time (t_s) on the DMF content of the running solution containing 25 mM SDS at pH 11. CZE conditions: migration distance, 50 cm; electric field, 429 V/cm. Compounds: (a) Chl-c_1; (b) Chl-c_2; (c) Pheo-c_1; (d) Pheo-c_2. t_0 = Migration time of methanol.

Fig. 3. Variation of apparent capacity factors (k'_{app}) of (a) Chl-c_1, (b) Chl-c_2, (c) Pheo-c_1 and (s) Pheo-c_2 with the addition of DMF to the running solution (pH 11). The SDS content was maintained at 25 mM. CZE conditions as in Fig. 2.

DMF to aqueous micellar solution (dielectric constants at 25°C, water 78.39 and DMF, 36.71 D; viscosities at 25°C, water 0.890 and DMF 0.802 cP).

The capacity factor could not be calculated for either chlorophyll in the running solution containing DMF, because a strict measurement of the velocity of the micelles was not carried out. The apparent capacity factor, k'_{app}, was calculated for each compound from the equation

$$k'_{app} = (t_s - t_0)/t_0 \qquad (1)$$

as a qualitative measure of the interaction between the solute and the micelle (k'_{app} has no strict relationship to the capacity factor usually employed in MECC).

The k'_{app} value of each compound decreased with increasing DMF content, as shown in Fig. 3. The decrease in k'_{app} with increasing DMF content is considered to result from the enhancement of the solubility in the DMF-containing bulk solution phase in a micellar solution. The migration of each chlorophyll compound is af-
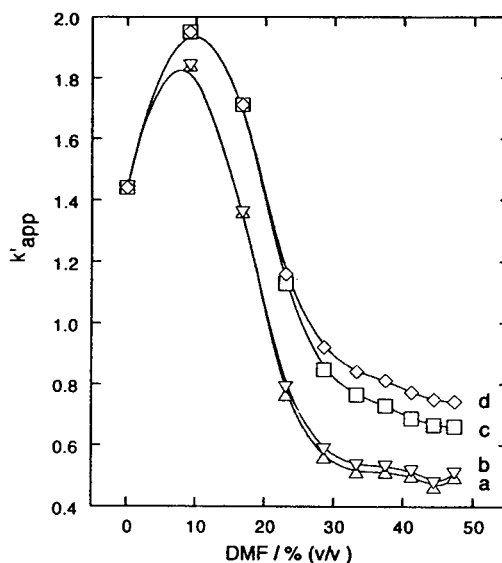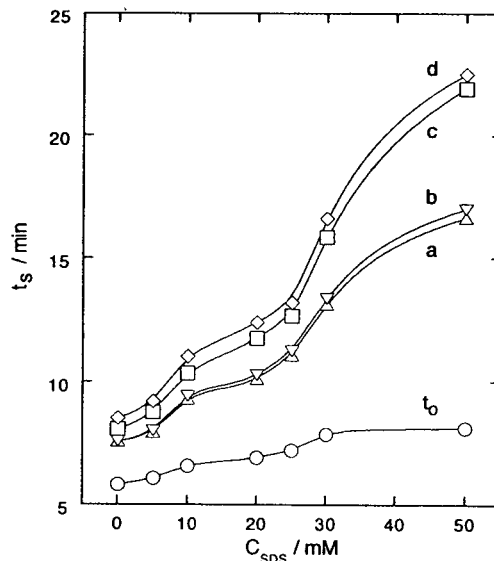

Fig. 4. Dependence of the migration times of Chl-c_1 and Chl-c_2 on the concentration of SDS in the running solution containing 33.3% (v/v) DMF. CZE conditions as in Fig. 2. Compounds: (a) Chl-c_1; (b) Chl-c_2; (c) Pheo-c_1; (d) Pheo-c_2. t_0 = Migration time of methanol.

fected by the partitioning between the SDS micelles and bulk solution.

Fig 4 shows the effect of the concentration of SDS (C_{SDS}) in the running solution on the migration time (t_S) of chlorophyll compounds, the DMF content being maintained at 33.3% (v/v). The t_S value of each chlorophyll compound increased with increasing C_{SDS}, which was considered to result from the increase in their distribution to the SDS micelles. The resolution for the Chl-c_1–Chl-c_2 pair was also enhanced with increasing C_{SDS}. The separation of Chl-c_1 and Chl-c_2 was successful with $C_{SDS} \gtrsim 10$ mM.

3.3. Application

The applicability of the present CZE conditions was demonstrated for the analysis of an acetone extract from brown seaweed *Undaria*
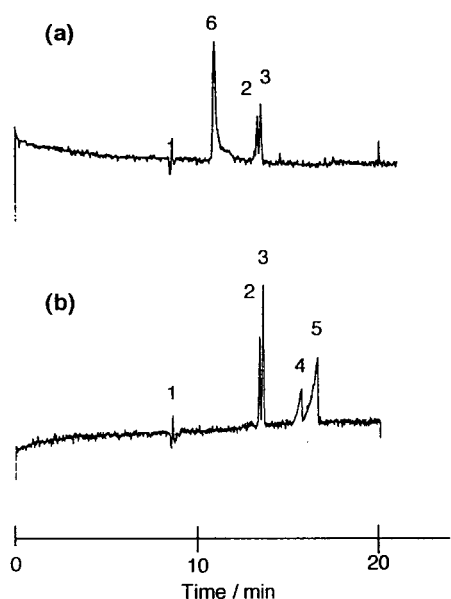
pinnatifida. The running solution contained SDS at 30 mM in a 10:6 (v/v) mixture of CAPS buffer (pH 11) and DMF (DMF content, 37.5%). The electropherograms obtained for the seaweed sample and a standard mixture are compared in Fig. 5. Three peaks are clearly found for the sample; they are assigned to Chl-a, Chl-c_1 and Chl-c_2, although no quantitative measurement was carried out.

4. Conclusions

According to the migration studies on different forms of chlorophyll-c, the main contribution to the separation of Chl-c_1 and Chl-c_2 is their partitioning phenomena between the micelles and the bulk solution, and the addition of DMF to the solution is the key to the successful resolution of the chlorophylls. MECC using a DMF-containing micellar solution is a promising tool for the separation and determination of Chl-c_1 and Chl-c_2, in particular.

Fig. 5. MECC separation of (a) chlorophyll pigments in acetone extracts from *Undaria pinnatifida* and (b) standard mixture. Running solution, 30 mM SDS in a mixture of 20 mM CAPS buffer (pH 11) and DMF (10:6, v/v); capillary, 50 cm (effective length for separation) × 50 μm I.D.; electric field, 429 V/cm; detection at 430 nm. Peaks: 1 = methanol; 2 = Chl-c_1; 3 = Chl-c_2; 4 = Pheo-c_1; 5 = Pheo-c_2; 6 = Chl-a.

References

[1] R.F.C. Montoura and C.A. Llewellyn, *Anal. Chim. Acta*, 151 (1983) 297.
[2] A.P. Murray, C.F. Gibbs, A.R. Longmore and D.J. Elet, *Mar. Chem.*, 19 (1986) 211.
[3] N. Suzuki, K. Saitoh and K. Adachi, *J. Chromatogr.*, 408 (1987) 181.
[4] S.W. Jeffrey, *Biochim. Biophys. Acta*, 177 (1969) 456.
[5] S.W. Jeffrey, *Biochim. Biophys. Acta*, 279 (1972) 15.
[6] K. Saitoh, I. Awaka and N. Suzuki, *J. Chromatogr.*, 653 (1993) 247.
[7] S. Terabe, K. Otsuka, K. Ichikawa, A. Tsuchiya and T. Ando, *Anal. Chem.*, 56 (1984) 111.
[8] K. Saitoh, C. Kiyohara and N. Suzuki, *J. High Resolut. Chromatogr.*, 14 (1991) 245.
[9] K. Saitoh, C. Kiyohara and N. Suzuki, *Anal. Sci.*, 7S (1991) 269.
[10] C. Kiyohara, K. Saitoh and N. Suzuki, *J. Chromatogr.*, 646 (1993) 397.
[11] Y.J. Yao, H.K. Lee and S.F.Y. Li, *J. Chromatogr.*, 637 (1993) 195.

ELSEVIER

Improved sensitivity by on-line isotachophoretic preconcentration in the capillary zone electrophoretic determination of peptide-like solutes[☆]

Dirk T. Witte*, Sofia Någård, Marita Larsson

Department of Bioanalytical Chemistry, Astra Hässle AB, S-431 83 Mölndal, Sweden

Abstract

Isotachophoresis (ITP) was studied as an on-line preconcentration technique in combination with capillary zone electrophoresis for the determination of positively charged peptide-like solutes. This paper shows the effects of different electrolyte compositions, different ITP times and various injection volumes on the time necessary for stacking and destacking. Injection volumes as large as 1.4 μl resulted in a resolution between the test peptides similar to that of pure capillary zone electrophoresis. Calibration graphs were linear in a range from 15 μmol/l down to 30 nmol/l after injections of 1.4 μl. Initial studies of plasma extracts appeared promising, although in the long term the overall efficiency decreased.

1. Introduction

Capillary electrophoresis (CE) is a rapidly developing separation technique with a high separation power [1]. Recently a review of quantitative aspects of CE with regard to the analysis of pharmaceuticals was published [2]. However, CE is still considered insensitive in terms of concentration when compared with the more established method of high-performance liquid chromatography (HPLC). This is one of the major limitations with regard to the applicability of CE for bioanalytical purposes. The insensitivity is largely caused by the small vol-

ume of the detection cell used in CE and the small injection volume, nanolitres compared with microlitres in HPLC, which can be applied without sacrificing the separation efficiency [3–5]. It should be noted that when measured in terms of mass quantified, CE is capable of extremely low detection limits, as much as 2–3 times better than HPLC [6].

To improve the sensitivity on the detector side, different approaches can be followed. One of them is based on laser-induced fluorescence (LIF) detection [7]. Another approach was the development of a Z-shaped flow cell for UV detection [8,9]. Since Olivares et al. [10] published the first paper in 1987 on the coupling of CE with mass spectrometry (MS), much work has been done to make this detection technique applicable for CE [11–13].

Apart from optimization on the detector side,

---

* Corresponding author.
☆ Presented at the *6th International Symposium on High Performance Capillary Electrophoresis*, San Diego, CA, 31 January–3 February 1994.

a highly efficient way to improve the sensitivity of CE is the injection of larger volumes in combination with on-line preconcentration techniques [14]. One of the advantages of on-line versus off-line preconcentration techniques is that the complete process can be easily automated. Different electrophoretic analyte focusing techniques can be used to increase the sample volume in CE. These are based on local differences in the electrical field strength to permit stacking of the analyte ions [15,16].

Field-amplified injection is a promising approach for preconcentration where the analyte is dissolved in a sample matrix, usually water, which has a lower conductivity than the background electrolyte. Owing to this difference in conductivity, stacking takes place at the boundary between the injected sample and the background electrolyte [17].

Isotachophoresis (ITP) is another stacking technique which can be used on-line with CE [15,16,18–21]. ITP is carried out in a discontinuous buffer system. The analytes are injected between a leading electrolyte (higher mobility than the analytes) and a terminating electrolyte (lower mobility). When a voltage drop is applied over the capillary, a steady-state migration configuration will be established in which the analytes migrate as consecutive zones which are not diluted by the background electrolyte as in capillary zone electrophoresis (CZE). During ITP, the concentration of the analyte zones will adapt to the higher concentration of the leading ion [22] and therefore stacking of the analytes will occur. As ITP results in consecutive zones of analytes, on-line ITP can be used as a preconcentration stacking technique before destacking and CZE separation take place.

In this work, the on-line combination of ITP with CZE in a single capillary was investigated for the determination of peptide-like solutes. The experiments were based on a combined ITP–CZE system initially described by Foret et al. [16], which was also studied for the determination of positively charged solutes with injection volumes up to 700 nl [23]. The effects of different electrolyte compositions, ITP times and injection volumes on the time necessary for stacking and destacking are described. Theoret-

ical models for different ITP–CZE systems have been presented, which show that both stacking and destacking can be controlled [24,25]. One of these models [25] is compared with results obtained in practice.

The present study demonstrates that injection volumes as large as 1.4 μl result in a resolution similar to that with small sample volumes in CZE, and that even larger injection volumes can be analysed by using recently published techniques [26,27] based on back-pressure. Preliminary results show that the combination of ITP–CZE with a back-pressure system can be used for bioanalytical work in the analysis of plasma extracts obtained through solid-phase extraction.

2. Experimental

2.1. Equipment

Experiments were performed with a PRINCE programmable injector for CE equipped with a high-voltage power supply (Lauerlabs, Emmen, Netherlands). Injections were made hydrodynamically and the volume was calculated according to Poiseuille with viscosities taken from the literature [28]. For detection a Spectra 100 UV detector (Spectra-Physics, San Jose, CA, USA) with an on-column cell (Linear Instruments, Reno, NV, USA) was used at 216 nm. For integration an SP-4270 integrator (Spectra-Physics) was used. The I.D. of the fused-silica capillary was 100 μm and the effective and total lengths were 57.5 and 72.5 cm, respectively (Polymicro Technologies, Phoenix, AZ, USA) unless stated otherwise. The capillary was thermostated by means of a forced air flow at 30°C unless other temperatures are indicated. To prevent electroosmotic flow, the capillary was coated with linear polyacrylamide [29], which also hindered the adsorption of the peptides on the capillary wall.

2.2. Chemicals

Ultra-high-purity grade water (18 MΩ cm resistivity) was obtained from an ELGA (High Wycombe, UK) purification system. Ammonium
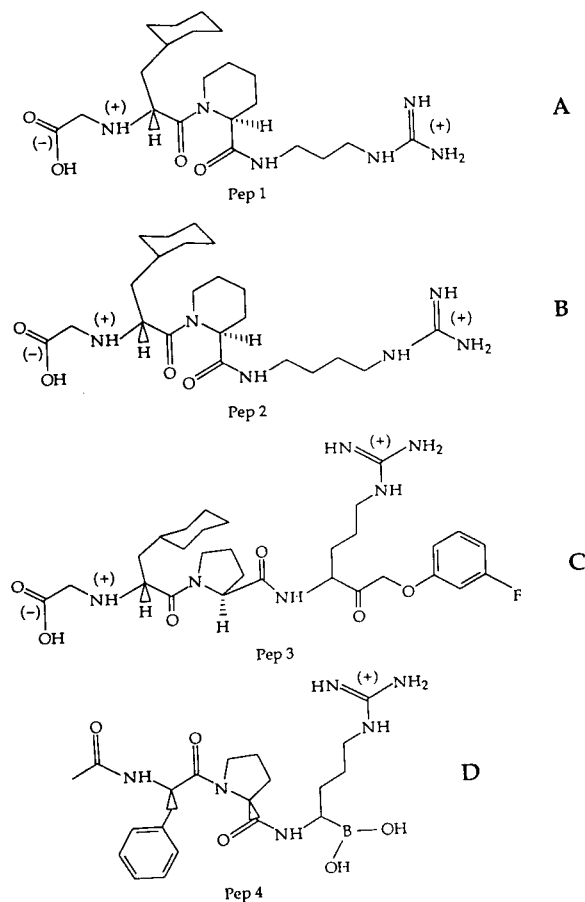
Fig. 1. Structures of test substances. (A) Pep1; (B) Pep2; (C) Pep3; (D) Pep4; (+) and (−) denote the actual charge at pH 3.6. The net charge is +1.

acetate (MicroSelect grade) and 6-aminocaproic acid (EACA) (puriss) were purchased from Fluka (Buchs, Switzerland). Acrylamide, ammonium peroxodisulfate and N,N,N′,N′-tetramethylethylenediamine (TEMED), used for the preparation of the coating, were obtained from Bio-Rad (Richmond, CA, USA). 3-Methacryl-

Table 2
pK_a values together with the absolute and effective mobilities (μ) of the ions present in the electrolytes at pH 3.6

Electrolyte	pK_a	Absolute μ ($m^2/V \cdot s$)	Effective μ ($m^2/V \cdot s$)
LE[a]	9.2 [31]	$72 \cdot 10^{-9}$ [31]	$72 \cdot 10^{-9}$
TE[a]	4.7 [23]	$362 \cdot 10^{-9}$ [23]	$7 \cdot 10^{-9}$
BGE[a]	4.3 [32]	$30 \cdot 10^{-9}$ [32]	$28 \cdot 10^{-9}$

[a] See Table 1.

oxypropyltrimethoxysilane, used to silylate the capillary before coating, was obtained from ABCR (Karlsruhe, Germany). Glacial acetic acid (HAc) was purchased from Merck (Darmstadt, Germany) and acetonitrile (HPLC grade) from Rathburn (Walkerburn, UK).

Fig. 1 shows the structures of the peptide-like test substances. Two of the standard solutes (Pep1 and Pep2) were supplied by the Department of Medicinal Chemistry at Astra Hässle (Mölndal, Sweden). Pep3 was synthesized at

Table 3
Effective mobilities for Pep 1–4

Analyte	Voltage (kV)	Effective mobility (10^{-9} $m^2/V \cdot s$)
Pep1	20	14.4
	30	21.3
Pep2	20	13.9
	30	20.7
Pep3	20	15.0
	30	19.7
Pep4	20	17.0
	30	22.2

Capillary bore, 100 μm; effective length, 57.5 cm; total length, 72.5 cm; electrolyte, BGE; temperature, 30°C for Pep1 and Pep2, 35°C for Pep3 and Pep4.

Table 1
Composition of electrolyte solutions

Electrolyte	Composition
Leading electrolyte (LE)	10 mmol/l ammonium acetate (NH₄Ac), pH 3.6[a]
Terminating electrolyte (TE)	50 mmol/l acetic acid (HAc), pH 3.1
Background electrolyte (BGE)	20 mmol/l 6-aminocaproic acid (EACA), pH 3.6[a]

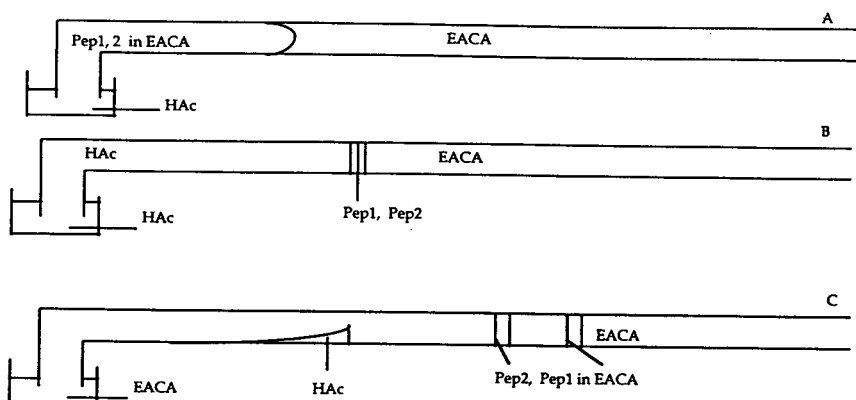[a] The BGE and LE were adjusted to pH 3.6 with acetic acid.

Fig. 2. ITP–CZE system with HAc as a terminating electrolyte and EACA as a background and leading electrolyte. (A) Start of ITP; (B) stacking complete; (C) controlled desacking followed by CZE.

Ferring Research Institute (Chilworth, Southampton, UK) [30] and Pep4 (DuP 714) at BACHEM Feinchemikalien (Bubendorf, Switzerland). It should be mentioned that the UV absorbances of Pep1 and Pep2 are lower than those for Pep3 and Pep4.

Tables 1 and 2 show the compositions of the electrolytes studied. Acetate was used as the counter ion and has an absolute mobility of $-42 \cdot 10^{-9}$ m^2/V·s [22].

2.3. Calibration

A calibration graph for Pep3 with Pep4 (1.5 μmol/l) as an internal standard was obtained by injections of 1.4 μl into a capillary with an I.D.
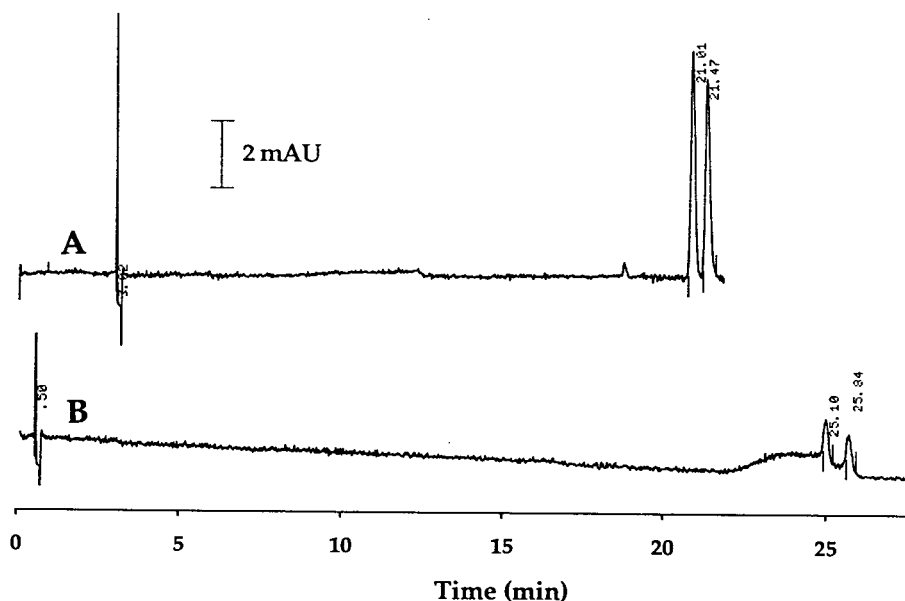

Fig. 3. Electropherograms obtained with HAc as a terminating electrolyte. Injection volume, 636 nl. ITP for (A) 3 and (B) 0.5 min; CE at 30 μA. The UV trace starts at the start of the ITP. ITP–CZE system as in Fig. 2.

of 100 μm and effective and total lengths of 47 and 64 cm, respectively. The studied samples contained Pep3 in a concentration range from 15 μmol/l down to 30 nmol/l. This injection volume results in a sample plug length of 18 cm. Three repetitive injections were made at each concentration level. The combined system, with HAc as terminating ion and NH_4^+ as leading ion, was used and ITP was initiated for 2 min at a constant current of 30 μA with the capillary end placed in a vial with terminating electrolyte at the anode side. The CZE separation took place at a constant voltage of 20 kV.

3. Results and discussion

3.1. Capillary zone electrophoresis

The electrophoretic behaviour of Pep1, Pep2, Pep3 and Pep4 in zone electrophoresis was studied with 18 mmol/l EACA buffer as background electrolyte. The differences in mobility between Pep1–4 are small, as shown in Table 3 and the differences observed between 20 or 30 kV might be due to heating of the capillary at the higher voltage. A voltage difference of 30 kV generated a current of approximately 60 μA and the total power generated was almost 2 W. This may result in heating in relatively wide-bore capillaries (100 μm) which are only cooled by an air-driven system such as that in our system. This results in a lower viscosity of the electrolyte and a higher mobility of the analytes. However, in general it also results in a higher diffusion and therefore a decrease in overall efficiency [31,32]. The decrease in efficiency is the main reason to avoid heating of the capillary. The effect of different injection volumes on the resolution of Pep1 and Pep2 was studied and showed that an increase in injection volume from 3.6 to 18 nl results in a decrease in resolution from 1.5 to almost 1.1.

3.2. Acetic acid as terminating electrolyte

Pep1 and Pep2 were dissolved in an aqueous solution of EACA, which was the background

electrolyte (BGE) in these experiments. The steps involved in the ITP–CZE sequence are depicted in Fig. 2. The main advantage of this system is that the time needed to collect the zone easily can be calculated according to Eq. 1 as proposed in literature [25]. The major disadvantage is that the vial at the inlet side of the capillary has to be replaced during the run, which prolongs the run time.
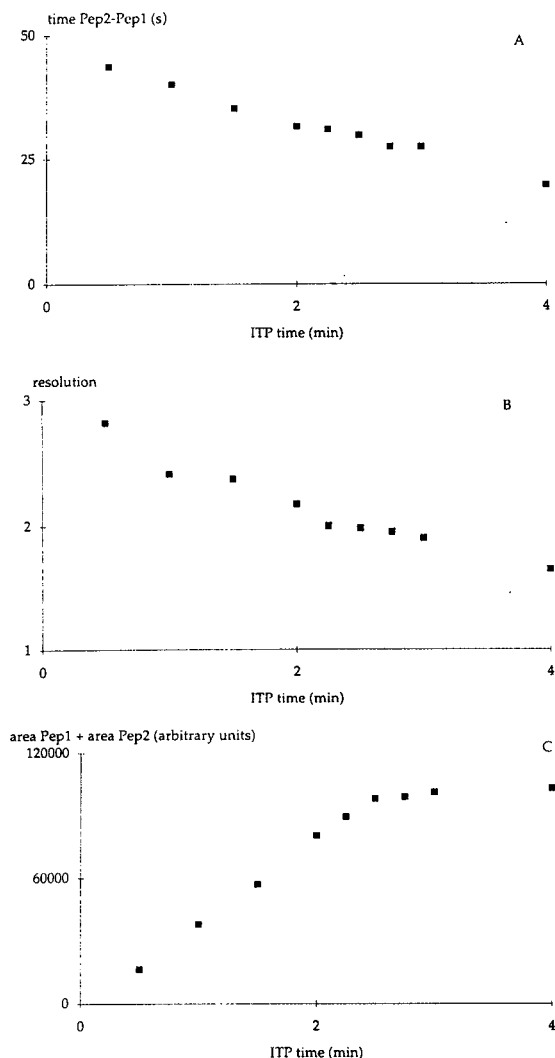

Fig. 4. Effect of different ITP times at a constant current of 30 μA on (A) difference in migration time between Pep1 and Pep2, (B) resolution between Pep1 and Pep2 and (C) total peak area. Injection volume, 636 nl of 7.5 μmol/l of Pep1 and Pep2. ITP–CZE system as in Fig. 2.

The fact that time and place of stacking and destacking need to be controlled can be seen in the Figs. 3 and 4. If destacking is initiated too early the whole plug will not be collected (Fig. 3B). If the destacking is too late the separation of the linked zones, which takes place during the CZE, is not complete. Fig. 4A shows that the difference in migration time (Δt) between Pep1 and Pep2 decreases if the stacked zones move further through the capillary. The same can be seen in Fig. 4B with regard to the resolution. The total migration time (ITP + CZE) of the analytes decreased on increasing the ITP time, i.e., the zones moved faster in the ITP mode than in the CZE mode.

The speed of the boundary between the leading and terminating electrolytes was measured at a constant current of 30 μA. The capillary was filled with leading electrolyte with the anode side placed in the terminating electrolyte. From the time it takes for the boundary to reach the detector one can calculate the speed of this boundary. At a fixed current this speed is assumed to remain constant. The observed speed of the zones during ITP was about 6.4 mm/s at a constant current of 30 μA. This means that a plug of 636 nl (8.1 cm) should be collected in about 2 min, but Fig. 4C shows that it takes longer than 2 min for the whole plug to be collected. This long collection path may be due to at least two reasons: the injected plug has a parabolic profile, which makes it longer than 8

cm, or already during the ITP those molecules of Pep1 and Pep2 which have not yet been collected move towards the cathode. For the second reason it is better to use Eq. 1 to calculate the time it takes to collect the whole sample. This equation takes into account the speed of the peptide-like solutes. The calculated time to collect the zone is 3.0 min at a constant current of 30 μA, which is in good agreement with our observations. From this equation, it can also be seen that a leading ion with a higher mobility, e.g., NH_4^+, results in a shorter collection time compared with a leading ion with a lower mobility, e.g., $EACAH^+$.

$$t = F[N_{le}\mu_{le}(1 - \mu_r/\mu_{le})]/I(\mu_{le} - \mu_s) \qquad (1)$$

where t = time to collect the injected plug (s); F = Faraday constant, 96 500 C/mol; N_{le} = number of moles of the leading ion present in the injected plug; μ_{le} = absolute mobility of leading ion ($cm^2/V \cdot s$); μ_r = absolute mobility of counter ion ($cm^2/V \cdot s$); μ_s = effective mobility of sample ion ($cm^2/V \cdot s$); I = current (A).

3.3. Ammonium as leading ion

Injections were made of mixtures of Pep1 and Pep2 in 10 mmol/l ammonium acetate in a capillary which contained EACA as a background electrolyte. The advantage of this system
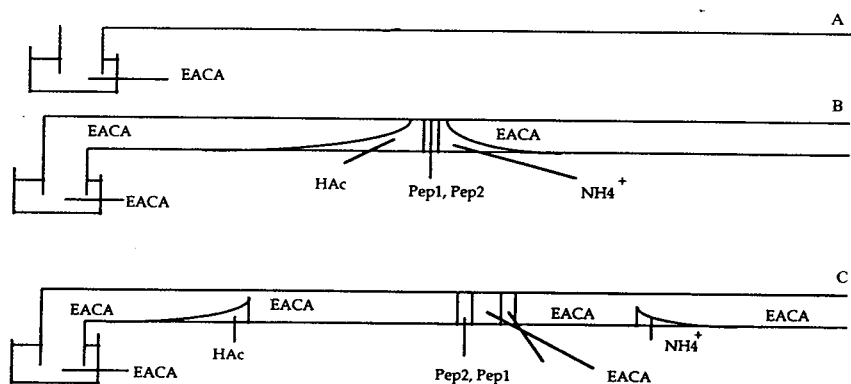

Fig. 5. ITP–CZE system with HAc as a terminating electrolyte, NH_4^+ as a leading electrolyte and EACA as a background electrolyte. (A) Start of ITP; (B) stacking complete; (C) uncontrolled destacking followed by CZE.

is that the inlet vial does not have to be replaced during the run and that the leading ion is fast, making the analysis quicker. The main disadvantage is that it is more complicated to estimate the position of the zone inside the capillary, as Eq. 1 is not valid. The ITP–CZE transition takes place smoothly. The basic steps in this system are outlined in Fig. 5. Fig. 6 shows electropherograms obtained after injections of different volumes. Δt (Fig. 7A) and the resolution (Fig. 7B) decrease at larger volumes because a larger part of the capillary is used for stacking. An increase

in the volume injected was accompanied by a decrease in migration time. The separation length (Table 4) was calculated from Δt, assuming that the solute mobilities were constant during CZE. Table 4 also shows that the sum of the length of the injected plug and the calculated capillary length used for separation is smaller than the effective length of the capillary. This means that the destacking takes place after collection of the whole plug and also that the plug length directly influences the place of and time for destacking.
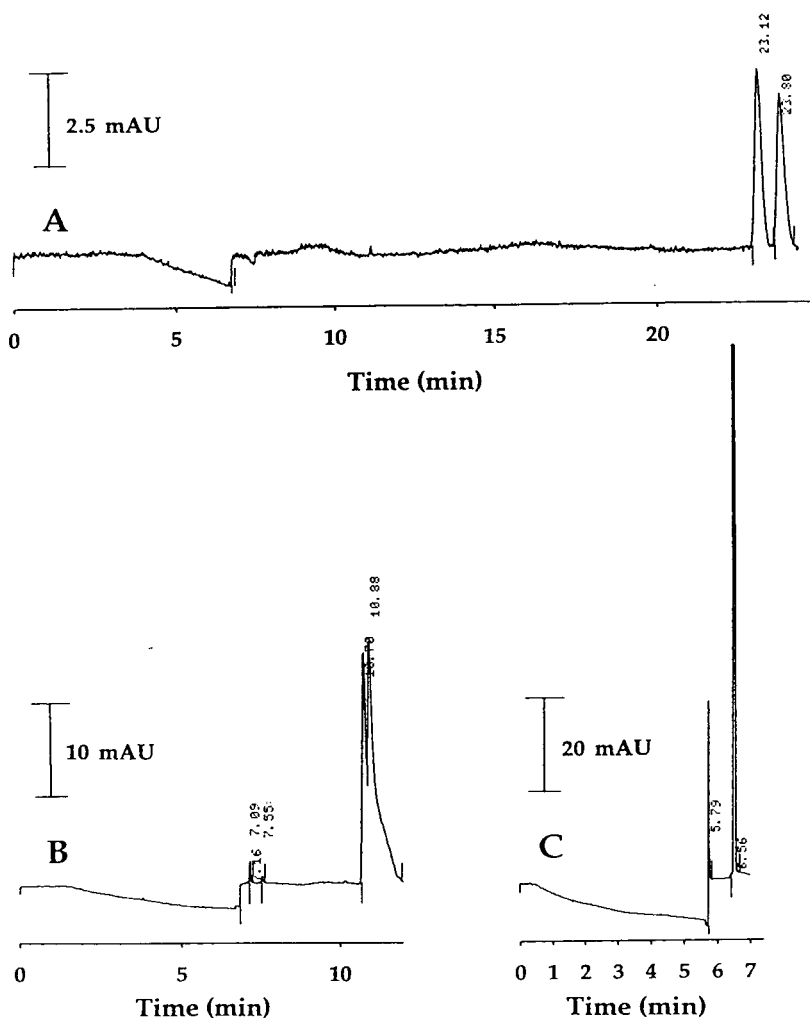

Fig. 6. Electropherograms after injections of (A) 636, (B) 3180 and (C) 4134 nl of a mixture of 7.5 μmol/l of Pep1 and Pep2 dissolved in 10 mmol/l ammonium acetate. ITP–CZE system as in Fig. 5.
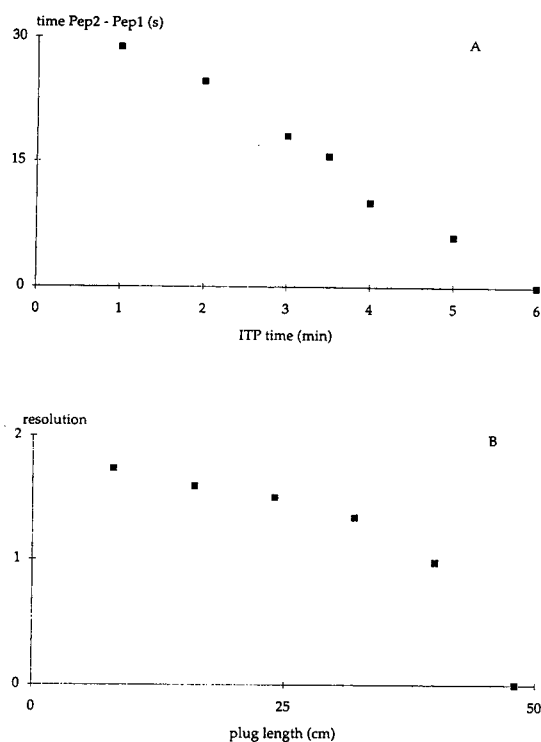
Fig. 7. Effect of different plug lengths (different injection volumes) on (A) difference in migration time between Pep1 and Pep2 and (B) resolution between Pep1 and Pep2. ITP–CZE system as in Fig. 5.

3.4. Combined ammonium acetate–acetic acid system

A system containing the combined electrolytes, as outlined in Fig. 8, would have the benefits from both systems described above. If NH_4^+ and HAc are used the system will be fast owing to the presence of a leading ion with a high mobility. In a combined system a volume of 636 nl will take 0.7 min to be collected instead of almost 3 min in the system without NH_4^+ owing to the faster moving leading ion present in the injected plug (Eq. 1). Owing to the presence of acetic acid, the concentrated zone will be easy to locate inside the capillary, as the ITP–CZE transition will not take place as long as HAc is present at the anode side. During zone electrophoretic separation the analytes will move through the capillary with EACA as a background electrolyte, which has a mobility much closer to that of the analytes compared with NH_4^+. With regard to peak shape and efficiency, it is favourable to choose a background electrolyte having a mobility close to that of the analytes of interest [33]. Still the main disadvantage of this combined system will be that the inlet vial has to be changed during the run. Fig.

Table 4
Calculated length left for separation at different injection volumes

Injection volume (nl)	Plug length[a] (cm)	Migration time (min)		Δt[b] (s)	CZE length[c] (cm)	Calculated length[d] (cm)
		Pep1	Pep2			
636	8	23.06	23.72	39.6	54	62
1272	16	17.58	18.00	25.2	35	51
1908	24	15.09	15.40	18.6	25	49
2544	32	12.61	12.84	13.8	19	51
3180	40	9.89	10.01	7.2	10	50
3816	49	7.36	7.39	1.8	2	51

Constant total amount of 4.8 pmol of each of Pep1 and Pep2; electrolyte system as in Fig. 5.
[a] Length of the injected plug.
[b] Δt = time(Pep2) − time(Pep1).
[c] The length used for separation during CZE after ITP, calculated from $57.5\Delta t/\Delta t_{CZE}$, assuming a constant selectivity. Δt_{CZE} = time(Pep2) − time(Pep1) measured in a CZE separation on a capillary with effective length of 57.5 cm (Δt_{CZE} = 42 s).
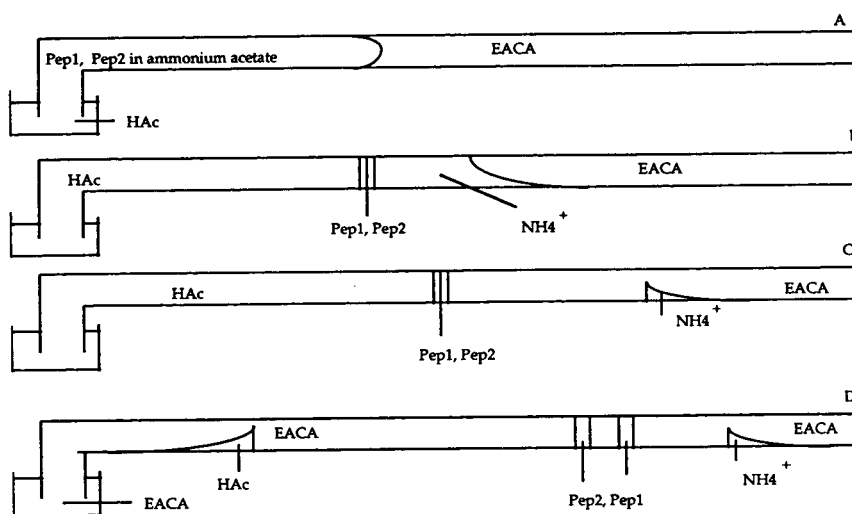[d] Plug length + CZE length; effective capillary length = 57.5 cm.

Fig. 8. ITP–CZE system with HAc as a terminating electrolyte, NH_4^+ as a leading electrolyte (LE) and EACA as a background electrolyte (BGE). (A) Start of the ITP; (B) stacking complete; (C) stacking continues; fast LE moves into slower BGE; (D) controlled destacking followed by CZE.
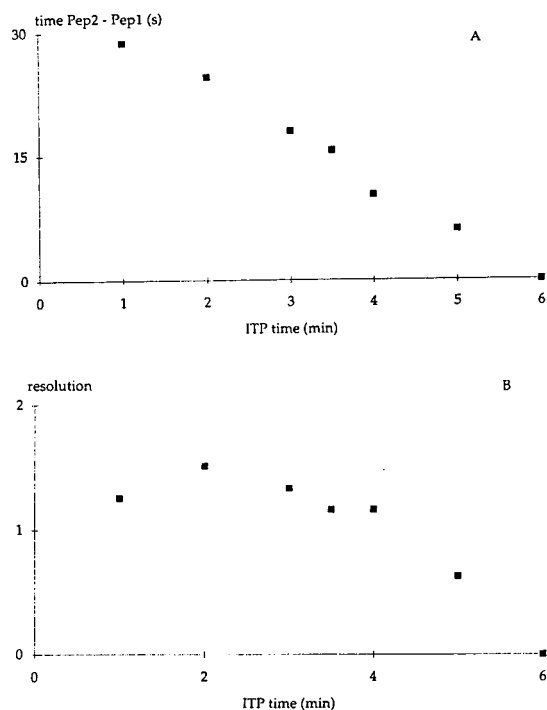

Fig. 9. Effect of different ITP times at a constant current of 30 μA on (A) difference in migration time between Pep1 and Pep2 and (B) resolution between Pep1 and Pep2. Injection volume, 1055 nl of 7.5 μmol/l of Pep1 and Pep2. ITP–CZE systems as in Fig. 8.

9A shows the influence of different ITP times on Δt and Fig. 9B shows a corresponding decrease in resolution. The total migration time decreased as the ITP period was prolonged.

3.5. Calibration graph

A calibration graph was obtained with the combined NH_4^+–HAc system showing the ratio of the area of Pep3 to that of Pep4 at different Pep3 concentrations. Fig. 10 shows electropherograms of the highest (15 μmol/l) and lowest (30 nmol/l) concentration of Pep3 studied. At the lowest concentration, 30 nmol/l with a signal-to-noise ratio of about 4, the R.S.D. of the ratio of the peak area of Pep3 to that of Pep4 was 11% ($n = 3$). The equation for the calibration graph is given in Table 5.

3.6. Analysis of plasma extracts

The plasma samples were extracted with a simple solid-phase extraction (SPE) method. A volume of 500 μl of the elution solvent from the SPE was evaporated under vacuum in a laboratory-made evaporation centrifuge. The residue was dissolved in 500 μl of EACA buffer–ace-
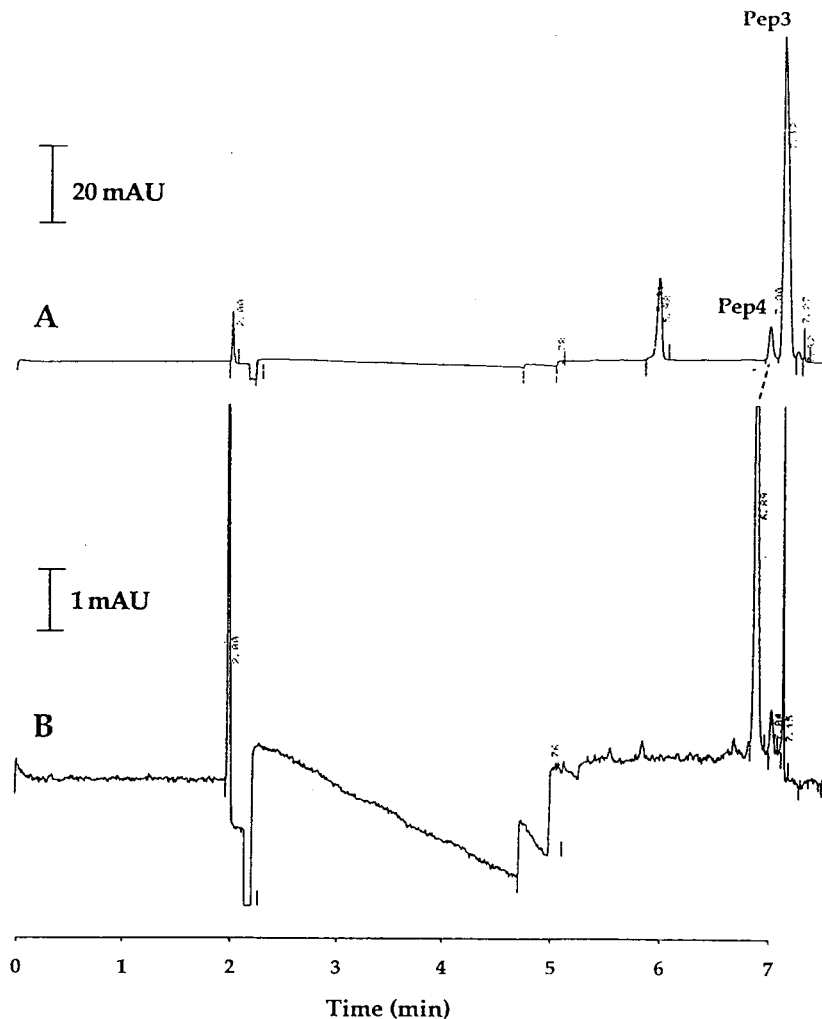
Fig. 10. Electropherogram of (A) 15 μmol/l and (B) 30 nmol/l of Pep3. Concentration of internal standard, 1.5 μmol/l. The UV trace starts at the start of the ITP. ITP–CZE system as in Fig. 8.

Table 5
Calibration graph

Equation	$y = ax + b$
Slope	$a = 1.076$
y-Intercept	$b = 0.004$
Correlation coefficient	0.9995

Equation of the calibration graph for Pep3 with Pep4 as internal standard (y = area counts in 0.50 μV·s; $x = \mu$mol/l). Concentration range, 30 nmol/l–15 μmol/l; injection volume, 1.4 μl; plug length, 18 cm. ITP–CZE system as in Fig. 8.

tonitrile (90:10, v/v). The acetonitrile was added to the electrolyte to prevent plasma components from sticking to the capillary. Without an organic modifier the efficiency of the CZE separation deteriorated seriously after only a few injections. Between runs the capillary was first rinsed for 2 min with acetonitrile followed by rinsing for 2 min with electrolyte.

To be able to measure pharmacological concentrations of Pep1, large volumes have to be injected. Therefore, injections of 1.5 μl were

Table 6
Instrumental sequence for the analysis of plasma extracts

Event	Time/action	Position	Inlet vial
Injection	2 min/75 mbar	33 cm[a]	Sample[b]
ITP	10 min/ 10 μA	36 cm	TE[c]
Δp	4 min/ −30 mbar	10 cm	Water
CZE	15 min/ 25 kV	–	BGE[d]
Clean 1	2 min/ 2000 mbar	–	Acetonitrile
Clean 2	2 min/ 2000 mbar	–	BGE
Replenish outlet	2 min/ 2 ml/min	–	BGE

Injection volume, 1500 nl; capillary bore, 75 μm; effective length, 50 cm; total length, 65 cm; temperature, 35°C; BGE, 10 mmol/l 6-aminocaproic acid (pH 3.6)–acetonitrile (90:10, v/v).
[a] Calculated front of the sample plug, measured from inlet side of the capillary.
[b] 3.18 μmol/l Pep1 and 2.18 μmol/l Pep2 in BGE.
[c] 50 mmol/l acetic acid (pH 3.1).
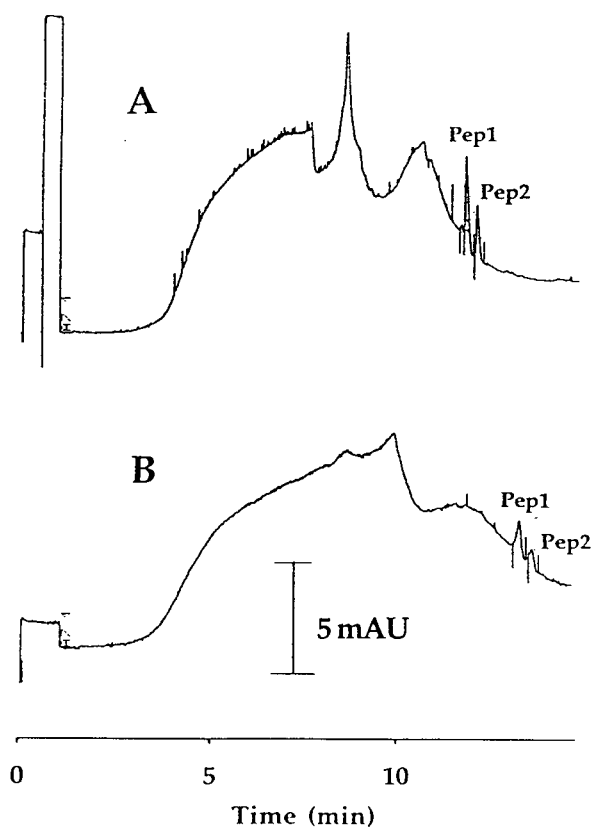[d] 10 mmol/l 6-aminocaproic acid (pH 3.6)–acetonitrile (90:10, v/v).

Fig. 11. Electropherograms (A) for the first injection and (B) after more than 100 injections of a plasma extract. The UV trace starts at the start of the CZE. ITP–CZE system as in Table 6.

made. After stacking of this long plug the resulting zones were too close to the detector to achieve sufficient separation during CZE. The zones needed to be transferred towards the inlet side of the capillary and under-pressure was therefore applied at the inlet side [26,27]. During the air-driven transfer, the sharp boundaries are disturbed by the parabolic flow profile. However, after moving the sample zone backwards, there is still a small zone of terminating electrolyte present in the capillary end. This facilitates refocusing of the zones, just before the CZE separation starts. Extension of the capillary length would result in a lower field strength and a longer operating time. The exact instrumental sequence used for the plasma extracts is given in Table 6. Fig. 11 shows the first injection and an injection after more than 100 runs. The differences in peak height and peak shape can be seen.

4. Conclusions

On-capillary ITP preconcentration permits the application of larger injection volumes for CE. All electrolyte systems presented in this paper were useful for the test substances. The main limitation with regard to the injection volume is the capillary length left for the separation after

the stacking. When destacking is initiated too close to the detector no separation, or insufficient separation, takes place if the stacked zones were not transferred towards the inlet side of the capillary. Early studies of plasma extracts with the combined system and a back-pressure technique appeared promising. Future work will focus on the improvement of the separation of analytes from plasma components and on the stability of the coating. We shall also try to decrease the overall run time, which is more than 37 min for the plasma samples in the system described. One should keep in mind that the injection of larger volumes in combination with the back-pressure system results in longer run times. Therefore, a balance between the required minimum detectable concentration and maximum acceptable run time should be found for each individual application.

Acknowledgement

References

[1] M. Albin, P.D. Grossman and S.E. Moring, *Anal. Chem.*, 65 (1993) 489A.
[2] K.D. Altria, *J. Chromatogr.*, 646 (1993) 245.
[3] D.M. Goodall, S.J. Williams and D.K. Lloyd, *Trends Anal. Chem.*, 10 (1991) 272.
[4] C.C. Campos and C.F. Simpson, *J. Chromatogr. Sci.*, 30 (1992) 53.
[5] F. Foret and P. Bocek, *Electrophoresis*, 11 (1990) 661.
[6] B.L. Karger, *Nature*, 339 (1989) 641.
[7] Y.F. Cheng and N.J. Dovichi, *Science*, 242 (1988) 562.
[8] J.P. Chervet, R.E.J. van Soest and M. Ursem, *J. Chromatogr.*, 543 (1991) 439.
[9] S.E. Moring, R.T. Reel and R.E.J. van Soest, *Anal. Chem.*, 65 (1993) 3454.
[10] J.A. Olivares, N.T. Nguyen, C.R. Yonker and R.D. Smith, *Anal. Chem.*, 59 (1987) 1230.
[11] W.A. Niessen, U.R. Tjaden and J. van der Greef, *J. Chromatogr.*, 636 (1993) 3.
[12] R. Kostiainen, E.J.F. Franssen and A.P. Bruins, *J. Chromatogr.*, 647 (1993) 361.
[13] N.J. Reinhoud, E. Schröder, U.R. Tjaden, W.M.A. Niessen, M.C. ten Noever de Braauw and J. van der Greef, *J. Chromatogr.*, 516 (1990) 147.
[14] C. Schwer, *LC · GC Int.*, 6 (1993) 630.
[15] D.S. Stegehuis, H. Irth, U.R. Tjaden and J. van der Greef, *J. Chromatogr.*, 538 (1991) 393.
[16] F. Foret, E. Szökö and B.L. Karger, *J. Chromatogr.*, 608 (1992) 3.
[17] R.-L. Chien and D.S. Burgi, *Anal. Chem.*, 64 (1992) 1046.
[18] T. Hirokawa, A. Ohmori and Y. Kiso, *J. Chromatogr.*, 634 (1993) 101.
[19] L. Krivankova, P. Gebauer, W. Thormann, R.A. Mosher and P. Bocek, *J. Chromatogr.*, 638 (1993) 119.
[20] D. Kaniansky, J. Marak, V. Madajova and E. Simunicova, *J. Chromatogr.*, 638 (1993) 137.
[21] F. Foret, V. Sustacek and P. Bocek. *J. Microcol. Sep.*, 2 (1990) 229.
[22] P. Bocek, M. Deml, P. Gebauer and V. Dolnik, in B.J. Radola (Editor), *Analytical Isotachophoresis*, VCH, Weinheim, 1988, pp. 40–57.
[23] M. Larsson and S. Någård, *J. Microcol. Sep.*, 6 (1994) 107.
[24] P. Gebauer, W. Thormann and P. Bocek, *J. Chromatogr.*, 608 (1992) 47.
[25] F. Foret, E. Szökö and B.L. Kager, *Electrophoresis*, 14 (1993) 417.
[26] N.J. Reinhoud, U.R. Tjaden and J. van der Greef, *J. Chromatogr.*, 641 (1993) 155.
[27] N.J. Reinhoud, U.R. Tjaden and J. van der Greef, *J. Chromatogr. A*, 653 (1993) 303.
[28] R.C. Weast (Editor), *Handbook of Chemistry and Physics*, CRC Press, Boca Raton, 68th ed., 1988, p. F-37.
[29] S. Hjertén, *J. Chromatogr.*, 347 (1985) 191.
[30] D.M. Jones, B. Atrash and M. Szelke, *Eur. Pat.*, EP 0530 167-A1, 1993.
[31] J.H. Knox and K.A. McCormack, *Chromatographia*, 38 (1994) 207.
[32] J.H. Knox and K.A. McCormack, *Chromatographia*, 38 (1994) 215.
[33] S. Hjertén, *Electrophoresis*, 11 (1990) 665.

Comparative stability study of thymidine and (dideoxy-D-erythro-hexopyranosyl)thymine analogues monitored by capillary electrophoresis

A. Van Schepdael*, K. Smets, F. Vandendriessche, A. Van Aerschot, P. Herdewijn, E. Roets, J. Hoogmartens

Laboratorium voor Farmaceutische Chemie en Analyse van Geneesmiddelen, Faculteit Farmaceutische Wetenschappen, K.U. Leuven, Van Evenstraat 4, 3000 Leuven, Belgium

Abstract

A capillary zone electrophoretic method was developed for monitoring the stability of hexopyranosyl analogues of thymidine. The final analytical conditions adopted were as follows: capillary, fused silica, 70 cm (62 cm to detector) × 75 μm I.D.; background electrolyte, 20 mM sodium tetraborate (pH 10.5); voltage, 30 kV; detection, UV at 262 nm; and temperature, 15°C. The performance of this analytical system is discussed. The stability study at three pH values allows to compare the stability of the analogues with that of thymidine and to investigate the influence of the configuration of the N-glycosidic bond on the rate constant of degradation.

1. Introduction

The compounds studied are depicted in Fig. 1: **1** = 1 - (2,3 - dideoxy - β - D - *erythro* - hexopyranosyl)thymine; **2** = 1 - (2,3 - dideoxy - α - D - *erythro* - hexopyranosyl)thymine; **3** = 1 - (2,4 - dideoxy - β - D - *erythro* - hexopyranosyl)thymine; **4** = 1 - (2,4 - dideoxy - α - D - *erythro* - hexopyranosyl) - thymine; **5** = 1 - (3,4 - dideoxy - β - D - *erythro* - hexopyranosyl)thymine. They are all modified nucleosides containing a thymine base but carrying a hexopyranosyl instead of a pentofuranosyl sugar. The sugar is deoxygenated in either positions 2 and 3, 2 and 4 or 3 and 4. The synthesis, the incorporation into oligonucleotides, the en-zymatic stability of these oligonucleotides and their base-pairing properties are described else-where [1–3]. As knowledge of the stability of the N-glycosidic bond in nucleosides is important with respect to their possible pharmaceutical use and to the stability of modified nucleosides in oligonucleotides (depyrimidination can lead to mutations), we undertook a comparative stability study of compounds **1–5** and thymidine (Thd).

2. Experimental

2.1. Reagents

Thymidine and thymine were purchased from Acros Chimica (Beerse, Belgium). The synthesis of **1–5** is described elsewhere [1–3]. All reagents
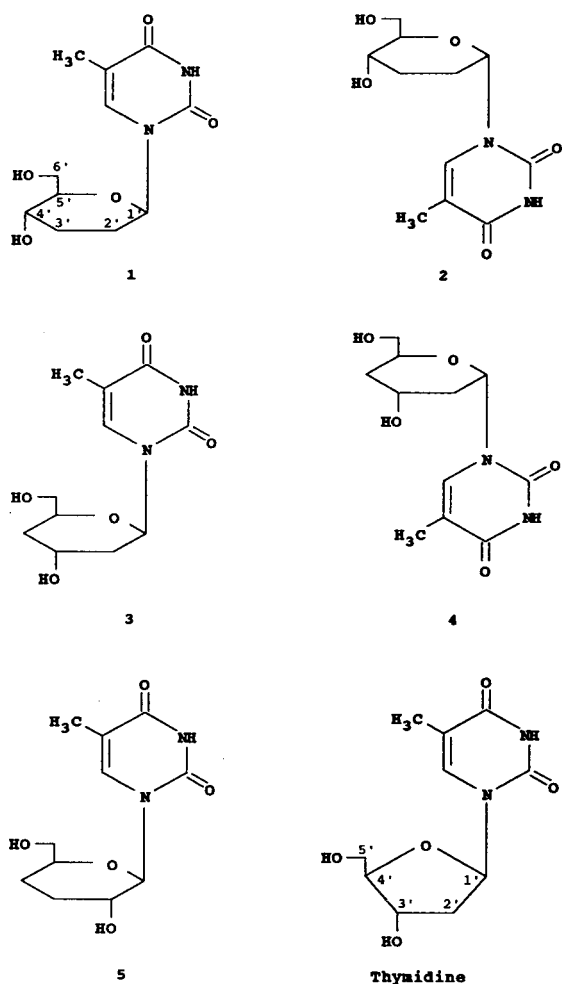
---

* Corresponding author.

Fig. 1. Structures of thymidine and hexopyranosyl analogues.

were of analytical-reagent grade. Buffers for capillary electrophoresis (CE) were prepared using water obtained from a Milli-Q[50] system (Millipore, Milford, MA, USA).

2.2. Capillary electrophoresis (CE)

CE was performed using a Spectraphoresis 500 instrument (Thermoseparation Products, Fremont, CA, USA) coupled to a 3396 Series II integrator (Hewlett-Packard, Avondale, PA, USA). Fused-silica capillaries were obtained from Polymicro Technologies (Phoenix, AZ,

USA). The pH of the sodium tetraborate buffer was adjusted with 1 M sodium hydroxide solution. Peak areas were corrected.

2.3. Stability study

All samples were incubated at 101°C in a Memmert (Schwabach, Germany) oven using buffers of different pH and with the ionic strength adjusted to 0.06 with KCl. A 0.01 M sodium dihydrogencitrate solution was brought to pH 1.2 with $1M$ HCl for the acid buffer. A 0.01 M neutral buffer of pH 6.8 and a 0.01 M alkaline buffer of pH 12.0 were prepared using the appropriate sodium phosphate species. Solutions (10^{-3} M) of the samples in the buffer were incubated for appropriate times and frozen. Just before CE analysis, acid samples were neutralized with 0.09 M KOH and alkaline samples with 0.01 M HCl.

3. Results and discussion

3.1. Analysis

Although a previous comparative stability study had been performed with liquid chromatography as the analytical technique, it suffered from some problems such as a decrease in retention of the analytes [4], so CE was preferred, also because of its proven speed in the determination of thymidine [5]. There are many papers on the CE of nucleosides, covering both capillary zone electrophoresis [6–11] and micellar electrokinetic capillary chromatography [12–18]. For this study the method used for thymidine [5] was slightly modified to improve the symmetry of the peaks and the stability of the compounds in the electrophoretic system. To achieve these respective aims, the pH was raised to 10.5 and the temperature of analysis was lowered to 15°C. The final analytical conditions were as follows: capillary, fused silica, 70 cm (62 cm to the detector) × 75 μm I.D.; background electrolyte, 20 mM sodium tetraborate (pH 10.5); voltage, 30 kV; detection, UV at 262 nm; temperature, 15°C; and current, limited to 150

μA. Samples were injected hydrodynamically for 3 s. Symmetry factors, resolution, migration times, mobilities, R.S.D.s of peak areas and migration times and numbers of theoretical plates are given in Table 1 for each of the compounds investigated. Table 2 collects calibration data accumulated in the concentration range 10^{-5}–10^{-3} M, using four calibration points (12 analyses in total). The correlation coefficient was 0.9999 in all instances. No reference substance was available for the unknown, so its mass was expressed as hexopyranosyl nucleoside in mass balance calculations. The detection limit (signal-to-noise ratio = 3) for Thd with this method was found to be 10^{-6} M. As six injections of this solution yielded an R.S.D. of 11%, we also regard this as the limit of quantification (LOQ). The equipment injects 4 nl s^{-1} of hydrodynamic injection on a capillary of 75 μm I.D. which gives an LOQ of 3 pg. The intra-day repeatability ($n = 10$) for Thd was 0.9% (migration time) and 1.4% (peak area). The inter-day repeatability ($n = 9$; 9 days) was 8.6% (migration time) and 3.1% (peak area).

3.2. Stability study

Fig. 2 shows representative electropherograms of samples degraded at pH 1.2. As described before [19], the degradation of thymidine yields not only thymine but also its anomer and pyranosyl isomers. This indicates that in the case of thymidine, degradation does not take place

Table 2
Calibration data for thymine, thymidine and hexopyranosyl analogues

Compound	a	b	$S_{y,x}$
Thymine	509	−222	205
Thymidine	372	279	631
1	390	101	226
2	389	263	316
3	359	426	333
4	394	217	193
5	377	128	556

a = Slope; b = intercept; y = corrected peak area; x = concentration (μg · ml^{-1}).

only by destabilization of the N-glycosidic bond after protonation of the heterocyclic base. Part of the mechanistic pathway must also involve sugar protonation and ring opening, which allows the sugar to reorganize into the anomer or pyranosyl isomer.

Nucleosides **1** and **3** only show the formation of thymine. Formation of their α-anomers **2** and **4**, respectively, might have taken place but the latter had probably already degraded because of their much higher degradation rate (a factor of 14 and 20 respectively; see Table 3). Nucleosides **2** and **4** show the formation of their (more stable) β-anomers **1** and **3** respectively. During degradation of **2**, small amounts of two unknown compounds were observed. These could be sugar isomers because their degree of formation is comparable to that of the anomer and this pattern of degradation is known for thymidine.

Table 1
Electrophoretic parameters for thymidine and hexopyranosyl analogues

Compound	Symmetry factor	Resolution with thymine	Migration time (min)	Mobility (cm^2 kV^{-1} min^{-1})	RSD (%)		No. of theoretical plates
					Peak area ($n = 3$)	Migration time ($n = 3$)	
Thymidine	1.0	15.8	7.0	23.3	1.2	0.3	139 000
1	1.1	4.7	6.8	24.0	0.3	0.8	81 500
2	1.0	11.5	6.3	25.9	1.0	0.2	165 500
3	1.0	17.7	7.5	21.8	1.1	1.6	178 600
4	1.0	15.8	6.7	24.4	0.6	0.4	186 000
5	0.9	13.5	7.3	22.4	0.6	0.1	176 400

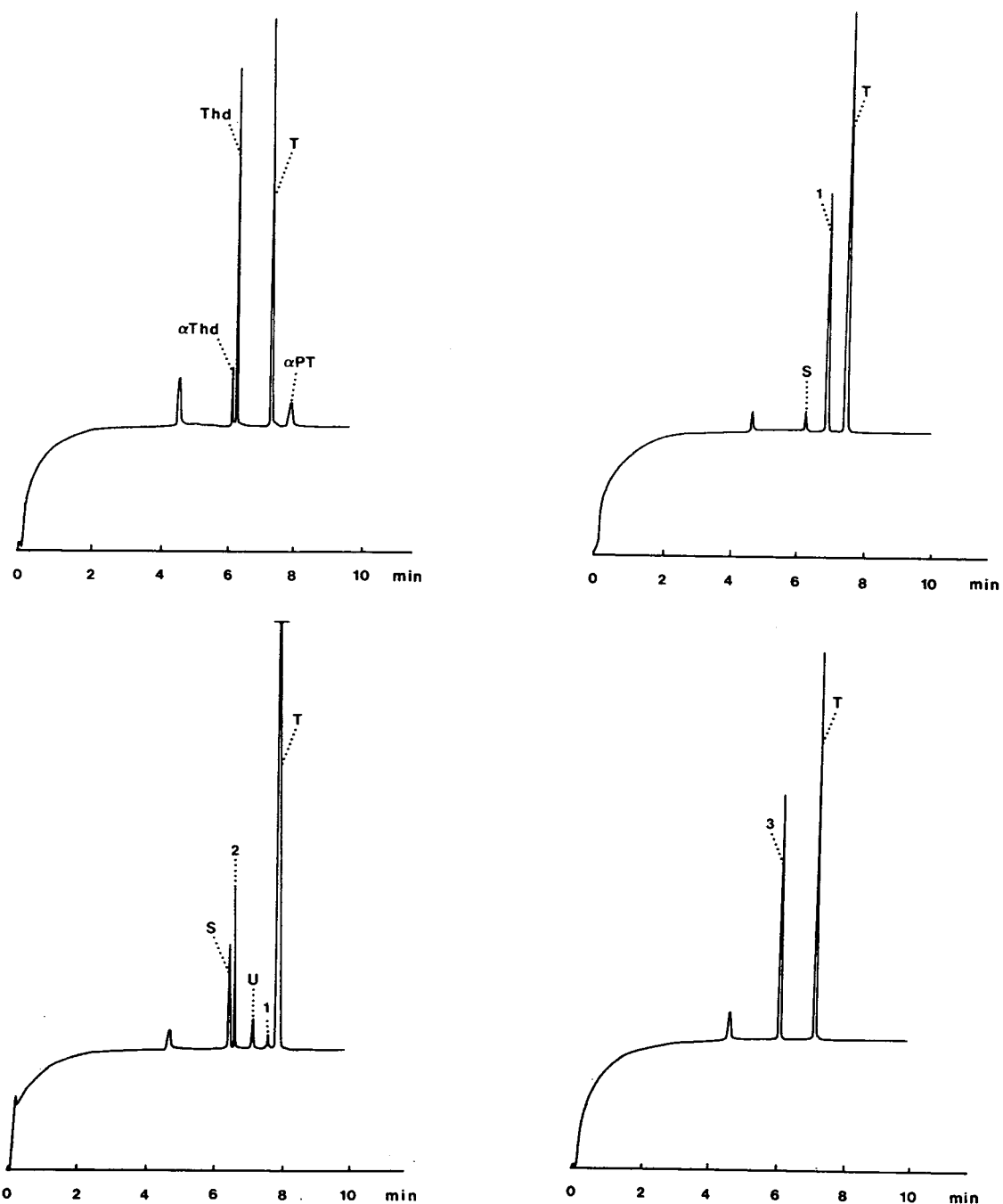For conditions, see text. Repeatability experiments were performed on 10^{-4} M solutions.

Fig. 2. Electropherograms of thymidine and hexopyranosyl analogues degraded at pH 1.22 and 101°C for the following periods: thymidine, 30 h; **1**, 216 h; **2**, 30 h; **3**, 216 h; **4**, 21 h; **5** 216 h. S = Synthetic impurity; U = unknown. αThd = 1-(2-deoxy-α-D-erythro-pentofuranosyl)thymine; Thd = thymidine; T = thymine; αPT = 1-(2-deoxy-α-D-erythro-pentopyranosyl)thymine. For analytical conditions, see text.
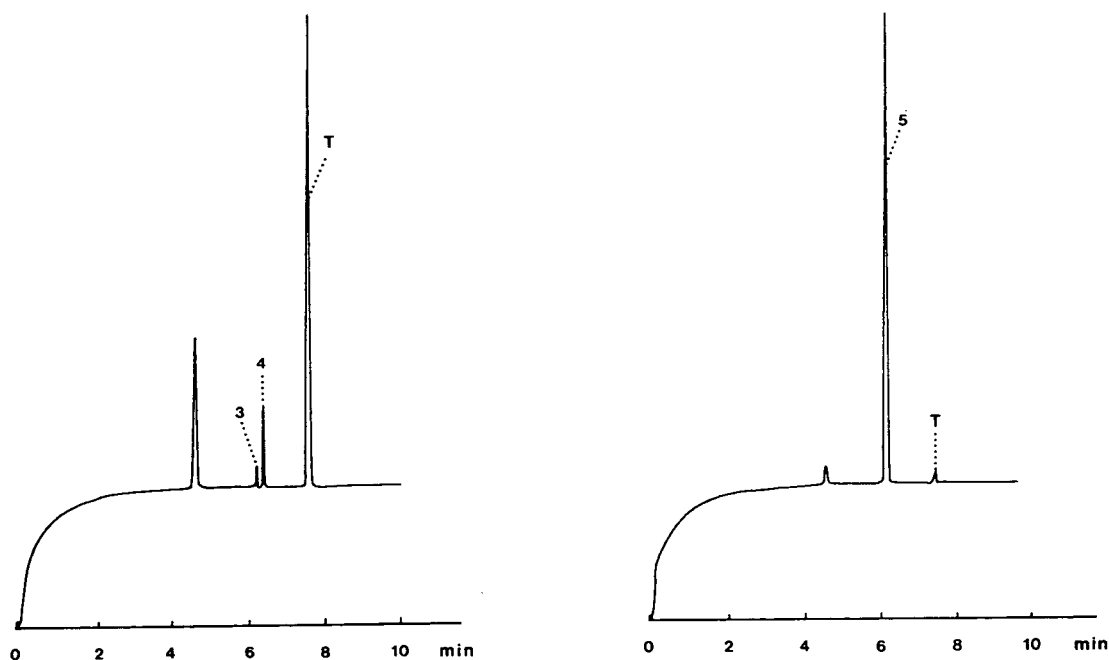
Fig. 2 (continued).

The hexopyranosyl sugar possibly isomerizes into a hexofuranosyl sugar. This could explain why nucleosides 3, 4 and 5 apparently do not form any isomers: they do not possess a 4'-hydroxyl group. On the other hand, for 1 no other degradation products apart from thymine could be distinguished. Nucleoside 5 was only degraded 1.7% after 216 h and is thus much more stable.

The degradation of the nucleosides seems to take place through the same mechanism as for thymidine in view of the degradation compounds formed and in view of relative stabilities of the different analogues. The higher stability of 5 can indeed be due to the presence of its 2'-OH, which destabilizes the positively charged intermediate and/or decreases protonation of the sugar oxygen. Mass balance calculations performed on the acid-degraded series fitted for all compounds except for 3, even after repeating the experiment (30% loss of total mass after 216 h). The reason for this is not clear. The amount of synthetic impurities present in 1 and 2 did not change during the experiment and these sub-

stances were therefore not taken into account during mass balance calculations. Degradation experiments at neutral and alkaline pH revealed the formation of thymine as the sole degradation product.

Table 3 contains the observed pseudo-first-order degradation rate constants at the three pH values tested. The hydrolysis rate increases towards the acidic region.

In contrast to thymidine [4], the difference in the hydrolysis rates of α- versus β-anomers (1 vs. 2, 3 vs. 4) is significant in acidic, neutral and alkaline media. As discussed before [4], thymidine anomers display the same kinetics in acidic media for stereoelectronic and conformational reasons. The N-glycosidic bond bears the same relationship to the lone pair of electrons of the ring oxygen whether it is in the α- or β-position [20]. It was also stated [20] that the puckering of the ring usually adjusts so that the substituent is quasi-axial in either case.

^{1}H NMR measurements at various temperatures [21] and molecular mechanics analysis [22] have further shown the flexibility of the

Table 3
Observed rate constants of degradation of thymidine and hexopyranosyl analogues at 101°C and $\mu = 0.06$

Compound	pH 1.2					pH 6.8					pH 12.0				
	k (h^{-1})	N	x	n	y	k (h^{-1})	N	x	n	y	k (h^{-1})	N	x	n	y
Thymidine	0.04516 ± 0.00088	16	8	1	2	0.0044 ± 0.00016	13	8	1	2	0.0015 ± 0.00013	32	8	2	<1
1	0.00637 ± 0.00021	30	8	2	2	0.0025 ± 0.00017	16	8	1	1	0.00052 ± 0.000047	32	8	2	<1
2	0.09110 ± 0.0020	32	8	2	4	0.0171 ± 0.00053	32	8	2	5	0.0039 ± 0.00014	28	8	2	1
3	0.00766 ± 0.00024	28	8	2	2	0.0032 ± 0.000073	32	8	2	2	0.00071 ± 0.0067	32	8	2	<1
4	0.15426 ± 0.00403	28	7	2	6	0.011 ± 0.00071	28	7	2	4	0.0067 ± 0.00037	14	7	1	2
5	m					s					s				

N = No. of electropherograms; x = No. of points on the time axis; n = No. of experiments; y = duration of experiments in half-lives; m = degradation negligible after 216 h; s = stable after 309 h.

thymidine sugar ring, which can adopt different conformations. Few reports exist on the conformation of glucopyranosyl-containing nucleosides. When 2,3-dideoxy-β-D-glucopyranosyl rings with either adenine or thymine as base were incorporated into an oligonucleotide, they were shown by NMR to be in the most stable chair conformation with all substituents equatorial and with both the adenine and the thymine base in the *anti* conformation [23]. This is consistent with a report [24] that large heterocyclic structures on the anomeric carbon of glucopyranosides, especially when they carry a positive charge, are in an equatorial position. This is due to a reverse anomeric and steric effect. When comparing the hydrolysis of acetals with an α- or β-bond, the α-isomer seems to cleave more easily because its conformation adapts better to a situation in which the ring oxygen lone pair is antiperiplanar with respect to the exocyclic bond [25,26].

Table 3 also allows one to compare the stability of thymidine containing a pentofuranosyl sugar with the anologue carrying a hexopyranosyl sugar (3). The latter is more stable than its pentofuranosyl counterpart towards acid hydrolysis. This has also been noted in the field of carbohydrates, where glycosides of 2-deoxyhexoses are much more stable than those of the corresponding 2-deoxypentoses and hexofuranosides of deoxy sugars are much more acid labile than the hexopyranosides [27].

Acknowledgements

References

[1] K. Augustyns, A. Van Aerschot and P. Herdewijn, *Nucleosides Nucleotides*, 10 (1991) 587.

[2] K. Augustyns, F. Vandendriessche, A. Van Aerschot, R. Busson, C. Urbanke and P. Herdewijn, *Nucleic Acids Res.*, 20 (1992) 4711.

[3] K. Augustyns, A. Van Aerschot, C. Urbanke and P. Herdewijn, *Bull. Soc. Chim. Belg.*, 101 (1992) 119.

[4] A. Van Schepdael, F. Heerinckx, A. Van Aerschot, P. Herdewijn, E. Roets and J. Hoogmartens, *Nucleosides Nucleotides*, 13 (1994) 1113.

[5] A. Van Schepdael, M. Vandewyer, A. Van Aerschot, P. Herdewijn, E. Roets and J. Hoogmartens, *J. Chromatogr.*, 648 (1993) 299.

[6] W.G. Kuhr and E.S. Yeung, *Anal. Chem.*, 60 (1988) 2642.

[7] D.J. Rose and J.W. Jorgenson, *J. Chromatogr.*, 438 (1988) 23.

[8] R.D. Smith, J.A. Olivares, N.T. Nguyen and H.R. Udseth, *Anal. Chem.*, 60 (1988) 436.

[9] R.D. Smith, H.R. Udseth, J.A. Loo, B.W. Wright and G.A. Ross, *Talanta*, 36 (1989) 161.

[10] V. Sustacek, F. Foret and P. Bocek, *J. Chromatogr.*, 480 (1989) 271.

[11] T. Grune, G.A. Ross, H. Schmidt, W. Siems and D. Perrett, *J. Chromatogr.*, 636 (1993) 105.

[12] A.S. Cohen, S. Terabe, J.A. Smith and B.L. Karger, *Anal. Chem.*, 59 (1987) 1021.

[13] A.S. Cohen, A. Paulus and B.L. Karger, *Chromatographia*, 24 (1987) 15.

[14] K.H. Row, W.H. Griest and M.P. Maskarinec, *J. Chromatogr.*, 409 (1987) 193.

[15] W.H. Griest, K.H. Row and M.P. Maskarinec, *Sep. Sci. Technol.*, 23 (1988) 1905.

[16] M.J. Gordon, X. Huang, S.L. Pentoney and R.N. Zare, *Science*, 242 (1988) 224.

[17] A.-F. Lecoq, C. Leuratti, E. Marafante and S. Di Biase, *J. High Resolut. Chromatogr.*, 14 (1991) 667.

[18] A.-F. Lecoq, L. Montanarella and S. Di Biase, *J. Microcol. Sep.*, 5 (1993) 105.

[19] J. Cadet and R. Téoule, *J. Am. Chem. Soc.*, 96 (1974) 6517.

[20] A.J. Kirby, *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*, Springer, Berlin, 1983, p. 61.

[21] N. Hicks, O.W. Howarth and D.W. Hutchinson, *Carbohydr. Res.*, 216 (1991) 1.

[22] P.C. Yates and S.V. Kirby, *Struct. Chem.*, 4 (1993) 299.

[23] G. Otting, M. Billeter, K. Wüthrich, H.-J. Roth, C. Leumann and A. Eschenmoser, *Helv. Chim. Acta*, 76 (1993) 2701.

[24] A.J. Kirby, *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*, Springer, Berlin, 1983, p. 16.

[25] A.J. Kirby, *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*, Springer, Berlin, 1983, p. 84.

[26] P. Deslongchamps, *Stereoelectronic Effects in Organic Chemistry*, Pergamon Press, Oxford, 1986, p. 39.

[27] G.A. Adams, in L. Whistler, J.N. BeMiller and W.L. Wolfrom (Editors), *Methods in Carbohydrate Chemistry*, Academic Press, New York, 1965, p. 286.

Short communication

Purification of (−)-epigallocatechin from enzymatic hydrolysate of its gallate using high-speed counter-current chromatography

Qizhen Du[a], Mingjun Li[a], Qi Cheng[a], Thian You Zhang[b], Yoichiro Ito[c],*

[a]*Tea Research Institute, Chinese Academy of Agricultural Sciences, Hangzhou, Zhejiang 310008, China*
[b]*Beijing Institute of New Technology Application, Beijing 100035, China*
[c]*Laboratory of Biophysical Chemistry, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA*

First received 13 June 1994; revised manuscript received 23 September 1994

Abstract

Epigallocatechin gallate (EGCG) was hydrolyzed at various concentrations of tannase under pH 6.0 at 35°C, and the reaction mixtures were separated by high-speed counter-current chromatography with a two-phase solvent system composed of hexane–ethyl acetate–water (1:13:20). The best results were obtained when 2 mg/ml of the enzyme buffer solution was added to 0.3 M EGCG buffers at a rate of 0.1 ml/min. Using 10 mg of the enzyme, 342 mg of epigallocatechin were obtained at a purity of 99.1%.

1. Introduction

(−)-Epigallocatechin (EGC) is an important natural product from tea leaves, since it possesses various medicinal properties including antioxidant effect, inhibition of tumor growth and prevention of cardiovascular sclerosis [1–3]. However, purification of EGC from tea extract requires two steps, i.e. Sephadex LH-20 column chromatography followed by HPLC [4]. We are currently preparing (−)-epigallocatechin gallate (EGCG) which can be hydrolyzed by tannase to produce EGC. For this purpose we conducted a series of experiments to maximize a yield of EGC from the hydrolysis of EGCG by tannase. In order to purify EGC from the hydrolysate, we

selected high-speed counter-current chromatography (HSCCC), because of its inherent advantage over conventional liquid chromatography by eliminating loss of samples by irreversible adsorption onto the solid support [5–7].

2. Experimental

2.1. Apparatus

HSCCC experiments were performed using a coil planet centrifuge equipped with a multilayer coil column that was designed and fabricated at the Beijing Institute of New Technology Application, China. The multilayer coil was prepared by winding a 1.6 mm I.D. polytetrafluoroethylene (PTFE) tube coaxially onto the column

* Corresponding author.

holder hub. The total column capacity measured 230 ml. The HSCCC centrifuge was rotated at 800 rpm with an 8 cm revolution radius. The system was equipped with an FMI pump (Zhejiang Instrument Factory, Hangzhou, China), a variable-wavelength UV detector (UV-752 made in Shanghai, China), a recorder and an injection valve.

2.2. Reagents

All of the organic solvents were of an analytical grade and purchased from Shanghai Chemical Factory, Shanghai, China. Tannase (tannin acylhydrolase, EC.3.1.1.20), 5000 U/g, was purchased from Kikkoman, Japan. The enzyme was stored at 4°C. EGCG was prepared using a new chromatographic method (Chinese patent pending) in our laboratory and stored at −20°C. Its purity was determined to be greater than 99% by HPLC analysis.

2.3. Tannase hydrolysis reaction

According to Thomas and Murtagh [8], tannase has an optimum pH at 6.0 and optimum temperature at 35°C. The tannase hydrolysis reaction reached the equilibrium in about 2 minutes though the enzyme concentration were different. We have selected five enzyme concentrations of 0.4, 0.8, 1.2, 1.6 and 2.0 mg/ml to initiate the reaction under pH 6.0 at 35°C (Fig. 1).

Five tubes, each containing 5 ml of 0.1 M NaH$_2$PO$_4$ buffers (pH 6.0) —in which the enzyme concentrations are 0.4, 0.8, 1.2, 1.6 and 2.0 mg/ml —were incubated in a shaking water bath

at 35°C. Then, 0.2 ml of 0.3 M EGCG in a buffer was added to each of the tubes at 2-min intervals until the hydrolysis process was halted. TLC was used to monitor the degree of hydrolysis at each step.

2.4. TLC analysis

Polyacrylamide plates used in thin-layer chromatography (TLC) were purchased from Huangyie Chemical Factory, China. Acetone–water (2:1) was used for development, and three analytes —EGCG, EGC and gallic acid— were detected by purple spots on the TLC plate after the color reaction with 0.1% FeCl$_3$. Each analyte in CCC fractions was identified using the respective standard samples.

2.5. HSCCC procedure

HSCCC experiments were performed with a two-phase solvent system composed of hexane–ethyl acetate–water (1:13:20, v/v/v). The solvent mixture was thoroughly equilibrated in a separatory funnel at room temperature and the two phases were separated. Both the upper and the lower phases were degassed separately by ultrasonication.

In each separation, the multilayer coil column was first entirely filled with the upper stationary phase. Then the lower mobile phase was pumped into the inlet of the column at a flow rate of 1.6 ml/min, while the apparatus was rotated at 800 rpm. After the mobile phase front emerged and the two phases had established the hydrodynamic equilibrium throughout the column, the sample solution was injected through the in-
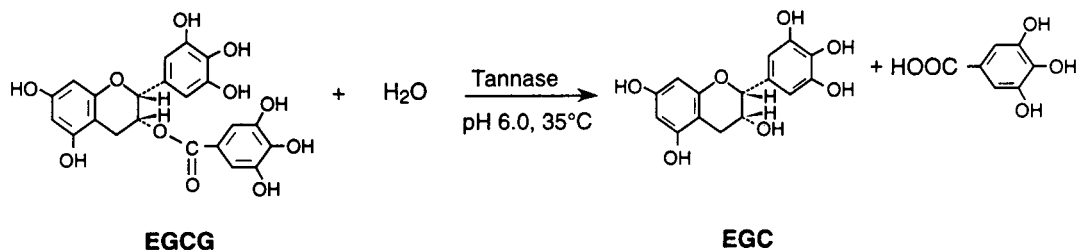

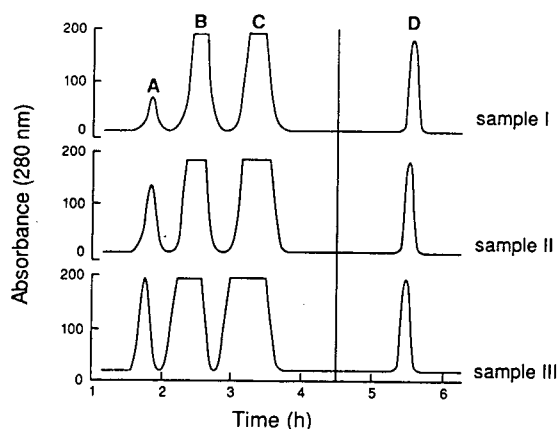Fig. 1. Hydrolysis reaction of EGCG by tannase.

Fig. 2. Chromatograms of three reaction mixtures. Samples I, II and III were obtained from 2, 6 and 10 mg of tannase, respectively (see Table 1). Sample sizes: I = 165 mg/6.2 ml; II = 360 mg/7.6 ml; III = 550 mg/9.0 ml.

jection valve. The effluent from the outlet of the column was continuously monitored with a UV detector at 280 nm and the peak fractions were collected (Fig. 2). After three peaks were eluted, the centrifuge was stopped and the column contents were fractionated by continuously eluting the column with the mobile phase.

3. Results and discussion

Table 1 shows the results of hydrolysis studies of the reaction mixture for five different enzyme concentrations. These data indicate that 2, 4, 6, 8 and 10 mg tannase can be used to hydrolyze 0.30 mmol EGCG in 10 min, 0.48 mmol EGCG in 16 min, 0.72 mmol EGCG in 24 min, 0.96 mmol EGCG in 32 min, and 1.14 mmol EGCG

Table 1
Analyses of reaction mixtures by five different tannase concentrations

Reaction time (min)	EGCG added (mmol)	Initial tannase concentrations in the reaction mixture (mg/5 ml)				
		2	4	6	8	10
2	0.60	+	+	+	+	+
4	0.12	+	+	+	+	+
6	0.18	+	+	+	+	+
8	0.24	+	+	+	+	+
10	0.30	+	+	+	+	+
12	0.36	+/−	+	+	+	+
14	0.42	+/−	+	+	+	+
16	0.48	+/−	+	+	+	+
18	0.54	+/−	+/−	+	+	+
20	0.60	+/−	+/−	+	+	+
22	0.66	+/−	+/−	+	+	+
24	0.72	+/−	+/−	+	+	+
26	0.78	+/−	+/−	+/−	+	+
28	0.84	+/−	+/−	+/−	+	+
30	0.90	+/−	+/−	+/−	+	+
32	0.96	+/−	+/−	+/−	+	+
34	1.02	+/−	+/−	+/−	+/−	+
36	1.08	+/−	+/−	+/−	+/−	+
38	1.14	+/−	+/−	+/−	+/−	+
40	1.20	+/−	+/−	+/−	+/−	+/−

+ = Hydrolyzed completely, two spots appearing on TLC plate; +/− = hydrolyzed incompletely, the third spot (EGCG) detected on TLC plate.

in 38 min, respectively. All five reaction mixtures were subjected to HSCCC separation after cooling down to room temperature. Fig. 2 shows chromatograms of three reaction mixtures obtained by HSCCC. We found that peak resolution between three eluted analytes (A, B and C) was reduced as the sample volume was increased, reaching almost its limit for sample III (Fig. 2).

As indicated earlier, peaks A, B and C were eluted with the mobile phase while peak D was collected from the column after stopping the centrifugation. In TLC analysis, fractions corresponding to peaks B, C and D each produced a single spot and were identified as gallic acid, EGC, and EGCG, respectively. Each analyte produced a monochromatic purple color reaction with 0.1% $FeCl_3$ on the TLC plate. Fraction corresponding to peak A showed a negative color reaction to 0.1% $FeCl_3$ suggesting that it is probably the enzyme protein.

HSCCC of sample III produced 342 mg of EGC with a high purity of 99.1% as determined by HPLC. The overall results of the present studies demonstrated that HSCCC can be effectively used for the preparation of EGC from the EGCG hydrolysate by tannase.

References

[1] Y. Hara, *New Food Ind.*, 32, No. 2 (1990) 33.
[2] Y. Hara, S. Matsuzaki and K. Nakamura, *Nippon Eiyo, Shokuryo Gakkaishi*, 42, No. 1 (1989) 39.
[3] J. Bao, X. Hong, J. Lou, H. Zhang and L. Li, *Proceedings of International Tea-Quality–Human Health Symposium*, Tea Research Institute, Chinese Academy of Agricultural Sciences, Hangzhou, 1987, p. 239.
[4] Y. Takino, H. Tanizawa, H. Ikeda, and S. Fujiki, *Jap. Pat.*, 59:26385 (1984).
[5] A. Marston, I. Slacanin and K. Hostettmann, *Phytochem. Anal.*, 1 (1990) 3.
[6] Y. Ito, *J. Chromatogr.*, 188 (1980) 43.
[7] Y. Ito and W.D. Conway, *Anal. Chem.*, 56 (1984) 534A.
[8] R.L. Thomas and K. Murtagh, *J. Food Sci.*, 50 (1985) 1126.

INFORMATION FOR AUTHORS

(Detailed *Instructions to Authors* were published in *J. Chromatogr. A*, Vol. 657, pp. 463–469. A free reprint can be obtained by application to the publisher, Elsevier Science B.V., P.O. Box 330, 1000 AH Amsterdam, Netherlands.)

Types of Contributions. The following types of papers are published: Regular research papers (full-length papers), Review articles, Short Communications and Discussions. Short Communications are usually descriptions of short investigations, or they can report minor technical improvements of previously published procedures; they reflect the same quality of research as full-length papers, but should preferably not exceed five printed pages. Discussions (one or two pages) should explain, amplify, correct or otherwise comment substantively upon an article recently published in the journal. For Review articles, see inside front cover under Submission of Papers.

Submission. Every paper must be accompanied by a letter from the senior author, stating that he/she is submitting the paper for publication in the *Journal of Chromatography A* or *B*.

Manuscripts. Manuscripts should be typed in **double spacing** on consecutively numbered pages of uniform size. The manuscript should be preceded by a sheet of manuscript paper carrying the title of the paper and the name and full postal address of the person to whom the proofs are to be sent. As a rule, papers should be divided into sections, headed by a caption (*e.g.*, Abstract, Introduction, Experimental, Results, Discussion, etc.). All illustrations, photographs, tables, etc., should be on separate sheets.

Abstract. All articles should have an abstract of 50–100 words which clearly and briefly indicates what is new, different and significant. No references should be given.

Introduction. Every paper must have a concise introduction mentioning what has been done before on the topic described, and stating clearly what is new in the paper now submitted.

Experimental conditions should preferably be given on a *separate* sheet, headed "Conditions". These conditions will, if appropriate, be printed in a block, directly following the heading "Experimental".

Illustrations. The figures should be submitted in a form suitable for reproduction, drawn in Indian ink on drawing or tracing paper. Each illustration should have a caption, all the *captions* being typed (with double spacing) together on a *separate sheet*. If structures are given in the text, the original drawings should be provided. Coloured illustrations are reproduced at the author's expense, the cost being determined by the number of pages and by the number of colours needed. The written permission of the author and publisher must be obtained for the use of any figure already published. Its source must be indicated in the legend.

References. References should be numbered in the order in which they are cited in the text, and listed in numerical sequence on a separate sheet at the end of the article. Please check a recent issue for the layout of the reference list. Abbreviations for the titles of journals should follow the system used by *Chemical Abstracts*. Articles not yet published should be given as "in press" (journal should be specified), "submitted for publication" (journal should be specified), "in preparation" or "personal communication".

Vols. 1–651 of the *Journal of Chromatography*; *Journal of Chromatography, Biomedical Applications* and *Journal of Chromatography, Symposium Volumes* should be cited as *J. Chromatogr.* From Vol. 652 on, *Journal of Chromatography A* (incl. Symposium Volumes) should be cited as *J. Chromatogr. A* and *Journal of Chromatography B: Biomedical Applications* as *J. Chromatogr. B*.

Dispatch. Before sending the manuscript to the Editor please check that the envelope contains four copies of the paper complete with references, captions and figures. One of the sets of figures must be the originals suitable for direct reproduction. Please also ensure that permission to publish has been obtained from your institute.
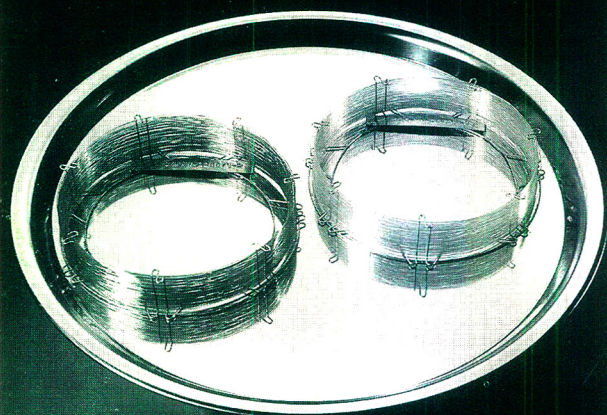
Proofs. One set of proofs will be sent to the author to be carefully checked for printer's errors. Corrections must be restricted to instances in which the proof is at variance with the manuscript.

Reprints. Fifty reprints will be supplied free of charge. Additional reprints can be ordered by the authors. An order form containing price quotations will be sent to the authors together with the proofs of their article.

Advertisements. The Editors of the journal accept no responsibility for the contents of the advertisements. Advertisement rates are available on request. Advertising orders and enquiries can be sent to the Advertising Manager, Elsevier Science B.V., Advertising Department, P.O. Box 211, 1000 AE Amsterdam, Netherlands; courier shipments to: Van de Sande Bakhuyzenstraat 4, 1061 AG Amsterdam, Netherlands; Tel. (+31-20) 515 3220/515 3222, Telefax (+31-20) 6833 041, Telex 16479 els vi nl. *UK*: T.G. Scott & Son Ltd., Tim Blake, Portland House, 21 Narborough Road, Cosby, Leics. LE9 5TA, UK; Tel. (+44-533) 753 333, Telefax (+44-533) 750 522. *USA and Canada*: Weston Media Associates, Daniel S. Lipner, P.O. Box 1110, Greens Farms, CT 06436-1110, USA; Tel. (+1-203) 261 2500, Telefax (+1-203) 261 0101.