Vol. 296, No. 1, September 30, 1994 ISSN 0003-2670

ANALYTICA CHIMICA ACTA

An international journal devoted to all branches of analytical chemistry

ELSEVIER

ANALYTICA CHIMICA ACTA

Scope. *Analytica Chimica Acta* publishes original papers, rapid publication letters and reviews dealing with every aspect of modern analytical chemistry. Reviews are normally written by invitation of the editors, who welcome suggestions for subjects. Letters can be published within **four months** of submission. For information on the Letters section, see inside back cover.

Submission of Papers

Americas

Prof. Harry L. Pardue
Department of Chemistry
1393 BRWN Bldg, Purdue University
West Lafayette, IN 47907–1393
USA

Tel:(+1-317) 494 5320
Fax:(+1-317) 496 1200

Computer Techniques

Prof. J.T. Clerc
Universität Bern
Pharmazeutisches Institut
Baltzerstrasse 5, CH-3012 Bern
Switzerland

Tel:(+41-31) 6314191
Fax:(+41-31) 6314198

Prof. Sarah C. Rutan
Department of Chemistry
Virginia Commonwealth University
P.O. Box 2006
Richmond, VA 23284-2006
USA
Tel:(+1-804) 367 7517
Fax:(+1-804) 367 8599

Other Papers

Prof. Alan Townshend
Department of Chemistry
The University
Hull HU6 7RX
Great Britain

Tel:(+44-482) 465027
Fax:(+44-482) 466410

Prof. Willem E. van der Linden
Laboratory for Chemical Analysis
Department of Chemical Technology
Twente University of Technology
P.O. Box 217, 7500 AE Enschede
The Netherlands
Tel:(+31-53) 892629
Fax:(+31-53) 356024

Prof. Paul Worsfold
Dept. of Environmental Sciences
University of Plymouth
Plymouth PL4 8AA
Great Britain

Tel:(+44-752) 233006
Fax:(+44-752) 233009

Submission of an article is understood to imply that the article is original and unpublished and is not being considered for publication elsewhere. *Anal. Chim. Acta* accepts papers in English only. There are no page charges. Manuscripts should conform in layout and style to the papers published in this issue. See inside back cover for "Information for Authors".

Publication. *Analytica Chimica Acta* appears in 16 volumes in 1994 (Vols. 281–296). *Vibrational Spectroscopy* appears in 2 volumes in 1994 (Vols. 6 and 7). Subscriptions are accepted on a prepaid basis only, unless different terms have been previously agreed upon. It is possible to order a combined subscription (*Anal. Chim. Acta* **and** *Vib. Spectrosc.*).
Our p.p.h. (postage, packing and handling) charge includes surface delivery of all issues, except to subscribers in the U.S.A., Canada, Australia, New Zealand, China, India, Israel, South Africa, Malaysia, Thailand, Singapore, South Korea, Taiwan, Pakistan, Hong Kong, Brazil, Argentina and Mexico, who receive all issues by air delivery (S.A.L.–Surface Air Lifted) at no extra cost. For Japan, air delivery requires 25% additional charge of the normal postage and handling charge; for all other countries airmail and S.A.L. charges are available upon request.

Subscription orders. Subscription prices are available upon request from the publisher. Subscription orders can be entered only by calendar year and should be sent to: Elsevier Science B.V., Journals Department, P.O. Box 211, 1000 AE Amsterdam, The Netherlands. Tel: (+31-20) 5803 642, Telex: 18582, Telefax: (+31-20) 5803598, to which requests for sample copies can also be sent. Claims for issues not received should be made within six months of publication of the issues. If not they cannot be honoured free of charge. Readers in the U.S.A. and Canada can contact the following address: Elsevier Science Inc., Journal Information Center, 655 Avenue of the Americas, New York, NY 10010, U.S.A. Tel: (+1-212) 633 3750, Telefax: (+1-212) 633 3990, for further information, or a free sample copy of this or any other Elsevier Science journal.
Advertisements. Advertisement rates are available from the publisher on request.

US mailing notice – Analytica Chimica Acta (ISSN 0003-2670) is published 3 times a month (total 48 issues) by Elsevier Science B.V. (Molenwerf 1, Postbus 211, 1000 AE Amsterdam). Annual subscription price in the USA US$ 3035.75 (valid in North, Central and South America), including air speed delivery. Second class postage paid at Jamaica, NY 11431. *USA Postmasters*: Send address changes to Anal. Chim. Acta, Publications Expediting, Inc., 200 Meacham Av., Elmont, NY 11003. Airfreight and mailing in the USA by Publication Expediting.

ANALYTICA CHIMICA ACTA

An international journal devoted to all branches of analytical chemistry

(Full texts are incorporated in CJELSEVIER, a file in the Chemical Journals Online database available on STN International; Abstracted, indexed in: Aluminum Abstracts; Anal. Abstr.; Biol. Abstr.; BIOSIS; Chem. Abstr.; Curr. Contents Phys. Chem. Earth Sci.; Engineered Materials Abstracts; Excerpta Medica; Index Med.; Life Sci.; Mass Spectrom. Bull.; Material Business Alerts; Metals Abstracts; Sci. Citation Index)

VOL. 296 NO. 1 CONTENTS SEPTEMBER 30, 1994

ANALYTICA CHIMICA ACTA
VOL. 296 (1994)

ANALYTICA CHIMICA ACTA

An international journal devoted to all branches of analytical chemistry
Revue internationale consacrée à tous les domaines de la chimie analytique
Internationale Zeitschrift für alle Gebiete der analytischen Chemie

---

Anal. Chim. Acta, Vol. 296 (1994)

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

0003-2670/94/$07.00

PRINTED IN THE NETHERLANDS

ELSEVIER

Analytica Chimica Acta 296 (1994) 1–20

ANALYTICA CHIMICA ACTA

Aspects of pseudorank estimation methods based on an estimate of the size of the measurement error

N.M. Faber *, L.M.C. Buydens, G. Kateman

Department of Analytical Chemistry, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, Netherlands

Received 2nd March 1994

Abstract

The estimation of the pseudorank of a matrix, i.e., the rank of a matrix in the absence of measurement error, is a major problem in multivariate data analysis. In the practice of analytical chemistry it is often even the only problem. An important example is the determination of the purity of a chromatographic peak. In this paper we discuss three pseudorank estimation methods that make use of prior knowledge about the size of the measurement error. The first method (Method A) is based on the standard errors in the diagonal elements of the row-echelon form of the matrix, the second method (Method B) is based on the eigenvalues of principal component analysis (PCA) and the third method (a *t*-test) is based on the singular values. Methods A and B are modifications of methods that are well known in analytical chemistry. However, these methods cannot provide significance levels for the estimated pseudorank. This holds for the original methods as well as the present modifications. The main reason for introducing these modifications is that in this way relationships are established between the *t*-test and methods that are already known. The aspects that are covered in this paper include the sampling distribution of the test statistic, the number of degrees of freedom to be used in the test, the adequacy of theoretical predictions and the bias that results from random measurement noise. The object of this paper is to demonstrate that using prior knowledge about the size of the measurement error may yield powerful pseudorank estimation methods. This is illustrated by comparing the significance levels obtained by the *t*-test and Malinowski's *F*-test. The *t*-test yields sharper significance levels for experimental data obtained from the literature as well as simulated data. This can be satisfactorily explained by the larger number of degrees of freedom that is employed in this test. The viability of the new *t*-test is supported by a thorough evaluation of the test data by a large number of conventional methods. As a remarkable by-product of the present investigation we find that a plot of the singular values yields a promising graphical pseudorank estimation method. Graphical methods have proved their use in the past in the case that the size of the measurement error is unknown. This new graphical method therefore provides a natural complement to the *t*-test.

Keywords: Principal component analysis; Pseudorank estimation

* Corresponding author.

SSDI 0003-2670(94)00201-V

1. Introduction

The estimation of the pseudorank of a matrix, i.e., the rank of a matrix in the absence of measurement error, is a major problem in multivariate data analysis. It typically arises when highly redundant data produced by modern analytical instruments are to be compressed to a more relevant format. This problem is not likely to be solved in the near future by a better instrumentation since for many applications the main interest is focused on very fast data acquisition, e.g., for the on-line monitoring of an industrial process. Therefore many methods have been proposed over the years to tackle this problem mathematically and the field of analytical chemistry has been especially fertile in producing such methods [1].

A useful and general classification of pseudorank estimation methods is based on the required prior knowledge about size and/or distribution of the measurement error. Methods that require such input are called *parametric* whereas methods that work without this prior knowledge are called *non-parametric*. Parametric methods are only reliable if it is safe to make the necessary assumptions concerning the noise. Otherwise they will not work. Interest in these methods has greatly faded with the introduction of Malinowski's methods based on error functions [2] and Wold's cross-validation [3]. Both methods are non-parametric and make use of a very popular method for the compression and subsequent analysis of multivariate data, i.e., principal component analysis (PCA).

Characteristic for these and many other methods in extensive use today is their inability to establish a *significance level* for the estimated pseudorank. A notable exception is Malinowski's F-test [4,5] which is recently developed from the concept of reduced eigenvalues [6]. It is a parametric method that however works without knowledge about the size of the measurement error. The only assumption being implicitly made is that the residuals are Gaussian distributed. (However, from classical analysis of variance it is well known that even this assumption is not necessarily restrictive [7].) This method is therefore essentially different from a number of very recently introduced methods for which significance levels can be obtained from simulation studies or resampling methods. Examples are canonical correlation [8] which may also be applied if only one data matrix is available [9], an algebraic method based on the Wronskian determinant [10], a non-parametric method based on resampling the zero-component region in the data matrix [11] and a non-parametric method based on the residuals of consecutive PC models [12].

In this paper we will discuss three parametric pseudorank estimation methods that explicitly require knowledge about the size of the measurement error. The first method (Method A) is based on the standard errors in the *diagonal elements of the row-echelon form* of a matrix [13]. The original formulation of this method [13] allows for a complication that critically depends on the nature of the test data matrix, i.e., the data matrix under consideration. Consequently we will propose a possible solution to this problem. The second method (Method B) constitutes a modification of the method of Hugus and El-Awady [14,1]. In the method of Hugus and El-Awady the *eigenvalues* of PCA of the test data matrix are compared to their standard error. In the current modification the eigenvalues of the test data matrix are corrected before they are compared to their standard error. The correction emulates the value that should be expected if an eigenvalue were caused by measurement error and is obtained as the dominant eigenvalue of a 'reference' matrix. In a previous paper we showed that an appropriately sized random matrix can be used as a suitable reference matrix [15]. Furthermore, we make use of the previously derived standard errors in the eigenvalues of PCA [16]. The third method (a t-test) is based on results obtained by Goodman and Haberman [17] for the *singular values* of a matrix. This method comes down to comparing the singular values of the test data matrix to the associated reference value in a similar way as Method B. It will be shown that in all these methods the data matrix is reduced to a 'canonical' form that is suitable for revealing the pseudorank. However, only the t-test is able to give *significance levels* for the estimated pseudorank.

The main reason for also introducing and discussing methods A and B is that in this way a relationship is established between the t-test and methods that are already introduced in analytical chemistry. This should lead to an improved understanding of the working of the proposed t-test and parametric methods in general. The aspects that are covered in this paper include the sampling distribution of the test statistic, the number of degrees of freedom to be used in the test, the adequacy of theoretical predictions and the bias that results from random measurement noise.

It is to be expected that prior knowledge about the size of the measurement error should yield a method that gives sharper significance levels than a method that does not use this extra knowledge. This will be illustrated by comparing the significance levels obtained by the t-test and Malinowski's F-test for literature data as well as simulated data. It is found that Malinowski's F-test gives rather conservative confidence levels. This can be explained by the small number of degrees of freedom employed in this test. The number of degrees of freedom associated with PCA will be further discussed in the Appendix. Support for the viability of the new t-test may also come from a thorough evaluation of the test data by a large number of methods that are currently in use in analytical chemistry. Thus we will also pay considerable attention to the background of these conventional methods.

In the remaining part of this paper we will assume that no data preprocessing has taken place and that the data matrix under consideration is open. Data preprocessing is absolutely necessary if the measurement error is heteroscedastic [18]. The consequences of closure and mean centering for the estimated rank have recently been discussed by Pell et al. [19]. Finally we will assume that pure data for the individual contributing sources are not available. Otherwise the Kalman filter (KF) approach developed at our laboratory [20] provides an excellent parametric pseudorank estimation method.

The following notation will be adapted throughout this paper. Bold upper-case letters will denote matrices, e.g., $\mathbf{M}$. Bold lower-case letters will denote column vectors, e.g., $\mathbf{u}$. Matrix and vector transposition are indicated by a superior 'T', e.g., $\mathbf{M}^T$ and $\mathbf{u}^T$. Italic letters (upper-case as well as lower-case) will denote scalars, e.g., M_{ij} is the element in row i and column j of M. The elements of diagonal matrices, e.g., Θ_{nn} and Λ_{nn}, are denoted by lower case letters, e.g., θ_n and λ_n, where the index indicates the position on the diagonal.

2. Theory

All parametric methods discussed in this section except Malinowski's F-test have in common that they are based on standard errors derived by the method of error propagation. Thus before introducing the parametric methods we will outline the principle behind the derivations. Since the parametric methods to be discussed primarily rely on a dependable estimate of the measurement error in the data matrix $\mathbf{M}$, $\sigma(\mathbf{M})$, we will also briefly discuss how it could be estimated in practice.

2.1. The method of error propagation

The method of error propagation deals with the way in which uncertainties are carried over or propagated from the data points to the estimated parameters. The parameters are written as a function of the data and this function is approximated by a truncated Taylor expansion. The function is expanded around the errorless values and truncation usually proceeds after the linear or quadratic term. It follows that the function should be differentiable in a sufficiently small neighbourhood of the errorless values. The method works well if the measured data points are unbiased estimates of the true data points and the errors are small. The characteristics and limitations of this method are discussed in detail by Moran and Kowalski [21]. We emphasize that for the methods described in this paper the error propagation is carried out to first-order. As a result the derived standard errors can only be expected to be accurate if the standard deviation of the measurement error is small.

2.2. Estimation of the standard deviation of the measurement error in a data matrix

It is evident that in general one should use all the information available in order to ensure that the estimate of the size of the error is accurate. For example, Bubert and Jenett [22] recommend to extend the data matrix obtained from Auger electron spectrometry with sputter cycles in the lower and higher energy regions where no lines can be detected. Essentially the same approach can be followed for chromatographic data. Here the zero-component regions should provide the necessary information [11].

2.3. Pseudorank estimation method based on the standard errors in the diagonal elements of the row-echelon form of the data matrix (Method A)

The oldest methods developed in (analytical) chemistry for pseudorank estimation are based on mathematical definitions of matrix rank in terms of the largest non-zero submatrix [23–25] or the number of non-zero rows in the row-echelon form of the matrix [13]. In this paper we will only consider the method of Wallace and Katz [13] although it should be clear that the main disadvantage connected with the other submatrix methods (excessive computing time) is greatly alleviated by the use of modern high-speed computers.

The method of Wallace and Katz – in the sequel referred to as method A – consists of setting up, in addition to the data matrix **M**, a companion matrix **E**, whose elements E_{ij} are the estimated error of M_{ij}. **M** is reduced to row-echelon form by Gaussian elimination [26] with *complete* pivoting:

$$M'_{ij} = M_{ij} - \frac{M_{i1}}{M_{11}} M_{1j} \tag{1}$$

The elements of **E** are transformed during the reduction of **M** by computing new values based on the propagation of errors in **M**:

$$E'_{ij} = \left[E_{ij}^2 + E_{1j}^2 \left(\frac{M_{i1}}{M_{11}} \right)^2 + E_{i1}^2 \left(\frac{M_{ij}}{M_{11}} \right)^2 + E_{11}^2 \left(\frac{M_{i1} M_{1j}}{M_{11}^2} \right)^2 \right]^{1/2} \tag{2}$$

Complete pivoting is used in order to minimize the rate of propagation of errors. The pseudorank is now determined by investigating for each dimension the following ratio

$$\rho_n(\varepsilon) = \left| \frac{M'_{nn}}{E'_{nn}} \right| \tag{3}$$

Thus given a good estimate of the amount of measurement error it can be established whether a row is zero in the statistical sense by examining the diagonal elements of the transformed matrices **M'** and **E'**. Two decision rules have been found in the literature. Wallace and Katz consider a diagonal element of the reduced data matrix to be significant if it is three times its estimated error whereas Halket [27] proposes a ratio of one.

There is a complication not accommodated for by the simultaneous transformation of the companion matrix. This complication is best illustrated by the following example given by Golub and Van Loan [26]

$$\begin{pmatrix} 1 & -1 & -1 & -1 & . & . \\ 0 & 1 & -1 & -1 & . & . \\ 0 & 0 & 1 & -1 & . & . \\ 0 & 0 & 0 & 1 & . & . \\ . & . & . & . & . & . \end{pmatrix}$$

Suppose for the sake of simplicity that this matrix is the data matrix we start with. Then application of the decision rule that the ratio should exceed, say, k would immediately lead to the conclusion that **M** is full rank if the estimated standard deviation of the measurement error is less than $1/k$. However applying the matrix to the vector whose elements are 1, 1/2, 1/4, 1/8,... shows that the columns of the matrix are nearly dependant because this weighted sum of the

columns is nearly zero. The matrix is very ill-conditioned and a much smaller perturbation then expected may cause the resulting matrix to be singular. As a result we need an independent test for invertibility. Such a test is easily constructed by using a well known result from numerical analysis that says that the smallest singular value of $\mathbf{M}$ is the L_2-norm distance of $\mathbf{M}$ to the set of all rank-deficient matrices [26]. Therefore we propose to compute, in addition to the ratios of Eq. 3, the following index τ:

$$\tau_n = \mathrm{cond}_2(\mathbf{M}'_n)\frac{\|\mathbf{E}'_n\|_2}{\|\mathbf{M}'_n\|_2} \tag{4}$$

where $\mathrm{cond}_2(\bullet)$ is the L_2-condition number, $\|\bullet\|_2$ is the L_2-norm and $\mathbf{M}'_n$ and $\mathbf{E}'_n$ denote the $n \times n$ leading principal submatrix of $\mathbf{M}'$ and $\mathbf{E}'$ respectively. The pseudo rank of $\mathbf{M}'$ should be at least n if $\tau_n < 1$.

2.4. Pseudorank estimation method based on the standard errors in the eigenvalues of PCA (Method B)

PCA is a method that finds new (orthogonal) base vectors that span the space of the matrix in an optimal way. The new base vectors are constructed in such a way as to explain successively the maximum amount of variation in the data. According to Malinowski [1] the data matrix can be reconstructed using only the significant dimensions found by PCA. The remaining dimensions will only contain measurement error. In the terminology of Malinowski the significant dimensions are denoted as *primary* and the remaining ones as *secondary*.

PCA is directly related to the singular value decomposition (SVD) of $\mathbf{M}$. Let s be equal to r or c whichever is smaller. The SVD decomposes $\mathbf{M}$ into a product of three matrices:

$$\mathbf{M} = \mathbf{U}\boldsymbol{\Theta}\mathbf{V}^{\mathrm{T}} \tag{5}$$

where $\mathbf{U}$ is an $r \times s$ orthogonal matrix whose columns $\mathbf{u}_n$ are the left singular vectors, $\mathbf{V}$ is an $s \times c$ orthogonal matrix whose columns $\mathbf{v}_n$ are the right singular vectors and $\boldsymbol{\Theta}$ is an $s \times s$ diagonal matrix with elements $\theta_1 \geq \theta_2 \geq \ldots \geq \theta_s$. These elements, the singular values, are the (positive) square roots of the eigenvalues λ_n of the cross-product matrices $\mathbf{M}\mathbf{M}^{\mathrm{T}}$ and $\mathbf{M}^{\mathrm{T}}\mathbf{M}$:

$$\lambda_n = \mathbf{u}_n^{\mathrm{T}}(\mathbf{M}\mathbf{M}^{\mathrm{T}})\mathbf{u}_n = \mathbf{v}_n^{\mathrm{T}}(\mathbf{M}^{\mathrm{T}}\mathbf{M})\mathbf{v}_n \tag{6}$$

The singular vectors $\mathbf{u}_n$ and $\mathbf{v}_n$ are seen to be eigenvectors of the cross-product matrices. The eigenvalue decompositions of $\mathbf{M}\mathbf{M}^{\mathrm{T}}$ and $\mathbf{M}^{\mathrm{T}}\mathbf{M}$ are often referred to as Q-mode and R-mode PCA respectively.

Hugus and El-Awady [14] have developed a test based on the following expression for the standard errors (R-mode analysis):

$$\sigma(\lambda_n) = \left(\sum_{k=1}^{c} V_{kn}^2 \sum_{l=1}^{c} V_{ln}^2 \sum_{i=1}^{r} \left(M_{ik}^2 \sigma(M_{il})^2 + M_{il}^2 \sigma(M_{ik})^2 \right)(1+\delta_{kl}) \right)^{1/2} \tag{7}$$

Here, δ_{kl} is the well-known Kronecker delta. In their test a PC is considered to be significant if the associated eigenvalue is larger than its standard error. (It should be noted that Bubert and Jenett [22] employ a critical ratio of three.) In a previous paper [16] we showed that the expression of Hugus and El-Awady is incorrect and should be replaced by

$$\sigma(\lambda_n) = 2\lambda_n^{1/2}\sigma(\mathbf{M}) \tag{8}$$

Furthermore, in the test of Hugus and El-Awady the eigenvalues are directly compared to their standard error. The underlying assumption is that an eigenvalue resulting from measurement error should be zero. However, measurement error also contributes to the variation in the data and one should try to take this fact into account. Thus we propose to test the eigenvalues for significance after correcting them for the value that could be the result of chance effect alone. In a previous publication [15] we showed that the secondary eigenvalues of the test data matrix can very well be estimated by the eigenvalues of a pure random matrix. This random matrix should preferably have the same *number of degrees of freedom* as the test data matrix. It is well known that the number of degrees of freedom left *after* extracting the nth PC from the test data is $(r-n)$ $(c-n)$. Since we need the number of degrees of

freedom *before* extracting the nth PC, the number of degrees of freedom of the reference matrix is therefore found to be $(r-n+1)\ (c-n+1)$. Since no parameters can be estimated from a random matrix, the number of degrees of freedom automatically equals the total number of data points. Thus the *size* of the reference matrix should be $(r-n+1) \times (c-n+1)$. This leads to the following correction procedure before testing the significance of the eigenvalues. The first eigenvalue of the test data matrix is corrected by subtracting the first eigenvalue of an $r \times c$ random matrix, the second eigenvalue of the test data matrix is corrected by subtracting the first eigenvalue of an $(r-1) \times (c-1)$ random matrix, and so on. In general, the correction for the nth eigenvalue under scrutiny is found as the dominant eigenvalue of an $(r-n+1) \times (c-n+1)$ random matrix. Accurate reference values are obtained by averaging the eigenvalues of a large number of random matrices. Tables with accurate reference eigenvalues have, for example, been published by Mandel [28] and are easily extended by Monte Carlo (MC) simulations [15]. As a modification to the method of Hugus and El-Awady – in the sequel referred to as method B – we therefore propose to examine the following ratio

$$\rho_n(\lambda) = \frac{\lambda_n - \lambda_{n,\mathrm{ref}}}{\sigma(\lambda_n)} \tag{9}$$

where $\lambda_{n,\mathrm{ref}}$ denotes the nth reference eigenvalue. It is to be expected that only the ratio associated with the last significant PC (complete model) should be consistent with the ratio found by equation (3) since here the data reduction proceeds in an entirely different way. Furthermore these values should only agree as long as the reference values $\lambda_{n,\mathrm{ref}}$ are negligible since we do not apply a correction in Method A.

In a previous publication we showed that the standard errors predicted by Eq. 8 overestimate the true standard errors [16]. The overestimate is negligible for the large primary eigenvalues but may be notable for the small ones. (The true standard errors were estimated by MC simulations.) Thus especially for the high-numbered primary eigenvalues the ratios calculated by Eq. 9 are an underestimate of the true ratios and consequently method B is expected to give conservative estimates of the pseudorank. This will, however, only constitute a problem if for the specific application at hand a false negative declaration, i.e., a primary PC is deemed non-significant, causes more harm than a false positive declaration, i.e., a secondary PC is deemed significant. It should be noticed that for many PCA based methods e.g., iterative target testing factor analysis (ITTFA) the incomplete model will lead to erroneous results. In that case the conservative estimate could still provide a useful lower bound.

There is still one point we want to discuss with respect to Eq. 9. Goodman and Haberman [17] have shown that the eigenvalues of PCA are biased as a result of random measurement noise. Although they only give the relevant expression for a one-dimensional PC model their result is easily generalized to an n^*-dimensional PC model by invoking Malinowski's error functions. We will show in a future publication that the bias in eigenvalue λ_n can be predicted as $\mathrm{bias}(\lambda_n) = (r + c - n^*)\sigma(\mathbf{M})^2$. The adequacy of this *theoretical* prediction, however, depends on the signal-to-noise ratio (*SNR*). For low *SNR* this prediction is not accurate enough in order to construct a confidence interval for the eigenvalue in absence of noise. It will be shown in this paper that an accurate *empirical* 'bias correction' is always provided by the reference eigenvalue $\lambda_{n,\mathrm{ref}}$.

2.5. Pseudorank estimation method based on the standard errors in the singular values of SVD (t-test)

Methods A and B are both characterized by comparing a ratio with a *fixed critical value* (one or three). This procedure is not in the spirit of hypothesis testing in statistics. In statistics a hypothesis is formulated about a test statistic and the validity of the hypothesis is derived from the sampling distribution of the test statistic, the appropriate number of degrees of freedom and a certain (predetermined) significance level. The number of degrees of freedom and thus also the critical value of the test statistic should depend on the test data at hand. It is extremely difficult

to devise such a procedure for the statistics given by Methods A and B because their sampling distribution is unknown. For Method B this is caused by the fact that the numerator and denominator in Eq. 9 are not independent. In general it is possible to infer the sampling distribution of a statistic from MC simulations. However, we will not pursue this line because it is possible to derive a significance test [1] from Method B in a straightforward manner without resorting to (additional) simulations.

In the case that the variance in an estimated parameter depends on the parameter itself, the standard procedure in statistics consists of 'stabilizing' the variance by transforming the parameter in such a way that the transformed parameter is independent of the associated variance [29]. It immediately follows from Eq. 8 that the standard error in the singular values is constant and equal to $\sigma(\mathbf{M})$ (see also [17,16]). Thus the stabilized parameters are simply given by the singular values. [2] Furthermore, the singular values are linear functions of the data. Thus given Gaussian distributed measurement errors, the sampling distribution of the singular values is also given by the Gaussian distribution [17]. The assumption of Gaussian distributed noise is often not justified in practice. However, it is well known that deviations from normality can be neglected if the number of observations (i.e., in our case the number of matrix elements) is sufficiently large. As a general guide, a number of at least 50 can be considered to be large enough [30,31]. If we have, in addition, some prior knowledge which suggests that the distribution of the matrix elements resembles the Gaussian in some way, e.g., symmetry, then this would allow us to regard a smaller number as large enough. It is interesting to note that very recently Booksh and Kowalski [32] have demonstrated a considerable 'normalizing' effect for the generalized rank annihilation method (GRAM), a calibration method that is based on PCA.

It is possible to examine the statistic that is obtained by simply replacing the eigenvalues in Eq. 9 by the corresponding singular values. It is easily shown that the resulting statistic is always larger. However, the Eq. 9 statistic does not take into account that the eigenvalues of the reference matrix vary in a similar way as the eigenvalues of the test data matrix. Thus an additional modification is necessary before the test statistic is complete. Approximating the standard errors in both singular values by $\sigma(\mathbf{M})$ yields as a possible test statistic

$$\rho_n(\theta) = \frac{\theta_n - \theta_{n,\mathrm{ref}}}{\sqrt{2}\,\sigma(\mathbf{M})} \tag{10}$$

If the test data matrix is large enough the ratio given by Eq. 10 is approximately distributed as Student's t independent of the distribution of the measurement error (large sample assumption). The number of degrees of freedom associated with this test statistic is determined by the size of the reference matrix, i.e., $\nu = (r - n + 1)(c - n + 1)$. The ratio of Eq. 10 is designed to test the null-hypothesis

$$\mathrm{H}_0: \theta_n = \theta_{n,\mathrm{ref}}$$

against the alternative hypothesis (a one-sided test)

$$\mathrm{H}_1: \theta_n > \theta_{n,\mathrm{ref}}$$

Hence we reject H_0 at the α level of significance if the realization of $\rho_n(\theta)$ is greater than or equal to the tabulated $t_\nu(1-\alpha)$. Analogous to the F-test of Malinowski (that is to be briefly discussed next) the proposed significance test starts at the high-numbered PCs. First, the singular value with $n = s$ is tested against the singular value of an $(r - s + 1) \times 1$ matrix. If the calculated ρ is less than the tabulated t, the singular value under test is added to the secondary set. Next, we examine the ratio for $n = s - 1$, and so

[1] We make a distinction here between pseudorank estimation methods: only methods that are able to provide a significance level are denoted as significance tests.

[2] It is important to note that Lawson and Hanson [30] give deterministic error bounds for the singular values of a perturbed matrix. If M is perturbed by a matrix E, an upper bound for the error in a singular value is given as the largest singular value of E. It should be clear that these error bounds are not statistical in nature (no assumptions are made about the elements in E) and therefore not accurate enough for the purpose of this paper.

on. The process of testing and adding to the null set is repeated until the calculated ρ exceeds the tabulated t.

It is seen that the variability of both singular values is taken into account in Eq. 10 in a pessimistic fashion since the standard error in the singular values of the reference matrix is smaller than $\sigma(\mathbf{M})$. (An overestimate by a factor of two should be expected [16].) The conservative character of the proposed t-test should guard the user against the consequences of, for example, violating the large sample assumption in practice. However, a thorough evaluation should demonstrate whether the test still has enough discriminating power or that it is useless.

2.6. Reduced eigenvalues and Malinowski's F-test

Malinowski discovered that the following function of the eigenvalues of PCA

$$REV_n = \frac{\lambda_n}{(r-n+1)(c-n+1)} \tag{11}$$

is constant for the secondary PCs [6]. Using these 'reduced' eigenvalues (*REV*s) an F-test was developed [4,5]:

$$F(\nu_1,\nu_2) = \frac{\sum_{j=n+1}^{s}(r-j+1)(c-j+1)}{(r-n+1)(c-n+1)} \times \frac{\lambda_n}{\sum_{j=n+1}^{s}\lambda_j} \tag{12}$$

with degrees of freedom $\nu_1 = 1$ and $\nu_2 = s - n$. The procedure consists of testing λ_n against the pool of $(s-n)$ remaining eigenvalues. One starts with eigenvalue λ_{s-1} and works backwards through the list of eigenvalues. If an eigenvalue is found to be insignificant, it is pooled with the remaining error eigenvalues, the counter n is lowered by one and the next eigenvalue is considered. It is seen that the number of degrees of freedom is taken as the number of eigenvalues involved in the test. The number of degrees of freedom associated with PCA is further discussed in the Appendix. It was found that in general testing on the 5% level tends to underestimate whereas testing on the 10% level tends to overestimate the pseudorank of the matrix [4].

3. Experimental

The methods discussed in the preceding section are evaluated by analyzing data obtained from the literature as well as simulated data.

3.1. Literature data

Investigating data from the literature in order to test a new method is useful since these data should be readily available for other researchers thus making the present results reproducible. Some of the data sets considered in this section are based on computer simulations. They should

Table 1
Characterization of literature test data

Data	FIAL80 [a]	GUTM68 [b]	HAVE85 [c]	MALI82 [d]	RITT76 [e]	WEIN70 [f]	WEIN71 [g]
Size	20 × 5	9 × 5	10 × 8	20 × 10	17 × 7	14 × 9	22 × 6
Pseudorank	3	2	3	5	2	3	2
$\hat{\sigma}(\mathbf{M})^h$	0.45	0.049	0.013	0.59	0.14	0.65	0.56

[a] Computer simulated powder diffraction intensities [33].
[b] Half-wave potentials of metal ions in various solvents [34]. Two PCs are deemed significant by the tests discussed below while Howery reports that three PCs are needed to adequately reproduce the data [35].
[c] Potentiometric data [36].
[d] Simulated mass spectra [37].
[e] Mass spectra for which the row corresponding to contaminating nitrogen is deleted [38].
[f] Chemical shifts in various solvents [39].
[g] Chemical shifts in various solvents [40]. Two significant PCs are found by the tests discussed below while Weiner et al. report that three PCs are needed for the reproduction of the data within the experimental error (0.5Hz).
[h] Estimated as $\hat{\sigma}(\mathbf{M}) = \Sigma_{j=n^*+1}^{s}\lambda_j/(r-n^*)(c-n^*)$ where n* denotes the pseudorank.

be particularly useful for the purpose of this paper since simulated data can be expected to be well-behaved with respect to the 'measurement' error. The selection of these data sets is primarily based on the following consideration: if we want to discover whether a procedure has failed to indicate the correct pseudorank, we should apply it to data for which a reliable estimate for the pseudorank is available. In Table 1 we have described the literature test data that meet this requirement. The selected data sets cover a wide range of experimental techniques. The first row gives the data matrix under investigation. In the second row the size of the matrices ($r \times c$) is given. It is seen that in general the number of data points is rather small. The obvious reason is that it is not practical to publish large data matrices in journal articles. The number of data points ranges between 45 for data set GUTM68 and 200 for data set MALI82. The unfavourable size of some of the data sets may lead to a small number of degrees of freedom and critically influence the outcome of the proposed t-test. In the third row we have listed the estimated pseudorank n^*. The quoted estimate is found by a large number of methods from which the following are the most widely used: cross-validation [2], reduced eigenvalues [6], imbedded error function [3], indicator function [3] and the eigenvalue ratio [41–43]. With two exceptions, i.e., GUTM68 and WEIN70, the determined pseudorank agrees with the value reported earlier. (These exceptions show that data reproduction – formerly a popular method – only provides a reliable pseudorank if $\sigma(\mathbf{M})$ can be estimated accurately.) In the last row we give the estimated standard deviation of the measurement error that is based on the residuals of the correct PC model. These values will be used as input for the parametric methods since for many of these data sets a reliable estimate independent of the data is not available. Exceptions are formed, for example, by data sets WEIN70 and WEIN71 for which a measurement error of 0.5 Hz is reported. One of these datasets, i.e., WEIN70, will be investigated using both values, i.e., 0.5 and 0.65, in order to evaluate the robustness of the parametric methods with respect to an inaccurate estimate of $\sigma(\mathbf{M})$.

3.2. Simulated data

In a previous investigation [16] we constructed a dilution series by simulating a number of multicomponent systems for which the signal of one component was systematically lowered. This way the usefulness of theoretical results like Eq. 8 was tested. A three-component LC-UV data matrix was simulated by multiplying Gaussian functions and the (normalized) UV spectra of adenine, cytidine and guanine taken from the work of Zscheile et al. [44]. The size of the resulting matrices was 36×36. Artificial Gaussian noise with standard deviation 0.5 mAU was added. In this paper we will restrict the discussion to dilutions where theoretical predictions should be expected to start to break down. For these dilutions the peakheights of adenine and guanine are 1000 mAU while the peakheight of cytidine is only 6, 5 and 3 mAU respectively. The resulting data sets are denoted as EXP1, EXP2 and EXP3. Details about the simulations are summarized in Table 2. A plot of data matrix EXP1 is shown in Fig. 1. The unfavourable ratio of peakheights and high overlap of the pure component responses (in both instrumental modes) are apparent. These data

Table 2
Characterization of simulated test data [a]

	Adenine	Cytidine	Guanine
Peak positions, μ	9	18	27
Standard deviation peaks, σ	5	5	5
Peakheights h for EXP1 in mAU	1000	6	1000
Peakheights h for EXP2 in mAU	1000	5	1000
Peakheights h for EXP3 in mAU	1000	3	1000
Number of spectra		36	
Number of wavelengths		36	
$\sigma(\mathbf{M})$ in mAU		0.5	

[a] The elements of the data matrices are generated as $M_{ij} = \sum_{k=1}^{K} C_{ik} S_{jk} + N(0,\sigma(\mathbf{M}))$ where K is the number of components (i.e., 3 in our case), C_{ik} is the value of the elution profile of component k at time i, S_{jk} denotes the absorbance of component k at wavelength j and $N(0,\sigma(\mathbf{M}))$ is a normally distributed number with zero mean and standard deviation $\sigma(\mathbf{M})$. The elements of the elution profiles are calculated as $C_{ik} = h_k \cdot \exp\,[-1/2(i-\mu_k)^2/\sigma_k^2]$ where the symbols have the meaning as indicated above.

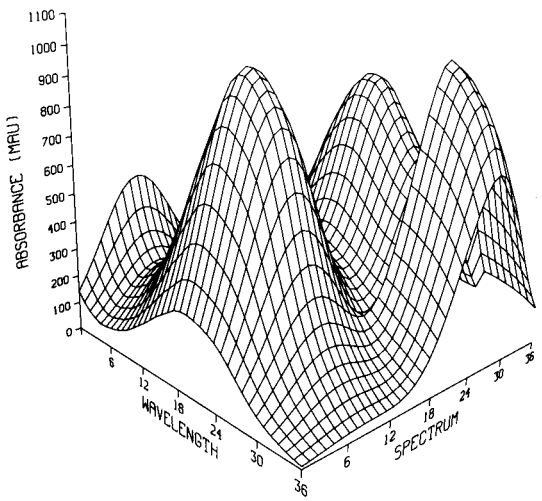


Fig. 1. Simulated three-component LC-UV data matrix EXP1.

matrices should therefore constitute an interesting test case for the methods discussed in this paper. Only data matrices EXP2 and EXP3 have been analyzed before in [16]. For data set EXP2 it was found that the prediction of the standard error in the eigenvalue was excellent for the two main components but wrong (too high) by 20% for the dilute component. However, given the low value of the relevant eigenvalue ratio ($\lambda_3/\lambda_4 = 1.77$) this result was seen as very promising. For data set EXP3 the true standard error for the dilute component was overestimated by 85%. Additional results showed that the third dimension for this data set primarily consists of noise.

3.3. Calculations

The computer program is written in Fortran77 and all calculations are performed in double precision arithmetic on a HDS-EX60 mainframe computer. Built-in subroutines and functions of the IMSL library [45] are used. The SVD of the data sets is performed by subroutine DLSVRR. Significance levels for the F-test are calculated from the output of function DFDF as $\%\alpha = 100 \times (1 - \mathrm{DFDF}(F,\nu_1,\nu_2))$. Occasionally very large F-values may cause a floating point underflow in the evaluation of DFDF. This problem is solved by setting $\%\alpha$ to zero if $F > 100$. Significance levels for the t-test are calculated from the output of function DTDF as $\%\alpha = 100 \times (1 - \mathrm{DTDF}(t,\nu))$.

4. Results and discussion

4.1. Literature data

Before presenting the results for the parametric methods introduced in the theoretical section, we want to discuss the performance of three conventional pseudorank estimation methods. These methods are based on (functions of) the eigenvalues and are often used graphically. Additionally, we will show that the singular values are a promising alternative for these conventional methods.

Table 3
Eigenvalues of PCA for literature test data. The numbers marked in bold indicate the estimated pseudorank

n	FIAL80	GUTM68	HAVE85	MALI82	RITT76	WEIN70	WEIN71
1	2.60×10^5	1.89×10^2	9.67×10	2.56×10^5	2.43×10^3	1.02×10^7	4.90×10^5
2	1.87×10^4	$\mathbf{2.38 \times 10^{-1}}$	4.61×10^{-1}	2.12×10^4	$\mathbf{3.34 \times 10^2}$	4.77×10^2	$\mathbf{5.54 \times 10^3}$
3	$\mathbf{2.53 \times 10^3}$	3.17×10^{-2}	$\mathbf{4.10 \times 10^{-2}}$	1.77×10^4	7.60×10^{-1}	$\mathbf{9.55 \times 10}$	1.13×10
4	3.53	1.51×10^{-2}	3.30×10^{-3}	1.02×10^4	3.35×10^{-1}	1.16×10	8.52
5	3.43	3.05×10^{-3}	1.87×10^{-3}	$\mathbf{2.42 \times 10^3}$	2.56×10^{-1}	8.79	3.73
6			3.40×10^{-4}	1.00×10	1.20×10^{-1}	3.94	1.54
7			2.98×10^{-4}	5.87	5.76×10^{-2}	1.56	
8			1.15×10^{-6}	5.20		1.36	
9				3.58		0.70	
10				1.33			

Table 4
Indicator function for literature test data. The numbers marked in bold indicate the estimated pseudorank

n	FIAL80	GUTM68 ($\times 10^{-2}$)	HAVE85 ($\times 10^{-3}$)	MALI82	RITT76 ($\times 10^{-2}$)	WEIN70 ($\times 10^{-1}$)	WEIN71
1	1.02	0.56	1.74	0.209	5.04	0.36	0.284
2	0.72	**0.48**	0.78	0.215	**0.54**	0.23	**0.033**
3	**0.10**	0.80	**0.43**	0.194	0.67	**0.16**	0.051
4	0.41	1.84	0.50	0.126	1.03	0.19	0.086
5	–	–	0.51	**0.020**	1.81	0.23	0.264
6			0.97	0.028	5.82	0.33	–
7			0.34	0.046	–	0.68	
8			–	0.088		2.24	
9				0.258		–	
10				–			

In Table 3 the eigenvalues of PCA are listed. Using the simple argument that primary eigenvalues should be relatively large one easily arrives at the values for the pseudorank given in Table 1. (It is interesting to note that support for the two-dimensional model for WEIN71 comes from comparing the eigenvalue pattern with that for WEIN70.) However, in general the jump between primary and secondary eigenvalues is not so prominent and several methods have been introduced to aid in the decision making process.

The results for two of these methods, the indicator function [2] and the reduced eigenvalues [6], are shown in Tables 4 and 5. Malinowski has postulated that the indicator function should exhibit a minimum for the true dimension of the data matrix [3]. Thus we arrive at the same pseudorank estimates using the indicator function. It is often stated as a disadvantage of this method that the minimum is shallow. This is also the case here for data set WEIN70 but the point is that we can not quantify such a statement without knowing the standard errors in the indicator function. It is possible to derive these standard errors from Eq. 8. This could lead to theoretical evidence for the postulated minimum. However, there is already considerable experimental evidence in the literature that indicates that the minimum is significant (the method is very successful) and we will not pursue this line. According to Malinowski the reduced eigenvalues should be constant for the secondary PCs while the values for the primary PCs should be larger. In another publication [15] we have shown by simulations of random matrices that it depends on the ratio of the rows and the columns of the matrix, the so-called divergence coefficient, whether the reduced eigenvalues are (approximately) constant. This numerical result is explained by a theoretical result of multivariate statistics for the joint probability density function (pdf) of the eigenvalues of a random matrix. (The joint pdf gives the probability of finding *any* eigenvalue in a certain range.) The shape of the joint pdf depends primarily on the divergence coefficient of the matrix. The consequences for the reduced eigenvalues are that different patterns should be expected depending on the value of the divergence coefficient. We have found, for example, that a divergence coefficient of (approximately) one leads to a large probability of finding a very small eigenvalue. This is confirmed here for data set HAVE85. (This very small eigenvalue is believed to be the reason for the dip in the indicator function.) For the other data sets the divergence coefficient ranges between 1.5 and 4. This is the range where the reduced eigenvalues of random matrices were shown to be approximately constant [15]. Thus for the data sets under investigation the method

[3] Occasionally a local minimum is observed as already noted by Malinowski. For the present data sets we find a local minimum for HAVE85 and MALI82. The local minimum for HAVE85 is discussed later. The local minimum for MALI82 is caused by the small ratio between the second and third eigenvalue.

Table 5
Reduced eigenvalues for literature test data. The numbers marked in bold indicate the estimated pseudorank

n	FIAL80	GUTM68	HAVE85	MALI82	RITT76	WEIN70	WEIN71
1	2.60×10^{3}	4.20	1.21	1.28×10^{3}	2.05×10	8.13×10^{4}	3.71×10^{3}
2	2.45×10^{2}	$\mathbf{7.43 \times 10^{-3}}$	7.32×10^{-3}	1.24×10^{2}	**3.48**	4.59	$\mathbf{5.27 \times 10}$
3	$\mathbf{4.69 \times 10}$	1.51×10^{-3}	$\mathbf{8.54 \times 10^{-4}}$	1.23×10^{2}	1.01×10^{-2}	**1.14**	0.14
4	0.10	1.26×10^{-3}	9.42×10^{-5}	8.60×10	0.60×10^{-2}	0.18	0.15
5	0.21	0.61×10^{-3}	7.80×10^{-5}	$\mathbf{2.52 \times 10}$	0.66×10^{-2}	0.18	0.10
6			2.27×10^{-5}	0.13	0.50×10^{-2}	0.11	0.09
7			3.72×10^{-5}	0.11	0.52×10^{-2}	0.07	
8			0.04×10^{-5}	0.13		0.10	
9				0.15		0.12	
10				0.12			

should work and this is confirmed by the results given in Table 5.

In Table 6 we have listed the singular values of the literature test data. The reason for showing the singular values is as follows. In the theoretical section it is argued that the error in the singular values is constant (and equal to the original measurement error). Contrary to the primary singular values the secondary singular values only consist of measurement error. Thus it is to be expected that the distance between the secondary singular values is bounded by the size of the measurement error whereas the distance between the primary singular values is also affected by the size of the systematic variation. Since the size of the systematic variation should be larger than the size of the measurement noise for the data to be analyzable at all, it seems logical to simply inspect the distance between the singular values. It is seen that for all data sets the distance between the secondary singular values is of the same order of magnitude. Furthermore, it is easily verified that the order of magnitude is given by the standard deviations in Table 1. This leads to the conclusion that plotting the singular values yields a promising graphical pseudorank estimation method: the singular values should (approximately) lie on a straight line for the secondary PCs whereas for the primary PCs the curve deviates upwards. It is worth mentioning that the logarithm of the eigenvalues is reported to yield a straight line for the secondary eigenvalues [46]. However, a systematic evaluation [15] showed that a straight line should only be expected for the low-numbered secondary PCs. This numerical result has now been explained since the logarithm and the square root may transform the eigenvalues in a similar way over a restricted range (they are both weak transformations) [4]. It should be kept in mind that we are using qualitative arguments here which we will try to quantify by means of the proposed t-test on the singular values. It is evident that such a quantification should be based on an estimate of the size of the measurement error.

The working and use of the parametric methods is illustrated by discussing the results in detail for data set WEIN70. This data set has a large number of degrees of freedom, i.e., $(r - n^*)(c - n^*) = (14 - 3)(9 - 3) = 66$, and has already been treated extensively in the literature (see e.g., [4] and [43]). Additional results (not shown here) obtained for the χ^2-test as well as the number of 3σ-misfits [14] are further evidence for the suitability of this data set for the evaluation of pseudorank estimation methods. The results of Methods A and B are summarized in Table 7. It should be noted that all calculations are performed with the reported value for the standard deviation of the measurement error (0.5 Hz) which

[4] In multivariate statistics standard errors in the eigenvalues as a result of sampling errors have been derived (see [16] for a detailed discussion). These standard errors also depend on the size of the eigenvalues but now Equation (8) is no longer appropriate. Now the stabilizing transform is given by the logarithm. This motivates inspecting the logarithm of the eigenvalues in the case that sampling errors play a role.

is 30% smaller that the standard deviation estimated from the residuals of the 3-dimensional PC model (0.65 Hz). The first column gives the PC under investigation. The second and third column give the diagonal elements of the reduced matrices, $\mathbf{M}'$ and $\mathbf{E}'$. The resulting ratio calculated from Eq. 3 is given in the next column. Four PCs are deemed significant if we use the (fixed) critical value of three. It is clear that three significant PCs would have been found if the input standard

Table 6
Singular values for literature test data. The numbers marked in bold indicate the estimated pseudorank

n	FIAL80	GUTM68	HAVE85	MALI82	RITT76	WEIN70	WEIN71
1	5.10×10^2	1.374×10	9.835	5.06×10^2	4.93×10	3.20×10^3	7.00×10^2
2	1.37×10^2	**0.488**	0.679	1.46×10^2	**1.83×10**	2.18×10	**7.44×10**
3	**5.03×10**	0.178	**0.202**	1.33×10^2	0.87	**9.77**	3.35
4	1.88	0.123	0.057	1.01×10^2	0.58	3.41	2.92
5	1.85	0.055	0.043	**4.92×10**	0.51	2.97	1.93
6			0.018	3.16	0.35	1.98	1.24
7			0.017	2.42	0.24	1.25	
8			0.001	2.28		1.17	
9				1.89		0.84	
10				1.15			

Table 7
Results of method A and B for literature data matrix WEIN70. The numbers marked in bold indicate the estimated pseudorank

n	Method A			Method B			
	diag($\mathbf{M}'$)	diag($\mathbf{E}'$) [a]	$\rho(\varepsilon)$	λ	λ_{ref} [a]	$\sigma(\lambda)$	$\rho(\lambda)$
1	4.56×10^2	0.50	9.11×10^2	1.02×10^7	9.28	3.20×10^3	3.20×10^3
2	1.49×10	0.55	2.71×10	4.77×10^3	8.34	2.18×10	2.15×10
3	7.11	0.83	8.61	9.55×10	7.38	9.77	**9.02**
4	−2.99	0.98	**3.04**	1.16×10	6.42	3.41	1.53
5	−2.22	1.24	1.79	8.79	5.47	2.97	1.12
6	−2.01	0.90	2.23	3.94	4.50	1.98	−0.28
7	1.55	1.74	0.89	1.56	3.55	1.25	−1.59
8	1.55	2.21	0.70	1.36	2.56	1.17	−1.03
9	−0.15	1.82	0.08	0.70	1.49	0.84	−0.94

[a] Calculated with $\hat{\sigma}(\mathbf{M}) = 0.50$.

Table 8
Results of t-test and F-test for literature data matrix WEIN70. The numbers marked in bold indicate the estimated pseudorank

n	t-test						F-test				
	Θ	Θ_{ref} [a]	$\sigma(\Theta_{ref})$	t	ν	%α	REV	F	ν_1	ν_2	%α
1	3.20×10^3	3.04	0.26	4.52×10^3	126	0.0	8.13×10^4	5.20×10^4	1	8	0.0
2	2.18×10	2.88	0.27	2.68×10	104	0.0	4.59	1.04×10	1	7	1.4
3	9.77	2.70	0.27	1.00×10	84	**0.0**	1.14	7.96	1	6	**3.0**
4	3.41	2.52	0.28	1.26	66	10.6	0.18	1.40	1	5	29.0
5	2.97	2.32	0.29	0.92	50	18.1	0.18	1.86	1	4	24.4
6	1.98	2.10	0.29	−0.17	36	56.7	0.11	1.33	1	3	33.2
7	1.25	1.86	0.31	−0.86	24	80.1	0.07	0.63	1	2	51.0
8	1.17	1.57	0.32	−0.57	14	71.1	0.10	0.83	1	1	53.0
9	0.84	1.17	0.35	−0.47	6	67.3	0.12	–	–	–	–

[a] Calculated with $\hat{\sigma}(\mathbf{M}) = 0.50$.

deviation would have been larger by only 2%. In the next two columns we give the eigenvalues of PCA and the appropriate reference eigenvalues. The standard error in the eigenvalues predicted from Eq. 8 is given in the next column. The resulting ratio calculated from Eq. 9 is given in the last column. Due to the considerable correction by the reference eigenvalue we now find only three significant PCs. It is clear that the outcome would only change if we would underestimate $\sigma(\mathbf{M})$ by a factor of two. It is seen that Method B is more robust with respect to errors in the input value of $\sigma(\mathbf{M})$ than Method A.

The results for the t-test and Malinowski's F-test are given in Table 8. The first three columns give the PC under consideration, the test singular value and the reference singular value respectively. In the next column we have listed the standard error in the reference singular values. Over the whole range this value is (much) smaller than $\sigma(\mathbf{M}) = 0.50$. This means that the t-values listed in the next column are conservative as indicated before. The degrees of freedom to be used in the test are given in the next column. The resulting significance levels clearly indicate the presence of three significant PCs and nearly a fourth one on the 10% level. However, the input value of $\sigma(\mathbf{M})$ is rather small compared to the value estimated from the residuals of the correct PC model and these results are therefore promising. (It follows that the robustness found earlier for Method B is misleading.) The results for the F-test are shown next. They have already been discussed in detail by Malinowski [4]. Three PCs are deemed significant at the 5% level of significance. Thus both tests agree about the true dimension of the data. However, there is a sharp contrast with respect to the significance levels. The significance level supplied by the t-test is essentially zero while the F-test attaches an uncertainty of 3% to the model. The reason for the discrepancy is that the t-test uses much more degrees of freedom than the F-test (see Appendix). Although the t-test should also give conservative estimates of the significance of a PC model it is seems to be less conservative than the F-test. Using the extra knowledge about the measurement error has indeed led to a sharper significance level. In the case that an important decision has to be made based on the result of a significance test the improvement obtained by the t-test may be appreciable.

In Table 9 we have summarized the results of the various parametric methods for the last significant PCs. It should be noted that for all data sets the value of $\sigma(\mathbf{M})$ taken from Table 1 is inserted in the relevant expressions. (As a result the first secondary PC is deemed non-significant in all cases.) The first column lists the data set under consideration. The next two columns gives the ratios calculated from Eqs. 3 and 9, respectively. It is seen that the agreement is very well in all cases. This result is remarkable since the evaluation of Eq. 9 involves the correction by a reference value obtained from random matrices. The next three columns list the results for the t-test. The t-values are extremely high in all cases leading to very small significance levels. The last three columns summarize the results for the F-test. In all cases the F-values are larger than the

Table 9
Results of various pseudorank estimation methods [a] for last significant PC of literature test data

Data	Method A	Method B	t-test			F-test		
	$\rho(\varepsilon)$	$\rho(\lambda)$	t	ν	$\%\alpha$	F	ν_2	$\%\alpha$
FIAL80	4.03×10	5.54×10	7.49×10	54	0.0	3.37×10^2	2	0.3
GUTM68	4.37	4.18	4.24	32	8.9×10^{-2}	5.66	3	9.8
HAVE85	7.19	7.19	7.89	48	0.0	1.25×10	5	1.7
MALI82	4.16×10	4.16×10	5.51×10	96	0.0	1.99×10^2	5	3.2×10^{-3}
RITT76	6.01×10	6.38×10	1.22×10^2	96	0.0	4.66×10^2	5	4.0×10^{-4}
WEIN70	6.58	6.52	6.79	84	0.0	7.96	6	3.0
WEIN71	6.26×10	6.64×10	8.98×10	105	0.0	4.00×10^2	4	3.7×10^{-3}

[a] Evaluated with $\hat{\sigma}(\mathbf{M})$ given in Table 1.

Table 10
Reference singular values [a] for literature test data. The numbers marked in bold indicate the estimated pseudorank for the test data. Reference values that are higher than the test values (see Table 3) are underlined

n	FIAL80	GUTM68	HAVE85	MALI82	RITT76	WEIN70	WEIN71
1	2.72	2.180	0.068	4.12	0.87	3.95	3.61
2	2.54	**0.196**	0.063	3.94	**0.82**	3.74	**3.41**
3	**2.34**	0.171	**0.058**	3.77	0.77	**3.52**	3.20
4	2.11	0.142	0.053	3.58	0.71	3.28	2.96
5	1.79	0.104	0.047	**3.37**	0.65	3.02	2.67
6			0.040	3.17	0.57	2.74	2.28
7			0.032	2.92	0.47	2.42	
8			0.021	2.66		2.04	
9				2.34		1.52	
10				1.91			

[a] Calculated with $\hat{\sigma}(\mathbf{M})$ given in Table 1.

t-values (this is not a general rule). However, as a result of the much smaller number of degrees of freedom the resulting significance levels are much larger than those found by the t-test. This is without consequence for the estimated pseudorank except for the (extremely small) data set GUTM68. For this data set Malinowski's F-test gives one significant PC at the 5% level and two significant PCs at the 10% level. This result is in agreement with the conclusion of Malinowski about the tendency to underestimate or overestimate the pseudorank at the 5 and 10% level respectively.

In Table 10 we give the reference singular values for the literature data matrices. Using these numbers it is possible to reproduce the results for the t-test given in Table 9. In another publication [15] we show that the secondary eigenvalues of the test data matrix are approached from above by the eigenvalues of the reference matrix. As a result we find that the test singular values tend to be smaller than their reference values. The differences are small, however, especially for the first secondary PC. The tendency of the reference values to be too large contributes to the conservative character of the t-test.

For the literature data sets we have found an excellent agreement between the results of the F-test and the t-test with respect to the estimated pseudorank. The reason is that these data sets are not selected for their discriminating ability. It is possible to investigate the difference in sensitivity in more detail by performing the following 'Gedankenexperiment' on the data. A hypothetical matrix is constructed from the test matrix by lowering the last significant singular value while keeping all other things fixed. The size of the hypothetical singular value is determined by the significance level it would give for a certain test. It is to be expected that in order to find a predetermined significance level, say 1%, the size

Table 11
Comparison of actual and critical values for the eigenvalue ratio (ER) at different confidence levels

Data	ER (actual value)	t-test			F-test		
		ER (1%)	ER (5%)	ER (10%)	ER (1%)	ER (5%)	ER (10%)
FIAL80	7.17×10^2	4.22	3.29	2.84	2.10×10^2	3.95×10	1.82×10
GUTM68	7.51	4.20	3.09	2.59	4.52×10	1.34×10	7.35
HAVE85	1.24×10	3.19	2.41	2.05	1.62×10	6.57	4.04
MALI82	2.42×10^2	2.85	2.26	1.98	1.98×10	8.04	4.94
RITT76	4.39×10^2	2.23	1.76	1.54	1.53×10	6.23	3.82
WEIN70	8.21	2.80	2.19	1.91	1.42×10	6.18	3.90
WEIN71	4.92×10^2	2.48	1.99	1.75	2.61×10	9.48	5.58

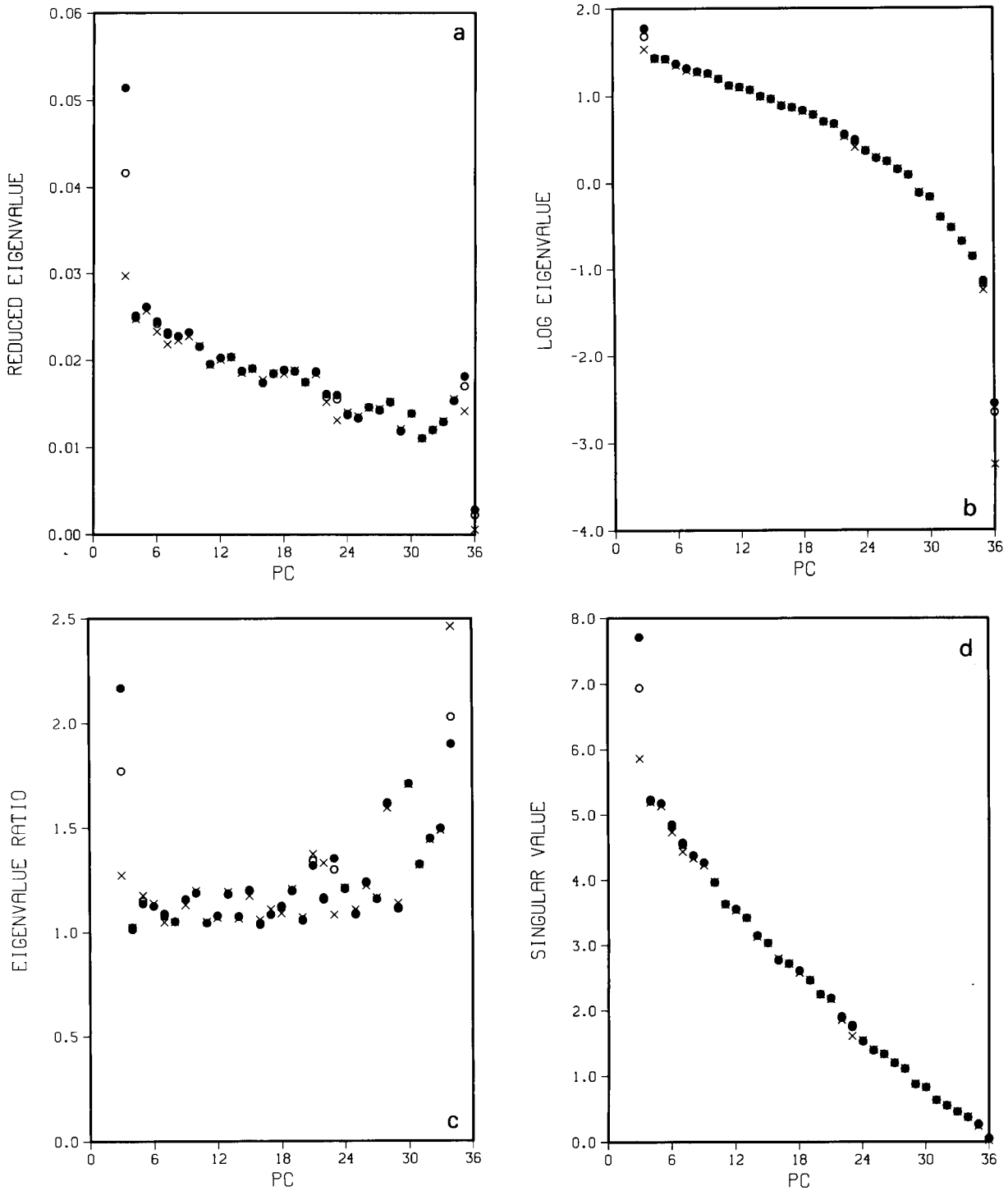


Fig. 2. (a) Reduced eigenvalues, (b) logarithm of eigenvalues, (c) eigenvalue ratios and (d) singular values for EXP1 (●), EXP2 (○) and EXP3 (×).

of this singular value is smaller for the *t*-test than for the *F*-test. Since there is an arbitrary difference in scale between the different data sets, we have listed in Table 11 the ratio of the last significant (hypothetical) and the first non-significant (actual) eigenvalue that would result in significance levels of 1, 5 and 10% respectively for both tests. The first column contains the dataset under consideration. The second column gives the eigenvalue ratio that is actually found. The next three columns give the eigenvalue ratios that would enable testing at the 1, 5 and 10% respectively by the *t*-test. The last three columns give the same results for the *F*-test. From these numbers it is easily discerned that the *t*-test is more sensitive than the *F*-test. The situation is especially favourable for the *t*-test if testing at the 1% level is required. The difference in sensitivity decreases rapidly with increasing significance level. It is interesting to compare the values found for WEIN70 to the critical region found by Hirsch et al. [43] for this data set from extensive simulations: 'If one wishes to ensure the detection of all significant factors and is not concerned that too many factors might be accepted, one should use an ER (eigenvalue ratio) test, probably with a critical value in the range 2.0 to 2.5.' It is remarkable that (approximately) the same critical region is found here by a test that is designed to be conservative.

4.2. Simulated data

In this section we will restrict ourselves to the comparison of the *t*-test and Malinowski's *F*-test. But before presenting these results we discuss the outcome of a large number of conventional methods. In this way we hope to discover what we can reasonably expect from the two significance tests.

In Fig. 2 we show the reduced eigenvalues, the eigenvalue ratios, the logarithm of the eigenvalues and the singular values for the simulated three-component systems. (The values for the first two PCs are not included for visual clarity.) The reduced eigenvalues slowly decrease for the secondary PCs [5]. The logarithm of the eigenvalues lie (approximately) on a straight line for the low-numbered non-significant PCs. This is further illustrated by the eigenvalue ratios being (approximately) constant in that region. (Plotting the eigenvalue ratios instead of the logarithm of the eigenvalues has the advantage of leading to a more practical scale.) These trends are in agreement with the results found earlier [15] for random matrices with an equal number of rows and columns. It is seen that the singular values lie (approximately) on a straight line for all non-significant PCs. This lends credit to the use of the singular values for visual inspection. The graphical methods strongly indicate the presence of three significant PCs for EXP1 and EXP2. For EXP3 no evidence for the presence of the dilute component can be deduced from these plots. The results for EXP2 should be contrasted to the minimum found at the second PC for both imbedded error function and indicator function (not shown here). It should be emphasized that finding a minimum for the imbedded error function can be satisfactorily explained. It is simply the point where less error (random *and* systematic) is introduced in the model by *not* including a PC that in fact contains systematic variation. Cross-validation [3] confirms the choice of a two-dimensional model by giving the ratios 0.04, 0.02 and 1.02 for the first three PCs (cut-off value is one). Since much structure is present in the pure component responses used to construct the data we also investigated the eigenvectors. The first-order autocorrelation function has proved to be a very sensitive method for this kind of data [47]. For the third PC the time constants found are 4.63 and 3.24 for the left and right singular vector respectively while the time constants are 0.47 and 0.58 for the corresponding vectors of the first secondary PC. This result is in excellent agreement with the cut-off value of 0.60 proposed by Shrager [47]. However, for data matrix EXP3 the values 0.75 and 0.69 are found for the third PC and basing a decision on this method becomes difficult. From all the conventional methods ap-

[5] The fact that the reduced eigenvalues are not constant is of little consequence for the application of the *F*-test. The *F*-test guards against violations of assumptions by the small number of degrees of freedom.

Table 12
Results of t-test and F-test for simulated data matrix EXP1. The numbers marked in bold indicate the estimated pseudorank

n	t-test						F-test				
	Θ	Θ_{ref} [a]	$\sigma(\Theta_{ref})$	t	ν	$\%\alpha$	REV	F	$\nu 1$	$\nu 2$	$\%\alpha$
1	4.23×10^3	5.75	0.22	5.97×10^3	1296	0.0	1.38×10^4	4.62×10^2	1	35	0.0
2	6.68×10^2	5.66	0.22	9.36×10^2	1225	0.0	3.64×10^2	1.54×10^4	1	34	**0.0**
3	7.71	5.58	0.22	3.01	1156	**0.1**	5.14×10^{-2}	2.44	1	33	12.8
4	5.24	5.49	0.22	−0.35	1089	63.7	2.52×10^{-2}	1.22	1	32	27.9
5	5.18	5.40	0.22	−0.31	1024	62.2	2.62×10^{-2}	1.30	1	31	26.4
6	4.85	5.31	0.23	−0.65	961	74.2	2.45×10^{-2}	1.24	1	30	27.4

[a] Calculated with $\hat{\sigma}(\mathbf{M}) = 0.50$.

plied to this matrix only a plot of the singular vectors indicates the presence of the third dimension (see Fig. 7 in [16]). The human eye is seen to be an excellent pattern recognizer. The application of this method is, however, restricted to ordered data [48]. In the light of the above results it is reasonable to demand that the significance tests yield certainly three significant PCs for EXP1 and EXP2 and possibly two for EXP3.

In Table 12 we have summarized the relevant results for data matrix EXP1 (for explanation of the symbols see Table 8). It is seen that the t-test correctly yields three significant PCs whereas the F-test fails. The failure is caused by the combined effect of the relatively small F-value and the small number of degrees of freedom. The results for data matrix EXP2 (not shown here) are very similar. Only the numbers for the third PC, which corresponds to the dilute component, change significantly. The t-value for the third PC is decreased to 1.92, yielding a significance level of 2.8%. Inspecting the plots in [16] may lead to the conclusion that this is somewhat pessimistic. Finally, for data matrix EXP3 we find a t-value of 0.40 which is not significant ($\%\alpha = 34.5$). The results for the simulated data sets support the results obtained for the literature data. The t-test makes effective use of the extra knowledge. The estimated pseudoranks are summarized in Table 13.

5. Conclusions

In this paper three methods are discussed with respect to their usefulness as a parametric pseudorank estimation method. These methods explicitly need an estimate of the size of the measurement error. Thus they are restrictive in their use but should be expected to be more sensitive than methods that cannot employ this extra knowledge. Method A (based on the row-echelon form of the matrix) and Method B (based on the eigenvalues of PCA) have the disadvantage of not supplying significance levels for the estimated pseudorank. (Furthermore, Method A is shown to be sensitive for an incorrect estimate of the size of the measurement error.) Thus they offer little advantage over a large number of methods that is already available and does not need the extra knowledge. The third method (a t-test) makes good use of the extra knowledge and does supply a significance level. As a result of the assumptions made in order to construct the test statistic, the t-test should be expected to give conservative estimates of the pseudorank. In or-

Table 13
Results of various pseudorank estimation methods for simulated test data

Data	Cross-validation	Imbedded error function	F-test	Indicator function	t-test	Reduced eigenvalue	Logarithm eigenvalue	Eigenvalue ratio	Singular value
EXP1	2	2	2	3	3	3	3	3	3
EXP2	2	2	2	2	3	3	3	3	3
EXP3	2	2	2	2	2	2	2	2	2

der to gauge the sensitivity of this t-test, a comparison is carried out with Malinowski's F-test for data obtained from the literature and simulated data. For the data matrices obtained from the literature the estimated pseudorank agrees very well for both significance tests. However, the t-test gives sharper confidence levels as a result of the larger number of degrees of freedom involved in the test. For the simulated data matrices Malinowski's F-test fails to indicate the correct dimension in cases where the t-test still yields sharp confidence levels. It is concluded that prior knowledge of the size of the measurement is put to effective use by the currently developed t-test. Additional support for the viability of the new t-test comes from a thorough analysis of the test data by a large number of conventional methods. Finally, as a remarkable by-product of the current research we have found that a plot of the singular values yields a promising graphical pseudorank estimation method. (This is only remarkable, since two modern textbooks on PCA do not mention this possibility [49,50].) Graphical methods have proved their use in the past in cases where the size of the measurement error is unknown. This new graphical method therefore provides a natural complement to the t-test.

Appendix 1

In essence Malinowski's theory deals with the number of degrees of freedom associated with secondary PCs. It is interesting to compare his results, viz. Eq. 11, with the theory developed by Mandel [28]. Mandel argues that an eigenvalue explains a portion of the sum of squares associated to the data. In order to arrive at the portion variance explained by an eigenvalue, the eigenvalue should be divided by an appropriate number of degrees of freedom. The consequence of this reasoning is as follows: for a secondary PC the expectation of the eigenvalue divided by the appropriate number of degrees of freedom should be an unbiased estimator of the variance of the measurement error $\sigma(\mathbf{M})^2$. Alternatively, dividing the expectation of the eigenvalue λ_n by the variance of the measurement error should yield an unbiased estimator of the appropriate number of degrees of freedom, ν_n:

$$\nu_n = \frac{E[\lambda_n]}{\sigma(\mathbf{M})^2} \tag{13}$$

where $E[\cdot]$ denotes expected value. In fact these expected values are the reference values discussed earlier. Since the variance accounted for by the secondary PCs should be constant the denominator in Eq. 11 should be proportional to the number of degrees of freedom associated with the PC under scrutiny. The proportionality constant N (normalization) is found by observing that the number of degrees of freedom summed over the secondary PCs should add up to the total number of degrees of freedom left after extracting n^* components, i.e., building the correct model:

$$N = \frac{(r-n^*)(c-n^*)}{\sum_{j=n^*+1}^{s}(r-j+1)(c-j+1)} \tag{14}$$

Hence $\nu_n = N(r-n+1)(c-n+1)$, found this way, should equal ν_n found from evaluating Eq. 13. We will return to this question in more detail in another publication [15]. It is tempting to evaluate Eq. 12 using the number of degrees of freedom associated with the sources of variance that are tested instead of using the number of sources as the number of degrees of freedom, i.e., take $\nu_1 = N(r-n+1)(c-n+1)$ and $\nu_2 = \Sigma_{n+1}^{s} N(r-j+1)(c-j+1)$ instead of $\nu_1 = 1$ and $\nu_2 = s-n$. In general this should lead to larger numbers of degrees of freedom and consequently the resulting F-test should yield an increased discriminating ability. Some obvious disadvantages are connected with this 'alternative F-test'. In general the resulting numbers will not be integral as already indicated by Mandel [28]. This problem is easily solved by rounding the numbers to the nearest integer. The confidence levels do not change very much by this operation, especially for large data matrices. A second disadvantage is the presence of the correct dimension of the PC model in Eq. 14. This problem cannot be solved since it leads to circular reasoning: the

pseudorank needs to be known in order to estimate it.

References

[1] E.R. Malinowski, Factor Analysis in Chemistry, Wiley, New York, 1991.
[2] E.R. Malinowski, Anal. Chem., 49 (1977) 612.
[3] S. Wold, Technometrics, 20 (1978) 397.
[4] E.R. Malinowski, J. Chemom., 3 (1988) 49.
[5] E.R. Malinowski, J. Chemom., 4 (1990) 102.
[6] E.R. Malinowski, J. Chemom., 1 (1987) 33.
[7] L. Ståhle and S. Wold, Chemom. Intell. Lab. Syst., 6 (1989) 259.
[8] X.M. Tu, D.S. Burdick, D.W. Millican and L.B. McGown, Anal. Chem., 61 (1989) 2219.
[9] X.M. Tu, J. Chemom., 5 (1991) 333.
[10] S.P. Koinis, A.T. Tsatsas and D.F. Katakis, J. Chemom., 5 (1991) 21.
[11] Y.-Z. Liang, O.M. Kvalheim and A. Höskuldsson, J. Chemom., 7 (1993) 277.
[12] V. Tomišić and V. Simeon, J. Chemom., 7 (1993) 381.
[13] R.M. Wallace and S.M. Katz, J. Phys. Chem., 68 (1964) 3890.
[14] Z.Z. Hugus and A.A. El-Awady, J. Phys. Chem., 75 (1971) 2954.
[15] N.M. Faber, L.M.C. Buydens and G. Kateman, Chemom. Intell. Lab. Syst., in press.
[16] N.M. Faber, L.M.C. Buydens and G. Kateman, J. Chemom., 7(1993) 495.
[17] L.A. Goodman and S.J. Haberman, JASA, 85 (1990) 139.
[18] R.N. Cochrane and F.H. Horne, Anal. Chem., 49 (1977) 846.
[19] R.J. Pell, M.B. Seasholtz and B.R. Kowalski, J. Chemom., 6 (1992) 57.
[20] C.B.M. Didden and H.N.J. Poulisse, Anal. Lett., 13 (1980) 921.
[21] M.G. Moran and B.R. Kowalski, Anal. Chem., 56 (1984) 562.
[22] H. Bubert and H. Jenett, Z. Anal. Chem., 335 (1989) 643.
[23] R.M. Wallace, J. Phys. Chem., 64 (1960) 899.
[24] G. Weber, Nature, 190 (1961) 27.
[25] S. Ainsworth, J. Phys. Chem., 65 (1961) 1968.
[26] G.H. Golub and C.F. Van Loan, Matrix Computations, Hopkins University Press, Baltimore, 1983.
[27] J.M. Halket, J. Chromatogr., 175 (1979) 229.
[28] J. Mandel, Technometrics, 13 (1971) 1.
[29] P.J. Bickel and K.A. Doksum, Mathematical Statistics, Holden-Day, San Francisco, CA, 1977.
[30] C.L. Lawson and R.J. Hanson, Solving Least Squares Problems, Prentice-Hall, Englewood Cliffs, 1974.
[31] J.R. Green and D. Margerison, Statistical Treatment of Experimental Data, Elsevier, Amsterdam, 1978.
[32] K. Booksh and B.R. Kowalski, J. Chemom., 8 (1994) 45.
[33] J. Fiala, Anal. Chem., 52 (1980) 1300.
[34] V. Gutmann, Co-ordination Chemistry in Non-aqeous Solutions, Springer Verlag, Vienna, 1968, p. 33
[35] D.G. Howery, Bull. Chem. Soc. Jpn, 45 (1972) 2643.
[36] J. Havel and M. Meloun, Talanta, 32 (1985) 171.
[37] E.R. Malinowski, Anal. Chim. Acta, 134 (1982) 129.
[38] G.L. Ritter, S.R. Lowry, T.L. Isenhour and C.L. Wilkins, Anal. Chem., 48 (1976) 591.
[39] P.H. Weiner, E.R. Malinowski and A.R. Levinstone, J. Phys. Chem. 74 (1970) 4537.
[40] P.H. Weiner and E.R. Malinowski, J. Phys. Chem., 75 (1971) 1207.
[41] H.B. Woodruff, P.C. Tway and L.J. Cline Love, Anal. Chem., 53 (1981) 81.
[42] R.A. Hearmon, J.H. Scrivens, K.R. Jennings and M.J. Farncombe, Chemom. Intell. Lab. Syst., 1 (1987) 167.
[43] R.F. Hirsch, G. Lam Wu and P.C. Tway, Chemom. Intell. Lab. Syst., 1 (1987) 265.
[44] F.P. Zscheile, H.C. Murray, G.A. Baker and R.G. Peddicord, Anal. Chem., 34 (1962) 1776.
[45] IMSL MATH/LIBRARY User's Manual, version 1.1, IMSL, Inc., Houston, TX, 1989
[46] N. Ohta, Anal. Chem., 45 (1973) 553.
[47] R.I. Shrager, SIAM J. Alg. Disc. Meth., 5 (1984) 351.
[48] C. Shen, T.J. Vickers and C.K. Mann, J. Chemom., 5 (1991) 417.
[49] I.T. Joliffe, Principal Component Analysis, Springer Verlag, New York, 1986.
[50] J.E. Jackson, A User's Guide to Principal Components, Wiley, New York, 1991.

ELSEVIER

Analytica Chimica Acta 296 (1994) 21–31

ANALYTICA CHIMICA ACTA

Developing an expert system for diagnosis of problem gas chromatographic data

Hai Du, Sharbari Lahiri, Guosheng Huang, Martin J. Stillman *

Department of Chemistry, University of Western Ontario, London, Ontario N6A 5B7, Canada

Received 11th September 1993; revised manuscript received 17th January 1994

Abstract

An expert system with a user-friendly interface has been developed to aid in the diagnosis of faulty analysis by gas chromatography as a first step towards totally automated analysis. The system advises users on the identity of faulty components and the possibility of improper operation based on the appearance of the chromatograms that result from the experiment. The knowledge used in this expert system has been assembled into a knowledge domain matrix and was coded using case histories. One case serves as an example: "A leaking syringe will result in loss of signal intensity and the appearance of false peaks due to contamination", is coded to give two entries in the knowledge domain matrix the cause "leaking syringe" is "True" for both "sensitivity loss" and "syringe related peaks". The knowledge base can be readily expanded through use of Table Edit and Rule Edit (programs developed in our laboratory). The inference engine (EAengine) can apply forward, backward or mixed chaining strategies to obtain a cause for the symptoms that are observed. The system is written in Microsoft Visual BASIC 2.0 within Window 3.x, calling the inference engine as a dynamic link library (DLL). The user can enter observations easily using a powerful new graphical-based interface. Gas chromatograph operation and control systems that use Windows 3.1 may be in progress at the same time. The diagnostic program may be suspended while a more urgent task is run. Communication between a Varian Star Workstation and this system was realized through the dynamic data exchange (DDE) facilities of the Windows 3.1 shell.

Keywords: Expert systems; Gas chromatography

1. Introduction

The successful interpretation of the analytical data of environmental samples for trace contamination by organic compounds requires the highest quality data [1]. The gas chromatography (GC) technique is unsurpassed in resolving power, sensitivity, and speed of analysis for determination of many organic and inorganic compounds of varying volatilities, stabilities, and chemical reactivities [2,3]. In combination with mass spectrometric (MS) detection, the trace organic compounds in complex mixtures can be readily identified and quantified at nanogram or even picogram levels from a single environmental sample following chromatographic separation. Full automation of

* Corresponding author.

SSDI 0003-2670(94)00187-Q

GC analytical procedures and analysis will ensure that the analytical results will be more reliable, and will free experts from routine work, letting them focus on more complicated problems. Automated diagnosis of instrumental and operational conditions is a necessary first step of total automation. Naturally, expert system technology must be used.

An expert system embodies in a computer the knowledge-based component of an expert's skill in such a form that the system can offer intelligent advice and, on demand, justify its own line of reasoning [4–6]. Expert systems offer a way of making expensive and rare expert knowledge available to all members of a user community. Analytical laboratories handling environmental samples work with difficult matrices, yet are required to provide results at the limits of the sensitivity of the technique. The availability of expert knowledge on a continuous basis from an expert system is a cost effective way of improving productivity.

A large number of expert systems are being developed for use in the analytical laboratory aimed in many cases at reducing well-known bottlenecks in the analysis of complex samples. Within the European Community project ESCA ("Expert Systems for Chemical Analysis"), a large group of scientists is working on expert systems for use with the liquid chromatographic (LC) techniques, particularly as applied to pharmaceutical analysis [7–10]. Expert systems are also being developed in other fields: arson analysis by GC–MS [11], voltammetric analysis of trace metals [12], validation of biochemical data [13], and control of bioprocesses [14]. However, in the area of problem diagnosis for chromatography instruments, only an expert system for troubleshooting LC assay methods has been reported [15]. In our laboratory, an expert system for the analysis of environmental samples by GC–MS is being developed to be used with a Varian GC 3400 with flame ionization detector and ion trap mass spectrometer coupled with a 8100 autosampler (the Saturn II GC–MS instrument, Varian), which contains a diagnostic module. Now, we describe development of a multimodule expert system that will allow fully automated diagnosis of a GC component fault and evaluation of a gas chromatogram or a GC system in run-time. This paper describes the first stages in development of this system, a GC diagnostic expert system, GCdiagnosis, which will allow users to diagnose faulty analysis by gas chromatography.

Our group has developed a number of expert systems for use in the analytical laboratory [16–22]. We believe that the user interface must be written within a graphical operating system so that interaction between the user and the knowledge base is carried out as efficiently as possible. Our expert systems are developed within the Windows 3.x shell or operating system. Our development strategy includes: acquiring domain knowledge, coding the knowledge into a knowledge domain matrix, and finally converting the matrix into knowledge base files.

The acquisition of knowledge is a key phase and represents a significant bottleneck in the development of knowledge-based systems. Although much textual information to aid in diagnosing analysis by gas chromatography is available, it is not the heuristic knowledge used by domain experts. Even under the same situations, methods of recognition and diagnosis naturally vary between domain experts. During development of this present system, an explanation-based learning approach was applied to the acquisition of knowledge. The KDS3 expert system shell (KDS, Wilmeter, IL) was used as a tool to check the logic of the knowledge base [23]. This report describes the assembly of knowledge into a knowledge domain matrix in the manner described by us previously [19], the design and development of a generic graphical-based user interface for diagnosis, and finally, implementation of the complete expert system.

2. Experimental

2.1. Computer environment

All computer work was carried out on an Intel 80486-based computer, VEISA ALR (Advanced Logic Research, Irvine, CA), containing 17 mega-

bytes of RAM, two 1.44-megabyte floppy disks, a 200-megabyte hard disk, and VGA graphic monitor.

Expert systems often require that the following types of software be selected: an expert system shell program, and programming language. In this study, software used by the prototype system includes Microsoft Window 3.x, Microsoft Visual Basic (programming language and graphics), EAshell, an expert system shell being developed in our laboratory [24], and the Varian Star Workstation. In addition, usually a complex expert system can be divided into several rather simple modules, and each module can be developed separately [25]. When all of the modules are completed, a global expert system may be obtained by linking these modules. The Windows 3.1 provides a simple method of realizing the communication between these modules by using DDE and multitasking functions.

2.2. Inference engine

Forward and backward chaining are two fundamental forms of inference. Forward chaining, also called data directed reasoning, begins with the facts that the user enters, the engine tests the knowledge base successively to find all possible causes that require these facts to be confirmed. Backward chaining or goal directed reasoning starts with the suspected goal, then the engine attempts to find evidence for this goal from the knowledge base. Although there are a large number of expert system shells available, our requirement for a small system that runs within the MS Windows environment (so that the system could coexist with the GC instrument control and data acquisition software was) restrictive. We employed our own EAshell system, a shell that has been designed for use with MS Windows applications through DLL (Dynamic Link Library) functions [24]. EAengine was used as the inference engine. EAengine uses knowledge supplied in the form of production rules and supports three inference strategies, forward, backward and mixed chaining (that is the combination of forward and backward chaining). The mixed chaining begins with the forward chaining. When a situation arises in which the backward chaining is the best solution to the problem, the inference engine continues with the backward chaining searching. When the backward chaining fails to find evidence for a goal, the engine automatically changes back to forward chaining.

The knowledge base file consists of three major sections: a goal section, a production rule section and a user query section. The goal section contains information about goals for backward chaining and corresponding graphical descriptions. A rule is a list data structure that contains the rule name, rule premise, and rule conclusion.

2.3. User interface

In order to provide the best answer, the expert system must gather enough facts. For this reason, the system was programmed in such a way that each symptom is accompanied by a diagram or textual description. First-time users or users with little experience will spend some time reading the descriptions while regular users and analysts with more experience may ignore the text. Generally, graphs are a necessary component for a diagnostic system. The graphical images are used to describe the physical components of the system and symptoms of the GC system. Windows is a graphical user interface, and it is possible to input a graphical description of a physical component by using a video camera. The graph can be stored as a bitmap file that can be loaded and displayed when needed.

3. Results and discussion

3.1. Components in the GCdiagnosis system

An expert system is a very complex computer program [21]. Fig. 1 shows the three key components in this system: (1) the user interface, (2) the inference engine, and (3) the knowledge acquisition system. We show the connection between the heuristic knowledge that is coded in terms of a case, the region of the knowledge matrix that holds the coded knowledge and, finally, the rules generated for use with EAengine. The user inter-

face incorporates high resolution images as symptom definitions, or as descriptions of the physical components of a GC system. The domain knowledge is represented by cases that are coded into a knowledge domain matrix in terms of symptoms and causes linked by logical connections. The inference engine will match facts with the heuristic rules and knowledge stored in the knowledge base. Acquiring domain knowledge, coding the knowledge into the knowledge domain matrix, converting the matrix into knowledge base files, and developing the program in Windows 3.x have been accomplished as part of the development of this system.

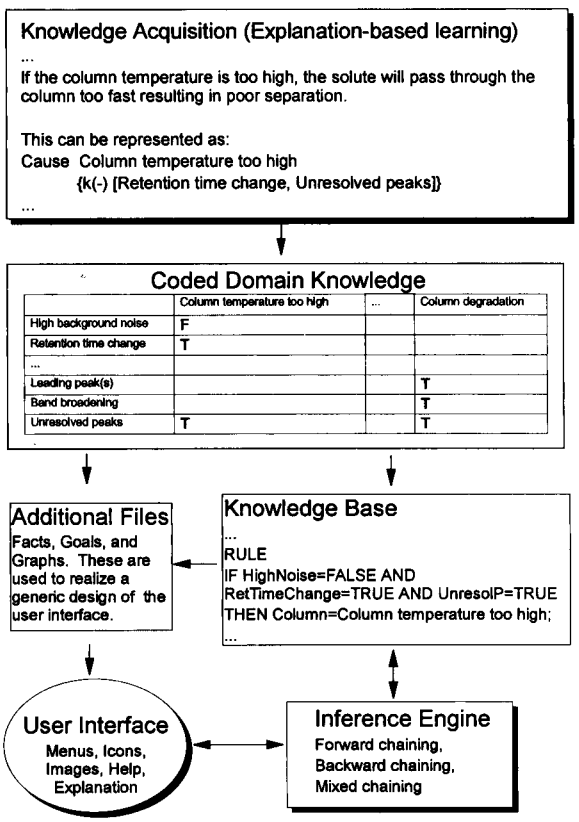


Fig. 1. Components in the GCdiagnosis system. The system consists of three key components: (1) the knowledge acquisition, (2) the inference engine, and (3) the user interface. The knowledge shown here for the effect of a high temperature on the column was compiled in terms of an explanation-based learning method and was represented as cause X (a number). Then the knowledge is coded into the knowledge domain matrix that is converted into production rules used by the inference engine (EAengine). The user interface incorporates high resolution images as both symptom definitions and descriptions of the physical components of a GC system.

Generally, developing an expert system requires that the knowledge engineer and expert must collaborate in cataloguing and assembling the knowledge base. In an alternative and more productive approach for small systems, the analytical chemist acts as both the expert and the knowledge engineer. In this case the expert must be able to assemble and code the knowledge. In this work, an explanation-based learning approach has been used in the knowledge acquisition [23], which may be considered as a variant of the latter method. Gas chromatography has become a mature area, in which many quantitative or qualitative relations have been used to describe the physico-chemical behaviour of gas chromatography, and a lot of diagnostic cases are available. It is possible to compile heuristic knowledge based on the model described above and selected cases of GC diagnosis.

3.2. Knowledge presentation in the form of a knowledge domain matrix

The method used for knowledge presentation is determined by the size and type of the knowledge base [26]. In this study, the knowledge can be represented in the form of a matrix, called the knowledge domain matrix. In this form of coding, the first or primary logical connections are made from case histories, then secondary connections are made between all existing conditions and conclusions. The basis of this technique has been described previously [19]. Fig. 2 shows the major parts of the current knowledge base. Fifteen symptoms frequently encountered in routine GC operations and twenty-eight independent causes were selected as cases. Symptoms are classified as either action-related or chromatogram-related. When a symptom describes operational conditions it is action-related; and when a symptom is observed in the chromatogram it is considered to be the chromatogram-related. For example, "High background noise" is chromatogram-related, while "Syringe/injection related peak" is action-related. The matrix contains related pairs

Goals:	Injection					Column															FuelGas	Detector				CarrierGas				Injector			
Symptoms / Causes	1. Leaking syringe	2. Dirt in the syringe	3. Sample size too large	4. Double injection	5. Injection too slow	6. Sample decomposition	7. Late elution peak(s)	8. Column temperature too high	9. Column temperature too low	10. Adsorption of sample on active sites	11. Reaction of sample with GC parts	12. Unstable column temperature	13. Column bleed	14. Column bleed (severe)	15. Leaking at column exit	16. Leaking severe	17. Column degradation	18. Contaminated column	19. Wrong fuel gas flow rate	20. Dust burned in flame	21. Unstable detector temperature	22. Flame not ignited in FID	23. Contaminated detector	24. Contaminated carrier gas	25. Carrier gas flow too low	26. Carrier gas flow too high	27. Carrier gas flow too high (severe)	28. Dirt in the injector	29. Dirt in the injector (severe)	30. Injector temperature too low	31. Leaking at septum
1. Sensitivity loss	T					T				F	T		F		T	T			T		T	F	T				F			T
2. Retention time change			F			F		T	T	T	T	T	F	T		T	T								T	T	T	F		F	T
3. High background noise								F							T			T			F		T	T		T	T	T	T		
4. Ghost peak(s)		T				T	T				T							F					T				F	T		T
5. Peak tailing									T	T	T							T					F				F	T	T	F
6. Baseline drift												T	T	T	T				F		T		T							
7. Syringe/injection related peak signal	T	T		T	T																									
8. Leading peak			T														T							T						
9. Unresolved peaks								T									T							F	T	T			T	
10. Band broadening					T		T	F	T	F							T												T	
11. Can not zero baseline												F	T	T	F				T				T								
12. Irregular spikes																				T											
13. Split peak				T																											
14. No peak(s)																T						T									
15. Sudden drop-off normal peak																										F	T				

Fig. 2. The current knowledge base for GCdiagnosis. The symptom set and cause set are located on the first column and the first row, respectively, and the logical connection, T(rue) or F(alse), is placed in the related cell.

of causes 13 and 14, 26 and 27, or 28 and 29 that are same but are coded to enable different fuzzy quantitation. For example, light column bleeding causes a baseline drift, the chromatogram cannot be zeroed, but when severe column bleeding occurs, retention times will be changed because the column phase ratio, β, is significantly changed (decreased). The symptom set and cause set are located on the first column and the first row, respectively, and the logical connection, T(rue) or F(alse), is placed in the related cell. The matrix is divided into six divisions or goals, each of which represents a sub-domain of the knowledge base. For example, causes 1 to 5 belong to the goal "*Injection*". The major advantages in the application of the matrix representation are: (1) it is easy to follow the reasoning path that form the basis of decisions made by the EAshell engine; (2) it makes the knowledge representation clear and knowledge refinement and expansion easy; and (3) it makes the knowledge more transferable. For instance, in the mixed chaining mode, when the user enters the fact "*High background noise (is True)*", only conclusions, 15, 18, 23, 24, 26–29, are satisfied. The engine picks the next question only from amongst these conclusions for further questioning. This procedure will continue until one answer is obtained or a list of possible causes that are ranked based on the numbers of matching facts. If the developer wants to change or expand the knowledge base, this can be done by adding new columns or/and rows, and filling in specific cells in the matrix. This knowledge presentation method makes knowledge refinement easy. In addition, the matrix can handle fuzzy quantitation. For example, if severe peak tailing is considered as a more informative symptom, then it may be added into the symptom set as a new symptom.

3.3. Production rules

Knowledge used in an expert system should be transferable to other expert system in the future. Because the knowledge is an independent component of the expert system and it is unrelated to the chosen programming techniques, this is possible. Different versions of an expert system may be programmed by different techniques, but the knowledge can be the same. The techniques used to implement GCdiagnosis reaches this feature because we first encode the domain knowledge in a knowledge matrix. Conversion of the knowledge domain matrix to production rules is straightforward. Knowledge presentation in the form of rules is a popular method in chemical applications of expert systems mainly because rules are a natural extension of scientific method [27]. A tool kit (Rule-Editor) was developed to convert the matrix into the KBF rule file for use with EAshell. Key parts of the KBF file are:

```
I   Goal section
        FIND
        CarrierGas, Injection, Injector, Column, Detector, FuelGas;
II  Production rule section
        RULE 1
        IF
        HighNoise = TRUE AND GhostPeak = TRUE AND PeakTail = FALSE
        THEN
        CarrierGas = Contaminated carrier gas;
            :
III User query section
        ASK HighNoise
        "High background noise?"
        OPTION
        TRUE, FALSE;
            :
```

Where CarrierGas, Injection, Injector, Column, Detector, and FuelGas represent the six major components in the model [23], respectively. This rule (named 1) is applicable if *high background noise* is observed in the chromatogram, a *ghost peak* is found, and there is no *peak tailing* observed. The conclusion of this rule is that the carrier gas is contaminated. When running the mixed chaining mode the user query section provides an interface between EAengine and a symptom variable, its text description, and its answer choices.

The development of an expert system is a dynamic process. This requires that the expert system must be designed to be easy to maintain.

When new knowledge is added to the knowledge base, the program still must be able to handle this new knowledge base with same user interface (except for adding new functions to the program). During development of GCdiagnosis we kept this in mind and used data files to make the program independent of knowledge base expansion. Besides the GCdiagnosis program and knowledge base, the following data files are used as parts of program:

goals file (a text file): used to store information about the connection between each goal and its graphical description;

facts file (a text file): used to store information about the connection between each symptom, its answer options and its graphical description;

subdirectory *graph*: used to store all graphical descriptions related to the KBF file, each graphical description has the Window bitmap format (with.BMP extension name) or Window metafile format (with.WMF extension name).

Through these files, the program can be updated without re-coding. When the knowledge base is updated, the corresponding changes are added into these files not into the program itself.

Fig. 3. GCdiagnosis system within Varian Star Workstation frame. The system can be easily integrated into the Star Workstation frame, and becomes a module of the Star Workstation.

3.4. Integration of the GCdiagnosis system into the Star Workstation software

The GCdiagnosis system can be installed as an icon inside the Varian Star Workstation frame within the Program Manager, Fig. 3, which allows users rapid access to the system even when the GC is under computer control. The Star Workstation is a powerful tool that helps users accomplish their laboratory work more efficiently, while completely controlling the Varian GC with the Autosampler. The GC results are collected from the instrument and stored on the hard disk. Each data file contains the raw data, calculated results, method, analysis conditions, and instrument run conditions for the GC. The Report module converts the data files into ASCII-coded files. Chromatographic performance can be evaluated with the System Suitability macro. A report is generated documenting the column efficiency (theoretical plates), resolution, capacity factor (k'), and the peak tailing from a representative chromatogram.

Dynamic data exchange (DDE) is a mechanism supported by the Microsoft Windows operating system that enables two applications to communicate with each other by continuously and automatically exchanging data. DDE builds a bridge between the GCdiagnosis system and the Star Workstation, which naturally makes the system become a module of the Star Workstation. In design, the GCdiagnosis system is defined as a destination application that initiates the communication, the results generated by the System Suitability macro are considered as a source application, Fig. 4. When a GC experiment is completed, the user may use this procedure to automatically determine if there are problems with GC data in the chromatogram. For example, if the results from System Suitability macro show that peak tailing is greater than 1.15, running the GCdiagnosis system is recommended.

In the area of gas chromatography, running test mixtures has been one method to diagnose the GC system. For instance, by evaluating the

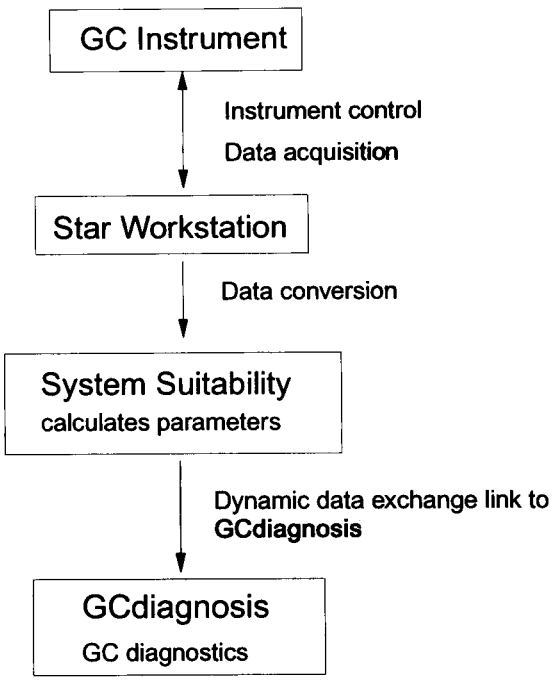


Fig. 4. Communication between the Star Workstation and GCdiagnosis system is realized through dynamic data exchange (DDE).

numbers of theoretical plates of hydrocarbon solutes and the shapes of those peaks under a particular set of conditions on that apparatus, poor injection technique and gas flow problems can be diagnosed [28]. Through DDE, the GCdiagnosis system can access the results of the System Suitability macro automatically, this makes it practical to diagnose faults detected from the chromatogram automatically. At present, only the calculated results of the column efficiency, resolution, and the peak tailing are used, and a data file recording those results is created so that changes in system performance can be tracked, and the behaviour of solutes under a particular set of conditions can be established.

4. Implementation

4.1. Implementation of the inference engine

The GCdiagnosis system has been designed to be generic for use with other analytical instruments, because there are many common aspects in all diagnostic problems. For example, a description of the physical components of the instrument is necessary, and generally the diagnosis is carried out based on symptoms observed from experimental results. In this system, two picture boxes are used to display the graphical descriptions of symptoms and physical components of the GC system. One box is in the forward chaining window, and another is in the backward chaining window. Use of these boxes allows the program to handle pictures easily. The main window contains three menus, "Files", "Inferences" and "Help", Fig. 5. A mouse click on the "File" prompts an open file command, which will open a specific knowledge base depending on user's requirement and load it into the memory. To start the GCdiagnosis for a GC system configured with a flame ionization detector, a user chooses the GCFID.KBF file. If an ion trap mass spectrometer is the detector then the GCITD.KBF file is chosen.

The inference engine menu includes a list of the available modes of the inference engine: the mixed chaining mode, the forward chaining mode and the backward chaining mode. Generally, mixed chaining is highly recommended. This engine mode requires that the user answers a series of questions prompted by the computer, where the order of the questions is based on the dy-

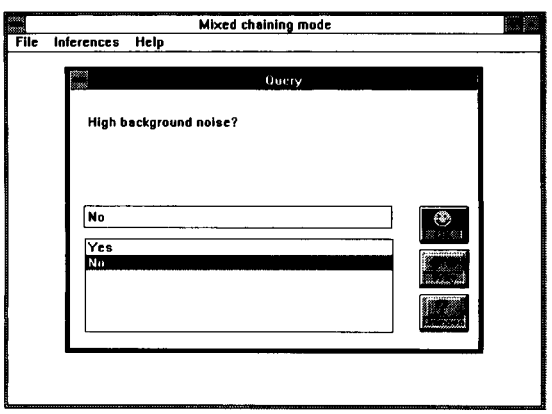


Fig. 5. Design of the main window and mixed chaining task. The main window contains three menus, "Files", "Inferences" and "Help".

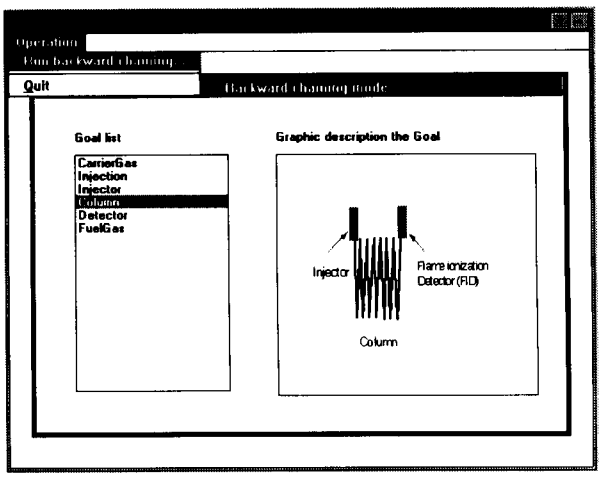


Fig. 6. Implementation of the backward chaining mode, showing six goals representing the six major components.

namic results of statistical analysis of the knowledge domain matrix. When users are certain all symptoms can be identified correctly the forward chaining mode is the best choice in order to save time. Sometimes users can run backward chaining if they suspect something is wrong with a specific component of a GC system. Another purpose of backward chaining is to double check possible symptoms: when users run forward or mixed chaining and obtain a result, they can run backward chaining as a check.

Fig. 6 shows the backward chaining window. There are six goals within the goal list, representing the carrier gas section, the injection section, the injector section, the column section, the detector section, and the fuel gas section, respectively. Clicking one of these list items activates a graphic description that displays as a picture box next to the list. When users are certain that a specific goal is the suspected fault section, running the backward chaining mode results in EAshell engine carrying out the mixed chaining mode based on a sub-knowledge base, a logic division related only to that goal. An analyst might prefer this mode when he/she guesses the source of the problem with the GC data.

The major components on the forward chaining window are similar to that of the backward chaining window except that one additional answer list and one OK button are added, Fig. 7. Users choose a symptom from the symptom list first and then a corresponding answer from the answer list. Clicking the OK button will tell the system that this fact is accepted. This procedure is repeated until all facts are input. When users are sure their observations are complete, this method is recommended.

4.2. Implementation of GCdiagnosis

The graphical user interface design allows users to enter their observations easily. For instance, in routine analysis if an unexpected small broad peak is observed in the chromatogram between two sharp peaks, and if the mixed chaining mode is chosen, the user just needs to answer the questions prompted by the system until the most probable cause, here "*Late eluting peak*", is confirmed. The following is the answer sequence.

High background noise?
No
Retention time change?
No
Leading peak?
No
Peak tailing?
No
Sensitivity loss?
No
Baseline drift?
Ba
No

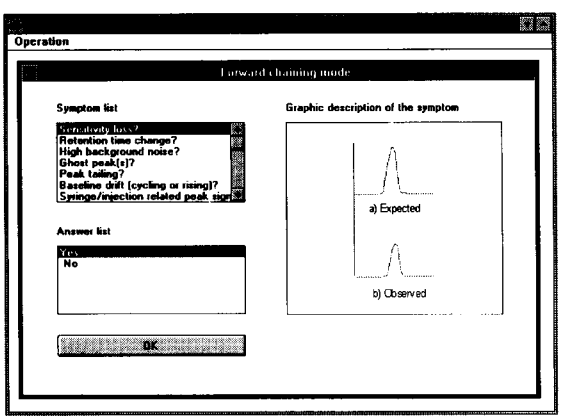


Fig. 7. Implementation of the forward chaining mode. Users can enter their observations by choosing a symptom from the symptom list first and then an answer from the answer list.

No peak(s)?
No
Ghost peak?
Yes
Syringe/injection related peak?
No
Band broadening?
Yes

It is known that a variety of factors can give rise to a "*ghost peak*", such as contaminated carrier gas, dirty syringe, sample decomposition, reaction of the sample with the parts of the GC system, and a late eluting peak. In order to identify the most probable cause correctly, the system requires more information to enable it to distinguish between these factors. The system will keep questioning the user based on the dynamic results of the statistical analysis of the knowledge domain matrix until the most probable cause has been found. When running the backward chaining mode by choosing the goal "*Column*", the inference engine will reason in terms of the subknowledge base related to that goal, and after answering three questions, "*Sensitivity loss (is) No*", "*Ghost peak(s) (is) Yes*" and "*Band broadening (is) Yes*", the same conclusion is obtained. For the forward chaining mode, users enter their observations: "*Ghost peak (is Yes)*" and "*Band broadening (is Yes)*", then run the inference engine, the same conclusion is obtained.

5. Conclusions

Development of expert systems is a dynamic process; the knowledge base needs to be upgraded continuously. Using a generic design method developed in this study, the GCdiagnosis program is independent of knowledge base expansion. Knowledge presentation in the form of a matrix makes it easy to expand the knowledge base. The present expert system will be upgraded in several respects in the near future.

The MS Windows environment prövides users with a friendly interface and makes the system easy to manage and more valuable for real-world applications. With the help of the graphical user interface, the operator can use the mouse to select specific symptoms and enter answers. Under the Windows environment, the graphic description of symptoms and physical parts of instrument can be easily entered by video camera, and it is easy to realize the communication between GCdiagnosis system and the Star Workstation through DDE.

Acknowledgments

Part of this work was presented at the Technology Transfer Conference, Ontario Ministry of the Environment (MOE), Toronto, November 5 and 6, 1992. The authors gratefully acknowledge the financial support from the Ontario Ministry of the Environment and NSERC of Canada through both operating and equipment strategic grants (to MJS). We wish to thank Qiwei Zhu, Dr. Shengping Ma and Dr. Devon Cancilla for useful discussions. EAshell was developed with funding from MOE, NSERC of Canada (through a strategic grant) and Environment Canada (through a GLURF grant). We are also grateful for funding from the University of Western Ontario President's Fund in support of a summer stipend to H. Du.

References

[1] J.G. Grasselli, Anal. Chem., 64 (1992) 677A.
[2] F.W. McLafferty and J.A. Michnowicz, CHEMTECH, 22 (1992) 182.
[3] D. Barcelo, Anal. Chim. Acta., 263 (1992) 1.
[4] T.H. Pierce and B.A. Hohne (Eds.), Artificial Intelligence Application in Chemistry, ACS Symposium Series Vol. 306, ACS, Washington, DC, 1986.
[5] B.A. Hohne and T.H. Pierce (Eds.), Expert System Application in Chemistry, ACS Symposium Series, Vol. 408, ACS, Washington, DC, 1989.
[6] J.M. Hushon (Ed.), Expert Systems for Environmental Applications, ACS Symposium Series, Vol. 431, ACS, Washington, DC, 1990.
[7] D. Goulder, T. Blaffert, A. Blokland, L. Buydens, A. Chhabra, A. Cleland, N. Dunand, H. Hindriks, G. Kateman, H. van Leeuwen, D. Massart, M. Mulholland, G. Musch, P. Naish, A. Peeters, G. Postma, P. Schoenmakers, M. de Smet, B. Vandeginste and J. Vink, Chromatographia, 26 (1988) 237.
[8] M. De Smet, G. Musch, A. Peeters, L. Buydens and D. Massart, J. Chromatogr., 485 (1989) 237.

[9] P.J. Schoenmakers and N. Dunand, J. Chromatogr., 485 (1989) 219.
[10] M. Mulholland, N. Dunand, A. Cleland, J.A. van Leeuwen and B.G.M. Vandeginste, J. Chromatogr., 485 (1989) 283.
[11] G. Holzer, W. Bertsch and Q.W. Zhang, Anal. Chim. Acta, 259 (1992) 225.
[12] (a) M. Esteban, I. Ruisanchez, M.S. Larrechi and F.X. Rius, Anal. Chim. Acta, 268 (1992) 95. (b) M. Esteban, I. Ruisanchez, M.S. Larrechi and F.X. Rius, Anal. Chim. Acta, 268 (1992) 107.
[13] P.M. Valdiguie, E. Rogari and H. Philippe, Clin. Chem., 38 (1992) 83.
[14] B. Hitzmann, A. Lubbert and K. Schugerl, Biotechnol. Bioeng., 39 (1992) 33.
[15] K. Tsuji and K.M. Jenkins, J. Chromatogr., 485 (1989) 297.
[16] W.R. Browett and M.J. Stillman, Prog. Anal. Spectrosc., 12 (1989) 73.
[17] W.R. Browett, T.A. Cox and M.J. Stillman, ACS Symposium Series, Vol. 408, ACS, Washington, DC, 1989, p. 210.
[18] M.J. Stillman, G. Huang, S. Lahiri and Q. Zhu, J. Liebowitz (Eds.), Proceedings of the 1st World Congress on Expert Systems, Pergamon, Oxford, 1991, p. 2645.
[19] S. Lahiri and M.J. Stillman, Anal. Chem., 64 (1992) 283A.
[20] D.A. Cancilla, G. Huang, S. Ma and M.J. Stillman, Proceedings of the 40th ASMS Conference, Washington, DC, June, 1992, p. 993.
[21] M.J. Stillman, Encyclopaedia of Analytical Science, Academic Press, London, 1994, in press.
[22] (a) S. Lahiri and M.J. Stillman, Proceedings of the Technology Transfer Conference, Ontario Ministry of the Environment (MOE), Toronto, November 5 and 6, 1992, BP44. (b) Q. Zhu and M.J. Stillman, ibid, BP44A. (c) H. Du, S. Lahiri, G. Huang and M.J. Stillman, ibid, P44B. (d) M.J. Stillman, S. Lahiri and G. Huang, ibid, BP44C. (e) S. Lahiri and M.J. Stillman, Proceedings of the Technology Transfer Conference, MOE, Toronto, November 25 and 26, 1991, p. 685. (f) M.J. Stillman, G. Huang, S. Lahiri and Q. Zhu, ibid, p. 698. (g) Q. Zhu and M.J. Stillman, ibid, p. 704. (h) M.J. Stillman, S. Lahiri and Q. Zhu, Proceedings of the Technology Transfer Conference, MOE, Toronto, November 19 and 20, 1990, p. 628. (i) M.J. Stillman, M. Moussa, and Z. Gasyna, Proceedings of the Technology Transfer Conference, MOE, Toronto, November 20 and 21, 1989, p. 254. (j) M.J. Stillman, T. Cox, and W.R. Browett, Proceedings of the Technology Transfer Conference, MOE, Toronto, November 28 and 29, 1988, p. 195.
[23] Hai Du and M.J. Stillman, Anal. Chim. Acta, 296 (1994) 43.
[24] G. Huang and M.J. Stillman, unpublished results.
[25] P. Conti, T. Hamoir, M. De Smet, H. Piryns, N.V. Driessche, F. Marris, H. Hindriks, P.J. Schoenmakers and D.L. Massart, Chemom. Intell. Lab. Syst., 11 (1991) 27.
[26] F.A. Settle, Jr. and M.A. Pleva, Chemom. Intell. Lab. Syst., 11 (1991) 13.
[27] Z. Hippe, Artificial Intelligence in Chemistry, Elsevier, Amsterdam, 1991.
[28] W. Jennings, Analytical Gas Chromatography, Academic Press, San Diego, CA, 1987.

ELSEVIER

Analytica Chimica Acta 296 (1994) 33–41

ANALYTICA CHIMICA ACTA

Knowledge acquisition for fault diagnosis in gas chromatography

Hai Du, Martin J. Stillman *

Department of Chemistry, University of Western Ontario, London, Ontario N6A 5B7, Canada

Received 17th January 1994

Abstract

Knowledge acquisition is frequently considered to be a major bottleneck in the development of an expert system. An explanation-based method was applied to knowledge compilation in a gas chromatographic diagnostic system, GCdiagnosis. Initially, generally accepted gas chromatography fundamental variables were identified to provide the necessary understanding of the physico-chemical processes of a gas chromatography system with flame ionization detection. A theoretical model was built, which consisted of six components, namely carrier gas, injection, injector, column, fuel gas and detector. Interactions between each component were defined. Through the operation of this model, the knowledge used for diagnosing problem GC data can be generated by tracing changes in a specific component that is found to be faulty. Consistency verification of the compiled knowledge was carried out using a prototype. The knowledge is available for extension in future versions of GCdiagnosis.

Keywords: Gas chromatography; Expert systems; Knowledge acquisition

1. Introduction

Continuous monitoring of the environment is extremely costly, requiring analysis of great many samples on a routine basis. Automation of the necessary analytical and interpretative steps in the analytical laboratory offers a solution to these high costs. Expert systems can be used to replace the technician or analyst for routine tasks. We have previously described the development of an expert system for use in metal analysis by atomic absorption spectrometry [1,2] and an expert system for use in interpreting results from analysis by gas chromatography–mass spectrometry (GC–MS) [3]. In this study we describe the theoretical background used to build the knowledge base in a diagnostic program to be used in the total automation of analyses by gas chromatography.

An expert system is a program that can offer intelligent advice and, on demand, justify its own line of reasoning [4–6]. This can lead to increased productivity with greater automation. Generally, there are two ways of developing an expert system: either the knowledge engineer can learn the expertise or the expert can learn how to code the knowledge [7]. In our laboratory, we have chosen the latter method as our technique to develop a number of expert systems that are designed for use in the analytical laboratory [1–3,8]. We have also used this approach in the development of a GC diagnostic expert system, GCdiagnosis [9].

* Corresponding author.

SSDI 0003-2670(94)00232-B

The first step in building an expert system is knowledge acquisition. Generally, an expert system manipulates facts and symbolic relationships that are represented in some form of knowledge base. Knowledge acquisition is an important step, yet it is frequently a major bottleneck in development of an expert system. In this phase, the knowledge extracted from domain experts (experienced chromatographers), system developers (analytical chemists), and documented information that aids in diagnosing gas chromatography [10–16], is combined to create the knowledge base. In this present study, we have recognized that experienced chromatographers use heuristic knowledge based on their experience, to solve problems. When this knowledge is not available for a new case, they use their knowledge of past examples or cases of problems combined with their theoretical knowledge of the properties of the components of the system together with the expressions controlling the physico-chemical process of the system to analyse the case to reach a solution. Experts then reorganize and update their existing heuristic knowledge. Obviously, it is necessary to apply suitable methods that can access this knowledge. It is also necessary that the acquired knowledge is represented in a specific format so that analytical chemists can accumulate their experience of gas chromatography practice. In addition, a well-organized knowledge base will allow the knowledge engineer to more easily understand the domain expertise. Because an expert system is a dynamic system, it must be possible to expand the knowledge base on a continuous basis.

All these problems prompted us to examine an explanation-based learning approach. In this study we applied an explanation-based learning to the analysis of problems in analysis by gas chromatography for knowledge acquisition purposes. Usually, an explanation-based learning method involves goal concepts, training examples, domain theory, and operationality criteria [17]. This paper describes the use of an explanation-based reasoning method for knowledge compilation for use in the diagnostic expert system, GC-diagnosis.

2. Explanation-based learning method

2.1. Descriptions of related theory

Many quantitative or qualitative relations have been used [10–15] to describe the physico-chemical characteristics of gas chromatography. Those described in the following are the more important fundamental variables in the description of the physico-chemical processes of a GC system with a flame ionization detector: S = peak area measured by flame ionization detector; R = the resolution between two adjacent bands; k = the capacity factor of a solute; β = the column phase ratio; n = number of theoretical plates; u_0 = linear velocity of a carrier gas. The equations used to calculate these variables are listed in the Appendix.

2.2. A model of the knowledge domain

We can describe the processes that take place during gas chromatographic separation and detection by flame ionization detector (FID) as follows:

(1) When a sample is introduced into the injector it is carried downstream by the carrier gas.

(2) A solute moves at a speed that is proportional to the carrier gas velocity, but is slower, and depends on the strength of its interactions with the stationary phase.

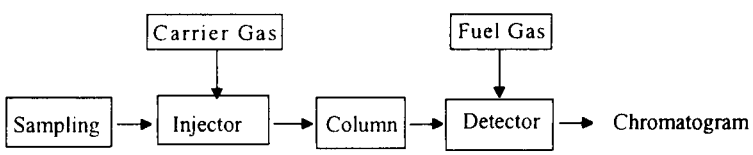


Fig. 1. Components of a GC system with flame ionization detector. Each component has its own inputs and outputs.

(3) A flame ionization detector provides a signal that indicates the presence of the compound. Based on the observed peak area and retention time of the compound, both quantitative and qualitative analysis can be made.

On the basis of this description the simple model shown in Fig. 1 was built to describe the basic structure and function of GC. This model represents the basic structure of the knowledge base. The model consists of six major components: carrier gas, sample and injection, injector, column, fuel gas and detector, each of which has a clear input-output definition. This description is essential in the development of a knowledge base for an expert system, as we shall see below, we can cluster goals and subgoals together for later use by the inference engine. The relation between inputs and outputs of a component can be represented as

$$\text{Component X}(I_1, I_2, \ldots I_m | O_1, O_2, \ldots O_n)$$

Where I_i are inputs and O_j are outputs of component X. The input of one component may be the output of another component. For this study we have defined the following components in terms of associated I_i and O_j parameters

Carrier gas (|Carrier gas flow)
Fuel gas (|Fuel gas flow)
Sampling (Sample, Syringe, Injection|Sample)
Injector (Carrier gas flow, Sample|Carrier gas, Gas phase sample)
Column (Carrier gas, Migration of gas sample|Elution of sample peaks)
Detector (Fuel gas, Elution of sample peaks|Chromatogram)

This description clearly defines the relationships between different components, which permits forward and backward reasoning during knowledge compilation.

2.3. Causal analysis

Because the gas chromatography system is an integrated unit, whenever a new value of one variable is produced by a component that is faulty, this value will be propagated to all connected components. Therefore, by tracing changes due to a faulty component, symptoms caused by the faulty component can be generated and modelled as a predicate in the relationship

$$\text{Cause Y}\{V_1, V_2, \ldots [\text{set of symptoms}]\}$$

Where Y is a specific cause and V_i are changes (high, low and unsteady) of the variables involved or qualitative relationships resulting from the learning set. In order to compensate for uncontrollable variables we use selected past cases of GC diagnosis as a learning set to obtain the

Table 1
Symptoms and causes selected for causal analysis. There are fifteen symptoms and twenty nine related causes that are frequently encountered in routine operation of the gas chromatograph

Symptoms	Causes	
1. Sensitivity loss	1. Leaking syringe	16. Unstable detector temperature
2. Retention time change	2. Dirt in the syringe	17. Double injection
3. High background noise	3. Sample decomposition	18. Column bleed
4. Ghost peaks	4. Dust burned in flame	19. Carrier gas flow too high
5. Peak tailing	5. Contaminated column	20. Contaminated carrier gas
6. Baseline drift	6. Column temperature too low	21. Column degradation
7. Syringe/injection related peak signal	7. Sample size too large	22. Leaking at column exit
8. Leading peak	8. Reaction of sample with GC parts	23. Leaking at septum
9. Unresolved peaks	9. Column temperature too high	24. Severe leaking
10. Band broadening	10. Wrong fuel gas flow rate	25. Injection too slow
11. Cannot zero baseline	11. Unstable column temperature	26. Adsorption of sample on active sites
12. Irregular spikes	12. Carrier gas flow too low	27. Late elute peak
13. Split peaks	13. Injector temperature too low	28. Dirt in the injector
14. No peak(s)	14. Flame not ignited in FID	
15. Sudden drop-off of normal peak	15. Contaminated detector	

qualitative relationships. In this way, qualitative analysis, which is based on related equations or relations, can be performed within related components. This is the basis of the explanation-based system needed to describe the knowledge domain to be coded into the knowledge base of GCdiagnosis.

3. Definition of the knowledge domain in terms of cases

Fifteen symptoms and twenty-eight related causes were selected as cases (Table 1). The requirements for selection were: (1) the chromatographic symptoms should be chromatogram-related, (2) the symptoms and causes should be those frequently encountered in routine operation of the GC, and (3) the set should contain a minimum number of symptoms, i.e., this minimum set of symptoms defined all related causes completely and correctly. Symptoms are classified as chromatogram-related or action-related. A symptom is considered to be chromatogram-related when it is observed in the chromatogram; when an operation condition needs to be changed it is action-related. For example, "High background noise" is chromatogram-related, and "Syringe/injection related peak signal" is action-related.

4. Computer environment

All computer work was carried out on an 80486-based computer, VEISA ALR (Advanced Logic Research, Irvine, CA), configured with 17 Megabytes of RAM, two 1.44-Megabyte floppy disks, a 200-Megabyte hard disk, and VGA graphic monitor. The software package KDS3 (KDS, Wilmeter, IL) was used to verify the consistency of the compiled knowledge.

5. Results and discussion

5.1. Relations between variables and symptoms

The changes in the values of the fundamental variables in our model were condensed by representing the current status in terms of either high (+), low (−), or unsteady (±). This operation set, {−, ±, +}, defines the values of variables. The major concern about using high, low, and unsteady as the operation set is that when a component in this type of coupled system is faulty, we may not be able to determine the precise values of related variables or parameters. Therefore, further calculations and analysis are impossible. In this situation, using unsteady is a good choice that makes the further calculations and analysis possible.

In diagnosis of problem GC data only changes in these parameters are of interest. Generally, one change in the GC system can cause one or more fundamental variable changes, and produces one or more symptoms. For example, given an increase in the column temperature, using the equations related to the capacity factor under isothermal condition and resolution between two adjacent bands (see Appendix), it can be concluded that this change will cause the capacity

Table 2
Relations between changes of fundamental variables and symptoms

Variables	Components	Symptoms
$S(-$ or $\pm)$	Detector	Sensitivity loss, no peaks, or unstable detector signals
$k(-)$	Column	Retention time change, $R(-)$
$k(+)$	Column	Retention time change, band broadening
$k(\pm)$	Column	Retention time change, baseline drift
$\beta(-)$	Column	$k(-)$, $n(-)$
$u_0(-)$	Carrier gas, column, detector	Retention time change, or band broadening
$u_0(+)$	Detector, carrier gas, column	Retention time change, no peaks, or unresolved peaks
$n(-)$	Column	Retention time change, $R(-)$, or band broadening
$R(-)$	Column	Unresolved peaks

factor decrease, $k = (-)$, and the resolution decrease, $R = (-)$. This indicates that an increase in column temperature will make components in a mixture pass through the column faster, resulting in bands which will exhibit poor separation. The relations between the changes of variables and symptoms are listed in Table 2. In Table 2, the relations between fundamental variables and components that may be described by these variables was established based on the relationships between different components and fundamental variables described previously, in order to facilitate operation of the theoretical model.

5.2. Qualitative relations

The efficiency of the separation of solutes by gas chromatography is affected both by controllable variables and uncontrollable variables. The controllable variables include the choice of mobile phase, the flow rate, the column chemistry, and the temperature. The uncontrollable variables include the signal noise, drift, and the column performance, variables that arise from uncertain factors. The uncontrollable variables become especially important when any component malfunction occurs. Therefore, analyzing the physico-chemical processes that occur in the gas chromatographic systems during the separation and detection of volatile compounds requires the integration of a range of knowledge and reasoning techniques. The selection of relations that define these processes and an understanding of how the primitive devices interact to change the chromatogram requires an accurate knowledge of the GC-FID system.

The fundamental variables can be used to describe the physico-chemical processes that take place in the GC-FID system. However, gas chromatography theory is complicated by the presence of non-linear effects and uncontrollable variables. In order to compensate for the uncontrollable variables we use selected past cases of GC diagnosis as a learning set. For example, a large sample size can cause a leading peak because the solute vapour pressure is too low to maintain sufficient solute in the vapour phase; a tailing peak may result from gas flow problems, or from active sites that may be present in the inlet and on the column; and the injector temperature will determine the rate of vaporization of sample introduced by the syringe. It is expected that an injection band profile with a very steep leading edge will be followed by an exponential decay. The decay may result from slow vaporization. Table 3 lists the relations used in this study.

5.3. Causal analysis for knowledge compilation

The gas chromatography system is an integrated unit, that is, a change in one component's properties will affect the properties of other components. Further, it is assumed that changes in one component can only affect the components downstream. For example, the component, Carrier gas, involves four components: carrier gas,

Table 3
Relations resulted from selected past cases of GC diagnosis

Fault components	Symptoms
1. Sample size too large	Leading peak or Peak split
2. Column temperature too low, $k(+)$	Peak tailing
3. Active sites in injector or on the column	Peak tailing and Retention time change
4. Injector temperature too low	Band broadening and Peak tailing
5. Sample decomposition	Ghost peaks, Peak split
6. Impurities or dirt	Noise, Ghost peak, or Peak split
7. Carrier gas flow too low, $u_0(-)$	Leading peaks
8. Injection action	Profile of sample stream [a]

[a] The injection action will affect the column sample input pulse. In analytical applications, only a small amount can be introduced as δ Dirac pulse [18]. If the injection is slow, it will result in wider input pulse. If the injection is double clicked, the input pulse will become double pulses.

Table 4
Results of causal analysis.

1. Contaminated carrier gas	{impurities, PN [a](+) [High background noise, Ghost peaks]}
2. Leaking syringe	{$S(-)$ [Sensitivity loss, Syringe/injection related peak signal]}
3. Dirt in the syringe	{PN(+) [Ghost peaks, Syringe/injection related peak signal]}
4. Sample decomposition	{$S(-)$, PN(+), split peak (±) [Sensitivity loss, Ghost peaks, Split peaks (±)]}
5. Late elution peak(s)	{PN(+), $k(+)$ [Ghost peaks, Band broadening]}
6. Sample size too large	{leading peak, split peak (±) [Leading peaks, Split peaks(±)]}
7. Dust burned in flame	{spikes [Irregular spikes]}
8. Column temperature too high	{$k(-)$ [Retention time change, Unresolved peaks]}
9. Column temperature too low	{peak tailing, $k(+)$ [Retention time change, Peak tailing, Band broadening]}
10. Adsorption of sample by active sites	{retention time change and peak tailing [Retention time change, Peak tailing]}
11. Reaction of sample with GC parts	{$S(-)$, PN(+), adsorption [Sensitivity loss, Retention time change, Ghost peaks, Peak tailing]}
12. Wrong fuel gas flow rate	{noise, canot zero baseline [Sensitivity loss, Cannot zero baseline]}
13. Unstable column tempearature	{$k(\pm)$ [Retention time change, Baseline drift]}
14. Unstable detector temperature	{$S(-$ and $\pm)$ [Sensitivity loss, Baseline drift]}
15. Carrier gas flow too low	{$u_o(-)$, leading peak [Retention time change, Leading peaks, Band broadening (±)]}
16. Carrier gas flow too high	{$u_o(+)$, $n(-)$ [Retention time change, High background noise, Unresolved peaks]}
17. Carrier gas flow too high (severe)	{$u_o(+)$, $n(-)$ [Retention time change, High background noise, Unresolved peaks, Sudden drop of normal peak]}
18. Injector temperature too low	{slow vaporization of sample [Peak tailing, Unresolved peaks, Band broadening]}
19. Flame not ignited in FID	{no peak signal [No peak(s)]}
20. Contaminated detector	{$S(-)$, noise [Sensitivity loss, High background noise, Baseline drift, Cannot zero baseline, No peak(s) (±)]}
21. Dirt in injector	{noise, peak split (±) [high background noise, split peaks (±)]}
22. Dirt in injector	{noise, PN(+) [High background noise, Ghost peaks, Peak tailing]}
23. Double injection	{profile of sample stream [Syringe/injection related peak signal, Split peaks]}
24. Injection to slow	{profile of sample stream [Syringe/injection related peak signal, Band broadening]}
25. Column bleed	{stationary phase bleed [Baseline drift, Cannot zero baseline]}
26. Column bleed (severe)	{stationary phase bleed, $\beta(-)$ [Retention time change, Baseline drift, Cannot zero baseline]}
27. Leaking at septum	{$S(-)$, $u_o(-)$, impurities [Sensitivity loss, Retention time change, Ghost peaks]}

Table 4 (continued)

28. Leaking at column exit	{$S(\pm)$, imputities [Sensitivity loss, High background noise, Baseline drift]}
29. Severe leaking	{$S(-)$, $u_o(-)$, out of detector limit [b] [Sensitivity loss, Retention time change, No peak(s)]}
30. Column degradation	{$\beta(-)$ [Retention time change, Leading peaks, Unresolved peaks, Band broadening]}
31. Contaminated column	{noise, active sites [High background noise, Peak tailing]}

[a] PN means peak number in the chromatogram.
[b] It means that the concentration is lower than the detector limit due to leaking

injector, column and detector. But when the cause of the problem is leaking at septum, then only components downstream, the Injector, the Column and the Detector, will be affected. In terms of the relations discussed above, it is possible to derive information from analysis of the chromatogram that results whenever one of the components is faulty.

For example, for the causes leaking at septum, and leaking at column exit, the common factor is that leaking results in escape of the sample into air, that is, $S(-)$, so sensitivity loss will be a common symptom for leaking. Leaking at septum also causes loss of carrier gas, $u_0(-)$, and which then results in contamination of the injector. Meanwhile leaking at column exit will produce a sample stream passing through the ionization flame that is unstable due to exposure of the stream to air that may bring impurities into the flame. These two causes can then be described as follows:

Cause 27	Leaking at septum {
	$S(-)$, $u_0(-)$, Impurity deposits
	[Sensitivity loss, Retention time change, Ghost peak]
	}

Cause 28	Leaking at the column exit {
	$S(-)$, Impurities from air, Unstable flame
	[Sensitivity loss, High background noise, Baseline drift]
	}

Here the cause number is the corresponding number listed in Table 4, where the results of the complete causal analysis are listed. The causes, 16 and 17, 21 and 22, or 25 and 26 are same but are coded to enable different fuzzy quantitation. For instance, usually high carrier gas flow will cause retention time changes and unresolved peaks; if the flow is very much too high, the flame of the detector will be extinguished by the strong flow. It is obvious that (i) impurities deposited inside the injector will cause ghost peak(s), (ii) a decrease in carrier gas flow will result in longer gas hold-up times, (iii) impurities from the surrounding atmosphere are one source of high background noise, and (iv) an unstable flame can generate unstable signals that make the baseline of the chromatogram drift.

Recently, the term, grey analytical systems, has been proposed [19,20] for multicomponent systems with incomplete knowledge. In our opinion, a gas chromatography system may be considered as a grey system, because of the presence of both controllable and uncontrollable variables. This system is not black, because the theory for the separation of solutes by gas chromatography is well known. It is not white either, because there exist uncontrollable variables. With the development of GC theory and practice, this system will become more white. As a consequence, the model, equations, and relations described here will be modified to reflect this development. Explanation-based learning, a deductive method that uses domain theory, goal concepts, a learning set, and operational criteria to describe the targeted system, is a good choice for the description of a GC system, at least it is good for knowledge acquisition purpose of gas chromatographic diagnosis.

5.4. Consistency verification

Each of the results listed in Table 3 can be called a primitive rule, which consists of symp-

toms and a cause in a predicate. Generally it is impossible to correctly diagnose a fault based on only one observed symptom, because there is logical connection between most symptoms and more than one cause. For example, when the symptom Ghost peak is observed, the causes may be Contaminated carrier gas, Dirt in the syringe, Late elution peak, Reaction of sample with GC parts, Dirt in injector, or Leaking at septum. This means that it is necessary to add some negative observations into the primitive rule set (such as the symptom Sensitivity loss is not observed) in order to form a complete rule set. In this study, we can use any symptom not listed in a predicate as a negative observation, because the knowledge base here is only specific to fault diagnosis in a GC system in which the relationships between its component are all well defined through the proposed model. In practice, only some of them are selected [9].

The major purpose of consistency verification is to check whether the compiled knowledge base is consistent, that is, are the primitive rules independent of each other so that one rule can be totally distinguished from others? If not the inference engine will produce incorrect conclusions. Another purpose of this study was to select a tool for managing the structure of a knowledge base when it was expanded. Inconsistencies may be introduced by oversimplification of the relationships between each of the components in a GC-FID system. Although it may be relatively easy to check for inconsistencies within a small knowledge base, it is very difficult to check a large knowledge base. Obviously, a computer-based tool can verify the knowledge consistency automatically, solving the problem.

KDS3 is a suitable tool for this task. KDS3 is an inductive expert system shell. We have previously used this shell to develop a prototype for diagnosing the cause of problem AAS data using a three level construction technique to build the knowledge base [8b]. In terms of modification of the knowledge base and inference ability [8b], the performance of the KDS shell is outstanding. These capabilities allow the prototype system developed by KDS to be highly efficient. Completion of the compilation of the knowledge base using KDS3 indicates that the relationships between causes and symptoms described in this paper are not in conflict.

6. Conclusions

In this study, we started with the generally accepted gas chromatography theory and a model that describes the physico-chemical processes of a GC system with ionization flame detector. Our efforts focused on identification of key relations that can be used to describe the separation process correctly. Using this model together with a learning set deduced from past cases of GC diagnosis, a heuristic knowledge base for GC diagnosis was compiled. The results indicate that it is possible to compile heuristic knowledge based on a model describing the physico-chemical process of gas chromatography and refined past cases of GC diagnosis. From our experience, we believe: (1) use of explanation-based methods generate knowledge efficiently, (2) the results (Table 3) can make the knowledge most accessible to both the expert and the knowledge engineer, and (3) the knowledge is easily convertable to knowledge domain matrix presentation. Logical checks carried out on the knowledge base shows that the relationship obtained based on this method are consistent. With increasingly accurate knowledge of chromatography, the model will be expanded and refined. This study will provide analytical chemists with a tool to accumulate their experience of gas chromatography practice.

A graphical-based system has been developed within the MS Windows environment. The system provides users with a friendly interface and the operator can use the mouse to select specific symptom set [9].

Acknowledgments

The authors gratefully acknowledge the financial support from the Ontario Ministry of the Environment and NSERC of Canada through both operating and equipment strategic grants (to

MJS). We are also grateful for funding from the University of Western Ontario (the President's Fund in support of a summer stipend for H. Du).

Appendix 1

The column phase ratio:

$$\beta = \frac{V_l}{V_m}$$

The capacity factor of a solute:

$$k = \frac{t_R - t_m}{t_m} = \frac{t'_R}{t_m}$$

Under isothermal conditions: $\log k = a + b/T$

The distribution constant: $K = k\beta$

The Van Deemter equation:

$h = A + B/u_0 + C \times u_0$

Theoretical plate numbers: $n = L/h$

The resolution of two adjacent bands 1 and 2:

$$R = \frac{n^{1/2}}{4}\left(\frac{\alpha - 1}{\alpha}\right)\left(\frac{k}{k+1}\right)$$

$$\alpha = t'_{R(2)}/t'_{R(1)} = k_2/k_1$$

The response of the flame ionization detector: $S \propto C_m$

Symbols used: T = the column temperature in Kelvin; V_m = the volume in the column occupied by the mobile phase in cubic micrometer; V_l = the volume in the column occupied by the stationary phase in cubic micrometer; t_m = the time required for elution of a nonretained substance (gas hold-up time) in seconds; t_R = the retention time, the time for the solute to emerge from the column after injection in seconds; t'_R = the adjusted retention time in second; L = the column length in millimeters; C_m = the concentration of a solute in ppm or ppb. a, b, A, B and C are constants.

References

[1] (a) W.R. Browett and M.J. Stillman, Prog. Anal. Spectrosc., 12 (1989) 73; (b) W.R. Browett, T.A. Cox and M.J. Stillman, ACS Symp. Ser., 408 (1989) 210.
[2] S. Lahiri and M.J. Stillman, Anal. Chem., 64 (1992) 283A.
[3] D.A. Cancilla, G. Huang, S. Ma and M.J. Stillman, Proceedings of the 40th ASMS Conference, Washington, DC, June, 1992, p. 993.
[4] T.H. Pierce and B.A. Hohne (Eds.), Artificial Intelligence Application in Chemistry, ACS Symp. Ser. 306, Washington, DC, 1986.
[5] B.A. Hohne and T.H. Pierce (Eds.), Expert System Application in Chemistry, ACS Symp. Ser. 408, Washington, DC, 1989.
[6] J.M. Hushon (Ed.), Expert Systems for Environmental Applications, ACS Symp. Ser. 431, Washington, DC, 1990.
[7] M. Mulholland, N. Dunand, A. Cleland, J.A. van Leeuwen and B.G.M. Vandeginste, J. Chromatogr., 485 (1989) 283
[8] (a) M.J. Stillman, Encyclopaedia of Analytical Science, Academic Press, London, 1994, in press.
[9] H. Du, S. Lahiri, G. Huang and M.J. Stillman, Anal. Chim. Acta, 296 (1994) 21.
[10] L.G. Blomberg, Adv. Chromatogr., 26 (1987) 229.
[11] J.Q. Walker, M.T. Jackson, Jr. and J.B. Maynard, Chromatographic System – Maintenance and Troubleshooting, Academic Press, New York, 1972.
[12] W. Jennings, Analytical Gas Chromatography, Academic Press, San Diego, CA, 1987.
[13] G. Guiochon and C.L. Guillemin, Quantitative Gas Chromatography, Elsevier, Amsterdam, 1988.
[14] R.E. Clement (Ed.), Gas Chromatography, Wiley, New York, 1990.
[15] R.M. Smith, Chromatography in Analytical Chemistry, Wiley, Chichester, 1988.
[16] Troubleshooting Guide for Capillary Gas Chromatography, Supelco, Bellefonte, PA, 1988.
[17] R.A. Michalski et al. (Eds.), Machine Learning, Vol. 2, Tioga Publishing, Palo Alto, CA, 1986.
[18] A.M. Katti and G.A. Guiochon, Adv. Chromatogr., 31 (1992) 1.
[19] Y. Xie, Y. Liang and R. Yu, Anal. Chim. Acta, 276 (1993) 455.
[20] Y. Liang, O.M. Kvalheim and R. Manne, Chemom. Intell. Lab. Syst., 18 (1993) 235.

ELSEVIER

Analytica Chimica Acta 296 (1994) 43–50

ANALYTICA CHIMICA ACTA

Simultaneous determination of 5′-deoxy-5-fluorouridine, 5-fluorouracil and 5,6-dihydro-5-fluorouracil in serum by liquid chromatography with diode array UV detection

A. Guerrieri [b], F. Palmisano [a,*], C.G. Zambonin [a], M. De Lena [c]

[a] *Analytical Chemistry Laboratory, Department of Chemistry, University of Bari, via Orabona No. 4, 70126 Bari, Italy*

[b] *Analytical Chemistry Laboratory, Department of Chemistry, University of Basilicata, via N. Sauro No. 85, 85100 Potenza, Italy*

[c] *Istituto Oncologico, via Amendola No. 209, 70126 Bari, Italy*

Received 8 February 1994; revised manuscript received 28 June 1994

Abstract

5,6-Dihydro-5-fluorouracil (5-FUH_2), the main catabolite of the anticancer drug 5-fluorouracil (5-FU), can be quantitated in human serum simultaneously with the pro-drug 5′-deoxy-5-fluorouridine (5′-dFUR), the parent drug 5-FU and the anabolites 5-fluorouridine and 2′-deoxy-5-fluorouridine by gradient elution-reversed phase liquid chromatography coupled to diode array UV detection. The analytes were extracted (after protein precipitation with ammonium sulphate) with ethyl acetate–isopropanol (90:10, v/v). The average recovery for 5-FUH_2 was 78.0% at 0.5 μg/ml level. 5-FUH_2 response was linear over three concentration decades and its detection limit was 200 ng/ml of serum. The analytical procedure has been used for simultaneous monitoring of 5′-dFUR, 5-FU and 5-FUH_2 levels in the serum of a patient under 5′-dFUR chemotherapy.

Keywords: Liquid chromatography; Fluoropyrimidine drugs

1. Introduction

Fluorinated pyrimidines and their nucleosides are well known to show a significant cytotoxic activity. Among them, 5-fluorouracil (5-FU), considered as one of the most active anticancer drugs [1,2], is widely used for the treatment of some solid tumours of breast, colon and rectum. Unfortunately, 5-FU is also known to cause a number of undesired toxic effects. To improve the therapeutic index of 5-FU, several fluorinated pyrimidine analogues have been designed as 5-FU pro-drugs. 5′-Deoxy-5-fluorouridine (5′-dFUR) is a relatively new synthesized fluoropyrimidine nucleoside [3] capable of releasing the active 5-FU preferentially [4–7] in tumour cells. 5′-dFUR shows an high antineoplastic activity against a broad spectrum of human tumours and presents [8] a lower toxicity and a higher therapeutic index with respect to 5-FU.

Since 5′-dFUR is a pro-drug of 5-FU, the mechanism of 5′-dFUR cytotoxicity is strictly similar to that reported for 5-FU [7–10]. After its release, 5-FU can be anabolised by two possible pathways. It can be reversibly converted, by uridine phosphorylase, to 2′-deoxy-5-fluorouridine (2′-dFUR) which is phosphorylated, by thymidine kinase, to 2′-deoxy-5-fluorouridine-5′-monophosphate (2′-dFUMP) which inhibits DNA synthesis by blocking the thymidilate synthetase system. The alternative anabolic pathway

* Corresponding author.

SSDI 0003-2670(94)00292-4

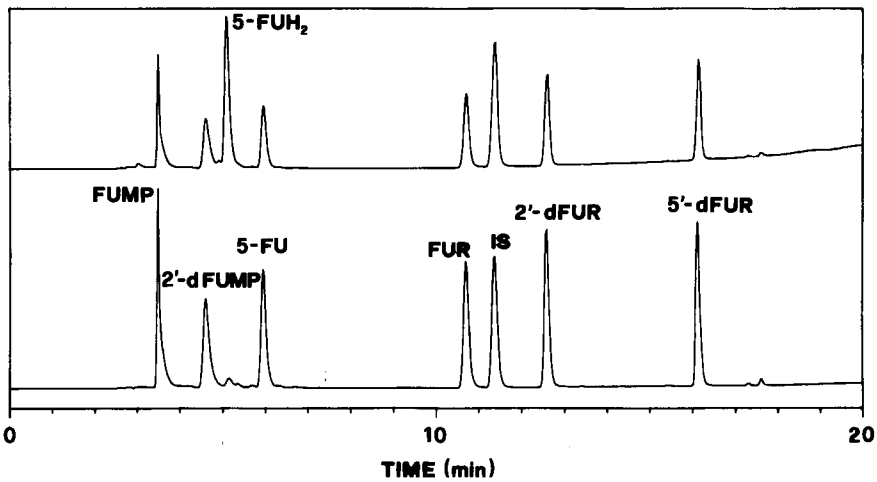


Fig. 1. Dual wavelength detection chromatogram of a mixture of fluoropyrimidine standards. Injected amounts: FUMP (500 ng), 5-FUH_2 (800 ng), 5-FU (200 ng), 2′-dFUMP, FUR, IS (5-BrU), 2′-dFUR, 5′-dFUR (all 400 ng). Detection wavelengths: 269 and 220 nm, lower and upper chromatogram, respectively. Detection wavelength band-width: 4 nm; reference signal: 450 nm (80 nm band-width); absorbance axis: 160 mAU full scale.

involves the conversion of 5-FU to 5-fluorouridine monophosphate (FUMP) either directly by a phosphoribosyl transferase or stepwise via 5-fluorouridine (FUR); FUMP is successively diphosphorylated and reduced to 2′-deoxy-5-fluorouridine diphosphate (2′-dFUDP) which is finally dephosphorylated to 2′-dFUMP.

Although extensive studies have clarified the importance of anabolic pathways on drug cytotoxicity, the knowledge of the role of catabolic pathways on 5-FU cytotoxicity is still limited [11–15]. Liver seems the major site of 5-FU catabolism the initial step being the reductive degradation to 5,6-dihydro-5-fluorouracil (5-FUH_2) by dihydrouracil dehydrogenase (DUD), followed by ring splitting to α-fluoro-β-ureidopropionic acid (FUPA), a compound readily excreted in urine. In the successive catabolic step, FUPA is catabolized to α-fluoro-β-alanine with final release of ammonia, urea and carbon dioxide. Contrary to earlier reports stating that none of these degradation products had a significant toxicity and/or antitumour activity, more recent works suggest that 5-FUH_2 possesses a significant cytotoxic activity. In fact, 5-FUH_2 seems to produce inhibition of thymidilate synthetase activity in Ehrlich ascites tumour cells [16] suggesting that it may also contribute to 5-FU toxicity, possibly by being anabolized to 5-FU. Since the reductive degradation by DUD is the rate-limiting step [17–19] in 5-FU catabolism, monitoring 5-FUH_2 levels can be usefully employed to estimate 5-FU availability for its anabolic route [20] as well as to evaluate the efficiency in DUD inhibition by certain drugs in the attempt of enhancing [21] 5-FU chemotherapeutic efficacy. Furthermore, 5-FUH_2 levels determination after 5-FU (or 5′-dFUR) administration in humans can help to design a more complete pharmacokinetic model of 5-FU and to set up a more effective dosage regimen [22].

Several analytical procedures have been developed for fluoropyrimidines analysis (see Ref. [23] for a review). ^{19}F-NMR, for example, allows detection [24,25] of all catabolites in the biofluids of patients under 5′-dFUR or 5-FU chemotherapy; this technique does not require any extraction or derivatization of biological samples but is limited by its low sensitivity. Gas–liquid chromatography (GLC) [26], capillary gas chromatography (CGC) coupled to nitrogen–phosphorus detection (NPD) [27] or electron-capture detection (ECD) [28] and gas chromatography–mass spectrometry (GC–MS) [29] methods have been also developed for the simultaneous determination of 5-FU and 5-FUH_2 in plasma but, unfortunately, these techniques do not possess a linear range as wide as required for pharmacokinetic studies and the possibility of a simultaneous determination of 5′-dFUR and/or some 5-FU anabolites is not mentioned. Liquid chromatography (LC) is the most widely applied technique in the determination of 5-FU, 5′dFUR and related nucleosides and nucleotides [23,30]; however, few studies

have been up to now addressed to 5-FUH_2 analysis. Sommadossi et al. [31] reported a reversed-phase (RP) ion-pair LC method for 5-FU anabolites and catabolites (including 5-FUH_2). However, the full analyte resolution was obtained by two columns in series and the necessary ion-pair formation required alkaline mobile phases not compatible with silica based columns; moreover, detection was based on radioactivity measurements and hence the method could be used for labelled drugs only. An LC–MS method capable of simultaneous determination of 5-FU and 5-FUH_2 has been also developed [32] even if it suffers of the same limitations already described for GC ones. Finally, capillary electrophoresis (CE) of several fluoropyrimidines including 5′-dFUR, 5-FU and 5-FUH_2 has been reported recently [33].

Recently, a simple and sensitive procedure [30] capable of simultaneous determination of 5′-dFUR, 5-FU, and some of its main anabolites in body fluids have been reported. Based on RP-LC separation of fluoropyrimidines and UV detection, the method appeared to be useful for monitoring the bioavailability of 5′-dFUR and 5-FU in patients undergoing anticancer chemotherapy. In this paper, the potential of diode array UV detection in RP-LC is explored in order to include 5-FUH_2 among the detectable 5′-dFUR metabolites. The results indicate the possibility of a simultaneous determination of several fluoropyrimidines by a single chromatographic run with wide linearity and sufficiently low detection limits. To prove the usefulness of the proposed method, serum samples of a patient under 5′-dFUR chemotherapy are analyzed. The capability of the proposed method to follow the time course of 5′-dFUR, 5-FU and 5-FUH_2 is demonstrated.

2. Experimental

2.1. Chemicals

5-FUH_2 was kindly donated by Hoffmann-La Roche AG (Basel). 5-FUH_2 purity was tested by LC. Stock and diluted solutions were prepared just before use. Other chemicals has been described previously [30].

2.2. Apparatus

The chromatographic system was the same as already described in Ref. [30]. It consisted of a Perkin-Elmer (Norwalk, CT) Model 3B dual pump module equipped with a Rheodyne (Cotati, CA) 7125 injection valve and a 5-μm Supelcosil LC-18-S, ODS column (250×4.6 mm i.d.) (Supelco, Bellefonte, PA). A 5-μm Supelguard LC-18-S pre-column (20×4.6 mm i.d.) (Supelco) was used to protect the analytical column. The detector was an HP 1040A photodiode array spectrophotometer (Hewlett-Packard, Palo Alto, CA) interfaced to an HP 85 computer equipped with an HP 9121 dual disk drive and an HP 7470A plotter. A Perkin-Elmer LCI-100 laboratory computing integrator directly connected to the analog output of the photodiode-array detector was also used.

2.3. Chromatographic and detection conditions

The mobile phase, the gradient program and the chromatographic conditions have been previously described [30]. A binary gradient composed of phosphate buffer (0.05 M, pH 6.5)–methanol (40:60, v/v) (solvent A) and phosphate buffer (0.05 M, pH 6.5) (solvent B) was used. The gradient program was: 15 min concave from 6.7% A to 66.6% A; 5 min isocratic at 66.6% A; 2 min linear to 6.7% A; equilibration time: 10 min. The concave segment was generated according to the following equation:

$$P(t) = P(i) + [P(f) - P(i)](t/T)^2$$

where: $P(t) = A\%/(A\% + B\%)$ ratio at time t; $P(i)$, $P(f)$ are the initial and final $A\%/(A\% + B\%)$ ratio, respectively; T = length of the concave segment.

The flow rate was 1 ml/min, the injection volume was 20 μl and the temperature was ambient.

Unless otherwise stated, spectra were acquired in the 210–400 nm range at the apex and on the ascending or descending part of each peak. 5′-dFUR, 5FU and its anabolites [30] were monitored at 269 nm (4 nm band-width, reference signal 450, 80 nm band-width) whereas 5-FUH_2 detection was accomplished by the technique of peak suppression [34] at 220 nm (4 nm band-width, reference signal 242, 6 nm band-width). The threshold value was 1 mAU (absorbance unit). The technique of spectra overlaying after normalization was used to check peak purity.

2.4. Sample collection and extraction procedure

Serum samples from patients with advanced colorectal cancer under 5′-dFUR chemotherapy were col-

lected at the Istituto Oncologico (Bari, Italy). To prevent 5-FUH_2 instability, serum samples were stored and processed according to Ref. [20].

Serum deproteinization and liquid–liquid extraction (LLE) procedures, already described in Ref. [30], were used to extract 5-FUH_2 from serum samples. 250 μl of serum and 20 μl of internal standard solution were mixed in a tapered tube and 500 μl of a saturated ammonium sulphate solution were added, followed by a brief vortex-mixing. 5 ml of an ethyl acetate–isopropanol mixture (90:10, v/v) were successively added and the resulting mixture was vigorously shaken for 5 min, centrifuged at 4000 rpm for 5 min and the organic phase carefully transferred into a tapered tube. The extract so obtained was evaporated to dryness at room temperature under a gentle stream of nitrogen with a Visiprep Vacuum Manifold coupled to a Visidry Drying Attachement (Supelco) that permitted simultaneous processing of at least 12 extracts. Typically, the extract was reconstituted with 250 μl of mobile phase and 20 μl were injected.

2.5. Quantitation

Quantitation was performed by the internal standard method. Calibration curves were obtained using 12 standards prepared by spiking drug-free serum samples with 2.5 μg/ml of 5-bromouracil (5-BrU, the internal standard) and variable amounts of 5′-dFUR, 5-FU and 5-FUH_2 in order to cover the concentration ranges from 50 ng/ml to 200 μg/ml, 25 ng/ml to 20 μg/ml and from 200 ng/ml to 20 μg/ml, respectively.

Serum 5′-dFUR, 5-FU and 5-FUH_2 concentrations have been calculated by the "multiple calibration curve" [35] approach as already described in Ref. [30].

3. Results and discussion

Fig. 1 shows the dual wavelength detection chromatogram of a mixture of standard fluoropyrimidine derivatives involved in the metabolism of 5′-dFUR. As can be seen, the gradient program, already able to resolve [30] the pro-drug 5′-dFUR, the drug 5-FU and some of the nucleosides and nucleotides involved in the anabolic activation of 5-FU, i.e., FUR, 2′-dFUR,

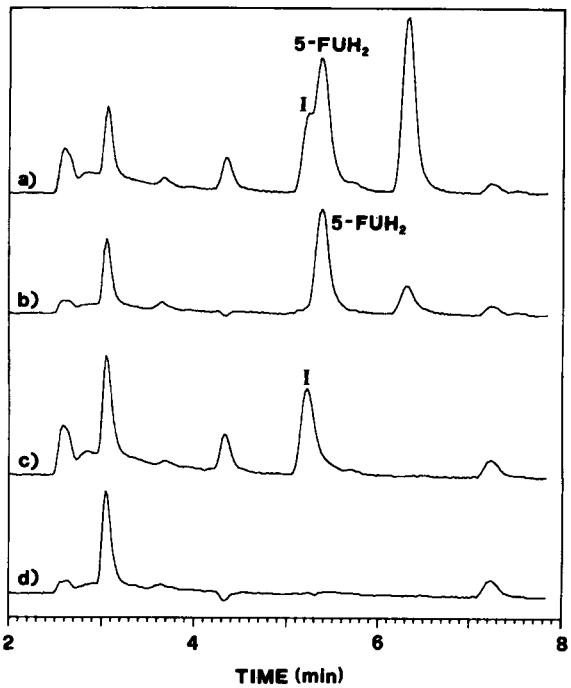


Fig. 2. Chromatograms relevant to extracts of drug-free serum spiked (**a–b**) and unspiked (**c–d**) with 5-FUH_2 at 5 μg/ml level. Chromatograms **b** and **d** were acquired by the "peak suppression" technique (see text) by using a pilot signal at 220 nm (4 nm band-width) vs. a reference signal at 242 nm (6 nm band-width) whereas chromatograms **a** and **c** were acquired by using the same pilot signal but a reference signal at 450 nm (80 nm band-width). **I** indicates an unknown endogeneous interferent. Absorbance axis: 20 mAU full scale.

FUMP and 2′-dFUMP, is also able to resolve the catabolite 5-FUH_2.

A drug-free serum sample spiked with 5-FUH_2 and processed as described in the Experimental section gave the chromatogram shown in Fig. 2a. As can be inferred from Fig. 2c showing the chromatogram relevant to an extract of drug-free serum, an unknown endogeneous component **I** coeluted with 5-FUH_2 making difficult an accurate quantitation of the analyte of interest. Attempts to improve 5-FUH_2 and **I** resolution by changes in the gradient program and/or mobile phase composition revealed unsuccessful as well as the use of different detection wavelengths (both the analyte and the interferent had very similar UV spectra). The interference was minimized by "peak suppression" or "signal subtraction" technique [34], i.e., by choosing a reference wavelength where the interference shows exactly the same absorbance as at the detec-

Table 1
Percentage recoveries of 5-FUH_2 from spiked serum samples

Drug level (μg/ml)	Mean recovery (%)	R.S.D. % ($n=5$)
0.5	78.0	5.5
5.0	83.3	4.2
50	84.0	5.5

tion wavelength of 5-FUH_2. As can be seen from the chromatogram of Fig. 2b, the use of a reference signal at 242 nm (6 nm band-width) for 5-FUH_2 detection at 220 nm (4 nm band-width) resulted in a nearly complete suppression of the chromatographic interference. This conclusion is additionally supported by the inspection of the "blank chromatogram" in Fig. 2d (acquired with the same technique) clearly showing the absence of interfering peaks in the elution region of 5-FUH_2.

The LLE procedure already developed for 5'-dFUR, 5-FU and other anabolites [30] revealed also satisfactory results for the extraction of 5-FUH_2 from serum. Table 1 reports the percentage recoveries of 5-FUH_2 from serum samples spiked at different concentration levels, calculated by direct comparison of peak areas of the analyte and standard dissolved in mobile phase. According to a "one-way analysis of variance" at $P=0.05$, the mean recoveries in Table 1 are not significantly different, i.e., the mean recovery of 5-FUH_2 is independent of its serum concentration in the range explored.

5-FUH_2 response in serum resulted linear over about three concentration decades. Calibration curves showed a correlation coefficient better than 0.999 and an intercept not significantly different from zero at 95% confidence level. A typical calibration curve for 5-FUH_2 gave a slope of 0.517 ± 0.006 mAU ml/μg and an intercept of -0.011 ± 0.059 mAU; the standard error of the fit, $S_{y/x}$, was 0.0311. The within-day ($n=5$) and between-days ($n=6$) coefficients of variation for 5-FUH_2 (at 1 μg/ml level) were 4.3% and 9.0%, respectively.

Fig. 3 shows a chromatogram of a drug-free serum sample spiked with 5-FUH_2 at a concentration level near to its detection limits. At a signal-to-noise ratio of 3 (noise calculated peak to peak in a blank chromatogram at the elution time of 5-FUH_2), the detection limit of 5-FUH_2 was in fact around 200 ng/ml using a 250 μl sample size, a 100 μl reconstitution volume and a 20 μl injection volume. Such limit of detection seems adequate for the monitoring of 5-FUH_2 in serum of patient under doxifluridine chemotherapy [36,37]. A further improvement might be eventually achieved by using a larger sample size and a larger injection volume. Detection limits for 5'-dFUR, 5-FU and other derivatives are those already reported [30].

To demonstrate the usefulness of the proposed method for "real" samples analysis, serum specimens taken from patients under 5'-dFUR chemotherapy have been processed and analyzed; Fig. 4 shows a typical

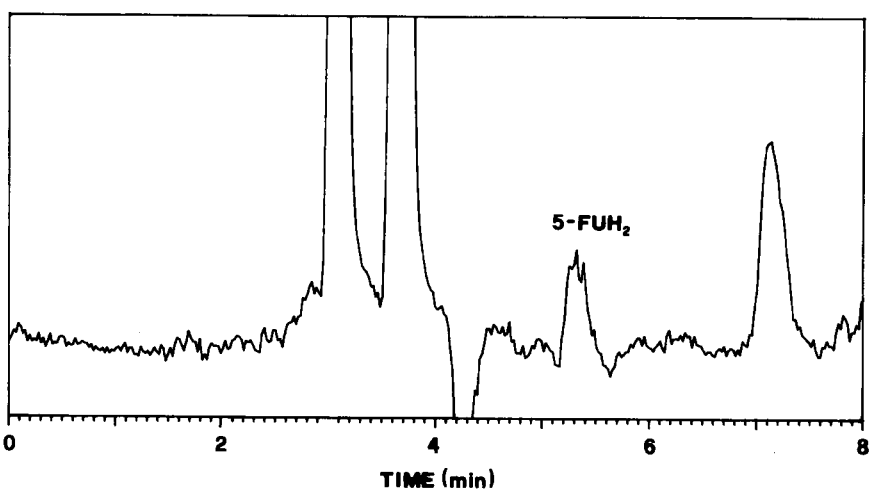


Fig. 3. Chromatogram of the extract of a drug-free serum spiked with 5-FUH_2 at a concentration level (200 ng/ml) near to its detection limit. Peak shown corresponds to 1.6 ng injected. Detection wavelength: 220 nm (4 nm band-width) vs. 242 nm (6 nm band-width). Absorbance axis: 0.5 mAU full scale.

dual wavelength detection chromatogram. As can be seen, 5′-dFUR, 5-FU and 5-FUH_2 deriving from 5-FU catabolic degradation can be easily separated and quantitated in a single chromatographic run.

Although a pharmacokinetic study is outside the scope of this work, an example of the obtainable concentration–time profiles of 5′-dFUR, 5-FU and 5-FUH_2 in a patient under 5′-dFUR chemotherapy is shown in Fig. 5 just to demonstrate the potential and the soundness of the described approach. In fact, Fig. 5 shows that, for the patient under study, 5′-dFUR and 5-FU display a two-compartment open model elimination as already observed in other studies [22,38]. Fig. 5 also shows that 5-FUH_2 reached its peak concentration well after the time at which 5′-dFUR and 5-FU attained their ones; moreover, the apparent elimination half-life of 5-FUH_2 appears greater than those of 5′-dFUR and 5-FU suggesting that the elimination of 5-FUH_2 is mainly controlled by its intrinsic elimination rate constant [39].

Table 2 summarizes the most salient features of the presently available procedures for 5-FUH_2 analysis. As can be seen, GC methods offer the lowest detection limits but with the limitations already discussed in the Introduction. Moreover, although 5-FU and 5-FUH_2 could be determined in one single run by applying a temperature program, two isocratic runs were preferred for routine measurement [27,36,37]. In contrast, LC techniques seem capable of a simultaneous determination of a large number of fluoropyrimidine derivatives.

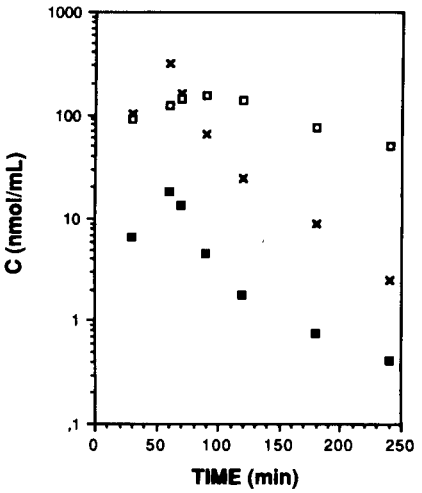


Fig. 5. Concentration–time profiles for 5′-dFUR (×), 5-FU (■) and 5-FUH_2 (□) in a patient under 5′-dFUR treatment (4 g/m^2 by i.v. infusion over 60 min period). Time start from the beginning of the infusion.

In conclusion, the method described in this paper could offer the right compromise between amount of information produced, sensitivity, linear range and simplicity of operation. Although capable of full resolution of several fluoropyrimidines, the method suffers of

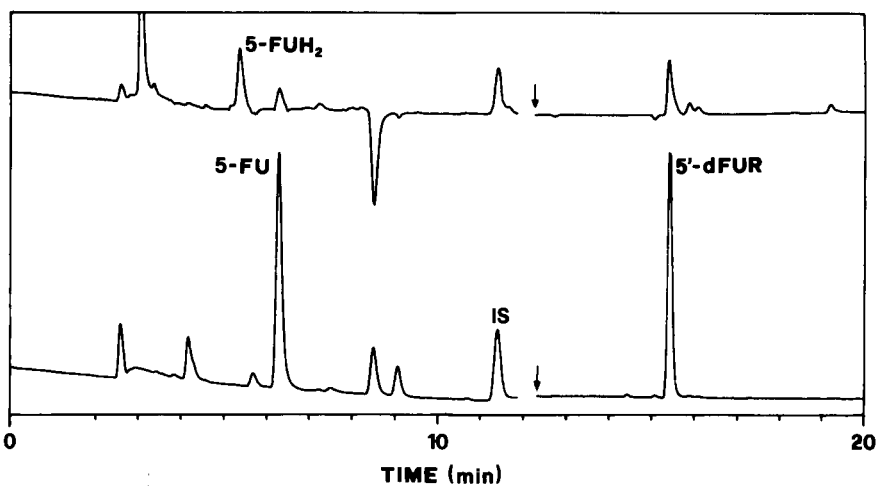


Fig. 4. Dual wavelength detection chromatogram of the extract of a serum sample taken from a patient under 5′-dFUR chemotherapy (4 g/m^2 by i.v. infusion over 60 min period). Sampling time: 120 min from the start of the drug infusion. 5′-dFUR, 5-FU and 5-FUH_2 calculated serum concentrations were 6.14, 0.236 and 17.94 μg/ml, respectively. Detection wavelength: lower chromatogram: 269 (4 nm band-width) vs. 450 nm (80 nm band-width); upper chromatogram: 220 (4 nm band-width) vs. 242 nm (6 nm band-width). Arrows indicate a change in detector sensitivity from 25 to 350 mAU full scale.

Table 2

Method	5-FUH_2 detection limit (ng/ml)	Analytes simultaneously detected	Note	Ref.
^{19}F NMR	1300	5′-dFUR, 5-FU and all catabolites in plasma	No extraction or derivatization of biological sample	24, 25
GC-ECD	6.6	5-FU and 5-FUH_2 in plasma	Plasma LLE required	27–28,
GC-NPD	25 [a]		Linear range: two decades	36–37
GC–EI-MS	80	5-FU and 5-FUH_2 in plasma	Plasma LLE and pentyl derivatization required	29
GC–CI-MS [b]	10			
LC	not reported	5-FU, all anabolites and catabolites in suspended rat hepatocytes	Full chromatographic resolution required two in series columns and alkaline mobile phase. Detection was by radioactivity	31
LC–MS	not reported	5-FU and 5-FUH_2 in blood	5-FU detection limit: 400 ng/ml. Linear range: 2 decades	32
CE-UV detection	not reported	5′-dFUR, 5-FU and 5-FUH_2 separation claimed	Application to biological samples not given	33
LC-diode array UV detection	200	5′-dFUR, 5-FU, 5-FUH_2 FUR, 2′-dFUR in serum	Serum LLE required. Linear range: three decades. Diode array detection	This work

[a] Calculated according to an absolute detection limit of 0.5 ng.
[b] EI-MS: electron impact mass spectrometry; CI-MS: chemical ionization mass spectrometry.

some limitations, i.e., the impossibility of a simultaneous extraction of fluoropyrimidine nucleotides FUMP and 2′-dFUMP. To overcome these limitations, work is in progress by using extraction techniques similar to those very recently developed in this laboratory [40].

Acknowledgements

F. Hoffmann-La Roche AG (Basel) is gratefully acknowledged for providing us a standard of 5-FUH_2. Work was carried out with financial assistance from MURST and "Associazione Italiana Ricerca sul Cancro" (AIRC, Milan).

References

[1] R. Duschinsky, E. Pleven and C. Heidelberger, J. Am. Chem. Soc., 79 (1957) 4559.
[2] C. Heidelberger, N.K. Chandhuri, P. Dannenberg, D. Mooren, L. Griesbach, R. Duschinsky, R.J. Schnitzer, E. Pleven and J. Scheiner, Nature (London), 179 (1957) 663.
[3] A.F. Cook, M.J. Holman, M.J. Kramer and P.W. Trown, J. Med. Chem., 22 (1979) 1330.
[4] H. Ishitsuka, M. Miwa and K. Takemoto, Gann, 71 (1980) 112.
[5] S. Suzuki, Y. Hongu, H. Fukazawa, S. Ichihara and H. Shimizu, Gann, 71 (1980) 238.
[6] A. Kono, Y. Hora and S. Sugata, Chem. Pharm. Bull., 31 (1980) 112.
[7] R.D. Armstrong and R.B. Diasio, Cancer Res., 41 (1981) 4891.
[8] R.D. Armstrong, J. Gesmonde, T. Wu and E. Cadman, Cancer Treat. Rep., 67 (1983) 541.
[9] R.D. Armstrong and R.B. Diasio, Cancer Res., 40 (1980) 3333.
[10] R.D. Armstrong, K.M. Connolly, A.M. Kaplan and E. Cadman, Cancer Chemother. Pharmacol., 11 (1983) 102.
[11] C.E. Myers, Pharmacol. Rev., 33 (1981) 1.
[12] N.K. Chaudhuri, B.J. Montag and C. Heidelberger, Cancer Res., 18 (1958) 318.
[13] N.K. Chaudhuri, K.L. Mukherjee and C. Heidelberger, Biochem. Pharmacol., 1 (1959) 328.
[14] K.L. Mukherjee and C. Heidelberger, J. Biol. Chem., 235 (1960) 433.
[15] K.L. Mukherjee, J.W. Boohar, D. Wentland, F.J. Aasfield and C. Heidelberger, Cancer Res., 23 (1963) 49.
[16] R.B. Diasio, J.D. Schuetz, H.J. Wallace and J.P. Sommadossi, Cancer Res., 45 (1985) 4900.
[17] P. Fritzon, J. Biol. Chem., 237 (1962) 150.
[18] E.S. Canellakis, J. Biol. Chem., 221 (1956) 315.
[19] F.M.N. Naguib, M.H. El Kouni and S. Cha, Cancer Res., 45 (1985) 5405.
[20] N. Van Den Bosch, O. Driessen, E.A. Van Der Velde and C. Erkelens, Ther. Drug Monitor., 9 (1987) 443.
[21] G.C. Daher, F.N.M. Naguib, M.H. El Kouini, R. Zhang, S.J. Soong and R.B. Diasio, Biochem. Pharmacol., 41 (1991) 1887.

[22] J.P. Sommadossi, C. Aubert, J.P. Cano, J. Gouveia, P. Ribaud and G. Mathè, Cancer Res., 43 (1983) 930.
[23] U.R. Tjaden and E.A. De Bruijn, J. Chromatogr., 531 (1990) 235.
[24] A.N. Stevens, P.G. Morris, R.A. Iles, P.W. Sheldon and J.R. Griffiths, Br. J. Cancer, 50 (1984) 113.
[25] J.P. Beteille, A. Lopez, M. Bon, M.C. Malet-Martino and R. Martino, Anal. Chim. Acta, 171 (1985) 225.
[26] B.J. Mc Dermott, H.W. Van Den Berg and R.F. Murphy, Biochem. Soc. Trans., 7 (1979) 65.
[27] E.A. De Bruijn, O. Driessen, N. Van Den Bosch, E. Van Strijen, P.H.Th.J. Slee, A.T. Van Oosterom and U.R. Tjaden, J. Chromatogr., 278 (1983) 283.
[28] E.A. De Bruijn, U.R. Tjaden, A.T. Van Oosterom, P. Leeflang and P.A. Leclercq, J. Cromatogr., 279 (1983) 603.
[29] C. Aubert, J.P. Sommadossi, P. Coassolo, J.P. Cano and J.P. Rigault, Biomed. Mass Spectrom., 9 (1982) 336.
[30] F. Palmisano, F. Berardi, M. De Lena, A. Guerrieri, V. Lorusso and P.G. Zambonin, Chromatographia, 33 (1992) 413.
[31] J.P. Sommadossi, D.A. Gewirtz, R.B. Diasio, C. Aubert, J.P. Cano and I.D. Goldman, J. Biol. Chem., 257 (1982) 8171.
[32] E. Araki, M. Iigo, T. Shirasaka, M. Tera, T. Murata and T. Kitsuwa, Nippon Iyo Masu Superkutoru Gakkai Koenshu, 15 (1990) 185 (C.A., 115 (1991) 63880a).
[33] E.A. De Bruijn, G. Pattyn, F. David and P. Sandra, J. High Resolut. Chromatogr. Chromatogr. Commun., 14 (1991) 627.
[34] L. Huber, Applications of diode-array detection in high performance liquid chromatography, appendix A, pp. 82, Hewlett-Packard Co., Publication Number 12-5953-2330, 1989.
[35] D.G. Mitchell, W.N. Mills, J.S. Garden and M. Zaeb, Anal. Chem., 49 (1977) 1655.
[36] E.A. de Bruijn, A.T. van Oosterom, U.R. Tjaden, H.J.E.M. Reeuwijk and H.M. Pinedo, Cancer Res., 45 (1985) 5931.
[37] E.A. De Bruijn, L. Remeyer, U.R. Tjaden, C. Erkelens, L.M. De Brauw and C.J.H. Van De Velde, Biochem. Pharmacol., 35 (1986) 2461.
[38] J.P. Sommadossi and J.P. Cano, J. Chromatogr., 225 (1981) 516.
[39] P.G. Welling, Pharmacokinetics, ACS Monograph 185, American Chemical Society, Washington, DC, 1986.
[40] A. Guerrieri, F. Palmisano, P.G. Zambonin, M. De Lena and V. Lorusso, J. Chromatogr., 617 (1993) 71.

ELSEVIER

Analytica Chimica Acta 296 (1994) 51–62

ANALYTICA CHIMICA ACTA

A robust liquid chromatographic method for measurement of medium components during penicillin fermentations

L.H. Christensen [a], G. Mandrup [b], J. Nielsen [a,*], J. Villadsen [a]

[a] *Department of Biotechnology, Technical University of Denmark, DK-2800 Lyngby, Denmark*
[b] *Novo Nordisk A / S, Process Control Laboratories, DK-4400 Kalundborg, Denmark*

Received 3rd January 1994; revised manuscript received 19th April 1994

Abstract

A reversed-phase gradient liquid chromatographic (LC) method was developed for measuring medium components formed during a penicillin V fed-batch fermentation. The method is based on a high concentration of sodium sulphate in the mobile phase and requires no chemical derivations of the fermentation sample. It is used for the quantitative analysis of *p*-hydroxyphenoxyacetic acid, isopenicillin N, 6-aminopenicillanic acid (6-APA), *o*-hydroxyphenoxyacetic acid, phenoxyacetic acid, two isomers of penicilloic V acid, *p*-hydroxypenicillin V and penicillin V. The linear working range of the method is from 0.00125 to 0.125 g/l except for penicillin V which has a linear response from 0.005 to 0.500 g/l. For analyte concentrations below 0.010 g/l the response is influenced by the analyte concentration of the previously analyzed sample. The kinetics for the epimerisation of the two penicilloic V acid isomers was quantified. Decreasing values of the total peak area for the two isomers was observed during the epimerisation and to account for this a weight factor ψ was introduced. Different methods for preparing standards of penicilloic V acid are discussed. The performance of the LC method was demonstrated during monitoring of 250 h of a fed-batch fermentation carried out on a complex medium. A peak for an as yet unidentified fermentation product was discovered and monitored during the entire fed-batch fermentation. Finally, the importance of a proper sample storage is discussed.

Keywords: Liquid chromatography; Penicillin fermentation

1. Introduction

In a submerged penicillin V fermentation approximately 10–20% of the penicillin originally secreted by the mold is degraded to by-products. With phenoxyacetic acid as the sidechain precursor penicillin V is the major product, but several intermediates of the penicillin biosynthesis may leak into the medium. These intermediates include isopenicillin N [1], 6-aminopenicillanic acid (6-APA) [2], and cyclic α-aminoadipate in the form of 6-oxo-piperidine-2-carboxylic acid [3]. Certain strains oxidize the sidechain precursor forming *o*-hydroxyphenylacetic acid [1], *p*-hydroxyphenoxyacetic acid [4] and correspondingly *p*-hydroxypenicillin V. The penicillins secreted into the medium undergo hydrolysis to penicilloic

* Corresponding author.

SSDI 0003-2670(94)00193-P

acid, a reaction which may be followed by a slow decarboxylation to penilloic acid. Consequently, at the end of a penicillin fermentation the medium contains a number of structurally similar products.

In a fast screening program for the antibacterial activity of a large number of strains a biological assay is the most convenient analytical method [5]. When evaluating the economic benefits of introducing a new industrial production strain a much more detailed analysis of the product spectrum is required. Not only will some of the by-products have no antibacterial effect, but more important, demands from regulatory bodies for high purity of the final product makes it necessary to assert a strict control of both the nature and the quantity of by-products.

Several accurate spectrophotometric assays for the quantitative analysis of penicillins have already been reported [6–8]. The majority of these are easily automated using flow-injection analysis (FIA) [9,10]. High measurement frequencies (10–20 samples per hour) to be used for on-line monitoring and control of fed-batch penicillin fermentations is hereby possible [10]. The automated spectrophotometric methods are fast, but most of these methods are less selective than liquid chromatographic (LC) methods [2]. The comparatively low analysis frequency of an LC method (1–2 samples per hour) is justified by the high level of detailed information acquired in each analysis.

This paper describes a reversed-phase, gradient LC method with spectrophotometric detection for the simultaneous measurement of 9 reaction components of the penicillin fermentation: isopenicillin N, *p*-hydroxyphenoxyacetic acid, 6-aminopenicillanic acid (6-APA), phenoxyacetic acid, *o*-hydroxyphenoxyacetic acid, *p*-hydroxypenicillin V, penicillin V, penilloic V acid and two isomers of penicilloic V acid. The method requires no chemical pre- or postcolumn derivatization of the components. The robustness and accuracy of the method has been validated during more than 3 years of daily use and fine tuning in two collaborating laboratories. Chemical and enzymatic methods for the preparation of penicilloic V acid were compared, and the kinetics for the reversible epimerisation of the penicilloic V acid epimers has been quantified. Application of the method on samples retrieved from a fed-batch penicillin V fermentation with corn-steep-liquor has been included.

2. Experimental

2.1. Methods

Reversed-phase LC methods for analyzing sidechain precursors and penicillin V in fermentation media using mixtures of phosphate and acetonitrile–methanol as mobile phase has been used by several research groups [1,2,11]. However, in the process of setting up a reversed-phase gradient method we discovered that the use of phosphate buffer and acetonitrile was incompatible with the tested C_{18} separation columns, i.e., unacceptably low column lifetimes (< 150 injections) were observed even during analysis of purified standards. Based on experiences with a reversed-phase method for the analysis of insulin [12] a gradient method with high concentrations of sodium sulphate added to the mobile phase was developed. The performance of this LC method was optimized with respect to the following separation parameters: composition of the mobile phases, the gradient programme, pH of the buffer solution, and column temperature. It was found that the performance of the analysis does not depend on the brand of the C_{18} column as was the case when the method was used for the analysis of insulin [12]. Secondly, the developed protocol extended the C_{18} column lifetime to more than 1500 sample injections.

The performance of the developed reversed-phase gradient method was tested in terms of its sensitivity, reproducibility and linear range of detection with respect to the following components: 6-APA, *p*-hydroxyphenoxyacetic acid, phenoxyacetic acid, *o*-hydroxyphenoxyacetic acid, *p*-hydroxypenicillin V, penilloic V acid and penicillin V. However, quantification of penicilloic V acid was rendered difficult by the presence of two peaks in the chromatogram. Nachtmann [2] reported the presence of two peaks for penicilloic

G acid in a reversed-phase C_{18} method used for analyzing medium samples from penicillin G fermentations. The two peaks for penicilloic G acid were designated the α- and the β-form, a reference to their relative position in the chromatogram. Using LC, ^{13}C and ^{1}H NMR techniques it was later shown that the two forms of penicilloic G acid arise due to a reversible epimerisation of 5*R*,6*R*-penicilloic G acid into 5*S*,6*R*-penicilloic G acid via the imine tautomer of penamaldic acid [13]. The reaction scheme for this epimerisation process is shown in Fig. 1.

Monitoring the epimerisation process with the developed gradient method showed that the decrease in the peak area for the 5*R*,6*R*-isomer ($A_{5R,6R}$) was partly balanced by a comparatively smaller increase in peak area for the 5*S*,6*R*-isomer ($A_{5S,6R}$). The ratio (ψ) between the increase of $A_{5S,6R}$ and the simultaneous decrease of $A_{5R,6R}$ cannot be explained by differences in the specific absorbance of the two isomers since no change in absorbance with time was observed for a non-equilibrium mixture of penicilloic V acid. Recordings with an array diode UV detector showed that both penicilloic V acid peaks have maximum absorbance at 210 nm and a local maximum at 268 nm. Since the value of ψ only varied little for different brands of columns it was suspected that the purity of the penicilloic V acid standard might have a major impact on the value of ψ. Penicilloic V acid was therefore produced from a penicillin V solution either through enzymatic treatment with β-lactamase (Penase®) or by alkaline hydrolysis. In both cases only the 5*R*,6*R*-form is formed. This process was followed by an spontaneous epimerisation reaction producing an equilibrium mixture of the two isomers. Using a homogeneous solution of β-lactamase one has to consider the risk of enzyme adsorption to the column material. Hence, only immobilized enzymes were used in the protocol for the enzymatic production of penicilloic V acid standards.

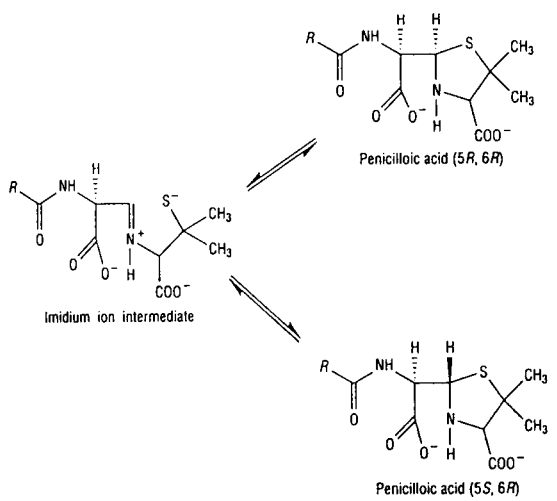


Fig. 1. Reaction scheme for the epimerisation of the 5*R*,6*R*-form of penicilloic acid via the imine tautomer of penamaldic acid (taken from [14]).

The robustness of the method was tested by measurements during 250 h of fed-batch fermentation. The samples from this fermentation contained high concentrations of complex media components including peptides and amino acids, but still no pretreatment was necessary. Three sets of identical samples were retrieved during the fermentation. One set was analyzed immediately after retrieval whereas the two remaining sets were stored frozen for three weeks and subsequently analyzed in two different laboratories.

2.2. Instrumentation

The LC equipment (from Waters Assoc.) consisted of two pumps (type 510), an autosampler/autoinjection system (type WISP 712), a column oven (type TCM), and a UV–visible spectrophotometer (type 481). The autoinjection system was equipped with a built-in cooling unit keeping the sample temperature at 4–5°C. Dataaquisition and pump control was performed by the 810 Baseline software from Waters Assoc.

2.3. Reagents

All reagents were prepared with Milli-Q water and microfiltrated using a filter with a poresize of 0.45 μm (Millipore Cat. No. HVLP4700). Eluents were degassed for 10–15 min by means of ultrasound.

The buffer eluent (eluent A) consisted of 0.2 M Na_2SO_4 (Merck No. 6649), 0.04 M H_3PO_4 (Merck No. 573) and between 1.3–5.0% (v/v)

Table 1
The composition of the mobile phase as a function of time. Linear refers to a linear increase or decrease of the concentration with respect to time

Time (min)	% (v/v) A	% (v/v) B	Gradient
0	95	5	Constant
5	95	5	Constant
20	60	40	Linear
25	60	40	Constant
30	95	5	Linear
45	95	5	Constant

acetonitrile (Merck No. 14291). The pH of the eluent was adjusted to 5.5 using 10 M NaOH or ethanolamine.

The hydrophobic eluent (hereafter called eluent B) was prepared by mixing 500 ml water with 500 ml acetonitrile (Merck No. 14291).

2.4. LC procedure

The mobile phase and the separation column were both heated to 40°C and the flow of mobile phase was adjusted to 1.00 ml/min. The chosen concentration versus time gradient is shown in Table 1. Two types of C_{18} columns were tested: (1) Hibar LiChrosorb C_{18}, 250 × 4 mm i.d. with 5-μm packing (Merck Cat. No. 50333) and (2) Nova-Pak C_{18}, 300 × 3.9 mm i.d. with 4-μm packing (Waters Assoc. Part. No. 11695). A Guard-pak μBondapack C_{18} precolumn (Waters Part. No. 88070) or a Nova-Pak C_{18} precolumn (Waters Part. No. 15220) was mounted between the injection point and the oven inlet. A sample volume of 10 μl was injected onto the column and detection was at 228 nm.

2.5. Preparation of standards

Standards of isopenicillin N, *p*-hydroxyphenoxyacetic acid, 6-aminopenicillanic acid, phenoxyacetic acid, *o*-hydroxyphenoxyacetic acid, penilloic V acid, potassium penicillin V, sodium penicillin G and *p*-hydroxypenicillin V were supplied by Novo Nordisk. The standard for *p*-hydroxypenicillin V was a low purity product containing between 88–93% (w/w) potassium penicillin V. The purity of each *p*-hydroxypenicillin V batch was determined by quantifying the extent of contamination by potassium penicillin V.

Two different methods were used for producing penicilloic V acid standards from potassium penicillin V.

(1) Using immobilized β-lactamase. 5 ml of 5% (v/v) glutardialdehyde in 0.1 M borate buffer (pH 8.5) was added to 200 mg of aminopropyl-CPG bead (Prod. No. 27791, Fluka) and allowed to react for 1 h with intermittent mixing. The beads were washed with five portions of 5 ml of phosphate buffer (0.1 M, pH 7.4). 5 mg of enzyme (Penase, Leo Pharmaceutical Products, Denmark) was dissolved in 1.5 ml phosphate buffer (0.1 M, pH 7.4) and this solution was added to the beads. The mixture was allowed to react for 4–5 h at room temperature. The beads were then washed with three portions of 5 ml phosphate buffer (0.1 M, pH 7.4) and packed into a column (15 × 2.0 mm i.d.). A solution containing 0.053 g/l potassium penicillin V dissolved in 45 mM phosphate buffer (pH 6.5) was pumped through the enzyme column at a flow rate of 0.5 ml/min.

(2) By alkaline hydrolysis. 52.9 mg of potassium penicillin V was dissolved in 40 ml of distilled water and 5.0 ml of 2.0 M NaOH was added. After 15 min 40 ml phosphate buffer (45 mM and pH 6.5) was added together with 5 ml of 2 M H_2SO_4. The pH was adjusted to 6.5 and the volume made up to 1.00 l with phosphate buffer (45 mM, pH 6.5).

2.6. Fed-batch fermentation

The fed-batch fermentation was carried out in a 41-l Chemap bioreactor (25–33 kg medium) using a high yielding strain of *Penicillium chrysogenum* supplied from Novo Nordisk. The temperature was fixed at 25°C, the headspace pressure at 0.5 bar overpressure, and the pH was controlled at 6.50 by addition of 2 M NaOH or 2 M H_2SO_4. Finally, the level of dissolved oxygen, pO_2, was maintained above 60% saturation by controlling the agitation rate. The gas flow rate was set at 27 l/min for the first 100 h and then increased to 32 l/min. Initially the medium contained 100 g/l corn-steep liquor (48.5% (w/w) dry matter), 5.7 g/l sucrose, 5.7 g/l phenoxy-

acetic acid, 12.0 g/l $(NH_4)_2SO_4$, 1.0 g/l KH_2PO_4, 0.06 g/l $CaCl_2 \cdot 2H_2O$. This medium was inoculated with 10% (w/w) of vegetative mycelium of age 50 h grown from 10^8 spores l^{-1} in a 7-l Chemap bioreactor. After a batch growth period of 9 h fed-batch operation was initiated with a feed containing 454.5 g/l glucose, 46.7 g/l $(NH_4)_2SO_4$ and 33.3 g/l phenoxyacetic acid. The

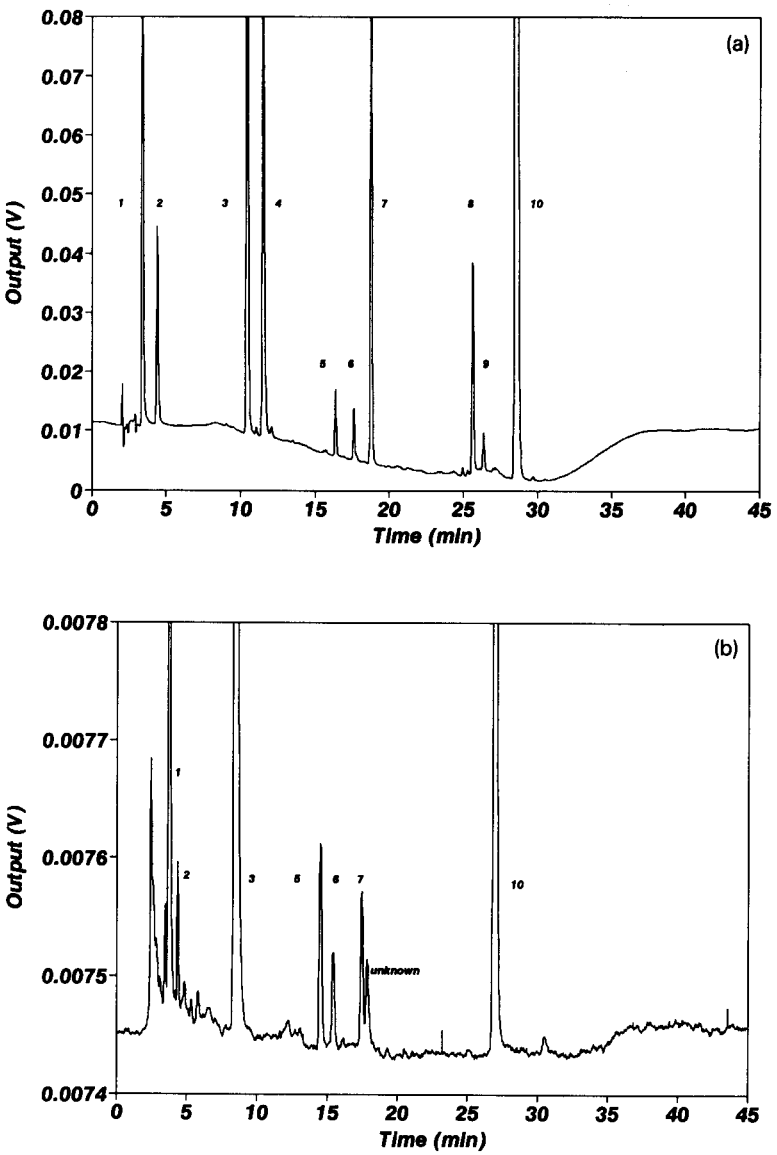


Fig. 2. (a) Chromatogram for a standard solution analyzed on a Nova-Pak C_{18} column (Waters Assoc) using eluent A containing 12.5 ml acetonitrile per l. (1) *p*-hydroxyphenoxyacetic acid, (2) 6-aminopenicillanic acid, (3) phenoxyacetic acid, (4) *o*-hydroxyphenoxyacetic acid, (5) 5*S*,6*R*-penicilloic V acid, (6) 5*R*,6*R*-penicilloic V acid, (7) *p*-hydroxypenicillin V, (8 + 9) penilloic V acid, (10) penicillin V. Due to a limited supply of isopenicillin N it was not possible to include this component in the standard. (b) Chromatogram for a fed-batch fermentation sample containing corn-steep liquor. The sample is retrieved after 249 h of fermentation and diluted with 45 mM phosphate buffer (pH 6.5) by a factor of 100. Eluent A contained 50 ml acetonitrile per l. Numbering of peaks as in Fig. 2a.

feed rate was as follows: 0.045 kg/h (9–50 h), 0.078 kg/h (50–150 h), 0.080 kg/h (150–235 h), and 0.090 (235–250 h). The density of the feed stock was 1.18 kg/l. After 89 h of fermentation 2.0 kg of a casein–peptone solution with a total nitrogen content of 1.84 moles was added in order to study the influence of the amino acid concentration on the penicillin V production.

2.7. Sample retrieval

Cell free medium samples were retrieved by means of an in situ membrane probe [15]. The first 20-ml sample was discharged in order to rinse the tubings and the void volume of the membrane (3.7 ml) from old sample. Subsequently, 10 ml sample was collected and stored at 3 5°C for less than 24 h.

3. Results

The discussion of the results is divided into three parts. First the reproducibility and accuracy of the developed gradient method is discussed. Next the epimerisation of penicilloic acid standards is discussed. Here a comparison of the purity of penicilloic acid standards produced by chemical and enzymatic methods is included. Finally, application of the LC procedure for off-line measurements of automatically retrieved samples from a fed-batch penicillin fermentation is described.

3.1. Performance of the method

Chromatograms for a standard containing all analytes (except for isopenicillin N) and for a 100 times diluted fermentation sample are shown in Fig. 2a and b, respectively. The performance of the C_{18} separation column in terms of the reproducibility of retention times and peak areas is to some extent a function of the column history. Small improvements in performance are often observed during the first week of usage of a new column or when using ethanolamine instead of NaOH for adjusting the pH of eluent A. Addition of ethanolamine or several injections of samples with high concentrations of analytes is believed to shield free silanol groups on the stationary phase from further interaction with analytes. The standard deviation of the method is therefore stated for new columns where the first 15 injections are discarded and the following 10 injections are used to calculate the relative standard deviation of the peak areas and retention times. The results hereof are listed in Table 2 for two different C_{18} columns. Using new columns with no addition of ethanolamine to eluent A the values in Table 2 are believed to represent a worst case situation.

Although the method has been applied mainly

Table 2
Retention times and relative standard deviation ($n = 10$) of the peak area (AU) for two C_{18} columns using eluent A with 12.5 ml acetonitrile per l. The concentration versus time gradient is stated in Table 1

Component	Hibar LiChrosorb C_{18} (Merck)		Nova-Pak C_{18} (Waters Assoc.)	
	Retention time (min)	R.S.D. (%) AU	Retention time (min)	R.S.D. (%) AU
Isopenicillin N	3.98 ± 0.07	3.4	3.04 ± 0.06	3.6
p-Hydroxyphenoxyacetic acid	4.52 ± 0.04	1.3	3.41 ± 0.03	1.0
6-Aminopenicillanic acid	5.70 ± 0.04	1.1	4.31 ± 0.05	1.8
Phenoxyacetic acid	11.62 ± 0.02	1.9	10.38 ± 0.05	1.8
o-Hydroxyphenoxyacetic acid	12.83 ± 0.03	1.4	11.40 ± 0.07	1.6
5*S*,6*R*-Penicilloic V acid	17.39 ± 0.02	2.9	16.77 ± 0.11	2.7
5*R*,6*R*-Penicilloic V acid	17.87 ± 0.02		18.38 ± 0.13	
p-Hydroxypenicillin V	20.51 ± 0.02	1.1	19.12 ± 0.15	1.3
Sodium penicillin G	26.47 ± 0.01	0.6	–	–
Penilloic V acid	–	–	26.19 ± 0.19	1.3
Potassium penicillin V	29.77 ± 0.03	1.8	29.38 ± 0.32	0.7

Table 3
Calculated values for the slope, intercept and its standard deviation for six components obtained by single, consecutive injection of samples containing increasing concentrations of analyte. The experiment was carried out with a Nova-Pak C_{18} column (Waters Assoc.) using eluent A containing 12.5 ml acetonitrile per l. Due to a limited supply of isopenicillin N it was not possible to quantify the linear range for this component

Component	Slope, AU/(g/l)	Std. error, slope (% rel.)	Intercept (AU)	Std. error, intercept (AU)
p-Hydroxyphenoxyacetic acid	22931837	0.38	−2758	9532
6-Aminopenicillanic acid	5136562	0.42	−2538	2340
Phenoxyacetic acid	19024969	0.38	−159	7823
Penicilloic V acids	7672974	0.53	297	3434
Penilloic V acid	5473785	0.58	1343	3476
Potassium penicillin V	9209043	0.52	−1048	21001

to penicillin V fermentations a standard of penicillin G was included in order to illustrate the flexibility of the method.

A precise quantification of isopenicillin N, *p*-hydroxyphenoxyacetic acid, and 6-APA is often hampered by their low resolution from the sample front containing low-molecular-weight components. Lowering the concentrations of acetonitrile (1.25%, v/v) in eluent A results in good separation of the sample front and isopenicillin N, 6-APA and *p*-hydroxyphenoxyacetic acid. This resolution is obtained at the expense of an increased risk of microbial growth in eluent A. With eluent A containing four times as much acetonitrile (5.0%, v/v) there is poor resolution of isopenicillin N, *p*-hydroxyphenoxyacetic acid and 6-APA from the sample front, but microbial growth in eluent A is then not a problem.

The penilloic V acid standards produced two peaks in the chromatogram (Nos. 8 and 9). This separation of the penilloic V acid into two peaks cannot be explained at the present and for the time being the peak area for penilloic acid is quantified as the total area of peaks 8 and 9.

The linear range of the LC method was determined by single, consecutive injections of samples containing increasing concentrations of analyte (Table 3). The analyte concentrations ranged from 0.00125 to 0.500 g/l (2.00 g/l in the case of penicillin V). The experiment was then repeated but this time using decreasing concentrations of the standards (Table 4). Linear regression analysis on these two sets of data showed good linearity in the range of 0.00125–0.125 g/l for *p*-hydroxyphenoxyacetic acid, 6-aminopenicillanic acid, penicilloic V acid, penilloic V acid and phenoxyacetic acid. Measurements of penicillin V produced a linear standard curve for concentrations between 0.005 and 0.5 g/l. The results of the linear regression analysis are listed in Tables 3 and 4. For concentrations below 0.010 g/l the measurements of penicilloic V acid, phenoxyacetic acid and 6-aminopenicillanic acid may exhibit hysteresis – the analytical result depends on

Table 4
Same as Table 3 but this time the standard curve was obtained from single injection of standards with decreasing concentrations of analyte. For experimental details, see text to Table 3

Component	slope, AU/(g/l)	Std. error, slope (% rel.)	Intercept (AU)	Std. error, intercept (AU)
p-Hydroxyphenoxyacetic acid	23390414	0.44	4287	9532
6-Aminopenicillanic acid	5170955	0.4	−499	2238
Phenoxyacetic acid	19194504	0.41	2400	8774
Penicilloic V acid	7919152	0.71	1032	4755
Penilloic V acid	5479208	0.49	2883	2958
Potassium penicillin V	9225145	0.31	−5851	12504

the analyte concentration in the previously examined sample.

3.2. Epimerisation of penicilloic acid

Penicilloic V acid standards were prepared from a penicillin V solution which was pumped through a column containing immobilized β-lactamase and thereby completely converted into one penicilloic acid isomer – presumably the 5*R*,6*R*-isomer. This isomer was gradually converted into a mixture of isomers by a reversible epimerisation process. Fig. 3 shows the peak area of 5*S*,6*R*-penicilloic acid plotted against the peak area of 5*R*,6*R*-penicilloic V acid during the epimerisation of the enzymatically produced penicilloic V acid dissolved in a 45 mM phosphate buffer (pH 6.5) and kept at 25°C. The curve in Fig. 3 can be approximated by a straight line with a slope of −0.90. This slope implies that the specific peak area (AU/g/l) for the 5*R*,6*R*-isomer, apparently, should be 1.11 times higher than that of the 5*S*,6*R*-isomer. The fact that the value remains constant during the entire epimerisation makes it difficult to believe that its deviation from unity is to be explained by insufficient stability of the penicilloic V acid when kept in a buffer solution of pH 6.5. Penicilloic V acid dissolved in acidic solutions can decarboxylate to penilloic V acid, but no peak for penilloic acid was observed during the reported epimerisation experiments. Since it is unlikely that samples always contain an equilibrium mixture of penicilloic V acids one needs to include a weight factor (ψ) when calculating the total peak area for the penicilloic V acid peaks in accordance with Eq. 1.

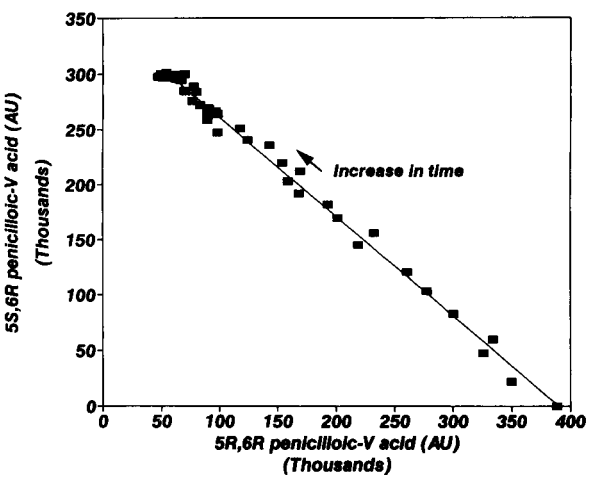


Fig. 3. The peak area of 5*S*,6*R*-penicilloic V acid as function of the peak area for 5*R*,6*R*-penicilloic V acid. The arrow indicates the increase in time.

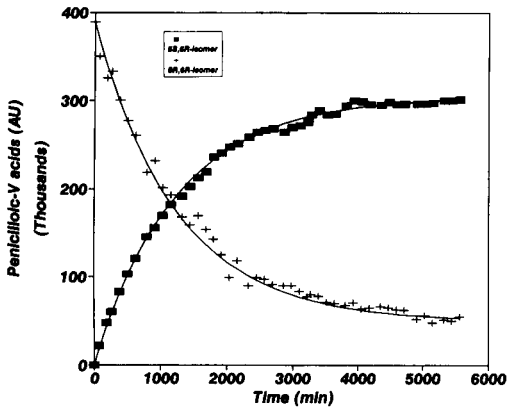


Fig. 4. The peak area versus time profiles for the epimerisation of penicilloic V acid in 45 mM phosphate buffer (pH 6.5) at 25°C. (■) 5*S*,6*R*-isomer, (+) 5*R*,6*R*-isomer. The analysis was carried out on a Nova-Pak C_{18} column using eluent A containing 50 ml acetonitrile per l. The penicilloic V acid was produced by using immobilized β-lactamase resulting in a ψ value of 1.11. Curves are model simulations.

$$A_{penicilloicV\ acid} = A_{5R,6R-form} + \psi \cdot A_{5S,6R-form} \quad (1)$$

Measurements of the ψ value for different C_{18} columns showed only minor variations (Nova-Pak C_{18}, 1.11 and LiCrosorb C_{18}, 1.21). ψ Values calculated from epimerisation experiments performed with penicilloic V acid prepared enzymatically are, however, significantly different from ψ values calculated from experiments based on alkaline hydrolysis of penicillin V. High ψ values (1.43–1.45) are, generally, observed for penicilloic V acid produced by alkaline hydrolysis. Iodiometric titration of penicilloic acid produced by alkaline hydrolysis revealed that the purity of this standard is approximately 80% (w/w). Ghebre-Sellassie [13] found quantitative conversion between the two isomers of penicilloic G acid when using ^{1}H NMR spectroscopy for monitoring of

the epimerisation. In addition, they found similar concentration versus time profiles when comparing the ^{1}H NMR results with measurements obtained by an LC method. This shows that the epimerisation is quantitative and no degradation of penicilloic V acid takes places during the process. The most likely explanation for ψ factors above 1.0 seems to be column recoveries below 100% for the 5*S*,6*R*-isomer.

The kinetics for the epimerisation in a 45 mM phosphate buffer (25°C) was quantified by monitoring the peak area of both epimers for more than 5000 min (Fig. 4). First, assume that the total concentration (c_0) of penicilloic V acids remains constant during the experiment, i.e., the concentration of the imine tautomer is neglected. Secondly, it has been reported that the epimerisation can be described by first order reversible kinetics [13]. If the concentration of the 5*R*,6*R*-isomer at time zero equals c_0 then its concentration versus time profile is given by Eq. 2a. Furthermore, the ratio between the two first order rate constants, k_1 and k_2, can be calculated from equilibrium data as shown in Eq. 3. Since the peak area is a linear function of the analyte concentration one can reformulate Eqs. 2a and 2b in terms of normalized peak areas as shown in Eqs. 4a and 4b. The constant column recovery of $1/\psi$ for the 5*S*,6*R*-isomer is accounted for in Eq. 4b.

$$c_{5R,6R} = \frac{c_0}{k_1 + k_2}\left(k_2 + k_1 e^{-(k_1+k_2)\cdot t}\right) \quad (2a)$$

$$c_{5S,6R} = c_0 - c_{5R,6R} \quad (2b)$$

$$K_{eq} = \frac{k_1}{k_2} = \frac{c^{eq}_{5S,6R}}{c^{eq}_{5R,6R}} = \frac{c_0 - c^{eq}_{5R,6R}}{c^{eq}_{5R,6R}} \quad (3)$$

$$\frac{A_{5R,6R}}{A^o_{5R,6R}} = \frac{1}{k_1 + k_2}\left(k_2 + k_1 e^{-(k_1+k_2)\cdot t}\right) \quad (4a)$$

$$\frac{A_{5S,6R}}{A^0_{5R,6R}} = \frac{1}{\psi}\cdot\left(1 - \frac{A_{5R,6R}}{A^0_{5R,6R}}\right) \quad (4b)$$

Fitting the values of k_1 and k_2 to the experimental data of Fig. 4 one arrives at values of $k_1 = 0.042$ h^{-1} and $k_2 = 0.006$ h^{-1}. This implies that an equilibrium mixture of penicilloic V acids (at 25°C, pH 6.5) contains seven times more of the 5*S*,6*R*-isomer as compared to the 5*R*,6*R*-isomer of penicilloic V acid. Ghebre-Sellasie et al., [13] reported an equilibrium ratio of four for the epimerisation of penicilloic G acids in slightly alkaline media (pH 8.7) whereas experiments by Fong et al. [16] found an equilibrium ratio of seven for a mixture of the amoxicillin penicilloic acids (50 mM phosphate buffer, pH 5.9, room temperature).

The slow epimerisation rate at room temperature means that several days will pass before an equilibrium mixture of penicilloic V acids is obtained. Heating the freshly prepared solution of penicilloic acid to 40°C increases both rate constants by a factor of approximately 12 ($k_1 = 0.50$ h^{-1} and $k_2 = 0.07$ h^{-1}). Concentrations within 5% of the equilibrium concentrations can then in principle be reached within a period of approximately 5 h. However, as noted by Möller et al. [11] heating a penicilloic V acid solution to 40°C can cause partial decarboxylation of the penicilloic V acid forming small amounts of penilloic V acid. Using the developed LC method it was found that when a non-equilibrium solution of penicilloic V acid is kept at 40°C for 3.5 h 9% (w/w) of the penicilloic V acid is converted into penilloic V acid. Heating of the enzymatically prepared penicilloic V acid is obviously a convenient way of rapidly obtaining substantial amounts of both penicilloic V acid epimers. However, the method does not seem suited for quantitative analysis unless the extent of decarboxylation to penilloic acid is carefully monitored.

3.3. *Measurements during a fed-batch fermentations*

The LC method was applied for measurement of *p*-hydroxyphenoxyacetic acid, phenoxyacetic acid, *p*-hydroxypenicillin V and penicillin V during a fed-batch penicillin V fermentation using corn-steep liquor as a source of amino acids, lactate and phosphate (Figs. 5 and 6). From measurements during similar fermentations with the same strain it is known that the level of *p*-hydroxyphenoxyacetic acid, 6-aminopenicillanic acid and isopenicillin N are in the range of 0.5–1.0 g/l at the end of a fed-batch fermentation. The se-

cretion of *o*-hydroxyphenoxyacetic acid has not been observed.

As seen in Figs. 5 and 6 the volumetric productivity of penicillin V remains high for the first 75 h. The discontinuity after 89 h is caused by a pulse addition of 2.0 kg of casein–peptone solution serving as an extra supply of amino acids. The volumetric productivity of penicillin for the following 100 h is slightly lower than in the first 75 h. After 175 h of fermentation the penicillin V concentration starts to level off since the microbial penicillin V production decreases whereas the rate of penicilloic acid formation by hydrolysis of the secreted penicillin is high. A similar trend is observed for *p*-hydroxypenicillin V. The phenoxyacetic acid concentration remains between 5.4 and 6.2 g/l for the first 75 h but declines rapidly as the rate of microbial assimilation increases above the feeding rate. From mass balances it is found that the stoichiometric ratio between the amount of synthesized penicillin V (including penicilloic V acid) and the utilized amount of phenoxyacetic acid varies between 0.7 and 0.9 mole/mole. These numbers are significantly below the theoretical value of 1.0 mole/mole indicating formation of as yet unidentified products or to a metabolic degradation of phenoxyacetic acid. Penilloic V acid was not present

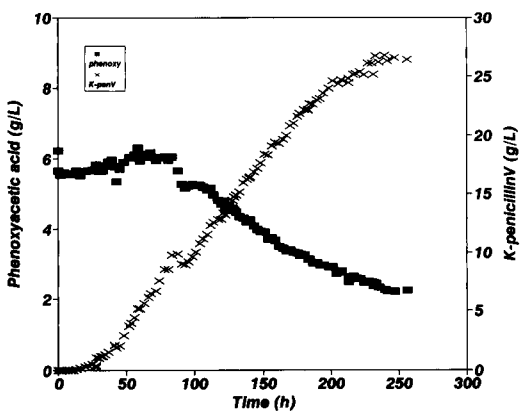


Fig. 5. Time profiles for the concentration of phenoxyacetic acid (■) and potassium penicillin V (×) for a fed-batch penicillin V fermentation with continuous feeding of glucose, $(NH_4)_2SO_4$ and phenoxyacetic acid. The measurements were performed with a LiChrosorb C_{18} column and eluent A contained 50 ml acetonitrile per l.

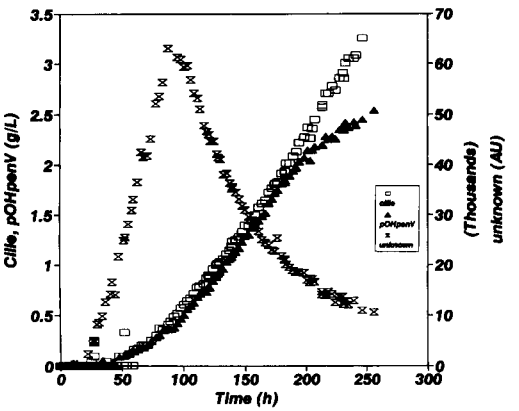


Fig. 6. Concentration versus time profiles for the concentration of *p*-hydroxypenicillin V (▲), the total concentration of penicilloic acids (□) and the peak area of an unknown compound (⋈). The measurements were performed with a LiChrosorb C_{18} column and eluent A containing 50 ml acetonitrile per l. The total peak area for penicilloic acid was obtained by using a weight factor (ψ) of 1.42.

in any of the fermentation samples and the concentration of *p*-hydroxyphenoxyacetic acid never exceeded 0.5 g/l (results not shown).

Besides the quantified compounds an extra peak appeared in the chromatogram with a retention time slightly larger than that of *p*-hydroxypenicillin V. After 90 h this peak had a maximum area which thereafter slowly decreased during the rest of the fermentation. The observed decrease in the concentration of this unknown compound cannot be accounted for solely by the dilution of the medium due to continuous feeding of substrate. Hence the compound is continuously degraded by microbial activity or by chemical reactions. Attempts to isolate several milligrams of this unknown compound by preparative LC methods for further analysis have so far been unsuccessful.

The concentration of penicillin V was measured immediately after retrieval of the samples and these results were compared with the concentrations found in samples stored at −18°C for 3 weeks (see Fig. 7). In samples containing more than 10 g/l potassium penicillin V there is a significant loss of penicillin V to penicilloic acid. For fermentation samples with pH 6.5 the crystallisation of water may change the pH of the

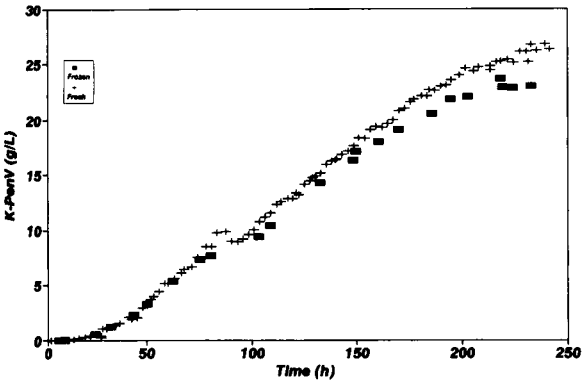


Fig. 7. The measured concentration of pencillin V immediately after retrieval of the samples (+) and after storage in the freezer for 3 weeks (■).

supernatant thereby leading to an increased rate of penicillin V hydrolysis. It is therefore recommended to dilute all fermentation samples with distilled water before freezing. The exact degree of dilution may depend on the overall composition of the medium. In this study dilution to penicillin V concentrations below 10 g/l was adequate.

Two identical sets of frozen samples were prepared and subsequently analyzed in two laboratories for four major components in the medium from the penicillin V fermentation: phenoxyacetic acid, penicilloic V acids, *p*-hydroxypenicillin V and penicillin V. The penicillin V concentrations are reported as pooled values expressing the total concentration of penicillin V and penicilloic V acids stated in units of penicillin V equivalents. The result of a linear regression analysis on the two sets of data are shown in Table 5 indicating a very high correlation between the results obtained in two laboratories. This clearly demonstrates the robustness of the reported LC method.

4. Conclusions

A reversed-phase gradient method based on high concentrations of sodium sulphate in the mobile phase has been developed. One advantage of the developed protocol is an increased lifetime of the C_{18} separation columns. Secondly, high resolution is achieved for all major products, by- and degradation products as well as the sidechain precursors present in the spent medium from a penicillin V fermentation. The high selectivity and the long column lifetime is obtained without any pretreatment of the fermentation samples except for a dilution step. Hence, the method is particularly well suited for routine monitoring of penicillin V fermentation processes and the subsequent down-stream processing steps.

Monitoring of the epimerisation process of penicilloic V acid showed that the total peak area for the two isomers of penicilloic V acid decreases during the epimerisation even though no decarboxylation to penilloic V acid was observed. Column–analyte interactions seems the only reasonable way of accounting for recoveries below 100% for the 5*S*,6*R*-penicilloic V acid epimer. In order to compensate for the latter a weight factor (ψ) was introduced for the calculation of the total penicilloic V acid peak area. The validity of introducing a ψ factor in the protocol is justified by the previously reported excellent agreement between the results obtained by the LC method and measurements performed by an automated, iodometric method [8]. At present it seems necessary to determine the exact value of ψ for each partic-

Table 5
Results of correlation analysis between two sets of identical, frozen samples analyzed in two different laboratories. Each set consisted of 25 samples retrieved at regular time intervals during a fed-batch penicillin V fermentation

Component	Slope	Std. error slope	Intercept	Std. error intercept
Phenoxyacetic acid	0.99	0.02	0.064	0.12
p-Hydroxypenicillin V	1	0.01	−0.01	0.04
Penicillin V + Penicilloic acid	1.03	0.02	−0.03	0.69

ular set of separation column and type of protocol used for the production of penicilloic V acid standards. The reported test indicates that the use of immobilized β-lactamase with no heating of the hydrolyzed solution gives the highest purity of penicilloic V acid as well as a ψ factor close to 1.

Freezing undiluted fermentation samples may lead to partial degradation of penicillin V to penicilloic V acid, in particular in samples with penicillin V concentrations in excess of 10 g/l. It is therefore recommended to analyse the samples immediately after their retrieval. If this is not possible one should dilute the samples with distilled water before storing them in the freezer.

References

[1] M.W. Adlard, B.M. Gordon, T. Keshavarz, C. Bucke, M.D. Lilly, A.T. Bull and G. Holt, Biotechnol. Tech., 5 (1991) 121–126.
[2] F. Nachtmann, Chromatographia, 12 (1979) 380–385.
[3] S.P. Brundidge, F.C.A. Gaeta, D.J. Hook, C. Sapino, R.P. Elander and R.B. Morin, J. Antibiotics, 33 (1980) 1348–1351.
[4] R.L. Hussey, W.G. Mascher and A.L. Lagu, Chromatographia, 15 (1983) 999–1003.
[5] C. Ball and M.P. McGonagle, J. Appl. Bacteriol., 45 (1978) 67–74.
[6] S.C. Pan, Anal. Chem., 26 (1954) 1438–1444.
[7] K.A. Holm, Anal. Chem., 44 (1972) 795–799.
[8] M. Carlsen, L.H. Christensen and J. Nielsen, Anal. Chim. Acta, 274 (1993) 117–123.
[9] I. Schneider, Anal. Chim. Acta, 166 (1984) 293–295.
[10] M. Carlsen, C. Johansen, R.W. Min, J. Nielsen, H. Meier and F. Lantreibecq, Anal. Chim. Acta, 279 (1993) 51–58.
[11] J. Möller, R. Hiddessen, J. Niehoff and K. Schügerl, Anal. Chim. Acta, 190 (1986) 195–203.
[12] G. Mandrup, Trends Anal. Chem., 11 (1992) 389–395.
[13] I. Ghebre-Sellassie, S.L. Hem and A.M. Knevel, J. Pharm. Sci., 73 (1984) 125–128.
[14] A. Shah, C.D. Orford, D. Perry and M.W. Adlard, LC–GC Int., 6 (1993) 224–234.
[15] L.H. Christensen, J. Nielsen and J. Villadsen, Chem. Eng. Sci., 46 (1991) 3304–3307.
[16] G.W.K. Fong, R.N. Johnson and B.T. Kho, J. Chromatogr., 255 (1983) 199–207.

ELSEVIER

Analytica Chimica Acta 296 (1994) 63–68

ANALYTICA CHIMICA ACTA

Determination of uranium in apatite minerals by inductively coupled plasma atomic emission spectrometry after solvent extraction and separation with 3-phenyl-4-benzoyl-5-isoxazolone into diisobutyl ketone

Osamu Fujino [a], Shigeo Umetani [b,*], Masakazu Matsui [b]

[a] *Research Institute for Science and Technology, Kinki University, Kowakae, Higashiosaka 577, Japan*
[b] *Institute for Chemical Research, Kyoto University, Uji, Kyoto 611, Japan*

Received 4th January 1994; revised manuscript received 20th April 1994

Abstract

Solvent extraction with inductively coupled plasma atomic emission spectrometry (ICP-AES) was applied to the determination of uranium in apatite minerals. The combination of 3-phenyl-4-benzoyl-5-isoxazolone (HPBI), a highly acidic extracting reagent ($pK_a = 1.23$), and diisobutyl ketone (DIBK) make it possible to extract uranium from a very low pH region where uranium is separated from the interfering elements and the precipitation of calcium phosphate is avoided. Apatite minerals were treated with concentrated nitric acid. After removing a small quantity of insoluble residue, uranium was extracted with 0.06 M HPBI into DIBK at pH 0.3 in order to separate it from the interfering elements. Uranium was determined by ICP-AES and direct injection of the DIBK phase. A wavelength of 367.01 nm was selected which gave the highest intensity ratio, i.e., the emission intensity of uranium in DIBK to the background one, and the best detection limit (0.02 mg/l). The uranium content in the apatite was found to be $(1.06 \pm 0.12) \times 10^{-2}\%$.

Keywords: Inductively coupled plasma spectrometry; Atomic emission spectrometry; Apatite minerals; Uranium

1. Introduction

In previous publications we have described a highly sensitive and precise quantitative analysis method for the determination of rare earth elements [1–3] in rare earth ores such as xenotime, monazite [4,5] and apatite [6]. Xenotime mainly contains yttrium and heavy lanthanides, monazite light lanthanides and apatite calcium as a macrocomponent. Analyses were performed using graphite furnace atomic absorption spectrometry or inductively coupled plasma atomic emission spectrometry (ICP-AES). The distribution of rare earth elements was also studied. The ores also contain uranium and thorium which are used as

* Corresponding author.

SSDI 0003-2670(94)00195-R

fuel for atomic energy. Especially monazite is known to be a resource of thorium.

Uranium is a radioactive element which is present in nature. Several μg/l of uranium have been found in natural water, and mg/kg amounts in rocks. Uranium is also widely distributed, but at low concentrations, in other natural samples. Since comparatively large quantities of uranium are found in apatite minerals [6] containing calcium phosphate as its macro-component and approximately 1% of rare earth elements, the distribution of uranium is of great interest. In order to establish a highly sensitive and precise quantitative method for the determination of uranium, ICP-AES, used in many fields as the method of choice, and solvent extraction separation using 3-phenyl-4-benzoyl-5-isoxazolone (HPBI) [7–9] and diisobutyl ketone as the organic phase have been examined in detail. As a result, ICP-AES proved to be a good quantitative analysis method for determination of uranium in apatite minerals.

2. Experimental

2.1. Apparatus and reagents

The instrumentation employed was a Jarrell-Ash Model 575 inductively coupled plasma emission spectrometer with a 0.75-m focal length Czerny-Turner type, holographic grating, 1800 grooves mm^{-1}, and a cross-flow type nebulizer. Slit-widths are: entrance 10 μm and exit 10 μm. The RF-power is supplied by a 27 MHz generator.

A standard solution of uranium (1 mg/l) was prepared by dissolving 0.89079 g of uranyl acetate $[UO_2(CH_3CO_2)_2 \cdot 2H_2O]$ in 10 ml concentrated nitric acid and diluted to 500 ml with pure water. Rare earth element solutions were prepared by dissolving their high-purity oxides (99.99–99.9999%) or sulfate for cerium in hot concentrated nitric acid. HPBI was synthesized according to the literature [8,9]. All other solvents and reagents were of reagent-grade or Suprapure materials and were used without further purification.

Table 1
Instrumental operating conditions

Plasma	Water	Organic solvent
RF power (kW)	1.3	2.0
Carrier gas (kg/cm^2)	1.5	1.3
Plasma gas (l/min)	0.8	1.6
Coolant gas (l/min)	12.8	20.0
Observation height (mm)	14.5	18.0

2.2. Analytical procedure

Apatite minerals were dissolved in hot concentrated nitric acid: 1–10 g of powdered sample in a covered 100 ml beaker was heated with 10 ml of concentrated nitric acid at 150–200°C for 30–60 min. After cooling, the sample solution was diluted with pure water, passed through a filter paper (No. 5c) to remove any residue and then diluted to 100 ml (10–100 mg sample ml^{-1}) with pure water.

The apatite mineral sample (apatite 0.1–0.2 g) solution was placed in a separatory funnel and the pH was adjusted to 0.3. 5 ml of 0.06 M HPBI in DIBK was added to the funnel, and the mixture was shaken for 10–20 min. The organic phase was transferred to a centrifuge tube and separated. The organic phase obtained was injected directly to the plasma of an ICP-AES spectrometer under the conditions shown in Table 1.

3. Results and discussion

3.1. Optimum measuring conditions and calibration curves

The influence of RF power, the flow rate of carrier, coolant and plasma gas on the emission intensity of uranium in an aqueous or a DIBK solution, and its distribution in a plasma flame have been examined. The optimum measuring conditions based on the results are shown in Table 1. Under these measuring conditions, the detection limit of uranium in DIBK was found to be 0.02 mg/l and the reproducibility was 0.1–0.2%. In the case of DIBK, the output of RF

power was set at 2.0 kW to prevent the plasma flame from extinguishing and adhesion of carbon to the plasma torch. Sensitivity and reproducibility were almost identical as for the aqueous solution. The calibration graph was linear in the concentration range of 0.5–50 mg/l of uranium. The relative standard deviation was 1% for ten replicate measurements of 0.5 mg/l uranium.

3.2. *Interference of coexisting elements and selection of wavelength*

For precisely measuring the emission intensity of uranium by ICP-AES, it is necessary to find an appropriate wavelength with high sensitivity and low interference by coexisting elements. Here, some wavelengths were selected meeting the above requirements based on the previously announced data [10].

The emission intensity values of the elements were supposed to be present in apatite minerals at a level higher than 1×10^{-3}% [6]. Calcium, phosphorous, silicon, aluminum, iron, strontium, barium, manganese, magnesium, sodium, potassium, thorium and rare earth elements were measured. The elements which are likely to interfere and their degrees of interference, except for calcium and phosphorous, are shown in Table 2. It shows that most rare earth elements and thorium would interfere with uranium when present in a ten-fold or larger excess. However, the contents of each of these elements in the apatite minerals determined by ICP-MS were found to be less than the value for uranium [11]. Consequently, we do not need to pay attention to the interference caused by these coexisting elements.

However, apatite minerals contain calcium phosphate as a macro-component being present in a several thousand times excess with respect to uranium. Here, the effect of coexisting calcium was examined in detail with respect to wavelengths 263.55, 367.01, 385.96, 393.20 and 409.01 nm, which are reported to be suitable for uranium [10]. Among these wavelengths, the results for 367.01 nm are shown in Fig. 1 as a typical example. The emission at 367.01 nm increased when adding synthetic hydroxyapatite although it does not contain uranium. The emission intensity of 10 mg/l of uranium decreased when synthetic hydroxyapatite was added and then started to increase over 10 mg/ml of synthetic hydroxyapatite. Spectral interference by calcium in the

Table 2
Spectral interference of coexisting elements in the determination of uranium in apatite minerals

Wavelength (nm)	Spectral interference of coexisting elements (%) [a]						
	× 10 [b]		× 100 [b]				
	5–50%	51–100%	5–10%	11–30%	31–60%	61–100%	> 100%
263.55	Er,Lu		Ce,Pr Eu,Ho Yb	Th,Sm Dy,Er Tm	Tb	Lu	Eu
367.01		Th	Y,Eu	Sc,Ce Pr,Sm Gd,Tb Dy,Ho Er,Yb	Nd,Tm		
385.96	Sc,Nd Dy		Th,La Ce,Gd Tm	Sm,Tb Ho,Er	Pr,Eu	Nd,Dy	Sc
393.20	Th,Dy		La	Sm,Tb Ho	Ce,Pr Eu,Er	Nd	Dy
409.01	Th		La,Dy Ho,Tm	Ce,Sm Gd,Tb	Pr,Nd Er		

[a] Degree of interference with respect to the emission intensity of uranium.
[b] Amount of interfering substance with respect to uranium.

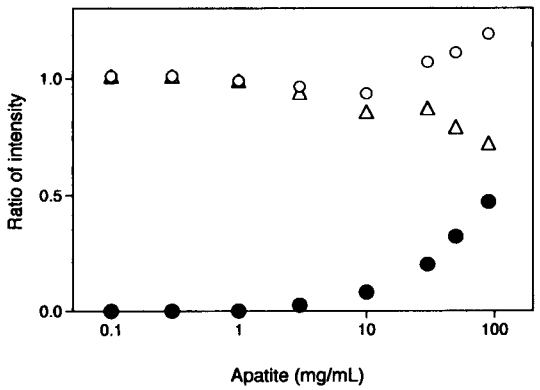


Fig. 1. Effect of synthetic hydroxyapatite on the emission intensity of uranium at 367.01 nm: (●) apatite alone; (○) uranium (10 mg/l) and apatite; (△) first curve subtracted from second curve.

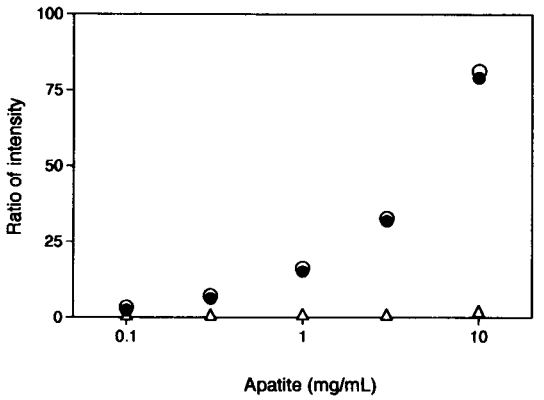


Fig. 2. Effect of synthetic hydroxyapatite on the emission intensity of uranium at 393.20 nm. Key to symbols as in Fig. 1.

macro-component (hydroxyapatite) enhances the emission intensity of uranium. Negative interference by the macro-component itself reduces the emission intensity of uranium. These interferences seem to exist simultaneously. Since this negative interference appeared frequently at the wavelengths examined, it is assumed that some unexpected changes in temperature or profile of a plasma flame and/or in the distribution of uranium in a plasma flame might be caused by a coexisting component of a high concentration. Spectral interference by calcium was also observed at other wavelengths examined and showed a similar tendency as seen in Fig. 1. However, the magnitude of the spectral interference for the various wavelengths was quite different, especially for 393.20 nm as can be seen in Fig. 2, where the spectral interference with the emission intensity of uranium (10 mg/l) was extremely higher than that at 367.01 nm. From the above results, it is clear that it is necessary to separate uranium from a macro-component prior to the measurement in order to precisely quantify uranium by ICP-AES. For this purpose, the solvent extraction method was adopted and the various examinations described below were carried out.

3.3. Selection of the extraction reagent

The sample solutions of apatite minerals used in the present work were treated with hot concentrated nitric acid. When the pH of the solution was higher than 1.5, a precipitate of phosphate was deposited and the extraction of uranium became difficult. Therefore, it is necessary to use an extraction reagent with which uranium can be quantitatively extracted from a 0.1 M or more acidic solution, developing no precipitation and allowing the separation from interfering elements. In the previous work [12], 1-phenyl-3-methyl-4-trifluoroacetyl-5-pyrazolone (HPMTFP) was employed as an extraction reagent to separate thorium from the interfering elements in phosphate minerals by ICP-AES. Although HPMTFP whose pK_a value is 2.56 [13] was found to be powerful enough for this purpose, an even more powerful extractant is required for the extraction and separation of uranium. For this purpose, 3-phenyl-4-benzoyl-5-isoxazolone (Fig. 3) has been chosen as an extraction reagent. HPBI, structurally similar to β-diketones, is a highly acidic extracting reagent readily derived from 3-phenyl-5-isoxazolone and benzoic anhydride [8,9]. The acid dissociation constant of HPBI was determined by a liquid–liquid distribution method

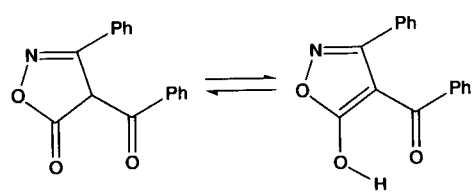


Fig. 3. Structure of 3-phenyl-4-benzoyl-5-isoxazolone (HPBI).

to be 1.23 [9]. Many metal ions are extractable from a lower pH region owing to its strong acidity. The pK_a value for 2-thenoyltrifluoroacetone, one of the most representative β-diketones, is 6.23 [14] and for 1-phenyl-3-methyl-4-benzoyl-5-pyrazolone, the well-known acidic and powerful extractant, it is 3.92 [13]. It is clear that HPBI would be most suitable for the extraction of uranium from a very low pH region required for the present work.

3.4. Selection of organic solvents for the plasma flame

It is known that some organic solvents extinguish the plasma flame or cause an adhesion of carbon deposit to the plasma torch. As a result of a detailed examination [3], the organic solvents which have 8–9 or more carbon atoms proved to be suitable for direct injection to ICP-AES. However, when the number of carbons is excessive, the viscosity becomes higher and the suction speed of the sample into an introduction instrument decreases. Therefore, we selected diisobutyl ketone (DIBK, $C_9H_{18}O$), dibutyl ether ($C_8H_{18}O$), *n*-decane ($C_{10}H_{22}$) and cyclooctane (C_8H_{16}), having 8–10 carbon atoms, and their combustion behaviour was examined. The results showed that none of the solvents caused either extinguishing of the plasma flame or carbon adhesion to the plasma torch.

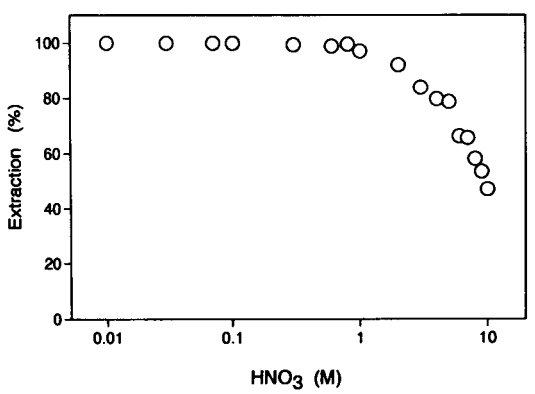


Fig. 4. Extraction of uranium with HPBI into DIBK from nitric acid solution. HPBI, 0.06 M; uranium, 10 mg/l; $V_a / V_o = 2$.

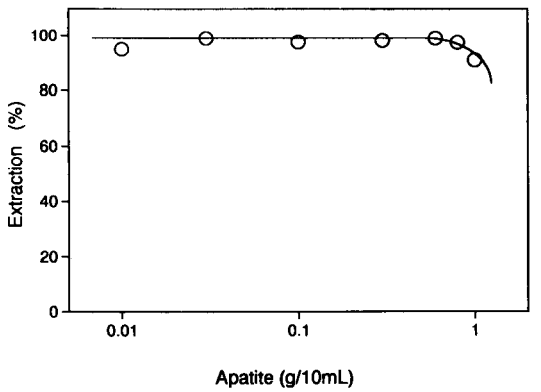


Fig. 5. Extraction of uranium with HPBI into DIBK from 0.5 M nitric acid solution containing synthetic apatite. HPBI, 0.06 M; uranium, 10 mg/l; $V_a / V_o = 2$.

The solubility of HPBI in DIBK, dibutyl ether, *n*-decane and cyclooctane was examined and found to be 0.06, 0.01, 0.008 and 0.003 M, respectively. HPBI was fairly soluble in DIBK and remaining DIBK in the nebulizer was readily removed by water after injection. Therefore, we decided to employ DIBK as an organic solvent.

3.5. Extraction of uranium

The influence of the nitric acid concentration on the extraction of uranium (10 mg/l) into DIBK containing 0.06 M HPBI was examined. The results are illustrated in Fig. 4, where the volume ratio (V_a/V_o) of the aqueous (V_a) and the organic (V_o) phase was set at two. As seen in Fig. 4, uranium was quantitatively extracted at a pH

Table 3
I_{BG}/I_W, I_U/I_{BG} and detection limit (D.L.) for DIBK, water and uranium in DIBK at various wavelength

wavelength (nm)	I_{BG} [a] / I_W [b]	I_U [c] / I_{BG}	D.L. of U (mg/l)
263.55	0.51	5.6	0.05
367.01	0.36	16.0	0.02
385.96	10.2	0.8	0.75
393.20	0.5	7.0	0.07
409.01	1.75	2.7	0.18

[a] Background height for DIBK.
[b] Background height for water.
[c] Emission intensity of uranium (10 mg/l).

Table 4
Determination of uranium in apatite minerals by the standard addition and calibration curve method

sample	sample weight (mg)	U added (μg)	U found (μg)	U in sample (10^{-2} %)
Apatite [a]	100	0	10.2	1.02
		25	36.5	1.15
		50	59.5	0.95
		75	83.5	0.85
	150	0	10.3	1.03
		25	37.3	1.23
		50	62.0	1.20
		75	84.3	0.93
	200	0	10.1	1.01
		25	37.0	1.20
		50	61.2	1.12
Average				1.06 ± 0.12 (0.99) [b]

[a] From Florida, USA.
[b] ICP-MS [11].

higher than 0 (1 M HNO_3).

The extraction behaviour of uranium, calcium and phosphate ions from an artificial sample solution of apatite mineral was examined. The volume ratio was again set at two. As can be seen in Fig. 5, the extraction percentage of uranium was almost 100% when the content of apatite mineral was lower than 0.8 g/10 ml, it decreased as the apatite content increased. Calcium and phosphate were not detected in the DIBK phase by ICP-AES irrespective of the amount of apatite mineral examined. It has been shown that uranium can be completely separated from the macro-component of apatite minerals by the extraction method with HPBI.

3.6. Selection of a wavelength for atomic emission intensity measurement of uranium in DIBK after solvent extraction

The background emission intensities for DIBK (I_{BG}) and water (I_W), and the emission intensity for uranium in DIBK (I_U) was measured at the wavelength mentioned in Table 2. In Table 3, the ratios (I_{BG}/I_W, I_U/I_{BG}) and the detection limits for uranium are summarized. The background intensities at 385.96 and 409.01 nm are relatively higher. The highest intensity ratio, the emission intensity (I_U) of uranium (10 mg/l) in DIBK to the background intensity (I_{BG}), and the best detection limit were obtained at 367.01 nm. Therefore, the wavelength of 367.01 nm was selected as being most suitable for the determination of uranium in DIBK.

3.7. Analysis of apatite minerals

The uranium contents in apatite minerals were determined by the proposed method. Based on the procedure mentioned above, a known quantity of uranium was added to the real sample solution and measured. The quantitative results are shown in Table 4. These results were also quite comparable with those measured by ICP-MS [11].

References

[1] O. Fujino, K. Araki, Y. Kometani, M. Sugiyama and M. Matsui, Bunseki Kagaku, 34 (1985) 386.
[2] O. Fujino, T. Atano, M. Sugiyama and M. Matsui, Bunseki Kagaku, 33 (1984) 604.
[3] O. Fujino, Y. Kometani, M. Sugiyama and M. Matsui, Bunseki Kagaku, 33 (1984) 465.
[4] T. Kano and H. Yanagida, Rare Earths – Properties and Applications, Gihodo, Tokyo, 1983, Chap. 2, pp. 32–50.
[5] N.E. Topp, Chemistry of the Rare-Earth Elements, Elsevier, Amsterdam, 1965, Chap. 2.
[6] D. McConnel, Apatite, Springer Verlag, New York, 1973.
[7] A. Jothi and G.N. Rao, Talanta, 37 (1990) 431.
[8] F. Korte and K. Störiko, Chem. Ber., 94 (1961) 1956.
[9] Q.T.H. Le, S. Umetani, H. Takahara and M. Matsui, Anal. Chim. Acta, 272 (1993) 293
[10] P.W.J.M. Boumans, Line Coincidence Table for Inductively Coupled Atomic Emission Spectrometry, Pergamon, Oxford, 1981.
[11] O. Fujino, in preparation.
[12] O. Fujino, M. Matsui, S. Umetani and H. Hiraki, Nippon Kagaku Kaishi, (1989) 39.
[13] S. Umetani and M. Matsui, Bull. Chem. Soc. Jpn., 56 (1983) 3426.
[14] C. Keller and H. Schreck, J. Inorg. Nucl. Chem., 31 (1969) 1121.

ELSEVIER

Analytica Chimica Acta 296 (1994) 69–76

ANALYTICA CHIMICA ACTA

Cathodic stripping potentiometric determination of selenium in biological and environmental materials

S.B. Adeloju *, T.M. Young

Centre for Electrochemical Research and Analytical Technology, Department of Chemistry, University of Western Sydney, Nepean, P.O. Box 10, Kingswood, NSW 2747, Australia

Received 1st November 1993; revised manuscript received 13th April 1994

Abstract

The use of cathodic stripping potentiometry for the reliable determination of trace and ultra trace concentrations of selenium in environmental and biological samples on a glassy carbon mercury film electrode is described. The optimum conditions for the method include 3 M HCl as supporting electrolyte, an electrolysis potential of −100 mV vs. SCE, a constant reduction current of −20 μA and the decomposition of the samples by dry ashing with magnesium nitrate as an ashing aid. Under these conditions, the detection limit is 0.8 μg/l with an electrolysis time of 5 min or 0.04 μg/l with 60 min deposition. The relative standard deviation for the measurement at this level is 6% ($n = 7$). The presence of inorganic and organic substances such as lead, copper, cadmium, zinc, CTMAB, LAS, LPC and Triton X-100 caused some suppression of the selenium peak, but these effects were easily circumvented by the use of standard additions method and a UV-irradiation procedure. The UV-irradiation of the digested environmental and biological materials was also effective in reducing the required dry ashing period to 1 h, as well as in improving the sensitivity and accuracy of the method.

Keywords: Cathodic stripping potentiometry; Selenium; Environmental analysis; Dry ashing

1. Introduction

Selenium is known to be an essential trace element for plants, animals and human, but it is also toxic when ingested in excessive amounts. Several health effects resulting from the deficiency of selenium in diet or its excessive ingestion have been reported in the literature [1–5]. The reliable diagnosis of some of these health effects and the related concern or interest in the levels of this element in the environment often requires its reliable determination in plants, soils, sediments, natural waters and biological materials.

The analytical techniques that have been used for the determination of selenium in these materials include spectrophotometry, colorimetry, fluorimetry, electrothermal atomic absorption spectroscopy, atomic absorption spectroscopy with hydride generation (HG-AAS), chromatography, mass spectrometry, inductively coupled plasma atomic emission spectroscopy, neutron activation

* Corresponding author.

SSDI 0003-2670(94)00182-L

analysis (NAA), x-ray fluorescence and electroanalytical methods including variants of polarography, stripping voltammetry and potentiometry [6–14]. Of these NAA, HG-AAS and the electroanalytical methods offer the necessary sensitivity required for the reliable determination of the low concentrations of selenium commonly found in biological and environmental materials. In particular, the instrumentation required for electroanalytical techniques are cheap and offer the necessary versatility, high sensitivity, simplicity and selectivity required for trace and ultra-trace determination of the element.

Although electroanalytical techniques such as polarography and cathodic stripping voltammetry (CSV) have been widely used for the determination of selenium in various sample materials, the use of cathodic stripping potentiometry (or constant current stripping analysis as it is also known) in this area is still relatively new. Yet the technique is similar in every aspects to CSV, except that it utilizes the application of a constant current for the stripping of the preconcentrated mercury selenide from the electrode. Furthermore, the technique has been demonstrated in the past to be suited to selenium determination [15]. However, cathodic stripping potentiometry (CSP) has only gained a very limited practical use for the determination of selenium in solid biological and environmental materials. To date, there is a considerable amount of uncertainty on the suitability of the technique for the reliable determination of selenium in these materials.

In this study, an attempt is made to develop a practically useful cathodic stripping potentiometric method for the reliable determination of trace and ultra-trace concentrations of selenium in solid biological and environmental materials. Factors investigated for the improvement of the existing CSP technique and its utilization for these sample materials include acid concentration, effects of inorganic and organic substances, and instrumental parameters such as electrolysis time, electrolysis potential and magnitude of constant current. In particular, the use of a sample preparation strategy which involves the combined utilization of a rapid dry ashing and a UV-irradiation procedure for the reliable determination of selenium in biological and environmental materials is carefully investigated. The successful application of the method to a range of biological and environmental materials such as natural waters, fish tissue, horse kidney and human diet will be demonstrated.

2. Experimental

2.1. Instrumentation

All stripping potentiograms were recorded on a computerised Radiometer TraceLab potentiometric stripping system. The working electrode was a glassy carbon mercury film electrode (GCMFE) with a diameter of 3 mm. The auxiliary electrode was a platinum wire (isolated with a bridge filled with the supporting electrolyte) and the reference electrode was a saturated calomel electrode (SCE). The cells used for all measurements were made of polypropylene. A Raytech portable dual ultra-violet lamp was employed for all sample irradiation.

2.2. Reagents

All chemicals were of analytical reagent grade. All dilutions and sample preparations were made with Milli Q water. The mercury(II) stock solution (800 mg/l) was prepared by dissolving mercuric chloride in 1.3 M HCl, while the selenium(IV) stock solution was prepared by dissolving selenium metal with aqua regia and making to required volume with 6 M HCl. The fish tissue, human diet and horse kidney reference materials were obtained from IAEA (Vienna). The river water samples were collected from the Nepean river in Western Sydney, Australia.

2.3. Glassware

All glassware, polypropylene sample cups and polyethylene bottles were filled with an acid mixture containing 1 M HCl and 1 M HNO_3, left to stand for at least one week and were rinsed several times with Milli-Q water prior to use.

2.4. Procedures

Preparation of mercury film

The glassy carbon electrode was polished initially with 5 μm aluminium oxide and lint free paper prior to the plating of mercury. A deposition potential of −900 mV vs. SCE was then applied to the electrode for 1 min in a solution containing 800 mg/l Hg. The GCMFE was reconditioned during the analysis with the presence of 20 mg/l Hg and a rest potential of −500 mV was applied to the electrode when standing in acidic solutions.

Measurements

The TraceLab system was programmed to apply a deposition potential of −100 mV vs. SCE, a standard deposition time of 100 s, a resting potential of −500 mV and a constant current of −20 μA during the stripping step.

Sample decomposition

0.5 g of sample was weighed accurately into a pre-cleaned crucible and heated gently on a water bath until most of the moisture and volatile organic matter was removed. Once the sample was dried it was wet digested with 2 ml of concentrated nitric acid and 2.5 ml of 80% w/v magnesium nitrate before dry ashing in a muffle furnace for 1 h at 500°C. The ashed sample was then removed from the furnace and allowed to cool prior to the addition of 50 ml of 6 M hydrochloric acid. The resulting mixture was heated gently on a hot-plate for 30 min to dissolve the ashed sample, release the selenium from any insoluble residue and convert all the selenium to the Se(IV) form. The digest was then transferred quantitatively into a 100 ml volumetric flask and made up to volume with Milli-Q water. When required, the sample was irradiated with the UV lamp for 4 h on shortwave (< 2500 Å) and 2 h longwave (3200–4000 Å) prior to measurement by CSP. Evaporation losses were corrected by adjusting sample volume to the original value.

Water sample

The sample was collected in a HDPE container, filtered through a 0.45 μm Nylon 66 filter, acidified to pH 1 using HNO_3 and stored below 4°C until required.

3. Results and discussion

3.1. Optimization of selenium CSP response

It has been previously established [16,17] that hydrochloric acid is a suitable supporting electrolyte for the reliable determination of selenium by CSV. In this medium, the associated electrode processes involve the preconcentration of selenium (IV) as mercury selenide:

$$Se^{4+} + Hg + 4e^- \rightarrow HgSe$$

and the subsequent reduction of this substance during the reductive stripping step:

$$HgSe + 2H^+ + 2e^- \rightarrow H_2Se + Hg$$

The influence of hydrochloric acid concentration on the CSP determination of selenium is illustrated in Fig. 1. Interestingly, unlike the CSV measurement of selenium on a hanging mercury drop electrode [16], the sensitivity of the CSP response increased with increasing acid concentration up to 5 M. The most sensitive response was obtained in 4–5 M HCl. However, in order to reduce acid consumption and to make the technique more amenable to digested samples which usually requires the reduction of selenium(VI) to

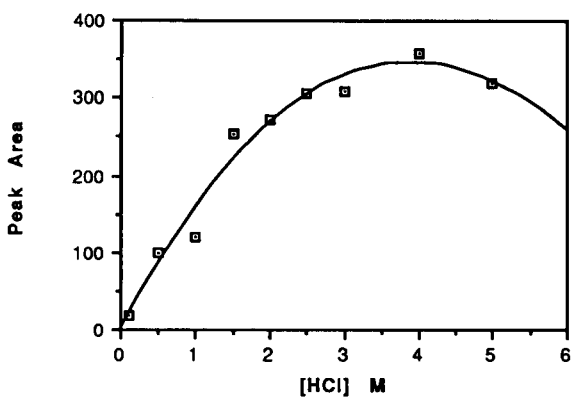


Fig. 1. Influence of hydrochloric acid concentration on selenium stripping peak. 20 μg/l selenium, t_{dep}: 100 s, −20 μA constant current and E_{dep}: −100 mV vs. SCE.

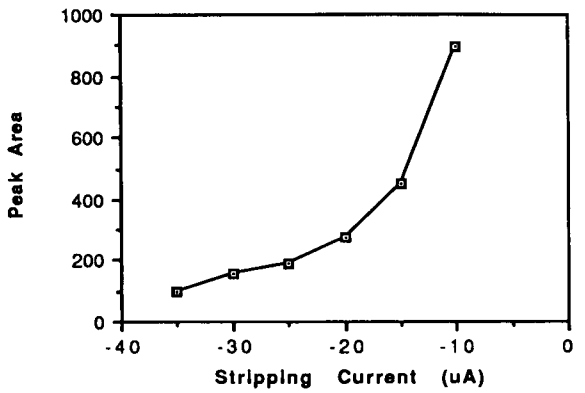


Fig. 2. Influence of applied constant stripping current on CSP response of selenium. 3 M HCl, other conditions as in Fig. 1.

selenium(IV) by pretreatment with 6 M HCl, the use of the 3 M hydrochloric acid as supporting electrolyte is preferred. This will enable a relatively simple dilution of the pretreated samples prior to the CSP determination and will have no significant effect on the sensitivity of the technique. Although the linear range is lower than can be accomplished with 5 M HCl, the 0–150 μg/l range obtained with the lower acid concentration is sufficient for most biological and environmental materials.

Fig. 2 shows the influence of the magnitude of the constant current employed during the stripping step on the sensitivity of the selenium response. The application of a constant current of -10 μA gave the most sensitive response for selenium. However, the resolution of the response obtained with this stripping current was inadequate. In contrast, the application of a stripping current of -20 μA gave a well resolved selenium response, but resulted in some loss in sensitivity. A constant current of -20 μA was therefore employed for all other work in view of the improved selenium peak resolution.

The influence of the electrolysis potential on the sensitivity of the selenium response is illustrated in Fig. 3a. Although the optimum response was obtained with the use of an electrolysis potential of 0 mV vs. SCE, this was however irreproducible and poorly defined. The use of an electrolysis potential of -100 mV gave a well resolved, reproducible and reasonably sensitive response and was adopted for all subsequent measurements.

The increasing electrolysis time resulted, as illustrated in Fig. 3b, in a corresponding increase in the sensitivity of the selenium response up to 20 min. Beyond this electrolysis time, the amount of selenium deposited is limited by the saturation of the electrode surface, which resulted in the non-linear relationship between the t_{dep} and peak area with electrolysis time >20 min range. However, the lowest detectable amount of selenium by CSP is highly dependent on the chosen electrolysis time. For example, the lowest detectable amount of selenium with an electrolysis period of 5 min is 0.8 μg/l (R.S.D. = 6%), but this can be lowered to as little as 0.04 μg/l (R.S.D. = 12%)

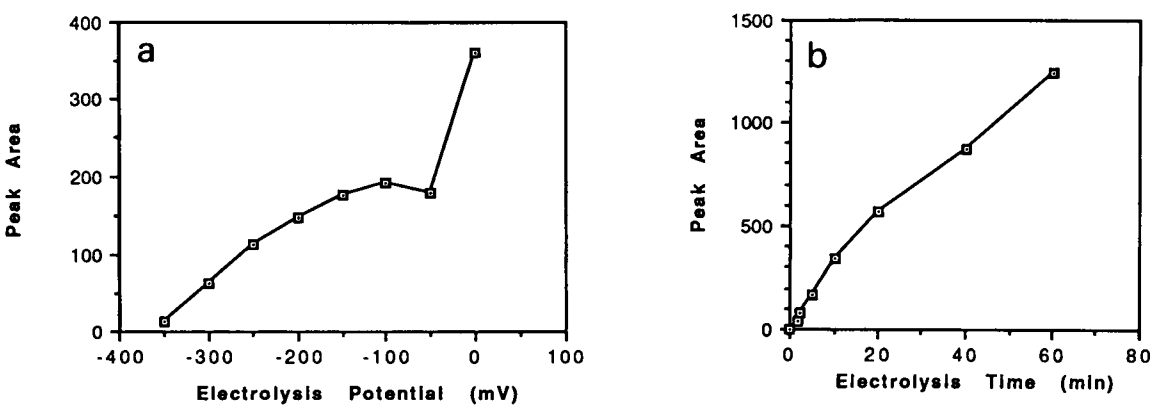


Fig. 3. Dependence of selenium stripping peak on choice of (a) electrolysis potential and (b) electrolysis time. 3 M HCl; other conditions as in Fig. 1, except that 5 μg/l selenium was used for (b).

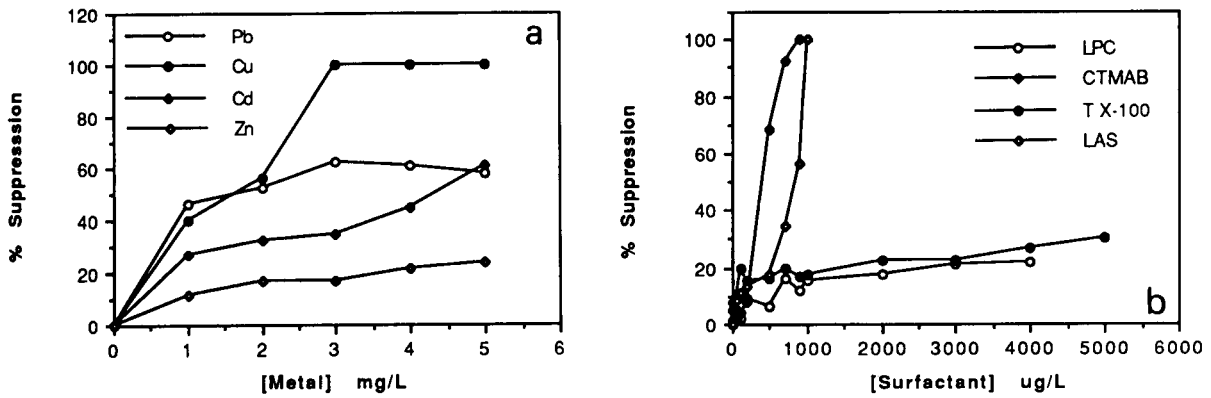


Fig. 4. Suppressive effect of (a) heavy metals and (b) surface active substances on selenium response. 20 μg/l Se, other conditions as in Figs. 1 and 3.

with the use of an electrolysis time of 60 min. It is important to highlight the significant differences between these values and those obtained by the calculation of the detection limit, based on three times standard deviation. From a practical viewpoint, the data for the lowest detectable amount of the selenium standard is more relevant and useful. It is therefore recommended that this value should be given more consideration when employing CSP for selenium determination. Usually at the calculated detection limits no selenium response was observed, whereas a selenium response is clearly evident at the specified lowest detectable amount.

3.2. Interferences

The presence of inorganic and organic substances in samples can have some effect on the selenium response, as has been previously demonstrated for CSV [16]. Fig. 4a shows the effect of some of the more common heavy metals on the CSP response for selenium. With the exception of copper, the presence of relatively high concentrations ($\geq$ 5 mg/l) of these metals resulted in no more than 60% suppression of the selenium peak. The observed peak suppression by these heavy metals may be due to the formation of intermetallic compounds, as has been previously suggested [16]. However, in real samples the concentrations of some of these metals, particularly cadmium and lead, will be considerably lower than 5 mg/l and therefore the degree of peak suppression will be much lower. This view is clearly demonstrated by the results in Fig. 4a which shows that the selenium peak suppression is less than 20% when a five fold amount of the interferant is present, less than 40% in the pres-

Table 1
Influence of electrolysis time on the sensitivity of the Se response

Electrolysis time (min)	Lowest detectable amount (μg/l)	Linear concentration range (μg/l)	Linear correlation coefficient	Detection limit [a] (μg/l)	Sensitivity [area/(μg/l)]
1.7	2	0–150	0.989	1.3	11.0
2.5	1.4	0–100	0.989	0.7	15.5
5.0	0.8	0–100	0.976	0.2	33.2
10	0.5	0– 70	0.971	0.1	67.7
20	0.2	0– 30	0.963	0.04	112.7
40	0.09	0– 20	0.925	0.03	175.4
60	0.04	0– 10	0.983	0.01	249.4

[a] Based on 3 times standard deviation.

Table 2
Influence of hydrochloric acid concentration on the analytical characteristics of the selenium CSP response

Hydrochloric acid concentration	Linear concentration range (μg/l)	Sensitivity [area/(μg/l)]	Correlation coefficient
0.1	0– 50	1.0	0.937
0.5	0– 70	4.9	0.994
1.0	0– 70	6.0	0.990
2.0	0– 70	13.1	0.964
2.5	0–150	15.3	0.990
3.0	0–150	15.5	0.989
5.0	0–300	16.0	0.989

ence of a 25 fold amount of interferant and less than 60% when a 100 fold amount of interferant is present. Evidently, the CSP determination of selenium in samples containing up to 100 fold or more of the inorganic interferants is still possible, provided that the quantification is made by use of standard additions.

The suppressive effect of surface active substances on the CSP determination of selenium is illustrated in Fig. 4b. These surface active substances interfered with the selenium measurement by blocking the electrode surface and, hence, reducing the access for the analyte to be deposited onto the electrode surface. Cetyltrimethyl ammonium bromide (CTMAB), lauryl pyridium chloride (LPC), lauryl sodium dodecyl benzene (LAS) and Triton X-100 (TX-100) had a significant suppression effect on the selenium response. The most pronounced suppression was observed in the presence of cetyltrimethyl ammonium bromide and lauryl sodium dodecyl benzene due to their higher surface activity. However, when applied to solid biological and environmental materials the organic substances will create no problem as these are readily removed during the sample decomposition and/or with UV irradiation.

3.3. Analytical applications

The achievable linear concentration range for the CSP determination of selenium, as with other stripping techniques, can be influenced by the electrolysis time and the hydrochloric acid concentration. This view is clearly supported by the data in Tables 1 and 2. In low hydrochloric acid concentrations (≤ 2 M), the selenium response was linear with increasing concentration within 0 and 70 μg/l, whereas in higher hydrochloric acid concentrations the linear range was extended to ≥ 150 μg/l. The best linear concentration range of 0–300 μg/l was obtained in 5 M hydrochloric acid solution, but the sensitivity of the selenium peak was about the same as that obtained in 3 M hydrochloric acid (Fig. 1). The data in Table 1 also indicate that the linear concentration range for selenium can be extended by reducing the electrolysis time, whereas the sensitivity and the detection limit can be significantly improved by increasing t_{dep}. The observed reduction of the linear concentration range with the increasing electrolysis time is due to the increased tendency for the saturation of the electrode surface with mercury selenide at the longer electrolysis time.

The application of CSP to the reliable and accurate determination of selenium in biological and environmental materials was accomplished

Table 3
Selenium concentrations found in standard reference materials by cathodic stripping potentiometry

Reference Sample	Type of sample	Selenium Content		
		Certified (μg/g) [a]	Value found with UV pretreatment (μg/g) [a]	Value found without UV pretreatment (μg/g) [a]
IAEA/Monaco MA-A-2 (TM)	Fish tissue ($n = 5$)	1.7 ± 0.3	1.68 ± 0.31	1.38 ± 0.29
IAEA H8	Horse kidney ($n = 4$)	4.7 ± 0.3	4.68 ± 0.30	4.22 ± 1.70
IAEA H9	Mixed human diet ($n = 4$)	0.11 ± 0.01	0.10 ± 0.01	0.080 ± 0.020

[a] Error quoted as mean deviation.

by modifying a previously reported dry ashing method for the CSV determination of the element in such samples [17]. Fig. 5 shows that the dry ashing method which is based on the use of magnesium nitrate as an ashing aid and subsequent UV-irradiation of the sample was suitable

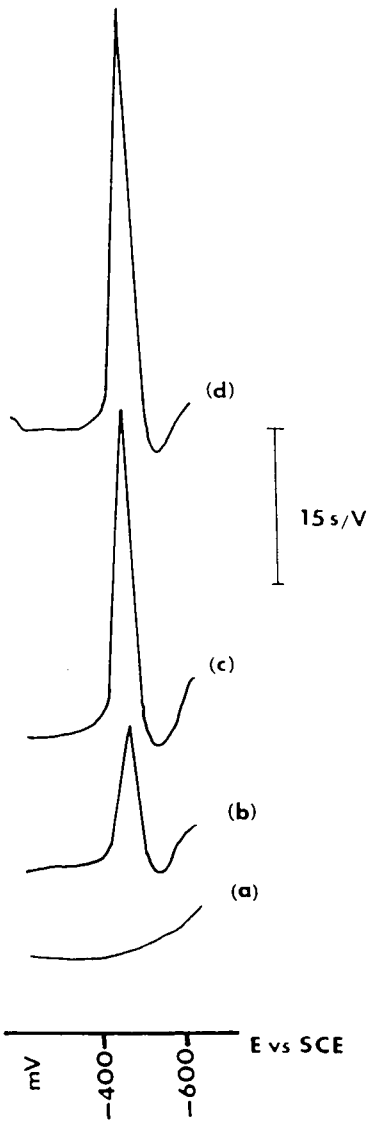


Fig. 5. Quantification of selenium in dry ashed and UV-irradiated fish tissue sample using the standard additions method. E_{dep}: −100 mV vs. SCE, t_{dep}: 5 min and constant current: −20 μA; (a) reagent blank, (b) fish tissue, (c) fish tissue +10 μg/l Se, (d) fish tissue +20 μg/l Se.

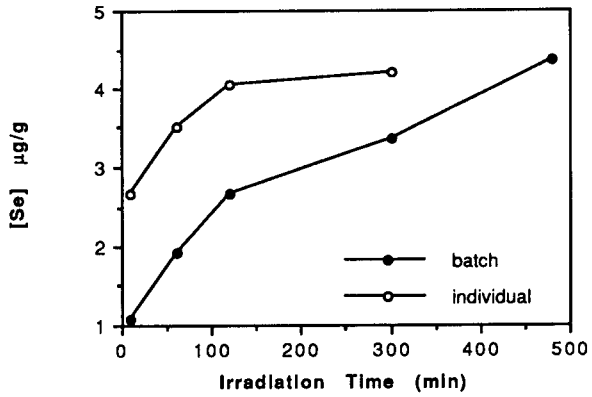


Fig. 6. Influence of UV-irradiation on selenium concentrations found in dry-ashed horse kidney sample. Other conditions as in Fig. 5.

for the reliable CSP determination of selenium. The data in Table 3 highlights the need for the UV irradiation of the dry ashed samples by its improvement of the accuracy of the CSP method. Generally, the peak resolution and precision of the selenium response improved considerably with the UV irradiation and the concentration of selenium found in the samples was more reliable, as demonstrated in Table 3. Fig. 6 illustrates the influence of the UV-irradiation on the concentration of selenium found in the dry ashed horse kidney sample. As can be expected, the use of large batch volume of sample required longer UV-irradiation period for the removal of the dissolved organic interferant than with small aliquots of sample taken in individual vessels. These results suggest that the short dry ashing period (1 h) employed in this study does not completely remove the soluble organic substances in the decomposed horse kidney sample. The UV-irradiation of the dry ashed samples for 6 h in the individual vessels ensures the reliable determination of selenium in environmental and biological materials by CSP. This is also clearly demonstrated by the data in Table 3 which indicate that the selenium concentration obtained for the dry ashed fish tissue, horse kidney and mixed human diet reference materials agreed favourably with the certified values when UV-irradiation of the sample is employed.

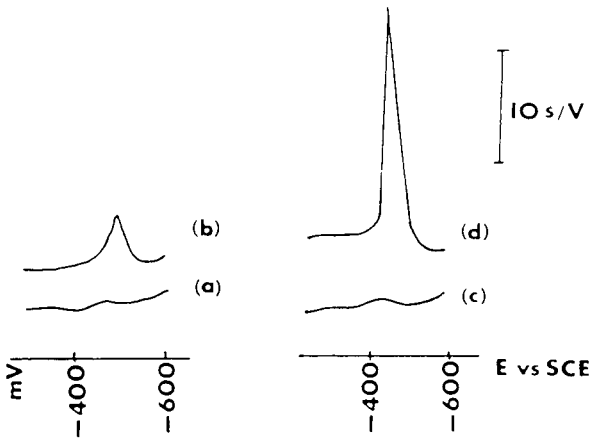


Fig. 7. CSP determination of dissolved selenium in a river water sample. (a) water sample, no heating, (b) as (a) +20 μg/l Se, (c) water sample with heating, (d) as (c) +20 μg/l Se. Electrolysis time: 60 min.

The successful application of the CSP method to the determination of dissolved selenium in a river water sample is demonstrated in Fig. 7, based on the use of standard additions procedure. The resolution of the selenium response obtained in this sample was improved when the water was filtered, acidified and heated in the presence of hydrochloric acid (as described for the conversion of the element to Se(IV) form) prior to the measurement. This is possibly due to the thermodynamically favourable but slow conversion of Se(VI) species to the electroactive Se(IV) species according to the equation [18]

$$H_2SeO_4 + 2H^+ + 2e^- \rightarrow H_2SeO_3 + H_2O$$

This process also appears to partially reduce the residual organic matter in the sample. The mean concentration of selenium found by CSP in the river water samples was 0.18 ± 0.06 μg/l ($n = 3$, error is mean deviation).

4. Conclusion

The combined use of cathodic stripping potentiometry and dry ashing with magnesium nitrate is suitable for the reliable determination of selenium in biological and environmental materials. As little as 0.8 μg/l can be detected by the method with an electrolysis time of 5 min or 0.04 μg/l with 60 min. Interferences from inorganic and organic substances were adequately overcome by use of standard additions for quantification and the removal of organic substances by UV irradiation. The determination of ultra-trace amounts of selenium in natural waters by CSP was also successfully demonstrated.

Acknowledgments

The authors are grateful to the University of Western Sydney, Nepean for the provision of research support and a postgraduate research scholarship for this work.

References

[1] D. Purves, Fundamental Aspects of Pollution Control and Environmental Science 7: Trace Elemental Contamination of the Environment, Elsevier, New York, 1985.
[2] G. Bluhm, Trace Elements in Medicine, 7 (1990) 139.
[3] C.E. Casey, Proc. Nutrition Soc., 47 (1988) 55.
[4] J. Parizek, Food Chem. Toxicol., 28 (1990) 763.
[5] J.E. Oldfield, Am. Inst. Nutrition., J. Nutrition, 117 (1987) 2002.
[6] L.S Clesceri, A.E Greenberg and R. Rhodes Trussell (Eds.), Standard Methods for the Examination of Water and Wastewater, 17th edn., American Public Health Association, 1989.
[7] K. Helrich (Ed.), Official Methods of Analysis (AOAC), Vol. 1, 15th edn., 1990.
[8] J.W. Forstrom, J.J. Zakowski and A.L. Tappel, Biochemistry, 17 (1978) 2639.
[9] E. Steinnes, Int. J. Appl. Radiat. Isotop., 18 (1967) 731.
[10] M.W. Blades, J.A. Dalziel and C.M. Elson, J. Assoc. Off. Anal. Chem, 59 (1976) 1234.
[11] S. Forbes, G.P. Bound and T.S. West, Talanta, 26 (1979) 473.
[12] J.K. Christensen, L. Kryger, J. Mortensen and J. Rasmussen, Anal. Chim. Acta, 121 (1980) 71.
[13] C.M.G. van den Berg and S.H. Khan, Anal. Chim. Acta, 231 (1990) 221.
[14] J. Wang and W.J. Sun, J. Electroanal. Chem., 291 (1990) 59.
[15] C. Hua, D. Jagner and L. Renman, Anal. Chim. Acta, 197 (1987) 257.
[16] S.B. Adeloju, A.M. Bond, M.H. Briggs and H.C. Hughes, Anal. Chem., 55 (1983) 2076.
[17] S.B. Adeloju, Analyst, 114 (1989) 455.
[18] D.F. Shriver, P.W. Atkins and C.H. Langford, Inorganic Chemistry, Oxford University Press, Melbourne, 1990, pp. 392.

ELSEVIER

Analytica Chimica Acta 296 (1994) 77–86

ANALYTICA CHIMICA ACTA

Square-wave voltammetric determination of lead(II) with a Nafion®/2,2-bipyridyl mercury film electrode

Jyh-Myng Zen *, Sen-Yuan Huang

Department of Chemistry, National Chung-Hsing University, Taichung, Taiwan

Received 4th January 1994; revised manuscript received 19th April 1994

Abstract

A chemically modified electrode (CME) system consisting of a Nafion-coated mercury film electrode containing appropriate amounts of 2,2-bipyridyl for practical use in the analysis of lead(II) was investigated. The relative advantages of using square-wave stripping voltammetry (SWSV) in combination with the CME for the determination of the reversible lead(II)–2,2-bipyridyl complex was assessed. The main advantages of the CME are better mechanical stability of the mercury film, and improved resistance to interference from both surface-active compounds and those metals which commonly interfere in anodic stripping measurements. For a 5-min preconcentration period in the presence of oxygen, a linear calibration curve from 1 to 100 $\mu g\ l^{-1}$ was obtained, and the detection limit was 0.1 $\mu g\ l^{-1}$. A lower detection limit can be obtained when longer preconcentration times are used. The CME surface could be regenerated by exposure to acid. For 10 successive preconcentration/determination/renewal cycles, the SWSV response could be reproduced with 5% relative standard deviation. Rapid and convenient acid renewal allows the use of an individual modified electrode surface in continuous-monitoring environmental or clinical applications.

Keywords: Stripping voltammetry; Lead; 2,2-Bipyridyl; Nafion; Mercury film electrode

1. Introduction

Due to new concerns regarding lead poisoning in children [1], there is a growing need for improved analytical methods for monitoring this metal. Numerous analytical methods are available for trace measurements of lead(II) [1]. Among these, electrochemical stripping analysis has been widely used for this task because of its remarkable sensitivity and suitability for decentralized testing [2–4]. However, improved electrochemical techniques are desired to address interference problems, to shorten the analysis time, or to allow the use of an individual electrode surface in continuous-monitoring applications. The present investigation was prompted by a desire to develop an alternative method for the determination of lead(II) due to the problems mentioned below. A sensitive adsorptive stripping scheme for lead(II) in the presence of oxine was reported [5], but did not possess the desired selectivity (particularly due to severe copper interference). A recent article by Wang et al. [4] reported a very sensitive

* Corresponding author.

SSDI 0003-2670(94)00194-Q

and selective stripping procedure, based on the adsorptive accumulation of the lead(II)–*o*-cresolphthalexon complex onto mercury electrode, for trace measurements of lead(II) in the presence of a large excess of copper. However, this procedure apparently can not be applied to the continuous-monitoring applications.

We report here an interesting chemically modified electrode (CME), designated as NC(Bpy) MFE, for the determination of lead(II) in the presence of a large excess of copper. CMEs have received a great deal of attention [6,7], particularly to enhance the sensitivity and selectivity of electrochemical techniques [8]. Generally, such procedures involve the immobilization of the chemically active species on the electrode surface. The target analyte is preconcentrated to the modifier, which has a particular affinity for the analyte. In this way, the sensitivity and selectivity of the CME are determined by the chemical reactivity of the modifier rather than the oxidation/reduction potential of the analyte. Many modification methods have been used to introduce modifiers onto the electrode surface for adsorptive stripping voltammetry [9]. Among these, polymeric coatings have attracted great interest for ease of preparation, chemical and physical stability, and complete surface coverage. Of these, perfluorinated Nafion films have been extensively examined because of their utility in entrapping positively charged analytes [10,11]. In effect, Nafion-coated mercury film electrode (NCMFE) has been proved to be very promising in preventing interference from surface-active compounds [12]. Unfortunately, considering the interference effect of the common ions, the NCMFE generally offers no more advantage than the mercury film electrode (MFE). Whereas, by incorporating a suitable modifying group capable of binding only a few metals, the formulation of a number of highly selective determinations should be possible.

Recently, our group is developing a new type of NCMFE by incorporating suitable chelating agents into the electrode. For example, a CME system consisted of a NCMFE containing appropriate amounts of dimethylglyoxime (DMG), designed as NC(DMG)MFE, was used for practical use in analysis of nickel(II) [13]. In this way, we can take advantage of three effects: the preconcentration effect of the ion-exchange polymer (Nafion), the selective chelating effect of the chelating agent (DMG), and the accumulation effect of the electrode (Hg). The electrodes thus developed were found to possess the properties anticipated and to be suitable for voltammetric quantitation of traces of nickel(II), which were preconcentrated selectively on the NC(DMG) MFE surface from a variety of complex chemical systems. Using the same strategy as the NC(DMG)MFE, 2,2-bipyridyl (Bpy) is the selective complexing reagent used to fabricate the CME in this work. The improvements offered by using SWSV on the NC(Bpy)MFEs for monitoring trace lead(II) are illustrated so as to establish the advantages of the technique in continuous-monitoring applications. Such a technique may be useful for children screening and prevention programs, for assessing industrial exposure to lead(II), and for monitoring trace lead(II) in (drinking and waste) water systems.

2. Experimental

2.1. Reagents and chemicals

Doubly-distilled deionized water was used to prepare all solutions. Nafion perfluorinated ion-exchange powder, 5 wt% solution in a mixture of lower aliphatic alcohols and 10% water, was obtained from Aldrich. Mercury(II) nitrate, 2,2-bipyridyl, Triton X-100, ammonium chloride, sodium acetate, and potassium dihydrogenphosphate were of analytical grade. All buffers and supporting electrolyte solutions were prepared from Merck (Suprapur reagents). Lead(II) standard solution (1 g l^{-1}) and all other standard metal solutions used in interference studies were also obtained from Merck.

2.2. Apparatus

All voltammetry experiments were performed with the BAS-100B electrochemical analyzer. The three-electrode system consists of a NC(Bpy)MFE

working electrode, a Ag/AgCl (3 M NaCl) reference electrode, and a platinum wire counter electrode.

2.3. Procedure

The glassy carbon disk electrode ($A = 0.0685$ cm^2, BAS) was polished with BAS polishing kit and rinsed with deionized water, then further cleaned ultrasonically in 1:1 nitric acid and deionized water successively. Nafion coating solutions (the total volume of the coating solution is always 4 μl, with different proportions of Nafion and Bpy) were spin coated onto the electrode surface at 2800 rev min^{-1}. A uniform thin film was formed by evaporating the solvent after about 5 min of spinning.

Mercury was deposited onto the glassy carbon/(Nafion, Bpy) substrate by adding 5×10^{-4} M mercury(II) nitrate to the supporting electrolyte medium at -800 mV vs. Ag/AgCl for 6 min as described by Hoyer et al. [12]. The supporting electrolyte medium contains 0.07 M acetate buffer (pH 4.0) prepared from sodium acetate and nitric acid.

Unless otherwise stated, a medium containing pH 4.0 potassium dihydrogenphosphate buffer was used in the electrochemical experiments.

Drinking water was collected from a water fountain at Chung-Hsing University. In order to fit into the linear detection range of square-wave (SW) voltammetry, the sample was diluted by a factor of 5. Electroplating waste solution was first acidified to pH < 2.0 with an appropriate amount of H_2SO_4 and then filtered with a 0.45-μm sterilized membrane. The dilution factors of electroplating waste solution used for lead(II) measurement was 10.

3. Results and discussion

3.1. Electrochemical behaviour

Fig. 1 shows SW voltammograms obtained with bare glassy carbon electrode (1), NCMFE (3), and NC(Bpy)MFE (4) in a phosphate buffer solution (pH 4.0) containing 0.1 mg l^{-1} lead(II) after 5 min accumulation at -0.6 V. The oxidation of the lead(II)–Bpy complex yields a well-defined SW peak current with a peak potential of -0.49 V (Fig. 1(3)). This lead(II) wave was not observed with the bare glassy carbon electrode (Fig. 1(1)), and a relatively smaller peak current was visible with the NCMFE (Fig. 1(3)). The signal was found to increase with lead(II) concentration and chemical deposition time using the NC(Bpy)MFE. The SW forward and reverse current voltammograms of the oxidation of the lead(II)–Bpy complex are shown in Fig. 2A, which clearly demonstrates that the above reaction is reversible. It is well known that the sensitivity of SW stripping of adsorbed species is proportional to the degree of reversibility of the electrochemical reaction [14,15]. Therefore, the species that undergo reversible reactions benefit more by the application of SW than those that undergo quasi-reversible or totally irreversible reactions. By carrying out the experiments with square-wave stripping voltammetry (SWSV), differential pulse stripping voltammme-

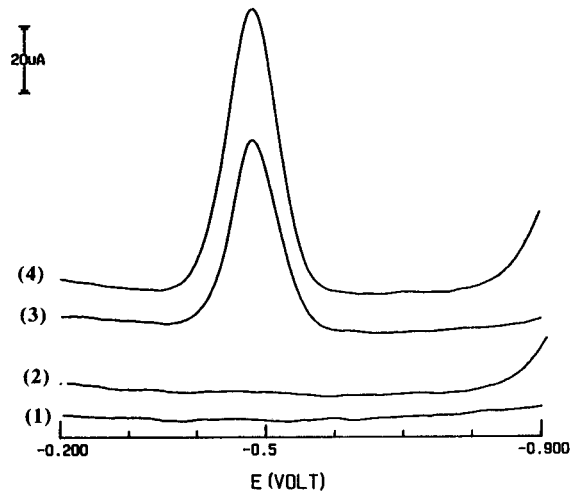


Fig. 1. SW voltammograms for lead(II) on different electrodes. (1) Blank glassy carbon electrode in pH 4 phosphate buffer solution containing 0.1 mg l^{-1} lead(II) after 5 min accumulation at -0.9 V. (2) NC(Bpy)MFE in blank pH 4 phosphate buffer solution. (3) NCMFE: conditions as in (1). (4) NC(Bpy)MFE: conditions as in (1). Effective scan rate 640 mV/s, modulation amplitude 50 mV, modulation frequency 160 Hz.

try (DPSV), and linear sweep stripping voltammetry (LSSV) at the same effective scan rate, the above prediction agrees well with the observed behaviour as shown in Fig. 2B. Comparing the sensitivity between SWSV and LSSV, the reversible lead(II)–Bpy oxidation shows an obvious improvement. In terms of background rejection, LSSV gives the highest sloping baselines which causes a difficulty in the measurement of low concentrations. SWSV competes with DPSV for background rejection ability, but has at least 4 times superior sensitivity than DPSV in practical applications. Since the reduction of oxygen is a totally irreversible reaction, the SW waveform is expected not to respond to it [14,15], thus providing a means of carrying out the analysis without deoxygenating the sample. The peak currents of 0.1 mg l^{-1} lead(II) measured by SWSV both before and after deoxygenation of the sample were found to be almost identical. The electrochemical study in this work was therefore carried out in the presence of oxygen.

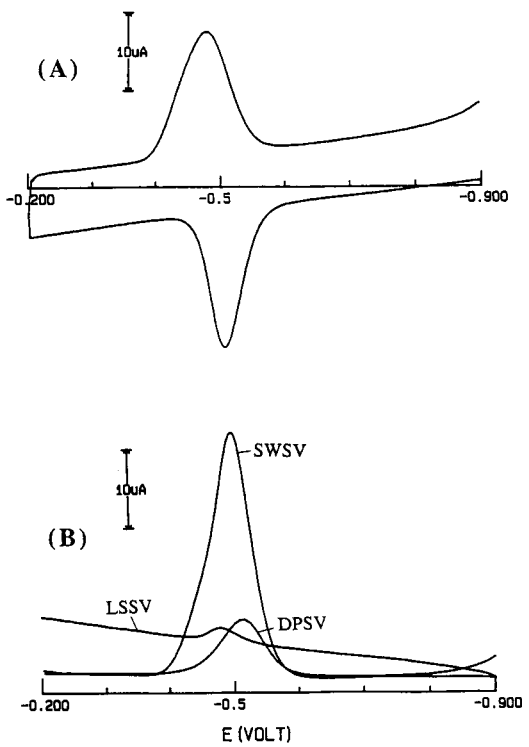


Fig. 2. (A) The SW forward and reverse current voltammograms for the oxidation of the lead(II)–Bpy complex. Conditions: modulation amplitude, 50 mV; modulation frequency, 100 Hz. (B) Comparison between SWSV, DPSV, and LSSV for 0.1 mg l^{-1} lead(II) obtained with the NC(Bpy)MFE at the same effective scan rate of 200 mV/s. Conditions: (SWSV) as in (A); (DPSV) pulse height 50 mV, pulse width 20 mV.

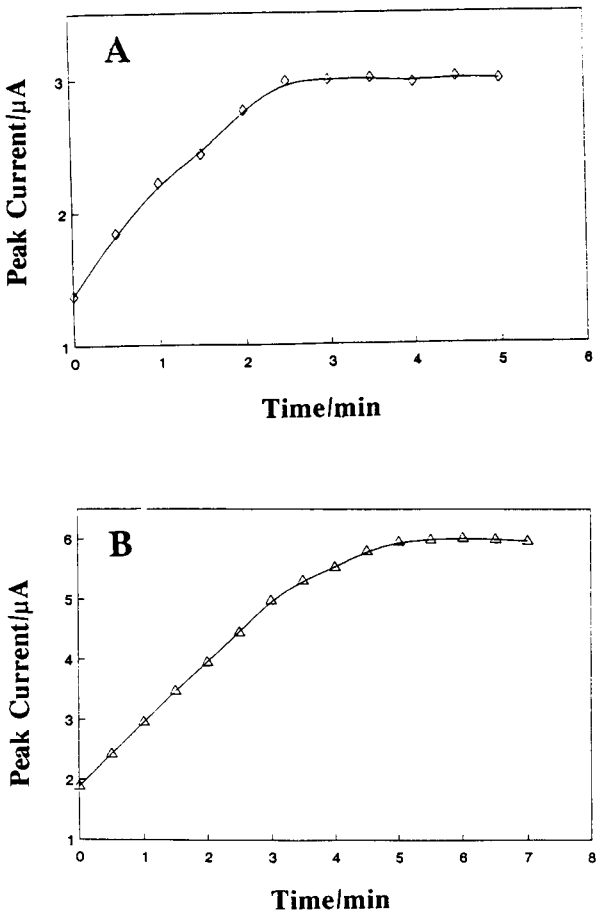


Fig. 3. Peak currents of lead(II) determination obtained at the NC(Bpy)MFE by successive determinations every 30 s. [Bpy]: A, 1 wt%; B, 3 wt%.

The equilibrium process of the above electrodes was further studied by the following experiment. Peak currents of lead(II) were measured at the CMEs with different amounts of Bpy in the films every 30 s. Fig. 3 shows the result for CME prepared from coating solution that contain 1 ml of Nafion + 1 ml of 1 wt% Bpy (A), and that of 1 ml of Nafion + 1 ml of 3 wt% Bpy (B). Due to the gradual increasing concentration of lead(II) in the film through exchange between Nafion and

lead(II) and through chelating between Bpy and lead(II), the observed peak currents are expected to increase with time and eventually level off after saturation. As can be seen, the time required to level off increases from 2.5 to 5 min as the content of Bpy in the coating solution increases from 1 to 3 wt%. This phenomenon is apparently a good indication of the proper chelating process between Bpy and lead(II).

3.2. Optimum conditions for electrochemical determinations

In order to arrive at the optimum conditions for the determination of lead(II), there are two aspects that should be considered: the electrode and the detection. For the electrode aspect, the principal factors governing the performance of the NC(Bpy)MFE are the concentration of Bpy in the film, the thickness of the Nafion/Bpy film, and the deposition of mercury. For the detection aspect, the factors consist of the pH of the solution, the preconcentration time, the preconcentration potential, and the SW parameters.

CMEs prepared from coating solutions that contain 1 ml of Nafion + 1 ml of Bpy (1–5 wt% in ethanol) were examined under identical conditions as shown in Fig. 4A. In all cases, lead(II) responses could be obtained and the peak current increased as the content of Bpy in the coating solution increased and reached a maximum at around 3 wt%. Apparently, the chelating ability of Bpy with lead(II) functions properly in the CME. The current starts to decrease when the Bpy content is higher than 3 wt%. This can be explained by the films gradually losing adhesion with the glassy carbon surface due to overloading of the lead(II)–Bpy complex. Electrodes prepared by the optimum combination of coating solution of 1 ml of Nafion + 1 ml of Bpy (3 wt% in ethanol) were therefore used in all subsequent work. The thickness of the Nafion/Bpy film directly controls the amount of Bpy in the film and hence the electrode performance. Nevertheless, the optimum film thickness apparently depends on both the diffusion process of the lead(II)–Bpy complex in the film and the maximum loading that does not affect the adhesion of the film to

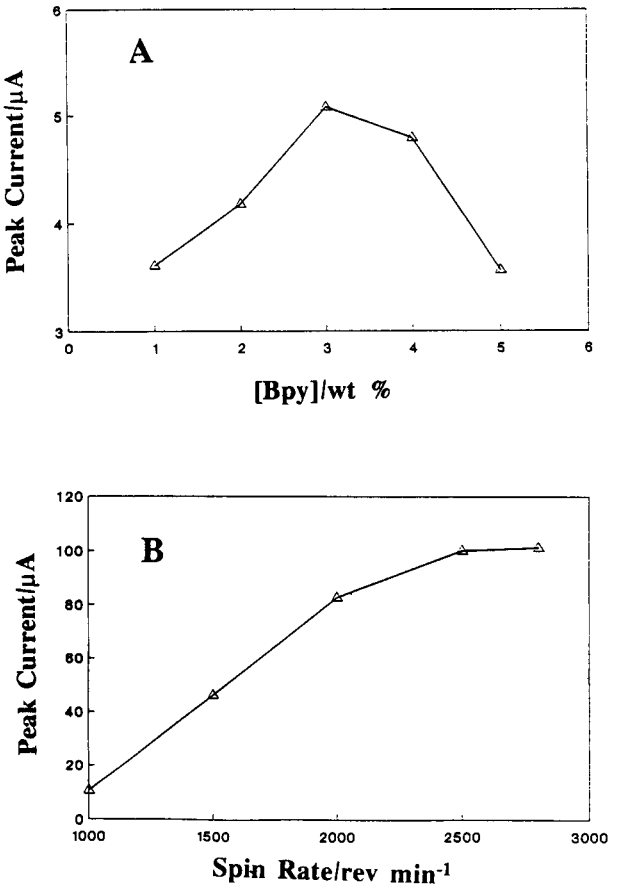


Fig. 4. Effect of the concentration of Bpy in the film (A) and the film thickness (B) to the peak current of 0.1 mg l^{-1} lead(II) obtained at the NC(Bpy)MFE. Conditions: (A) SWSV modulation amplitude, 25 mV; modulation frequency, 15 Hz. Preconcentration time (t_p) = 3 min and preconcentration potential (p_p) = −0.6 V. (B) SWSV modulation amplitude, 50 mV; modulation frequency, 160 Hz. Other conditions as in (A).

the glassy carbon surface. The film thickness was varied by preparing the electrodes with the optimum coating solution, 1 ml of Nafion + 1 ml of Bpy (3 wt% in ethanol), at different spin-coating rates as shown in Fig. 4B. The peak current levels off around 2500 rev. min^{-1} and starts to decrease rapidly when the spin-coating rate is lower than 2000 rev. min^{-1}. Since the film thickness increases as the spin-coating rate decreases, this

phenomenon might also be explained by the films gradually losing their adhesion to the glassy carbon surface due to the overloading of the lead(II)–Bpy complex. Electrodes prepared with the optimum coating solution at the 2800 rev. min^{-1} spin-coating rate were therefore used in all subsequent work.

Mercury electrodeposited onto the glassy carbon/(Nafion, Bpy) substrate from mercury(II)-containing solutions could clearly be seen as a greyish deposit formed. The amount of mercury plated onto the CME depends on the deposition time. It was reported earlier that the reduction of the mercuric ions and the growth of the mercury phase might take place at the glassy carbon/Nafion interface [12]. On the other hand, there is still a possibility that the mercury may fill up the pores of the Nafion/Bpy coating and extend beyond it. The results were found to be similar to our previous study of using the NC(DMG)MFE in the determination of nickel(II) [13]. A deposition time of 6 min was chosen in all subsequent work. In summary, as to the electrode aspect, the optimum conditions for electrode preparation are as follows: coating solution of 1 ml of Nafion + 1 ml of the Bpy (3 wt% in ethanol), 2800 rev. min^{-1} spin-coating rate, and 6 min of mercury deposition time.

As to the detection aspect, the optimum pH ranges for 0.1 mg l^{-1} lead(II) accumulation were observed when the sample was slightly acidic, pH 3–5, as shown in Fig. 5A. Among various buffer solutions, the phosphate buffer solutions were found to be the best. In alkaline medium, because of the hydrolysis of lead(II) and lead(II)–Bpy complex, the preconcentration of lead(II) onto the modified electrode surface is thus less effective. The use of preconcentration solution more alkaline than pH 9 often resulted in the precipitation of lead(II) hydroxide. The effect of preconcentration potential on the SW response for lead(II) is shown in Fig. 5B. As can be seen, between −0.3 and −0.6 V vs. Ag/AgCl, the peak current increases as the potential of the electrode becomes more negative. This behaviour is explained by the fact that lead(II) bears a positive charge, as a result the adsorption of lead(II) is favoured at more negative potentials.

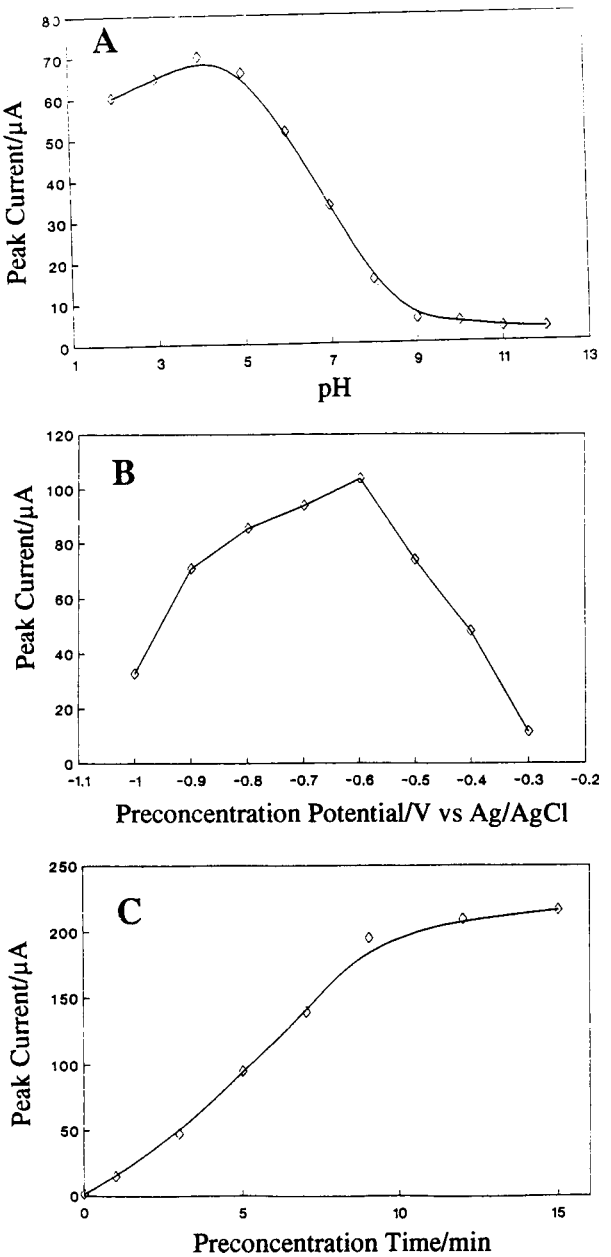


Fig. 5. Effect of pH (A), preconcentration potential (B), and preconcentration time (C) on the peak current for 0.1 mg l^{-1} lead(II) obtained at the NC(Bpy)MFE. Conditions for the SWSV measurement: (A) $t_p = 5$ min; $p_p = -0.9$ V. (B) $t_p = 5$ min in pH 4 phosphate buffer solution. (C) $p_p = -0.6$ V in pH 4 phosphate buffer solution.

However, the peak current starts to drop as the potential is more negative than −0.6 V. A preconcentration potential of −0.6 V was therefore chosen in all subsequent work. Note that preconcentration of lead(II) takes place at −0.6 V, which is on the negative side of the lead reduction peak; therefore, part of the preconcentration mechanism of lead is based on electrolysis. Nevertheless, the contribution from the chelating process between Bpy and lead(II) was believed to be the major factor as indicated by many evidences in this study. Further study is needed to quantitatively determine the mechanism of the system. The dependence of the peak current on the preconcentration time for lead(II) is shown in Fig. 5C. The rate of lead(II) uptake is dependent on concentration. For 0.1 mg l^{-1} of lead(II), the peak current increases as the preconcentration time increases and starts to level off around 9 min. For a lower concentration of lead(II), it takes longer for the peak current to level off (not shown). This phenomenon is as expected and further confirms the chelating process between Bpy and lead(II) in the CME. Apparently, since electrodes were prepared at 2800 rev. min^{-1} spin-coating rate with 4 μl of the optimum coating solution, which is composed of 1 ml of Nafion + 1 ml of the Bpy (3 wt% in ethanol), the amount of Bpy on the CME was actually fixed. Therefore, in order to chelate with all the Bpy sites available, a longer time is needed for lower concentrations of lead(II).

The SW parameters that were investigated were the frequency, the pulse height and the pulse increment. These parameters are interrelated and have a combined effect on the response. The response for lead(II) increases with SW frequency, but at frequencies higher than 160 Hz sloping background current renders the measurement difficult. Increase in the pulse height causes an increase in the lead(II) peak up to 50 mV and the peak potential shifts to the positive direction with increasing frequency. The scan increment together with the frequency defines an effective scan rate; hence, increase of either the frequency or the scan increment results in an increase in the effective scan rate. Overall, the best signal to background current characteristics can be obtained by the following instrumental settings: modulation amplitude, 50 mV; modulation frequency, 160 Hz; effective scan rate 640 mV/s.

3.3. Analytical characteristics of CME

The optimum condition for lead(II) accumulation was observed when the sample was slightly acidic (pH 3–5). In fact, this behaviour proved to be quite useful, as it provided a chemical method for removal of the deposit. A 5-s immersion into 1 M hydrochloric acid was generally sufficient to remove the preceding lead(II) deposition completely as was observed after visual and electrochemical inspection. For 10 successive preconcentration/determination/renewal cycles, the SW voltammetric response could be reproduced with 5% relative standard deviation.

The NC(Bpy)MFE can be used for the determination of lead(II) at the μg l^{-1} level even in the presence of oxygen. Calibration data were obtained in phosphate buffer spiked with lead(II) following 5-min preconcentration. Fresh sample solutions were used for each individual concentration and preconcentration time, and at least six determinations for each data point (R.S.D. = 1–5%). The calibration plots thus obtained show a good linear behaviour with slope (μA/μg l^{-1}), intercept (mA), and correlation coefficients of 0.8112, 0.0269 and 0.9989, respectively. The linear range for 5-min preconcentration was investigated and the results indicated excellent linearity from 1 to 100 μg l^{-1} lead(II). The sensitivity started to decrease when the concentration of lead(II) was higher than 100 μg l^{-1}.

The detection limit is defined as the concentration of the analyte resulting in a signal three times the standard deviation of the blank, as recommended by IUPAC [16]. For a preconcentration time of 5 min, the detection limit was 0.1 μg l^{-1} for lead(II). Of course, with shorter preconcentration times, higher concentrations were required for detection. On the contrary, an even lower detection limit could be achieved for lead(II) provided that the preconcentration time is longer than 5 min.

3.4. Interferences

Perhaps the main advantage of the NC(Bpy) MFE for lead(II) determination is its inherent selectivity. Any electroactive species that forms a precipitate with Bpy might present an obstacle in the determination of lead(II). However, the chemical conditions in effect during the preconcentration step might easily be optimized to favour the formation of the lead(II) complex over that of most of the other metals. Therefore, the number of species interacting in this manner is rather limited. For 0.02 mg l^{-1} lead(II), the results show that over a 1000-fold excess concentration of copper(II), aluminum(III), cobalt(II), calcium(II), magnesium(II), zinc(II), mercury(II), nickel(II), palladium(II) and titanium(II) did not influence the lead(II) response. Bismuth(III) and thallium(III) were found to interfere at a 100-fold excess, and cadmium(II) and tin(II) interfered at a 10-fold excess. Note that in the presence of only a 4-fold excess of thallium and tin, the lead(II) cannot be measured selectively by the conventional anodic stripping procedure due to severe peak overlap [4]. The lack of copper interference is of great significance, considering its major interference in the oxine-based adsorptive stripping scheme [5], and the presence of copper in numerous relevant samples. Fig. 6 compares adsorptive stripping measurements of lead(II), in the absence (a) and presence (b) of a large excess of copper, utilizing the NC(DMG)MFE (A) and the NC(Bpy)MFE (B). While the lead(II)-DMG peak is largely obscured by the copper(II)–DMG response, the large excess of copper has no effect upon the lead(II)–Bpy peak.

The interference effects caused by surface-active compounds in ASV by using MFE was well recognized [12]. Coating of the MFE with permselective membranes has been introduced as a means of circumventing the organic interferences in ASV [11,12]. The function of the membrane is to prevent the organic interferences from reaching the interface at which the deposition/stripping takes place. The effect of surface-active organic compounds on the SW stripping response for lead(II) at the NC(Bpy)MFE was examined. For comparison, the effect of the surfactants on both the NCMFE and the conventional MFE was also recorded. Triton X-100 was used to simulate the effect of a typical nonionic surfactant, and Fig. 7 shows how the 0.02 mg l^{-1} lead(II) strip-

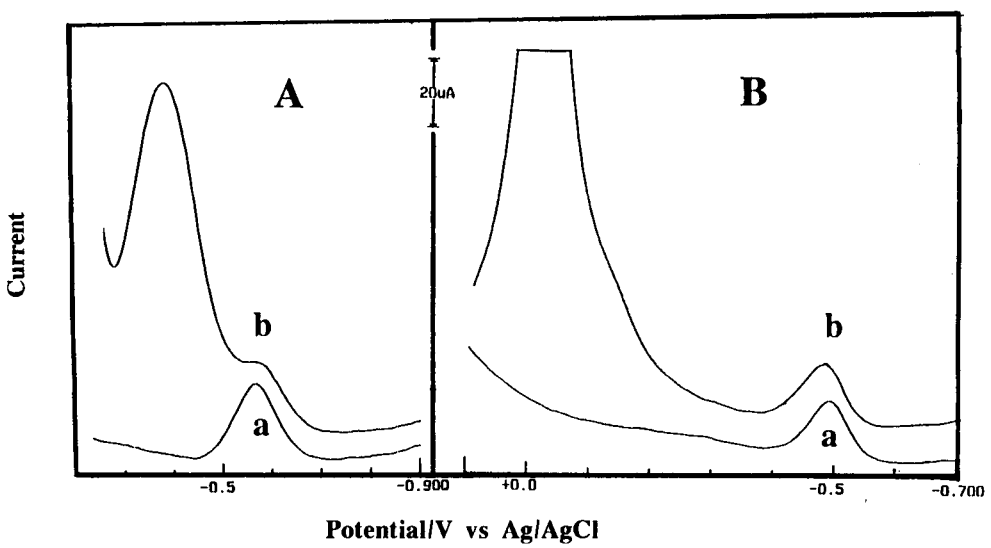


Fig. 6. Effect of copper(II) upon the response for lead(II) at the NC(DMG)MFE (A) and the NC(Bpy)MFE (B). (a) $[Pb^{2+}] = 0.1$ mg l^{-1} for (A) and 0.02 mg l^{-1} for (B); (b) same as (a) but after additions of 1 mg l^{-1} copper (II). Conditions: $t_p = 5$ min; $p_p = -0.6$ V; (A) in pH 9 buffer solution; (B) in pH 4 buffer solution.

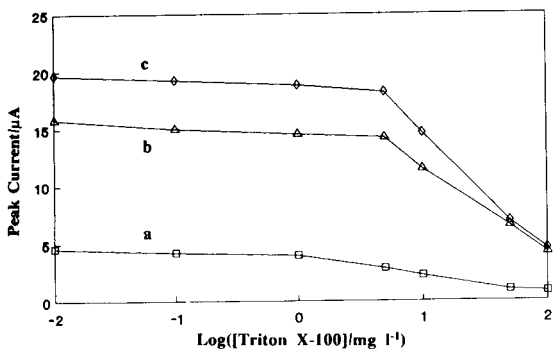


Fig. 7. Effect of the surfactant Triton X-100 at different concentrations on the SWSV response for 20 $\mu g\ l^{-1}$ lead(II) on the MFE (a), the NCMFE (b), and the NC(Bpy)MFE (c). Other conditions as in Fig. 6.

ping peak current is affected for different concentrations of Triton X-100. As can be seen, the detection can tolerate the presence of Triton X-100 up to about 5 mg l^{-1} for both the NCMFE and the NC(Bpy)MFE. Compared to similar experiments carried out with MFE, the tolerance was improved by a factor of 5. The peak current for the NC(Bpy)MFE is higher than the NCMFE is once again a good indication of the proper chelating process between Bpy and lead(II). The peak current for the NC(Bpy)MFE is much higher than the MFE, which is a good indication of the proper ion-exchange process of Nafion.

3.5. Practical applications

The analytical utility of the NC(Bpy)MFE was assessed by applying it to the determination of lead(II) in drinking water and electroplating waste solution. The results are summarized in Table 1, indicating good accuracy for all the samples. Real values of lead(II) content in drinking water and electroplating waste solution were calculated from the dilution factors of each sample during preparation.

Table 1
Determination of lead(II) in drinking water and electroplating waste solution

	Drinking water	Electroplating waste solution
Detected value, [a] original ($\mu g\ l^{-1}$)	5.89 ± 0.42	19.96 ± 1.44
Spike ($\mu g\ l^{-1}$)	10.0	10.0
Detected value, [a] after spike ($\mu g\ l^{-1}$)	15.93 ± 0.48	30.16 ± 2.38
Recovery (%)	100.4	102
Real value [a] ($\mu g\ l^{-1}$)	29.45 ± 2.1	199.6 ± 14.4

[a] $n = 6$.

4. Conclusions

In this article it has been demonstrated that SWSV in conjunction with the NC(Bpy)MFEs is a potentially interesting development in the field of anodic stripping analysis. The CME offers considerably higher resistance to organic interferences than the MFE and the mechanical stability of the mercury film is also largely improved. The NC (Bpy)MFE's possession of good sensitivity, selectivity and the capability to be easily regenerated with exposure to acid may make the technique very attractive in the trace determination of lead(II) in continuous-monitoring applications. Future work is being carried out in order to assess the applicability of the technique in flowing streams.

Acknowledgments

The authors gratefully acknowledge the financial support by the National Science Council of the Republic of China under grant NSC 82-0208-M005-053.

References

[1] R.L. Boeckx, Anal. Chem., 58 (1986) 274A.
[2] J. Wang, Stripping Analysis: Principles, Instrumentation and Application, VCH, Deerfield Beach, FL, 1985.
[3] J. Boone, T. Hearn and S. Lewis, Clin. Chem., 25 (1979) 389.
[4] J. Wang, J. Lu and C. Yarnitzky, Anal. Chim. Acta, 280 (1993) 61.
[5] C.M.G. van den Berg, Anal. Chim. Acta, 215 (1986) 111.
[6] R.W. Murray, in A.J. Bard (Ed.), Electroanalytical Chemistry, Vol. 13, Marcel Dekker, New York, 1984.
[7] R.W. Murray, A.G. Ewing and R.A. Durst, Anal. Chem., 59 (1987) 379A.

[8] S. Dong and Y. Wang, Electroanalysis, 1 (1989) 99.
[9] J. Wang, in A.J. Bard (Ed.), Electroanalytical Chemistry, Vol. 16, Marcel Dekker, New York, 1989.
[10] M.N. Szentirmay and C.R. Martin, Anal. Chem., 56 (1984) 1898.
[11] C.R. Martin, I. Rubinstein and A.J. Bard, J. Am. Chem. Soc., 104 (1982) 4817.
[12] B. Hoyer, T.M. Florence and G.E. Batley, Anal. Chem., 59 (1987) 1608
[13] J.-M. Zen and M.-L. Lee, Anal. Chem., 65 (1993) 3238.
[14] M. Lovric and M. Branica, J. Electroanal. Chem., 226 (1987) 239.
[15] M. Lovric and S. Komorsky-Lovric, J. Electroanal. Chem., 248 (1988) 239.
[16] J.N. Miller, Analyst, 116 (1991) 3.

ELSEVIER

Analytica Chimica Acta 296 (1994) 87–97

ANALYTICA CHIMICA ACTA

Matrix isopotential synchronous fluorescence Direct determination of gentisic acid in urine

José Antonio Murillo Pulgarín *, Aurelia Alañón Molina

Department of Analytical Chemistry and Foods Technology, University Castilla-La Mancha, 13071 Ciudad Real, Spain

Received 4 January 1994; revised manuscript received 20 April 1994

Abstract

A new fluorimetric technique is described for the determination of compounds in samples with unknown background fluorescence. The technique consists of performing synchronous scans through a trajectory which joins points of equal intensity of a fluorescent matrix tridimensional spectrum. The method is called "matrix isopotential synchronous fluorescence". This technique can be improved by application of derivatives. In order to demonstrate the validity, applicability and simplicity of the method for determining analytes in complex fluorescent matrices, it is applied to the analysis of gentisic acid in urine. Better sensitivity and reproducibility are achieved in these matrices than with the fluorimetric methods described in the literature, without the need for tedious pre-separation.

Keywords: Fluorimetry; Matrix isopotential synchronous fluorescence; Gentisic acid

1. Introduction

In luminescence studies, high sensitivity and selectivity are generally expected. However, problems of selectivity and ultimately of sensitivity can occur in multicomponent analysis because of the overlap of the broad-band spectra observed. As a consequence, a number of procedures has been proposed to increase selectivity in quantitative luminescence spectrometry.

Conventional fluorescence involves generating an emission spectrum by scanning the emission wavelength, λ_{em}, when the sample is irradiated at a single excitation wavelength, λ_{ex}. Similarly, an excitation spectrum is obtained by scanning the excitation wavelength while recording the emission signal at a single wavelength. Another possibility is to vary λ_{ex} and λ_{em} simultaneously. This technique has several variants, depending on the scan-rates of the two monochromators.

If the scan-rate is constant for both monochromators, and therefore a constant wavelength interval, $\Delta\lambda$, is kept between λ_{em} and λ_{ex}, the technique is known as conventional synchronous fluorescence spectroscopy. This technique is very simple and is the most frequently used of all synchronous modes. It was introduced by Lloyd [1], who applied it to a mixture of C_{20} polynuclear hydrocarbons from the pyrolysis and combustion of organic materials.

* Corresponding author.

SSDI 0003-2670(94)00197-T

When the fluorescence of a compound is perturbed by Rayleigh scattering, the synchronous technique allows an improvement in the measurement over conventional techniques, if appropriate $\Delta\lambda$ values are selected. The main difficulty encountered in the application of the synchronous scan technique is that the best wavelength interval value must be known beforehand for optimum results and, in some multicomponent systems, several different wavelength interval values might be necessary to achieve complete identification. The best wavelength interval values can readily be determined from a tridimensional plot. For this reason, it is useful to generate the total luminescence spectrum beforehand, with the aim to fully characterise the fluorescent properties of the samples.

If the excitation and emission wavelengths are varied simultaneously but at different rates, the technique is known as variable angle synchronous fluorescence spectroscopy. These different rates allow the construction of planes at angles between $+90°$ and $-90°$ to the excitation X-axis throughout the whole spectrum.

This technique is, logically, more selective, and provides considerable flexibility. It was proposed by Kubic et al. [2] and applied to the forensic examination of automobile engine lubricants. The fluorescence behaviour of 61 automobile lubricants and 45 engine oils has been studied by this technique [3,4]. Recently it has been applied to the simultaneous determination of vitamin B_2 and vitamin B_6 [5], polynuclear aromatic hydrocarbons [6] and fluorescein derivatives [7].

The angle of the scan trajectory can also be varied continuously through the emission–excitation matrix to describe any desired path. This technique, known as non-linear variable angle synchronous fluorescence spectroscopy has been applied to two pharmaceutical determinations: chlorpromazine in the presence of its chief degradation product, chlorpromazine sulphoxide, and oxytetracycline in the presence of the additives vitamin c, thiamine, nicotinamide and riboflavin [8], etc.

A special example of the technique just described is that of constant energy synchronous fluorescence, proposed by Inman and Wineforder [9], in which the energy between the measured excitation and emission radiation is kept constant. This technique reduces solvent interference by Raman scattering similarly to the way in which constant wavelength synchronous fluorimetry reduces Rayleigh scattering interference.

This technique has been little used up to now, mainly due to the lack of commercially available instruments. It has been applied to the spectral resolution of mixtures of polyaromatic hydrocarbon isomers and polyaromatic hydrocarbon alkyl homologues [9,10], extending it to low temperature conditions by reducing the sample temperature to that of liquid nitrogen or helium [11,12].

The combination of synchronous and derivative fluorimetry was first suggested by John and Soutar [13] to enhance minor spectral features and to allow more reliable identification of spectra. Although direct synchronous spectra are often sufficiently resolved for analytical purposes, the second derivative of the synchronous spectrum could be helpful in some cases in differentiating closely spaced bands. The main advantages of this technique are selectivity, sensitivity, rapidity, simplicity and low cost.

This technique has hardly been used, but recently in this laboratory the simultaneous determination of methicillin and nafcillin was realized by first derivative fluorescence and synchronous fluorescence spectroscopy [14].

In this study, we propose a new fluorimetric technique which uses a cut made in the total fluorescence spectrum through a trajectory joining points of equal intensity (isopotential trajectory). This trajectory is obtained by means of a program developed in BASIC [15]. We have called the technique "matrix isopotential synchronous fluorescence" (MISF).

Gentisic acid (2,5-dihydroxybenzoic acid or 5-hydroxysalicylic acid) is a minor metabolite of salicylic acid and aspirin and has been employed in the palliative treatment of rheumatoid arthritis [16]. Therefore the determination of gentisic acid in biological fluids is of interest both for emergency and for routine testing, and in pharmacokinetic investigations. Human urine is composed of numerous organic substances, only some of them are fluorescent. These compounds, however, are

sufficient to provide a high background fluorescence and to interfere with the direct determination of gentisic acid in urine. In this case a prior extraction procedure is necessary [17,18]. To demonstrate the usefulness of the proposed technique gentisic acid in urine was determined.

2. Theory

One of the major disadvantages of fluorimetric methods is that it is not easy to remove background fluorescence. This means that direct determinations of fluorescent compounds, for example, in such important matrices as biological fluids (urine, serum, plasma, etc.) have to be approached by means of fluorescence derivatization or by prior extraction procedures, with the disadvantages involved.

If the sample in which we wish to perform a spectrofluorimetric determination has an almost invariable composition, even though its fluorescence intensity may vary, it is possible to maintain a constant background signal if a cut is made in the matrix total luminescence spectrum following one of the trajectories which joins points of equal fluorescence intensity (isopotential trajectory) from an initial point to final excitation and emission wavelengths.

Fig. 1a shows a total luminescence spectrum of a theoretical matrix, assuming the excitation and emission spectra to be "Gaussian" curves and an isopotential trajectory "A" that passes through points $\lambda_{ex} = 320$ nm and $\lambda_{em} = 489$ nm. Let us assume a fluorescent analyte with a total luminescence spectrum like the one shown in Fig. 1b, with an excitation maximum at 320 nm and an emission maximum at 489 nm. If this analyte is to be found in the presence of the previous matrix, the mixture will present a spectrum which will be the sum of the two of them (Fig. 1c). As can be seen in Fig. 1d, the spectrum of the matrix following this trajectory is a horizontal line of constant intensity.

By means of the application of the preceding trajectory (Fig. 1a) to the total luminescence spectrum of a analyte in the presence of the fluorescent matrix (Fig. 1c), the same spectrum is obtained (Fig. 1f) as that of the analyte in isolation (Fig. 1b and e), increased by a constant equivalent to the matrix intensity in that isopotential trajectory (Fig. 1d).

The intensity of the sample will be (Fig. 1c):

$$I_{s_{(\lambda_{ex},\lambda_{em})}} = I_{a_{(\lambda_{ex},\lambda_{em})}} + I_{b_{(\lambda_{ex},\lambda_{em})}}$$

where $I_{s_{(\lambda_{ex},\lambda_{em})}}$ is the fluorescence intensity of a sample at an excitation wavelength of λ_{ex} and an emission wavelength λ_{em}; $I_{a_{(\lambda_{ex},\lambda_{em})}}$ is the fluorescence intensity of analyte at an excitation wavelength of λ_{ex} and an emission wavelength λ_{em}; and $I_{b_{(\lambda_{ex},\lambda_{em})}}$ is the background fluorescence intensity at an excitation wavelength of λ_{ex} and an emission wavelength λ_{em}.

By definition, if the trajectory is isopotential $I_{b_{(\lambda_{ex},\lambda_{em})}}$ is constant (k) during the whole scan of wavelengths through this trajectory (Fig. 1d):

$$I_{s_{(\lambda_{ex},\lambda_{em})}} = I_{a_{(\lambda_{ex},\lambda_{em})}} + k \qquad \text{(Fig. 1f)}$$

Although the quantitative composition may vary in different determinations, the constant may be ascertained as the value towards which at least one of the extremes of the spectrum of the analyte to be determined tends asymptotically, and in consequence, it is possible to discover the fluorescence intensity corresponding only to the analyte to be determined.

$$\lim_{\lambda_{ex} \to \lambda_{exL}^{-}} I_{s_{(\lambda_{ex},\lambda_{em})}} = \lim_{\lambda_{ex} \to \lambda_{exL}^{-}} I_{a_{(\lambda_{ex},\lambda_{em})}} + k$$

or

$$\lim_{\lambda_{ex} \to \lambda_{exI}^{+}} I_{s_{(\lambda_{ex},\lambda_{em})}} = \lim_{\lambda_{ex} \to \lambda_{exI}^{+}} I_{a_{(\lambda_{ex},\lambda_{em})}} + k$$

where $\lambda_{ex\,L}$ = last excitation wavelength; $\lambda_{ex\,I}$ = initial excitation wavelength; $\lim_{\lambda_{ex} \to \lambda_{exL}^{-}} I_{s_{(\lambda_{ex},\lambda_{em})}}$ = left limit of the MISF spectrum of the sample (that is, the value to which the right extreme of the spectrum tends; and $\lim_{\lambda_{ex} \to \lambda_{exI}^{+}} I_{s_{(\lambda_{ex},\lambda_{em})}}$ = right limit of the MISF spectrum of the sample (that is, the value to which the left extreme of the spectrum tends. If

$$\lim_{\lambda_{ex} \to \lambda_{exL}^{-}} I_{a_{(\lambda_{ex},\lambda_{em})}} \to 0 \qquad \text{(Fig. 1e)}$$

or

$$\lim_{\lambda_{ex} \to \lambda_{exI}^{+}} I_{a_{(\lambda_{ex},\lambda_{em})}} \to 0 \qquad \text{(Fig. 1e)}$$

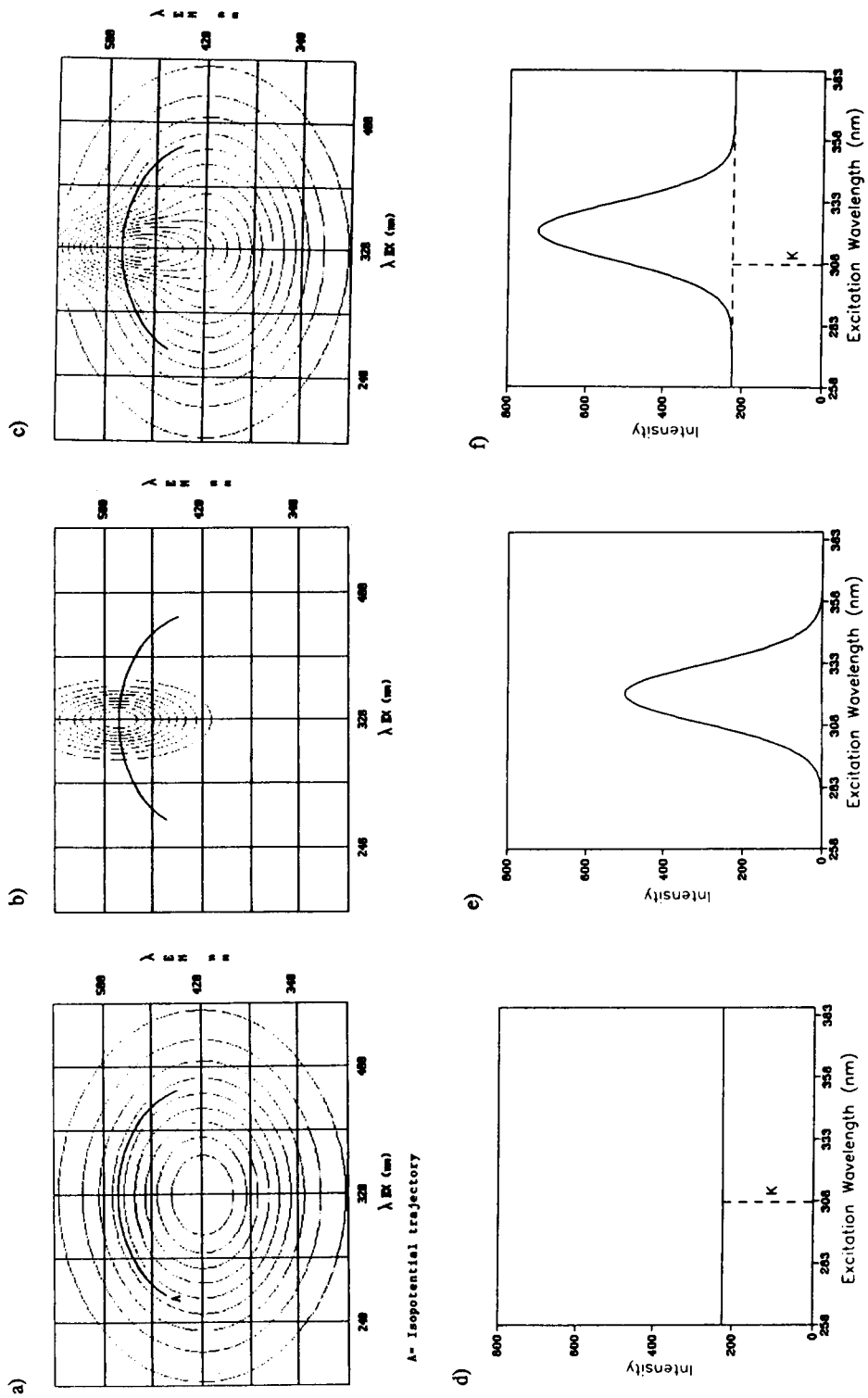


Fig. 1. Total luminescence spectrum and an isopotential trajectory of a theoretical matrix (a), a theoretical analyte (b) and the sum of both (c). MISF spectrum according to the preceding trajectory of matrix (d), analyte (e) and the sum of both (f).

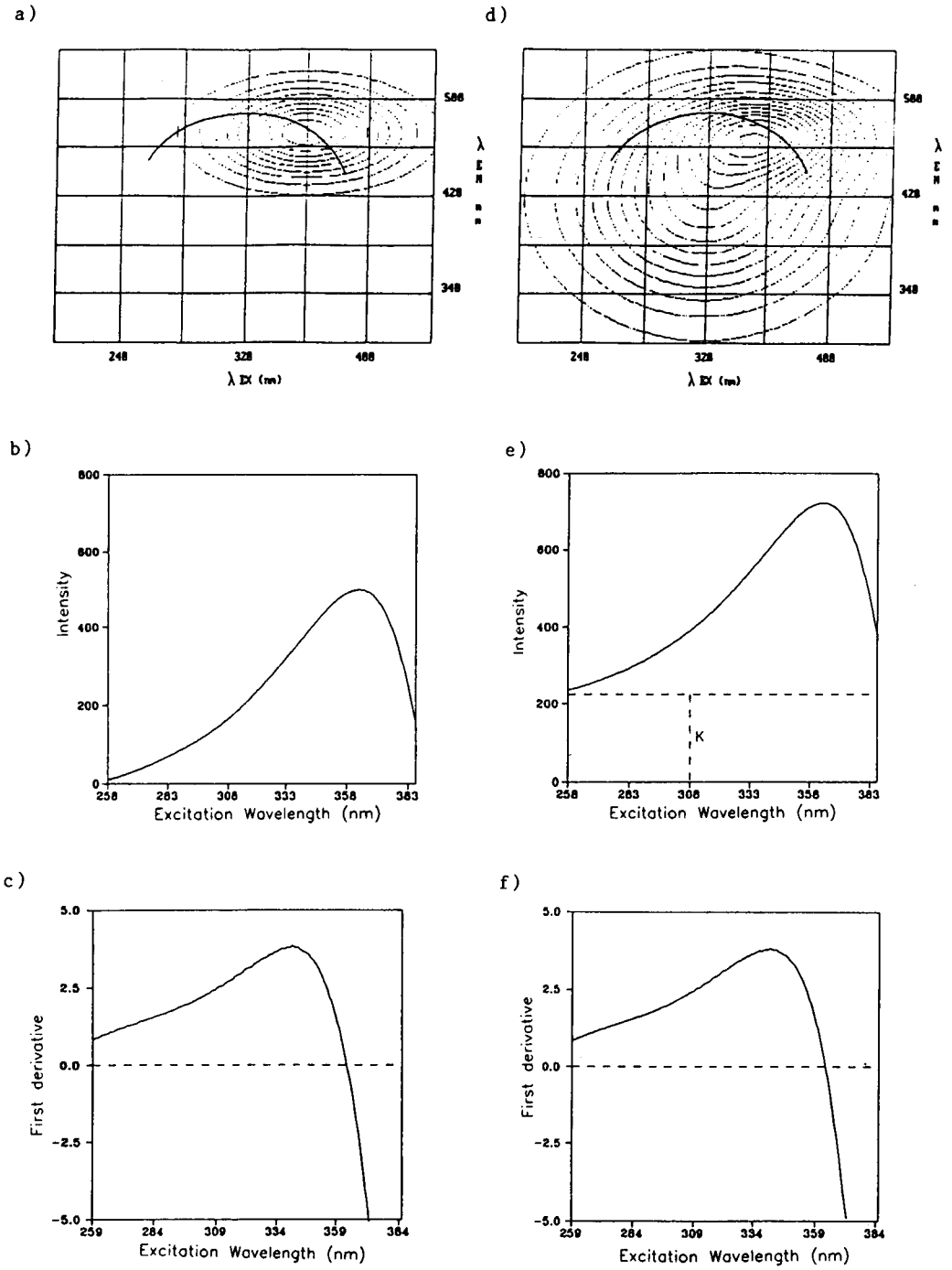


Fig. 2. Total luminescence spectrum and an isopotential trajectory of a theoretical analyte (a) and a mixture of theoretical analyte and matrix (d). MISF spectrum according to the preceding trajectory of analyte (b) and of mixture (e). First derivative MISF spectra of analyte (c) and mixture (f).

then

$$k = \lim_{\lambda_{ex} \to \lambda_{exL}^{-}} I_{s_{(\lambda_{ex},\lambda_{em})}} \quad \text{(Fig. 1f)}$$

or

$$k = \lim_{\lambda_{ex} \to \lambda_{exI}^{+}} I_{s_{(\lambda_{ex},\lambda_{em})}} \quad \text{(Fig. 1f)}$$

If this is not possible because the spectrum of the analyte in isolation according to this trajectory does not adopt the value zero at the extremes (Fig. 2a and b) as in the previous case, the problem of not knowing the matrix constant value can be avoided by applying the first derivative (Fig. 2c), that is to say:

$$\frac{dI_{s_{(\lambda_{ex},\lambda_{em})}}}{d\lambda_{ex}} = \frac{dI_{a_{(\lambda_{ex},\lambda_{em})}}}{d\lambda_{ex}} + \frac{dI_{b_{(\lambda_{ex},\lambda_{em})}}}{d\lambda_{ex}}$$

as the trajectory is isopotential (Fig. 1d), then

$$\frac{dI_{b_{(\lambda_{ex},\lambda_{em})}}}{d\lambda_{ex}} = 0$$

and

$$\frac{dI_{s_{(\lambda_{ex},\lambda_{em})}}}{d\lambda_{ex}} = \frac{dI_{a_{(\lambda_{ex},\lambda_{em})}}}{d\lambda_{ex}} \quad \text{(Fig. 2c and f)}$$

For this reason the value of the first derivative of the spectrum of the sample coincides exactly with the matrix isopotential synchronous fluorescence (MISF) spectrum of the analyte.

It is always possible to find the matrix trajectory which passes through the maximum fluorescence excitation and emission wavelengths of the component to be determined, and so the same sensitivity is achieved as in a direct determination in the absence of background fluorescence.

The only condition which is necessary and sufficient for the application of this new technique is that the fluorescence of the matrix and of the analyte should be additive, as is the case with all techniques applied to samples containing more than one fluorescent compound. Even if this condition is not fulfilled, this technique could be used by comparing the fluorescence values with a calibration graph prepared on the matrix itself instead of a calibration graph constructed with the analyte in isolation.

This new technique was applied to the determination of gentisic acid in urine.

3. Experimental

3.1. Apparatus

All fluorimetric measurements were performed on a Perkin Elmer LS 50 equipped with a xenon lamp, connected to an Ataio S 3000 ST 386 computer fitted with Perkin-Elmer FL Data Manager software (design for handling fluorescence data on a personal computer) and Epson FX-850 printer.

Thermostatic equipment and a Crison Model 2001 pH-meter with a glass-saturated calomel combination electrode and a Selecta centrifuge (Model Mixtaxel) were also used.

3.2. Software

A program was developed which enables us to obtain the values of λ_{ex} and λ_{em} for any constant fluorescence intensity value from a tridimensional spectrum. As the values obtained for a particular curve are not equidistant, the Lagrange interpolation method is applied to all points, which are placed in an order according to excitation wavelengths at 0.5 nm intervals. Once the trajectory has been defined, the spectrum is obtained by means of the Ftotal program [15], so that the spectra obtained display the same format as those performed directly from the Perkin-Elmer LS50 spectrofluorimeter.

3.3. Reagents

All experiments were performed with analytical reagent grade chemicals, pure solvents and Milli-Q water. The standard gentisic acid was obtained from Sigma. A stock solution of gentisic acid is prepared in a 500-ml volumetric flask by dissolving 25 mg in water. This stock solution is used to prepare standard solutions by suitable dilutions.

A 0.05 M buffer solution of pH 6.0 was prepared by mixing adequate amounts of sodium dihydrogenphosphate with sodium hydroxide.

Urine samples were obtained from fasting and healthy people in the morning.

3.4. Procedure

Sample preparation.

Urine was centrifuged daily for 10 min at 2500 *g*. 4 ml of clear supernatant solution were transferred into a 100-ml flask and diluted to the mark with water. This solution is stored and maintained below 5°C.

For the preparation of a calibration graph, a suitable aliquot containing 0.50–6.30 μg of gentisic acid was transferred into a 25-ml volumetric flask. 5.0 ml of phosphate buffer solution (pH 6.0) and 2.0 ml of solution of urine were added to every volumetric flask. The flasks were diluted to the mark with water.

Several solutions prepared in the same way were also obtained, with different urine samples and in the absence of gentisic acid.

Similarly, another calibration graph was obtained for gentisic acid in aqueous solution.

For every sample, 61 excitation spectra of 240 nm width were performed varying the emission wavelength by 4 nm steps. The total luminescence spectra were obtained by means of Ftotal [15]. Then the adequate trajectory and matrix isopotential synchronous spectra were obtained for gentisic acid by means of the same program. The determination of gentisic acid is possible by measuring the fluorescence intensity with regard to the value of the final extreme.

4. Results and discussion

4.1. Determination of gentisic acid in urine

Matrix isopotential synchronous fluorescence was applied to the analysis of gentisic acid in urine, a very fluorescent matrix, under the optimum chemical conditions established by Salinas et al. [17] and Konstantianos and Ioannou [18]

As can be observed in Fig. 3, the fluorescence

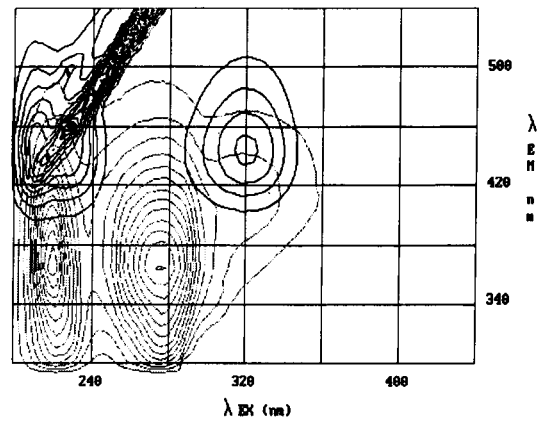


Fig. 3. Total fluorescence spectra of urine. Broken line, dilution 1:312; Solid line, 0.100 μg/ml of gentisic acid.

maxima characteristic of gentisic acid are located in the UV region where the urine shows two broad peaks which prevent determination of this compound without prior separation.

In order to determine gentisic acid in urine, first we checked that the qualitative composition of the fluorescent metabolites of the urine was practically constant. In order to do this, the fluorescence spectra of twenty samples of urine from healthy persons of both sexes and different diets (25–35 years of age) were compared.

Urine was centrifuged daily for 10 min at 2500 *g*. 4 ml of clear supernatant solution were transferred into a 100-ml flask and diluted to the mark with water. Then, 2.0 ml of preceding solution was transferred into a 25-ml volumetric flask and 5.0 ml of phosphate buffer solution (pH 6.0) were added. The solution was diluted to the mark with water.

The total fluorescence spectrum was obtained in the form of a contour map in order to conduct the comparison. From the preceding study it was concluded that the different samples of urine display the same type of fluorescence, with hardly any variation to be observed in the form of the spectrum and in the location of the fluorescence maxima, although it is possible to observe variations in their intensity.

The spectrum corresponding to the arithmetic mean of the twenty total fluorescence spectra of the different urine samples was obtained by means

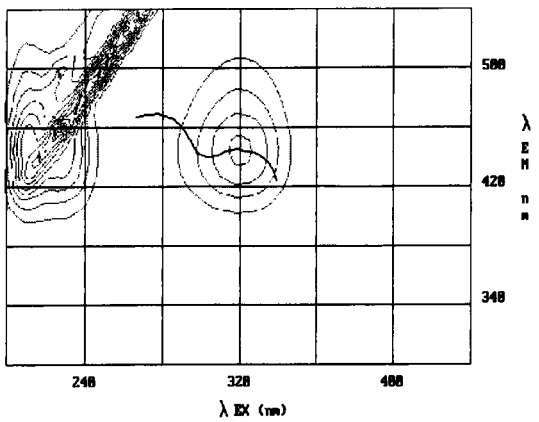


Fig. 4. Total fluorescence spectrum of gentisic acid (broken line) and isopotential trajectory of urine selected (solid line).

of the Ftotal program. In order to determine the gentisic acid we calculated the isopotential trajectory in the averaged urine spectrum, which passes through the excitation and emission maxima of the gentisic acid, as is shown in Fig. 4.

Urine sample solutions containing gentisic acid gave signals smaller than for aqueous standard solutions, owing to some type of binding with precipitated proteins.

Total luminescence spectra of gentisic acid were obtained in different urine samples in order to construct five calibration graphs and to conduct recovery experiments.

The MISF technique was applied to all samples and fluorescence intensity was measured with regard to the value of the final right extreme of the spectrum, where the fluorescence intensity of gentisic acid is zero.

In the same way we obtained the total luminescence spectra of gentisic acid in aqueous solution at the same concentrations, to which we applied the above mentioned trajectory (Fig. 4) in order to obtain their MISF spectra, and we also measured their intensity, like for the urine samples. When the results obtained in urine were compared with those obtained in aqueous solution, a recovery percentage of $70.2 \pm 0.5\%$ (intensity of gentisic acid in urine/intensity of gentisic acid in aqueous solution) was obtained, a value very similar to that obtained by Konstantianos et al. [18] when determining it in serum.

As can be observed in Fig. 5a, which shows the spectra of gentisic acid obtained in urine, the extreme value is constant and, in this case, corresponds to a value of 71.045 fluorescence units. Fig. 6a shows the five calibration graphs, and it is possible to observe high concordance among them, as is revealed in Table 1, which shows the results of the statistical analysis.

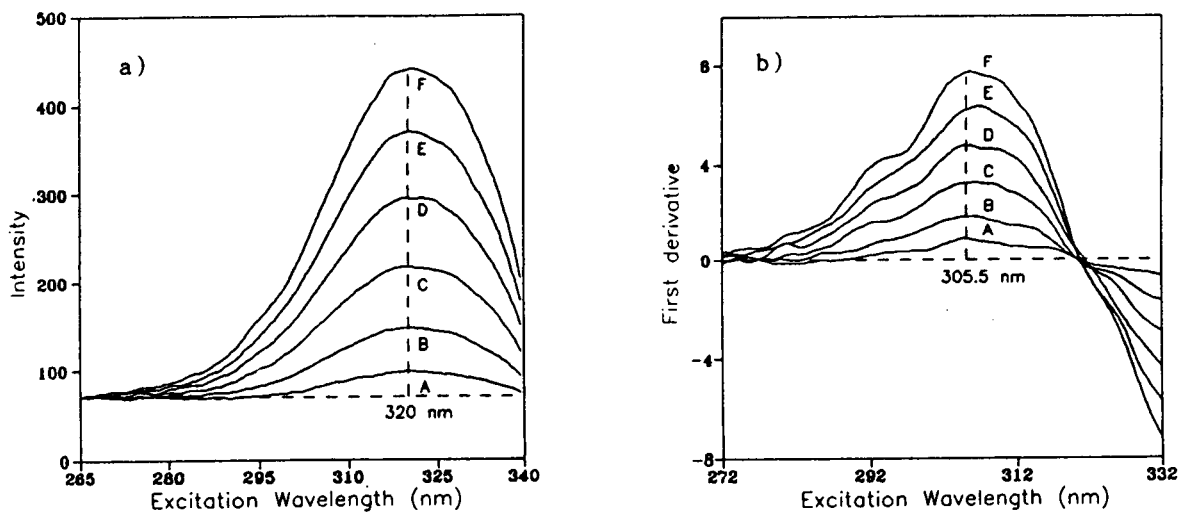


Fig. 5. (a) Set of MISF spectra of gentisic acid in urine. Gentisic acid concentrations: (A) 0.020; (B) 0.052; (C) 0.100; (D) 0.152; (E) 0.200 and (F) 0.252 μg/ml. (b) Their first derivative spectra.

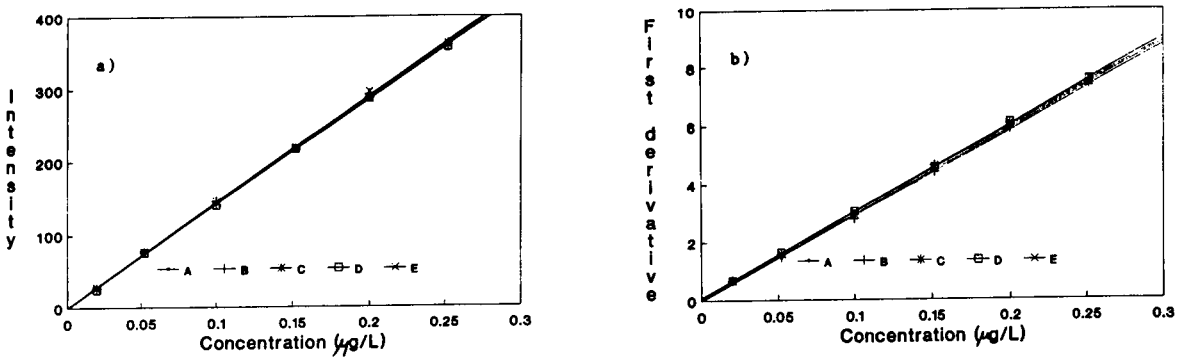


Fig. 6. Calibration graphs of gentisic acid in different urine samples obtained by measuring the fluorescence intensity with regard to the value of the final extreme of MISF spectra (a), and by measuring the first derivative spectra (b).

The Student *t*-test [19] was applied to all calibration graphs to study the significance of the intercept on the *Y*-axis, resulting in a negligible intercept for all the calibration graphs (Table 1). This shows that at the right extreme wavelength used by us the fluorescence intensity of gentisic acid tends towards zero.

The analysis of variance method [20] was also applied to check that the calibration graphs were independent of the urine sample used, and with satisfactory results.

In order to check the validity of the theoretical suppositions, the first derivative according to the Savitsky–Golay algorithm [21,22] was applied to the matrix isopotential synchronous spectra.

The number of points through which the derivative is obtained was optimized, with the conclusion that derivative spectra with a suitable signal-to-noise ratio are obtained with 15 points.

Fig. 5b shows the spectra derived from the calibration of gentisic acid in urine. It can be observed that they display a maximum at 305.5 nm, where maximum sensitivity is achieved. The first derivative signals of all the calibration graphs are displayed in Fig. 6b, which shows that this signal is independent of the urine sample used (background signal) since the graphs are coincident as can be seen from the application of the analysis of variance method.

In all cases the Student *t*-test [19] was applied,

Table 1
Statistical data of determination of gentisic acid in urine

Method used	Correlation coefficient	Intercept on *Y*-axis	Slope	Standard deviation		Experimental *t* value	New slope
				Intercept	Slope		
1 [a]	0.9999	−0.5	1433	1.5	10	0.334	1430
	0.9999	0.5	1442	1.9	12	0.228	1444
	0.9998	0.2	1429	1.9	13	0.101	1430
	0.9998	−1.3	1430	1.9	13	0.653	1423
	0.9997	0.1	1449	2.7	17	0.009	1450
2 [b]	0.9997	−0.019	29.12	0.054	0.35	0.347	29.02
	0.9996	−0.018	29.47	0.065	0.42	0.276	29.37
	0.9994	0.102	29.14	0.077	0.51	1.320	29.70
	0.9999	0.076	29.70	0.038	0.25	2.010	30.13
	1.0000	0.049	29.75	0.020	0.13	2.411	30.03

[a] 1 = Determination of gentisic acid by matrix isopotential synchronous fluorescence.
[b] 2 = Determination of gentisic acid by first derivative matrix isopotential synchronous fluorescence. Theoretical *t* value: 2.571.

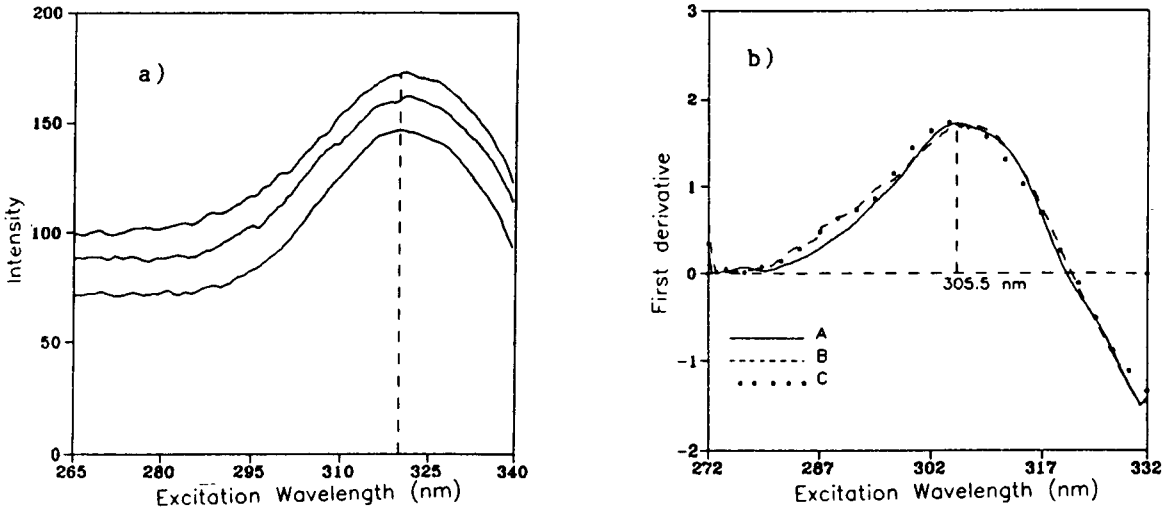


Fig. 7. (a) Set of MISF spectra of 0.052 μg/ml of gentisic acid in diferent urine samples. (b) Their first derivative spectra.

but negligible intercepts on the *Y*-axis were found.

In order to study the repeatability of the method, a series of ten solutions was prepared containing 0.100 μg/ml of gentisic acid in urine, with results of 8.9×10^{-4} and 1.3×10^{-3} μg/ml when matrix isopotential synchronous spectrofluorimetry and first derivative matrix isopotential synchronous spectrofluorimetry were used, respectively. These values are considerably better than those obtained by other procedures described in the literature [17,18].

Fig. 7 shows the effect of background fluorescence intensity (urine) on gentisic acid spectra and their derivatives. It can be seen that the spectra are identical in form although their intensities are different by a constant value, while derivation totally removes the background effect.

5. Conclusions

This study describes a new fluorimetric technique which we have called matrix isopotential synchronous fluorescence. The technique is particularly useful for removing fluorescence matrix background effects, and enables the determination of individual compounds in complex samples. In order to demonstrate its validity, it was applied to the determination of gentisic acid in urine with excellent reproducibility and a sensitivity equal to other modern fluorimetric methods [17,18], and without pre-separation step.

This technique can be combined with the application of derivatives to increase its field of application.

References

[1] J.B.F. Lloyd, Nature, 64 (1971) 231.
[2] T.A. Kubic, T. Kanabrocki and J. Dwyer, paper presented at the American Academy of Forensic Sciences 32nd Annual Meeting, 1980.
[3] T.A. Kubic, C.M. Lasher and J.J. Dwyer, Forensic Sci. Soc., 28 (1983) 186.
[4] T.A. Kubic and F.X.J. Sheehan, Forensic Sci. Soc., 28 (1983) 345.
[5] Y. Li, X. Huang, J. Xu and G. Chen, Fenxi Huaxue, 19 (1991) 538.
[6] M.T. Oms, R. Forteza, V. Cerda, F. García and A.L. Ramos, Int. J. Environ. Anal. Chem., 42 (1990) 1.
[7] M.T. Oms, R. Forteza, V. Cerda, S. Maspoch, J. Coello and M. Blanco, Anal. Chim. Acta, 233 (1990) 159.
[8] B.J. Clark, A.F. Fell, K.T. Milne, D.M.G. Pattie and H. Williams, Anal. Chim. Acta, 170 (1985) 35.
[9] E.L. Inman Jr. and J.D. Winefordner, Anal. Chem., 54 (1982) 2018.
[10] M.J. Kerkhoff, E.L. Inman Jr., E. Voigtman, L.P. Hart and J.D. Winefordner, Appl. Spectrosc., 38 (1984) 239.
[11] E.L. Inman Jr. and J.D. Winefordner, Anal. Chim. Acta, 141 (1982) 241.
[12] M.J. Kerkhoff, L.A. Files and J.D. Winefordner, Anal. Chem., 57 (1985) 1673.

[13] P. John and I. Soutar, Anal. Chem., 48 (1976) 520.
[14] J.A. Murillo Pulgarín and A. Alañón Molina, Talanta, 41 (1994) 21.
[15] J.A. Murillo Pulgarín and A. Alañón Molina, Comput. Chem., 17 (1993) 341.
[16] C.O. Wilson and O. Gisvold, Textbook of Organic, Medicinal and Pharmaceutical Chemistry, Lippincott, Philadelphia, PA, 4th edn., 1962, pp. 65–594.
[17] F. Salinas, A. Muñoz de la Peña, A.I. Durán-Merás and M.S. Durán, Analyst, 115 (1990) 1007.
[18] D.G. Konstantianos and P.C. Ioannou, Analyst, 117 (1992) 877.
[19] P.D. Lark, B.R. Craven and R.C.L. Bosworth, The Handling of Chemical Data, Pergamon, Oxford, 1968, pp. 136–205.
[20] G.E.P. Box, W.G. Hunter and J.S. Hunter, Statistics for Scientists, (1989) 169.
[21] A. Savitzky and M.J.E. Golay, Anal. Chem., 36 (1964) 1627.
[22] J. Steinier, Y. Termonia and J. Deltour, Anal. Chem., 44 (1972) 1906.

ELSEVIER

Analytica Chimica Acta 296 (1994) 99–105

ANALYTICA CHIMICA ACTA

Room temperature phosphorescence properties of 27 coumarin derivatives on filter paper

Long-Di Li *, Shi-Zhong Yang [1]

Department of Chemistry, Tsinghua University, Beijing 100084, China

Received 21st December 1993; revised manuscript received 20th April 1994

Abstract

Some coumarin derivatives have been used as fluorescence labelling agent. The room temperature phosphorescence (RTP) properties of 27 kinds of coumarin derivatives on filter paper have been studied. It is observed that most of them can give more intensive RTP signals when using 1.0 mol l^{-1} lead acetate as a heavy-atom perturber and that there is an obvious relation between the RTP characteristics and the substituting group in some of the derivatives. The fact that many of the coumarin derivatives can give RTP emission means that they can be expected to serve as RTP labelling agent.

Keywords: Phosphorimetry; Coumarin; Filter paper; Heavy atom perturbation; Substituent effect

1. Introduction

Phosphorescence and fluorescence belong to the field of luminescence spectroscopy. Although the sensitivity of phosphorimetry may be slightly lower than that of fluorimetry, phosphorimetry can get some information about the triplet state and provide a better selectivity, so it has developed into a useful analytical tool. However, the requirement for cryogenic conditions has limited the applications of the conventional low-temperature phosphorimetry (LTP).

Since its discovery in 1974 by Roth and coworkers [1,2], room temperature phosphorimetry (RTP) has attracted great interest and has become a practical technique for the detection of numerous organic compounds [3–5]. The solid substrate room temperature phosphorimetry (SS-RTP) is one of the important applications of RTP. Several solid substrates such as filter paper, silica gel, sodium acetate and inorganic compounds have been used; filter paper seems to be one of the best substrates. On the other hand, many coumarin derivatives are highly fluorescent and some of them have been used as laser dyes [6–8] or as fluorescence labelling agent for fatty acids and other organic acids, imides, mercapto compounds and phenols, etc. [9,10]. In 1989, Al-Kindy and Miller [11,12] suggested for the first

* Corresponding author.

[1] Present address: Dept. of Food Science, Xinjiang August 1st Agricultural College, Wurumqi 830052, Xinjiang, China.

SSDI 0003-2670(94)00196-S

time to use 6-coumarinsulphonyl chloride as labelling reagent for fluorescence and LTP in liquid chromatography. Following these studies, we reported its fluorescence behaviour in aqueous solution [13] and RTP properties on several solid substrates [14]. In this paper, the RTP spectral characteristics of 27 substituted coumarins on filter paper have been investigated. The result shows that most of them can give more intensive RTP signals when using $Pb(AC)_2$ as heavy-atom perturber and there is an obvious substituting group effect on the RTP properties. These results mean that some of the coumarin derivatives can be expected to serve not only as fluorescence labelling reagents, but also as an RTP one.

2. Experimental

The phosphorescence spectra were recorded on a Hitachi Model 850 fluorescence spectrophotometer equipped with a laboratory-made membrane substrate sample holder (Fig. 1).

The chemical structures and names of the 27 derivatives studied and the code used in this paper are shown in Table 1. The code 6, 8, 10, 12, 14, 15, 20–23, 25–27 and code 2–4, 7 and 11, compounds were purchased from Sigma and Aldrich, respectively. The code 24 compound was obtained from Fluka. The code 5 compound was synthesized by ourselves and other coumarin derivatives were purchased from Tianjin Institute of Chemical Reagents (China). These derivatives were used without further purification. Other reagents used for preparation of the solutions were analytical-reagents, grade and bidistilled water was used throughout the experiments.

5×10^{-4} mol l^{-1} solutions of the derivatives in ethanol were prepared by dissolution of the weighed amounts and appropriate dilution with ethanol. 1.0 mol l^{-1} $Pb(AC)_2$ solution, used as heavy-atom perturber, was prepared by dissolving weighed amounts $Pb(AC)_2 \cdot 3H_2O$ in acetic acid and appropriate dilution with water and the concentration of acetic acid was kept to 1.0 mol l^{-1} in the final solution.

The RTP properties of 27 coumarins on filter paper were investigated following the scheme given in Fig. 2. The spectral intensity measurements were repeated five times by using independent samples. The RTP lifetime was calculated according to the decay curve of the phosphorescence.

3. Results and discussion

3.1. Selection of the experimental conditions

Treatment of filter paper

The SS-RTP method using filter paper still suffers from two problems, namely the presence of background emission and sensitivity to moisture. Moreover, the heavy-atom perturbers used to increase the RTP of the analyte also enhance the paper background emission. In this paper, some chemical treatment methods (immersing for 12 h, then cleaning with water or used solvent) were used to reduce the background emission. The results show that the treatment method using 1.5% NaOH solution is the best one; it obtained a lower background value and the best precision (Table 2).

Selection of drying temperature and time

The sample drying operation was accomplished using a constant-temperature electrically

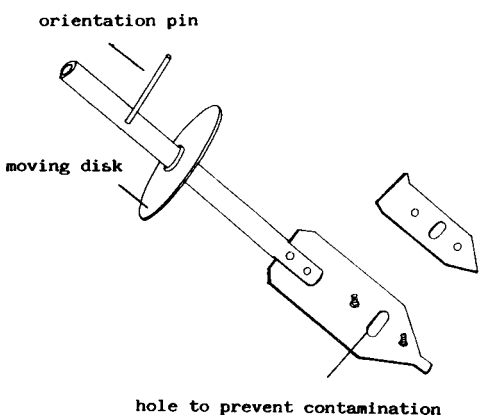


Fig. 1. The home-made membrane substrate sample holder. Set the sample holder in the Dewar holder of the phosphorescence accessory, then adjust the position of the orientation pin until the signal level becomes maximum and place at the same position each measurement.

heated plate. Because excessive heating in the drying process should be avoided, optimal drying temperature T, drying time t_1 and t_2, and delay time t_3 (Fig. 2) were identified by orthogonal factorial experiments. The results show that the effects of T, t_2, t_3 and t_2 are very significant, significant, significant and little significant, respectively, the appropriate experimental conditions being $T = 84$–89°C, $t_1 = t_2 = t_3 = 1$ min.

3.2. Effects of substituting group on the RTP properties

The RTP properties of 27 coumarin derivatives and fluorescence properties of 20 derivatives are shown in Table 3. Most of the derivatives studied can give more intensive RTP signals on filter paper when using 1.0 mol l^{-1} $Pb(AC)_2$ as heavy-atom perturber. Their RTP properties, $\lambda_{ex}/\lambda_{em}$ maximum wavelength, intensity and RTP lifetime τ_p are not only dependent upon the kind of heavy-atom perturber (e.g. by using 1.0 mol l^{-1} $Pb(NO_3)_2$ instead of 1.0 mol l^{-1} $Pb(AC)_2$ as the perturber, lower intensity and shorter lifetime RTP signals were observed; these data are not shown in Table 3), but also upon the chemical structure of the derivatives.

It is well known that the substituents which enhance π-electron mobility increase the fluores-

Table 1
Chemical structure, name and code of the 27 derivatives studied

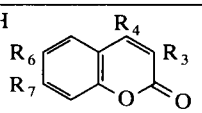
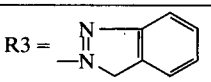
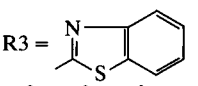
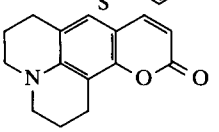
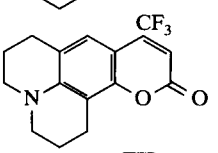
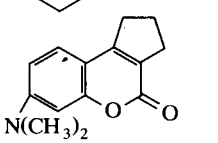
Code	Structure and name	Code	Structure and name
1	R3, R4, R6, R7 = H Coumarin 	16	 R7 = $-N(C_2H_5)_2$ Coumarin-7
2	R4 = –OH 4-Hydroxycoumarin	17	 R7 = $-N(C_2H_5)_2$ Coumarin-6
3	R7 = –OH 7-Hydroxycoumarin	18	 Coumarin-102
4	R3 = –COOH 3-Carboxylcoumarin	19	 Coumarin-153
5	R6 = $-SO_2Cl$ 6-Coumarinsulphonyl chloride	20	 Coumarin-138
6	R6 = R7 = –OH 6,7-Dihydroxycoumarin	21	6,7-Dihydroxy-4-methylcoumarin
7	R4 = $-CH_3$, R7 = –OH Coumarin-4	22	6,7-Dimethoxy-4-methylcoumarin
8	R4 = $-CH_3$, R7 = $-NH_2$ Coumarin-120	23	6,7-Diacetoxy-4-methylcoumarin
9	R4 = $-CH_3$, R7 = $-NH(C_2H_5)$ Coumarin-445	24	4-Bromomethyl-7-methoxycoumarin
10	R4 = $-CH_3$, R7 = $-N(CH_3)_2$ Coumarin-311	25	4-Bromomethyl-6,7-dimethoxycoumarin
11	R4 = $-CH_3$, R7 = $-N(C_2H_5)_2$ Coumarin-1	26	3-[2-(Diethylamino)ethyl]-7-hydroxy-4-methylcoumarin hydrochloride
12	R4 = $-CF_3$, R7 = $-NH_2$ Coumarin-151	27	7,8-Dihydroxy-6-methoxycoumarin 8-β-D-glucopyranoside
13	R4 = $-CF_3$, R7 = $-NH(C_2H_5)$ Coumarin-500		
14	R4 = $-CF_3$, R7 = $-N(CH_3)_2$ Coumarin-152		
15	R4 = $-CF_3$, R7 = $-N(C_2H_5)_2$ Coumarin-35		

Table 2
Comparison of the background emission of filter paper for various chemical treatment methods ($n = 6$)

	No heavy atom	1.0 mol l^{-1} $Pb(AC)_2$ as heavy-atom perturber							
		No. treatment	NaOH (6 mol l^{-1})	NaOH (1.5%)	HCl (6 mol l^{-1})	Ethanol	Acetone	β-CD [a] (satd.)	Water (80°C, 3 h)
I_{RTP}	1.62	17.10	8.75	12.48	16.85	13.84	15.78	15.38	13.88
C.V.(%)	13.6	6.8	20.7	3.7	7.0	3.9	6.0	4.2	4.4

[a] β-CD = β-cyclodextrin.

cence intensity and shift the maximum towards longer wavelenths. In general, electron-donating substituents tend to enhance fluorescence while electron-withdrawing substituents tend to diminish or to quench it. First, the monosubstituted compounds will be considered. Of the monosubstituted derivatives, 3-carboxylcoumarin gives the strongest intensity and the RTP lifetime τ_p is up to 6.25 ms. Both 4-hydroxy- and 7-hydroxycoumarin have the same substituent group; the non-bonding electrons on oxygen contributing the increase in the mobility of the π electrons of the benzene ring, but their $\lambda_{ex}/\lambda_{em}$ and RTP intensity are different because the –OH group is in a different position. For disubstituted compounds, when the group on position 4 is $-CH_3$, but the group on position 7 is changed from –OH, $-NH_2$, $-NH(C_2H_5)$, $-N(CH_3)_2$ to $-N(C_2H_5)_2$ (for code 7–11 compounds), the RTP $\lambda_{ex}/\lambda_{em}$ maxima shift regularly to longer wavelengths and the RTP intensity is also gradually increased (Fig. 3). These substituted derivatives are all more intense phosphores. It is interesting to note that after the substituted group on position 4 is changed from $-CH_3$ to $-CF_3$ (for code 12–15 compounds), the RTP intensity decreases with $\lambda_{ex}/\lambda_{em}$ maxima shifting towards longer wavelengths. However, the regularity mentioned above for 7-substituted derivatives is inverted and the RTP intensity is gradually decreasing in the same order.

When benzimidazolyl or benzothiazolyl was introduced or the electron conjugation system was extended (code 16–20 compound), the RTP signal of the derivative is enhanced and the $\lambda_{ex}/\lambda_{em}$ maxima shift towards longer wavelength. The RTP emission intensity of compound 19 is rather low due to the presence of a $-CF_3$ group in the molecule.

The effects of polysubstitution are much more complex since the overall effect on the π-electron mobility produced by different substituents must be considered. Consider derivatives 21, 22 and 23, which have a methyl group in position 4 but two –OH, $-OCH_3$ or $-OOCCH_3$ groups in positions 6 or 7, respectively. Results shown in Table 3 indicate that the RTP emission intensity is increasing in the order $-OOCCH_3 > -OCH_3 >$ –OH, but the $\lambda_{ex}/\lambda_{em}$ maxima shift towards shorter wavelength in the same order.

It is also found that RTP spectra of some the of coumarin derivatives such as 6-coumarinsulph-

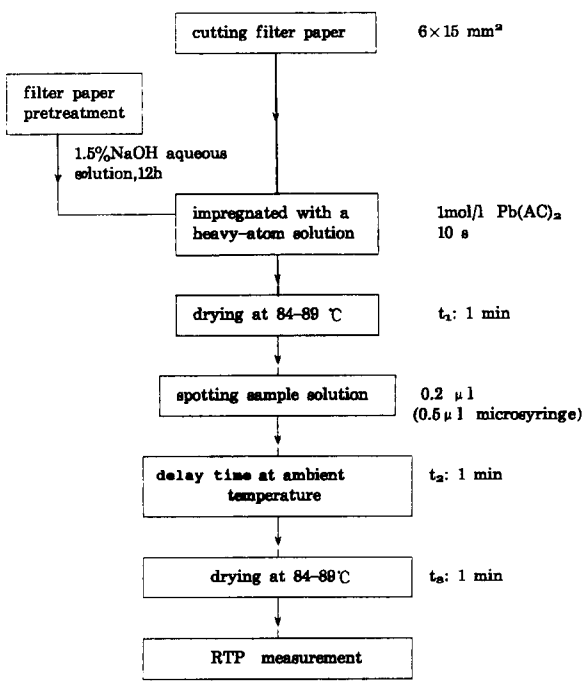


Fig. 2. Scheme of the RTP analysis procedure using filter paper.

onyl chloride, 4-methyl-7-diethylaminocoumarin and 4-methyl-7-dimethylaminocoumarin, etc., show a wide-range and structureless or less structured spectrum when the concentration of the sample solution is higher. Fig. 4 shows the RTP spectra of 6-coumarinsulphonyl chloride obtained by using 0.2 μl of 1.0×10^{-2} mol l^{-1} ethanol solution of the compound and scanning repeatedly with an interval of 4 min. The RTP spectrum shows at first wide-range structureless characteristics and after long time irradiation during repeated scanning, the RTP spectrum returns to the original one the $\lambda_{ex}/\lambda_{em}$ maxima of which are 320/496 nm. This behaviour might be caused by the formation of an excimer or a dimer of the derivative at high concentrations.

It is important to note that some coumarin derivatives which were suggested for use as fluorescence labelling agent recently, such as 6-coumarinsulphonyl chloride [8], 4-bromomethyl-7-methoxycoumarin [15–17], 4-bromomethyl-6,7-dimethoxycoumarin [18–20], etc., can also give more intensive RTP signals. This means that if they are used as phosphorescence labelling agents, maybe a much better selectivity will be obtained because the phosphorescence wavelength lies in the longer-wavelength region which exhibits a lower interference of the background. Some sam-

Table 3
RTP and fluorescence properties of 27 coumarin derivatives on filter paper

Code	Compound	Phosphorescence				Fluorescence	
		$\lambda_{ex}/\lambda_{em}$	I_p	I_b	τ_p(ms)	$\lambda_{ex}/\lambda_{em}$	I_r
1	Coumarin	306/499	64.2	15.9	2.94	–	0
2	4-Hydroxycoumarin	308/477	25.5	10.9	2.31	–	0
3	7-Hydroxycoumarin	336/511	100.2	7.5	2.75	330/389	79.0
4	3-Carboxylcoumarin	314/516	328.2	15.0	6.25	311/398	16.3
5	6-Coumarinsulphonyl chloride	320/496	86.1	10.2	2.56	–	0
6	6,7-Dihydroxycoumarin	294/486	12.3	12.2	–	–	0
7	Coumarin-4	333/497	131.1	6.6	3.12	326/383	53.0
8	Coumarin-120	373/508	189.4	1.4	3.25	360/424	271.9
9	Coumarin-445	380/517	204.3	1.2	3.31	374/440	237.0
10	Coumarin-311	385/526	247.8	1.2	3.31	379/444	250.0
11	Coumarin-1	392/529	303.9	0.9	3.50	385/446	243.5
12	Coumarin-151	398/602	51.5	0.9	2.44	380/476	166.0
13	Coumarin-500	407/614	49.8	0.6	1.86	396/486	134.2
14	Coumarin-152	422/622	40.4	0.3	2.25	406/497	117.1
15	Coumarin-35	426/632	43.0	0.1	2.19	415/500	134.0
16	Coumarin-7	474/616	110.0	−0.3	2.38	464/508	332.4
17	Coumarin-6	478/631	67.2	−0.5	2.50	473/514	459.8
18	Coumarin-102	417/561	141.1	0.4	3.06	402/470	154.6
19	Coumarin-153	440/656	21.9	−0.4	2.06	434/523	69
20	Coumarin-138	388/534	160.1	0.9	3.38	379/443	208.9
21	6,7-Dihydroxy-4-methylcoumarin	365/540	6.7	4.5			
22	6,7-Dimethoxy-4-methylcoumarin	335/525	95.4	1.5			
23	6,7-Diacetoxy-methylcoumarin	318/497	decays	rapidly			
24	4-Bromomethyl-7-methoxycoumarin	322/520	30.9	9.0			
25	4-Bromomethyl-6,7-dimethoxy-coumarin	358/520	42.0	3.0			
26	3-[2(-Diethylamino)ethyl]-7-hydroxy-4-methylcoumarin hydrochloride	332/511	28.8	3.0			
27	7,8-Dihydroxy-6-methoxycoumarin 8-β-D-glucopyranoside	385/550	21.3	2.2			

$n = 5$, $V = 0.2$ μl, $C = 5.0 \times 10^{-4}$ mol l^{-1} (17# is saturated solution).
I_p and I_b are RTP intensity of sample and blank respectively; τ_p is phosphorescence lifetime, I_f is fluorescence intensity.

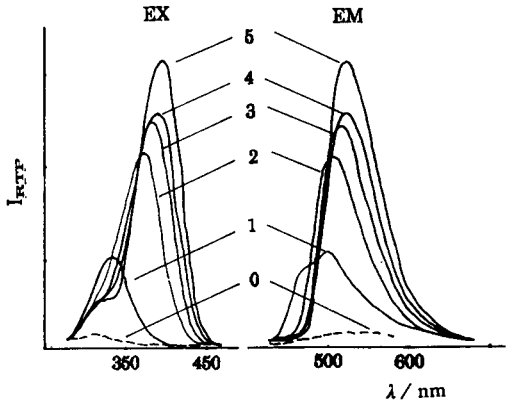


Fig. 3. RTP spectra of several coumarin derivatives on filter paper. 0 = Background; 1 = coumarin-4; 2 = coumarin-120; 3 = coumarin-445; 4 = coumarin-311; 5 = coumarin-1.

ples such as plasma and other biological liquids often have a higher fluorescence background.

3.3. Quantitative analysis potentiality

The RTP analytical curve for 6-coumarinsulphonyl chloride was obtained according to the procedure in Fig. 2. As shown in Fig. 5, the RTP signal intensity is proportional to the absolute amount of the analyte, and a good linear relationship is maintained up to 200 pmol in 0.2 μl sample solution. The detection limit calculated based on three times the standard deviation of background emission is 1.5 pmol.

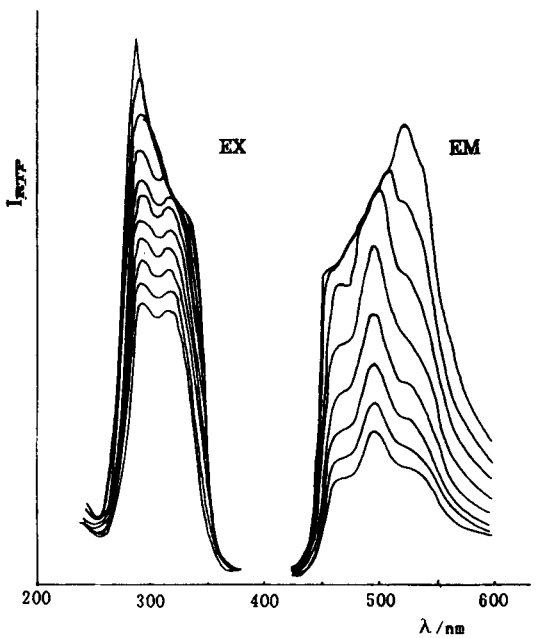


Fig. 4. Repeated scan RTP spectra of 6-coumarinsulphonyl chloride ($\lambda_{ex}/\lambda_{em}$ = 320/496 nm). Concentration = 1.0×10^{-2} mol/l; V = 0.2 μl; Δt = 4 min; scan speed = 200 nm/min.

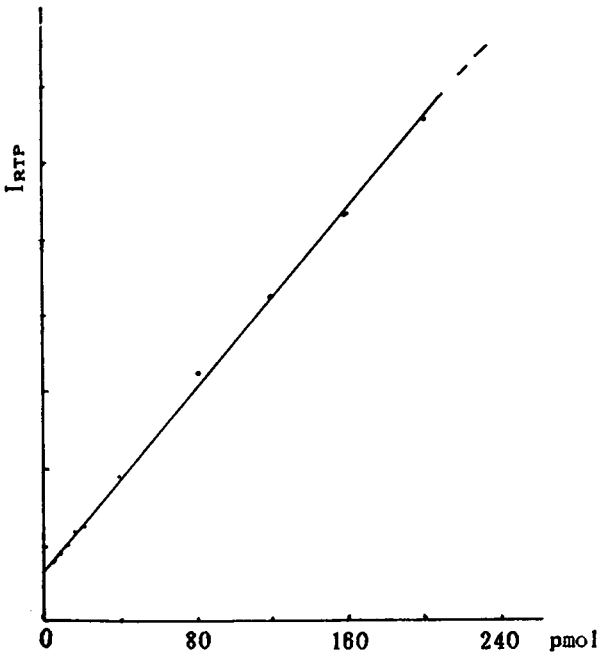


Fig. 5. The RTP analytical curve for 6-coumarinsulphonyl chloride. V = 0.2 μl; $\lambda_{ex}/\lambda_{em}$ = 320/496 nm).

4. Conclusion

The fact that many of the coumarin derivatives, especially those which have been used as fluorescence labelling reagents, can give a more intensive RTP signal means that they could be expected to serve as RTP labelling reagent.

Acknowledgement

This project was supported by the Chinese National Natural Science Foundation (Grant No. 29275200) and Beijing Natural Science Foundation (Grant No. 1922007).

References

[1] M. Roth, J. Chromatogr., 30 (1967) 276.
[2] S.L. Wellen, R.A. Paynter and J.D. Winefordner, Spectrochim. Acta, 30A (1974) 2133.
[3] T. Vo-Dinh, Room Temperature Phosphorimetry for Chemical Analysis, Wiley, New York, 1984.
[4] R.J. Hurtubise, Solid Surface Luminescence Analysis: Theory, Instrumentation and Applications, Marcel, Dekker, New York, 1981.
[5] R.J. Hurtubise, Phosphorimetry: Theory, Instrumentation and Application, VCH, New York, 1990.
[6] J. Kuhl, H. Telle, R. Schieder and U. Brinkmann, Opt. Commun., 24 (1978) 251.
[7] T.E. Bush and G.W. Scott, J. Phys. Chem., 85 (1981) 144.
[8] G. Jones II, W.R. Jackson and C. Choi, J. Phys. Chem., 89 (1985) 294.
[9] A. Hulshoff and H. Lingeman, in S.G. Schuman (Ed.), Molecular Luminescence Spectroscopy Methods and Application, Part 1, Wiley, New York, 1985, p. 625.
[10] Z.-Y. Qiu, H.-Z. Li and S.-B. Zeng, Sepu (Chromatography), 7 (1989) 340.
[11] S.M. Al-Kindy and J.N. Miller, Anal. Chim. Acta, 227 (1989) 145.
[12] S.M. Al-Kindy and J.N. Miller, Anal. Chim. Acta, 227 (1989) 155.
[13] S.-Z. Yang and L.-D. Li, Fenxi Huaxue, 20 (1992) 202.
[14] L.-D. Li and S.-Z. Yang, Fenxi Huaxue, 20 (1992) 1261.
[15] W. Dünges, Anal. Chem., 49 (1977) 442.
[16] J.H. Wolf and J. Korf, J. Chromatogr., 436 (1988) 437.
[17] C.G. Kindberg, C.M. Riley, J.F. Stobaugh and M. Slavik, J. Chromatogr., 473 (1989) 431.
[18] R. Farinotti, Ph. Siard, J. Bouson, S. Kirkiacharian, B. Valeur and G. Mahuzier, J. Chromatogr., 269 (1983) 81.
[19] T. Hiratsuka, J. Biochem., 101 (1987) 1457.
[20] J.G. Alvazer, X.G. Fang et al., J. Liq. Chromatogr., 13 (1990) 2727.

ELSEVIER

Analytica Chimica Acta 296 (1994) 107–113

ANALYTICA CHIMICA ACTA

A model for optical saturation thermal lens spectrometry

Guillermo Ramis-Ramos *, J.J. Baeza Baeza, E.F. Simó Alfonso

Departament de Química Analítica, Universitat de València, 46100 Burjassot (València), Spain

Received 5th January 1994; revised manuscript received 30th March 1994

Abstract

A model for thermal lens spectrometry (TLS) in partial optical saturation conditions is developed. The model predicts that optical saturation is approached to a significant degree in many TLS experiments, and that it can be reached to a high degree with the currently available laser and optical technologies. When optical saturation is approached, the sensitivity decreases and the signal depends on the rate constant of the slowest step of the decay process. Important consequences for the future development of photothermal spectrometric techniques can be derived.

Keywords: Thermal lens spectrometry; Optical saturation

1. Introduction

The thermal lens effect was first described by Gordon et al. [1]. One decade later, thermal lens spectrometry (TLS) was developed as an analytical technique by Dovichi and Harris [2–4], Carter et al. [5–8], Mori et al. [9], Nakanishi et al. [10] and other authors. Extremely low absorbances can be accurately measured by TLS. Advances have continued to be produced in the last few years, including the development of photothermal circular dichroism, and descriptions of colateral effects with analytical and physico-chemical potential applications. Several reviews have been published [11–15].

Two models describing the thermal lens effect have been put forward, i.e., the parabolic model [1,16] and the aberrant model [17–20]. Following the parabolic model, the TLS signal is given by:

$$\frac{\Delta I}{I} = \frac{I(0) - I(\infty)}{I(\infty)} = \theta + \frac{\theta^2}{2} \quad (1)$$

where

$$\theta = \frac{2.303 P(-\mathrm{d}n/\mathrm{d}T) A}{\lambda_{\mathrm{pr}} \kappa} \quad (2)$$

P is the pump power, $\mathrm{d}n/\mathrm{d}T$ is the temperature coefficient of the refractive index, A is the absorbance, λ_{pr} is the wavelength of the probe laser, and κ is the thermal conductivity of the medium. For weak TLS signals the quadratic term is frequently neglected. To produce good agreement between theory and data, Dovichi [12] and Carter and Harris [18] introduced a correction coefficient in the terms of Eq. 1:

$$\frac{\Delta I}{I} = 0.52\theta + 0.13\theta^2 \quad (3)$$

* Corresponding author.

0003-2670/94/$07.00
SSDI 0003-2670(94)00200-6

The coefficients of Eq. 3, however, have been found to vary with the nature of the solute [21,22]. In an attempt to rationalize the differences found for a series of solutes, we examined the possibility of approaching optical saturation. When optical saturation is approached, the excited state populates significantly, and the population of the fundamental state decreases by an equivalent amount. In absorptiometric techniques, the sensitivity is proportional to the population of the fundamental state and, therefore, a smaller sensitivity should be expected if optical saturation is approached.

Optical saturation has been demonstrated in atomic emission techniques and in molecular fluorescence with pulsed laser excitation [23]. As far as we know, however, optical saturation in absorptiometric techniques has not been discussed. Large radiation intensities are required to approach optical saturation and, therefore, spectrophotometry which is based in measuring the difference between two radiation intensities cannot be used in optical saturation conditions. Instead, photothermal spectroscopy, including TLS and photoacoustic spectroscopy, offers a chance of demonstrating optical saturation with an absorptiometric technique.

We have developed a model for TLS in partial optical saturation conditions. The model predicts that optical saturation is approached to a significant degree in many TLS experiments reported in the literature, and that it can be reached to a high degree with the currently available laser and optical technologies. Important consequences for the future development of the photothermal spectrometric techniques can be derived.

2. Theory

2.1. A kinetic model for TLS

In both spectrophotometry and TLS, the available radiant energy is partially converted to electronic energy by the absorbing species, and released to the medium as thermal energy. In photoacoustics, part of the energy is dissipated as a mechanical wavefront. In TLS, part of this thermal energy is accumulated within the irradiated region to build up the thermal gradient, and part of it is dissipated by thermal conduction and convection. The temperature rise for a Gaussian pump beam is characterized by the time constant; that for liquids ranges from 200-300 ms to 2-3 μs when the spot-size of the beam ranges from 100 to 1 μm [12]. After a period of time which is proportional to the square of the pump laser spot-size, a dynamic equilibrium is achieved, and all the absorbed energy is communicated to the bulk phase.

In all cases, the thermal energy is generated by an absorption-decay process which can be described as follows:

$$X + \phi \underset{k_s}{\overset{k_a}{\rightleftarrows}} X^* \tag{4}$$

$$X^* + \phi \xrightarrow{k_i} X + 2\phi \tag{5}$$

$$X^* \xrightarrow{k_q} X + q \tag{6}$$

where ϕ represents the molar concentration of photons (mol dm^{-3}), q is the thermal energy, X and X* are the ground and excited states of the absorbing species,respectively, k_a is the absorption rate constant, and k_s and k_i are the rate constants of the spontaneous and induced emission processes, respectively. As shown by Einstein, the rate constants for the absorption and induced emission processes are equal ($k_a = k_i$) [23–25]. For simplicity, photochemical changes have not been included in the model of Eqs. 3-6, and will be not considered in this work.

The decay rate constant, k_q, represents all the non-radiative decay processes which contribute to: (1) the degeneration of the electronic and vibrational energies as translational or thermal energy, (2) the depopulation of the excited state, and (3) the regeneration of the ground state population. These processes include external conversion, internal conversion and intersystem crossing. The rate constants of these processes are highly dependent on the structure and the environment of the molecule. The rate constant for internal conversion to the ground state varies from typically 10^5 to 10^7 s^{-1}. For organic

molecules, the rate constant for intersystem crossing is typically 10^6 to 10^9 s^{-1}, and increases in the presence of a heavy atom [23–25].

This gives rise to the question of which is the value of k_q which should be used to model thermal lens experiments when optical saturation conditions are approached. When an emission technique is used, the signal is a function of the population of the excited state and, therefore, the rate constant of the fastest decay process which is capable of depleting the excited state should be used to model the experiments. However, TLS is an absorptiometric technique and the sensitivity is proportional to the population of the ground state. Therefore, if optical saturation is approached, the slowest step of the sequential decay process will limit the sensitivity and should be used to model the experiments. Taking into account that the intersystem crossing should be combined with internal conversions to regenerate the ground state population, the latter should be the slowest step of the total decay process. Therefore, in the discussion that follows it will be assumed that k_q ranges from 10^5 to 10^7 s^{-1}, as does the rate constant for the internal conversion process.

2.2. Kinetic interpretation of the absorption process

Eqs. 4 and 5 can be used to establish the relationship between k_a and the molar absorptivity, ε. In the beam direction, the balance of photons is:

$$\frac{d\phi}{dt} = -k_a(x - x^*)\phi \tag{7}$$

where x and x^* are the molar concentrations of X and X*, respectively. The spontaneously emitted photons have been previously absorbed, and therefore, most of them cause the intensity in the forward beam direction to decrease. These photons, and the photons which have been absorbed and converted to thermal energy, cannot be distinguished from each other by the spectrophotometric detector. Both types of photons contribute positively to k_a. Only a small fraction of photons is spontaneously emitted in the forward beam direction and can reach the detector. This negative contribution to k_a has been neglected in Eq. 7.

Eq. 7 indicates that the concentration of photons which move along the beam direction decreases at a rate which is proportional to the difference of populations between the ground and the excited states, and to the photon concentration itself. During a differential of time, dt, the photons move along a pathlength differential, dz, some of them are absorbed and some induce the emission of other photons. The relationship between dt and dz is given by:

$$dt = \frac{n}{c}dz \tag{8}$$

where n is the refractive index of the medium and c is the speed of light in the vacuum; c/n is the speed of light in the medium. By substituting dt in Eq. 7 and rearranging:

$$\frac{d\phi}{\phi} = -\frac{k_a(x - x^*)n\,dz}{c} \tag{9}$$

Under the conditions used in spectrophotometry, the concentration of the excited species, x^*, is extremely small and can be neglected. Integration of Eq. 9 from ϕ_0 (incident photon concentration) to ϕ (emergent photon concentration) along the optical pathlength, b, leads to:

$$\ln\frac{\phi}{\phi_0} = -\frac{k_a x n b}{c} \tag{10}$$

From Eq. 10 and from the Lambert-Beer law it can be deduced that:

$$A = -\log\frac{I}{I_0} = -\log\frac{\phi}{\phi_0} = \varepsilon x b = \frac{k_a n x b}{(\ln 10)c} \tag{11}$$

where I_0 and I are the intensities of the incident and emergent beams. Finally:

$$k_a = \frac{(\ln 10)\varepsilon c}{n} \tag{12}$$

2.3. Power released as thermal energy per unit volume and unit time in TLS

From the absorption-decay processes of Eqs. 4-6, the following kinetic model can be derived:

$$\frac{dx}{dt} = -k_a\phi(x - x^*) + k_s x^* + k_q x^* \tag{13}$$

Eq. 13 can be rearranged as follows:

$$\frac{dx}{dt} + (2k_a\phi + k_d)x = (k_a\phi + k_d)x_0 \qquad (14)$$

where $k_d = k_s + k_q$ (the sum of the rate constants of the radiative and non-radiative decay processes, respectively). For simplicity, in the discussion which follows we will assume that the radiative decay processes are negligible, that is, $k_d \approx k_q$. Taking into account that the total concentration of the chromophore, $x_0 = x + x^*$, is constant, the integration of Eq. 14 is immediate, and we have:

$$x = x_0\frac{k_a\phi + k_d}{2k_a\phi + k_d}\left[1 + \frac{k_a\phi}{k_a\phi + k_d} \times \exp[-(2k_a\phi + k_d)t]\right] \qquad (15)$$

Rearranging Eq. 15:

$$x^* = x_0\frac{k_a\phi}{2k_a\phi + k_d}[1 - \exp[-(2k_a\phi + k_d)t]] \qquad (16)$$

Eq. 16 indicates that x^* increases from 0 (when $t = 0$) up to an asymptotic value. At a sufficiently large value of t a dynamic steady-state is established, and the exponential term can be neglected. Owing to the large values of k_d (see above), the stationary state should be reached to a high degree below the microsecond range. This is negligible when compared to the time required by a spectrophotometer to measure the absorbance, or for a thermal lens to be built up to a sufficiently high degree and measured (millisecond range). At steady-state the exponential term in Eq. 16 is zero, and x^* is maintained constant. The kinetic model for the release of thermal energy (Eq. 6) is therefore simplified to:

$$\frac{dq}{dt} = k_q x^* = x_0\frac{k_q k_a\phi}{2k_a\phi + k_d} \qquad (17)$$

where dq/dt is the number of moles of photons which are absorbed and transformed to thermal energy per volume unit per unit of time. To transform moles of photons to energy units we write:

$$\frac{dq}{dt} = x_0\frac{k_q k_a\phi}{2k_a\phi + k_d}\frac{hcN_A}{\lambda_{pu}} \qquad (18)$$

where h is the Planck constant, λ_{pu} is the wavelength of the pump radiation and N_A is Avogadro's number. The photon concentration, ϕ, can be expressed as a function of the pump power as follows:

$$\phi = \frac{P\lambda_{pu}n}{hc^2N_A\pi\omega_{pu}^2} \qquad (19)$$

where ω_{pu} is the average radius (distance from the axis to the intensity $1/e^2$ point) of the pump beam at the cell location. By substituting the values of ϕ and k_a given by Eqs. 19 and 12, respectively, in Eq. 18:

$$\frac{dq}{dt} = x_0 B\frac{(\ln 10)\varepsilon P}{\pi\omega_{pu}^2} \qquad (20)$$

where:

$$B = \frac{k_q}{\dfrac{2(\ln 10)\varepsilon\lambda_{pu}P}{\pi\omega_{pu}^2 hcN_A} + k_d} = \frac{k_q}{\dfrac{0.1226\varepsilon\lambda_{pu}P}{\omega_{pu}^2} + k_d} \qquad (21)$$

and the coefficient, 0.1226, is given in J^{-1} cm^{-1} mol. The coefficient 1.226×10^{-6} should be used instead when ε is given in cm^{-1} mol^{-1} dm^3, λ_{pu} in nm, P in mW and ω_{pu} in mm.

2.4. Modification of the parabolic model for TLS

Eqs. 20–21 give the energy that is absorbed and dissipated as thermal energy per unit volume per unit time. Next, Eq. 20 will be used to obtain the intensity of the thermal lens effect. The classical parabolic model referred to the mode-matched dual-beam co-linear configuration [1,17–19] will be used to develop the theory and in the discussion which follows. A significant amount of energy is assumed to always be present in the center of the probe beam at the far field, that is, it is assumed that in all instances the

detector system is far from saturation. Owing to the Gaussian distribution of the energy across the profile of the TEM_{00} beam [17–19], we can write:

$$Q(r,t) = x_0 B \frac{(\ln 10)\varepsilon P}{\pi \omega_{pu}^2} \exp\left(-\frac{2r^2}{\omega_{pu}^2}\right) \quad (22)$$

where r is the radial variable. The only difference between Eq. 22 and the corresponding expression used in the development of the classical parabolic model of TLS is the introduction of the variable B, which takes into account the possible influence of the rate constants of the model given by Eqs. 4–6. Integration of Eq. 22 to obtain the temperature change across the radial variable is well documented [1,17–19]. It results in:

$$\Delta T(r,t) = -x_0 B \frac{(\ln 10)\varepsilon P r^2}{4\kappa\pi\omega_{pu}^2}\left(1 + \frac{t_c}{2t}\right) \quad (23)$$

where t_c is the thermal lens time constant [12], which is given by:

$$t_c = \frac{\omega_{pu}^2 \rho c_p}{4\kappa} \quad (24)$$

where ρ and c_p are the density and the specific heat of the medium, respectively. Following the parabolic model, the focal length of the thermal lens, f, is given by:

$$\frac{1}{f} = \frac{dn}{dT}\frac{\partial^2 T}{\partial r^2} b \quad (25)$$

where b is the thickness of the sample. The focal length is related to the intensity of the thermal lens effect by means of:

$$\frac{\Delta I}{I} = -\frac{2z}{f} + \frac{z_c + z^2}{f^2} \quad (26)$$

where z is the distance from the probe beam waist to the cell. For a maximum sensitivity this distance should equal the confocal distance, z_c, which is given by:

$$z_c = \frac{\pi\omega^2}{\lambda_{pr}} \quad (27)$$

where ω is the probe beam spot size at the waist. Taking into account that:

$$\frac{\partial^2 T}{\partial r^2} = \frac{\partial^2 \Delta T}{\partial r^2} = -2x_0 B \frac{(\ln 10)\varepsilon P}{4\kappa\pi\omega_{pu}^2} \quad (28)$$

substituting in Eqs. 25–27 and introducing the Carter and Harris correction [12,18], the following expression is obtained:

$$\frac{\Delta I}{I} = 0.52B\theta + 0.13B^2\theta^2 \quad (29)$$

Eq. 29 coincides with Eq. 3 when $B = 1$, that is when:

$$\frac{0.1226\varepsilon\lambda_{pu}P}{\omega_{pu}^2} \ll k_d \approx k_q \quad (30)$$

2.5. The effects of approaching optical saturation in thermal lens spectrometry

From Eqs. 15 and 16, and assuming that the stationary state has been reached (see above), we can write:

$$B = \frac{x - x^*}{x_0} = \frac{1 - x^*/x}{1 + x^*/x} \quad (31)$$

Eq. 31 reveals that the variable B is the difference between the populations of the ground and excited states divided by the total concentration of the absorbing species. Therefore, when $B = 1$, $x^* \ll x$. The variable B can also be written as a function of the degree of optical saturation, x^*/x, which ranges from 0 to 1.

From Eqs. 21 and 31, we can write:

$$\frac{x^*}{x} = \frac{1}{\frac{2k_q\omega_{pu}^2}{0.1226\varepsilon\lambda_{pu}P} + 1} \quad (32)$$

which can be used to predict the degree of optical saturation that can be reached for a given set of experimental conditions. Thus, for instance, the necessary power to achieve 50% optical saturation is given by:

$$P_{0.5} = \frac{2\omega_{pu}^2 k_{pu}}{1.226 \times 10^{-6}\varepsilon\lambda_{pu}} \quad (33)$$

Eq. 33 indicates that for a beam of $\omega_{pu} = 0.1$ mm, with $\varepsilon = 20{,}000$ cm^{-1} mol^{-1} dm^3, and $\lambda_{pu} = 488$ nm, $P_{0.5}$ ranges from 167 mW up to 16.7 W when k_q ranges from 10^5 to 10^7 s^{-1}, and that $P_{0.5}$ decreases by a factor of 1:100 when the beam is focused in a factor of 10× ($w_{pu} = 0.01$ mm). Therefore, Eq. 32 predicts that optical saturation has been reached to a significant degree in many TLS experiments reported in the literature.

The effect of approaching optical saturation on the shape of a TLS spectrum has been simulated in Fig. 1. Similarly, the effect of increasing the pump power on the TLS signal for simulated species of increasing k_q values is shown in Fig. 2. Eq. 29 also predicts that, in the presence of significant optical saturation, both the sensitivity and the coefficient of the quadratic term of the calibration graph will be smaller than expected.

Finally, the possibility of reaching optical saturation in a large degree is considered. From Eq. 32 it is deduced that for x^*/x to approach 1, a powerful and tightly focused pump laser beam

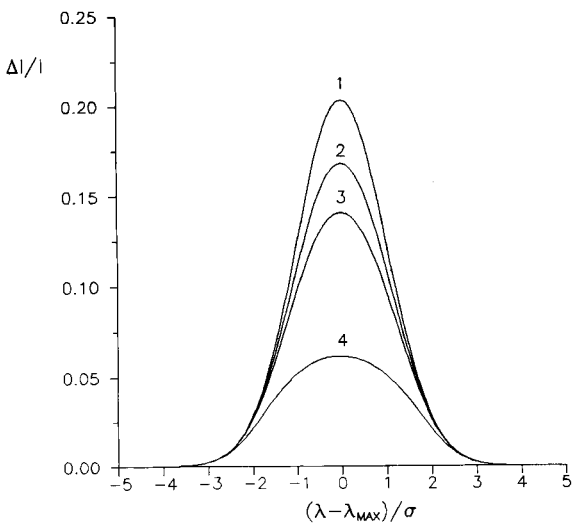


Fig. 1. Variation of the TLS signal with the wavelength of the pump radiation as predicted by Eq. 29. The x-axis is given in standard deviation units for a Gaussian absorption band. Assumed data were: $x_0 = 2\times10^{-7}$ mol dm^{-3}, $\varepsilon_{max} = 20000$ cm^{-1} mol^{-1} dm^3, $\lambda_{max} = 488$ nm, $\sigma = 10$ nm (standard deviation of the band), $\kappa = 5.95$ mW cm^{-1}°C^{-1} (water), $(-dn/dT) = 0.81\times10^{-4}$°C^{-1} (water), $\lambda_{pr} = 633$ nm, $\omega_{pu} = 0.1$ mm, $P = 100$ mW. Values of k_q: (1) 10^7 s^{-1}, (2) 10^6 s^{-1}, (3) 5×10^5 s^{-1}, and (4) 10^5 s^{-1}. Values of x^*/x when $\lambda_{pu} = \lambda_{max}$: (1) 1%, (2) 10%, (3) 19%, and (4) 55%.

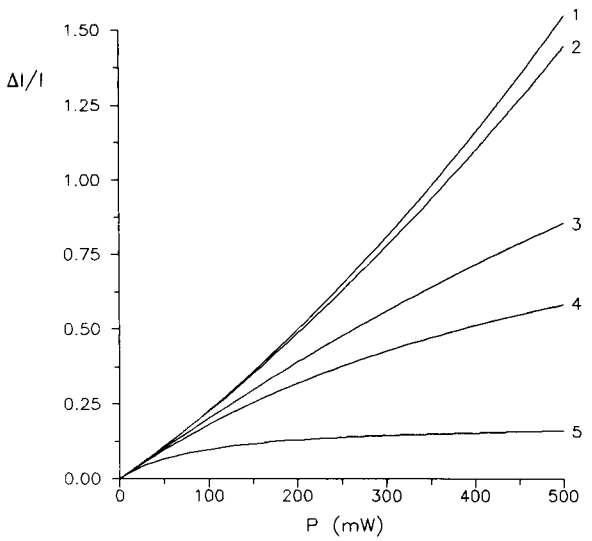


Fig. 2. Variation of the TLS signal with the pump power as predicted by Eq. 29. Assumed data were: $x_0 = 4\times10^{-7}$ M, $\lambda_{pu} = \lambda_{max}$, and other data as in Fig. 1. Values of k_q: (1) 10^8 s^{-1}, (2) 10^7 s^{-1}, (3) 10^6 s^{-1}, (4) 5×10^5 s^{-1}, and (5) 10^5 s^{-1}. Values of x^*/x when $P = 500$ mW: (1) 0.3%, (2) 3%, (3) 23%, (4) 37%, and (5) 75%.

should be used, and a non-fluorescent compound having a large molar absorptivity and a relatively small value of k_q should be pumped. In these conditions, from Eq. 21 we can write:

$$\frac{0.1226\varepsilon\lambda_{pu}P}{\omega_{pu}^2} \gg k_d \approx k_q \qquad (34)$$

and from Eqs. 2, 29 and 34:

$$\frac{\Delta I}{I} = \frac{9.77(-dn/dT)b\omega_{pu}^2}{\lambda_{pr}\lambda_{pu}\kappa} x_0 k_q \qquad (35)$$

where for simplicity the quadratic term has not been written, and where the coefficient (9.77) is given in J cm mol^{-1}. The coefficient 9.77×10^{12} should be used instead when ε is given in cm^{-1} mol^{-1} dm^3, (dn/dT) in °C^{-1}, b in cm, ω_{pu} in mm, the wavelengths in nm, κ in mW cm^{-1}°C^{-1}, x_0 in mol dm^{-3} and k_q in s^{-1}.

3. Discussion

On the basis of Eq. 35, optical saturation thermal lens spectrometry (OSTLS) appears as a dif-

ferent technique when compared to both conventional absorption spectrophotometry and TLS. First, if a high value of x^*/x is reached, the TLS signal would not depend on the pump power and on ε, and therefore, it would not depend on the absorbance either, but it would be directly proportional to the concentration of the absorbing species and to k_q. Therefore, OSTLS can be used to study the slowest step of the sequential decay process.

Secondly, through k_q, the OSTLS signal should exhibit a more complex dependence on the nature of the solute and the medium than predicted by the classical TLS theory. Thus, Eq. 35 indicates that the OSTLS signal should depend on those factors which can modify the rate of the slowest step of the decay process, such as the temperature, the rigidity of the absorbing species and the nature of the medium. Therefore in OSTLS, a chromophore can be used as a probe to measure the variations of the factors which make k_q vary. In this way, it could be possible to monitor environmental changes in complex systems, e.g., biochemical molecules, or selected sites in biological systems.

Also, control of the variables which modify the k_q values can be of interest to improve TLS sensitivity and selectivity, including the development of multicomponent TLS-OSTLS procedures. It is also interesting to observe that the sensitivity in OSTLS depends not only on the wavelength of the probe laser, but also on the wavelength of the pump radiation, the sensitivity increasing when probe and pump lasers of shorter wavelengths are used. Finally, the OSTLS signal should be insensitive to the noise and drift of the pump power.

It should be kept in mind, however, that approaching optical saturation to a high degree can be prevented by other effects which have not been considered in the model, e.g., diffusion and convection of the "hot" or intermediate molecules in gases and liquids, and volatilisation in liquids. Data obtained in the author's laboratory with several organic dyes in water and water-ethanol mixtures agree with the predictions of Eqs. 21 and 29 at least at moderate values of the degree of saturation, x^*/x. Work on these aspects is in progress [22].

Acknowledgements

This work was supported by the DGICYT of Spain, Project PB90/425.

References

[1] J.P. Gordon, R.C.C. Leite, R.S. Moore, S.P.S. Porto and J.R. Whinnery, J. Appl. Phys., 36 (1965) 3.
[2] N.J. Dovichi and J.M. Harris, Anal. Chem., 51 (1979) 728.
[3] N.J. Dovichi and J.M. Harris, Anal. Chem., 52 (1980) 2338.
[4] N.J. Dovichi and J.M. Harris, Anal. Chem., 53 (1981) 106.
[5] C.A. Carter, J.M. Brady and J.M. Harris, Appl. Spectrosc., 36 (1982) 309.
[6] C.A. Carter and J.M. Harris, Appl. Spectrosc., 37 (1983) 166.
[7] C.A. Carter and J.M. Harris, Anal. Chem., 55 (1983) 1256.
[8] C.A. Carter and J.M. Harris, Anal. Chem., 56 (1984) 922.
[9] K. Mori, T. Imasaka and N. Ishibashi, Anal. Chem., 54 (1982) 2034.
[10] K. Nakanishi, T. Imasaka and N. Ishibashi, Anal. Chem., 57 (1985) 1219.
[11] M.D. Morris and K. Peck, Anal. Chem., 58 (1986) 811A.
[12] N.J. Dovichi, CRC Crit. Rev. Anal. Chem., 17 (1987) 357.
[13] J. Georges and J.M. Mermet, Analusis, 16 (1988) 203.
[14] V.I. Grishko and I.G. Yudelevich, Zh. Anal. Khim., 44 (1989) 1753.
[15] G. Ramis-Ramos, Anal. Chim. Acta, 283 (1993) 623.
[16] C. Hu and J.R. Whinnery, Appl. Optics, 12 (1973) 72.
[17] S.J. Sheldon, L.V. Knight and J.M. Thorne, Appl. Optics, 21 (1982) 1663.
[18] C.A. Carter and J.M. Harris, Appl. Optics, 23 (1984) 476.
[19] T. Berthoud, N. Delorme and P. Mauchien, Anal. Chem., 57 (1985) 216.
[20] Jun Shen, R.D. Lowe and R.D. Snook, Chem. Phys., 165 (1992) 385.
[21] M. Terazima, M. Horiguchi and T. Azumi, Anal. Chem., 61 (1989) 883.
[22] G. Ramis-Ramos, Y. Martín Biosca, E.F. Simó Alfonso, J.J. Baeza Baeza and J.S. Esteve Romero, in preparation.
[23] J.D. Ingle and S.R. Crouch, Spectrochemical Analysis, Prentice Hall, New Jersey, 1988, pp. 316, 458 and 564.
[24] O. Svelto, Principles of Lasers, Plenum, New York, 3rd edn., 1989, pp. 3–4.
[25] W. Demtröder, Laser Spectroscopy, Springer, Berlin, 2nd corrected printing, 1982, pp. 13–16.

PUBLICATION SCHEDULE FOR 1994

	J	F	M	A	M	J	J	A	S	O	N	D
Anal. Chim. Acta	284/3 285/1–2 285/3	286/1 286/2 286/3	287/1–2 287/3 288/1–2	288/3 289/1 289/2	289/3 290/1–2 290/3	291/1–2 291/3 292/1–2	292/3 293/1–2 293/3	294/1 294/2 294/3	295/1–2 295/3 296/1	296/2 296/3 297/1–2	297/3 298/1 298/2	298/3 299/1 299/2
Vib. Spec.	6/2		6/3		7/1		7/2		7/3		8/1	

INFORMATION FOR AUTHORS

Detailed "Instructions to Authors" for *Analytica Chimica Acta* was published in Volume 289, No. 3, pp. 381–384. Free reprints of the "Instructions to Authors" of *Analytica Chimica Acta* and *Vibrational Spectroscopy* are available from the Editors or from: Elsevier Science B.V., P.O. Box 330, 1000 AH Amsterdam, The Netherlands. Telefax: (+31-20) 5862 459.

Manuscripts. The language of the journal is English. English linguistic improvement is provided as part of the normal editorial processing. Authors should submit three copies of the manuscript in clear double-spaced typing on one side of the paper only. *Vibrational Spectroscopy* also accepts papers in English only.

Rapid publication letters. *Letters* are short papers that describe innovative research. Criteria for letters are novelty, quality, significance, urgency and brevity. Submission data: max. of 2 printed pages (incl. Figs., Tables, Abstr., Refs.); short abstract (e.g., 3 lines); *no* proofs will be sent to the authors; submission on floppy disc; *no* revision will be possible.

Abstract. All papers, reviews and letters begin with an Abstract (50–250 words) which should comprise a factual account of the contents of the paper, with emphasis on new information.

Figures. Figures should be suitable for direct reproduction and as rich in contrast as possible. One original (or sharp glossy print) and two photostat (or other) copies are required. Attention should be given to line thickness, lettering (which should be kept to a minimum) and spacing on axes of graphs, to ensure suitability for reduction in size on printing. Axes of a graph should be clearly labelled, along the axes, outside the graph itself.

All figures should be numbered with Arabic numerals, and require descriptive legends which should be typed on a separate sheet of paper. Simple straight-line graphs are not acceptable, because they can readily be described in the text by means of an equation or a sentence. Claims of linearity should be supported by regression data that include slope, intercept, standard deviations of the slope and intercept, standard error and the number of data points; correlation coefficients are optional.

Photographs should be glossy prints and be as rich in contrast as possible; colour photographs cannot be accepted. Line diagrams are generally preferred to photographs of equipment. Computer outputs for reproduction as figures must be good quality on blank paper, and should preferably be submitted as glossy prints.

Nomenclature, abbreviations and symbols. In general, the recommendations of IUPAC should be followed, and attention should be given to the recommendations of the Analytical Chemistry Division in the journal *Pure and Applied Chemistry* (see also *IUPAC Compendium of Analytical Nomenclature, Definitive Rules*, 1987).

References. The references should be collected at the end of the paper, numbered in the order of their appearance in the text (*not* alphabetically) and typed on a separate sheet.

Reprints. Fifty reprints will be supplied free of charge. Additional reprints (minimum 100) can be ordered. An order form containing price quotations will be sent to the authors together with the proofs of their article.

Papers dealing with vibrational spectroscopy should be sent to: Dr J.G. Grasselli, 150 Greentree Road, Chagrin Falls, OH 44022, U.S.A. Telefax: (+1-216) 2473360 (Americas, Canada, Australia and New Zealand) or Dr J.H. van der Maas, Department of Analytical Molecular Spectrometry, Faculty of Chemistry, University of Utrecht, P.O. Box 80083, 3508 TB Utrecht, The Netherlands. Telefax: (+31-30) 518219 (all other countries).

0003-2670/94/$07.00

Chromatography of Mycotoxins

Techniques and Applications

edited by **V. Betina**

Journal of Chromatography Library Volume 54

This work comprises two parts, Part A: Techniques and Part B: Applications. In Part A the most important principles of sample preparation, extraction, clean-up, and of established and prospective chromatographic techniques are discussed in relation to mycotoxins. In Part B the most important data, scattered in the literature, on thin-layer, liquid, and gas chromatography of mycotoxins have been compiled. Mycotoxins are mostly arranged according to families, such as aflatoxins, trichothecenes, lactones etc. Chromatography of individual important mycotoxins and multi-mycotoxin chromatographic analyses are also included. Applications are presented in three chapters devoted to thin-layer, liquid, and gas chromatography of mycotoxins.

Contents: **PART A. TECHNIQUES.**

1. Sampling, Sample Preparation, Extraction and Clean-up *(V. Betina).* Introduction. Sampling and Sample Preparation. Sample Extraction and Clean-up. Illustrative Example. Conclusions.

2. Techniques of Thin Layer Chromatography *(R.D. Coker, A.E. John, J.A. Gibbs).* Introduction. Clean-up Methods. Normal Phase TLC. Reverse-phase TLC (RPTLC). High Performance Thin Layer Chromatography (HPTLC). Preparative TLC. Detection. Quantitative and Semi-Quantitative Evaluation. Illustrative Examples. Conclusions.

3. Techniques of Liquid Column Chromatography. *(P. Kuronen).* Introduction. Sample Pretreatment. Column Chromatography. Mini-Column Chromatography. High-Performance Liquid Chromatography. Conclusions.

4. Techniques of Gas Chromatography *(R.W. Beaver).* Introduction. Resolution in Gas Chromatography. Extracolumn Resolution. Conclusions.

5. Emerging Techniques: Immunoaffinity Chromatography *(A.A.G. Candlish, W.H. Stimson).* Introduction. Immunoaffinity Chromatography Theory. Practical Aspects and Instrumentation. Sample Preparation. Illustrative Examples.

6. Emerging Techniques: Enzyme-Linked Immunosorbent Assay (ELISA) as Alternatives to Chromatographic Methods *(C.M. Ward, A.P. Wilkinson, M.R.A. Morgan).* Introduction. Principles of ELISA. Sample Preparation. Instrumentation and Practice. Illustrative Examples. Conclusions.

PART B. APPLICATIONS.

7. Thin-Layer Chromatography of Mycotoxins *(V. Betina).* Introduction. Aflatoxins. Sterigmatocystin and Related Compounds. Trichothecenes. Small Lactones. Macrocyclic Lactones. Ochratoxins. Rubratoxins. Hydroxyanthraquinones. Epipolythiopiperazine-3,6-diones. Tremorgenic Mycotoxins. Alternaria Toxins. Citrinin. α-Cyclopiazonic Acid. PR Toxin and Roquefortine. Xanthomegnin, Viomellein and Vioxanthin. Naphtho-γ-pyrones. Secalonic Acids. TLC of Miscellaneous Toxins. Multi-Mycotoxin TLC. TLC in Chemotaxonomic Studies of Toxigenic Fungi. Conclusions.

8. Liquid Column Chromatography of Mycotoxins *(J.C. Frisvad, U. Thrane).* Introduction. Column Chromatography. Mini-Column Chromatography. High Performance Liquid Chromatography. Informative On-line Detection Methods. Conclusions.

9. Gas Chromatography of Mycotoxins *(P.M. Scott).* Introduction. Trichothecenes. Zearalenone. Moniliformin. Alternaria Toxins. Slaframine and Swainsonine. Patulin. Penicillic Acid. Sterigmatocystin. Aflatoxins. Ergot Alkaloids. Miscellaneous Mycotoxins. Conclusions.

Subject Index.

1993 xiv + 440 pages
Price: US $ 180.00 / Dfl. 315.00
ISBN 0-444-81521-X

ORDER INFORMATION

For USA and Canada
ELSEVIER SCIENCE INC.
P.O. Box 945
Madison Square Station
New York, NY 10160-0757
Fax: (212) 633 3880

In all other countries
ELSEVIER SCIENCE B.V.
P.O. Box 330
1000 AH Amsterdam
The Netherlands
Fax: (+31-20) 5862 845

US$ prices are valid only for the USA & Canada and are subject to exchange rate fluctuations; in all other countries the Dutch guilder price (Dfl.) is definitive. Customers in the European Union should add the appropriate VAT rate applicable in their country to the price(s). Books are sent postfree if prepaid.

ELSEVIER SCIENCE B.V.

0003-2670(19940930)296:1;1-#