ELSEVIER

Analytica Chimica Acta 475 (2003) 137–150

ANALYTICA CHIMICA ACTA

www.elsevier.com/locate/aca

Quantifying catecholamines using multi-way kinetic modelling

Rikke P.H. Nikolajsen [a,b,*], Karl S. Booksh [c], Åse M. Hansen [a], Rasmus Bro [b]

[a] *National Institute of Occupational Health, Copenhagen, Denmark*
[b] *Chemometrics Group, Food Technology, The Royal Veterinary and Agricultural University, Copenhagen, Denmark*
[c] *Department of Chemistry and Biochemistry, Arizona State University, Arizona, AZ, USA*

Received 21 December 2001; received in revised form 24 September 2002; accepted 8 October 2002

Abstract

A new method for quantifying adrenaline and noradrenaline concentrations from mixtures of catecholamine standards is described. The method derives selectivity from the different rates, at which the fluorescing 3,5,6-trihydroxyindole derivatives (lutines) of the catecholamines are formed and degraded for adrenaline and noradrenaline. The standards used had the concentration ranges 50–1200 nmol/l for adrenaline and 30–1400 nmol/l for noradrenaline. Fluorescence landscapes were measured at consecutive time points for every sample hereby creating a four-way data array. It is shown that the raw dataset can be dramatically reduced in size without loosing significant information hereby making calculations much faster and lessening instrumental performance requirements. The data follow a two-component four-way parallel factor analysis model (PARAFAC), from which quantitative information is also obtained. Two-component multilinear partial least squares regression (*N*-PLSR) was also employed for the quantification of the catecholamines. The results for PARAFAC and *N*-PLSR were very similar with root mean squared errors of cross-validation (RMSECV) being in the range 24–30 nmol/l. Several improvements of the method are suggested, and it is expected that the method will be suitable for determination of catecholamines in urine from healthy subjects.

Keywords: Catecholamines; PARAFAC; Kinetics; Trihydroxyindole; Urine; Stress

1. Introduction

Urinary catecholamine excretion is a measure of acute stress, and has been used as such in several occupational health investigations, for review see [1]. At present, high performance liquid chromatography (HPLC) methods are typically used for separation of adrenaline (A) and noradrenaline (NA) followed by quantification using fluorescence or electrochemical detection (see e.g. [2–4] and for review see [5]). A representative HPLC method using the lutine reaction as post-column derivatisation reaction [4] has a limit of detection (LOD) of 3.1 nmol/l for adrenaline and 5.2 nmol/l for noradrenaline. Development of a faster and cheaper method for determining adrenaline and noradrenaline in urine would facilitate that these biomarkers are used more often in the investigation and prevention of stress in the working environment. An analytical method for the determination of the catecholamines should firstly be able to distinguish the two analytes, and secondly be able to accurately determine small concentration changes in the urine from healthy subjects. The urinary concentrations in healthy subjects are in the ranges 0–100 nmol/l for adrenaline and 50–500 nmol/l for noradrenaline.

* Corresponding author. Present address: Mikroelektronik Centret (MIC), Technical University of Denmark, Ørsteds Plads, Building 345 Ø, Lyngby DK-2800, Denmark.
E-mail address: rpn@mic.dtu.dk (R.P.H. Nikolajsen).

0003-2670/02/$ – see front matter
PII: S0003-2670(02)01256-4

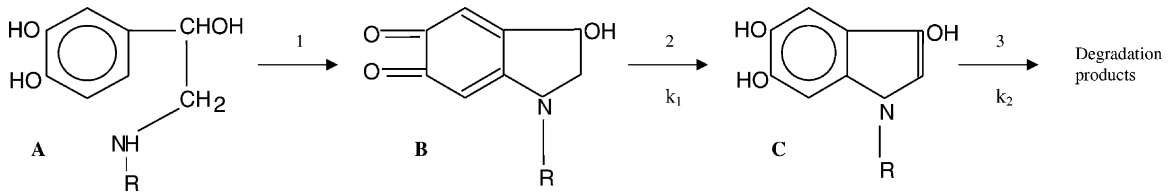

Fig. 1. The lutine reaction: catecholamine (**A**) is oxidised to adrenochrome/noradrenochrome (**B**) and tautomerises to lutine (**C**), which further oxidises to degradation products (*o*-quinones/melanins). R = CH_3 for adrenaline and R = H for noradrenaline. k_1 and k_2 designate the rate constants for reactions (2) and (3).

1.1. The lutine reaction

The fluorophore of adrenaline and noradrenaline is the same, namely catechol ($C_6H_6O_2$), and it has been shown that it is not possible to distinguish the fluorescence spectra of these two compounds [6]. Adrenaline and noradrenaline can be derivatised to the corresponding fluorescing lutines (3,5,6-trihydroxyindole derivatives[1]) (cf. Fig. 1).

As can been seen from Fig. 1, the lutines are intermediates. Adrenolutine and noradrenolutine have slightly different excitation and emission maxima, but more importantly the rates of formation and degradation are different; adrenolutine is formed and degraded faster than noradrenolutine. Several authors (e.g. [4,7–10]) have described the formation of lutines in strongly alkaline media after oxidation of the catecholamines to adrenochrome and noradrenochrome, respectively. Hexacyanoferrate(III) is typically used as oxidising agent [4,8–10], sometimes with zinc sulphate as catalyst [8,9]. The pH is of importance when oxidation takes place, since the rate of formation of noradrenochrome is very slow at acidic pH values. The alkaline solution inducing tautomerisation can be, e.g. strong sodium hydroxide or carbonate buffer. The fluorescence from the lutines rather quickly fades out unless an antioxidant is present; ascorbic acid was often preferred [4,8–10]. Adding the antioxidant at the same time destroys excess oxidising agent. The fluorescent lutines are further degraded (oxidised) to non-fluorescent melanins (if no antioxidant is present this degradation will be fast).

The lutine reaction was first described by Lund [7] in 1949–1950, and he stated “the difference in oxidisability between adrenaline and noradrenaline permits the determination of these two substances in the same sample”. However, most often the lutine reaction has been employed as a selective derivatisation reaction for catecholamines after they were cleaned-up from the urine matrix and separated by HPLC (e.g. [4,10,11]). A few authors have tried quantifying the catecholamines not using a traditional separation technique: Christenson and McGlothlin [12] used derivative fluorescence spectra of the lutines to determine the ratio of adrenaline and noradrenaline. Valcárcel et al. [13] proposed second-derivative synchronous fluorescence spectroscopy to differentiate and quantify the catecholamines in urine; this method was later adopted and automated by Cañizares and Luque de Castro [8]. Llavero et al. [9] utilised the lutine kinetics for quantification using initial rate and fluorescence amplitude. Finally, Wu et al. [14] and Cai et al. [15] employed the lutine reaction (only quantifying adrenaline) as a model system for their method based on detecting the intermediate product in successive reactions [14] and error–compensation algorithm [15]. It was chosen to base the method development described in this paper on the method by Llavero et al. [9].[2] The method presented here should have the advantage of using data with much more information (whole fluorescence landscapes, instead of measurements at one excitation and one emission wavelength) for the determinations implying a more stable method towards interferents, matrix effects and less prone to minor changes in the reaction conditions.

[1] From now on the name *lutines* will be used.

[2] Performance parameters: linear range for A: 0.5–2500 nmol/l and for NA 1–2800 nmol/l; no LOD given; repeatability 0.9% for A and 2.1% for NA (studied on replicates of 24-h urine samples).

1.2. Fluorescence spectroscopy and chemometrics

Fluorescence spectroscopy is a fast, cheap and sensitive analytical method, and is attractive to use in conjunction with chemometric methods. Parallel factor analysis (PARAFAC) [16–18] of fluorescence data has previously been applied for quantification as well as resolution of mixtures with good results [19,20]. Measurements of fluorescence landscapes as a function of time were studied using PARAFAC in work by Tan et al. [21] and Gui et al. [22]. Excitation/emission matrices were resolved by PARAFAC in order to study the kinetic system of simultaneous degradation of chlorophyll a and b extracted from spinach [21]. Kinetic fluorescence detection was employed to determine the concentrations of two overlapped components (glycine and glutamine) after thin-layer chromatographic separation; the data are analysed by trilinear decomposition (the same model as PARAFAC) [22].

Partial least squares regression (PLSR) [23] was also shown to work well for fluorescence data (e.g. [24–26]), when quantitative calibration models are the sole purpose of the analysis. No direct resolution of pure spectra or time profiles is obtained with this approach. Further, *N*-way partial least squares regression (*N*-PLSR) [27] on fluorescence landscapes was shown to have advantages, e.g. in terms of interpretability, over bilinear PLSR. *N*-PLSR was successfully employed for analysing kinetic-spectroscopic multi-way data [28–30].

The aim of this study was to develop an analytical method for mixtures of catecholamines, potentially suitable for urine samples, using lutine kinetics, fluorescence spectroscopy and chemometrics. The first task was to build a suitable semi-automated flow system connected to an excitation emission fluorimeter in order to collect the data. Secondly, to quantify the catecholamines using chemometric models.

2. Theory

2.1. Lutine kinetics

The different rate constants of the formation and degradation of lutines will be employed to differentiate and quantify catecholamines in mixtures. Two consecutive pseudo-first-order reactions are assumed to describe the development in lutine concentrations over time (see e.g. [14]). All reactants are added in excess which means that only the concentrations of adrenochrome and noradrenochrome are rate determining in reaction (2) (Fig. 1), and the lutine concentrations are rate determining in reaction (3) (Fig. 1). Rate expressions for the consecutive first-order reactions have previously been reported by, e.g. Wu et al. [14]. Eq. (1) states the according theoretical expression for the adrenolutine concentration [L] at time '*i*', having a starting concentration of adrenaline $[A]_0$ and the two rate constants $k_{A,1}$ and $k_{A,2}$ (cf. k_1 and k_2 in Fig. 1). An equivalent equation holds for noradrenolutine concentration with respect to $[NA]_0$, $k_{NA,1}$ and $k_{NA,2}$.

$$[L]_i = \frac{k_{A,1}[A]_0}{(k_{A,2} - k_{A,1})}(e^{-k_{A,1}t_i} - e^{-k_{A,2}t_i}) \quad (1)$$

Hence, it is expected that the fluorescence intensity of the lutines is linearly related to the catecholamine concentration. The oxidation of catecholamines to adrenochrome/noradrenochrome (cf. Fig. 1, reaction (1)) is not monitored. This reaction is simply run for a specified time interval for each sample, only changing the catecholamine concentrations, meaning that the same fraction of catecholamines should be oxidised in each run. Excess oxidant is immediately destroyed when the antioxidant is added (with the alkali). The lutines are the only compounds in Fig. 1 fluorescing in the area with excitation range roughly 390–410 nm and emission range roughly 500–530 nm.

2.2. PARAFAC

Fluorescence intensity measurements at several excitation and emission wavelengths can be described by a trilinear PARAFAC model [16–18]. Ideally then, PARAFAC resolves the true emission and excitation spectra of each compound in a mixture, when the correct number of components is chosen for the PARAFAC model. This usually requires that all excitation and emission spectra are linearly independent (i.e. no two are identical in shape) and that the concentration of the analytes vary independently. More relaxed conditions can also be provided, but in general it is problematic if any spectra or concentration profiles are identical [31,32].

Adrenolutine and noradrenolutine spectra and the concentrations chosen in this study fulfil these requirements. The development of the lutine fluorescence landscapes through time was monitored. It was assumed that each analyte has a distinct concentration profile (following Eq. (1)), resulting in the data ideally following a two-component quadrilinear PARAFAC model, written in Eq. (2).

$$x_{ijkl} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} d_{lf} + e_{ijkl} \quad (2)$$

where x_{ijkl} is the fluorescence intensity measured in sample i at emission wavelength j and excitation wavelength k and time l. There are I samples, J emission wavelengths, K excitation wavelengths and L times. F is the number of components and e_{ijkl} holds the residuals of the model. In a quadrilinear model there are four *modes* (in this case: one concentration, two spectral and one time). The model consists of four loading matrices (**A**, **B**, **C** and **D**), with as many columns (here: two) as there are components/analytes, holding the pure spectra/profiles of each mode. It holds that, e.g. **A** is of size $I \times F$ and has typical elements a_{if}.

In traditional second-order calibration, each score vector (column in the sample loading matrix) is taken to be indicative for one analyte and only that specific score vector is used for building the calibration model for that particular analyte (e.g. using univariate linear regression). However, as model errors and similar effects may degrade the performance of the model, it seems reasonable to use all score vectors for building the calibration model for any given analyte. In this particular case meaning that multivariate linear regression using both score vectors is applied for the prediction of each analyte (analogous to principal component regression, PCR). If including adrenaline scores in the prediction of noradrenaline concentration (and vice versa) improves the calibration result, this indicates that there are either chemical interactions between the components or mathematically derived interactions arising e.g. from lack of synchronisation in time curves due to experimental errors.

Hence, having built a PARAFAC model, the sample loadings can be used in a regression model, predicting reference concentrations, hereby producing regression coefficients for the prediction of new samples. In order to predict concentrations in new samples with similar composition (i.e. no additional fluorescing components present), all loadings except the sample loadings are fixed, while recalculating a PARAFAC model for a new sample set of fluorescence measurements. This results in new sample loadings, which can be converted to concentrations by multiplying with regression coefficients from the calibration. The second-order advantage (handling "uncalibrated" interferents) is retained even when new interferents are present during prediction, because the crucial part of the second-order advantage is related to the uniqueness of the PARAFAC model and not the type of subsequent calibration model [33,34]. Imagine, for example, that samples are to be predicted that contain two additional components (here, two additional fluorophores). By fitting a four-component PARAFAC model where the emission, excitation and time profiles of two of the components are fixed at the previously found estimates, then A and NA can be predicted. Hence, the second-order advantage also holds in these situations. As is always the case in second-order calibration, it is an underlying assumption that the loadings, used for calibration, are not influenced by the new interferents.

There are different routes to take for how to actually perform prediction when new interferents are present. As indicated above, the known spectral loadings may be fixed, but the scores of the calibration samples may also be fixed to avoid systematic errors to affect these. Alternatively, it is also feasible to model the data without fixing any parameters. This would enable an exploratory confirmation as to whether the prior model is actually valid. These different possibilities have not yet undergone systematic investigations in the literature.

2.3. N-PLSR

N-way partial least squares regression [27] is an extension of PLSR [23,35] to higher orders. The model for four-way-PLSR is shown in Eq. (3).

$$x_{ijkl} = \sum_{f=1}^{F} t_{if} w_{if}^{J} w_{kf}^{K} w_{lf}^{L} + e_{ijkl} \quad (3)$$

x_{ijkl} is the fluorescence intensity measured in sample i at emission wavelength j and excitation wavelength k and time l. The scalar F is the number of

components and e_{ijkl} holds the residuals of the model (which is not identical to the residual in the PARAFAC model). A slightly modified version of the *N*-PLSR model has been suggested recently [36], but the details of this algorithm will not be described here. It incorporates a so-called core array in the model of the four-way array, but the scores and loadings remain the same and so do the predictions. The algorithm (new as old) finds the scores, *t*, that give maximum covariance with the dependent variables (*y*) in a quadrilinear sense. The result of a four-way-PLSR is four matrices (***T*** holding the scores and three ***W*** matrices holding the weights), each with as many columns as there are PLSR components. The advantage of using *N*-PLSR over unfold-PLSR is a stabilisation of the decomposition, which potentially gives increased interpretability and better predictions.

Eqs. (2) and (3) are algebraically the same, but there is difference in the way the loadings are sought. PARAFAC is a curve resolution method, seeking to explain the variance in the data (in the four-way case as a set of "quadriads") in a least squares sense, without taking the dependent variable, *y*, into account. *N*-PLSR seeks the solution that will give the best prediction of *y*.

3. Experimental

3.1. Flow system

A flow system for running the lutine reaction was constructed using seven metering pumps (model SV653 from Valcor Scientific, http://www.valcor.com/valcor_scientific.htm). The pumps were connected to a valve controller (ValveLink 16, version 2.2, from AutoMate Scientific, http://www.autom8.com), which was linked to a computer. LabView (National Instruments, http://www.ni.com) was used to control the action of each individual pump (when to pump, for how long, at which speed). The pump program for LabView was written at Arizona State University (ASU). A schematic figure of the system is shown in Fig. 2. Manifolds and valves were purchased from Upchurch Scientific (http://www.upchurch.com). Manifold 1 was a seven-port manifold (part no. P-151; ports not used were blocked), while manifold 2 and 3 were three-way T- and Y-connectors (e.g. P-514 and P-713). The three valves (P-721) were manually operated shut-off valves, which ensured that no back-flow occurred. All tubing was Teflon (inner diameter: 0.085 inches–2 mm), the manifolds in PEEK and the valves

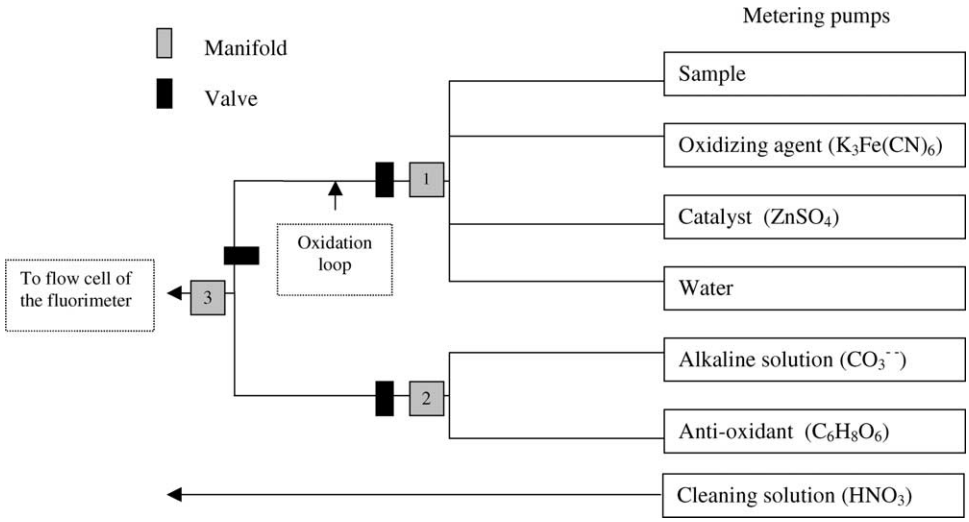

Fig. 2. Schematic illustration of the flow system consisting of seven pumps, three manifolds and three manual shut-off valves (preventing back-flow when pumps are stopped). Since the cleaning solution was strong acid, this pump was connected directly to the flow cell, when filling it with cleaning solution. This procedure ensured that the rest of the flow system would not be not polluted with acid.

Table 1
Solutions for the lutine reaction

Solution	Concentration	Solvent
Catecholamine stock solutions	573.14 μM A 721.04 μM NA	5 mM HCl
Oxidising agent, fresh solution every day	0.0031 M $K_3Fe(CN)_6$	0.1 M KH_2PO_4 adjusted to pH = 7 with 1 M NaOH
Catalyst	0.0017 M $ZnSO_4$	Water
Alkaline solution	1 M CO_3^- pH = 10.0	Prepared from solutions of Na_2CO_3 and $NaHCO_3$, adjusted with 1 M NaOH.
Antioxidant, fresh every day	0.011 M $C_6H_8O_6$	Water
Cleaning solution	About 12% HNO_3	Water

in EFTE[3] material. The total length of coil from sample container to flow cuvette is ca. 100 cm, i.e. about 15 ml volume, including the oxidation loop, which was 4 ml (27 cm). The flow system was connected to a 10 mm × 10 mm quartz flow cuvette[4] (volume ca. 3.5 ml) in the fluorimeter sample chamber. It is of importance that the tubing from manifold 3, Fig. 2 to the flow cuvette is as short as possible, since the lutine reaction starts immediately after mixing.

3.2. Fluorimeter

The measurements were performed on a fluorimeter constructed according to Muroski et al. [37]; later described and used by Jiji et al. [38,39]. The fluorimeter is a so-called single measurement excitation–emission matrix (EEM) fluorimeter, with which it is possible to measure fluorescence landscapes (EEMs) within a second. The light source was a 75 W xenon arc lamp. The excitation grating had 600 grooves/mm and was blazed at 300 nm, and the emission grating had 300 grooves/mm and was blazed at 500 nm. Both excitation and emission slits were 1.0 mm. The detector was a CCD[5] camera (model ST-6 OPTO-HEAD from Santa Barbara Instrument Group, http://www.sbig.com), which was maintained at −20.0 °C by a thermoelectric air cooler. Data collection was controlled by Kestrel Spec ST-6 for Windows, version 3.13 (Catalina Scientific, http://www.catalinasci.com/kestrelspec.htm).

3.3. Chemicals

Chemicals were all analytical grade and used as is without further purification. Milli-Q-water was used for all solutions. The chemicals used were: L-ascorbic acid ($C_6H_8O_6$), zinc sulphate ($ZnSO_4$), potassium hexacyanoferrate(III) ($K_3Fe(CN)_6$), monopotassium phosphate (KH_2PO_4), sodium carbonate, anhydrous (Na_2CO_3), sodium bicarbonate ($NaHCO_3$), sodium hydroxide (NaOH), hydrochloric acid, 37% (HCl) and nitric acid, 65% (HNO_3). About 10 mg of the free bases of the two catecholamines[6] (SIGMA, product numbers: E4250 and A7257) were each dissolved in 100 ml 5 mM hydrochloric acid, producing stock solutions with the concentrations 573.14 μM adrenaline and 721.04 μM noradrenaline. From these two stock solutions all sample solutions are prepared by dilution with 5 mM HCl. Table 1 shows the solutions used for the experiments.

3.4. Measurements

A set of 33 samples of adrenaline and/or noradrenaline dissolved in 5 nM HCl was made from the stock solutions. The concentrations in the samples are presented in Table 2. Adrenaline concentrations ranged from 30 to 1200 nmol/l and noradrenaline

[3] PEEK is polyeretherketone and ETFE (Tefzel) is ethylene-tetrafluoroethylene, both have excellent solvent resistance.
[4] The cuvette had black electrical tape on three corners, and on the side opposite to the in-coming light in order to reduce stray light interferences caused by internal reflection.
[5] CCD: charge-coupled device.
[6] Adrenaline's systematic name is: *d*-1-[3,4-dihydroxyphenyl]-2-methyl-aminoethanole and noradrenaline's: *d*-1-[3,4-dihydroxyphenyl]-2-aminoethanole.

Table 2
Concentrations of adrenaline and noradrenaline in the 33 samples measured

Concentration (nmol/l)		Number of replicates in final dataset of 49
Adrenaline	Noradrenaline	
1000	0	1
800	0	2
500	0	2
80	0	0
0	1400	1
0	1000	2
0	600	2
0	300	2
750	250	2
400	1100	2
250	250	2
50	100	2
100	300	2
400	700	2
900	400	2
200	800	2
500	200	2
350	700	2
400	400	2
200	500	2
200	400	2
600	600	2
700	550	2
60	60	2
300	100	2
60	30	1
1200	600	1
1000	1000	1
30	15	0
850	1200	0
300	50	0
400	200	0
600	800	0

All but two of the samples were measured in replicate, i.e. 64 measurements. Samples containing only one of the analytes are mentioned in the first row. The last column in the table states how many of the replicates measured (usually two) were included in the 'final dataset' with 27 samples (49 measurements) (cf. Section 4.1).

concentrations from 30 to 1400 nmol/l. The 33 samples were measured in random order (each in two replicates, except for two samples) at room temperature (24 +/− 1 °C) in the course of four consecutive days. Measuring the same sample at different days made the replicates. Some white precipitate had previously been observed to build-up on the inside of the cuvette during the time of several runs. Hence, a cleaning procedure was incorporated so that the cuvette was soaked in strong nitric acid during the oxidation, which seemed to prevent noticeable build-up.

In short, the procedure for one measurement was (cf. Fig. 2):

1. Flush the system.
2. Mix sample (catecholamine solution), oxidising agent and catalyst in manifold 1, and pump to oxidising loop (ca. 6 ml of catecholamine solution is used). At the same time cleaning solution is pumped to the flow cell.
3. Stop flow. The oxidation takes place in the oxidation loop during precisely 5 min. Before the 5 min are up, the flow cell is flushed with water and a dark current spectrum is collected.
4. Alkaline solution and antioxidant are mixed in manifold 2 before they are pumped to manifold 3, mixed with the solution from the oxidising loop, and pumped to the flow cell in the course of 10 s (flow rate to cell is ca. 42 ml/min). The fluorescence measurements were started by manually pushing a button when the last pump stroke was heard.

The pump program (the four steps described above) took 350 s after which 60 fluorescence landscapes were collected, which took about 20 min. The integration (exposure) time chosen for each landscape was 15 s, with 1-s delay between readout of the CCD and the beginning of a new exposure; the scan speed was 100 nm/min. Each image was 30 by 250 pixels (binning 3 × 8), which corresponds approximately to the wavelength ranges: 360–420 nm for excitations and 450–610 nm for emissions. Hence, the four-way raw data matrix has the size: 30 excitations and 250 emissions at 60 successive times, for each of the 64 samples/replicates.

3.5. Data analysis

The fluorescence landscapes (image files from Kestrel Spec) were converted to text files using a programme ('missy') developed for this purpose, and the text files were imported into MATLAB™, version 5.3, using a conversion m-file (programs were written at ASU). Design experiments were evaluated and bilinear PLSR calculations performed using The Unscrambler®, version 7.6, Camo Inc. The *N*-way

Toolbox, downloaded from http://www.models.kvl.dk, was used for all other calculations. Data are available at http://www.models.kvl.dk.

4. Results and discussion

4.1. Data pre-treatment

After importing the data into MATLAB™, the dark current spectrum from each measurement was subtracted from the respective fluorescence landscapes (i.e. the same dark current spectrum is subtracted from 60 landscapes). Cosmic ray spikes from the CCD detector were removed by manual inspection of each sample/replicate (data points containing spikes were set to missing; NaN in MATLAB), so they would not disturb modelling. Next, the emission range was truncated, so that only the wavelength area containing chemical information was included. Further, the mean of every three emission points was used, since insignificant information is lost and the data becomes easier to handle computationally. This resulted in a total reduction of the emission mode from 250 to 48 points, and the new emission range: 472–562 nm with steps of approximately 2 nm (the excitation mode is also in steps of 2 nm). In the wavelength ranges thus chosen, no scatter interferences (Rayleigh or Raman) are present.

Contour plots of some raw data (fluorescence landscapes seen from 'above') at different times for a pure adrenaline and noradrenaline sample, respectively are shown in Fig. 3. The adrenolutine concentration increases much faster than noradrenolutine concentration, which is seen when comparing the increase in peak intensities of the left plots with the ones to the right. Also, the concentration of adrenolutine is less 'at 20 min' compared to 'at 10 min', meaning that degradation of adrenolutine has already 'taken over'. It is also noteworthy that the excitation and emission maxima for the two compounds are different. In order to do a preliminary examination of the data, the dataset was unfolded and two-way two-component PLSR1 with adrenaline and then noradrenaline concentration as dependent variable was calculated. The two-component models seem to fit the data, and there is a clear linear relationship between the measured data and concentration of the analytes. However, as will be clear from Section 4.3, the data shows small irregularities, that affect PARAFAC modelling. On account of the plots from PLSR1 models for the whole dataset 15 outliers were removed. Very few of them were severe outliers, and would probably not disturb modelling significantly had they stayed in the dataset. However, in this initial investigation, the possible influence of a potential outlier is uninteresting, and hence removing these 15 measurements ensures that conclusions drawn do not pertain to trivial erratic causes. This resulted in no systematic removal of samples (i.e. not only the low concentration samples or the 'pure' samples were removed), cf. Table 2. Probable causes for potentially outlying samples are bad synchronisation (due to many manual interactions in the measurement procedure), and suspected errors in calculation of a few of the reference concentrations. Further it should be noted that the outlier removal was very conservative, and most of them were found to fit the final model (cf. Section 4.3).

The dataset used for further calculations had 27 samples (5 of which were left in only one replicate) compared to previously 33, and in the reduced dataset, adrenaline concentration range was 50–1200 nmol/l; for noradrenaline 30–1400 nmol/l. The calculations described below have this reduced dataset (30 excitations, 48 emissions, 60 time points and 49 samples/replicates) as "origin". Since the dataset only has 27 samples, no test set validation will be performed; instead segmented cross-validation will be used.

4.2. Data compression

Since the dataset is still of considerable size (ca. 34 MB), it is computationally demanding to fit, e.g. PARAFAC models. It is desirable to see which mode has the greatest influence on prediction power, and how much each mode can be reduced without decreasing the predictive ability (root mean squared error of cross-validation, RMSECV). A Box-Behnken experimental design [40] with three levels of each design variable, one centre sample and no replicates was used to test this. Four design variables were chosen, and this produced a total of 25 experiments. The design variables were:

1. Reduction in emission mode by using every k'th point for modelling ($k = 1, 5, 9$).

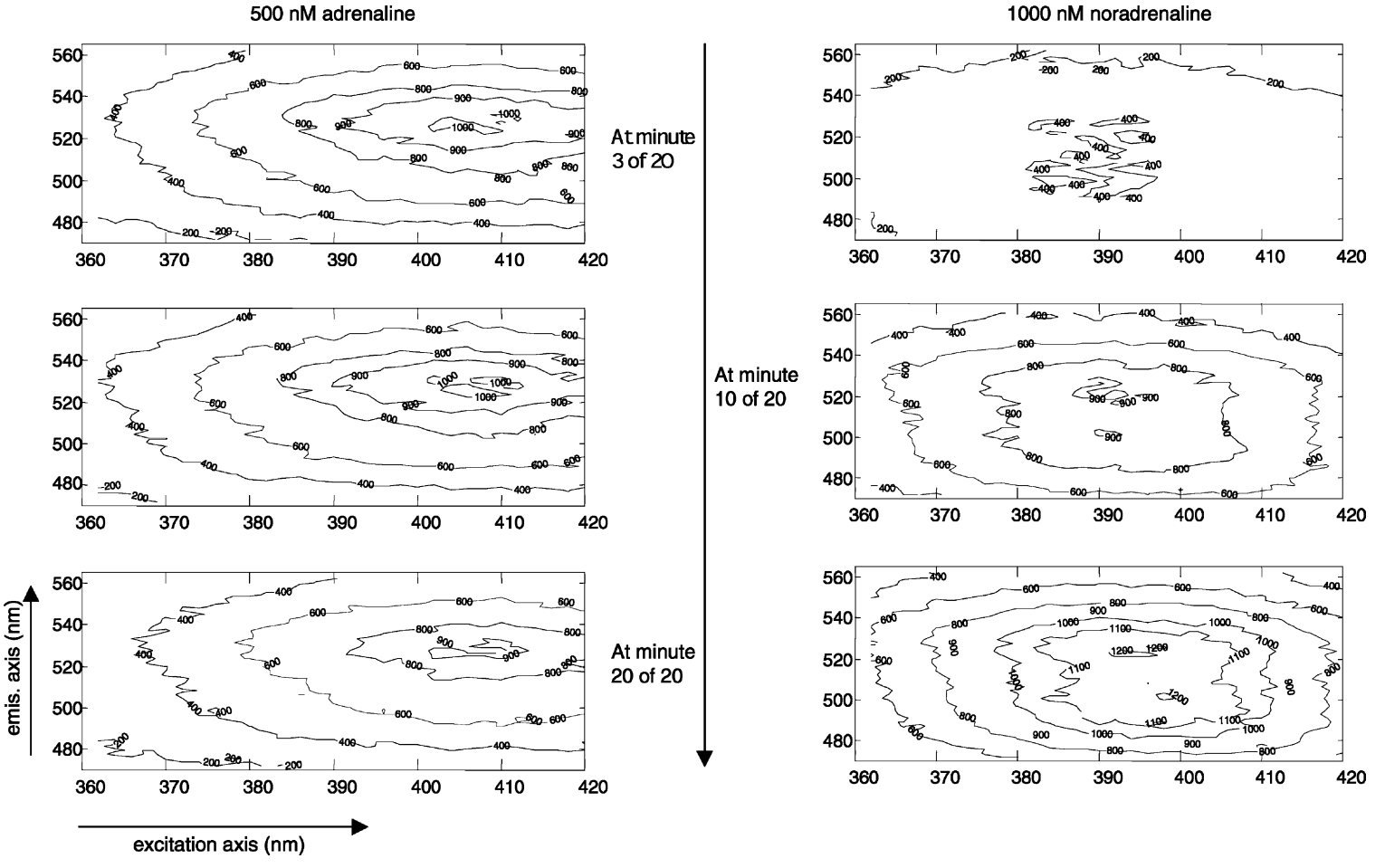

Fig. 3. Contour plots showing the development in lutine fluorescence (i.e. lutine concentration) through time for a 500 nM adrenaline sample (the three plots to the left) and 1000 nM noradrenaline sample (the three plots to the right). The numbers on the contours indicate the fluorescence intensity; vertical axes are emission wavelength (nm) and horizontal axes are excitation wavelengths (nm).

2. Reduction in excitation mode by using every l'th point for modelling ($l = 1, 5, 9$).
3. Reduction in time mode by using every m'th point for modelling ($m = 1, 5, 9$).
4. Another reduction in time mode only using time point 1 to n ($n = 10, 35, 60$). The reduction from point 3 was performed after this truncation of the time mode.

The responses chosen for evaluating the design experiments are RMSECV's (for adrenaline and noradrenaline, respectively) from PARAFAC modelling. The two RMSECV's for each design experiment were calculated on the basis of a two-component four-way PARAFAC model[7] with segmented cross-validation (seven segments; replicates in same segments). The known concentrations were regressed on to the resulting cross-validation scores, and RMSECV obtained from using the regression model on the left-out samples.

The designed experiment showed that design variables 1–3 had about the same influence, and that it is possible to use every fifth measurement in each of the modes and still not loose prediction ability. Truncating the time mode has a higher influence, and shows that using the first few time points is optimal for adrenolutine predictions, and using all time points[8] (or rather the whole time scale) is optimal for noradrenolutine. This is not surprising, considering that adrenolutine is produced much faster than noradrenolutine, and practically only adrenolutine information will be present in the first measurements. The redundancy of information shows that it is possible to lessen the instrumental requirements, meaning that lower resolution in the spectral modes and longer exposure times for the camera is possible. This would mean that the precision of the measurements could be improved (cf. Section 4.5).

Compressing the data by calculating a Tucker3 model (keeping the sample mode intact and reducing the other three modes to few, e.g. seven, variables) was tried as alternative to simple binning. However, this approach was not fruitful because computation of the Tucker3 model was very time consuming.

[7] Loadings from a PARAFAC model with no reduction in any mode were used as starting guesses.

[8] The best compromise is using 45 of the 60 time points—but there is very little difference in using 45 or all 60.

In order to have a dataset of reasonable size for calculations, and still have enough data to be able to visualise the findings so they relate to chemical properties, the following compromise is chosen: all 49 samples/replicates, 12 emission points, 10 excitations and 20 time points spanning the whole time scale (i.e. every fourth emission point and every third excitation and time point from the reduced dataset, cf. Section 4.1). The size of this dataset is less than 1 MB.

4.3. PARAFAC modelling

A two-component four-way model is fitted to the reduced dataset, again using segmented cross-validation (seven segments, replicates in same segment), which results in seven PARAFAC calibration models of different parts of the data. The calibration models converge within 55 iterations, and explain more than 99.5% of the variance in the data. When inspecting the loadings for emission, excitation and time mode, these are practically identical for all seven models, and are shown in Fig. 4. The models showed overlapping but not similar loadings for the two analytes in the fluorescence modes; they showed a steep increase in intensity for adrenolutine, and a much slower increase for noradrenolutine (the general trends can also be compared to raw data in Fig. 3). The residuals for the models are generally random and equally distributed between –100 and 100 fluorescence units. A three-component PARAFAC model was calculated in order to check whether a third component might be needed to explain other variation in the data (e.g. pick up some background variation thereby improving the model). This was not the case and the third component only appeared to describe noise. It seems that the two-component PARAFAC model qualitatively describes the data well. However, there is a systematic model error, which can be dealt with by taking the estimated score values of both analytes into account when predicting the concentration of each of the analytes (cf. next section).

In Eq. (4), the formula for RMSECV is seen; N is the number of samples, y_{ref} the reference values and y_{cv} the predictions from cross-validation.

$$\text{RMSECV} = \sqrt{\frac{\sum_1^N (y_{\text{CV}} - y_{\text{ref}})^2}{N}} \tag{4}$$

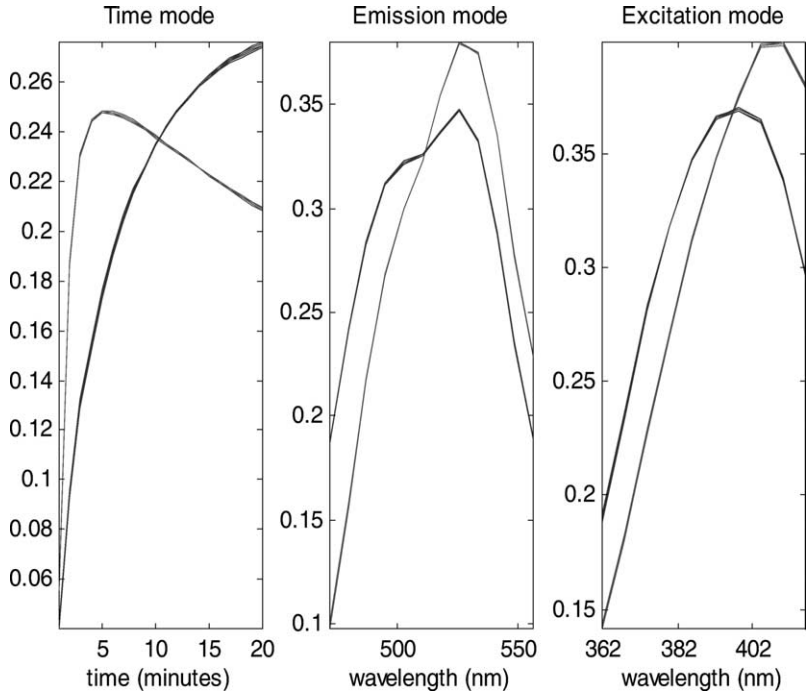

Fig. 4. Loadings from time, emission and excitation mode for the seven PARAFAC models from cross-validation.

The scores from the sample mode of PARAFAC models contain information on relative concentrations (cf. Section 2.2). Hence, it can be expected that each score vector is linearly related to the concentration of one analyte. Table 3 presents the RMSECV's for the model; four different methods for building a calibration model were investigated. First, one score vector was regressed against one analyte's reference concentration (with or without using an offset). Second, both scores were used, i.e. multiple linear regression (MLR) was performed for each of the analytes (with and without offset). From the results in Table 3, it is obvious that both scores are needed for quantifying both components. This can have several causes; one is that when both analytes are present they influence each other's rate constants. This would mean that the consecutive first-order reaction model introduced in Section 2 does not hold. Another explanation is simply that the measurement uncertainty and instrumental deficiencies (e.g. insufficient synchronisation) cause a slight confusion of the contribution from the two analytes to the scores. The second explanation is substantiated, as results very similar to those in Table 3 are obtained when predicting the whole sample set from a PARAFAC calibration model including only samples of pure analytes. Hence, the effect is most likely not due to chemical interactions. In any case, the apparent adequacy of the estimated pure time pro-

Table 3
Root mean squared error of cross-validation (RMSECV) for the two-component PARAFAC model (calculated in four different ways) and for the two-component four-way-PLSR model

Result	RMSECV (nmol/l)	
	Adrenaline	Noradrenaline
PARAFAC results		
One score = one analyte		
Without offset	131.4	79.7
With offset	72.8	74.7
MLR on score matrix		
Without offset	23.4	76.2
With offset	24.2	29.1
N-PLSR result	26.1	29.3

Good agreement is seen in the last two rows, meaning that PARAFAC and *N*-PLSR give similar results.

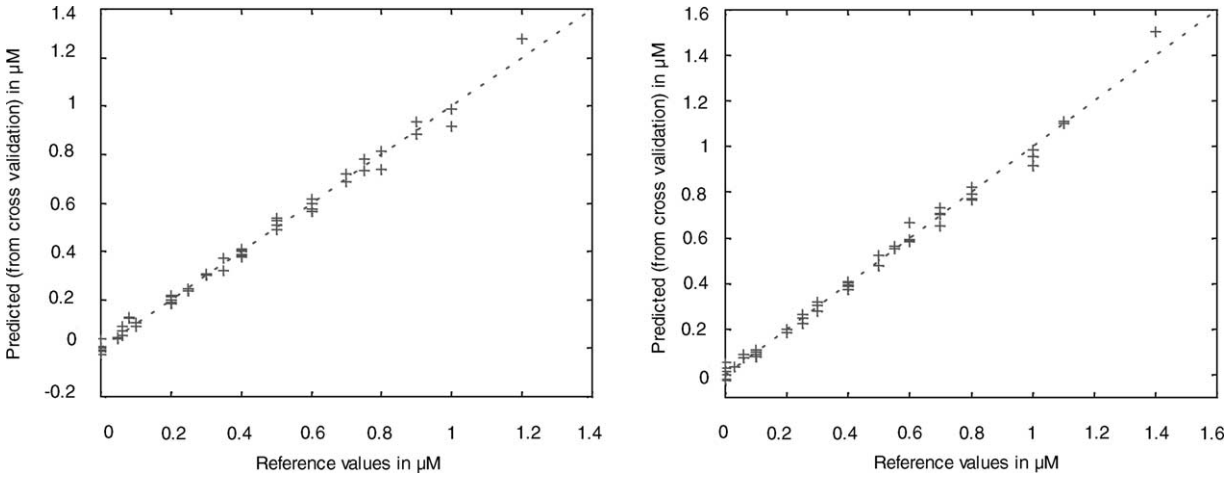

Fig. 5. Predicted (from cross-validation) vs. reference concentrations for *N*-PLSR1 models for adrenaline (left) and noradrenaline (right), respectively. The dotted lines are $y = x$.

files and spectra, as well as the quantitative results, makes it reasonable to accept the model. Further investigations may explain what can be the cause of the systematic error, but it remains that the best RMSECV with PARAFAC is less than 30 nmol/l for both analytes.

The 15 samples, which were removed as possible outliers in the previous calculations (cf. Section 4.1), were predicted using one of the cross-validation models, in order to see if there were indeed severe outliers in the dataset of 33 samples.[9] The root mean squared errors of prediction (RMSEP) calculated after performing MLR (including an offset) on the scores from the 15 samples are 103.9 nmol/l for adrenaline and 45.0 nmol/l for noradrenaline (to be compared with 24.2 and 29.1 nmol/l, cf. Table 3), which especially for adrenaline is a large increase. If instead 59 of the original 64 samples/replicates are predicted, RMSEP for both compounds are practically identical to the ones found for the dataset with 49 samples. Hence, only five measurements[10] do indeed deviate significantly from the others, and are considered as outliers.

[9] The outlying samples are now used as prediction set, and an RMSEP instead of an RMSECV can be calculated.

[10] Actually three whole samples, one of these was only measured in one replicate.

4.4. Comparison with N-PLSR

RMSECV's from two two-component *N*-PLSR1[11] are 26.1 and 29.3 nmol/l for adrenaline and noradrenaline, respectively. In both *N*-PLSR1 models, two components explain more than 99.5% of the variance in the data, and more than 99% variance of the *y*-data. Predicted (from cross-validation) versus reference plots are shown in Fig. 5. The loadings for the *N*-PLSR models are very similar to the PARAFAC loadings. However, loadings from *N*-PLSR models cannot be interpreted as can PARAFAC loadings, since PARAFAC's aim is curve resolution, i.e. the loadings have a 'physical meaning', which *N*-PLSR loadings do not have.

From this it seems that *N*-PLSR and PARAFAC will give similar results (cf. Table 3). However, the PARAFAC model still has the *second-order advantage* compared to the *N*-PLSR models. This means that a PARAFAC model in theory is stable even though the sample matrix might change, which can be expected to be of importance when analysing real urine samples. However, the matrix can be somewhat controlled by using a clean-up procedure (see Section 4.5) that results in a well-defined sample matrix. Thus, the final

[11] For the *N*-PLSR the same cross validation segments as in PARAFAC modelling were used.

choice between *N*-PLSR and PARAFAC has to take future results on real samples and choices with respect to pre-treatment of the samples into account and can thus not be made at this point.

The limit of detection for the presented method can be conservatively estimated on the basis of the RMSECV's in Table 3 to ca. 75 nmol/l for adrenaline and ca. 90 nmol/l for noradrenaline (LOD was calculated as $3 \times$ S.D.; RMSECV is an estimate for the mean standard deviation (S.D.)). Probably a more realistic value would be $3 \times$ S.D. for samples with low concentration. LOD estimated on the basis of two samples[12] is about 50 nmol/l for adrenaline and 70 nmol/l for noradrenaline (using *N*-PLSR predictions). Since the method needs to be improved to compete with existing methods (cf. Section 4.5), the LOD presented here is not nearly as low as can be expected for the optimised method. LOD is usually calculated using definitions from IUPAC (LOD is '3 standard deviations of 20 measurements of a low concentration sample') or ISO (guide no. 111843-1 and 111843-2). Using one of these definitions would be crucial for thorough evaluation of the improved method.

4.5. Practical considerations and future perspectives

The results show that the method still needs to be improved. The goal for the development of this method is to be able to measure samples in the concentration range 0–1000 nmol/l for each of the analytes with a capability of detection (or limit of detection) in the range 1–7 nmol/l. The relative uncertainty on a measurement should of course be as small as possible and at least less than 10% (which was the uncertainty found for the method by Hansen et al. [4]). This means that the present method needs to have an analytical range and a limit of detection about one order of magnitude lower than now, and the uncertainty of measurements need to be improved. Indeed, there are several realistic approaches to this end.

Since this is a kinetic method, timing is crucial. Generally, a higher degree of automation in the system and more precise pumps, and thereby better control of timing, is expected to give a quite dramatic improvement of the results. A temperature control for the whole flow system would also be preferable, since both kinetics and fluorescence measurements are highly temperature dependent. When the flow system has been optimised, the reaction conditions will of course have changed, and an optimisation of the chemical conditions (relative concentrations, reaction temperature, etc.) is necessary. The capability of detection might also be improved by increasing the exposure time for the CCD camera, since the designed experiment (Section 4.2) showed that as little as every fifth time point can be used without loosing valuable information. Additionally, using a better, hence more expensive, CCD detector will improve sensitivity further.[13] Finally, the use of a fluorescent standard would enable correction for changes in lamp intensity and detector response. The above mentioned points can be overcome by investing in new equipment, leaving only the problem of cleaning up urine samples for the procedure. There are numerous procedures developed for this purpose, which are presently used in the different HPLC analysis available. It should be possible to choose a suitable clean-up procedure and incorporate it in the flow system, so the method would be fully automated. Since the lutine reaction is specific for adrenaline and noradrenaline, it would probably be enough just to remove fluorescence background interference in the measurement area. At present, the speed of analysis is about the same as for most HPLC methods (about 1/2 h per sample), but it is expected that the analysis time can be reduced considerably, as there are several parameters that can be adjusted (e.g. speed of the kinetic reaction can be controlled with temperature, and maybe further truncation of the time mode will prove to be possible in a new system set-up). With the above mentioned improvements it is likely that this method will be able to compete with the most accurate of the present HPLC methods, both regarding speed of analysis and analytical performance parameters.

5. Conclusions

A method for determining adrenaline and noradrenaline concentrations from mixtures of cate-

[12] Concentrations (in nmol/l) in the two samples were: 50 A + 100 NA and 60 A + 60 NA.

[13] The CCD camera presently used is really designed for amateur astronomy, not for scientific spectroscopy.

cholamines was described. *N*-PLSR and PARAFAC models give similar results, RMSECV's being in the range 24–30 nmol/l. The raw data matrix can be dramatically reduced in size without loosing information. Several improvements of the method are possible, and it is expected that the method will be suitable for determination of catecholamines in urine from healthy subjects.

Acknowledgements

Michelle L. Nahorniak, Department of Chemistry and Biochemistry, Arizona State University (ASU), USA is thanked for helpful discussions and suggestions, and for writing the conversion m-file used. Brian K. Dable, Department of Chemistry and Biochemistry, ASU wrote the LabView programme used. The programme ('missy') was developed by Nathan Wilken, Senior Application System Analyst, Department of Information Technology, ASU. Rikke Nikolajsen furthermore wishes to thank Knud Højgaards Foundation, DK; the Danish Research Training Council (formerly: The Danish Research Academy) and the Danish National Institute of Occupational Health for funding. Rasmus Bro is grateful for financial support through LMC (Center for Advanced Food Studies) and AQM (Advanced Quality Monitoring) supported by the Danish Ministries of Research and Industry.

References

[1] J.K. Sluiter, M.H. Frings Dresen, T.F. Meijman, A.J. van der Beek, Occup. Environ. Med. 57 (2000) 298.
[2] G. Grossi, A.M. Bargossi, C. Lucarelli, R. Paradisi, C. Sprovieri, G. Provieri, J. Chromatogr. 541 (1991) 273.
[3] C. Sarzanini, E. Mentasi, M. Nerva, J Chromatogr. 671 (1994) 259.
[4] Å.M. Hansen, J. Kristiansen, J.L. Nielsen, K. Byrialsen, J.M. Christensen, Talanta 50 (1999) 367.
[5] R.P.H. Nikolajsen, Å.M. Hansen, Anal. Chim. Acta 449 (2001) 1.
[6] R.P.H. Nikolajsen, Å.M. Hansen, R. Bro, Luminescence 16 (2001) 91.
[7] A. Lund, Acta Pharmacol. 6 (1950) 137.
[8] P. Cañizares, M.D. Luque de Castro, Anal. Chim. Acta 317 (1995) 335.
[9] M.P. Llavero, S. Rubio, A. Gomez-Hens, D. Perez Bendito, Anal. Chim. Acta 229 (1990) 27.
[10] K.S. Boos, B. Wilmers, R. Sauerbrey, E. Schlimme, Chromatographia 24 (1987) 363.
[11] P.O. Edlund, D. Westerlund, J. Pharm. Biomed. Anal. 2 (1984) 315.
[12] R.H. Christenson, C.D. McGlothlin, Anal. Chem. 54 (1982) 2015.
[13] M. Valcárcel, A. Gomez-Hens, S. Rubio, Clin. Chem. 31 (1985) 1790.
[14] X. Wu, R. Cai, Z. Lin, J. Cheng, Anal. Chim. Acta 325 (1996) 105.
[15] R. Cai, X. Wu, Z. Liu, W. Ma, The Analyst 124 (1999) 751.
[16] R.A. Harshman, M.E. Lundy, Comput. Stat. Data Anal. 18 (1994) 39.
[17] S.E. Leurgans, R.T. Ross, Stat. Sci. 7 (1992) 289.
[18] R. Bro, Chemom. Intell. Lab. Syst. 38 (1997) 149.
[19] K.S. Booksh, A.R. Muroski, M.L. Myrick, Anal. Chem. 68 (1996) 3539.
[20] R. Bro, Chemom. Intell. Lab. Syst. 46 (1999) 133.
[21] Y. Tan, J.H. Jiang, H.L. Wu, H. Cui, R.Q. Yu, Anal. Chim. Acta 412 (2000) 195.
[22] M. Gui, S.C. Rutan, A. Agbodjan, Anal. Chem. 67 (1995) 3293.
[23] H. Martens, T. Næs, Multivariate Calibration, Wiley, Chichester, 1989.
[24] R.D.B. Jimenez, A.I.J. Abizanda, F.J. Moreno, J.J.A. Leon, Clin. Chim. Acta 249 (1996) 21.
[25] C. Leal, M. Granados, J.L. Beltran, R. Compano, M.D. Prat, The Analyst 122 (1997) 1293.
[26] Z.-P. Wang, L.-L. Shi, G.-S. Chen, K.L. Chen, Talanta 51 (2000) 315.
[27] R. Bro, J. Chemom. 10 (1996) 47.
[28] R. Bro, H. Heimdal, Chemom. Intell. Lab. Syst. 34 (1996) 85.
[29] S.R. Crouch, J. Coello, S. Maspoch, M. Porcel, Anal. Chim. Acta 424 (2000) 115.
[30] J. Coello, S. Maspoch, N. Villegas, Talanta 53 (2000) 627.
[31] N.D. Sidiropoulos, R. Bro, J. Chemom. 14 (2000) 229.
[32] J.B. Kruskal, in: R. Coppi, S. Bolasco (Eds.), Multiway Data Analysis, Elsevier, Amsterdam, 1989, p. 8.
[33] E. Sanchez, B.R. Kowalski, J. Chemom. 4 (1990) 29.
[34] K.S. Booksh, Z. Lin, Z. Wang, B.R. Kowalski, Anal. Chem. 66 (1994) 2561.
[35] S. Wold, P. Geladi, K. Esbensen, J. Öhman, J. Chemom. 1 (1987) 41.
[36] R. Bro, A.K. Smilde, S. de Jong, Chemom. Intell. Lab. Syst. 58 (2001) 3.
[37] A.R. Muroski, K.S. Booksh, M.L. Myrick, Anal. Chem. 68 (1996) 3534.
[38] R.D. Jiji, G.A. Cooper, K.S. Booksh, Anal. Chim. Acta 397 (1999) 61.
[39] R.D. Jiji, G.G. Andersson, K.S. Booksh, J. Chemom. 14 (2000) 171.
[40] D.C. Montgomery, Design and Analysis of Experiments, Wiley, New York, 1991.