

Robust DOSY NMR data analysis

R. Huo, R. Wehrens, L.M.C. Buydens *

Institute for Molecules and Materials, Analytical Chemistry, Radboud University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands

Received 2 January 2006; received in revised form 27 February 2006; accepted 20 March 2006

Available online 2 May 2006

Abstract

Multivariate methods based on curve resolution outperform single channel methods in the analysis of diffusion-ordered spectroscopy (DOSY) NMR data in terms of accuracy and ease of interpretation, especially for systems with large spectral overlap. In this paper, the focus is on the robustness of two multivariate methods, classical multivariate curve resolution (MCR) and MCR combined with non-linear least squares regression (MCR–NLR). Three important factors that influence the analysis are investigated: Peak shifts, phase shifts, and the difference of diffusion coefficients. Using controlled disturbances of a data set of a mixture of three discrete components, ATP, glucose, and SDS, it is shown that both multivariate methods outperform SPLMOD, one of the standard single channel methods. In particular, MCR–NLR is the more accurate as well as the more robust of the two multivariate methods. This implies that MCR–NLR is less sensitive to the quality of the data and may give good results in cases where other methods fail.

© 2006 Elsevier B.V. All rights reserved.

Keywords: DOSY; Multivariate curve resolution (MCR); Robust analysis; Experimental design

1. Introduction

Several forms of multivariate curve resolution (MCR) have been used to resolve DOSY NMR data [1–3]. It has been proven that MCR has the advantages of dealing with non-uniform gradients and overlap peak, as well as potentially handling polydisperse samples, due to the fact that MCR has no assumption of exponential decay. Previous research has shown that a combination of MCR and non-linear least squares regression (MCR–NLR) can yield better separation and interpretation of DOSY data [3] as it can effectively remove the rotational ambiguities of MCR. The development of the MCR–NLR algorithm can be regarded as improvement of classical MCR applying only non-negativity constraints. Apart from the hard model using exponential profile (MCR–NLR), there are other ways to put chemical information in the modelling, e.g. using other constraints in the MCR, such as unimodality (one maximum per decay profile) and local rank conditions (setting the zones on the chemical shift dimension where one component is present) [4]. However, in the case of DOSY data, we will only

investigate two extremes; a hard, mono-exponential model for the decay profiles and a completely free form of the decays.

For users of DOSY NMR, it is necessary to provide them with guidelines of situations in which the processing methods can be expected to work well. It is therefore essential to investigate how the performance of MCR–NLR and classical MCR for DOSY data reacts to variations in parameter values, e.g. the data quality and the difference of diffusion coefficients of the components in a mixture. Also, it is useful to compare single channel methods with the two multivariate techniques. This can be achieved by robustness tests [5,6].

Robustness tests have been widely applied in pharmaceutical industry for method validation, due to the strict requirement of regulatory authorities [6]. According to the guidelines presented by the International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use (ICH), “the robustness of an analytical procedure is a measure of its capacity to remain unaffected by small, but deliberate variations in method parameters and provides an indication of its reliability during normal usage” [7]. Therefore, a robustness test of a method is an experimental set-up to evaluate to what extent a (analytical) method is sensitive to small changes in different laboratories or under different circumstances. Usually, two-level factorial designs are carried out to determine whether

* Corresponding author. Fax: +31 24 3652653.

E-mail address: L.Buydens@science.ru.nl (L.M.C. Buydens).

factors or interactions between factors have significant effects on the response. However, they are only suitable for cases that the factors and the responses are linear related. When quadratic relations have to be considered, multi-level designs need to be applied [8].

There are several sources that can affect the performance of the DOSY data processing methods. In DOSY NMR experiments, the quality of DOSY data can be different if measured by different laboratories or NMR machines, e.g. containing different amounts of peak and phase shifts, and random noise [9]. This will have an influence on the resulting pure spectra and decay profiles, depending on how sensitive the processing methods are to the experimental artefacts. In addition, differences between the diffusion coefficients of the components may be too small to allow individual spectra to be resolved. In the literature, it has been stated that the routine data processing methods implemented in commercial NMR software, i.e. single channel methods, require the difference of diffusion coefficients to be at least a factor of two [10]. Therefore, the influence of this on the performance of the two MCR methods should also be investigated. In addition, the spectral overlap is an important factor in DOSY data analysis. However, this factor will not be considered here since our primary aim was to compare the two MCR methods for a given set of spectra. It should be included in future studies, although it may be difficult to “control” the degree of overlap in simulations. In this paper, the robustness of two multivariate methods for DOSY data processing, classical MCR and MCR–NLR, is examined by performing an experimental design. Three factors, phase shifts, peak shifts, and differences of diffusion coefficients, are considered. Realistic data sets will be constructed with different values for these three factors. The responses used will be the errors in both the calculated pure spectra and the decay profiles. Because there can be a quadratic relationship between the responses and the factors, a central composite design is employed [11]. This is a powerful multi-level design, specifically developed for constructing a second-order polynomial model. From the response surface of the model, one can search for the boundaries of the three factors within which the multivariate methods can yield good results. Extra data sets, constructed with the three factors within these boundaries, will be investigated, to reveal the difference in performance of the multivariate methods and SPLMOD, one of the routine single channel methods.

2. Theory

2.1. Processing DOSY data

The principles of processing DOSY NMR by single channel methods and the two MCR methods have been described elsewhere [2,3,10]. Only a brief review will be given here. The intensities of the DOSY NMR signals decay exponentially,

$$I(g^2) = \sum_{n=1}^N I_0(n) \exp[-D(n)(\Delta - \delta/3)K^2] \quad (1)$$

where N is the number of components in the sample, and $D(n)$ is the diffusion coefficient of the n th component (m^2/s). δ is the

duration of gradient pulses (s) and Δ is the diffusion time (s), both of which are experimental constants. K is the product of δ , γ , the gyromagnetic ratio of the ^1H nucleus ($\text{rad s}^{-1} \text{T}^{-1}$), and g , the gradient strength (T m^{-1}). $I_0(n)$ is the intensity at a gradient strength g of zero. Single channel methods apply mono- or multi-exponential fitting for each channel of the spectra to obtain diffusion coefficients for all components, according to Eq. (1). SPLMOD is a single channel method that analyse sums of pure exponentials by performing a least squares fit of Eq. (1) [12].

A DOSY data set is bilinear and hence it can also be analysed by MCR [13], which decomposes the data into two sub-matrices, C and S :

$$I = C \cdot S^T + E \quad (2)$$

C represents the exponential term of Eq. (1) that can be regarded as the matrix containing pure decay profiles, and S represents the intensity before gradients are applied, and contains the pure spectra. Consequently, MCR attempts to recover the pure components from a DOSY data set. This is achieved by the alternating least square (ALS) algorithm.¹² ALS is started with an estimate of pure decay profiles, e.g. obtained by the orthogonal projection approach (OPA) [14]. These are used to calculate estimates for the pure spectra:

$$S = I^T \cdot C \cdot (C^T \cdot C)^{-1} \quad (3)$$

New decay profiles can be calculated from the estimated S

$$C = I \cdot S \cdot (S^T \cdot S)^{-1} \quad (4)$$

The steps as described in Eqs. (3) and (4) iterate until the S and C reach convergence. Non-negativity constraints are applied in each ALS iteration step to restrict the calculated pure spectra and decay profiles to be positive. These constraints help to provide less ambiguous solutions that are chemically and physically meaningful. The procedure of ALS with non-negativity constraints is called classical MCR in this paper. The algorithm of MCR–NLR mainly follows the same procedure as classical MCR. The difference is that a nonlinear least square regression (NLR), e.g. the Levenberg–Marquardt algorithm [15], is applied to the calculated decay profiles obtained in each iteration, so that the decay profiles are forced to follow an exponential decay. The main drawback of classical MCR is that, despite the non-negativity constraints, there are many possible solutions fulfilling the convergence criterion and hence the iteration steps can end up with incorrect pure spectra and decay profiles. Embedding NLR into MCR can yield unique solutions.

2.2. Robustness tests

To perform robustness tests for evaluation of classical MCR and MCR–NLR methods, the first step is to select the factors that can have an impact on the results of the calculated pure spectra and decay profiles. The second step is to select the suitable experimental design. To find the boundaries of the factors within which the two methods are robust, a least squares model is built

to describe the relationship between the responses and the factor levels. Two-level designs can estimate the effect of the factors and their interactions. However, these designs can describe only linear model but do not consider possible quadratic effects. Multi-level designs can build a second order model, i.e. a curvilinear relationship is considered. A central composite design [8,11] is one of the economical multi-level designs to investigate the relationship between the factors and the responses. In this paper, three factors will be considered, so a three factor central composite design is applied. As presented in Table 1, it contains three parts. The first part is a two-level full factorial design, i.e. eight (2^3) trials to construct the data sets are included. These are the first eight experiments in Table 1. The second part is called star design, in which more levels are used to describe curvature of the responses. There are six trials for the star design, i.e. from 9 to 14. The third part is the central point, with all the factors at the nominal levels (trial 15) and it is normally replicated. Table 1 also presents the factor levels. The nominal level is described as “0”. The high level and low level, indicated by “−1” and “+1”, respectively, are considered to be located at a distance of 1 from the nominal level. For a three-factor central composite design, particularly, the start points are situated symmetrically at a distance of 1.682 from the nominal level [8], so these levels are represented by “−1.682” and “+1.682”.

The third step is to determine the responses. For evaluating the robustness of classical MCR and MCR–NLR, responses describing the performance of the two methods are the accuracy of the resolved spectra and decay profiles for the components in a mixture. Therefore, the responses are the chi-square error, χ^2 , measuring for the agreement between each calculated and true spectrum [16], and the relative difference between calculated and true diffusion coefficients E_d (%).

$$\chi^2 = \sum (S_{\text{cal}} - S_{\text{true}})^2 / S_{\text{true}} \quad (5)$$

$$E_d\% = (d_{\text{cal}} - d_{\text{true}}) / d_{\text{true}} \times 100\% \quad (6)$$

Both errors should be as close to zero as possible. Because some of the points in the true spectra contain zero values, an

Table 1
A three factor central composite design

Trial	Phase shifts (x_1)	Peak shifts (x_2)	Difference of D (x_3)
1	−1	−1	−1
2	+1	−1	−1
3	−1	+1	−1
4	+1	+1	−1
5	−1	−1	+1
6	+1	−1	+1
7	−1	+1	+1
8	+1	+1	+1
9	−1.682	0	0
10	+1.682	0	0
11	0	−1.682	0
12	0	+1.682	0
13	0	0	−1.682
14	0	0	+1.682
15	0	0	0

offset is added to the denominator in Eq. (5) to avoid zero division. In principle, the offset can be any number since in general it will not change the shape of a chi-square response surface but will only affect the scale. Here, a small offset of 0.1 is chosen, in order to yield the results that are chi-squares easily comparable to the traditional correlation coefficient. Therefore, the χ^2 is calculated by the following equation,

$$\chi^2 = \sum (S_{\text{cal}} - S_{\text{true}})^2 / (S_{\text{true}} + 0.1) \quad (7)$$

After the responses are obtained, a quadratic model with three factors is constructed:

$$y_{\text{response}} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{11}x_1^2 + b_{22}x_2^2 + b_{33}x_3^2 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 \quad (8)$$

In the equation, the b coefficients estimate measures of the main effects of the factors (b_1, b_2, b_3), the quadratic effects (b_{11}, b_{22}, b_{33}) and the interaction effects between the factors (b_{12}, b_{13}, b_{23}). x_1, x_2 , and x_3 represent the corresponding factor levels for peak shifts, phase shifts and difference of diffusion coefficients, respectively. The b coefficients can be calculated by multiple linear regression:

$$b = (X^T X)^{-1} X^T y \quad (9)$$

In this equation, X is a matrix containing the factor levels in each of the simulation. Hence, there are 15 rows for the simulations and 10 columns for the terms of the quadratic equation (Eq. (8)) in X matrix. Since an advantage of a experimental design is that the variables in a predictor matrix (X) are not correlated, the estimates for b coefficients do not influence each other. Consequently, testing significance of a factor can be achieved by testing each b coefficient separately. The coefficients are considered to be significant if the confidence interval does not include zero. After the b coefficients have been evaluated, a model will be refitted with only the significant b coefficients. Consequently, a series of plots of the factors against the fitted response can be created. Given a critical value for the y_{response} in the response plots, the boundary of the factors for robustness areas can be defined.

3. Experimental

3.1. Execution of robustness tests

Since previous research indicates that phase and peak shifts are generally present in experimental DOSY data [17], we will investigate how the phase and peak shifts affect the results obtained from MCR and MCR–NLR, respectively. Also, the difference between diffusion coefficients can influence the results and it will be used as the third factor for this robustness test. The level of the factors should be chosen symmetrically around the nominal level, which should correspond to a normal good-quality data set. Here, we simulated data in which phase

shifts and peak shifts were the same for each peak within the spectrum for a single gradient level but randomly different between spectra of different gradient levels in a single DOSY experiment. According to Table 1, there are 15 data sets with different factor levels that are simulated for a central composite design. To obtain more reliable results, each simulation is duplicated and therefore there are 30 data sets in total. Each of the data set is analysed by classical MCR and MCR–NLR, respectively. The two responses, χ^2 for the calculated spectra and the relative error for the calculated diffusion coefficients, are estimated. It is also possible to consider the recovery of the decay profiles rather than the errors of the diffusion coefficients. However, the diffusion coefficient is — for users — the more informative parameter. Finally, a quadratic model is constructed and the significance of the effects is evaluated.

3.2. Data

The data for the robustness test are constructed using real-life pure spectra and decays, obtained from an experimental DOSY data set of a mixture containing three main components, 50.0 mM adenosine 5'-triphosphate (ATP), 50.0 mM D-(+)-glucose, and 25.0 mM Sodium Dodecyl Sulfate (SDS) in D₂O. The experimental DOSY data set was recorded by referring to the experiment as being done in Ref. [10]. The data set were recorded on a Bruker Avance 500 MHz spectrometer. The maximum gradient applied is 60 G/cm. The diffusion time (Δ) is 0.4 s and the duration of gradient pulses (δ) is 3.4 ms. The original experimental data set contains 64 decaying ¹H NMR spectra with K values ranging from 3.90×10^{11} to $9.75 \times 10^{14} \text{ m}^{-2}$. In each spectrum there are 16384 data points in the chemical shift dimension of a spectral width from -0.55 to 11.3 ppm. Since the first few spectra recorded with low gradients are very likely to be distorted, the first 10 spectra are left out. In addition, the water signal in each spectrum is replaced by random noise and one of every two

spectral points is used to construct the spectra. Consequently, the experimental data set, used to find the pure spectra, has the size of 54×8192 and three main components, ATP, glucose, and SDS micelle. A 3-D stacked plot of the data set is depicted in Fig. 1. The first spectrum of this data set is shown in Fig. 2A. To obtain the best pure spectra of the components, the data set was first preprocessed to correct for phase and peak shifts, and lineshape distortion [3]. Then the pure spectra and decay profiles were calculated from the experimental data using the MCR–NLR method. The results are shown in Fig. 2B and C. The resolved pure spectra of the three components, ATP, glucose, and SDS, are in very good agreement with the corresponding true spectra (see Fig. 2B). Therefore, these calculated pure spectra are used to construct the data sets containing different levels of phase and peak shifts, as well as different diffusion behaviour of the components for the robustness test. It must be emphasised that any set of pure spectra can be used to simulate the required data sets, and that it does not matter how they are obtained. Moreover, each simulated data set also contains white noise with the standard deviation of 0.03% of the highest peak intensity.

The factor and the factor levels for the composite design are presented in Table 2. The construction of the data sets with different levels of the factors includes several steps. First, a reference peak is selected that is present in the spectra with all gradients. Random peak and phase shifts, according to the corresponding level of the experimental design, are added to this peak [18]. Next, this peak shift and phase shift are transferred to all other peaks in the same spectra, using FIDDLE [19]. Note that here FIDDLE is used to add these shifts, instead of removing them. Thus, all peaks in a spectrum contain the same peak and phase shifts as the selected peak but the shifts in individual spectra are different. For this data set, the peak between 0.4 and 0.8 ppm (with 190 spectral points), is selected as the reference peak (see Fig. 2A). To create different decay behaviour for the components, we use a fixed diffusion coefficient of

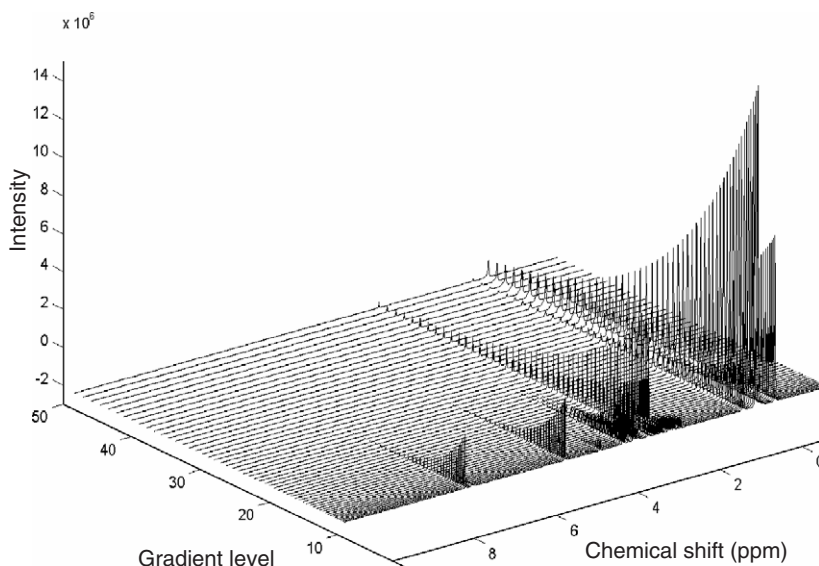


Fig. 1. A 3-D stacked plot of the DOSY NMR data set.

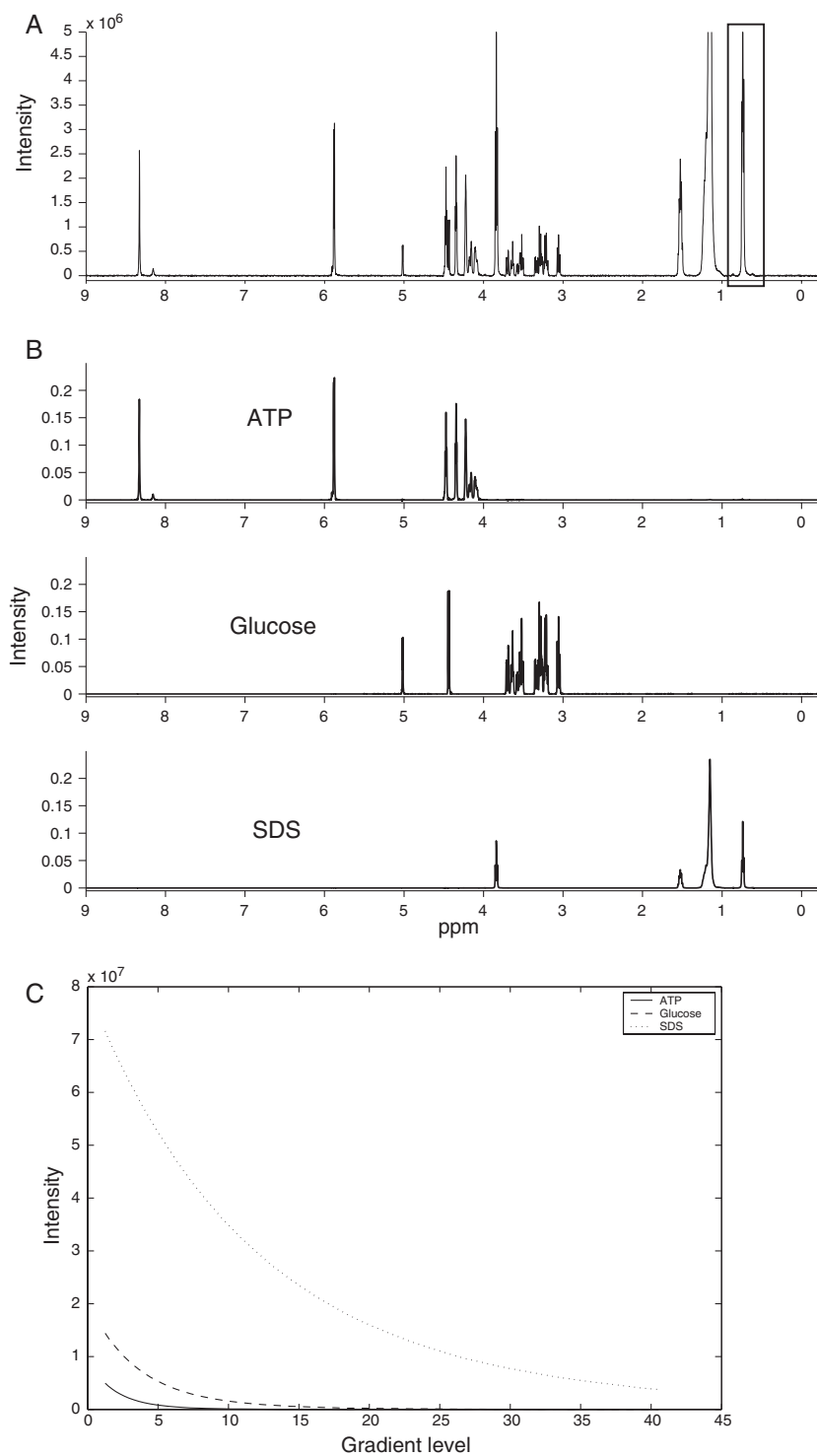


Fig. 2. (A) The first spectrum of the experimental data set for the mixture (the highlighted peak is used to create phase and peak shifts); (B) pure spectra of the three components and (C) the corresponding decay profiles obtained from the experimental data by MCR–NLR.

$2.72 \times 10^{-12} \text{ m}^2/\text{s}$ for SDS and change the diffusion coefficients of the other two components with the corresponding different factors. The nominal level for the difference between pairs of diffusion coefficients is a factor of 1.8, so the diffusion coefficients of ATP, glucose, and SDS at the nominal level are 8.83×10^{-12} , 4.91×10^{-12} , and $2.72 \times 10^{-12} \text{ m}^2/\text{s}$, respectively. The other data sets are constructed in the same way.

3.3. Software

All calculations are done by using MATLAB 7.0 on a SUN UNIX workstation. The results obtained from single channel methods are calculated by in-house modifications of SPLMOD. The classical MCR is the MCR function in PLS-Toolbox 2.1 [20], and the MATLAB code for the MCR–NLR algorithm is available

Table 2
Factors and factor levels to examine classical MCR and MCR–NLR

	Star point (–1.682)	Low (–1)	Nominal (0)	High (+1)	Star point (+1.682)
Phase shift (x_1)	–0.0917– 0.0917°	–0.287– 0.287°	–0.573– 0.573°	–0.859– 0.859°	–1.054– 1.054°
Peak shift (x_2) data point	–0.0023– 0.0023	–0.050– 0.050	–0.120– 0.120	–0.190– 0.190	–0.238– 0.238
Difference D factor (x_3)	2.305	2.100	1.800	1.500	1.292

on our website: <http://www.cac.science.ru.nl>, in the package MULVADO.

4. Results and discussions

4.1. Responses of the MCR and MCR–NLR methods

The χ^2 values of the calculated spectra and the relative errors of the calculated diffusion coefficients of the three components are presented in Fig. 3. The error bars represent the difference of the two replicate simulations; the mean values of the replicates are connected by lines to distinguish between MCR and MCR–NLR. The numbers at the x -axis correspond with the experiments for Table 1. Fig. 3A and B display the errors for resolving ATP, the first component, by classical MCR and MCR–NLR. From Fig. 3A, one can see that both classical MCR and MCR–NLR can achieve quite low values of χ^2 for the first five simulations. The resulting χ^2 of experiment 6 to 8 from the two methods are quite different. It can be seen that classical MCR results in higher chi-square values than MCR–NLR. This is also the case for experiment 10 and 14. All these simulated data have either a high level of experimental artefacts or very similar diffusion coefficients for the components. The errors of some of the calculated diffusion coefficients by MCR–NLR, shown in Fig. 3B, are slightly worse than those obtained from classical MCR. Nevertheless, the highest relative error of the calculated diffusion coefficient from MCR–NLR is not greater than 4.5%, which can still be regarded as a good result. From Fig. 3A and B, one can see that the worst result of the calculated spectra leads to a χ^2 of 0.7 (corresponding to correlation coefficient of 0.96) and the highest error for the diffusion coefficients is 4.3%.

The χ^2 values for the second component, glucose, displayed in Fig. 3C, show more or less the same patterns as ATP but the size of the errors is much larger. This is a component that is more difficult to resolve because there are two other components whose diffusion coefficients are similar to it. It is evident that classical MCR is not able to resolve this component since the χ^2 is very high for experiment 7, 8, 10, and 14. The simulated data in experiment 7, 8, and 14 contain a high level of similarity in the diffusion behaviour of the components, i.e. the differences are given by factors of 1.5 and 1.295, respectively. The simulated data for experiment 10, containing nominal level of diffusion difference and peak shifts but a high level of phase shifts, results in the highest χ^2 value of 14.6 (the correlation coefficient is 0.04) by classical MCR. This illustrates that classical MCR is not able to handle a high level of artefacts. This is

due to the problem of MCR that the solutions are not unique, i.e. MCR can end up with wrong incorrect resolution of a mixture without changing the error of the model. This is particularly true when a data set contains high level of disturbances. MCR–NLR yields much lower χ^2 values in all the experiments and hence it can be seen that MCR–NLR can resolve the components better, especially in difficult situations. This is because MCR–NLR can yield unique solutions with the fitting of decay profiles in each iteration of MCR. The errors of the diffusion coefficients for component glucose, presented in Fig. 3D, are quite low for MCR–NLR in all the cases while the errors obtained from classical MCR are high for experiment 10 and 14 (33.6% and 7.01%), reflecting the bad resolution of the spectra for the corresponding experiments. Nevertheless, it can be seen that in general the calculated diffusion coefficients are more stable than the calculated pure spectra for both methods.

Fig. 3E and F display the responses for the most easily resolved component SDS micelle. This component is easy because it has the smallest diffusion coefficient, i.e. the intensities of this component decay slowly, and has only one component (glucose) with a similar diffusion coefficient. As can be seen, in all the cases both classical MCR and MCR–NLR can resolve this component reasonably well, although the errors from classical MCR are slightly larger than those from MCR–NLR.

4.2. Interpreting the results of robustness tests

The relationship between the responses and the various levels of the factors can be evaluated by multiple linear regression, as described in Eqs. (8) and (9). In addition, the significance of the resulting regression coefficients (the effects) is checked in order to select the most important predictor variables for the model. To achieve this, the 95% confidence intervals for the true regression parameters are calculated:

$$b_i \pm t_{0.025, n-p} s_{b_i} \quad (i = 1 \dots 10) \quad (10)$$

where b_i is the corresponding regression coefficient and $t_{0.025, n-p}$ is the tabulated critical t -test value, n and p are row and column number of the predictor matrix X , respectively. There are two replicates for each simulation, so $t_{0.025, n-p}$ ($n=30$, $p=10$)=2.09. The $(s_{b_i})^2$ is the variances of the difference parameters, which are the diagonal elements of the variance–covariance matrix of the regression coefficients [21]. If zero is included in the confidence interval, the b_i is not significant at the 5% significance level.

Only the regression coefficients describing the relationship between χ^2 and the factors are shown here since, as can be seen in Fig. 3 (note the scale of the y -axis!), the factors have much more influence on the resolved spectra than on the diffusion coefficients. In fact, the effects of the factors on the diffusion coefficients are all insignificant (results not shown) and will not be discussed further. The ten regression coefficients b_i of the three components and their 95% confidence intervals, for classical MCR and MCR–NLR, are obtained by Eqs. (9) and (10), respectively (results not shown). The regression coefficients for ATP and glucose obtained from classical MCR are about three times higher than the corresponding values from MCR–NLR.

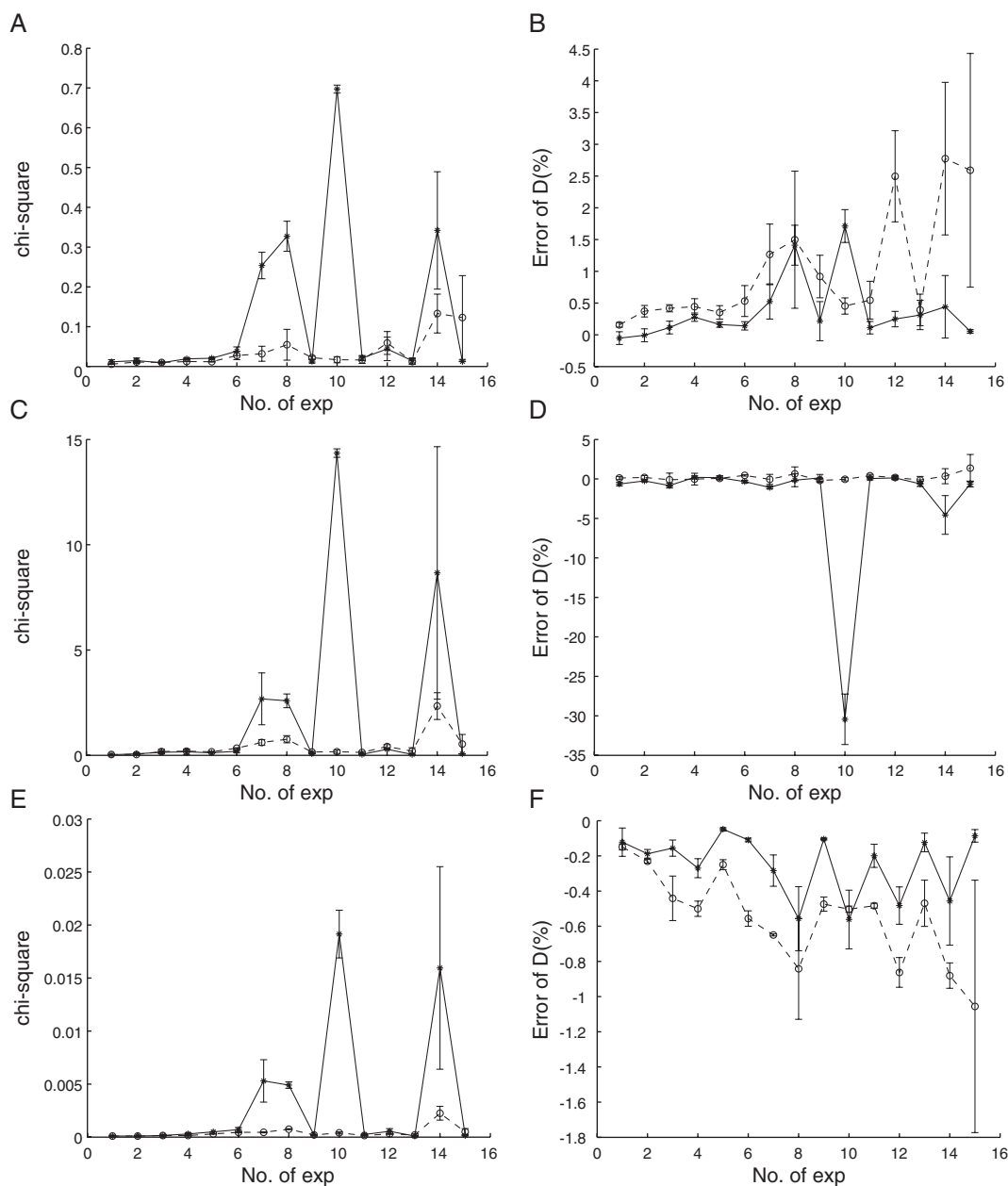


Fig. 3. Errors of the experiments obtained by MCR (solid line “—”) and MCR–NLR (dash line “- -”). (A) Chi-squares of the resolved spectrum of ATP; (B) errors of the diffusion coefficients of ATP; (C) chi-squares of the resolved spectrum of glucose; (D) errors of the diffusion coefficients of glucose; (E) chi-squares of the resolved spectrum of SDS; (F) errors of the diffusion coefficients of SDS.

This indicates that the factors have more negative effects on the separation of the pure components by classical MCR method. Since SDS is well resolved in all the cases, the b_i values for SDS will not be taken into account for further discussion.

According to the 95% confidence interval of each b_i , if ATP is resolved by classical MCR, two terms representing the main effects of phase shifts (x_1) and difference of D (x_3), and the quadratic term of phase shifts (x_1^2) are significant, whereas if separated by MCR–NLR, the terms for difference of D (x_3) and the quadratic terms, phase shifts (x_1^2) and peak shifts (x_2^2), are significant. For the separation of glucose, the term for the main effect of phase shifts (x_1) is significant by classical MCR and the

main effect of difference of D (x_3) is significant by MCR–NLR. Performing multiple linear regression by only the significant effects, the response surface for each case is expressed by the following models:

$$\text{For ATP by MCR : } \hat{y} = 0.042 + 0.0919x_1 + 0.0832x_3 + 0.0089x_1^2 \quad (11)$$

$$\text{For ATP by MCR–NLR : } \hat{y} = 0.0803 + 0.0212x_3 - 0.0273x_1^2 - 0.0210x_2^2 \quad (12)$$

$$\text{For glucose by MCR : } \hat{y} = 1.97 + 1.76x_1 \quad (13)$$

$$\text{For glucose by MCR–NLR : } \hat{y} = 0.418 + 0.368x_3 \quad (14)$$

From Eqs. (11) and (12), the highest χ^2 value is 0.36 and 0.12 for MCR and MCR–NLR, respectively. Both classical MCR and MCR–NLR lead to the χ^2 response surface for component ATP well below the quality cut-off of 1 (results not shown), indicating that both methods are robust for resolving this component.

The second component, glucose, is more difficult to resolve because it has two components that have similar diffusion behaviour. For example, if $x_3 = 1.682$, i.e. the difference D factor is 1.3, it means that the D of glucose is smaller than ATP with a factor of 1.3 and greater than SDS with a factor of 1.3. This “middle” component has more overlap problems on the dimension of diffusion coefficient. This can be illustrated by the χ^2 response lines shown in Fig. 3. Both of classical MCR and MCR–NLR result in higher χ^2 values for glucose than ATP. Again the χ^2 values obtained from classical MCR are higher. They increase distinctly with the level of phase shifts (see Fig. 3A), which is in agreement with Fig. 2C. From Fig. 2C, one can see that the high level of difference D factor (experiment 14) leads to large χ^2 value as well as large standard deviation of the two replicates. The b coefficient for difference D factor is not included in the model (Eq. (13)) because statistically, it is not significant. This is maybe due to the fact that there is large standard deviation between the responses of two replicates. Nevertheless, one should bear in mind that similar diffusion coefficients could lead to large errors of the results from classical MCR and it is in fact significant practically. One may argue that only two replicates may result in a freak accident where a method ends up in a local minimum and hence lead to large variability of the responses (or the errors). Although this could be the case, the large variability of the errors obtained from classical MCR for some of the data in fact is chemically meaningful; it occurs in the data generated at high noise level and small differences in diffusion coefficients (e.g. exp 14). To verify this, more replicates for some of the experiments like exp. 14 were performed, and the bad and extremely results with large variation obtained by classical MCR were confirmed. This is because in this situation the rotational ambiguity problem of classical MCR is more distinct, which can result in incorrectly resolved pure spectra and decay profiles without changing the residuals associated with the MCR model [22]. In Fig. 3B, the straight line is shown, describing the relationships between difference D factor and χ^2 obtained from MCR–NLR. The lower χ^2 values imply that MCR–NLR is able to deal with the phase and peak

Table 3
Factor levels of the data sets to evaluate single channel methods

Data	Phase shifts (x_1)	Peak shifts (x_2) (data point)	Difference D factor (x_3)
1	−0.50–0.50°	−0.17–0.17	1.70
2	−0.86–0.86°	−0.19–0.19	1.50

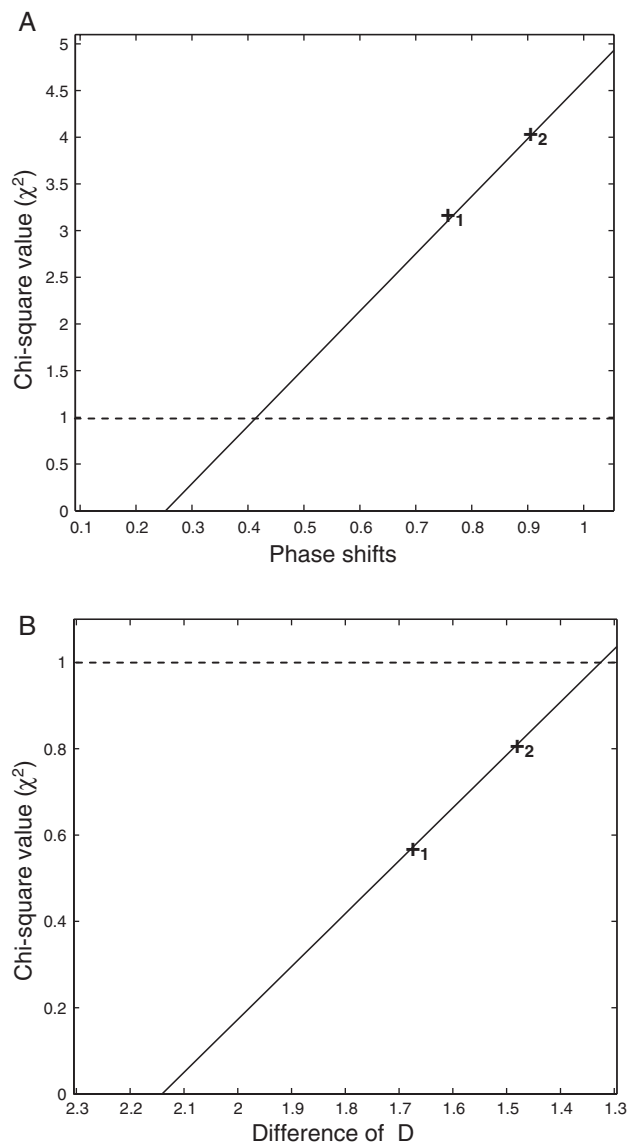


Fig. 4. Plots of χ^2 response of the component glucose varying with the different levels of, (A) phase shifts obtained from classical MCR; and (B) difference of D , obtained from MCR–NLR. The dash line indicates the quality cut-off of 1 for χ^2 values. (The points with “+” are those level to evaluate both MCR methods and the SPLMOD method.)

shifts, as well as the rotational ambiguities, better than classical MCR.

4.3. Comparison of the MCR methods with a single channel method

SPLMOD is one of the single channel methods implemented in the commercial NMR software XWINNMR by Bruker. To test the single channel method, two levels (both relatively low and high) of the three factors are selected so that the resulting χ^2 value is reasonable for MCR–NLR, i.e. $\chi^2 \leq 1$ (correlation coefficient > 0.93), but not necessarily for classical MCR. The factor levels of data sets are presented in Table 3 and they are indicated by “+1”, “+2”, in Fig. 4. In Table 3, the values for the phase shifts range, the peak shifts range, and the difference D

factor of the two situations are shown. Three replicates of the data sets are constructed for each situation. The first three replicate data sets are constructed with phase shifts (x_1) of range -0.5 – 0.5° , peak shifts (x_2) -0.17 – 0.17 data points, and the difference D factor (x_3) is 1.7. The second three replicated data sets are constructed using the high level of all the factors, i.e. phase shifts ranging -0.86 – 0.86° , peak shifts -0.19 – 0.19 data points and difference D factor 1.5. With these data, it is intended to evaluate how the single channel method SPLMOD reacts to different levels of the factors. In addition, these data sets are also analysed by the two multivariate methods for comparison. For each replicate data set in each situation, the diffusion coefficient for each channel (single peak or group of peaks) in the chemical

shift dimension is calculated by SPLMOD. Then the 2D DOSY plot of each data set is constructed. Because the 2D DOSY plots are similar for the replicates, only one for each set of the data is shown (see Fig. 5). Fig. 5A displays the first spectrum of the mixture, which is used to identify the corresponding peaks for each component. In the two 2D DOSY plots, Fig. 5B and C, the horizontal axis represents the chemical shifts dimension with the same scale as that in Fig. 5A. The vertical axis represents the dimension for diffusion coefficients, which is constructed by the resulting diffusion coefficients from SPLMOD. As can be seen in Fig. 5B, there are some variations in the calculated diffusion coefficients of the same component between the different channels. For example, the diffusion coefficient at 8.0 ppm is

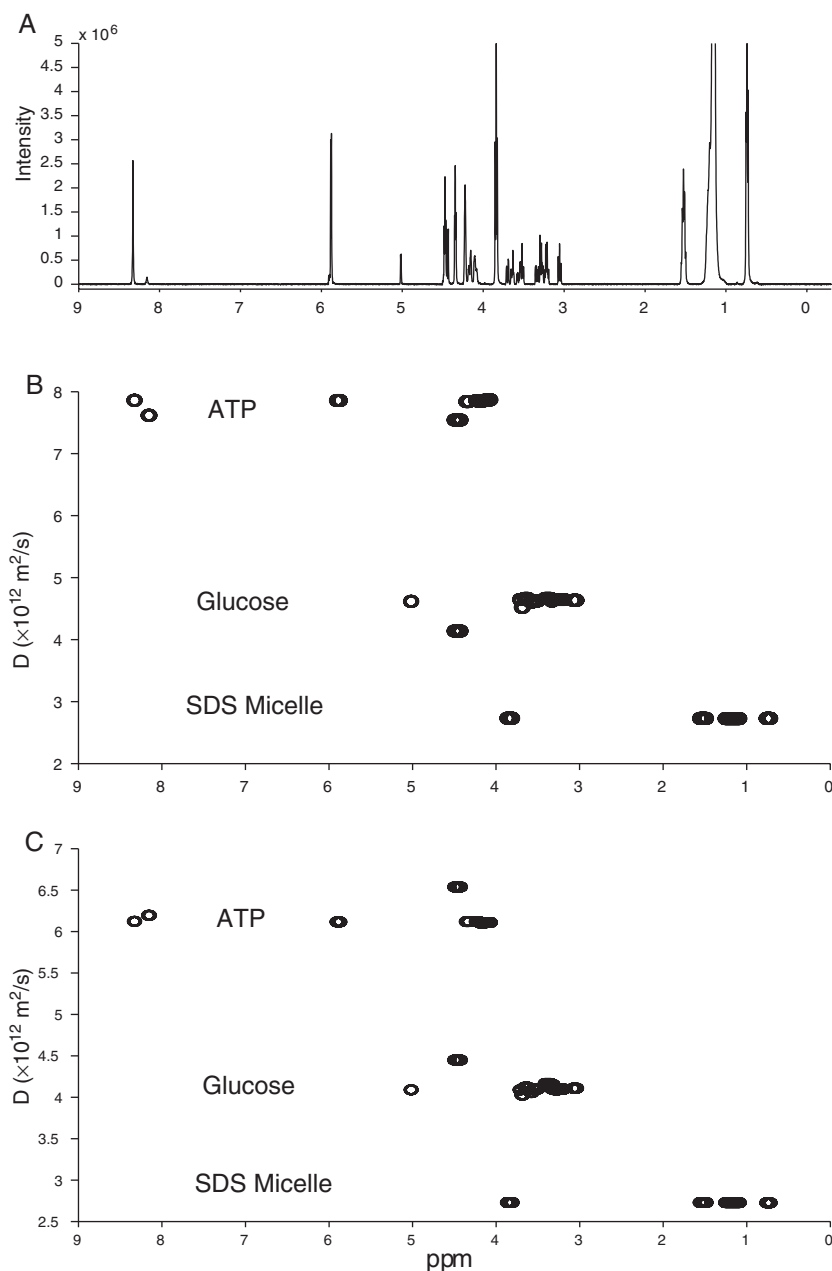


Fig. 5. 2-D DOSY plots obtained from a single channel method. (A) First spectrum of the mixture; (B) the DOSY plot for the data when $x_1 = 0.5^\circ$, $x_2 = 0.17$ data points, and $x_3 = 1.7$; (C) the DOSY plot when $x_1 = 0.86^\circ$, $x_2 = 0.19$ data points, and $x_3 = 1.5$.

deviating from other diffusion coefficients of ATP. The area around 4.5 ppm contains the peaks where two components, ATP and glucose, overlap. Two diffusion components are found by SPLMOD in this region, but both of the calculated diffusion coefficients are different from those of ATP and glucose. From this example, it can be seen that single channel methods can result in different diffusion coefficients for the same component under the effect of experimental artefacts. Moreover, the overlap area is difficult to resolve correctly, especially when the diffusion coefficients in it are similar. The second DOSY plot, presented in Fig. 5C, is for a replicate of the data set with high levels of the factors. With a larger amount of phase and peak shifts, one can see that there are more variations of the diffusion coefficient for the same components, especially for component ATP and glucose. The diffusion coefficients of the overlapping area (~ 4.5 ppm) are deviating further from those of ATP and glucose. It is evident that the overlapping peaks around 4.5 ppm are even more difficult to resolve when the difference D factor becomes smaller.

To compare the multivariate methods and the single channel method easily, the χ^2 values of the resolved pure spectra by SPLMOD are also calculated in the same way as before. Using the first spectrum of each data set, the resolved pure spectra of the components can be obtained by projecting the peaks with similar diffusion coefficients on the chemical shift dimension. Since some of the diffusion coefficients, e.g. those at 4.5 ppm, are quite different from others, they will not be considered to belong to the closest component if they have a 10% or higher difference from the mean of the main diffusion coefficients. In this way, the peaks assigned to the same component construct the resulting pure spectra obtained from SPLMOD. These pure spectra are normalized to the unit length of 1 so that they are comparable to the reference spectra and the χ^2 values can be calculated in a similar way of the multivariate methods. Table 4 presents the mean of the estimated χ^2 values of the three replicates for each situation, resulting from classical MCR, MCR–NLR, and SPLMOD, respectively. At the relatively low level of the factors (data1), the χ^2 values obtained from SPLMOD are all higher than those from the two multivariate methods. In particular, SPLMOD yields a high χ^2 value for component glucose, indicating single channel methods cannot resolve this component correctly. It can also be seen that both classical MCR and MCR–NLR achieve very low χ^2 for all the components, indicating the component of the mixture are well separated. When the factor levels are higher, the data is more difficult to resolve. This is reflected by the higher χ^2 values for data2. The classical MCR

Table 4
Comparison of the mean χ^2 values obtained from classical MCR, MCR–NLR, and SPLMOD

Component	Data1 (χ^2)			Data2 (χ^2)		
	MCR	MCR–NLR	SPLMOD	MCR	MCR–NLR	PLMOD
ATP	0.0274	0.0152	0.4575	0.3600	0.1506	0.4524
Glucose	0.1276	0.1886	2.0573	5.7688	1.0091	4.8294
SDS	0.0004	0.0002	0.0019	0.0096	0.000	0.0025

results in the highest χ^2 value for glucose, suggesting this component is not resolved correctly from the data with high factor level. This is because the rotational ambiguities problem of classical MCR is enlarged when a considerable amount of experimental artefacts (peak shifts, phase shifts etc.) exist in the data, and the overlap between the components in the diffusion coefficient dimension and chemical shift dimension become serious. Embedding NLR in each iteration of classical MCR overcomes the rotational ambiguities caused by the experimental artefacts and overlap problems. Consequently, MCR–NLR is able to resolve better pure spectra, which leads to reasonably low values of χ^2 , even for the difficult component glucose. The χ^2 value of glucose obtained from SPLMOD for data2 are also very high. The resulting diffusion coefficients of the overlapping peaks at 4.5 ppm leading to incorrect assignment of the peaks to the glucose spectrum are the main reason for this large error. From the comparison of the three processing methods, one can see that when classical MCR and single channel methods is not able to correctly resolve the data with a certain level of the factors, MCR–NLR still retains its robustness.

5. Conclusions

Robustness tests reveal that MCR–NLR can handle artefacts and similar diffusion coefficients of the components in DOSY data better than classical MCR and the single channel method SPLMOD. In general, multivariate analysis of DOSY data deals with the overlap problems in both of the chemical dimension and the diffusion coefficient dimension, while single channel method analyse each channel to obtain the diffusion coefficient separately. Therefore, the main problem of a single channel method is that it can easily fail when a DOSY data set contains overlapping areas, which are very commonly present in real-world data. Classical MCR can deal with the overlapping areas better, only if the data contain small amount of experimental artefacts and the diffusion coefficients are not so similar. If considering experimental artefacts and overlap problems, MCR–NLR proves to be the best choice to deal with them. In addition, preprocessing of the data set to correct for random shifts can improve the performance of MCR–NLR further [3]. To process DOSY data, it is suggested that the data should be first pre-processed to minimise the experimental artefacts and then analysed by MCR–NLR. Single channel methods can be used to analyse the data as a further proof of the results. In this paper, MCR–NLR manages to identify discrete samples in the situation that classical MCR fails. However, because classical MCR has no assumption of decay behaviour of the components, it can be the potential of analysing polydisperse samples, which will be one of the issues for future study.

Acknowledgements

This research is funded by the Dutch Technology Foundation STW (790.35.393). The authors thank Jos Joordens (Biophysical Chemistry, Radboud University Nijmegen) for recording the ^1H NMR DOSY data.

References

- [1] L.C.M. Van Gorkom, T.M. Hancewica, Analysis of DOSY and GPC–NMR experiments on polymers by multivariate curve resolution, *J. Magn. Reson.* 130 (1998) 125–130.
- [2] R. Huo, R. Wehrens, J. van Duynhoven, L.M.C. Buydens, Assessment of techniques for DOSY NMR data processing, *Anal. Chim. Acta* 490 (2003) 231–251.
- [3] R. Huo, R. Wehrens, L.M.C. Buydens, Improved DOSY NMR data processing by data enhancement and combination of multivariate curve resolution with non-linear least square fitting, *J. Magn. Reson.* 169 (2004) 257–269.
- [4] M. Esteban, C. Ariño, J.M. Díaz-Cruz, M.S. Díaz-Cruz, R. Tauler, Multivariate curve resolution with alternating least squares optimisation: a soft-modelling approach to metal complexation studies by voltammetric techniques, *Trends Anal. Chem.* 19 (2000) 49–61.
- [5] M. Zeiter, J.-M. Roger, V. Bellon-Maurel, D.N. Rutledge, Robustness of models developed by multivariate calibration. Part I. The assessment of robustness, *Trends Anal. Chem.* 23 (2004) 157–170.
- [6] Y. Vander Heyden, A. Nijhuis, J. Smeyers-Verbeke, B.G.M. Vandeginste, D.L. Massart, Guidance for robustness/ruggedness tests in method validation, *J. Pharm. Biomed. Anal.* 24 (2001) 723–753.
- [7] ICH Harmonised Tripartite Guideline Prepared Within the Third International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH), Text on Validation of Analytical Procedures, 1994.
- [8] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. de Jong, J. Smeyers-Verbeke, Handbook of chemometrics and qualimetrics, Data Handling in Science and Technology, vol. 20A, part A, Elsevier, Amsterdam, 1997–98.
- [9] B. Antalek, Using pulsed gradient spin echo NMR for chemical mixture analysis: how to obtain optimum results, *Concepts Magn. Reson.* 14 (2002) 225–258.
- [10] K.F. Morris, C.S. Johnson Jr., Resolution of discrete and continuous molecular size distributions by means of diffusion-ordered 2D NMR spectroscopy, *J. Am. Chem. Soc.* 115 (1993) 4291–4299.
- [11] V. Wsól, A.F. Fell, Central composite design as a powerful optimisation technique for enantioresolution of the rac-11-dihydrooracin (the principal metabolite of the potential cytostatic drug oracin), *J. Biochem. Biophys. Methods* 54 (2002) 377–390.
- [12] R.H. Vogel, Pslmod users manual, Data Analysis Group, EMBL-DA09, EMBL, Heidelberg, Germany, 1988.
- [13] R. Tauler, Multivariate curve resolution applied to second order data, *Chemom. Intell. Lab. Syst.* 30 (1995) 133–146.
- [14] F.C. Sánchez, B. Van den Bogaert, S.C. Rutan, D.L. Massart, Multivariate peak purity approaches, *Chemom. Intell. Lab. Syst.* 34 (1996) 139–171.
- [15] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes, Cambridge University Press, New York, 1998.
- [16] J.C. Miller, J.C. Miller, Statistics for Analytical Chemistry, 3rd Edition, Ellis Horwood Ltd, Chichester, UK, 1993.
- [17] R. Huo, R.A. van de Molengraaf, J.A. Pikkemaat, R. Wehrens, L.M.C. Buydens, Diagnostic analysis of experimental artefacts in DOSY NMR data by covariance matrix of the residuals, *J. Magn. Reson.* 172 (2005) 342–358.
- [18] H. Wijes, W.J. Melssen, H.J.A. in 't Zandt, M. van der Graaf, A. Heerschap, L.M.C. Buydens, Automatic correction for phase shifts, frequency shifts, and lineshape distortions across a series of single resonance lines in large spectral data sets, *J. Magn. Reson.* 144 (2000) 35–44.
- [19] G.A. Morris, Compensation of instrumental imperfections by deconvolution using an internal reference signal, *J. Magn. Reson.* 80 (1988) 547–552.
- [20] B.M. Wise, N.B. Gallagher, PLS_Toolbox 2.1: for use with MATLAB™, Eigenvector Research, Inc. 2000.
- [21] N.R. Draper, H. Smith, Applied Regression Analysis, 2nd edition, John Wiley, New York, 1981.
- [22] K. Smilde, H.C.J. Hoefsloot, H.A.L. Kiers, S. Bijlsma, H.F.M. Boelens, Sufficient conditions for unique solutions within a certain class of curve resolution models, *J. Chemom.* 15 (2001) 405–411.