Available online at www.sciencedirect.com
SCIENCE DIRECT®
Analytica Chimica Acta 490 (2003) 17–29
ANALYTICA CHIMICA ACTA
www.elsevier.com/locate/aca

Evaluation of synthetic liquid chromatography—diode array detection—mass spectrometry data for the determination of enzyme kinetics

Ernst Bezemer, Sarah Rutan*

Department of Chemistry, Virginia Commonwealth University, P.O. Box 842006, Richmond, VA 23284-2006, USA

Received 23 October 2002; received in revised form 11 March 2003; accepted 18 March 2003

Abstract

In this paper, we investigate the accuracy and precision of the results from diode array detector (DAD) data and mass spectrometry (MS) data as obtained subsequent to chromatographic separations using computer simulations. Special attention was given to simulations of multiple injections from a developing enzymatic reaction. These simulations result in three-way LC–DAD–MS kinetic data; LC–DAD and LC–MS data were also evaluated independently in this investigation. The noise characteristics of the MS detector prevent accurate determination of the individual reaction rate constants by the analysis method. Using the data from the DAD in combination with the MS detector results in improved estimation of the rate constants. The results also indicate that the higher resolving power of the MS information compensates for the lower signal-to-noise ratio in these data, compared to DAD data.

Keywords: Enzyme kinetics; LC–DAD–MS data; Computer simulations; Multivariate curve resolution–alternating least squares (MCR–ALS)

1. Introduction

Mass spectrometry (MS) detectors are commonly used with liquid chromatography (LC) instruments [1]. This coupling allows for the identification of unknown compounds found in the chromatograms. The structural information given by MS is far superior to the traditionally used UV-Vis detectors, such as single wavelength detectors and diode array detectors (DAD) [2]. However, these spectrophotometric detectors have highly reproducible responses and are, therefore, excellent for quantification [1]. Quantification is important for many LC applications. In our current research, the determination of rate constants from multiple injections of evolving chemical systems into a LC–DAD–MS instrument is of interest. In order to fully understand the complete three-way LC–DAD–MS: kinetic data analysis, reduced order datasets are studied first. This is achieved by initially investigating two-way data arrays to characterize the effects of combining spectroscopic and mass spectrometric detection methods. This is followed by a study of fitting kinetic data to the enzyme kinetic model to isolate possible complications resulting from the kinetic fitting routine. Lastly, all the individual analyses are combined into a single three-way analysis of a simulated enzyme kinetic experiment using a LC–DAD–MS instrument.

The evolving chemical system under investigation is the enzyme-catalyzed conversion of a substrate to

* Corresponding author. Tel.: +1-80-4828-7517; fax: +1-80-4828-8599.
E-mail address: srutan@saturn.vcu.edu (S. Rutan).

0003-2670/$ – see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0003-2670(03)00356-8

Scheme 1.

a product. The general scheme for this reaction is shown in Scheme 1, in which E is the enzyme; S, the substrate; ES, the enzyme substrate complex; and P, the product. This system is typically described by the Michaelis–Menten equation shown in Eq. (1) [3].

$$\left(\frac{\mathrm{dP}}{\mathrm{d}t}\right)_{t=0} = v_0 = \frac{v_{\max}[\mathrm{S}]}{K_\mathrm{m} + [\mathrm{S}]} \tag{1}$$

where v_0 is the initial rate of formation of product, $v_{\max}$ is the maximal conversion rate, K_m is the Michaelis constant and [S] is the initial substrate concentration.

The Michaelis–Menten equation is based on several assumptions, including that the enzyme–substrate complex concentration is at a steady state. For this assumption to be valid, the substrate has to be in large excess [3]. A more accurate and complete description would be to use the rate equations as given in Eqs. (2)–(5), which are not based on these assumptions and are valid for any substrate concentration at any time point.

$$\begin{aligned}\frac{\mathrm{d[S]}}{\mathrm{d}t} &= -k_1[\mathrm{E}][\mathrm{S}] + k_2[\mathrm{ES}] \\ &= -1 \times k_1[\mathrm{E}][\mathrm{S}] + 1 \times k_2[\mathrm{ES}] + 0 \times k_3[\mathrm{ES}]\end{aligned} \tag{2}$$

$$\begin{aligned}\frac{\mathrm{d[P]}}{\mathrm{d}t} &= k_3[\mathrm{ES}] \\ &= +0 \times k_1[\mathrm{E}][\mathrm{S}] + 0 \times k_2[\mathrm{ES}] + 1 \times k_3[\mathrm{ES}]\end{aligned} \tag{3}$$

$$\begin{aligned}\frac{\mathrm{d[E]}}{\mathrm{d}t} &= -k_1[\mathrm{E}][\mathrm{S}] + k_2[\mathrm{ES}] + k_3[\mathrm{ES}] \\ &= -1 \times k_1[\mathrm{E}][\mathrm{S}] + 1 \times k_2[\mathrm{ES}] + 1 \times k_3[\mathrm{ES}]\end{aligned} \tag{4}$$

$$\begin{aligned}\frac{\mathrm{d[ES]}}{\mathrm{d}t} &= k_1[\mathrm{E}][\mathrm{S}] - k_2[\mathrm{ES}] - k_3[\mathrm{ES}] \\ &= +1 \times k_1[\mathrm{E}][\mathrm{S}] - 1 \times k_2[\mathrm{ES}] - 1 \times k_3[\mathrm{ES}]\end{aligned} \tag{5}$$

These equations can be further expanded using Eqs. (6)–(8)

$$k_1[\mathrm{E}][\mathrm{S}] = k_1[\mathrm{E}]^1[\mathrm{S}]^1[\mathrm{ES}]^0[\mathrm{P}]^0 \tag{6}$$

Table 1

Chemical model	Reaction pathway matrix	Reaction order matrix
Enzyme kinetic model[a]	$\begin{bmatrix} -1 & +1 & 0 \\ 0 & 0 & +1 \\ -1 & +1 & +1 \\ +1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$
Zero-order kinetic model	$\begin{bmatrix} -1 \\ +1 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \end{bmatrix}$

[a] This model, based on the rate constants instead of the Michaelis–Menten parameters, $v_{\max}$ and K_m, does not require the steady state assumption to be made.

$$k_2[\mathrm{ES}] = k_2[\mathrm{E}]^0[\mathrm{S}]^0[\mathrm{ES}]^1[\mathrm{P}]^0 \tag{7}$$

$$k_3[\mathrm{ES}] = k_3[\mathrm{E}]^0[\mathrm{S}]^0[\mathrm{ES}]^1[\mathrm{P}]^0 \tag{8}$$

The coefficients of the linear combinations in each differential equation (Eqs. (2)–(5)) form the reaction pathway matrix as shown in Table 1. The exponents for each species in Eqs. (6)–(8) form the reaction order matrix shown also in Table 1.

The parameters in the Michaelis–Menten equation can be related to the micro-rate constants (k_1, k_2 and k_3) in Eqs. (2)–(8) by Eqs. (9) and (10) [3]

$$K_\mathrm{m} = \frac{k_2 + k_3}{k_1} \tag{9}$$

$$v_{\max} = k_3[\mathrm{E}]_\mathrm{T} \tag{10}$$

where $[\mathrm{E}]_\mathrm{T}$ is the total enzyme concentration.

We have previously described an approach for fitting kinetic data to any chemical model [4]. This approach encodes the information represented by Eqs. (2)–(8) (or alternatively any other chemical model) as two matrices, as described earlier. This method for describing kinetic systems has advantages relative to previously reported methods, as it can be easily generalized to, for example, inhibition of the enzyme by excess substrate, multiple substrate models or multiple enzyme models. These applications, however, are beyond the scope of this paper. In the present work, a kinetic fitting approach, using the matrix representation described earlier, is used to investigate the estimation of rate constants in Eqs. (2)–(8) from LC–DAD–MS kinetic data and to understand the influence of the addition of an MS detector to an

LC–DAD instrument on the accuracy and precision of the rate constants. The effect of this additional data on the chemometric resolution of chromatographically overlapped compounds has also been studied.

A multivariate curve resolution method using alternating least squares (MCR-ALS) was used for the analysis [5,6]. The MCR-ALS algorithm has been described previously and is based on Beer's law, as shown in Eq. (11) [7]

$$\boldsymbol{D} = \boldsymbol{C} \cdot \boldsymbol{S}^{\mathrm{T}} + \boldsymbol{E} \tag{11}$$

where $\boldsymbol{D}$ is the data matrix ($m_\lambda \times m_{\mathrm{rt}}$) containing absorbance measurements as a function of retention time (rt), wavelength (λ) and m_{rt} and m_λ are the number of data points of the spectrum and chromatogram, respectively, $\boldsymbol{C}$ ($m_{\mathrm{rt}} \times n$) contains the concentration profiles of each component, $\boldsymbol{S}^{\mathrm{T}}$ ($m_\lambda \times n$) is the transpose of the pure component spectra matrix (where n is the number of components) and $\boldsymbol{E}$ ($m_\lambda \times m_{\mathrm{rt}}$) contains the error. The $\boldsymbol{S}$ matrix contains the spectral information of each pure component, and can be row augmented to contain MS data as well as DAD data.

When an extra order of information is present, such as for reaction kinetic data, the model can be described as shown in Eq. (12)

$$\boldsymbol{D} = \boldsymbol{R} \cdot \boldsymbol{K} \cdot \boldsymbol{S}^{\mathrm{T}} + \boldsymbol{E} \tag{12}$$

where $\boldsymbol{D}$ is a three-way array ($m_{\mathrm{rt}} \times m_\lambda \times m_{\mathrm{rx}}$) containing absorbance measurements as a function of retention time, wavelength and reaction time, $\boldsymbol{K}$ ($m_{\mathrm{rx}} \times n$) contains the kinetic profiles for each component and $\boldsymbol{R}$ ($m_{\mathrm{rt}} \times n$) contains the retention profiles for each component. $\boldsymbol{R}$ is a two-way array when the trilinearity constraint is applied and a three-way array ($m_{\mathrm{rt}} \times m_{\mathrm{rx}} \times n$) when the bilinearity constraint is applied.

The data array, $\boldsymbol{D}$, is decomposed into $\boldsymbol{R}$, $\boldsymbol{K}$ and $\boldsymbol{S}$ by assuming the $\boldsymbol{R}$ and $\boldsymbol{K}$ can be combined into a $\boldsymbol{C}$ matrix according to Eq. (13) [8].

$$\boldsymbol{D} = (\boldsymbol{R} \cdot \boldsymbol{K}) \cdot \boldsymbol{S}^{\mathrm{T}} + \boldsymbol{E} = \boldsymbol{C} \cdot \boldsymbol{S}^{\mathrm{T}} + \boldsymbol{E} \tag{13}$$

The $\boldsymbol{C}$ and $\boldsymbol{S}^{\mathrm{T}}$ matrices are subsequently solved by alternating between Eqs. (14) and (15) until a minimum fit error is reached [7].

$$\boldsymbol{S}^{\mathrm{T}} = (\boldsymbol{C}^{\mathrm{T}} \cdot \boldsymbol{C})^{-1} \cdot \boldsymbol{C}^{\mathrm{T}} \cdot \boldsymbol{D} \tag{14}$$

$$\boldsymbol{C} = \boldsymbol{D} \cdot (\boldsymbol{S}^{\mathrm{T}} \cdot \boldsymbol{S})^{-1} \cdot \boldsymbol{S} \tag{15}$$

The search for this minimum may be guided by the use of chemically relevant constraints, such as non-negativity, unimodality, selectivity and trilinearity. All these constraints are commonly used with the ALS algorithm and have been described in detail in the literature [7,9].

The application of a kinetic model during the resolution of chromatographically overlapped compounds has been performed by other research groups [10–12]. However, none of these methods are able to flexibly describe any chemical kinetic model. The method recently described by our group has been shown to effectively encode any reaction mechanism comprised of elementary reaction steps and is used in this work [4].

In order to start the iterations, an initial estimate of either the concentration profiles ($\boldsymbol{C}$) or the pure component spectra ($\boldsymbol{S}$) has to be obtained. Evolving factor analysis (EFA) has been used here to generate the starting profiles for the concentration profiles [13]. EFA follows the evolution of the principal components as a function of retention time by applying singular value decomposition (SVD) on an increasing number of rows of the dataset. It makes the chemically valid assumption that the first component to elute from the chromatographic column is also the first component to stop eluting, and so on, for each additional component. It uses this assumption to correlate the forward and backward analyses and generates an initial estimate for the concentration profiles [13].

Many examples can be found in literature on the application of chemometric methods to overlapped LC–DAD data [14,15] and to resolve overlapped GC–MS data [16–19] and overlapped LC–MS data [20–23]. Others have used the MS information to determine reaction rates, with limited success [24]. Enzyme kinetics have also been characterized by MS [25]. The combination of DAD and MS data that are examined here, is rarely described, thus an investigation into the usability of this type of augmented data is warranted.

2. Experimental

All data were simulated in the MATLAB programming environment on various Athlon™ and Pentium IV™ computers. The data were generated using the model represented by Scheme 1, resulting in a total of four components. However, the enzyme and

enzyme–substrate complex are not often observed in typical enzyme kinetic experiments as the enzymes are typically denatured and precipitated before analysis. In addition, many analytical techniques have difficulty analyzing relatively small molecules, such as typical substrates and products, while at the same time determining the concentration of relatively large molecules such as enzymes. Moreover, only accurate mass determinations will permit differentiation between an enzyme and an enzyme with a substrate molecule bound to it. Therefore, these species are omitted in our analysis.

The retention profiles for the substrate and product were generated by Gaussian functions consisting of 50 points, as shown in Fig. 1A. The DAD-spectral profiles were generated by adding Gaussian functions to first-order decay curves, yielding spectra with 51 points, as shown in Fig. 1B. The simulated electrospray MS spectral profile was an all zero vector of 52 points with a single point with a value of one to indicate the parent ion. A typical MS spectrum (simulated for unit mass resolution) including noise is shown in Fig. 1C. The kinetic profiles were generated using the enzyme kinetic model with rate constants of $0.58\,(\mu M\,s)^{-1}$, $4.0\,s^{-1}$ and $0.29\,s^{-1}$, for k_1, k_2 and k_3, respectively. These numbers are comparable to typical enzyme parameters, such as those found for the CYP2D6 catalyzed O-demethylation of dextromethorphan, which has values of 7.5 μM and 17.3 nmol (nmol min)$^{-1}$ for the K_m and v_{max}, respectively [26]. The initial substrate concentration was 20 μM while the enzyme concentration was 0.03 μM.

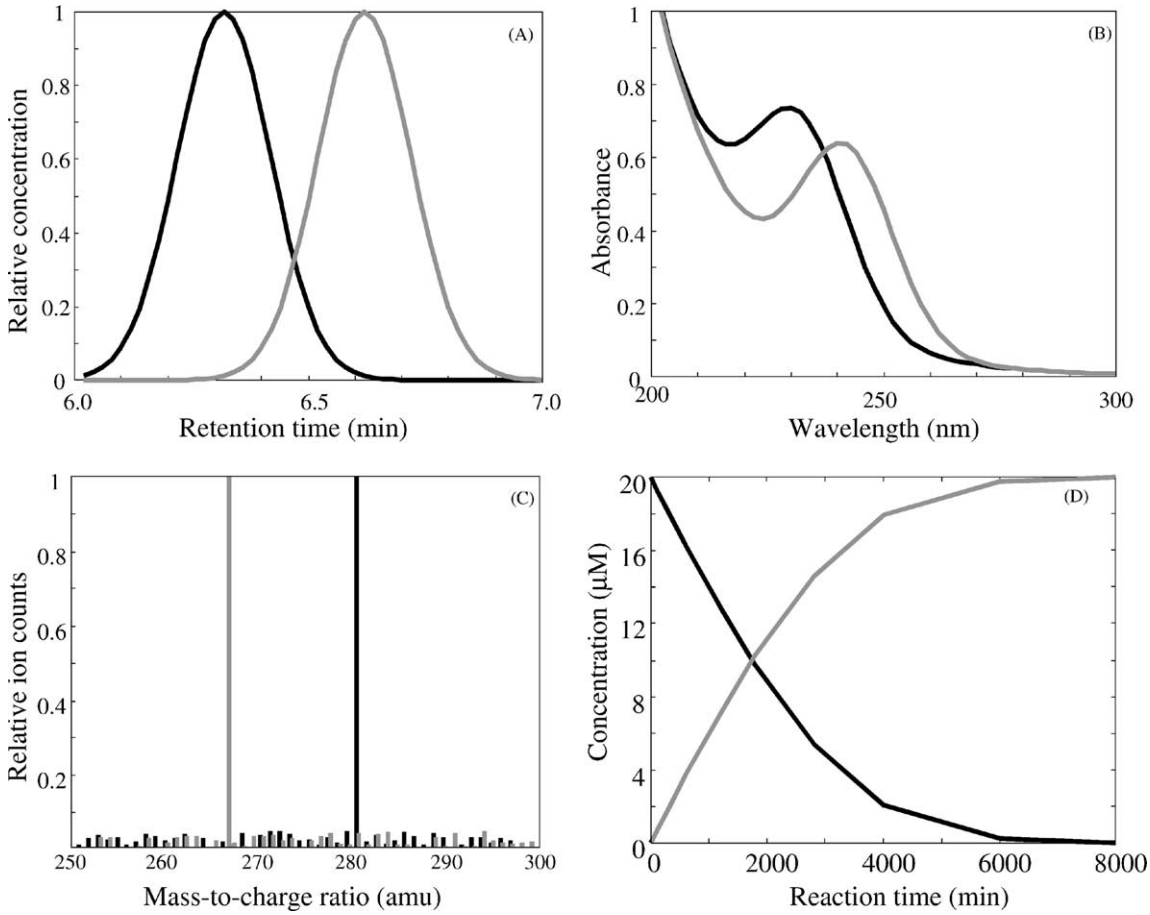

Fig. 1. (A) The synthetic retention profiles, (B) DAD spectral profiles, (C) MS spectral profiles and (D) kinetic profiles for the substrate (▬) and the product (▬).

A typical kinetic profile is shown in Fig. 1D. The 10 reaction time points were chosen between 0 to 8000 s at increasing sampling intervals (6, 120, 300, 600, 1200, 1800, 2800, 4000, 6000, 8000 s). While the actual time and concentration values for the simulations could be arbitrary, the values chosen here are typical for those observed in real enzyme kinetic experiments. The spectral profiles for the combined DAD and MS instrument were created by augmenting the DAD spectral profiles with the MS profiles, resulting in a single mixed spectral profile for each component.

These default profiles were adjusted to investigate the influence of various factors on the resolution results. The retention times were changed to modify the degree of chromatographic overlap between the substrate and the product peaks. The position of the Gaussian function in the spectral profile was changed to vary the spectral similarity between the two components. The position of the peak in the mass spectrum, corresponding to the mass-to-charge ratio of the product, was changed to determine the influence of differences in molecular weight. The initial substrate concentration was varied in a manner similar to that used in enzyme kinetic experiments and ranged from 10 to 200 μM. Two different kinetic model constraints were explored: the enzyme kinetic model shown in Scheme 1 and zero-order kinetics. The zero-order kinetic model was applied because when the substrate is in large excess relative to the enzyme concentration, the kinetic profiles are linear. Table 1 shows the reaction order and pathway matrices that describe the models that were used to fit enzyme kinetic simulated data.

The noise in the data was generated in a way that was analogous to that observed in real DAD and MS data. Normally distributed random numbers were added to the DAD spectral profiles. In order to simulate the chemical noise found in MS data, non-negative (clipped), normally distributed random numbers were added to the MS profiles. The default noise level for the simulations are 1 and 5% for the DAD and MS profiles, respectively. These noise levels were varied over the range from 0 to 15% for the noise level in the DAD and from 0 to 50% for the MS profiles. These values are relative to the maximum signal value in the corresponding spectrum.

The aging of the column and other instrument fluctuations were simulated by decreasing the retention time of the product, thereby increasing the chromatographic overlap with the substrate peak as the reaction progresses. In the three-way data consisting of LC–DAD–MS kinetic data this breaks the trilinearity of the data and could introduce errors in the results.

The data analysis was performed using in-house programs written in the MATLAB programming environment [27]. EFA, originally developed by Maeder and Zilian [13] and Keller and Massart [28], was adapted from a program written by the Tauler research group [29]. The MCR-ALS program used here was developed in this research group [8], but was based on some of the approaches used in the MCR-ALS program developed in Tauler's research group [30,31].

The initial estimates for the ALS algorithm were generated by EFA using two components. All profiles were constrained to non-negativity. The retention profiles were constrained by horizontal unimodality [9]. With this set of synthetic data, the retention and mass spectral profiles have selective regions in which only a single component contributes to data and even though the ALS algorithm allows for the use of selectivity constraints, we did not apply this constraint during the analysis to prevent improving the results by this a priori information that may not be available when analyzing real data.

When three-way reaction kinetic data were analyzed, EFA was applied to the sixth slice (which corresponds to approximately equal amounts of substrate and product), and the resulting chromatographic profiles were used as the initial guess for all slices of the data array. Besides the earlier mentioned constraints, trilinearity was also applied to the first component, which included no retention time shifts. The analysis was performed while the second component was constrained to either trilinearity or bilinearity in the spectral dimension to determine the influence of constraining a non-trilinear component to trilinearity.

The quality of ALS resolution results was evaluated according to Eq. (16). Although the fit error does not indicate the fit quality for each individual profile, earlier publications have shown that the total fit error follows the same trends as the fit quality for the individual component profiles [8].

$$\text{Fit error (\%)} = 100\sqrt{\frac{\sum_{t,s,k}(d_{t,s,k}^{\text{optimal}} - d_{t,s,k}^{\text{data}})^2}{\sum_{t,s,k}(d_{t,s,k}^{\text{data}})^2}} \quad (16)$$

3. Results and discussion

3.1. Analysis of two-way LC–DAD, LC–MS, and LC–DAD–MS data

3.1.1. Influence of the noise level

The influence of the noise level in the DAD and MS data was analyzed by changing the noise level in the DAD signal from 0 to 15% while changing the noise level of the MS data from 0 to 50%. All the other parameters for the synthetic dataset remained constant. For these experiments, the kinetic direction was ignored and only two-way data were analyzed, with an initial substrate concentration of 20 μM. Ten synthetic sets were created and the results were averaged to obtain more precise estimates of the effect of a particular noise level. Furthermore, three different two-way datasets were analyzed: LC–DAD, LC–MS and LC–DAD–MS.

3.1.2. Influence of the noise level on singular values

It is important to evaluate the influence of the noise level on the number of significant singular values as determined by SVD. Instead of plotting numerous SVD plots (log singular value versus singular value number), our attention was focused on the relative magnitude of the third singular value. The data were generated using two components, therefore, the third singular value (sv_3) should be due to noise only and should not be significantly greater than the fourth (sv_4) and higher singular values. In order to display the information found by SVD analysis, the ratio of the third singular value was calculated in relation to the second and the fourth singular values according to Eq. (17).

$$SV_3\text{-ratio} = \frac{sv_3 - sv_4}{sv_2 - sv_4} \tag{17}$$

This ratio indicates whether the third component is closer to the noise level (sv_4) or to the real component (sv_2). When the SV_3-ratio is small (i.e. <0.5), the third singular value is closer to the noise level, represented by sv_4, and generally indicates the presence of only two components. As the SV_3-ratio increases, the third singular value approaches the second singular value and could be falsely attributed to a third component present in the data. The results from the SVD-analysis of the LC–DAD–MS data are shown in Fig. 2. The SV_3-ratio increases approximately linearly with the increasing noise level in the MS profile. The reason that the third singular value becomes more significant as the amount of noise in the MS data increases is due to the fact that this noise is non-random and causes

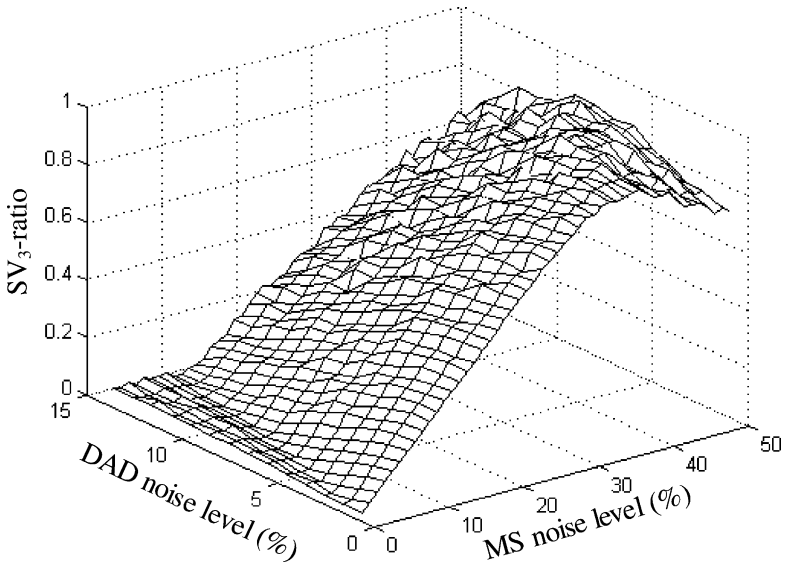

Fig. 2. The SV_3-ratio as a function of the DAD and the MS noise levels for a single simulated LC injection 30 min into the reaction (substrate and product concentration each at 10 μM).

a positive bias in the data. Although normally distributed random numbers were used to simulate the noise, the negative numbers were replaced by zeros, in an attempt to represent chemical noise in the mass spectrum. This effectively increases the rank of the data matrix. It is interesting to note that when a higher level of random noise is present in the DAD profile, a more reliable estimation of the number of components is obtained for the LC–DAD–MS experiment than for the corresponding LC–MS experiment. This is probably caused by the noise in the DAD profile masking the effect of the bias caused by the MS noise.

The cause of the lowering of the SV_3-ratio seen at higher noise levels (>40%) in the MS direction and at moderately low noise levels in the DAD direction, is that both the third and fourth singular values have become significant and thus the difference between the third and fourth singular value is decreasing. At this point the SV_4-ratio indicates that four components are present, due to the high levels of non-random noise in the mass spectral data.

The analysis of the LC–DAD data indicated clear evidence for the presence of two components only for all noise levels (data not shown), where even at the highest noise level of 15%, the SV_3-ratio is <0.05.

Analysis of the LC–MS data (not shown) indicates that when the MS noise level increases the SV_3-ratio increases and levels off at 0.4 above the 15% noise level. This is again due to the increase in the significance of the fourth singular value.

3.1.3. Influence of noise level on resolution results

The same variations in the noise levels that were used to study the SVD results were used to examine the effect of the noise on the resolution results. The results for the analysis of the LC–DAD–MS data are shown in Fig. 3. The trend in fit error as a function of the noise level is similar in both the MS and DAD directions. When the LC–DAD or LC–MS data are analyzed individually, the effect of increasing the noise level in either the MS or DAD profiles on the fit error follows the same trend.

3.1.4. Influence of signal overlap

The algorithm used here underwent a thorough investigation of its behavior with respect to retention time and DAD spectral overlap, as described in a previous publication [8]. In the present work, the addition of MS data provides a more complete picture of the behavior of the algorithm.

Increases in the degree of chromatographic peak overlap or the similarity between the two DAD spectral profiles result in a corresponding increase in the fit error. However, changing the MS profile has a more

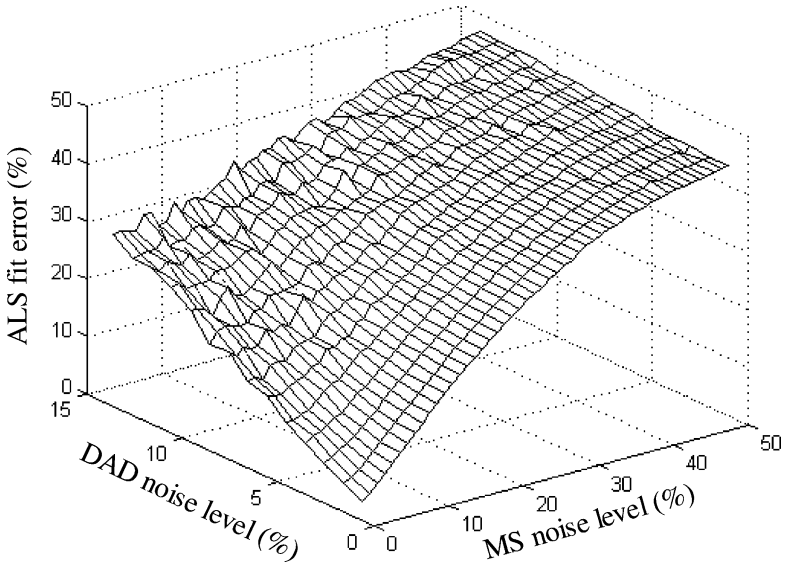

Fig. 3. The ALS fit error for LC–DAD–MS data as a function of the DAD noise level and the MS noise level for a single simulated LC injection 30 min into the reaction (substrate and product concentration of 10 μM).

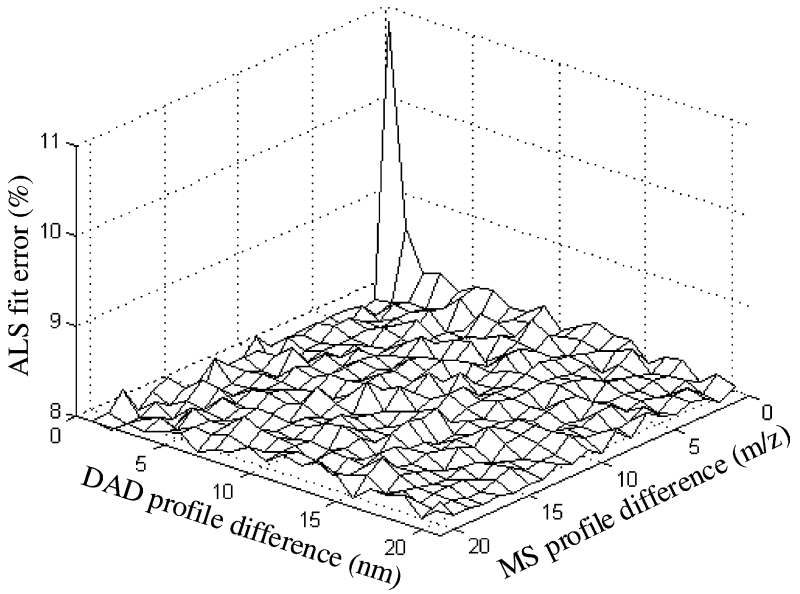

Fig. 4. The ALS fit error for LC–DAD–MS data as a function of difference in position of the Gaussian function in the DAD spectral profile and the difference in the molecular weight of the product and substrate ions for a single simulated LC injection 30 min into the reaction (substrate and product concentration of 10 μM).

abrupt influence on the fit results, as shown in Fig. 4. As soon as the MS spectra are different by 1 point (1 amu), the ALS algorithm benefits from the unique selectivity of that data. It results in complete resolution of the overlapped retention and spectral profiles and the fit error is no longer dependent on the degree of overlap of the DAD spectral profiles. This is the case even though the selectivity constraint was not applied during the analysis.

3.2. Analysis of one-way kinetic data

The kinetic fitting routine that has been incorporated into the ALS algorithm has been studied extensively using a range of different chemical kinetic models [4]. However, it has not previously been used to model of enzyme kinetics. The problem of model or parameter indistinguishability of the micro-rate constant description of enzyme kinetics is addressed here. Without the determination of the concentration profiles of the enzyme and the enzyme–substrate complex, and only the temporal profiles of the substrate and the product are available, the reaction can be accurately modeled by zero-order reactions in the case of a large excess of substrate. This section discusses how well the algorithm can estimate the rate constants from the available data when this simplified model is used.

For this study, the number of reaction time points in the simulated dataset remained at 10, however, the time intervals were changed to accommodate the longer duration of the reaction at higher substrate concentrations. The total reaction time was estimated by multiplying the initial substrate concentration by 100 and adding 5000 s. The time points were created using exponentially increasing time intervals.

3.2.1. Fits of individual kinetic profiles

When a consecutive first-order reaction is monitored by measuring the response of the product and/or the intermediate, k_1 and k_2 cannot be differentiated. This is known as the flip–flop effect [32]. When studying enzyme kinetics with one specific set of starting conditions, it might be possible that different combinations of rate constants would describe the same enzyme kinetic model to the same level of precision. Alternatively, different models may also fit the data to the same level of precision. To further explore these issues, the kinetic profiles are studied individually without

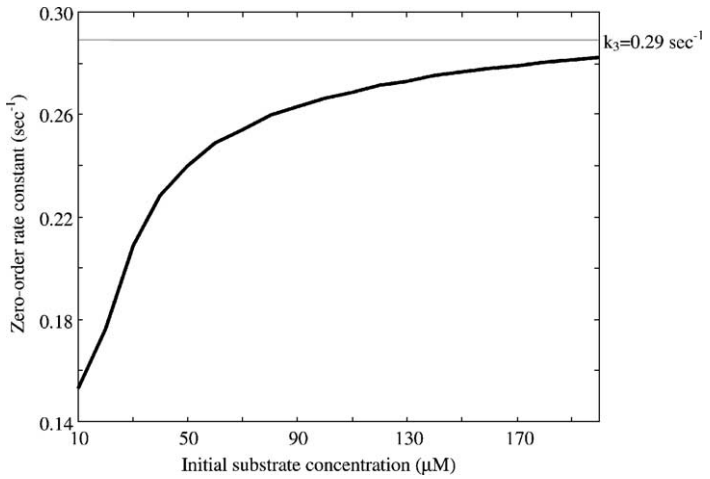

Fig. 5. The median zero-order rate constant (of 10 simulations) as estimated from fitting the enzyme kinetic concentration profiles of the substrate and the product as a function of the initial substrate concentration (——) and the true value for k_3 (——).

any of the corresponding chromatographic or spectral data. The rate constants used to simulate the enzyme kinetic profiles are kept constant at $0.58\,(\mu M\,s)^{-1}$, $4.0\,s^{-1}$ and $0.29\,s^{-1}$ for k_1, k_2 and k_3, respectively.

When using a zero-order reaction mechanism to fit the enzyme kinetic data, as the initial substrate concentration is increased, the concentration profiles will become linear and the zero-order rate constant will approach k_3, as shown in Fig. 5. The early points in the figure that correspond to low substrate concentrations, where a steady state is never reached, result in a poor fit to the kinetic profiles by the zero-order kinetic model and inaccurate estimates for the rate constant. At high substrate concentration levels, the rate constant for a

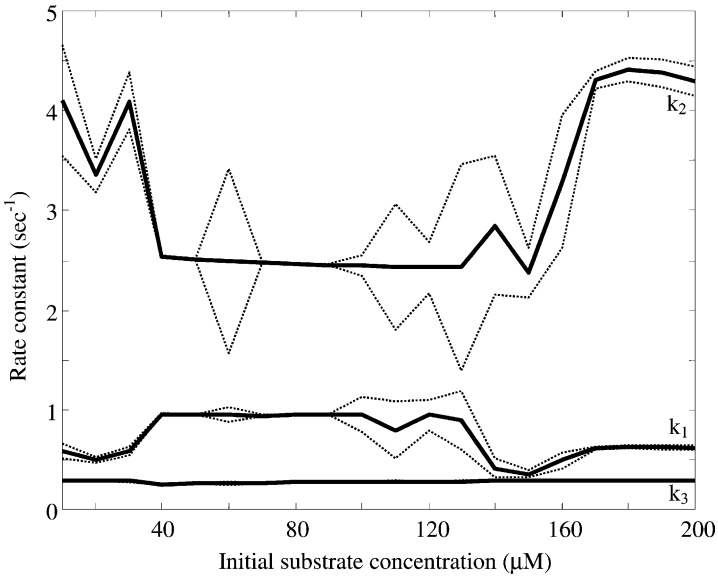

Fig. 6. The median (of 10 simulations) rate constants (——) for the enzyme kinetic model while fitting only the kinetic concentration profiles of the substrate and product, as a function of initial substrate concentration, ±1S.D. (.....).

fit of the enzyme kinetic data to a zero-order chemical model approaches the k_3, giving the true maximum substrate conversion rate (v_{max}). This is the assumption often used in Michaelis–Menten enzyme kinetic studies, where a single time point measurement is used to estimate v_0.

Fig. 6 shows the estimated rate constants as a function of the initial substrate concentration obtained by fitting these data to the correct enzyme kinetic model. All rate constants are fit accurately at lower (<40 μM) and higher (>160 μM) substrate concentrations. However, at intermediate substrate concentration levels, the kinetic fitting algorithm finds rate constants of 0.96, 2.53 and 0.26 s^{-1} for k_1, k_2 and k_3, respectively, which correspond to 16 nmol (nmol min)$^{-1}$ and 3.5 μM for v_{max} and K_m, respectively, instead of the true values of 17.3 nmol (nmol min)$^{-1}$ for v_{max} and 7.5 μM for K_m. The fitting routine finds a local minimum at that combination of rate constants. The larger standard deviations observed for k_1 and k_2 for some initial substrate concentrations are also indicative of some of the 10 simulations converging to the alternate minimum.

3.3. Fit of three-way data

3.3.1. Analysis of trilinear three-way data

As was seen in Section 3.2.1 with the analysis of kinetic data, the fitting algorithm might find a local minimum in the solution for the enzyme kinetic model. This, however, was the result of a single run of the simplex search algorithm. Many simplex optimizations are carried out during each ALS analysis. Having resolved the retention, spectral and kinetic profiles, the kinetic profiles are fit to the enzyme kinetic model, resulting in a set of rate constants. These rate constants are subsequently used to simulate the kinetic profiles and are used to recalculate the spectral and retention profiles. The rate constants found, are used in the next ALS iteration as the starting point for the simplex search, such that the kinetic profiles are slightly changed from the previous iteration. One would expect that the chances of the algorithm getting stuck in a local minimum during these multiple starts of the simplex algorithm within ALS, from different starting points at each ALS iteration, will be limited. In this section, we investigated just that LC–DAD–MS, LC–DAD and LC–MS three-way datasets were generated using different initial substrate concentrations

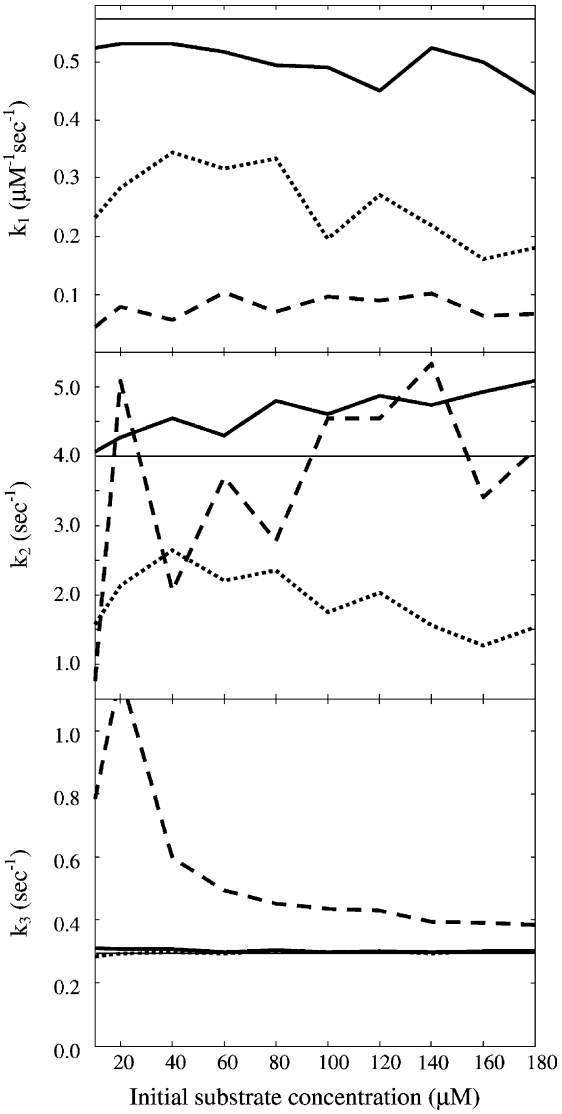

Fig. 7. The median rate constants k_1, k_2 and k_3 (of 10 simulations) from the analysis for the LC–DAD–MS data (——), LC–DAD data (- - -), LC–MS data (.....) and the true values (——) using ALS with the kinetic constraint from the analysis of three-way data for different initial substrate levels.

while maintaining a total enzyme concentration of 0.03 μM. Fig. 7 shows the median of the results of analyzing 10 different simulated three-way datasets using ALS with the implementation of the kinetic constraint. In comparing Fig. 7 with Fig. 6, it is immediately obvious that the algorithm no longer finds an alternate

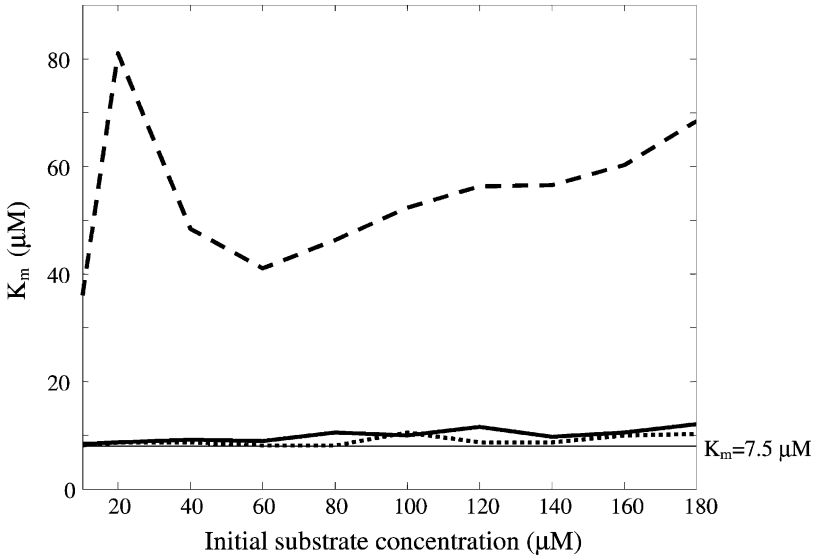

Fig. 8. The Michaelis constant (K_m) as calculated by $(k_2 + k_3)/k_1$ from the analysis for the LC–DAD–MS data (——), LC–DAD data (- - -), LC–MS data (.....) and the true value (——) using ALS with the kinetic constraint from the analysis of three-way data for different initial substrate levels.

solution for the rate constants in the 40–160 μM range for the initial substrate concentration.

A closer examination of the results reveals some more characteristics of this analysis. The analysis of the LC–MS data (dotted line) shows an underestimation of k_1 and k_2. However, the Michaelis constant ($K_m = (k_2 + k_3)/k_1$), as shown in Fig. 8, is still similar to the true value. The increased noise level of the MS data does not allow the algorithm to distinguish between the correct values for the micro-rate constants and those found here, but does allow for the correct estimation of Michaelis constant.

The analysis of the LC–DAD data (dashed line) is inaccurate over the entire range. This is caused by the incomplete resolution of the overlapped retention profiles. (Some of the substrate absorbance is wrongly attributed to the product.) Especially with a lower initial substrate concentration, slow product formation is artificially enhanced by this incomplete resolution and the observed k_3 is overestimated.

The three-way ALS resolution of the LC–DAD–MS kinetic data (solid line) takes advantage of the selectivity in the mass spectrum to better resolve the overlapped retention profiles, while at the same time the lower noise level of the DAD data allows the algorithm to more accurately determine the micro-rate constants.

3.3.2. Analysis of non-trilinear three-way data

Unless the experimental conditions are very carefully controlled, most multi-dimensional datasets will show non-trilinear characteristics caused by minor retention time shifts and other non-ideal behavior. Non-trilinear LC–DAD–MS, LC–DAD and LC–MS three-way datasets were generated, using an initial substrate concentration of 20 μM and total enzyme concentration of 0.03 μM. The chromatographic peak corresponding to the product was simulated to shift towards the substrate peak as the reaction progressed, resulting in lower chromatographic resolution and thus breaking the trilinear structure of the data. These shifts were generated to simulate column deterioration and temperature drift that may occur during real chromatographic experiments. The maximum chromatographic peak shift was increased systematically from 0 to 9 s, giving a total of nine shifted experimental datasets, as well as a dataset with no peak shifts. These maximum shifts are represented on the *x*-axis in Fig. 9. Each dataset was analyzed using two different combinations of ALS constraints.

Fig. 9 shows the results for the median of 10 simulations of LC–DAD–MS data, LC–DAD data and LC–MS data, analyzed using ALS and constraining all components to trilinearity. The observed rate con-

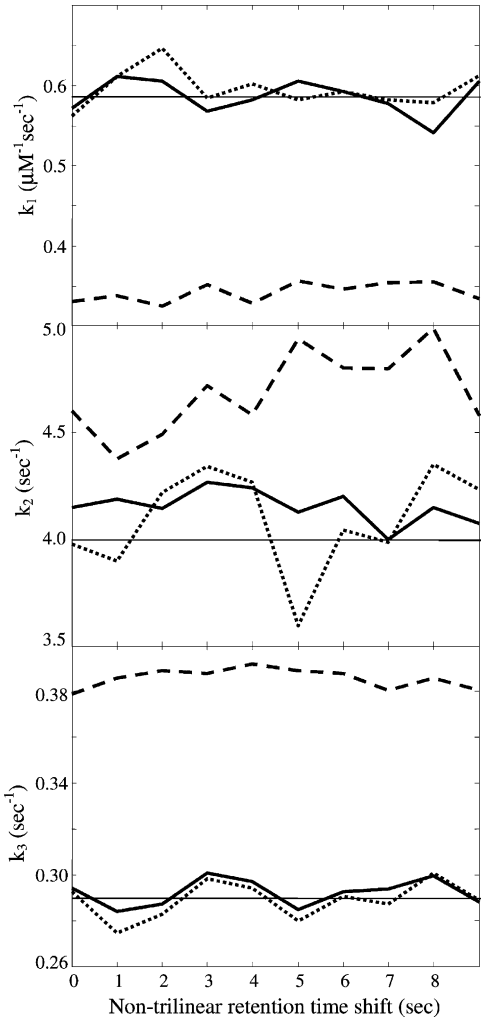

Fig. 9. The median rate constants k_1, k_2 and k_3 from the analysis for the LC–DAD–MS data (——), LC–DAD data (- - -), LC–MS data (.....) and the true values (——) using ALS while constraining all components to trilinearity from the analysis of three-way data with a non-trilinear retention time shift of the product peak, using an initial substrate concentration of 20 μM.

stants are independent of the retention time shift of the product signal. There is an increase in fit error and a broadening of the retention profile of the product component with an increase in the peak shift of the product as the reaction progresses. However, this does not significantly affect the resolved kinetic profiles for the substrate and the product and thus does not change the outcome of the kinetic fitting.

When the product signal was no longer constrained to trilinearity, the resulting rate constants (data not shown) demonstrate a nearly identical trend to that observed when both components were constrained by trilinearity. As a result of these studies we conclude that there is no effect of these chromatographic peak shifts on the values of the rate constants.

4. Conclusions

In this paper, we have examined the influence of the addition of MS data to LC–DAD data on the accuracy of the estimated rate constants obtained from fitting simulated enzyme kinetics data.

Incorrect rate constants are found for individual runs of the kinetic fitting routine results in non-optimal solutions due to algorithm converging on a local minimum. The use of multiple iterations within the ALS algorithm improves the convergence properties of the kinetic fitting routine and increases the accuracy of the resulting parameters. If the three-way data obtained from each different substrate level are combined in a four-way dataset, and analyzed by a four-way resolution algorithm, the remaining ambiguities should be eliminated.

Analyzing non-trilinear data increases the overall fit error when the data is fit to a trilinear model. However, the resolved kinetic profiles are not significantly changed. Therefore, the estimated rate constants are not significantly influenced by these non-trilinearities, even in the case of a significant shift of the chromatographic peaks.

Our analysis of the LC–DAD data indicates the importance of successful curve resolution for accurate estimation of the kinetic constants. When utilizing LC–DAD data, it is probably advisable to experiment with the selectivity constraint to improve the resolution results to obtain reasonable kinetic parameter estimates.

The higher noise level found in MS data reduces the accuracy in the estimation of the micro-rate constants using MS data alone, without spectroscopic data. The analysis of the LC–MS data, however, provides a good estimate for the Michaelis constant. The accuracy of the Michaelis constant is typically more important for real metabolism studies. In addition, MS data contain much more structural information than DAD data,

which is beneficial for the elucidation of reaction pathways. There is also an advantage in using the MS data for the resolution of overlapped chromatographic peaks as the unique selectivity found in MS data assists the chemometric algorithm in resolving the overlapped signals.

Using the combined LC–DAD–MS data results in the best estimates for the micro-rate constants (especially k_1 and k_2). The analysis of either LC–DAD–MS or LC–MS data provides accurate estimations of the K_m and v_{max} values.

The analysis of the higher order data resulting from the combination of hyphenated chromatographic instrumentation with spectroscopic and mass spectrometric detection shows promise for the characterization of reaction kinetics. The results from these studies provide useful guidance for the design of enzyme kinetic experiments in conjunction with multi-way chemometric methods.

Acknowledgements

This research was sponsored by the Jeffress Trust (J-517) and by grant CHE-0076290 from the National Science Foundation.

References

[1] D.C. Harris, Quantitative Chemical Analysis, 3rd ed., Freeman, New York, 1991.
[2] R.M. Silverstein, G.C. Bassler, T.C. Morrill, Spectral Identification of Organic Compounds, 3rd ed., Wiley, New York, 1991.
[3] L. Stryer, Biochemistry, 3rd ed., Freeman, New York, 1988.
[4] E. Bezemer, S.C. Rutan, Chemom. Intell. Lab. Syst. 59 (2001) 19–31.
[5] E. Sanchez, B.R. Kowalski, J. Chemom. 4 (1990) 29–45.
[6] K.S. Booksh, Z. Lin, Z. Wang, B.R. Kowalski, Anal. Chem. 66 (1994) 2561–2569.
[7] R. Gargallo, R. Tauler, F. Cuesta Sanchez, D.L. Massart, Trends Anal. Chem. 15 (1996) 279–286.
[8] E.C. Bezemer, S.C. Rutan, Chemom. Intell. Lab. Syst. 60 (2002) 239–251.
[9] A. de Juan, Y. Vander Heyden, R. Tauler, D.L. Massart, Anal. Chim. Acta 346 (1997) 307–318.
[10] E. Furusjo, L. Danielson, Anal. Chim. Acta 373 (1) (1998) 83–94.
[11] K. Héberger, A.P. Borosy, J. Chemom. 13 (1999) 473–489.
[12] S. Bijlsma, H.F.M. Boelens, A.K. Smilde, Appl. Spectr. 55 (2001) 77–83.
[13] M. Maeder, A. Zilian, Chemom. Intell. Lab. Syst. 3 (1988) 205–213.
[14] J. Chen, S.C. Rutan, Anal. Chim. Acta 335 (1996) 1–10.
[15] R. Tauler, D. Barcelo, Trends Anal. Chem. 12 (1993) 319–327.
[16] C. Demir, P. Hindmarch, R.G. Brereton, Analyst 121 (1996) 1443–1449.
[17] F. Gong, Y.Z. Liang, Q.S. Xu, J. Chen, J. Chromgr. A 905 (2001) 193–205.
[18] F. Gong, Y.Z. Liang, H. Cui, F.T. Chau, B.T.P. Chan, J. Chromgr. A 909 (2001) 237–247.
[19] B.J. Prazen, K.J. Johnson, A. Weber, R.E. Synovec, Anal. Chem. 73 (2001) 5677–5682.
[20] J.S. Salau, M. Honing, R. Tauler, D. Barcelo, J. Chromgr. A 795 (1998) 3–12.
[21] S. Dunkerley, J. Crosby, R.G. Brereton, K.D. Zissis, R.E.A. Escott, Analyst 123 (1998) 2021–2033.
[22] S. Dunkerley, R.G. Brereton, J. Crosby, Chemom. Intell. Lab. Syst. 48 (1999) 99–119.
[23] K.D. Zissis, S. Dunkerley, R.G. Brereton, Analyst 124 (1999) 971–979.
[24] E.C. Bezemer, S.C. Rutan, Anal. Chem. 73 (2001) 4403–4409.
[25] A.M. Di Guilmi, N. Mouz, Y. Petillot, E. Forest, O. Dideber, T. Vernet, Anal. Biochem. 284 (2000) 240–246.
[26] Q. Mei, C. Tang, Y. Lin, T.H. Rushmore, M. Shao, Drug Metab. Disposit. 30 (2002) 701–708.
[27] Mathworks, MATLAB, version 6.0, Natick, MA, 1999.
[28] H.R. Keller, D.L. Massart, Anal. Chim. Acta 246 (1991) 379–390.
[29] http://www.ub.es/gesq/mcr/mcr.htm.
[30] R. Tauler, A. Smilde, B.R. Kowalski, J. Chemom. 9 (1995) 31–58.
[31] R. Tauler, Chemom. Intell. Lab. Syst. 30 (1995) 133–146.
[32] J.W. Moore, R.G. Pearson, Kinetics and Mechanism, 3rd ed., Wiley, New York, 1981.