J. Am. Chem. Soc. **1996,** *118,* 10331−10332

Generalized Rank Annihilation Method Applied to a Single Multicomponent Pulsed Gradient Spin Echo NMR Data Set

Brian Antalek* and Willem Windig

Imaging Research and Advanced Development
Eastman Kodak Company
Rochester, New York 14650-2132

Received June 27, 1996

Characterizing mixtures can be a time intensive process, especially when faced with significant spectral overlap. Pulsed gradient spin echo (PGSE) NMR offers a means for resolving the NMR spectra of individual components in a multicomponent solution based upon their respective self-diffusion coefficients. The experiment provides a spectral data set in which the signal amplitude of every resonance associated with each component decays exponentially at the same rate. There are several ways to process the data set to obtain the spectra from the individual components as well as their respective self-diffusion coefficients. Diffusion ordered spectroscopy (DOSY)[1] uses programs which fit a multiexponential function to single or groups of "data channels", but is limited in separating highly overlapped spectral regions. Component resolved (CORE) NMR was recently introduced[2] and utilizes, through an iterative procedure, the whole spectrum and not just data channels for accomplishing component separation. This approach resolves highly overlapped regions but is time intensive. A third approach utilizes a multivariate statistical analysis based upon a procedure developed by Kubista[3] which uses the complete band shape for determining individual spectra and diffusivities.[4] This method requires two data sets where the spectral amplitudes of the components vary differently for each set. Schulze and Stilbs[4] deal with this requirement with use of a variant of the conventional PGSE NMR experiment that produces two data sets but compromises the integrity of the spectral band shape. We present an alternative approach based upon a technique similar to that of Kubista, which uses a *single* PGSE data set acquired under standard conditions that preserves spectral integrity.

PGSE NMR is based on the method of Stejskal and Tanner[5] and is derived from the nuclear spin echo concept of Hahn[6] and Carr and Purcell.[7] The diffusion coefficient is obtained from the attenuation of the spin echo under the influence of pulsed magnetic field gradients.[8]

The experiment was performed at 45 °C using the LED[9] pulse sequence by varying the gradient strength (g) and keeping the gradient width (δ) and all other timing parameters constant. A value of 400 ms was used for the diffusion time (Δ) and 3.5 ms for δ. The quantitative relationship between the echo attenuation and self-diffusion coefficient for component i is given by the following equation

$$\left(\frac{E}{E_o}\right)_i = e^{-D_i(\gamma g\delta)^2(\Delta-\delta/3)} \tag{1}$$

where γ is the gyromagnetic ratio of the ^{1}H nucleus, E is the measured signal amplitude, and E_o is the amplitude with no gradients.

Table 1. Representative Signal Decay Data

D		*A*		*B*	
27	8	27	8	9	4
9	4	9	4	3	2
3	2	3	2	1	1
1	1				

Kubista showed that it is possible to obtain an exact analytical mixture solution when two correlated data sets are available.[3,10]

$$\mathbf{A} = \mathbf{CP} \tag{2}$$

$$\mathbf{B} = \mathbf{C}\boldsymbol{\beta}\mathbf{P} \tag{3}$$

A and **B** are data matrices of size cv; c is the number of spectra, and v is the number of spectral data points. **C** (size cn) and **P** (size nv) are the concentration and pure spectra matrices of n components. The matrix β is a diagonal matrix (size nn). In these two data sets the pure spectra and the concentrations have a correlation coefficient of 1, they only differ by a certain scaling factor as defined by the matrix β.

Booksh and Kowalski[11] showed that the Kubista method can be expressed in terms of the generalized rank annihilation method (GRAM), which eliminates some problems and restrictions with the Kubista method. For this study, the algorithm of GRAM as proposed by Wilson *et al.*[12] is used.

The formulation of the problem in the form of eq 2 and 3 imply the use of two data sets, obtained under different experimental conditions.[4] This, however, is not necessary for PGSE NMR data, because the decaying profiles are exponential. This is illustrated with the data in Table 1. *D* represents a data set with two components, of which the exponentially decaying concentration profiles are given in the first and second column under *D*. The first column has a faster decay than the second column. The data set *A* is created using the first three "spectra" of data set *D*, and data set *B* is created by using the last three spectra of data set *D*. Because of the exponential character of the profiles, the first column of *A* is a constant (three) times the first column of column *B*, and the second column of *A* is a constant (two) times the second column of *B*. As a consequence, the data sets *A* and *B* fulfill the requirements as expressed in eqs 2 and 3. This means that in the case of exponentially decaying functions, one can use two different parts of the data set to create the two data sets necessary to apply the Kubista method.

For the experiments described in this paper, the two data sets for the data analysis were generated by using spectra 1−12 and 2−13 from the original data set with 13 spectra (each spectrum contains 4095 data points). The sample is comprised of 0.1% w/w TX-100, a nonionic surfactant ($(CH_3)_3CCH_2C(CH_3)_2$-$C_6H_4O(CH_2CH_2O)_xH$, $x = 10$ on average, Eastman Kodak Company, Rochester, NY) and 5% w/w gelatin (deionized, Eastman Gelatin, Peabody, MA). The measurements were carried out at 30 °C on a Varian Inova 400 wide bore (89 mm) spectrometer operating at 399.9 MHz and equipped with a standard pulsed field gradient (PFG) accessory.

For the data analysis, the function gram_wsk of the Matlab (The MathWorks, Inc., Cochituate Place, 24 Prime Park Way, Natich, MA 01760) Second Order Toolbox, version 1.1, of the Center for Process Analytical Chemistry (University of Washington, Box 351700, Seattle, WA 98195) was used. The

---

[1] Morris, K. F.; Johnson, C. S., Jr. *J. Am. Chem. Soc.* **1993**, *115*, 4291−4299.

[2] Stilbs, P.; Paulsen, K.; Griffiths, P. C. *J. Phys. Chem.* **1996**, *100*, 8180−8189.

[3] Kubista, M. *Chemom. Intell. Lab. Syst.* **1990**, *7*, 273−279.

[4] Schulze, D.; Stilbs, P. *J. Magn. Reson, Ser. A* **1993**, *105*, 54−58.

[5] Stejskal, E. O.; Tanner, J. E. *J. Chem. Phys.* **1965**, *42*, 288−292.

[6] Hahn, E. L. *Phys. Rev.* **1950**, *80*, 580−594.

[7] Carr, H. Y.; Purcell, E. M. *Phys. Rev.* **1954**, *94*, 630−638.

[8] Stilbs, P. *Prog. Nucl. Magn. Reson. Spectrosc.* **1987**, *19*, 1−45.

[9] Gibbs, S. J.; Johnson, C. S., Jr. *J. Magn. Reson.* **1991**, *93*, 395−402.

[10] Scarminio, I.; Kubista, M. *Anal. Chem.* **1993**, *65*, 409−416.

[11] Booksh, K. S.; Kowalski, B. R. *J. Chemom.* **1994**, *8*, 287−292.

[12] Wilson, B.; Sanchez, E.; Kowalski, B. R. *J. Chemom.* **1989**, *3*, 493−498.

S0002-7863(96)02172-5 CCC: $12.00

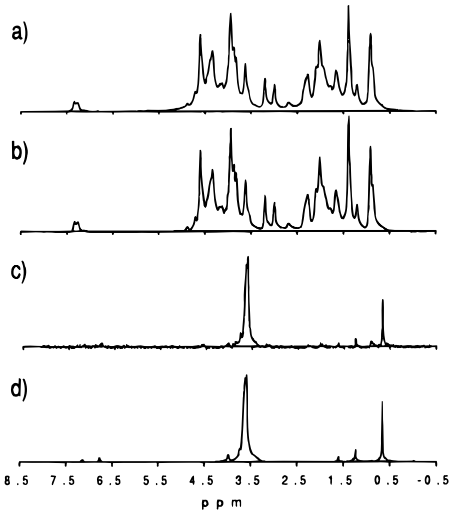

Figure 1. Spectra from the components of the described sample: (a) the first extracted spectrum; (b) a reference spectrum of gelatin; (c) the second extracted spectrum; (d) a reference spectrum of TX-100.

gram_wsk function applies the algorithm described by Wilson *et al.*[12] The computer configuration is a PENTIUM, 90 MHz, 64 MB of RAM. Processing time for the data set described above was less than 10 s.

Figure 1a−d represents the extracted spectra for the individual components along with their representative pure spectra from corresponding single-component solutions obtained using the same experimental parameters. The extracted spectrum for gelatin is virtually identical to the reference spectrum which is not surprising considering it is the dominant component in the solution. Although some noise can be observed in the extracted spectrum of TX-100, the similarity is striking, considering the low concentration.

Figure 2 represents the decay function of the surfactant TX-100. This plot clearly confirms the exponential character of the data. Since the signal behavior for each component is determined from the whole band shape and not just one resonance, more precise self-diffusion coefficients in principal may be obtained. We have obtained self-diffusion coefficients of 8.86 × 10^{-11} m^2/s and 1.47 × 10^{-11} m^2/s for TX-100 and gelatin, respectively. These values compare well with those measured directly from analysis of integral regions: 9.08 × 10^{-11} m^2/s for surfactant (0.6−0.7 ppm, *tert*-butyl moiety) and 1.40 × 10^{-11} m^2/s for gelatin (1.2−2.5 ppm). In this case, the gelatin signal was only attenuated by 40%. This is not enough to observe a significant multiexponential effect from molecular weight polydispersity.

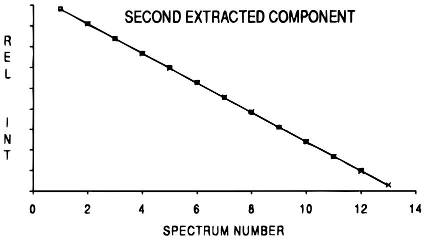

Figure 2. The concentration profile of the second component, TX-100, for the first and second data sets. The log of the relative intensity is plotted on the *y* axis. Data from both sets completely overlap.

In summary, the application of the GRAM technique has been shown to be successful as a fast-processing scheme for obtaining spectra from individual components in a mixture using a simple experiment. There are several aspects of the technique which make it amenable for *routine* mixture characterization including the ability to successfully handle highly overlapped spectra, the need for a relatively low number of spectra (10−15), and the ability for full spectral analysis (typically up to 16K data points each). Due to the fact that an exact solution is obtained the technique may have the potential to automatically extract the number of components present in the mixture with a set of appropriate flags in the routine. Full limitations have not yet been realized, however, since as the difference in relative diffusivities become smaller, criteria such as good signal-to-noise and exponential purity of the signal decay become increasingly important.

JA962172V