Mixture Analysis in Combinatorial Chemistry. Application of Diffusion-Resolved NMR Spectroscopy

Mengfen Lin and Michael J. Shapiro*

Preclinical Research, Sandoz Research Institute, Sandoz Pharmaceuticals Corporation, 59 Route 10, East Hanover, New Jersey 07936

Received July 10, 1996

Combinatorial chemistry has recently attracted great attention in the pharmaceutical industry. This technique quickly generates large numbers of structurally diverse compounds for use in a variety of high-capacity biological screens.[1−4] It is rapidly becoming a dominant trend in the search for lead compounds in drug discovery. One aspect of combinatorial chemistry is the synthesis of mixtures of structurally related compounds. This facilitates a high throughput in both synthesis and screening. One of the methodologies in combinatorial chemistry is the split and mix synthesis.[5−7] This technique produces a mixture of compounds as the final product. In developing the synthesis of very large libraries, smaller test systems are studied to optimize synthetic strategies. NMR spectroscopy has been widely used in chemical structural analysis after prior separation of mixture components. The utility of NMR for studying intact mixtures has not been extensively demonstrated. As part of our efforts in analysis of combinatorial chemistry reactions, we report here the development of a new method for the analysis of complex mixtures.

2D NMR methods such as TOCSY have been used in relatively simple mixture analysis.[8] However, for compounds that have their spin systems insulated by groups such as esters or ethers, TOCSY methodologies are not sufficient for complete analysis. Recently, the use of pulsed field gradient (PFG) technology has been demonstrated as an useful technique for mixture analysis.[9−12] In these experiments, signals arising from different compounds can be separated on the basis of their diffusion coefficients. In small organic molecule mixtures, the rate of diffusion in solution can be used to distinguish components on the basis of their molecular size variance. A high-resolution diffusion spectroscopic technique called "HR-DOSY" has also been reported to analyze a complex biological mixture.[13] However, in cases where the spectrum is severely overlapped, as is often the case in combinatorial mixtures, and the diffusion coefficients of each component are similar, this method alone is also insufficient for spectral assignment. The challenge of combinatorial generated mixtures clearly requires new methods of analysis. We have developed a new method, which involves the use of PFG and TOCSY, which we call diffusion encoded spectroscopy "DECODES" to address this mixture analysis problem.

A mixture of five esters was used to simulate the results of a split and pool combinatorial synthesis: (±)-*sec*-butyl acetate, propyl acetate, ethyl butyrylacetate, isopropyl butyrate, and butyl levulinate. The 1D ^{1}H NMR spectrum of this mixture is shown in Figure 1. The proton resonances in the region between 0.5 and 2 ppm are severely overlapped. A TOCSY experiment was conducted for the mixture, and each coupled spin can be identified, e.g., propyls and butyls. However, it is not possible from this data to link the spin systems through the silent ester moiety to spin connectivities and, hence, establish molecular identity.

Unlike any other 2D experiment, size-resolved or diffusion-resolved NMR assigns the resonances on the basis of the diffusion coefficient for each proton (or other spin) in the molecule and, therefore, can be used to distinguish resonances arising from different molecules.[9,14] However, in the region between 0.5 and 2 ppm, where the ^{1}H resonances are overlapped, components cannot be resolved by diffusion alone. This is a limitation in the present PFG methodology. TOCSY, on the other hand, can separate coupled spin systems in the second dimension, resolving potential overlapping resonances. Therefore, the combination of PFG and TOCSY "decodes" the spin systems, allowing individual components in complicated mixtures to be assigned.

To illustrate this concept, we have applied this to the same mixture of five esters. The pulse sequence used in DECODES is LEDmlev-17, in which a LED (longitudinal encode and decode) pulse sequence[15] is applied to spacially encode the molecule, and then mlev-17[16] is used for spin lock, as shown in Figure 2. In our DECODES experiment, 14 2D TOCSY spectra were collected as a function of gradient amplitude. This is similar to a recently reported "DOSY-NOESY" experiment.[17] The intensity of each cross peak is attenuated in each spectrum due to diffusion and gradient strength attenuation as shown in eq 1. In this equation, D is the

$$I = I_0 \exp[-D(\Delta - \delta/3)\gamma^2 g^2 \delta^2] \qquad (1)$$

diffusion coefficient, γ is the gyromagnetic ratio of the nuclei detected, Δ is the diffusion delay time, g and δ

(1) Thompson, L. A.; Ellman, J. A. *Chem. Rev.* **1996**, *96*, 555−600.
(2) Gallop, M. A.; Barrett, R. W.; Dower, W. J.; Fodor, S. P. A.; Cordon, E. M. *J. Med. Chem.* **1994**, *37*, 1233−1251.
(3) Gordon, E. M.; Barrett, R. W.; Dower, W. J.; Fodor, S. P. A.; Gallop, M. A. *J. Med. Chem.* **1994**, *37*, 1385−1401.
(4) Borman, S. *Chem. Eng. News* **1996**, 29−54.
(5) Geysen, H. M.; Rodda, S. J.; Mason, T. J. *Mol. Immunol.* **1986**, *23*, 709−715.
(6) Sebestven, F.; Dibo, G.; Kovacs, A.; Furka, A. *Biomed. Chem. Lett.* **1993**, *3*, 413−418.
(7) Fruka, A.; Sebestyen, F.; Asgedom, M.; Dibo, G. *Int. J. Pept. Protein Res.* **1991**, *37*, 487−493.
(8) Johnson, K.; Barrientos, L. G.; Le, L.; Murthy, P. P. N. *Anal. Biochem.* **1995**, *231*, 421−431.
(9) Morris, K. F.; Johnson, C. S., Jr. *J. Am. Chem. Soc.* **1992**, *114*, 3139−3141.
(10) Morris, K. F.; Johnson, C. S., Jr. *J. Am. Chem. Soc.* **1993**, *115*, 4291−4299.
(11) Morris, K. F.; Stilbs, P.; Johnson, C. S., Jr. *Anal. Chem.* **1994**, *66*, 211−215.
(12) Lin, M.; Jayawickrama, D. A.; Rose, R. A.; Delvisio, J. A; Larive, C. K. *Anal. Chim. Acta* **1995**, *307*, 449−457.
(13) Barjat, H.; Morris, G. A.; Smart, S.; Swanson, A. G. ; Williams, S. C. *J. Magn. Reson. Ser. B* **1995**, *108*, 170−172.
(14) Stilbs, P.; Paulsen, K.; Griffiths, P. C. *J. Chem. Phys.* **1996**, *100*, 8180−8189.
(15) Gibbs, S. J.; Johnson, C. S., Jr. *J. Magn. Reson.* **1991**, *93*, 395−402.
(16) Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355−360.
(17) Gozansky, E. K., and Gorenstein, D. G. *J. Magn. Reson.* **1996**, *111*, 94−96.

S0022-3263(96)01315-1 CCC: $12.00

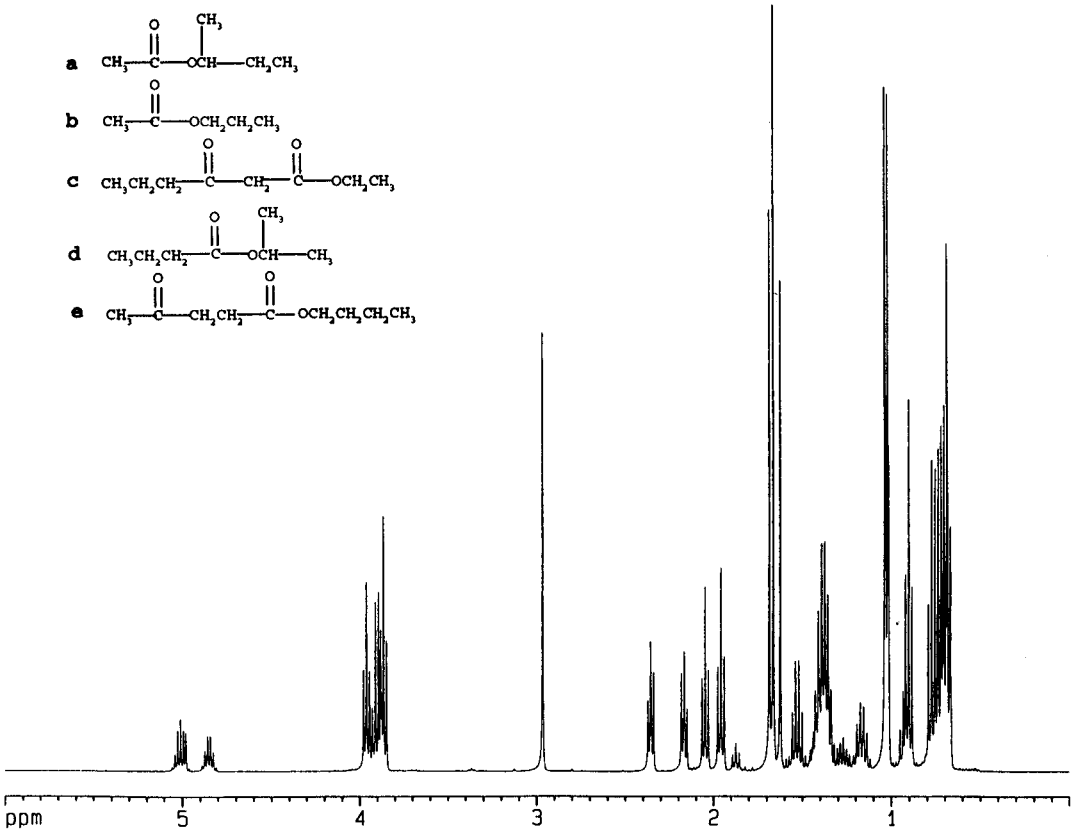

Figure 1. 1D ^{1}H NMR spectrum of the ester mixture in benzene at 298 K. The concentration of each ester is about 25 mM.

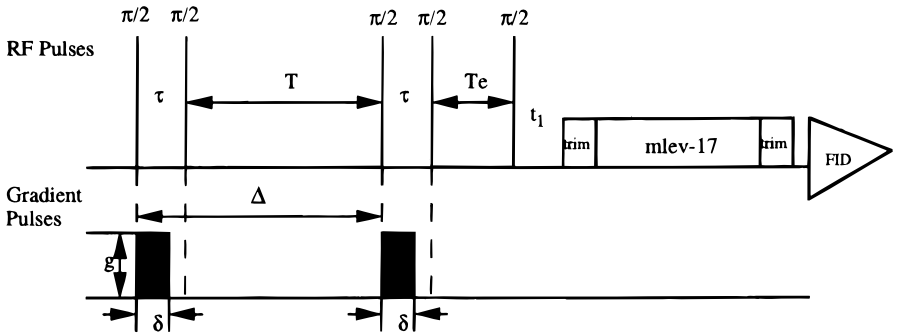

Figure 2. LEDmlev-17 pulse sequence. Δ is diffusion delay time; Te is eddy current delay time; δ is gradient pulse duration time; and g is gradient pulse amplitude.

Table 1. Diffusion Coefficient Measured for Each Component in the Ester Mixture

compd	diffusion coeff (m^2/s)	compd	diffusion coeff (m^2/s)
propyl acetate	$2.15 \pm 0.02 \times 10^{-9}$	isopropyl butyrate	$1.99 \pm 0.02 \times 10^{-9}$
butyl acetate	$2.27 \pm 0.04 \times 10^{-9}$	butyl levulinate	$1.67 \pm 0.05 \times 10^{-9}$
ethyl butyrate	$1.88 \pm 0.03 \times 10^{-9}$		

are the gradient pulse amplitude and duration time, respectively, and I_0 is the intensity of a peak at the limitation of g equals to zero. The delay time Δ and δ were kept as constant thoughout the experiment while g was attenuated from 0.03 to 0.4 T m^{-1} during the experiment. A typical TOCSY spectrum obtained in this manner is shown in Figure 3. As expected, the coupled spin systems are well separated.

Peaks are picked from the TOCSY spectrum, and the natural logarithms of the intensities of the peaks are plotted versus the square of the gradient strength. The plot yields a straight line, and the slope is determined by $-D(\Delta-\delta/3)\gamma^2\delta^2$. Since we have kept all of the delay times constant, the diffusion coefficient D can be easily calculated from the slope.

Figure 4 shows a typical ln I vs g^2 plot obtained for the cross peaks of butyl acetate and butyl levulinate, and the diffusion coefficients were calculated from the slopes. Shown in Table 1 are the calculated diffusion coefficients for each of the components. As long as there are resolved cross peaks in the spectrum, the diffusion coefficient can be calculated and molecular identity can be obtained. The use of TOCSY aids this possibility by providing multiple opportunities for such resolution to occur.

The DECODES method represents a new technique to perform mixture analysis without the need of prior

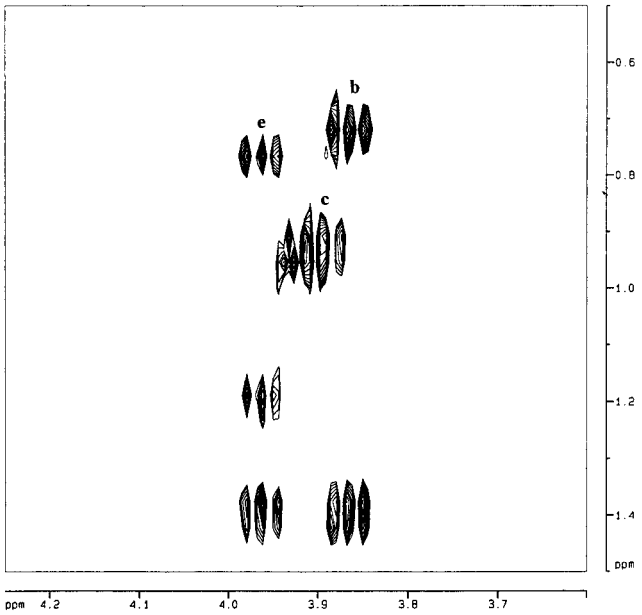

Figure 3. Selected region of the TOCSY spectrum of the ester mixture in benzene acquired with a gradient pulse strength of 0.165 T m^{-1}. Peaks arise from propyl acetate (b), ethyl butyrylacetate (c), and butyl levulinate (e).

separation. DECODES should prove useful in the identity of compounds from small split and mix synthetic

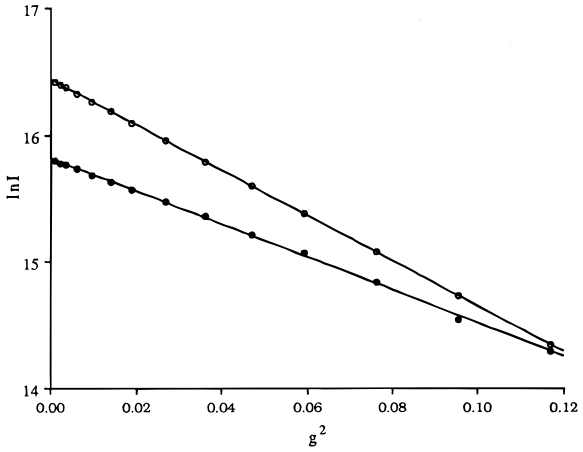

Figure 4. Plot of the natural logarithms of the intensities of peaks b (top circles) and e (bottom circles) labeled in Figure 3 vs g^2 (T^2 m^{-2}). The diffusion coefficients are calculated from the slopes.

pools as well as determination of diffusion coefficients in molecules with complicated overlapping spin systems.

Acknowledgment. We would like to acknowledge Jefferson Chin of Sandoz Research Institute for technical support. We also would like to thank Dr. Nirmala of Sandoz Research Institute and Dr. Mike Goekle of Bruker Instruments for useful discussions.

JO961315T