

Spectral Deconvolution of Chemical Mixtures by Covariance NMR

Fengli Zhang and Rafael Brüschweiler*^[a]

A method is presented for the deconvolution of the NMR spectrum of a chemical mixture without requiring physical separation of its components. The method, which is termed "Demix", is based on a principal component analysis of a series of one-dimensional (1D) spectra that are statistically modulated during

preparation and TOCSY mixing periods. The largest principal components correspond to the 1D NMR spectra of the scalar J-coupled spin networks of the individual components of the mixture. The method is demonstrated for aqueous mixtures of the amino acids Glu, Leu, Lys, and Val.

Introduction

The characterization of a mixture in terms of its components is a central task in analytical chemistry.^[1] Commonly used procedures first physically separate the chemical components, for example by chromatography, followed by chemical or spectroscopic analysis, including mass spectrometry, optical or NMR spectroscopy.^[1,2] Nonhyphenated approaches have become available using NMR: based on differential translational diffusion among the components, diffusion-ordered spectroscopy (DOSY) separates the NMR spectrum of the mixture along a second dimension.^[3,4]

Herein, an alternative nonhyphenated NMR-based method is presented, which is independent of the diffusion properties of the components. It deconvolutes the one-dimensional spectra of the different components due to differences in their scalar J-coupling networks and chemical shifts. It is based on the covariance matrix of a set of 1D NMR spectra with different preparation and mixing times, which are processed by principal component analysis.^[5] The largest principal components then correspond in good approximation to the one-dimensional spectra of the coupled spin networks of the components of the mixture.

Principal component analysis has been previously applied to a series of 1D in vivo NMR spectra for the quantification of the position, amplitude, and phase of a single peak^[6–11] and in food chemistry to 1D spectra of different samples.^[12] In contrast, herein, the variations in the 1D spectra were induced by variations in the pulse sequence in the spirit of multidimensional NMR spectroscopy.^[13]

Computational Methods

Covariance NMR spectroscopy identifies correlated changes of the magnetizations of spins from a set of 1D spectra. The 1D spectra were obtained by using the TOCSY preparation scheme:^[14] $(90^\circ)_\varphi$ - t_1 -Isotropic Mixing (τ_m)-Detection (t_2), where $(90^\circ)_\varphi$ corresponds to a 90° excitation pulse with phase φ , $t_1 = k\Delta t_1$ denotes a free evolution period, which is an integer (k) multiple of the time increment Δt_1 , preceding the isotropic mixing period of duration τ_m . The free induction decay (FID) is detected in quadrature mode at N_2 time points $t_2 = n\Delta t_2$ ($n =$

$0, \dots, N_2 - 1$) yielding the complex time-domain data matrix $s(k\Delta t_1, \tau_m, n\Delta t_2)$. For each duration of the evolution time $t_1 = k\Delta t_1$ ($k = 0, \dots, N_1 - 1$), FIDs were co-added with different mixing times τ_m ($m = 1, \dots, M$) and discrete Fourier transformed with respect to t_2 leading to the real $N_1 \times N_2$ data matrix, Equation (1)

$$S(k, l) = \text{Re} \sum_{m=1}^M \sum_{n=0}^{N_2-1} s(k\Delta t_1, \tau_m, n\Delta t_2) \exp\{-i2\pi nl/N_2\} \quad (1)$$

representing the N_1 absorption spectra with N_2 data points each, where Re defines the real part. This is, in fact, a set of spectra $S(t_1, \omega_2)$ in a mixed time-frequency domain, averaged over different mixing times. $S(k, l)$ is then converted into the $N_2 \times N_2$ covariance matrix \mathbf{C} with elements given by Equation (2)^[15]

$$C_{ij} = \frac{1}{N_1 - 1} \sum_{k=0}^{N_1-1} (S(k, i) - \langle S(i) \rangle)(S(k, j) - \langle S(j) \rangle) \quad (2)$$

where $\langle S(i) \rangle$ is the average spectrum given by $\langle S(i) \rangle = N_1^{-1} \sum_{k=0}^{N_1-1} S(k, i)$. Element C_{ij} is the mathematical covariance between the amplitudes at positions i and j of the 1D spectra corresponding to resonance frequencies $\nu_2(i)$ and $\nu_2(j)$, where $\nu_2(j) = -SW/2 + (SW/N_2)j$ with the spectral width $SW = 1/\Delta t_2$. In the next step, a principal component analysis^[5] was applied to the covariance matrix \mathbf{C} by diagonalization, $\mathbf{C}\mathbf{v}_u = \lambda_u \mathbf{v}_u$, where λ_u is the eigenvalue to the eigenvector \mathbf{v}_u . The eigenvectors to the dominant eigenvalues λ_u represent 1D spectra of the individual components of the mixture, as is demonstrated in the following examples. Because this method serves for the spectral deconvolution of mixtures, it is termed "Demix". It represents an application of covariance NMR spectroscopy^[15] that

[a] Dr. F. Zhang, Dr. R. Brüschweiler
Carlson School of Chemistry and Biochemistry, Clark University
950 Main Street, Worcester, MA 01610-1477 (USA)
Fax: (+1) 508-793-8861
E-mail: bruschwiler@nmr.clarku.edu

extracts the relevant information from the covariance matrix by means of a principal component analysis.

Results and Discussion

A mixture consisting of the three amino acids glutamic acid (Glu), lysine (Lys), and valine (Val) was prepared in a D₂O buffer. Approximate concentrations of the different amino acids were around 7.0 mM. A 1D proton NMR spectrum of the mixture is shown in Figure 1 (top) together with 1D spectra of the individual amino acids prepared as separate samples (right column).

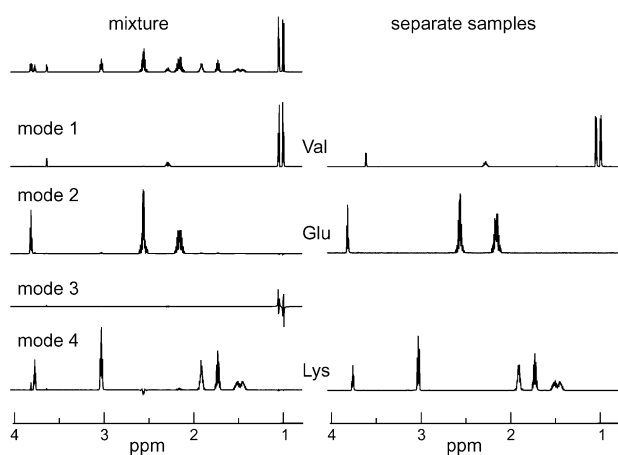


Figure 1. Deconvolution covariance NMR method applied to an aqueous three amino acid mixture containing Glu, Lys, and Val. The 1D NMR spectrum of the mixture is shown on top of the left column and the 1D spectra of the individual amino acids, recorded as separate samples, are plotted in the right column. The four largest eigenmodes of the covariance matrix defined in Equation (2) are plotted in the left column. Eigenmode 3 does not represent a complete amino acid in the mixture and can be easily identified by its large positive and negative amplitudes. All experiments were collected at 600 MHz (14 T) and 298 K.

A set of $N_1=1024$ 1D spectra was collected using the TOCSY pulse sequence shown above with MLEV-17 as the mixing sequence^[16] with an increment $\Delta t_1=50 \mu\text{s}$. The phase φ_1 was subjected to time-proportional phase incrementation (TPPI).^[17] For each t_1 , transients for five different mixing times were co-added with $\tau_m=45, 62, 76, 97, 120$ ms. These mixing times were chosen randomly with the purpose of ensuring rather uniform magnetization transfers within the different spin systems.

After Fourier transformation along t_2 , Equation (1), the 1D spectra consisting of $N_2=1024$ points each were converted into the covariance matrix according to Equation (2), which was subjected to diagonalization yielding 1024 eigenmodes with their corresponding eigenvalues. The four eigenmodes belonging to the four largest eigenvalues are plotted in the left column of Figure 1. Comparison with the right column corresponding to the spectra of separate amino acids shows that three out of the four largest eigenmodes correspond in very good approximation to the 1D spectra of the different amino acids. The third largest eigenmode (mode 3) is localized on the

methyl resonances of Val and does not correspond to an individual amino acid. This mode and the other $N_2-4=1020$ modes that do not belong to individual components of the mixture can be identified easily by their ratio of the variance of all positive vector elements to the variance of all negative vector elements, which is close to 1. In contrast, modes that correspond to individual amino acids have ratios of 2.5 or larger. Figure 2 demonstrates the method with a four amino acid mixture containing Glu, leucine (Leu), Lys, and Val. Despite the spectral overlap of leucine and lysine, the method performed well.

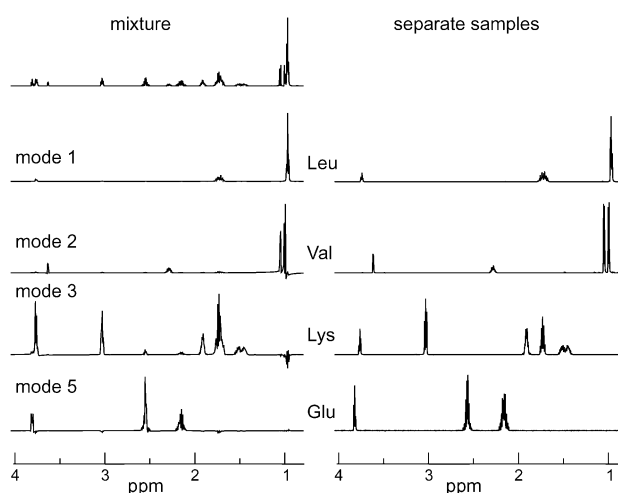


Figure 2. Demix method applied to an aqueous four amino acid mixture containing Glu, Leu, Lys, and Val analogous to Figure 1. Four of the five largest eigenmodes are shown in the left column corresponding to the individual amino acids. Eigenmode 4 (not shown), which is similar to eigenmode 3 in Figure 1, does not represent an entire amino acid.

The Demix method exploits the differential scalar J -coupling networks of the mixture's components along which magnetization is transferred during the TOCSY mixing period together with differential chemical shift evolution, which leads to a unique modulation of the resonance signals as a function of t_1 . The principal component analysis effectively captures independent patterns (modulations) in the covariance matrix and represents them in terms of distinct eigenmodes. The presence of similar J -coupling networks for two components is advantageous, since this requires a smaller number of mixing times to achieve uniform magnetization transfers among spins in the sum spectra. If a component has two (or more) separate decoupled spin networks, the 1D spectra of the different spin networks appear as separate modes. As is always the case when using diagonalization methods, in the event of accidental degeneracy (or near-degeneracy) any linear combination of degenerate eigenvectors is also an eigenvector. Because this may obscure the identity of the chemical components belonging to these eigenmodes, suitable rotation of the modes within the degenerate subspace might be useful in such cases. The purpose of the co-addition of spectra with different mixing times τ_m is the equalization of peak amplitudes to smooth the oscillatory τ_m dependence of the TOCSY-transfer amplitudes.

Conclusion

In summary, the method presented herein provides a straightforward deconvolution of the NMR spectrum of a mixture, and requires neither the physical separation of its components nor pulsed-field gradients or frequency-selective radio frequency pulses. The 1D eigenmodes can be efficiently screened against a database of 1D NMR spectra to identify the components of the mixture in a (semi-)automated manner. This method could find useful applications in the context of metabolomics/metabonomics.

Acknowledgments

This work was supported by NIH grant GM066041.

Keywords: covariance · mixtures · NMR spectroscopy · principal component analysis · TOCSY

- [1] D. A. Skoog, F. J. Holler, T. A. Nieman, *Principles of Instrumental Analysis*, 5th edition, Harcourt Brace and Company, **1998**.
- [2] *On-Line LC NMR and Related Techniques* (Ed.: K. Albert), Wiley, Chichester, UK, **2002**.
- [3] a) K. F. Morris, C. S. Johnson, Jr., *J. Am. Chem. Soc.* **1992**, 114, 3139–3141; b) C. S. Johnson, Jr., *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, 34, 203–256.
- [4] S. Viel, F. Ziarelli, S. Caldarelli, *Proc. Nat. Acad. Sci. USA* **2003**, 100, 9696–9698.
- [5] R. A. Johnson, D. W. Wichern, *Applied Multivariate Statistical Analysis*, 5th ed., Prentice Hall, New Jersey, **2002**.
- [6] R. Stoyanova, A. C. Kuesel, T. R. Brown, *J. Magn. Reson. Ser. A* **1995**, 115, 265–269.
- [7] T. R. Brown, R. Stoyanova, *J. Magn. Reson. Ser. B* **1996**, 112, 32–43.
- [8] R. Stoyanova, T. R. Brown, *J. Magn. Reson.* **2002**, 154, 163–175.
- [9] M. A. Elliott, G. A. Walter, A. Swift, K. Vandenborne, J. C. Schotland, J. S. Leigh, *Magn. Reson. Med.* **1999**, 41, 450–455.
- [10] H. Witjes, W. J. Melssen, H. J. A. I. Zandt, M. van der Graaf, A. Heerschap, L. M. C. Buydens, *J. Magn. Reson.* **2000**, 144, 35–44.
- [11] S. Van Huffel, Y. Wang, L. Vanhamme, P. Van Hecke, *J. Magn. Reson.* **2002**, 158, 1–14.
- [12] I. Duarte, A. Barros, P. S. Belton, R. Righelato, M. Spraul, E. Humpfer, A. M. Gil, *J. Agric. Food Chem.* **2002**, 50, 2475–2481.
- [13] R. R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon, Oxford, **1987**.
- [14] L. Braunschweiler, R. R. Ernst, *J. Magn. Reson.* **1983**, 53, 521–528.
- [15] R. Brüscheiler, F. Zhang, *J. Chem. Phys.* in press.
- [16] A. Bax, D. G. Davis, *J. Magn. Reson.* **1985**, 65, 355–360.
- [17] D. Marion, K. Wüthrich, *Biochem. Biophys. Res. Commun.* **1983**, 113, 967–974.

Received: November 27, 2003 [F 1073]