

High Resolution Diffusion-Ordered Spectroscopy (HR-DOSY)

Application Note

Authors

Péter Sándor
Agilent Technologies GmbH
D64289 Darmstadt
Germany

Abstract

This application note demonstrates the utility of the DOSY package in the VnmrJ 3.1 software for both 2D and 3D HR-DOSY analysis of complex mixtures, by first elucidating the presence of two components in an epoxy mixture (2D), and then proving that a minority component in a protein mixture was in fact a smaller derivative of the major component (3D).

Introduction

High Resolution Diffusion-Ordered Spectroscopy (HR-DOSY) is a very efficient tool for obtaining structural information on mixtures of small and medium size molecules. Unlike other physical methods used for mixture analysis, or NMR combined with liquid chromatography (HPLC-NMR) or mass spectrometry (MS-NMR), HR-DOSY does not require a separation technique. This is particularly important when separation is impossible (food samples, body liquids etc.), impractical, or just undesirable. Accurate determination of the diffusion coefficients may allow the separation of the NMR spectra of mixture components, or may serve as an experimental basis for predicting hydrodynamic radii and therefore molecular weights. Diffusion data are also often used to monitor solution-state aggregation or binding between ligands and macromolecules.

This application note demonstrates the utility of the VnmrJ 3.1 software of two different mixtures of complex molecules, one an epoxy resin mixture and the other a small protein. Excellent diffusion resolution was obtained between the epoxy monomer and octamer using 2D HR-DOSY, in spite of very intense signal overlap in the aromatic and methyl regions. 3D HR-DOSY analysis of the protein, which was suspected to contain a small amount of dimer, determined that the minority component in fact had slightly higher mobility and therefore lower molecular mass than the majority component.

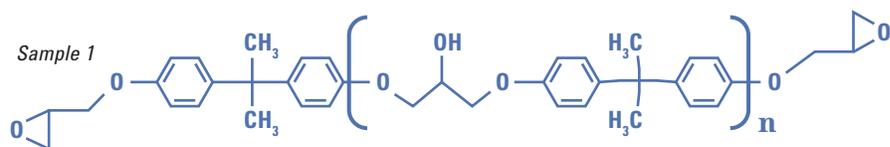


Agilent Technologies

Experimental

Samples

Sample 1, (kindly provided by Simon Welsh, International Paint, Felling, UK) is a mixture of two commercially available epoxy resins produced by The Dow Chemical Company. The two components are structurally very similar (D.E.R. 331, the “liquid resin”, with $n=0$ and EPIKOTE 1007 the “solid resin”, with $n\sim 7$) and show very intense signal overlap in the aromatic and methyl regions of the ^1H NMR spectrum, which is a particular challenge for diffusion spectroscopy. In the CDCl_3 solution the relative concentrations were chosen such that the intensities of the aromatic and methyl signals from the two species were comparable.



Sample 2 (courtesy of Nico van Nuland, University of Granada, Spain and Jean Jeneer NMR Center, VIB, Brussels, Belgium) is a 1 mM ^{15}N -labeled chimeric SH3-domain protein SPCp41¹ (75 amino acid residues) dissolved in 92% $\text{H}_2\text{O}/8\%$ D_2O , $\text{pH}=3.5$. Although the sample was found to be chromatographically pure, the ^{15}N - ^1H HMQC (Heteronuclear Multiple Quantum Coherence) spectra clearly identified two components in a ratio of 4:1, the main component being the expected and assigned original SH3 protein (effective concentration ~ 0.8 mM). 3D-DOSY-HMQC experiments were requested to confirm whether the minor species (effective concentration 0.2 mM) could be identified as a dimeric form of the main conformer.

Instruments and experimental conditions

Sample 1 was analyzed on an Agilent MR-400 NMR spectrometer operating at 399.86 MHz and equipped with a 5 mm OneNMR probe and a Performa I gradient amplifier (maximum gradient strength ~ 20 G/cm). Thirty-one values in the range of 2-20 G/cm were arrayed for the encoding/decoding gradient pairs of 3 milliseconds duration. To achieve proper signal attenuation, a relatively long (300 milliseconds) diffusion delay

had to be used. In chloroform solutions with low viscosity and long diffusion delays the probability of convection cannot be excluded, therefore a pulse field gradient double stimulated echo (Dpfgdste) pulse sequence was used that compensates for incidental thermal convections. The data were collected with active sample temperature regulation at 25° C.

Sample 2 was run on a VNMRs600 instrument operating at 599.715 MHz equipped with 3 RF channels, a room temperature 5 mm HCN probe as well as a Performa IV gradient amplifier (maximum gradient strength ~ 60 G/cm). Eight phase-sensitive, 2D gradient compensated stimulated echo ^{15}N - ^1H HMQC (Dgcstehmqc) experiments were run at 25° C with a diffusion delay of 100 msec, a gradient pulse width of 2 msec and with gradient amplitudes varied in the range of 5-55 G/cm. No gradients were used to select coherence pathways in the HMQC step. The spectral widths were set to 10 and 2.2 kHz in F2 and F1, and the acquired complex data points were 1024 and 128, respectively. Each free induction decay (FID) was an average of 64 transients. The FID size in F1 was doubled by linear prediction and zero filled to 1024 complex points prior to Fourier transformation. Matching cosine square window functions were applied both in F2 and F1.

Results and Discussion

Using DOSY to Determine Diffusion Coefficients

Diffusion pulse sequences are based on a gradient stimulated echo, or one of its numerous variants.

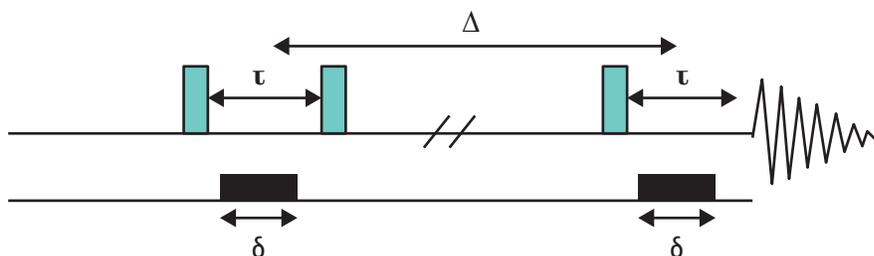
In DOSY experiments the strength of the diffusion-encoding gradient is arrayed, and decaying signal intensities (2D DOSY) or volume integrals (3D DOSY) are fitted to the Stejskal-Tanner expression:

$$S(G_{z1}) = S(0) \exp(-D_i \gamma^2 \delta^2 (G_{z1})^2 (\Delta - \delta/3))$$

where $S(G_{z1})$ and $S(0)$ are the signal intensities obtained with gradient strengths of G_{z1} and 0, respectively, D is the diffusion coefficient, γ is the gyromagnetic constant, δ is the gradient pulse duration and Δ is the diffusion delay.

The result of the 2D DOSY analysis is a 2D spectrum in which signals are dispersed according to chemical shifts in one dimension (F2) and diffusion constant in the other (F1). The cross-peak position is determined by the calculated diffusion coefficient (D), and its "line-width" in the diffusion dimension is determined by the standard deviation of the D values (Figure 1).

It is then not surprising that there is a very strong correlation between the data quality (accuracy and precision of the diffusion coefficients) and the "chromatographic" resolution of the DOSY experiments. Statistical errors (typically associated with insufficient signal-to-noise ratio) need to be avoided by proper parameter setting. Systematic errors, on the other hand, must be corrected wherever possible. When signal decays are not strictly exponential, even complex and powerful processing methods may have seriously limited applicability.



The physical basis for measuring diffusion by NMR

2D ^{31}P -DOSY of mixture of AMP and inorganic phosphate in D_2O diffusion delay = 300 msec

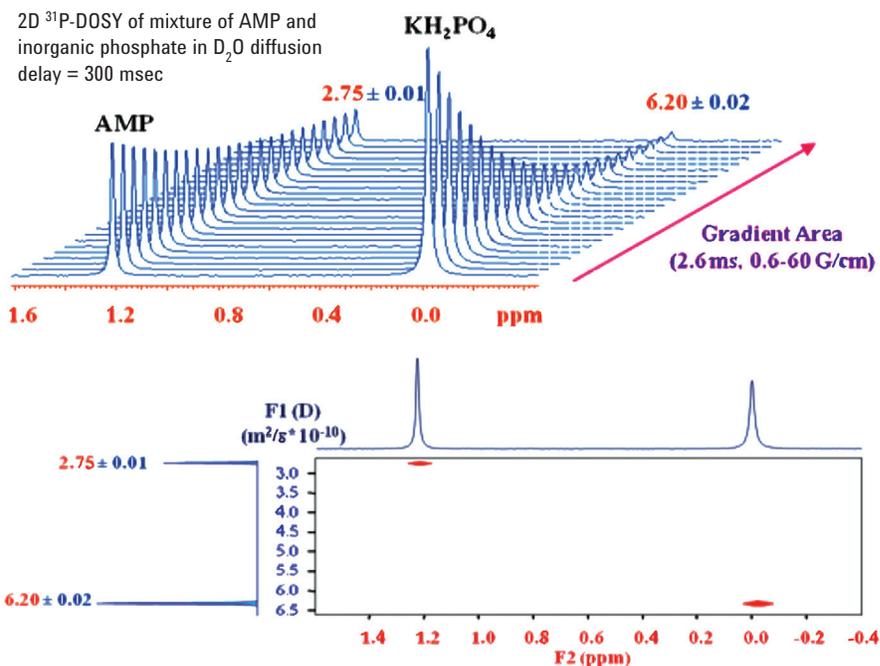


Figure 1. Top: signal attenuation as a function of increasing gradient area in a mixture of AMP and inorganic phosphate. The red numbers show the calculated D values with their standard deviation in blue. Bottom: The 2D display after DOSY processing.

Gradient calibration and correction for non-uniform gradients (NUG)

One of the possible sources of non-exponential signal attenuation is related to the gradient coil itself and is beyond operator control. Due to necessary compromises made in all probe designs, the field gradients produced are not perfectly linear along the gradient axis over the active volume of the sample. Fortunately, it is straightforward to correct for these effects of gradient non-uniformity by fitting the data to a modified Stejskal-Tanner equation that takes into account the actual gradient shape produced by the probe². A single experiment is used to determine the necessary correction to the Stejskal-Tanner equation that can then be used in processing all subsequent DOSY data. The calibration is done by the system administrator once per probe either on a standard doped D₂O sample, or on any sample with a known diffusion coefficient. The correction values are stored in the probe file and used automatically whenever new DOSY experiments are set up. This calibration not only produces the NUG correction factors, but the most accurate gradient amplitude as well. Both parameter sets are vital to obtain reliable and accurate diffusion data and to handle overlapping resonances correctly when multi-exponential fitting is used. The degree of gradient non-uniformity may vary with probe type and release date. NUG correction, however, may make high quality DOSY data available even with early versions of gradient probes.

2D DOSY analysis of an epoxy resin mixture

The experimental setup of the epoxy resin mixture (Sample 1) is summarized in Figure 2. The graphic displays the pulse sequence diagram. The input fields in the Pulse Sequence panel provide access to all diffusion-related parameters that need to be set or need to be optimized individually for every sample (including finding the best gradient area/diffusion delay combination).

When the data acquisition is complete, the FIDs are processed via the DOSY Process panel (Figure 3). This is not pulse sequence specific, as all 2D DOSY experiments share the same panel. The FID of every array element was zero filled to 16K data points, multiplied by an exponential decay function of 2 Hz, followed by Fourier transformation and baseline correction. The DOSY processing used NUG correction (that was preceded by a proper NUG calibration) with a triple exponential fitting and resulted in the 2D DOSY spectrum displayed in Figure 4.

The right-most block in the process panel allows the user to inspect the quality of data fitting on a peak-to-peak basis. The output is displayed in the graphics area (top of Figure 3) and shows the experimental and fitted signal attenuation as the function of the square of the gradient pulse area as well as the difference of the experimental and fitted data points. The latter is displayed on the bottom of the graphics window (multiplied by a factor of 100) and clearly shows that for the selected peak no systematic errors are present. If this inspection should indicate that one or more of the data points have significantly higher errors than the average, then the processing can be repeated by excluding the outliers. A mouse-click on the Show Diffusion Projection button displays the diffusion “chromatogram” (the vertical inset in Figure 4) that can be handled (expanded, plotted etc.) like an ordinary 1D NMR spectrum. The panel also allows the operator to toggle between the raw NMR data and the DOSY display without the necessity of reprocessing the data.

Easy-to-use acquisition setup for 2D DOSY

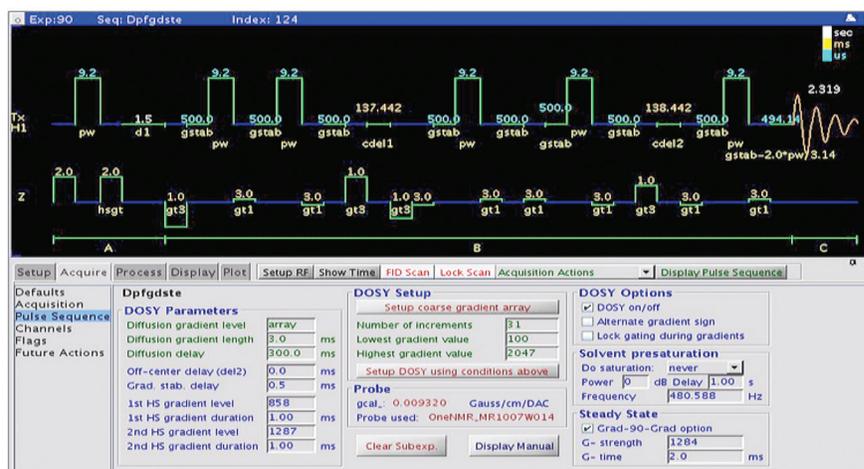


Figure 2. The pulse sequence schematics and the Pulse Sequence panel used for the epoxy mixture

Processing tools for 2D data analysis

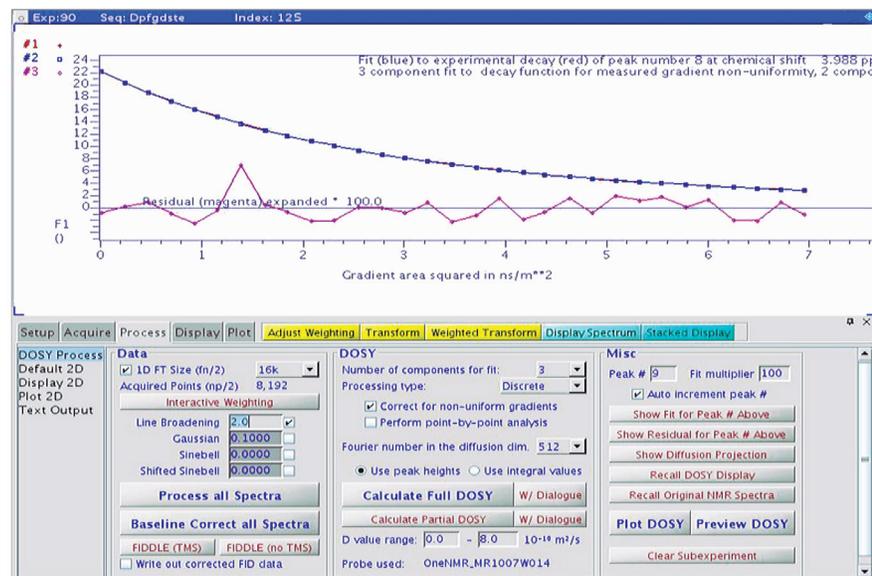


Figure 3. The 2D DOSY Process panel (bottom) and the results of the data fitting (top) of one of the aromatic signals. The top decaying curve shows the experimental signal amplitudes and the fitted curve. Below that, the differences between the experimental and calculated intensities are displayed after a multiplication of 100.

The 2D DOSY display in Figure 4 shows excellent diffusion resolution between the epoxy monomer and octamer. It is particularly remarkable that signals that are not resolved in the frequency domain (the aromatic, the methyl and the epoxy protons) exhibit the same diffusion property as those with no overlap (the -O-CH₂- region between 4 and 4.5 ppm). This data quality can only be achieved if the experimental parameters and conditions are chosen with special care and systematic errors are either avoided or properly corrected. The ultimate resolution power of diffusion experiments is a very complex issue³, where, among others, the concentration ratio of the mixture components, the ratio of the decay constants (D) and the signal-to-noise ratio of the NMR spectra play an important role.

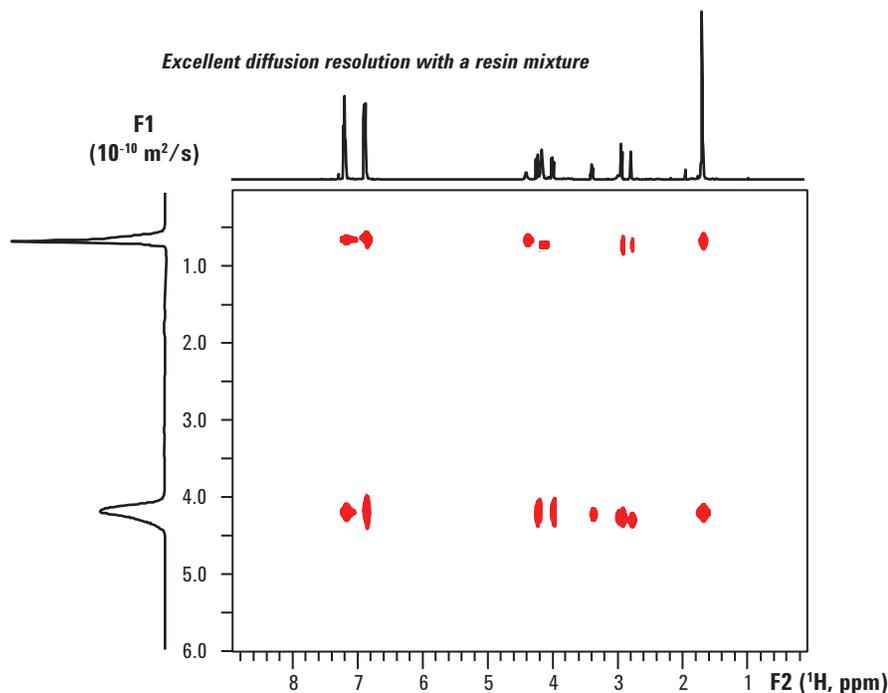


Figure 4. The 2D DOSY plot shows the results of the DOSY processing. On the top the high-resolution 1D spectrum; on the side the calculated "chromatogram" is displayed.

3D DOSY analysis of a protein mixture

Very complex mixtures and biological samples, with almost no exception, do not produce acceptable frequency resolution in 1D NMR spectra. The degree of signal overlap can then be eliminated or substantially reduced by combining the diffusion sequence with homo- or heteronuclear 2D correlation pulse sequences leading to 3D DOSY data sets.

The ^{15}N - ^1H HMQC spectra of protein SPCp41 (Sample 2) clearly indicated that it contains two molecules or two conformers (in a ratio of about 4:1) that could not be easily separated by chromatographic methods. Although most of the cross-peaks of the HMQC spectra did overlap, a few pairs of cross-peaks from the major and minor component showed sufficient frequency separation to make a diffusion study possible. Originally the major component was believed to be a monomer and the minor component a dimer, which is a hypothesis that diffusion measurements should be able to prove or negate.

Figure 5 displays the overall 2D ^{15}N - ^1H chemical shift correlation spectrum as well as four insets of cross-peak pairs that were separated enough to measure individual volume integrals reliably. In the insets, the numbers refer to the calculated diffusion coefficients of each cross-peak, together with the standard deviation of the D values in parenthesis. For all four cross-peak pairs, the minor component clearly shows slightly higher mobility and therefore slightly lower molecular mass than its more intense counterpart. Therefore the dimerization theory can clearly be excluded. Subsequent N-terminal sequencing revealed that the minor component is a derivative of the original SH3 protein with the first

Relative mobility of the major and minor components of a protein mixture

F1 (^{15}N , ppm)

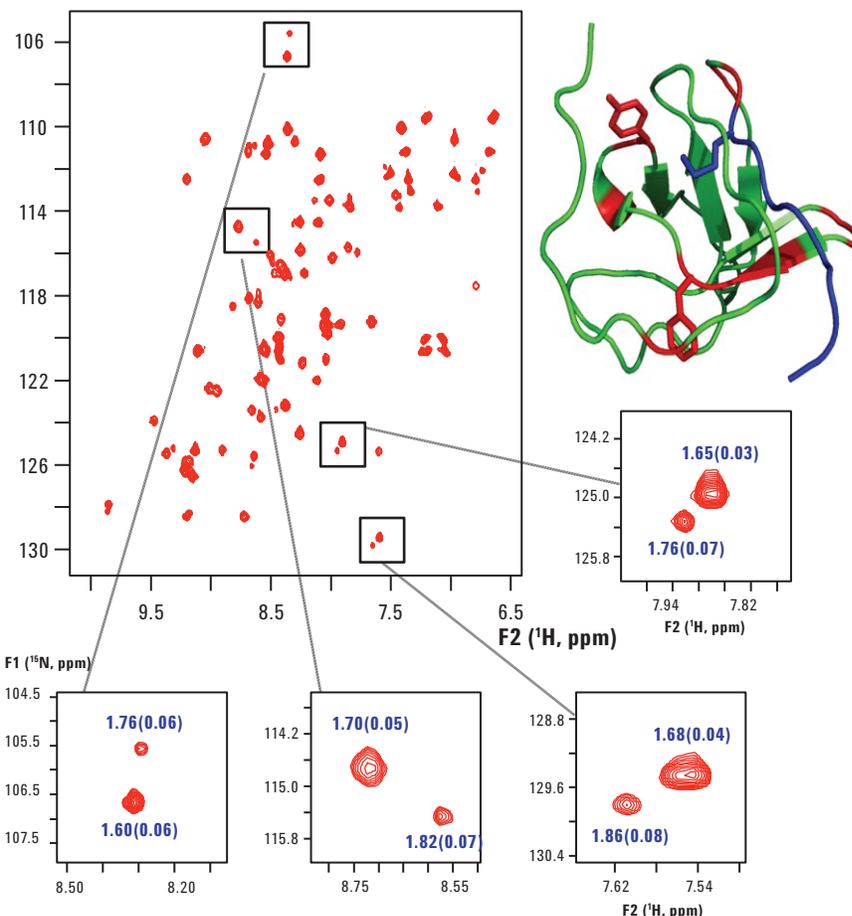


Figure 5. ^{15}N - ^1H HMQC spectrum of sample 2 with the lowest diffusion encoding/decoding gradient (top) together with four insets showing separated cross-peak pairs (bottom). The numbers in the insets display the calculated D values with their statistical error in parenthesis. The seven-amino acid sequence that is missing from the N-terminal in the minor component is indicated in blue in the structure drawing.

seven amino acids missing from the N-terminal. The protein structure in Figure 5 displays the missing sequence in blue and indicates the sites in red where the corresponding conformational changes result in separated cross-peaks in the HMQC spectra.

3D DOSY processing must always be preceded by cross-peak definition that can be done either automatically by the software, or manually under full

operator control. Cross-peaks can now be redefined, excluded, and combined via the Process panel and the interactive spectrum display, allowing the operator to include only the peaks of interest in the diffusion analysis (Figure 6). Once the peak definition is complete (often the most time consuming step of DOSY processing), the peak assignment can be saved in the FID and recalled anytime the data reprocessing is required.

Easily accessible processing tools for 3D DOSY

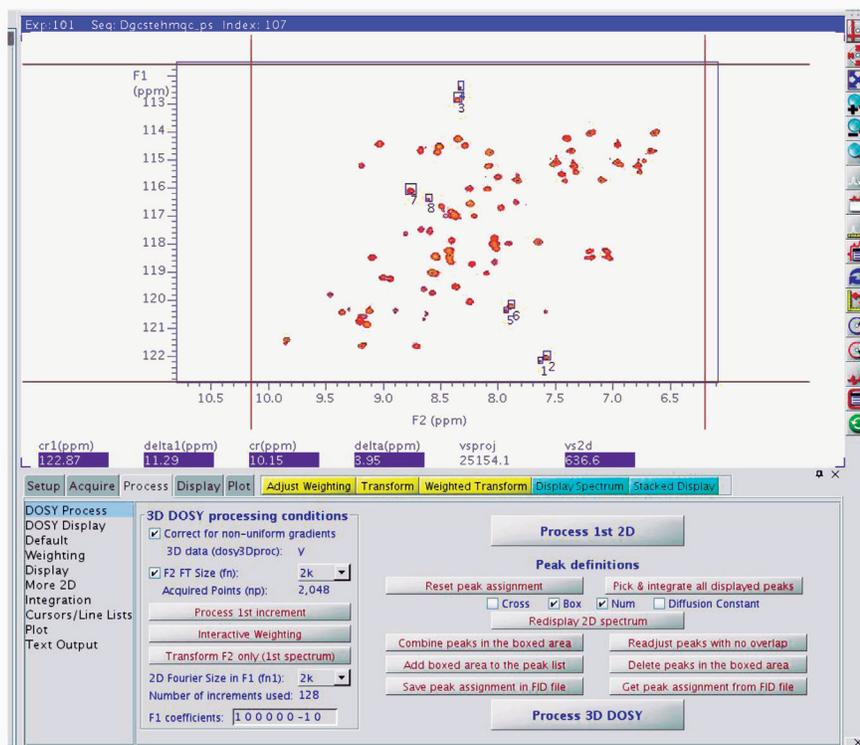


Figure 6. The DOSY Process panel and the interactive spectrum display used to analyze Sample 2.

Conclusions

The improved resolution embodied in the DOSY package in VnmrJ 3.1 provides higher accuracy and precision of diffusion coefficient measurements. Overlapping peaks on the chemical shift scale, which come from molecules with different diffusion coefficients, can now be separated using the multi-exponential fitting routines. The end benefit is more information content for analyses that were previously ambiguous or intractable. The package includes a total of 17 diffusion sequences, with some of them specifically designed for biological samples or allowing convection compensation. A separate DOSY manual (High Resolution Diffusion-Ordered Spectroscopy, Agilent Pub

No. 91001990) provides very detailed information covering all aspects of data acquisition, and an FID library helps inexperienced users to practice data processing.

The power of the new VnmrJ 3.1 DOSY package was demonstrated using HR-DOSY to analyze two complex mixtures successfully. The presence of two components in an epoxy mixture, a monomer and an octamer, was clearly shown using 2D DOSY, even though aromatic, methyl and epoxy protons signals were not resolved in the frequency domain. 3D DOSY analysis of a protein mixture gave convincing evidence that a minority component previously thought to be a dimer was in fact a smaller derivative of the major component, subsequently shown to be 7 amino acids shorter.

Acknowledgements

The DOSY package in VnmrJ 3.1 embodies major improvements over previous versions, primarily in the area of data processing. It was developed at and is licensed from the University of Manchester, UK.

References

1. A. M. Candell, F. Conjero-Lara, J. C. Martinez, N. van Nuland and M. Bruix. "The high-resolution NMR structure of a single-chain chimeric protein mimicking a SH3-peptide complex", *FEBS Letters* **581**, 687-692 (2007).
2. M. A. Connell, P. J. Bowyer, P. A. Bone, A. L. Davis, A. G. Swanson, M. Nilsson, and G. A. Morris, "Improving the accuracy of pulsed field gradient NMR diffusion experiments: Correction for gradient non-uniformity." *J. Magn. Reson.* **198**, 121-131 (2009).
3. M. Nilsson, M.A. Connell, A. Davis and G.A. Morris. "Biexponential Fitting of Diffusion-Ordered NMR Data: Practicalities and Limitations." *Anal. Chem.* **78**, 3040-3045 (2006).

For more information

Learn more:

www.agilent.com

Buy online:

www.agilent.com/chem/store

Find an Agilent customer center in your country:

www.agilent.com/chem/contactus

U.S. and Canada

1-800-227-9770

agilent_inquiries@agilent.com

Europe

info_agilent@agilent.com

Asia Pacific

adinquiry_aplsc@agilent.com

www.agilent.com

Product specifications and descriptions in this document are subject to change without notice.

© Agilent Technologies, Inc., 2011
Published in USA, April 8, 2011
Publication Number 5990-7600EN



Agilent Technologies