Easy, Robust and Accurate NMR Analysis of Biomolecules using BioPack

Technical Overview

Author

George Gray
Agilent Technologies
Santa Rosa, CA USA

Introduction

In the last fifteen years, scientists have progressed from working with smaller unlabeled proteins and homonuclear NMR experiments to a full suite of heteronuclear triple resonance experiments in order to fully explore protein structure. The size limit of NMR applicability has been expanded via TROSY-based experiments and deuterium labeling. At the same time, the process of backbone and side chain assignments of proteins and nucleotides has become more routine. These factors, together with the additional structural constraints through scalar and residual dipolar couplings, are allowing studies of proteins of unprecedented size. The major goal of protein NMR spectroscopy therefore has shifted from pulse sequence development to answering exciting biochemical questions. NMR hardware and software must now provide a full breadth of experimental approaches to enable the answers to these questions. The full power of NMR biomolecular structure studies must now be available to both seasoned and new NMR practitioners. Ease of use, reproducibility and accurate results are therefore a necessity. This Technical Overview demonstrates the power of Agilent BioPack software to meet these needs, providing a sophisticated turn-key approach in which the operator needs only select from over 350 sophisticated NMR pulse sequences, specify experiment time, and then generate high-quality and reproducible data.

Agilent Technologies

The Power of BioPack

BioPack enables you to get the biomolecular information that you need, whether or not you are an experienced NMR operator. A full suite of sophisticated pulse sequences (>350) is available to study proteins, RNA, or DNA complexes. There are experiments for backbone characterization, secondary and tertiary structure investigation, residual dipolar coupling measurement, relaxation time determination, and ligand binding. Most importantly, pulse sequences generated on one Agilent NMR instrument can be used reliably on any other Agilent instrument, without any adjustments.

The unique macro language used in VnmrJ software tailors the experiment setup to quickly provide the needed parameters for each experiment and start generating data. The operator simply chooses one of the many sophisticated experiments, selects the number of transients depending on sample concentration and numbers of increments for the desired resolution, and then starts the experiment. Both experienced operators and new users can start generating useful data right away.

Calibration is fully automatic, saving hours of manual experimentation optimizing selective pulses and water suppression parameters. All experiment pulse powers and pulse widths are automatically set (Figure 1). Shaped pulses, decoupling and spinlock waveforms are auto-generated based on the determined pw90's. The use of uniform names for parameters across all pulse sequences simplifies use, especially for new users. Gradients are automatically calibrated for coherence transfer (NH and CH), and Bloch-Siegert phase corrections are automatically determined. BioPack is also continually updated via the Agilent site on a regular and frequent basis, adding new experiments and new processing capabilities.

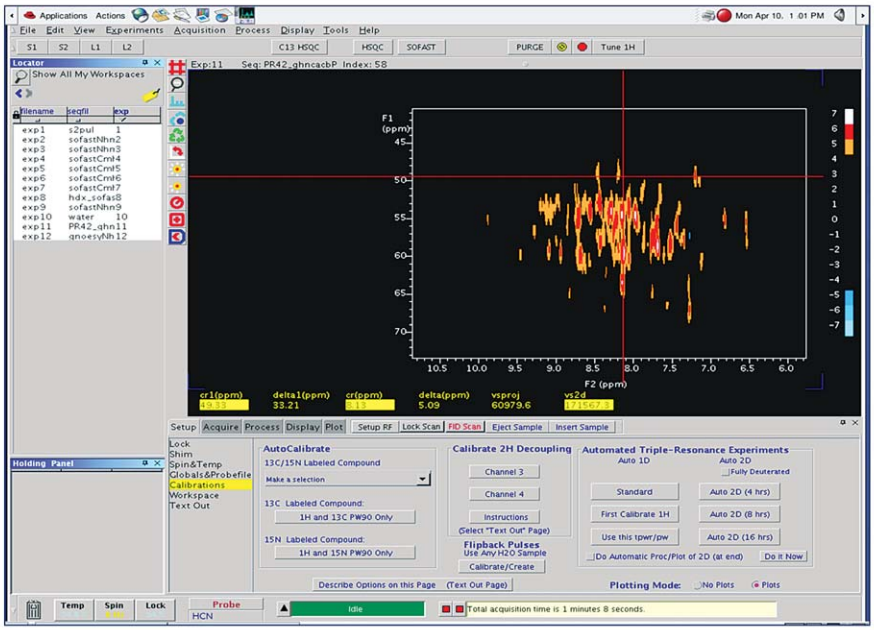

Figure 1. This page provides tools to accomplish automatic calibration for labeled and unlabeled proteins. This includes calibration of ^{2}H decoupling parameters and water flipback pulses. The Auto1D selections generate a series of 1D first increments that give an idea of relative sensitivity for a variety of pulse sequences. The Auto2D selections run ~37 2D experiments using triple-resonance pulse sequences to assess the suitability of these experiments for 3D.

Over 350 distinct NMR pulse sequences to choose from

BioPack contains pulse sequences for all major biomolecular applications, many of which have a multitude of options for further refinement of experimental requirements. Whether your focus is protein backbone assignment, determination of distances or dynamics of protein folding or the study of polynucleotides, BioPack has a pulse-sequence experiment to fit your needs, a few examples of which are shown in Table 1.

For homonuclear experiments, multi-frequency suppression can be achieved using Pbox. This is possible for the presat, wet and watergate experiments, as well as NOESY, TOCSY and ROESY (Table 1). It will saturate lines from a line list (including H_2O) during the relaxation delay. The PURGE option is also available for superior water suppression over standard presat[1]. This sequence is basically a modified presat with short spin-echos and gradients.

Table 1. Summary of selected pulse sequence experiments in BioPack

Experiment Type	Selected Examples*
Homonuclear 1D and 2D	PRESAT, watergate, WET, Multi-Frequency Presat NOESY, TOCSY, ROESY
Heteronuclear N15 Indirect Detection 2D	CLEANEX N15-HSQC, N15-HSQC with Homodecoupling, watergate N15-HSQC, N15-ultra fast HMQC
3D and 4D N15 HSQC and/or HMQC	N15-NOESYHSQC, 3D N15, C13-NOESYHSQC C13-HMQCNOESY-N15-HSQC(4D), 13C, 15N edited NOESY with TROSY(4D)
13C HSQC and/or HMQC indirect Detection	C13-HMQCNOESYHSQC(4D) C13-TOCSYHSQC C13-HMQC 13C,15N edited NOESY with TROSY
Triple-Resonance Assignment	C(CO)NH (or C(CC-TOCSY-CO)N-NH) HCACOCANH, HNCA (intra-residue-only), HADAMAC
Projection Reconstruction	H(CCO)NH (or H(CC-TOCSY-CO) N-NH), CBCANNH, HN(CO)CA, C13-NOESY-HSQC(SE)

Experiment Type	Selected Examples*
Triple-Resonance TROSY for Fully-Deuterated Proteins	HNCO 3D, HNCOCACB 3D, HN(CO)CA 3D (sequential), HN(CA)CO 4D
Non-Linear Sampling	All nD experiments
Triple-Resonance	Aromatics: aromatic proton-beta carbon correlation, Relaxation:HNCO_NOE Side-Chain Assignments: CACB-TOCSY-C(methyl)H(methyl)-SQ, etc.
Coupling Measurement	LR-JCH, N15-HMQCJ, S3 for J(N-CO)/(HN-CO) doublets, S3 2D J(CoCa) in a 1H-15N correlation
1D and 2D for Polynucleotides	PRESAT, WET, jump-return, Watergate, WET-NOESY, 13C TROSY, 15N HSQC(long-range)
3D Experiments for Polynucleotides	13C NOESY-HSQC, CPMG-NOESY, DE-H(C)CH-TOCSY, A-HNC-TOCSY-CH
Projection Reconstruction for Small Molecules	N15-NOESY-HSQC, C13-NOESY-HSQC(SE), N15-TOCSY-HSQC, C13-TOCSY-HSQC(SE)

*These examples represent only a small fraction of the pulse sequences available for each experiment type.

Triple resonance experiments

While the 1D and 2D water suppression experiments are widely used, triple resonance experiments are required for the sequential assignment of larger proteins (> 150 AA). They are called 'triple resonance' because three different RF channels (typically ^{13}C, ^{15}N and ^{1}H) are used. The experiments are performed on doubly labeled (^{13}C, ^{15}N) proteins. The most important advantage of the triple resonance spectra is their simplicity, because the spectra are simpler and the signals are spread over three dimensions. The problem of spectral overlap is therefore markedly reduced (this is the main reason that proteins of more than 20 kDa can be assigned with triple resonance experiments). Another advantage of triple resonance spectra is their high sensitivity, which is due to an efficient transfer of magnetization.

Projection reconstruction experiments

These experiments are automated in BioPack, significantly increasing the speed of acquisition of multi-dimensional NMR spectra (Table 1, Figure 2). The information from a small number of plane projections is used to extract chemical shift assignments. Projections at any desired angle of incidence are obtained by Fourier transformation of time-domain signals acquired when two or more evolution intervals are incremented simultaneously at different rates. A dedicated "experiment manager" in BioPack enables entry of acquisition parameters (number of scans, numbers of increments, etc.) for a series of different pulse sequences.

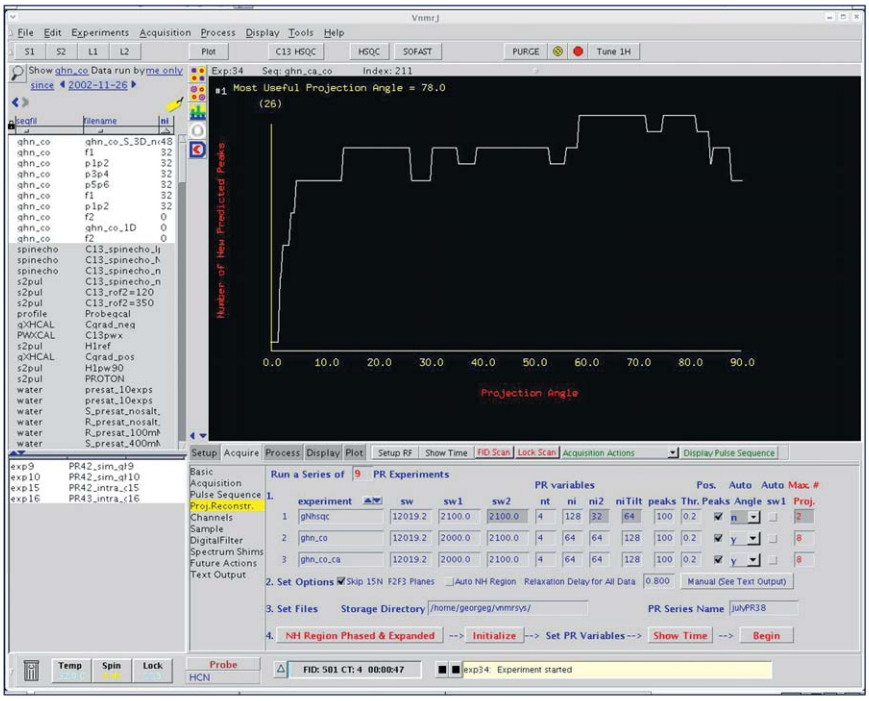

Figure 2. This automatic Projection Reconstruction manager permits easy setup of a series of user-selected triple-resonance experiments with automatic projection angle determination. This gives the most efficient data collection while still achieving chemical shift assignment.

Ultra-fast multi-dimensional NMR

One of the exciting features in BioPack is the capability to use ultra-fast methods based on the work of the Brutscher group in Grenoble [2, 3]. These methods of indirect detection utilize recycle delays as short as a few milliseconds. The key to the NH-detected experiments is band selection of the NH region using shaped excitation. All pulses on protons are designed to only affect the NH protons and not affect the state of magnetization of aliphatic or water protons. This preserves the +Z magnetization of the other protein protons, which serves as a reservoir of magnetization to rapidly relax the NH protons back to +Z during the acquisition time. The relaxation delay can then be set as short as 1 msec (or essentially zero, since the acquisition time serves as the relaxation delay).

The maximum acquisition time in any indirect dimension can be set for the desired resolution, providing high quality data with good resolution. The dramatic reduction of the relaxation delay from ~1 sec to ~1 msec can reduce the time to run an HSQC experiment from 2-3 minutes to only a few seconds. The main application for these experiments is the study of fast processes such as exchange and dynamics. Of course, they are also of great use in the case of short sample lifetime.

The pulse sequences are designed to preserve non-NH magnetization along +Z by using region-selective pulses. Each of these experiments has the same or better sensitivity performance as the standard BioPack experiments for normal (1 second) relaxation delays. They are more sensitive by a factor of 2–4 for short relaxation delays (for example, 100 msec).

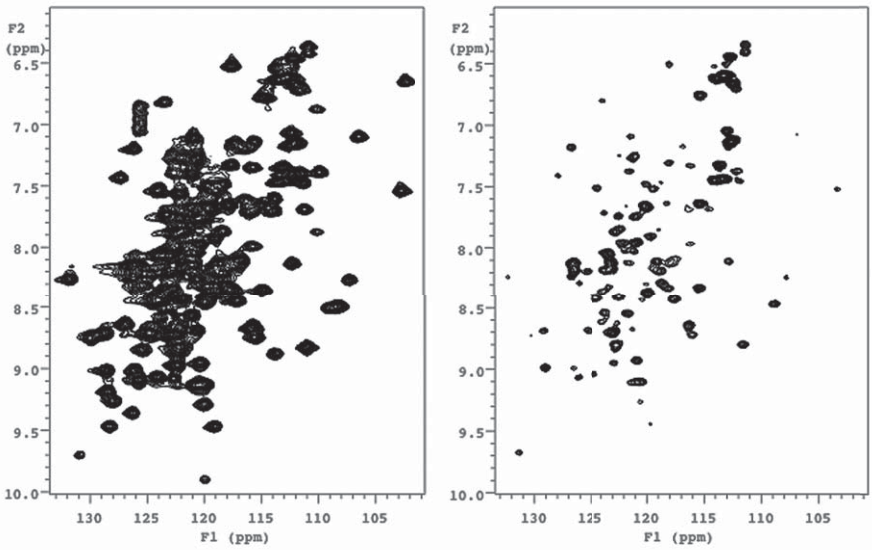

Figure 3 ^{15}N SOFAST-HMQC (left); Normal HSQC (right). Total acquisition time for each is 9 sec.

The impact of preservation of Z magnetization is demonstrated in Figure 3, which shows comparative data for a recycle time of 10 msec and total time of 9 seconds for a complete 2-D SOFAST-HMQC. The data from a normal HSQC done in the same time (Figure 3, right) has considerably lower sensitivity than the ^{15}N SOFAST data (Figure 3, left).

Fast and Ultra-Fast experiments are selected in BioPack using the drop-down Experiments menu in VnmrJ. The BioPack VnmrJ Calibrations page Setup folder includes an Automatic BEST (band-selective excitation short-transient) 2D button that will run a series of automatic 2-D experiments using triple resonance assignment experiments with automatic processing and plotting. Typical results are shown in Figure 4.

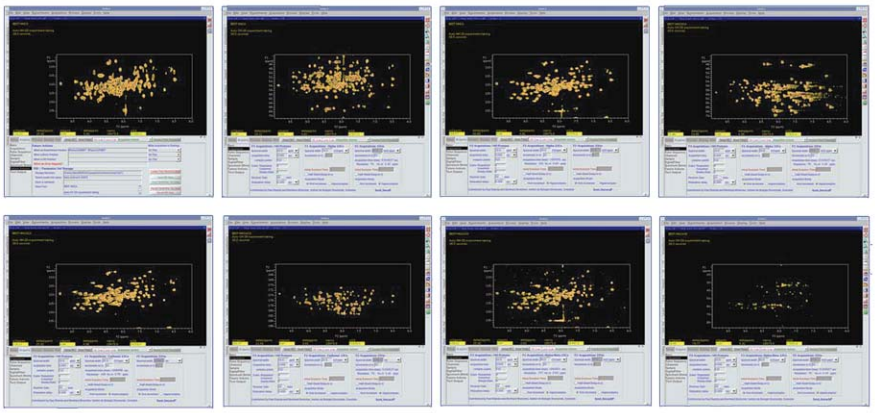

Figure 4. An example of BEST 2D protein experiments automated in BioPack, in which 13 experiments were completed in only 13 minutes (only 8 are shown).

Non-Linear Sampling in Multi-Dimensional NMR

Another of the leading-edge features in BioPack is the capability to do non-linear sampling (NLS) to significantly increase the speed of multi-dimensional NMR experiments. Normal multi-dimensional data sets are collected using regular (linear) incremental increase of (an) evolution time(s) because the "Fast Fourier Transform" algorithm is used for processing and this method requires regular sampling. A typical 3D data set may have several thousand fids acquired, because the total number of fids is proportional to the product of the number of evolution times used for each indirect dimension. However, other processing methods have recently become available (Multi-dimensional Fourier Transform, Maximum Entropy, Multi-Dimensional Decomposition, etc.) that are able to process non-linearly sampled data where the number of evolution times for each indirect dimension is fewer than that required for use of the Fast Fourier Transform. This non-linear approach can represent a significant timesaving in acquisition if only 5-25 % of the data needs to be sampled, and yet the data can be accurately processed.

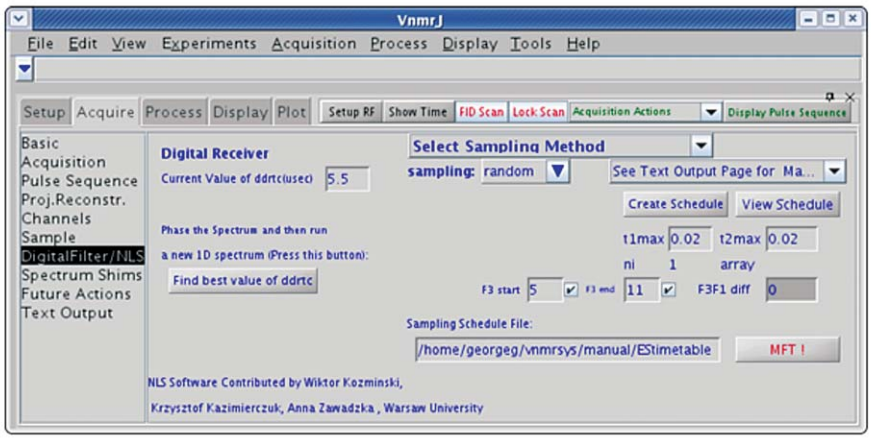

Figure 5. The Digital Filter/NLS menu page for the Kozminski Method.

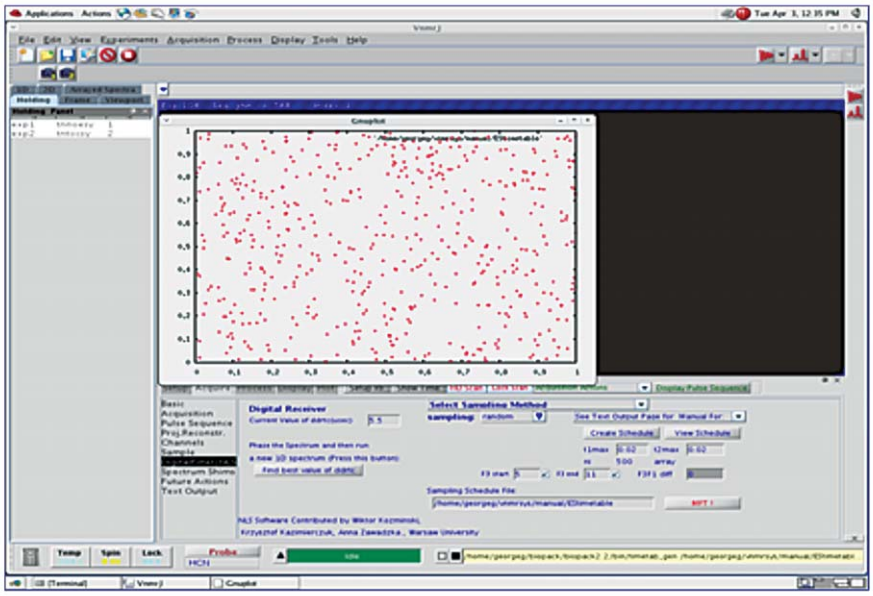

Figure 6. Sampling Schedule display via the "View Schedule" button.

The flexibility of the pulse programming language of VnmrJ software has allowed straightforward and generic implementations of acquisition of NLS data in BioPack. In particular, the evolution time(s) for a specific fid can be looked up from a text file, rather than being automatically calculated. All that is necessary is for the sequence to be modified, replacing the internally calculated evolution time(s) with ones from a text file. This approach is general and applies to any multi-dimensional NMR experiment.

The "Select Sampling Method" menu has three NLS options: Non-Linear (Sparse) Sampling (Kupce method), Non-Linear (Sparse) Sampling (Orekhov method), and Non-Linear (Explicit) Sampling (Kozminski method). Each of these three sequences is accessed via the Sampling page in the Acquire folder which has a menu to facilitate setting up an NLS experiment. An example of the menu for the Kozminski method is shown in Figure 5. The button "Create Schedule" creates a new sampling schedule using the specified "ni" and sampling method and stores it under the name and location specified by the Sampling Schedule File entry box. The button "View Schedule" displays a plot of time domain data points (Figure 6, for the case of random sampling).

NLS data may be processed within VnmrJ using software contributed by Eriks Kupce (CLEAN) or Wiktor Kozminski (MFT).

The NLS approach has been demonstrated to give up to a 20-fold increase in throughput by reducing the total time of the NMR experiment without any loss of spectral information. The ease of implementation for any new pulse sequence within VnmrJ and BioPack makes this powerful tool accessible to a wide variety of NMR users.

Conclusion

BioPack has a powerful suite of experiments that makes protein and polynucleotide NMR spectroscopy easy and reliable, while still delivering optimal and accurate results. This is not just a series of "canned" experiments, but a way of producing expert level results while greatly reducing time-costly mistakes and the need for local pulse sequence expertise. BioPack allows NMR scientists to focus more of their time on sample preparation and optimization, as well as analysis of the acquired data. While being automated enough to be used by a new user safely and productively, BioPack has many layers that allow an inspired user to fully understand the sequence of interest, the options presented, and the calibration processes.

References

1. A. Simpson and S. Brown, "Purge NMR: Effective and easy solvent suppression", J Magn Reson 75, 340-346 (2005).
2. P. Schanda and B. Brutscher, "Very Fast Two-Dimensional NMR Spectroscopy for Real-Time Investigation of Dynamic Events in Proteins on the Time Scale of Seconds", J Am Chem Soc 127, 8014-8015 (2005).
3. E. Lescop, P. Schanda, B. Brutscher, "A set of BEST triple resonance experiments for time-optimized protein resonance assignment." J Magn Reson 187, 163-169 (2007).

www.agilent.com/chem/nmr

Published in the U.S.A., September 12, 2011
5990-9067EN

Agilent Technologies