Journal of Magnetic Resonance 346 (2023) 107326

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

Less is more: A simple methyl-TROSY based pulse scheme offers improved sensitivity in applications to high molecular weight complexes

Nicolas Bolik-Coulon [a,b,c], Alexander I.M. Sever [b], Robert W. Harkness [a,b,c], James M. Aramini [a], Yuki Toyama [a,b,c], D. Flemming Hansen [d,*], Lewis E. Kay [a,b,c,e,*]

[a] Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada
[b] Department of Chemistry, University of Toronto, Toronto, ON M5S 3H6, Canada
[c] Department of Biochemistry, University of Toronto, Toronto, ON M5S 1A8, Canada
[d] Department of Structural and Molecular Biology, Division of Biosciences, University College London, London WC1E 6BT, United Kingdom
[e] Program in Molecular Medicine, Hospital for Sick Children Research Institute, Toronto, ON M5G 0A4, Canada

ARTICLE INFO

Article history:
Received 21 October 2022
Revised 3 November 2022
Accepted 4 November 2022
Available online 9 November 2022

Keywords:
HMQC
Delayed decoupling
Sensitivity gains
MDa protein complexes

ABSTRACT

The HMQC pulse sequence and variants thereof have been exploited in studies of high molecular weight protein complexes, taking advantage of the fact that fast and slow relaxing magnetization components are sequestered along two distinct magnetization transfer pathways. Despite the simplicity of the HMQC scheme an even shorter version can be designed, based on elimination of the terminal refocusing period, as a further means of increasing signal. Here we present such an experiment, and show that significant sensitivity gains, in some cases by factors of two or more, are realized in studies of proteins varying in molecular masses from 100 kDa to 1 MDa.

1. Introduction

Despite the tremendous progress in solution NMR spectroscopy over the past several decades [1–2], which has been fueled by advances in hardware, pulse sequences, processing methods, and sample preparation, sensitivity remains a central issue. This is particularly the case for applications to high molecular weight biomolecules where rapid relaxation of signals and spectral overlap are two formidable challenges. The development of TROSY-based approaches, first focusing on backbone amide ^{15}N-$^{1}H^{N}$ spin pairs [3], and subsequently on methyl groups [4], has extended the range of biomolecules that can be studied, offering new insights into the role of structure and dynamics in biomolecular function and misfunction [1,5]. The Kay laboratory has had a longstanding interest in the functional dynamics of large complexes such as oligomeric proteases and unfoldases that regulate protein homeostasis in the cell, and in the nucleosome core particle that is the fundamental building block of chromatin. Complexes on the order of 1 MDa have been studied [6] and the often high quality spectra generated allows detailed and quantitative analyses that were normally only possible in NMR studies of much smaller protein systems. However, even when reasonable spectra are obtained, there is always a distribution of peak intensities, with some correlations much weaker than others. It would be advantageous in these cases, and in applications to cellular complexes that are on the order of or larger than 1 MDa, especially, to develop approaches that increase spectral sensitivity even further.

The HMQC pulse scheme, a simple, yet powerful, experiment for correlating one bond coupled X-^{1}H spins [7,8] has found a prominent role in studies of large protein complexes when applied to methyl groups, as the transfer of magnetization involves a pair of pathways that comprise either fast or slow relaxing components [4]. In particular, labeling with $^{13}CH_3$ methyls in an otherwise deuterated protein [9] or nucleic acid [10] background minimizes mixing of the two pathways, with the slowly relaxing pathway dominating in applications to high molecular weight systems. We wondered if it would be possible to improve the sensitivity of the HMQC experiment still further by simply shortening it. In principle, this can be achieved by eliminating the final element that is responsible for refocusing anti-phase magnetization into in-phase ^{1}H magnetization prior to direct detection with X-nucleus decoupling. Such an approach had been introduced by Bax and Sarkar in a number of ^{13}C-detect heteronuclear correlation experiments, where the goal was to minimize the number of pulses

* Corresponding authors.
E-mail addresses: d.hansen@ucl.ac.uk (D.F. Hansen), kay@pound.med.utoronto.ca (L.E. Kay).

https://doi.org/10.1016/j.jmr.2022.107326
1090-7807/

rather than delays [11], as addition of pulses could result in sensitivity losses. In the present set of applications, involving studies of very high molecular weight systems, removal of refocusing elements can lead to significant enhancements in signal-to-noise, as illustrated below. A disadvantage, discussed in the original paper [11], is that decoupling during acquisition must be delayed until refocusing is complete, leading to spectra with distorted lineshapes that are not suitable for conventional quantitative analysis and necessitating post-acquisition manipulation to generate spectra with improved (*i.e.*, Lorentzian) peak shapes. Motivated by work from the Bax laboratory and by a recent paper introducing XL-ALSOFAST-HMQC for recording non-TROSY based HMQC spectra in fully protonated proteins that used a shortened refocusing element in combination with delayed decoupling [12] we asked whether suitable protein spectra could also be obtained by removing the element altogether and what the best approach might be for eliminating the distortions in the resulting lineshapes. Here we have examined spectra of a number of different high molecular weight protein systems that have been processed using a simple *J*-deconvolution procedure to remove peak distortions. These include the HtrA2 protease (equilibrium between 105 kDa trimer and a 210 kDa hexamer) [13], the nucleosome core particle (NCP, 210 kDa) [14], the half-proteasome ($\alpha_7\alpha_7$, 360 kDa) [15], and lumazine synthase (AaLS, 1 MDa) [16]; in some cases spectra have been recorded at temperatures as low as 7 °C to increase global tumbling times. Significant sensitivity gains are achieved for weak signals.

2. Materials and methods

2.1. Protein expression and purification

Protein L, HtrA2, NCP, and $\alpha_7\alpha_7$ samples were generated as described previously [10,13,15,17]. Sample conditions were: 1.4 mM protein L, 99.9 % D_2O, 50 mM sodium phosphate, pH 6.0 (uncorrected); 0.30 mM HtrA2 (protomer concentration), 20 mM HEPES-NaOH (pD 7.4), 1 mM EDTA, 99.9 % D_2O; 0.15 mM NCP, 20 mM sodium phosphate, pD 6, 0.05 % NaN_3, 99.9 % D_2O; 1 mM $\alpha_7\alpha_7$ (protomer concentration), 25 mM potassium phosphate, pH 6.8 (uncorrected), 50 mM NaCl, 1 mM EDTA, 0.03 % NaN_3 and 2 mM DTT, 99.9 % D_2O.

In order to produce a sample of lumazine synthase a gBlock™ gene fragment containing the full-length (residue 1-154) *Aquifex aeolicus* lumazine synthase (AaLS) gene was synthesized by IDT (codon-optimized). It was placed into a *Champion* pET vector (kanamycin resistance) without *N*-terminal or *C*-terminal affinity purification fusion tags using Gibson Assembly® (New England BioLabs Inc.). The coding portion of the plasmid was confirmed via DNA sequencing. The AaLS plasmid was transformed into *E. coli* BL21-CodonPlus(DE3)-RIPL competent cells via heat-shock. Cells were grown in D_2O M9 minimal media, using [^{2}H,^{12}C]-glucose as the sole carbon source. The sample was labeled as U-^{2}H, Ile-δ1-$^{13}CH_3$, Leu,Val-$^{13}CH_3$ (only one of the two methyls of Leu,Val was NMR labeled) and Met-ε-$^{13}CH_3$, as described previously [9,18], achieved with the addition of precursors (60 mg.L^{-1} α-ketobutyric acid, 100 mg.L^{-1} α-ketoisovaleric acid, and 100 mg.L^{-1} ε-$^{13}CH_3$ methionine) 1 h prior to induction of protein expression. Cells were grown at 37 °C until an OD_{600} of $\sim$ 0.8 was achieved, at which point protein synthesis was induced by adding 0.2 mM IPTG. The cells were then allowed to grow for a further $\sim$ 20 h at 25 °C. Subsequently, the cells were harvested via centrifugation and resuspended in lysis buffer (20 mM Tris•HCl, pH 8.0, + 1 mg.ml^{-1} lysozyme).

Purification of AaLS consisted of first lysing the cells by ultrasonication, followed by centrifugation at 9000 rpm for 15 min. The supernatant was filtered (0.45 μm pore size) and loaded onto a HiTrap® Q HP column equilibrated with 20 mM Tris•HCl (pH 8.0). AaLS was eluted using a 0-100% gradient over 14 column volumes of 20 mM Tris•HCl, 1 M NaCl (pH 8.0) [16]. Fractions which contained AaLS were collected and $(NH_4)_2SO_4$ was added to a concentration of 1 M. The solution was loaded onto a HiTrap® Butyl HP column equilibrated with 25 mM NaH_2PO_4, 1 M $(NH_4)_2SO_4$, 1 mM EDTA (pH 7.0). AaLS was eluted using a buffer of 25 mM NaH_2PO_4, 500 mM $(NH_4)_2SO_4$, 1 mM EDTA (pH 7.0). Fractions containing AaLS were collected and diluted 2-fold with 25 mM NaH_2PO_4, 200 mM NaCl, 1 mM EDTA (pH 7.8). The solution was then heated to 80 °C for 20 min to induce unfolding and aggregation of protein impurities, and subsequently centrifuged at 9000 rpm for 15 min to clear these aggregates [16]. The resulting supernatant was concentrated, and final purification of AaLS was achieved using a Superose 6 10/300 size-exclusion column equilibrated with 25 mM NaH_2PO_4, 200 mM NaCl, 1 mM EDTA (pH 7.8). Fractions containing AaLS were buffer exchanged into 99.9 % D_2O, 25 mM NaH_2PO_4, 100 mM NaCl, 1 mM EDTA, pD 7.5 using an Amicon Ultra-15 concentrator (molecular weight cutoff = 30 kDa) to produce a 0.2 mM NMR sample (protomer concentration). The concentration of AaLS was measured using a NanoDrop spectrophotometer and calculated from the extinction coefficient at 280 nm (13,980 M^{-1}.cm^{-1}) established from ExPASy's ProtParam tool (https://web.expasy.org/protparam/).

2.2. NMR spectroscopy

All NMR spectra were recorded on a 1 GHz Bruker Avance NEO spectrometer equipped with a cryogenically cooled *x,y,z*, pulsed field gradient triple-resonance TCI probe. HMQC spectra were recorded with the number of scans, acquisition times in t_1 and experimental times listed below. The delay τ for magnetization transfer (Fig. 1) was set to 1.8 ms unless otherwise specified and a recycle delay of 1.5 s was used in all applications. The following are acquisition details for each spectrum: HtrA2, 24 scans (25 °C, t_{1max} = 21.9 ms, experimental time is 2.4 hrs) and 8 scans (40 °C, t_{1max} = 21.9 ms, experimental time is 49 min); NCP, 32 scans (25 °C, t_{1max} = 19.9 ms, experimental time is 2.9 hrs and 7 °C, t_{1max} = 17.9 ms, experimental time is 2.6 hrs); $\alpha_7\alpha_7$, 32 scans (7 °C, t_{1max} = 17.9 ms, experimental time is 2.6 hrs) and 8 scans (40 °C, t_{1max} = 21.9 ms, experimental time is 49 min); AaLS, 32 scans (40 °C, t_{1max} = 17.3 ms, τ = 1.6 ms, experimental time is 2.5 hrs) and 40 scans (25 °C, t_{1max} = 14.9 ms, τ = 1.5 ms, experimental time is 2.7 hrs). Data were processed using the *nmrPipe* suite of programs [19], with peak intensities obtained using *Peakipy* (https://github.com/j-brady/peakipy). ^{1}H $R_{2,S}$ values were recorded as described previously [20].

For optimal performance, samples should be dissolved in D_2O. This has the benefit that suppression of residual water is, of course, easier than in applications involving H_2O-based buffers. However, there is an additional important benefit. As the goal is to apply the method to MDa complexes sensitivity becomes critical. Applications in H_2O will suffer from exchange transfer of dephased water magnetization to the protein and concomitant decrease in methyl magnetization from a saturation transfer effect.

3. Results and discussion

3.1. Lineshape of Fourier transformed delayed decoupled FID

Fig. 1a shows the pulse scheme that has been used to record delayed decoupled HMQC spectra (ddHMQC) of U-^{2}H, $^{13}CH_3$-labeled proteins. The method is very similar to the 'standard' HMQC except that a simultaneous pair of ^{1}H and ^{13}C 180° pulses is applied in the center of the 2τ period to ensure that ^{1}H chemical

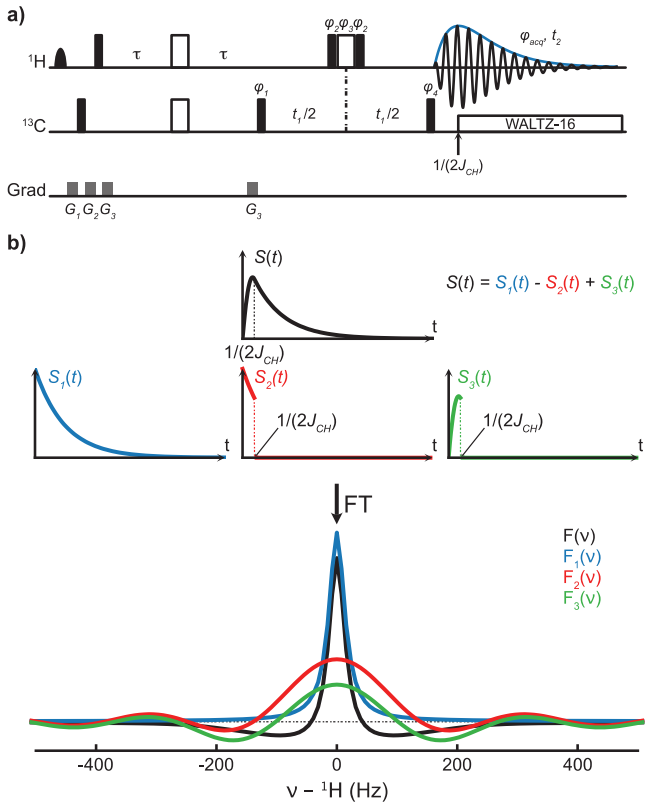


Fig. 1. (a) Pulse scheme of the ddHMQC experiment. All 90° (180°) non-selective pulses are denoted by black narrow (white wide) bars, applied with phase *x* unless indicated otherwise. ^{1}H and ^{13}C pulses are centered at 1 ppm and 20 ppm, respectively, with the exception of the shaped ^{1}H pulse at the start of the sequence which has an EBURP1 profile [34] (7 ms) and is applied on water. WALTZ-16 decoupling [27] (3.3 kHz field at 1 GHz) commences at a time of 1/(2J_{CH}) after the start of acquisition, once magnetization has refocused to in-phase. See Supporting Information for modified WALTZ16 decoupling scheme used. The value of τ is set to optimize sensitivity, typically less than 1/(4J_{CH}). The phase cycle is φ_1 = x,-x; φ_2 = 2(x),2(y),2(-x),2(-y); φ_3 = 2(y),2(-x),2(-y),2(x); φ_4 = 8(x),8(-x); φ_{acq} = 2(x,2(-x),x),2(-x,2(x),-x). A minimum phase cycle of two is required, although the number of scans should be set to a multiple of 4 with the modified WALTZ16 decoupling scheme. Gradient strengths (in % maximum) and durations are G_1=(10 %,1ms), G_2=(20 %,1ms), G_3=(30 %,0.5 ms). Suppression of the residual HDO signal is obtained with the selective water pulse - gradient scheme highlighted in the figure. When necessary, suppression can be achieved using a weak presaturation field (15–20 Hz) during the recycling delay. As described in Materials and Methods we strongly prefer to use D_2O solvents in studies of very high molecular weight systems to avoid saturation transfer effects from H_2O and, therefore, concomitant losses in sensitivity. (b) The resulting signal, $S(t)$ (black), can be decomposed into three components $S_i(t)$, $i \in \{1,2,3\}$, with $S(t) = S_1(t) - S_2(t) + S_3(t)$, as indicated. The Fourier transform (FT) of each component, $F_i(\upsilon)$ = FT($S_i(t)$), is shown along with $F(\upsilon)$ = FT($S(t)$), the resulting lineshape in the ^{1}H dimension (R_2 = 100 s^{-1}).

shift is refocused while concomitant evolution due to J_{CH} proceeds, and that signal acquisition begins immediately after the final 90° ^{13}C pulse, without first refocusing the anti-phase ^{1}H magnetization. As a result, decoupling is delayed by 1/(2J_{CH}) from the start of direct detection, where J_{CH} is the one-bond ^{1}H–^{13}C scalar coupling constant, approximately 125 Hz for all methyl groups considered here (124.5 ± 0.8 Hz for Leu, 125.2 ± 0.3 Hz for Ile-δ1, and 126.2 ± 0.3 Hz for Val [21] with the exception of Met where $J_{CH} \approx$ 140 Hz). As described previously [11–12], the delayed decoupling approach leads to a distorted lineshape in the ^{1}H dimension of ^{13}C–^{1}H spectra that can be calculated as the real part of the Fourier transform of the FID illustrated in Fig. 1b (black), where an on-resonance signal has been assumed with a decay rate, R_2, of 100 s^{-1} and J_{CH} = 125 Hz. In order to simplify the calculation $S(t)$ is divided into three time domains, $S_i(t)$, $i \in \{1,2,3\}$, such that $S(t) = S_1(t) - S_2(t) + S_3(t)$, Fig. 1b (middle), whose Fourier transforms can be calculated in a straightforward manner, as shown below. It follows that

$$F(\omega) = F_1(\omega) - F_2(\omega) + F_3(\omega) \tag{1.1}$$

where

$$F_1(\omega) = \int_0^\infty S_1(t)\exp(-i\omega t)dt = \int_0^\infty \exp(-R_2 t)\exp(-i\omega t)dt$$
$$F_2(\omega) = \int_0^{\frac{1}{2J_{CH}}} S_2(t)\exp(-i\omega t)dt = \int_0^{\frac{1}{2J_{CH}}} \exp(-R_2 t)\exp(-i\omega t)dt$$
$$F_3(\omega) = \int_0^{\frac{1}{2J_{CH}}} S_3(t)\exp(-i\omega t)dt = \int_0^{\frac{1}{2J_{CH}}} \sin(\pi J_{CH} t)\exp(-R_2 t)\exp(-i\omega t)dt \tag{1.2}$$

Evaluation of Eq. (1.2) gives

$$F_1(\omega) = \frac{R_2}{R_2^2+\omega^2}$$
$$F_2(\omega) = \frac{R_2}{R_2^2+\omega^2} + \frac{\left[-R_2\cos\left(\frac{\omega}{2J_{CH}}\right)+\omega\sin\left(\frac{\omega}{2J_{CH}}\right)\right]\exp\left(-\frac{R_2}{2J_{CH}}\right)}{R_2^2+\omega^2}$$
$$F_3(\omega) = \frac{1}{2}\left\{\frac{\omega-\pi J_{CH}}{D_1}\left[\sin\left(\frac{\omega}{2J_{CH}}\right)\exp\left(-\frac{R_2}{2J_{CH}}\right)-1\right] - \frac{R_2\cos\left(\frac{\omega}{2J_{CH}}\right)\exp\left(-\frac{R_2}{2J_{CH}}\right)}{D_1}\right\}$$
$$+\frac{1}{2}\left\{\frac{\omega+\pi J_{CH}}{D_2}\left[\sin\left(\frac{\omega}{2J_{CH}}\right)\exp\left(-\frac{R_2}{2J_{CH}}\right)+1\right] - \frac{R_2\cos\left(\frac{\omega}{2J_{CH}}\right)\exp\left(-\frac{R_2}{2J_{CH}}\right)}{D_2}\right\} \tag{2.1}$$

where

$$D_1 = (\omega - \pi J_{CH})^2 + R_2^2$$
$$D_2 = (\omega + \pi J_{CH})^2 + R_2^2. \tag{2.2}$$

Fig. 1b (bottom) plots $F_i(\nu)$, along with $F(\nu)$ ($\omega = 2\pi\nu$), with each frequency domain response color coded as for the corresponding time domain signal. $F_1(\nu)$ is a perfect Lorentzian line, since $S_1(t)$ is an exponentially decaying signal that extends to $t = \infty$, but $F_2(\nu)$ and $F_3(\nu)$ are more complex as they contain sinc wiggles, resulting from the truncation of each of the time domains at t = 1/(2J_{CH}) = 4 ms. In addition to the truncation artifacts, the $F_3(\nu)$ lineshape is further distorted by the fact that it is comprised of two multiplet components (at frequencies of ± J_{CH}/2 Hz), rather than single lines at 0 frequency for each of $F_1(\nu)$ and $F_2(\nu)$, with each of the two lines antiphase dispersive, as compared to the inphase absorptive components at ν = 0 for $F_1(\nu)$ and $F_2(\nu)$. The multiplet components of $F_3(\nu)$ are broad, as their lifetimes extend to only 4 ms, effectively giving rise to negative lobes on either side of a central peak along with additional wiggles that are due to truncation. Fig. 1b also shows $F(\nu)$ (black) that clearly illustrates the non-Lorentzian character of the resultant lineshape that is a linear combination of the three contributions, $F_i(\nu)$.

The lineshapes that are produced by the delayed decoupling approach are clearly not optimal as the broad negative tails from each peak will interfere with adjacent resonances (see below). The data can be manipulated in the time domain by dividing the FID by sin($\pi J_{CH}t$) for $0 < t < 1/(2J_{CH})$ so that $S(t)$ is transformed in a post-acquisition manner into $S_1(t)$, resulting in a pure Lorentzian lineshape after Fourier transformation of the time domain. This is an example of *J*-deconvolution, a procedure that dates back to work by Bothner-By and Dadok [22] who selected a doublet from a frequency domain spectrum, inverse transformed the spectral region to the time domain and then divided by cos(πJt) in this case to remove the scalar modulation, as is effectively done here. Simulations that we have done suggest that the deconvolution procedure is quite robust with respect to variability in J_{CH}. For example, a comparison of simulated data with J_{CH} = 140 Hz (*i.e.*, J_{CH} for Met methyl groups), followed by *J*-deconvolution using J_{CH} = 125 Hz, (*i.e.*, J_{CH} for Ile, Leu and Val methyl groups), and Fourier transformation shows that the mismatch leads to a small

increase in apparent linewidth, varying from 1 % to 6% for R_2 values between 10 s^{-1} and 150 s^{-1}.

A similar time domain manipulation has also been used recently in the context of the XL-ALSOFAST HMQC experiment [12], that is optimized for studies of fully protonated proteins with gradient coherence transfer selection, where delayed decoupling was also employed. In this scheme, however, direct detection does not proceed immediately after the final ^{13}C 90° pulse, but awaits the application of a gradient for coherence selection, so that the FID in this case is a mixture of in-phase and anti-phase magnetization at the start of direct acquisition.

A clear limitation of the *J*-deconvolution method, as pointed out in the original paper describing the approach [22], emerges from the near zero values of the signal that result from the sine/cosine modulations introduced by scalar coupled evolution. It, therefore, becomes necessary to divide by numbers close to zero at certain

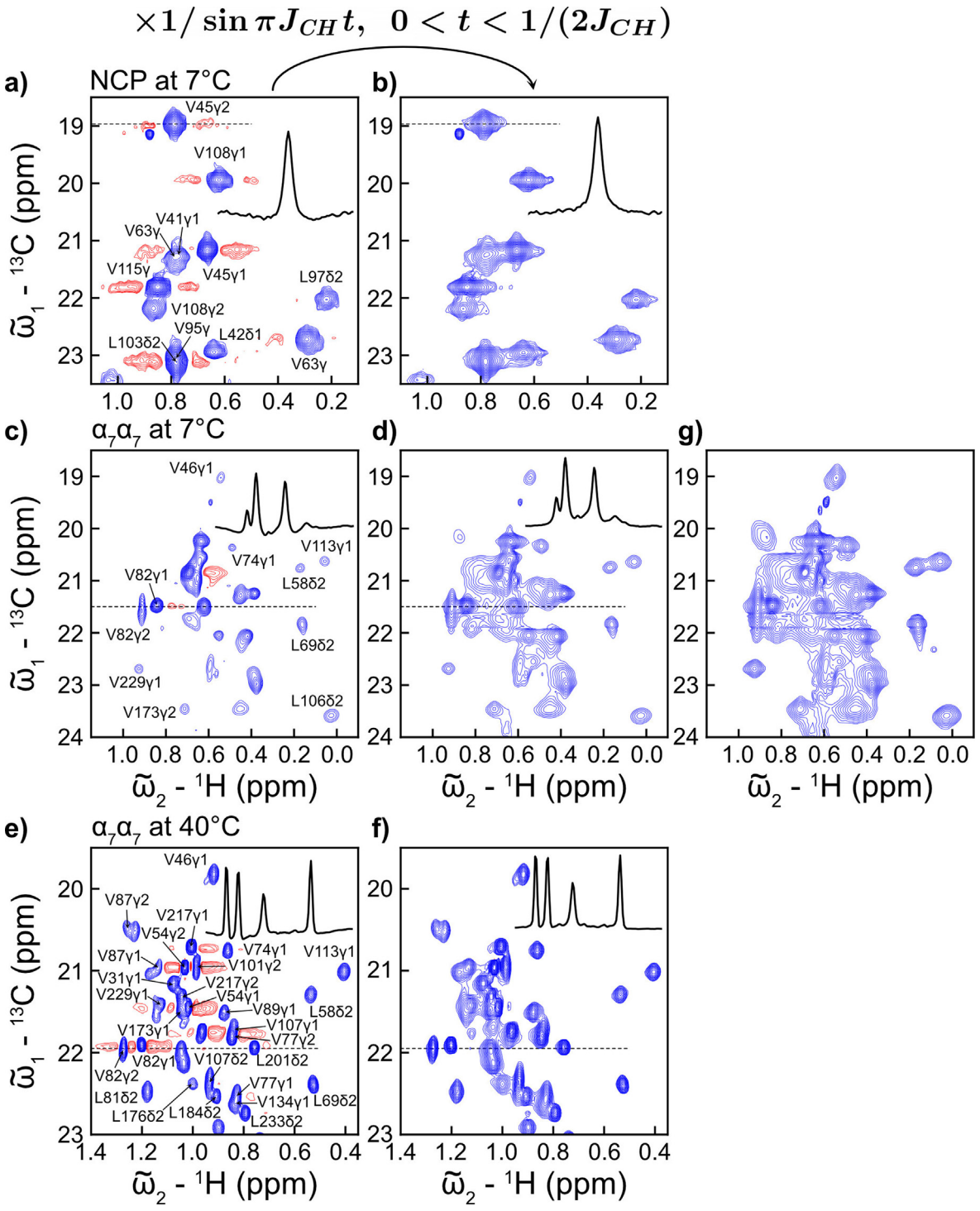


Fig. 2. Selected regions of ddHMQC spectra recorded on NCP (a,b) and $\alpha_7\alpha_7$ (c-g) samples (1 GHz) using the scheme of Fig. 1, either processed without (a,c,e) or with (b,d,f,g) *J*-deconvolution, contoured starting at the same level for datasets for a given sample (with the exception of panel g, same dataset as d, but where the contours start from a level that is twice lower). Traces of regions indicated by dashed lines are shown.

positions in the FID, that can result in large noise spikes. In some applications it is possible to judiciously choose acquisition times to avoid zero-crossings of magnetization. In the present case there are no zero-crossings as decoupling commences when the signal becomes in-phase at $t = 1/(2J_{CH})$ from the start of acquisition, however the signal at $t = 0$ is zero and it is, therefore, not possible to restore the first point using this procedure. This results in a DC offset for traces in the ^{1}H dimension that can be 'fixed' via standard baseline correction in the frequency domain, or by backwards linear prediction of the first point of the *J*-deconvoluted spectrum. A processing script that is based on the well-known *nmrPipe* software package [19] combined with an in-house written python based [23–24] *J*-deconvolution program is provided in Supporting Information. Notably, division by $\sin(\pi J_{CH}t)$ for small t values does increase the noise floor, in the applications considered here by 5–15 %, but this is significantly less than the sensitivity gains that can be realized in studies of molecular machines where R_2 values are typically very large (see below).

Fig. 2 highlights expanded regions from ddHMQC spectra (1 GHz), processed either without (a,c,e) or with *J*-deconvolution (b,d,f,g), recorded on highly deuterated nucleosome core particle (NCP) and $\alpha_7\alpha_7$ half proteasome samples. The NCP sample comprises two copies each of four histones (H2A, H2B, H3, and H4) along with a 147 basepair DNA fragment [14], where histone 2B (H2B) is labeled as Ile-δ1-$^{13}CH_3$, Leu,Val-$^{13}CH_3$ (referred to as ILV-$^{13}CH_3$-, or ILVM-$^{13}CH_3$ when Met-ε-$^{13}CH_3$ is labeled in what follows) and the other histones are NMR silent (fully deuterated; a,b), while the $\alpha_7\alpha_7$ sample was ILVM-$^{13}CH_3$ labeled (c-g). Only one of the two isopropyl methyl groups of Leu and Val is labeled as $^{13}CH_3$ (non-stereospecifically), while the other is $^{12}CD_3$, to minimize proton cross-relaxation [9]. In addition, the DNA of the NCP sample was deuterated and methyl groups added at C5 and A6 base positions [10]. Spectra were recorded at 7 °C (in addition to 40 °C for $\alpha_7\alpha_7$) to slow down the overall tumbling of the complexes. On the basis of previous estimates of the assumed isotropic correlation times of NCP and $\alpha_7\alpha_7$ complexes of 115 ns (45 °C) [10] and 120 ns (50 °C) [15], respectively, correlation times of 340 ns and 390(150) ns are calculated for NCP, 7 °C, and $\alpha_7\alpha_7$ at 7 °C(40 °C). Comparisons of spectra a and b (NCP), or c,d,g and e,f ($\alpha_7\alpha_7$), show that the *J*-deconvolution approach removes the tails of the lines, improving both resolution and sensitivity.

In order to quantify the sensitivity gains that are obtained by removing the final refocusing period in the ddHMQC pulse scheme, followed by *J*-deconvolution of the time domain data we have first recorded standard HMQC and ddHMQC spectra on a sample of ILV-$^{13}CH_3$ protein L, a 7.5 kDa protein with a correlation time of 5.3 ns at 25 °C [25]. Given the small size of the protein no sensitivity gain would be expected in the ddHMQC dataset over the HMQC correlation map. Fig. 3a compares intensity ratios of crosspeaks recorded using HMQC and ddHMQC schemes (1 GHz) and, as expected, the relative intensities of correlations are similar. Our control studies on protein L pointed out an issue with the efficacy of heteronuclear decoupling in the case of the ddHMQC. While ^{13}C decoupling is efficient when the signal is in-phase at the start of acquisition, we have found that modulation sidebands (on the order of 1 % or less of the main peak) are produced with delayed decoupling. These can be observed for small proteins in spectra recorded with high signal to noise (10,000:1, 4 scans in the case of protein L). We have examined a number of different decoupling sequences, including bi-level adiabatic decoupling with different adiabatic shapes [26], however the best performance was obtained using WALTZ-16 [27] (3.3 kHz B_1 field at 1 GHz). Application of a ^{13}C 90° purge pulse or a ^{13}C 180° pulse in alternate scans immediately prior to turning on the decoupler with no change in the receiver phase (*i.e.*, at $t_2 = 1/(2J_{CH})$ from the start of acquisition, using synchronous decoupling) so as to ensure only in-phase magnetization, did not improve the situation for reasons we currently do not understand. Notably, such artifacts were not observed in any of the applications to high molecular weight systems (see below) with the exception of crosspeaks from I0δ1 and I19δ1 in spectra recorded of the nucleosome particle at 25 °C (but not at 7 °C) that are both associated with the very flexible unfolded *N*-terminal tail of histone H2B. These, again, were less than 1 % of the main correlation. In our hands all sidebands can be eliminated by using a slightly modified WALTZ16 scheme whereby decoupling begins at different WALTZ4 elements within the WALTZ16 block. Thus, for scan j, decoupling starts at the (j%4)+1 WALTZ4 element, so that a minimum of 4 scans per FID is required. The decoupling file used (Bruker) is provided in Supporting Information.

While no sensitivity gain was observed for protein L, the situation, in contrast, is quite different for larger molecules. Panels b-d of Fig. 3 plot the ratios of peak intensities from ddHMQC and HMQC datasets as a function of the transverse relaxation rates of the slow component of methyl ^{1}H magnetization, $R_{2,S}$ (circles) for a number of different protein systems and temperatures. These include, HtrA2, a mitochondrial protease in equilibrium between

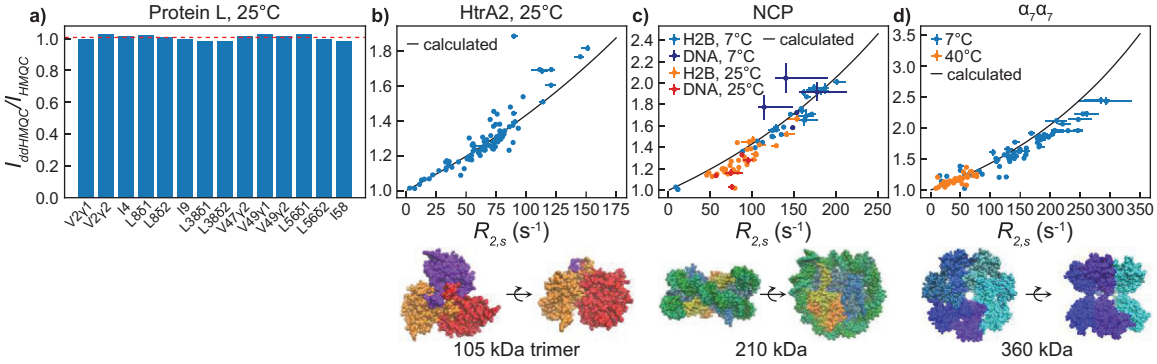
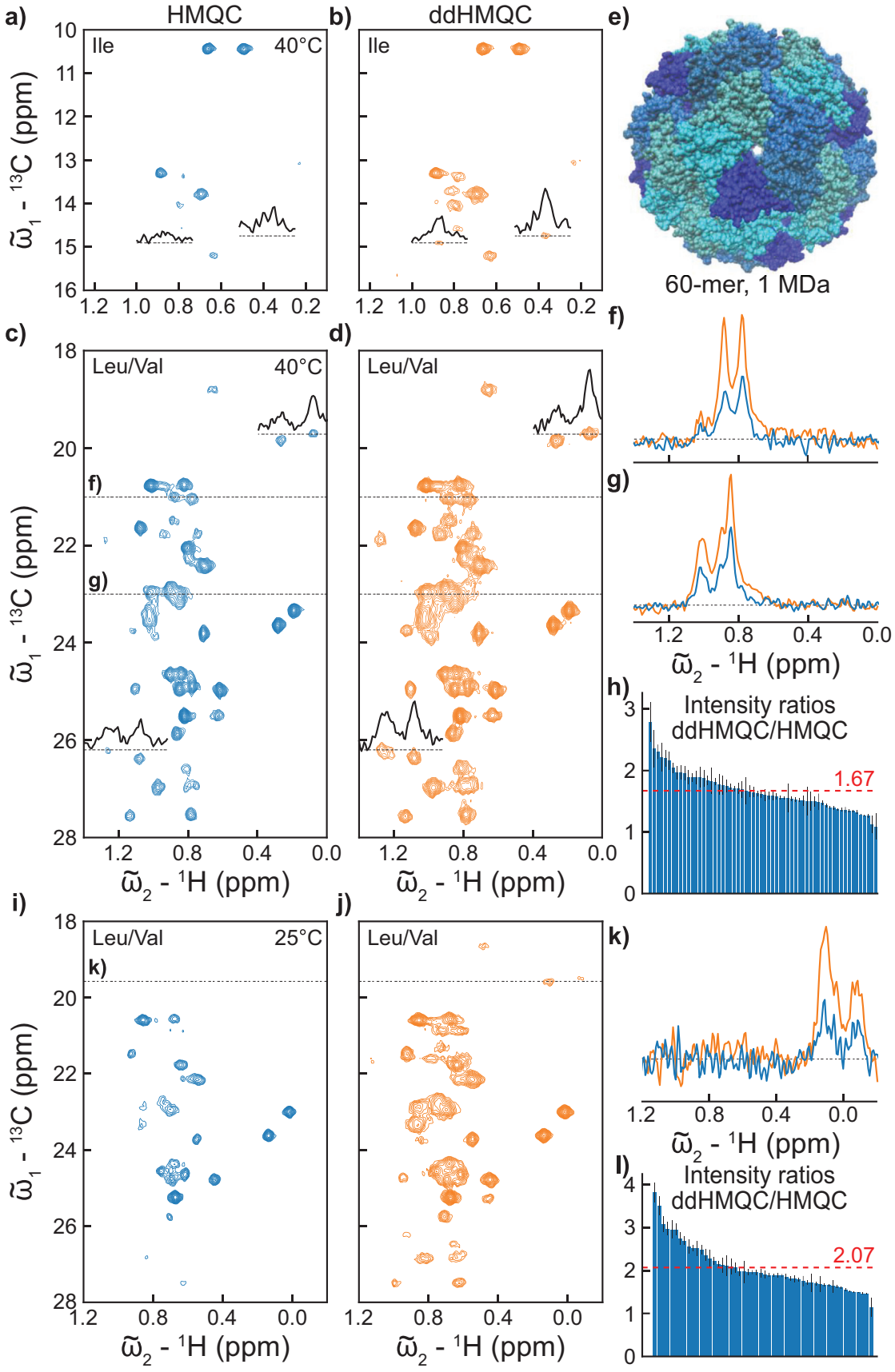


Fig. 3. Intensity ratios of correlations measured in ddHMQC and HMQC datasets (I_{ddHMQC}/I_{HMQC}) for molecules ranging in molecular mass from 7.5 kDa to 360 kDa. Although ratios of ≈ 1 are observed for protein L (a, dashed red line at $I_{ddHMQC}/I_{HMQC} = 1$), ratios increase for larger systems (b-d) and are correlated with ^{1}H transverse relaxation rates of the slowly relaxing components of ^{1}H magnetization, $R_{2,S}$. Cartoon diagrams of the structures of HtrA2 (pdb entry 1lcy [35]), NCP (pdb entry 6esf [36]), and $\alpha_7\alpha_7$ (pdb entry for Thermoplasma acidophilum proteasome 1pma [37]) are shown. In (c) intensity ratios from methyl groups of protein (referred to as H2B) and DNA (referred to as DNA) components are quantified. Noise levels (arbitrary units) are higher in ddHMQC spectra, with increases from 5.9 (HMQC) to 6.5 (ddHMQC, protein L, arbitrary units), 1.5 to 1.6 (HtrA2), 1.4 to 1.5 (NCP, 25 °C), 1.3 to 1.4 (NCP, 7 °C), 3.3 to 3.7 ($\alpha_7\alpha_7$, 40 °C) and 1.4 to 1.6 ($\alpha_7\alpha_7$, 7 °C). The increased noise floor is not included in the reported signal intensities in the paper. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)


a)
HMQC
Ile
40°C
b)
ddHMQC
Ile
e)
60-mer, 1 MDa
c)
Leu/Val
40°C
d)
Leu/Val
f)
g)
h)
Intensity ratios
ddHMQC/HMQC
1.67
i)
Leu/Val
25°C
k)
j)
Leu/Val
l)
Intensity ratios
ddHMQC/HMQC
2.07
ω1 - 13C (ppm)
ω2 - 1H (ppm)

trimeric (105 kDa) and hexameric conformers (25 °C,b; correlation time of the trimer approximately 110 ns [13]), 210 kDa NCP particles (7 °C and 25 °C,c), and 360 kDa $\alpha_7\alpha_7$, a dimer of disk shaped heptameric 21 kDa protomers (7 °C and 40 °C,d). Also highlighted are the three-dimensional structures of each complex. The solid line in each plot is the calculated expected intensity ratio after *J*-deconvolution of the ddHMQC data, given by $\exp(R_{2,S}2\tau)$, which takes into account the decay of magnetization during the final refocusing element of the HMQC scheme that is absent in the ddHMQC sequence. In the calculation we have assumed that the fast relaxing component of the magnetization, corresponding to the anti-TROSY component, decays to zero during the first 2τ element in the pulse scheme and the successive t_1 evolution period (see below). It is clear that the sensitivity gain grows as a function of $R_{2,S}$, and hence molecular weight.

As indicated above, both NCP and $\alpha_7\alpha_7$ spectra were recorded at 7 °C to increase their tumbling times as a mimic for even bigger structures. In our experience lowering the temperature often deteriorates spectral quality, not necessarily only because of the decreased tumbling rate, but also resulting from μs-ms timescale dynamics that are no longer averaged out at the lower temperatures. With this in mind we sought another system of increased molecular weight that could be studied at physiological temperatures. We have chosen lumazine synthase (AaLS), an enzyme that catalyzes the penultimate step in the biosynthesis of riboflavin (vitamin B2) in archaea, bacteria, and plants [16]. AaLS consists of 60 copies of 16 kDa protomers, arranged as a dodecamer of pentamers, giving rise to a virus-like spherical particle of approximately 1 MDa in molecular mass, with an assumed isotropic rotational correlation time that is calculated to be 370 ns and 540 ns at 40 °C and 25 °C, respectively, in D_2O solvent from HYDROPRO [28–29]. Fig. 4a-d illustrate spectral regions of ILVM-$^{13}CH_3$ AaLS (1 GHz, 40 °C) from both HMQC (a,c) and ddHMQC (b,d) datasets contoured at the same level. It is clear that the ddHMQC data includes correlations that are absent from the HMQC, and that there are very significant gains in sensitivity for a large number of peaks (compare traces in Fig. 4a-d and 4f,g), with an average signal gain of 1.7-fold (Fig. 4h). Larger gains (2.1-fold, on average) are obtained when the global tumbling is slowed down by lowering the temperature to 25 °C (Fig. 4i-l).

A comparison of HMQC and ddHMQC datasets of HtrA2 (protomers of 325 residues), recorded at 40 °C, 1 GHz and using a sample where only the *proR* methyl groups are $^{13}CH_3$ labeled [30], illustrates an interesting feature of the ddHMQC experiment, Fig. 5. In particular, when plotted at similar levels, ddHMQC spectra appear less well resolved than their HMQC counterparts. This reflects the fact that ddHMQC peak intensities are higher than those in HMQC spectra. Assuming that peak heights are, on average, *k* times larger, for example, the ratio of ddHMQC to HMQC linewidths close to the noise floor (*i.e.*, at much lower intensities than maximum) is approximately $k^{0.5}$ (see Supporting Information), leading to an apparent broadening of the ddHMQC correlations. A second contributing factor is that by shortening the HMQC pulse scheme to generate the ddHMQC it may well be that for "relatively small" proteins on the order of 100 kDa, such as HtrA2, the fast relaxing component is not completely eliminated during the abbreviated pulse scheme. Although the dominant component in spectra would remain the slowly relaxing one, the faster decaying component would contribute to the linewidth. This is illustrated in Fig. 5 where HMQC (a), ddHMQC (b), and purged ddHMQC (c) spectra for HtrA2 are highlighted (correlation time of approximately 70 ns at 40 °C), plotted at identical levels. In the case of the purged dataset (see Supporting Information) the fast relaxing magnetization is actively eliminated (Fig. S1), as described previously [31], giving rise to what appears to be a better resolved spectrum. However, the purge element, of duration $1/(4J_{CH})$, does degrade sensitivity because it both suppresses the fast relaxing component which may not decay to zero in this case, hence eliminating a contribution to the peak intensity, and also because of the additional delay in the sequence for purging (see Supporting Information) which affects the sensitivity of the slow relaxing component as well. As a result, the ratio of intensities of the purged ddHMQC to the HMQC is less than one in this case (0.94 ± 0.16). In general, therefore, the purge is not recommended.

As described above, significant sensitivity gains can be realized in the ddHMQC experiment, especially in applications to MDa complexes. In the case of AaLS, correlations that could not be observed in HMQC spectra (40 °C and 25 °C, 1 GHz) were present in the ddHMQC version, as highlighted in Fig. 4. It is important to emphasize, however, that the delayed decoupling/*J*-deconvolution approach is not flawless. A comparison of HMQC and ddHMQC spectra shows that the baselines of ddHMQC spectra can be slightly undulated, even after a 4th degree polynomial correction is applied in the final processing step (Figs. S2-S4). This reflects, somewhat, the fact that in the *J*-deconvolution scheme the effect of noise in the data is exacerbated by division by $\sin(\pi J_{CH}t)$ for $0 < t < 1/(2J_{CH})$. We have written a simple routine that corrects the baseline further (see Supporting Information) and the "corrected" spectra obtained in this manner are illustrated in Figs. S2-S4 as well. It is worth noting that such baseline distortions are also observed in HMQC spectra, but are not as pronounced. Second, as mentioned above, deconvolution does affect the noise floor, which is typically 5–15 % higher in ddHMQC datasets (see legend to Fig. 4 for values for each spectrum). In an effort to eliminate the *J*-deconvolution step we are in the process of developing artificial intelligence-based approaches [32] in which a trained neural network is able to properly reconstruct the initial $1/(2J_{CH})$ of the ddHMQC time domain. This work remains ongoing.

In summary, a simple pulse scheme (even shorter than the HMQC!) is described for methyl-TROSY based studies of high molecular weight complexes. The experiment is complementary to the XL-ALSOFAST scheme described by Gossert and coworkers [12] which is designed for non methyl-TROSY based studies of protonated protein systems, but with a focus here on applications involving molecular machines where sensitivity is critical, and, therefore, minimization of delays, as is possible by exploiting delayed decoupling, becomes essential. The idea of 'simplification through elimination' follows along the lines of a concatenation strategy introduced to complex pulse schemes that eliminates some pulses and delays, allowing evolution from more than one

Fig. 4. (a-d) Comparison of HMQC and ddHMQC spectra recorded on lumazine synthase, 40 °C, 1 GHz. (e) Cartoon representation of the structure, pdb entry 1hqk [38]. (f,g) Traces at specific ^{13}C frequencies are highlighted by dashed lines. The spectra are contoured starting at the same level. Several traces are also shown in panels a-d. (h) Intensity ratios of correlations measured in the ddHMQC and HMQC spectra at 40 °C. The average ratio is 1.67 and is shown with the dashed red line. (i,j) Comparison of HMQC and ddHMQC spectra (focusing on the Leu/Val region) recorded on lumazine synthase at 25 °C, 1 GHz, along with a trace (k) at the ^{13}C frequency highlighted by the dashed lines in the spectra. The spectra are contoured starting at the same level. (l) Intensity ratios of the correlations measured in the ddHMQC and HMQC spectra at 25 °C, with the average ratio of 2.07 shown with the dashed red line. Noise levels (arbitrary units) for ddHMQC:HMQC are 1.7:1.5 and 1.3:1.2 in spectra recorded at 40 °C and 25 °C. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

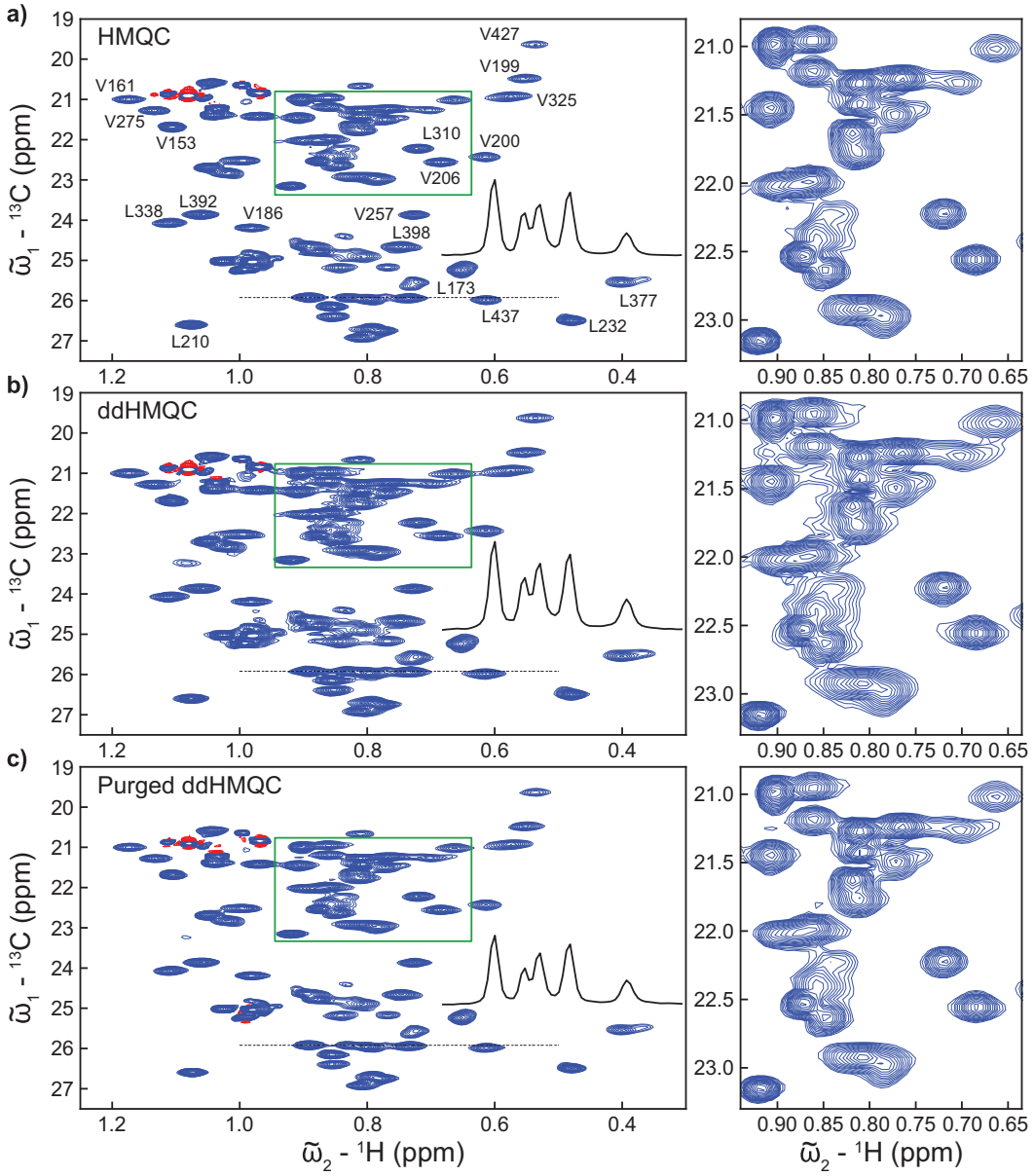


Fig. 5. Comparison of selected regions of HMQC (a), ddHMQC (b), and purged ddHMQC (c) spectra for HtrA2, 1 GHz, 40 °C, contoured starting at the same level. Regions in green boxes are expanded to the right. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

process to occur simultaneously in an effort to decrease relaxation losses [33]. The ddHMQC scheme leads to significant gains in applications to MDa complexes, such as AaLS, and it is anticipated that experiments based on this approach will prove extremely beneficial in further studies of even larger molecules.

Data availability

Data will be made available on request.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: L.E.K is an associate editor of this journal (JMR).

Acknowledgements

N.B.C, R.W.H., and Y.T. acknowledge post-doctoral support from the Canadian Institutes of Health Research (CIHR), as well as a Hospital for Sick Children Restracomp Fellowship (R.W.H.), a Japan Society for the Promotion of Science Overseas Research Fellowship (Y.T.), and a Uehara Memorial Foundation postdoctoral fellowship (Y.T.). A.I.M.S. is grateful for an Ontario Graduate Scholarship. The authors are grateful to Dr. Eriks Kupce, Bruker, for useful discussions about decoupling and to a reviewer for pointing out early literature on delayed decoupling. D.F.H. is supported by the Biotechnology and Biological Sciences Research Council UK (BBSRC) (ref: BB/T011831/1). L.E.K. acknowledges support from the CIHR (FND-503573) and the Natural Sciences and Engineering Council of Canada (2015-04347). For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmr.2022.107326.

References

[1] T.R. Alderson, L.E. Kay, NMR spectroscopy captures the essential role of dynamics in regulating biomolecular function, Cell 184 (3) (2021) 577–595.
[2] Biomolecular NMR spectroscopy. Chem. Rev. 122 (2022) 9265–10086.
[3] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, Proc. Natl. Acad. Sci. 94 (23) (1997) 12366–12371.
[4] V. Tugarinov, P.M. Hwang, J.E. Ollerenshaw, L.E. Kay, Cross-correlated relaxation enhanced ^{1}H–^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes, J. Am. Chem. Soc. 125 (34) (2003) 10420–10428.
[5] Y. Jiang, C.G. Kalodimos, NMR studies of large proteins, J. Mol. Biol. 429 (17) (2017) 2667–2676.
[6] A.M. Ruschak, L.E. Kay, Proteasome allostery as a population shift between interchanging conformers, Proc. Natl. Acad. Sci. 109 (50) (2012) e3454–e3462.
[7] L. Müller, Sensitivity enhanced detection of weak nuclei using heteronuclear multiple quantum coherence, J. Am. Chem. Soc. 101 (16) (1979) 4481–4484.
[8] A. Bax, R.H. Griffey, B.L. Hawkins, Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR, J. Magn. Reson. 55 (2) (1983) 301–315.
[9] V. Tugarinov, L.E. Kay, An Isotope Labeling Strategy for Methyl TROSY Spectroscopy, J. Biomol. NMR 28 (2) (2004) 165–172.
[10] G. Abramov, A. Velyvis, E. Rennella, L.E. Wong, L.E. Kay, A methyl-TROSY approach for NMR studies of high-molecular-weight DNA with application to the nucleosome core particle, Proc. Natl. Acad. Sci. 117 (23) (2020) 12836–12846.
[11] A. Bax, S.K. Sarkar, Elimination of refocusing pulses in NMR experiments, J. Magn. Reson. 60 (1) (1984) 170–176.
[12] P. Rößler, D. Mathieu, A.D. Gossert, Enabling NMR studies of high molecular weight systems without the need for deuteration: the XL-ALSOFAST experiment with delayed decoupling, Angew. Chem. Int. Ed. 59 (43) (2020) 19329–19337.
[13] Y. Toyama, R.W. Harkness, L.E. Kay, Dissecting the role of interprotomer cooperativity in the activation of oligomeric high-temperature requirement A2 protein, Proc. Natl. Acad. Sci. 118 (35) (2021), e2111257118.
[14] K. Luger, A.W. Mäder, R.K. Richmond, D.F. Sargent, T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8 Å resolution, Nature 389 (1997) 251–260.
[15] R. Sprangers, L.E. Kay, Quantitative dynamics and binding studies of the 20S proteasome by NMR, Nature 445 (7128) (2007) 618–622.
[16] E. Sasaki, D. Böhringer, M. van de Waterbeemd, M. Leibundgut, R. Zschoche, A. J.R. Heck, et al., Structure and assembly of scalable porous protein cages, Nat. Commun. 8 (1) (2017) 14663.
[17] A. Mittermaier, L.E. Kay, χ_1 torsion angle dynamics in proteins from dipolar couplings, J. Am. Chem. Soc. 123 (28) (2001) 6892–6903.
[18] I. Gelis, A.M.J.J. Bonvin, D. Keramisanou, M. Koukaki, G. Gouridis, S. Karamanou, et al., Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR, Cell 131 (4) (2007) 756–769.
[19] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (3) (1995) 277–293.
[20] V. Tugarinov, L.E. Kay, Relaxation rates of degenerate ^{1}H transitions in methyl groups of proteins as reporters of side-chain dynamics, J. Am. Chem. Soc. 128 (22) (2006) 7299–7308.
[21] A. Mittermaier, L.E. Kay, Effect of deuteration on some structural parameters of methyl groups in proteins as evaluated by residual dipolar couplings, J. Biomol. NMR 23 (2002) 35–45.
[22] A.A. Bothner-By, J. Dadok, Useful manipulations of the free induction decay, J. Magn. Reson. 72 (3) (1987) 540–543.
[23] J.J. Helmus, C.P. Jaroniec, Nmrglue: an open source Python package for the analysis of multidimensional NMR data, J. Biomol. NMR 55 (4) (2013) 355–367.
[24] S. van der Walt, S.C. Colbert, G. Varoquaux, The NumPy array: a structure for efficient numerical computation, Comput. Sci. Eng. 13 (2) (2011) 22–30.
[25] H. Sun, L.E. Kay, V. Tugarinov, An optimized relaxation-based coherence transfer NMR experiment for the measurement of side-chain order in methyl-protonated, highly deuterated proteins, J. Phys. Chem. B 115 (49) (2011) 14878–14884.
[26] Ē. Kupče, R. Freeman, G. Wider, K. Wüthrich, Suppression of cycling sidebands using Bi-level adiabatic decoupling, J Magn Reson A. 122 (1) (1996) 81–84.
[27] A.J. Shaka, J. Keeler, T. Frenkiel, R. Freeman, An improved sequence for broadband decoupling: WALTZ-16, J. Magn. Reson. 52 (2) (1983) 335–338.
[28] A. Ortega, D. Amorós, J. García de la Torre, Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models, Biophys. J. 101 (4) (2011) 892–898.
[29] J. García de la Torre, M.L. Huertas, B. Carrasco, Calculation of hydrodynamic properties of globular proteins from their atomic-level structure, Biophys. J. 78 (2) (2000) 719–730.
[30] P. Gans, O. Hamelin, R. Sounier, I. Ayala, M.A. Durá, C.D. Amero, et al., Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins, Angew. Chem. Int. Ed. 49 (11) (2010) 1958–1962.
[31] D.M. Korzhnev, K. Kloiber, V. Kanelis, V. Tugarinov, L.E. Kay, Probing slow dynamics in high molecular weight proteins by Methyl-TROSY NMR spectroscopy: application to a 723-Residue enzyme, J. Am. Chem. Soc. 126 (12) (2004) 3964–3973.
[32] G. Karunanithy, H.W. Mackenzie, D.F. Hansen, Virtual homonuclear decoupling in direct detection nuclear magnetic resonance experiments using deep neural networks, J. Am. Chem. Soc. 143 (41) (2021) 16935–16942.
[33] L.E. Kay, M. Ikura, A. Bax, The design and optimization of complex NMR experiments. Application to a triple-resonance pulse scheme correlating Hα, NH, and ^{15}N chemical Shifts in ^{15}N–^{13}C-labeled proteins, J. Magn. Reson. 91 (1991) 84–92.
[34] H. Geen, R. Freeman, Band-selective radiofrequency pulses, J. Magn. Reson. 93 (1) (1991) 93–141.
[35] W. Li, S.M. Srinivasula, J. Chai, P. Li, J.W. Wu, Z. Zhang, et al., Structural insights into the pro-apoptotic function of mitochondrial serine protease HtrA2/Omi, Nat. Struct. Biol. 9 (6) (2002) 436–441.
[36] S. Bilokapic, M. Strauss, M. Halic, Histone octamer rearranges to adapt to DNA unwrapping, Nat. Struct. Mol. Biol. 25 (1) (2018) 101–108.
[37] J. Löwe, D. Stock, B. Jap, P. Zwickl, W. Baumeister, R. Huber, Crystal Structure of the 20*S* Proteasome from the Archaeon *T. acidophilum* at 3.4 Å Resolution, Science 268 (5210) (1995) 533–539.
[38] X. Zhang, W. Meining, M. Fischer, A. Bacher, R. Ladenstein, X-ray structure analysis and crystallographic refinement of lumazine synthase from the hyperthermophile *Aquifex aeolicus* at 1.6 Å resolution: determinants of thermostability revealed from structural comparisons, J. Mol. Biol. 306 (5) (2001) 1099–1114.