Four-dimensional NOE-NOE spectroscopy of SARS-CoV-2 Main Protease to facilitate resonance assignment and structural analysis

Angus J. Robertson, Jinfa Ying, and Ad Bax

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

Correspondence: Ad Bax (bax@nih.gov)

This study is dedicated to Robert Kaptein on the occasion of his 80th birthday

ABSTRACT. Resonance assignment and structural studies of larger proteins by NMR can be challenging when exchange broadening, multiple stable conformations, and ^{1}H back-exchange of the fully deuterated chain pose problems. These difficulties arise for the SARS-CoV-2 Main Protease, a homodimer of 2×306 residues. We demonstrate that the combination of four-dimensional (4D) TROSY-NOESY-TROSY spectroscopy and 4D NOESY-NOESY-TROSY spectroscopy provides an effective tool for delineating the ^{1}H-^{1}H dipolar relaxation network. In combination with detailed structural information obtained from prior X-ray crystallography work, such data are particularly useful for extending and validating resonance assignments, as well as for probing structural features.

1. Introduction

The extension of conventional two-dimensional ^{1}H-^{1}H NMR spectroscopy of natural proteins (Wüthrich, 1986) to three-dimensional (3D) homonuclear NMR experiments offered the ability to simplify spectral analysis by removing resonance overlap (Vuister et al., 1988;Oschkinat et al., 1988) and by providing access to a direct, more detailed analysis of ^{1}H-^{1}H dipolar cross relaxation networks. In particular, the homonuclear 3D NOE-NOE experiment (Boelens et al., 1989;Breg et al., 1990) not only decreased resonance overlap, it also directly elucidated spin-diffusion pathways. This information complemented and validated the elegant relaxation matrix analysis of spin diffusion (Boelens et al., 1988).

Such homonuclear ^{1}H 3D experiments and analysis strategies were soon followed by a myriad of heteronuclear 3D experiments that required isotopic enrichment, and therefore cloning and bacterial overexpression (Marion et al., 1989b;Zuiderweg and Fesik, 1989;Ikura et al., 1990;Marion et al., 1989a;Wagner, 1993). Most of these heteronuclear experiments simply served to disperse the regular ^{1}H-^{1}H 2D spectrum into a third dimension, thereby removing spectral overlap but providing little or no new information on the all-important ^{1}H-^{1}H spin diffusion pathways. The 3D NOESY-HMQC experiment (Marion et al., 1989b;Zuiderweg and Fesik, 1989) subsequently was extended to four dimensions (4D), thereby dispersing the conventional 2D ^{1}H-^{1}H NOESY experiment into two additional dimensions that correspond to the chemical shifts of the nuclei to which each of the protons is covalently bound (Kay et al., 1990;Clore et al., 1991;Zuiderweg et al., 1991).

These multi-dimensional experiments provided a tremendous degree of spectral simplification, in particular after appropriate analysis software became available. However, it also quickly became clear that extension to large, slowly tumbling proteins was hampered by low signal-to-noise, caused by the relative inefficiency of the magnetization transfer steps when the dimensionality of a spectrum is increased. This decrease in sensitivity was remedied by generating the protein in a highly perdeuterated state, while keeping the solvent-exchangeable backbone amide protons protonated (Torchia et al., 1988;Lemaster and Richards, 1988). Combining the perdeuteration approach with both the triple resonance assignment strategy (Grzesiek et al., 1993) and the subsequently introduced powerful TROSY line-narrowing method (Pervushin et al., 1997) made it possible to assign and analyze the structure of quite large proteins, as exemplified by the 723-residue protein malate synthase G (Tugarinov et al., 2002;Tugarinov et al., 2005a). The sensitivity gained by perdeuteration, enabling the recording of 4D ^{15}N-separated NOE spectra, also was key in solving the structure of a HIV-1 accessory protein that had been too challenging for analysis by more conventional methods (Grzesiek et al., 1995). Zhu and co-workers introduced a TROSY-NOESY-TROSY version of Grzesiek's 4D NOESY experiment which, illustrated for a partially deuterated 27 kDa protein, yielded further improved sensitivity and intrinsic ^{1}H and ^{15}N line widths (Xia et al., 2000). Their implementation relied on eight-step phase cycling, thereby limiting digitization of the time domain data and unable to exploit the improved TROSY relaxation in the indirect dimensions. Diercks *et al.* introduced an elegant method to suppress diagonal signals from a 4D TROSY-NOESY-TROSY spectrum (Diercks et al., 2010), a particularly useful feature when spectral resolution is limited. However, we opted not to use this implementation as, at the long mixing times used, the diagonal resonances serve as convenient reference anchors during analysis and are sufficiently attenuated such as not to obscure nearby peaks in our 4D spectra that are of very high digital resolution.

In the present report, we merge the above-mentioned prior advances, 3D NOE-NOE and 3D ^{15}N-separated NOESY, into a 4D experiment. In combination with extensive perdeuteration and gradient-enhanced encoding to enable a four-step phase cycle as well as non-uniform sampling (NUS) (Rovnyak et al., 2004), the experiments take better advantage of the improved resolution afforded by 4D NMR. We demonstrate the utility of the experiments by applying them to the study of the main protease of SARS-CoV-2 (M^{pro}) which is the virus responsible for coronavirus-2019 disease (COVID-19). M^{pro}, also known as $3CL^{pro}$ or Nsp5, is a homodimeric cysteine protease of

2×306 residues that does not have closely related mammalian homologues and is therefore an intense target for drug development, with a promising inhibitor now entered in a phase I clinical trial (Boras et al., 2021). Its NMR analysis is challenging, not only for its large size (67.6 kD), but also because of the presence of a minor conformer associated with the cis-isomer of one of its 13 Pro residues (P184), the difficulty in back-exchanging all backbone amide protons when the protein is expressed in $^{2}H_2O$, and the presence of intermediate time scale motions that lead to exchange broadening in the vicinity of the protein's active site. Here we focus on an enzyme variant where the catalytic Cys145 residue has been mutated to Ala (M^{pro}_{C145A}), a construct that is stable for multiple weeks at the high concentrations required for NMR spectroscopy. The assignment process and a full structural analysis of the protein will be presented elsewhere. The focus of the present work is on technical innovations, including recording two types of 4D NOE-based NMR spectra that proved invaluable both for the validation of resonance assignments, as well as the subsequent structural analysis.

2. Methods and Experiments

2.1 Protein production

The gene encoding a C145A variant of M^{pro} (M^{pro}_{C145A}) with an N-terminal affinity (and solubility), tag was synthesized by GenScript (USA) and then cloned into a Pet24a+ plasmid between BamH1 and Xho1 restriction sites. The fusion protein encoded for 6His tag – GB1–SG rich linker–TEV cleavage site – M^{pro}_{C145A}, and was purified according to methods collectively developed by the COVID-19 NMR consortium (Altincekic, 2021). In brief, following cell culture and harvesting (see Supporting Information), the cell lysate was passed down a 6His-affinity column (IMAC) and eluted in a small volume; the solubility tag was cleaved off to generate a native N-terminus; the reaction mix was then passed through an IMAC column to remove uncleaved protein, before size separation on a Sephadex G75 column. The resulting, extensively perdeuterated, $^{2}H,^{15}N$-M^{pro}_{C145A} homo dimer was used for experiments and throughout the recombinant protein expression, extensive care was taken to achieve a high (>~98%) level of deuteration of the non-exchangeable hydrogens. For more detail, see Supporting Information.

2.2 Recording of NMR data

Spectra were acquired on a sample containing 1.8 mM (0.9 mM dimer) $^{2}H,^{15}N$-M^{pro}_{C145A} in 10 mM

sodium phosphate, pH 7.0, 0.5 mM TCEP, 3% v/v ^2H_2O and 0.3 mM sodium trimethylsilylpropanesulfonate (DSS; as an internal chemical shift reference), in a 300 µL Shigemi microcell. All experiments were recorded at 25 ˚C on an 800 MHz Bruker Avance III spectrometer, equipped with a 5-mm TCI probe containing a triple-axis gradient accessory, and running TopSpin software version 3.1.

Considering that the measurements are carried out for perdeuterated protein, the spectral windows in the indirect ^1H dimensions were limited to the region downfield from the water resonance, and EBURP2 pulses (Geen and Freeman, 1991) were used where needed to prevent excitation of protons outside of this region. For the 4D TROSY-NOESY-TROSY experiment (Fig. 1a), only a small fraction (0.54%) of the full data matrix, consisting of 1536* (^1H, t_4, 119.8 ms) × 90* (^{15}N, t_3, 35.1 ms) × 91* (^1H, t_2, 20.0 ms) × 90* (^{15}N, t_1, 35.1 ms) complex points, was recorded (SI Table S1) using nonuniform sampling for the indirect dimensions. Selection of the sampled data points was unweighted, randomly distributed, and without time-ordering. The sampling schedule (included in the uploaded data sets, see Data availability) was automatically generated by Topspin 3.1 when the experiment started, with the random seed left to its default value (54321) and the T_2 for each indirect dimension set to 1 second to avoid significant weighting of the sampling schedule. Specifically, a total of 31896 FIDs along the directly detected t_4 dimension were recorded. Using 4 scans per FID for phase cycling, and an interscan delay of 2.07 seconds, the total measurement time was approximately 88 hours, but three-fold shorter would have sufficed (see Concluding remarks).

Nonstandard processing was needed for the TROSY-NOESY-TROSY experiment because the spectrum was recorded with sensitivity-enhanced gradient selection in the ^{15}N t_1 evolution period that preceded the NOE mixing (Xia et al., 2000). Specifically, the 4D NUS data set was first sorted and expanded according to the sampling schedule using the *nusExpand.tcl* script within the NMRPipe software package (Delaglio et al., 1995). The expanded data was then converted to the NMRPipe format, with the quadrature mode for t_3 set to *Echo-AntiEcho*, while the quadrature mode for t_1 was temporarily set to *Complex*. After the conversion, the 4D matrix needs to be transposed to enable use of the NMRPipe macro *bruk_ranceA.M* to correctly reshuffle the data, turning the phase-modulated t_1 dimension into conventional amplitude modulated data prior to processing as regular, complex data. This transposition is accomplished by reading in the NMRPipe-formatted matrix with the z-axis along the t_2 dimension, application of the macro, and

restoring the data to its original axis order prior to regular processing, with the full script included with the raw time domain data sets (see Data availability). For the processing, the direct dimension was apodized with a squared, shifted sine bell window, spanning from 72º to 176.4º, whereas an additional 15 Hz exponential line broadening was used to better match the apodization window to the natural decay of the signal, thereby improving the signal to noise ratio (S/N). This was followed by zero filling and Fourier transformation. Subsequently, the indirect data points that were not experimentally sampled were reconstructed using the SMILE (Ying et al., 2017) program, and the reconstructed data was further processed in NMRPipe. To enhance the spectral resolution, by default the acquisition times in all indirect dimensions were extended by 50% during the SMILE reconstruction, leading to an effective sampling sparsity of 0.16%. The data matrix for the final reconstructed 4D spectrum consists of 614 (^{1}H, F_4, 6.3 Hz/point) × 512 (^{15}N, F_3, 5.0 Hz/point) × 512 (^{1}H, F_2, 8.9 Hz/point) × 512 (^{15}N, F_1, 5.0 Hz/point) real points (see SI Table S1).

The full time-domain data matrix of the 4D NOESY-NOESY-TROSY experiment (Fig. 1b) consists of 1536* (^{1}H, t_4, 95.8 ms) × 90* (^{15}N, t_3, 35.1 ms) × 60*(^{1}H, t_2, 12.0 ms) × 60* (^{1}H, t_1, 12.0 ms) complex points (SI Table S1). An unweighted, random NUS sampling scheme with a sparsity of 1.69% (corresponding to 43856 t_4 FIDs) was used to record a small subset of the data points. Using an interscan delay of 1.77 s and 4 scans per FID, the total experimental time was approximately 110 hours, but three-fold shorter would have sufficed (see Concluding remarks). The data was processed and reconstructed in the same manner as described above with 50% extension of all indirect dimensions during the SMILE reconstruction, resulting in an effective sparsity of 0.50% and a final spectral matrix size of 492 (^{1}H, F_4, 7.8 Hz/point) × 512 (^{15}N, F_3, 4.7 Hz/point) × 512 (^{1}H, F_2, 13.9 Hz/point) × 512 (^{1}H, F_1, 13.9 Hz/point) real points (SI Table S1). Note that since in the NOESY-NOESY-TROSY experiment the data was recorded using the Echo-AntiEcho mode (Kay et al., 1992) only in the t_3 dimension, immediately preceding acquisition, the *bruk_ranceA.M* macro was not needed after the conversion of the expanded NUS data. The residual in-phase axial peaks along the F_2 dimension were treated as real peaks and optimally reconstructed by SMILE to suppress the sampling artifacts of the axial signals from spreading to the regions with NOE peaks. The processing macros used for both 4D spectra are included with the raw time domain data (see Data availability).

2.3 Spectrum analysis

Spectra were processed using NMRPipe software (Delaglio et al., 1995); peak picking and spectrum analysis was performed using SPARKY software (Goddard and Kneller, 2008;Lee et al., 2015) as well as NMRDraw (Delaglio et al., 1995). Programs for visualization and analysis were written using freely available python libraries (Hunter, 2007;Harris et al., 2020), as well as NMR-specific python libraries (Helmus and Jaroniec, 2013).

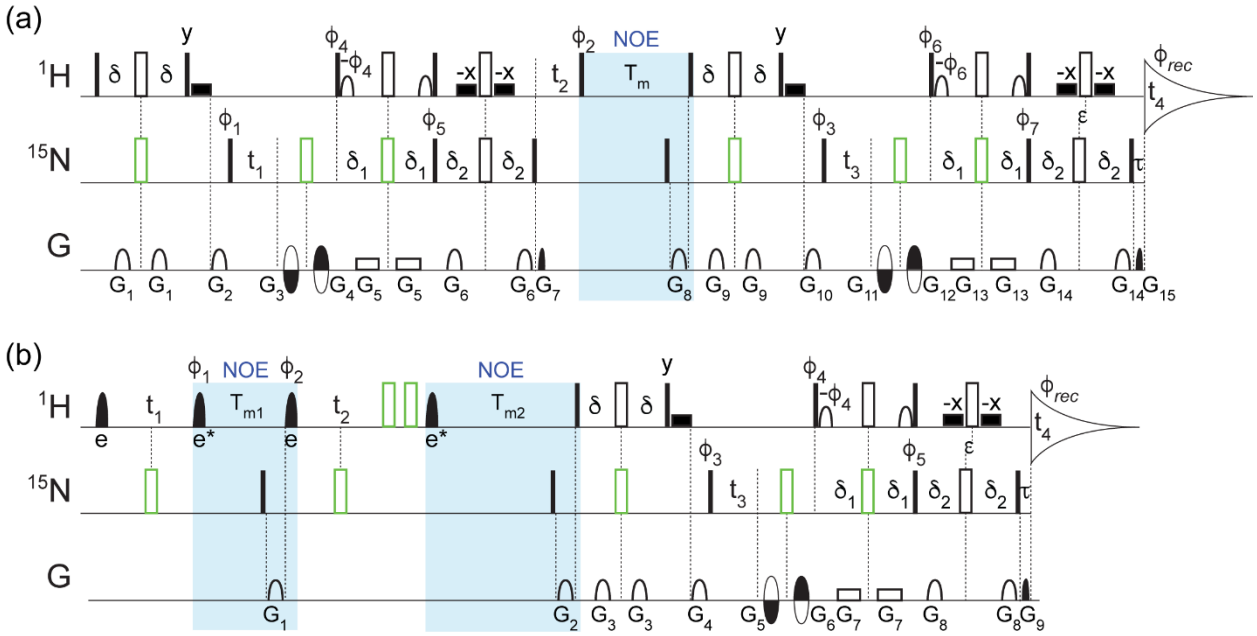

Figure 1. Pulse schemes for four-dimensional **(a)** TROSY-NOESY-TROSY and **(b)** NOESY-NOESY-TROSY experiments. The filled and open rectangular bars on the ^{1}H and ^{15}N channels represent 90° and 180° pulses, respectively. The open bars in green represent composite 180° pulses consisting of $90°_{x}207°_{y}90°_{x}$ for ^{1}H and $90°_{x}216°_{y}90°_{x}$ for ^{15}N. The filled shaped ^{1}H pulses correspond to selective EBURP2 (labeled e) and time-reversed EBURP2 (labeled e*) pulses (Geen and Freeman, 1991), while the open ^{1}H shaped pulses represent 90° water-flipback pulses (center lobe of a sinc profile, 1.1 ms duration at 800 MHz) (Grzesiek and Bax, 1993). The wide filled rectangular boxes denote 90° water-flipback pulses (also 1.1 ms duration at 800 MHz). Unless indicated otherwise, all pulses were applied along x. The following delays were used for the initial INEPT and TROSY transfers: $\delta = 2.1$ ms, $\delta_1 = 2.1$ ms, $\delta_2 = 2.5$ ms. The ^{1}H chemical shift evolution during the delay $\tau = 0.181$ ms was compensated by offsetting the last pair of ^{1}H and ^{15}N 180° pulses by $\varepsilon = \tau/2$ to avoid the linear phase error in the t_4 dimension. **(a)** For the TROSY-NOESY-TROSY: NOE mixing time $T_m = 200$ ms; phase cycling: $\phi_1 = y$, $\phi_2 = x, x, -x, -x$, $\phi_3 = y, -y$, $\phi_4 = y$, $\phi_5 = y$, $\phi_6 = y$, $\phi_7 = y$, $\phi_{rec} = y, -y, -y, y$; gradients were sine-bell or rectangular shaped (as depicted in the figure) with durations $G_{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15}$ = 0.977, 1.2, 0.4, 0.4, 0.986, 0.977, 0.081, 1.7, 0.977, 1.2, 0.4, 0.4, 0.986, 0.977, and 0.081 ms, z-strengths of 21.7, 28.7, -25.9, 32.9, 2.1, 25.9, 29.4, 30.8, 21.7, 28.7, -31.5, 38.5, 2.1, 25.9, and 35.0 G/cm, and additional x- and y-strength of -22.5, 27.5, and 25.0 G/cm for G_{11}, G_{12}, and G_{15}, respectively. The duration of decoding pulses G_7 (G_{15}) was empirically optimized for maximum signal and can differ from the

theoretical value derived from the gyromagnetic ratios of ^{15}N and ^{1}H and the encoding pulses G_3+G_4 (G_{11}+G_{12}) by several microseconds due to rise and fall times of short gradient pulses. Quadrature detection in t_3 (t_1) was achieved using the Echo-AntiEcho scheme (Kay et al., 1992) by inverting the encoding gradient G_{11} and G_{12} (G_3 and G_4) together with ϕ_6 and ϕ_7 (ϕ_4 and ϕ_5) to obtain the second FID for every t_3 (t_1) increment. The t_2 dimension was acquired using States-TPPI by incrementing ϕ_2 by 90°. **(b)** For the NOESY-NOESY-TROSY, the selective EBURP2 and time-reversed EBURP2 pulses have a duration of 1.0 ms at 800 MHz, centered at 8.3 ppm, thereby exciting the amide protons downfield from the water resonance without perturbing any upfield exchangeable protons, or residual aliphatic protons resulting from imperfect perdeuteration. NOE mixing times, T_{m1} = 50 ms; T_{m2} = 300 ms. Phase cycling: ϕ_1 = x, x, -x, -x, ϕ_2 = x-π/4, ϕ_3 = y, -y, ϕ_4 = y, ϕ_5 = y, ϕ_{rec} = y, -y, -y, y. Gradients were sine bell or rectangular shaped with durations $G_{1,2,3,4,5,6,7,8,9}$ = 1.7, 1.2, 0.977, 1.2, 0.4, 0.4, 0.986, 0.977, and 0.081 ms, z-strengths of 20.3, 30.8, 21.7, 28.7, -31.5, 38.5, 2.1, 25.9, and 35.0 G/cm, and additional x- and y-strengths of -22.5, 27.5 and 25.0 G/cm for G_5, G_6, and G_9, respectively. The duration of G_9 was empirically optimized for maximum signal. Quadrature detection in t_3 was achieved using the Echo-AntiEcho scheme by inverting the encoding gradient G_5 and G_6 together with the ϕ_4 and ϕ_5 to obtain the second FID for every t_3 increment, while States-TPPI was used to obtain quadrature in the t_1 dimension, by incrementing the ϕ_1 pulse phase by 90°, and for t_2 by incrementing ϕ_2 by 90°. Pulse sequence code and parameter files can be downloaded from https://zenodo.org/record/4625615.

3. Results and Discussion

Two types of complementary 4D NOE experiments were recorded: (1) 4D TROSY-NOESY-TROSY and (2) 4D NOESY-NOESY-TROSY (Fig. 1). While the former is very similar to the HMQC-NOESY-TROSY experiment used recently for a single α-helical domain with a long rotational correlation time (Barnes et al., 2019), the 4D NOESY-NOESY-TROSY experiment extends earlier work by Kaptein and co-workers (Boelens et al., 1989;Breg et al., 1990).

3.1 Recording and analysis of the 4D TROSY-NOESY-TROSY spectrum

The rotational correlation time of the C145A variant of M^{pro} (M^{pro}_{C145A}) at 25 ˚C is *ca.* 27 ns, and consequently, transverse relaxation is rapid for both ^{15}N and $^{1}H^{N}$ nuclei. For this reason, it proved beneficial to substitute a TROSY element for the HMQC or HSQC segment that was previously used for such measurements (Kay et al., 1990;Barnes et al., 2019). Even though the TROSY element only utilizes half of the amide $^{1}H^{N}$ magnetization present at the start of the pulse sequence, combining its ^{15}N evolution with sensitivity-enhanced gradient selection during the subsequent t_2 evolution period (Fig. 1a) limits the loss to √2, or even somewhat less when taking the gain from the ^{15}N Boltzmann magnetization into account (Pervushin et al., 1998). A 2D TROSY spectrum (Fig. S1) of this sample allowed identification of 261 backbone amide peaks out of 293 non-proline residues, suggesting the feasibility of implementing the TROSY version of the 4D NOESY experiment. Conformational exchange on a time scale that results in extensive line broadening and incomplete back exchange of amides when the protein was purified in $^{1}H_2O$ are the primary causes for the absence of the *ca.* 30 amide signals.

The high quality and S/N of the TROSY-HSQC spectrum (Fig. 2b), suggested the feasibility of implementing the TROSY version of the 4D NOESY experiment. Combined with the enhanced relaxation properties during t_1 and t_2 evolution of the TROSY-selected coherence, we found experimentally that spectral quality attainable for M^{pro}_{C145A} with the 4D TROSY-NOESY-TROSY was better than with the HMQC-NOESY-TROSY version of the experiment, consistent with the previous report that the TROSY implementation improved both the sensitivity and resolution over the 4D HSQC-NOESY-HSQC (Xia et al., 2000). Figure 2 shows expanded regions of six (F_1,F_2) cross-sections through the 4D spectrum, each orthogonal to the (F_3,F_4) frequencies of the six amide

correlations that are highlighted in a section of the regular 2D ^{1}H-^{15}N TROSY-HSQC spectrum of Figure 2b. A total of 231 peaks, out of the 261 peaks in the 2D TROSY spectrum, can be detected as (semi-)resolvable diagonal peaks in the projected ^{15}N-^{1}H (F_3,F_4) plane (data not shown). These numbers do not include the doubling of resonances associated with isomerization of P184.

The cross sections exemplify the power of 4D analysis for three types of secondary structure: α-helix (Fig. 2c), β-sheet (Fig. 2d), and a loop region (Fig. 2e). Due to the long NOE mixing time used in this experiment (200 ms), substantial spin diffusion occurs which results in numerous NOE correlations for each amide. For example, α-helical residues L232 and M235 not only show NOE interactions with one another, but also share NOE cross peaks to V233 and A234, with M235 even showing a weak cross peak to N231. Such correlations are particularly useful for validating the assignments obtained from the limited number of triple resonance backbone assignment experiments that are applicable to larger proteins such as M^{pro}.

The amides of L67 and Q69 in strand β4 only share a single NOE, to sequential residue V68, but they show valuable long-range NOEs to amide protons in strands β1 (C22) and β5 (L75). G195 and D197, located in the long loop that connects strand β13 to helix α6, have an NOE to one another as well as sequential NOEs, but show no long-range interactions, consistent with the X-ray structure (Douangamath et al., 2020). However, NOEs from L67 or Q69 to T21 or Q19 are not observed, despite close proximity, due to the minimal back exchange of amide protons in the β1 strand.

It is interesting to compare the diagonal peak intensities in these various cross sections of the TROSY-NOESY-TROSY spectrum. Diagonal intensity is a function of the amount of amide ^{1}H z magnetization present at the start of the pulse sequence, *i.e.*, it depends on the non-selective longitudinal relaxation time of the amide proton, but also on the attenuation of this magnetization during the NOE mixing time, in other words, on the selective longitudinal relaxation time which is dominated by J(0) spectral density terms. The latter dominate the differences in diagonal intensity seen in the various cross sections. For example, the helical amides of L232 and M235 rapidly lose their magnetization to their proximate sequential amide neighbors, separated by *ca.* 2.7 Å, that each are in close contact with other neighboring protons. By contrast, none of the L67, Q69, G195 and D197 amides are closer than 3.7 Å from any neighboring protonated amide in the 1.25 Å X-ray structure of M^{pro} (Douangamath et al., 2020), causing their diagonal intensities to

remain high.

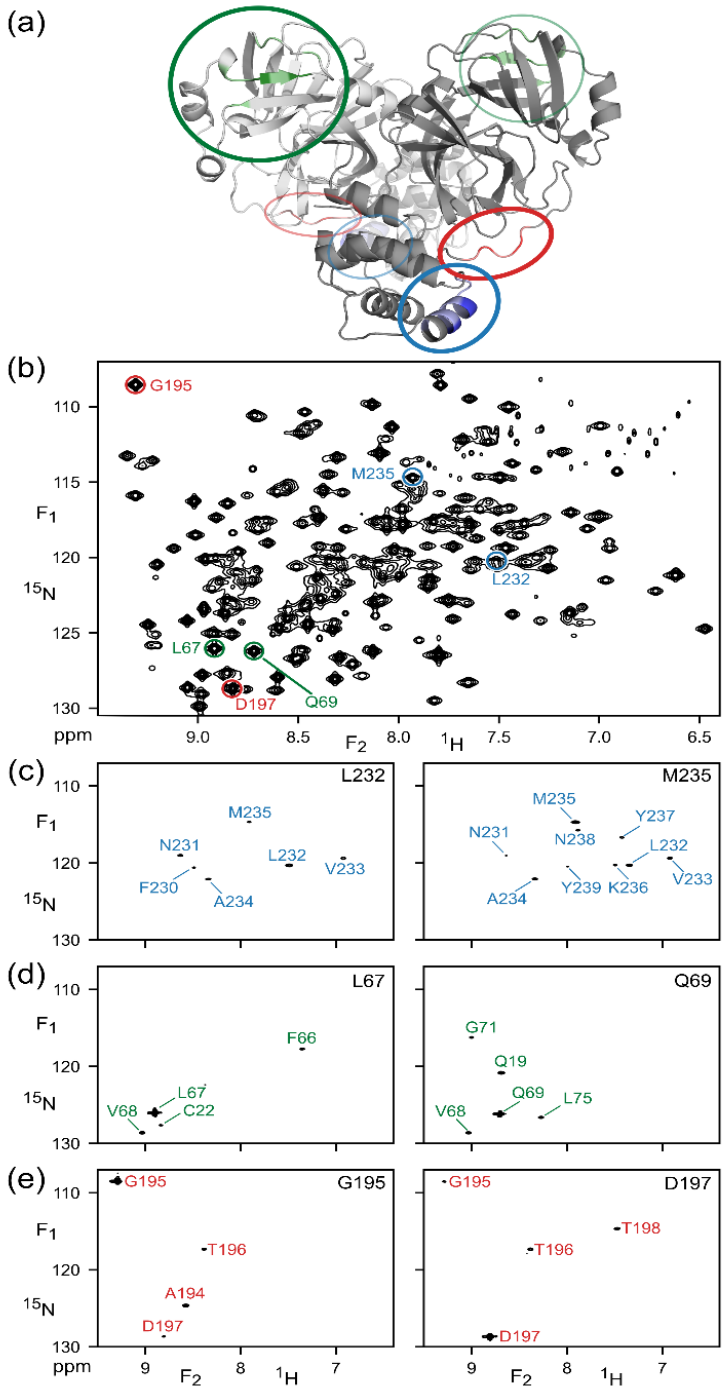

Figure 2. Illustration of amide-amide NOEs in perdeuterated, amide-protonated SARS-CoV-2 Main Protease, observed by 4D TROSY-NOESY-TROSY. (a) Ribbon diagram depicting the backbone homodimeric X-ray structure (PDB: 6LU7), with colors marking the regions that are highlighted in (F_1,F_2) cross sections taken at the (F_3,F_4) coordinates of (c) L232 and M235 (blue); (d) L67 and Q69 (green) and (e) G195 and D197 (red). These resonances are marked in (b), which represents the most crowded region of the 800-MHz ^{1}H-^{15}N TROSY-HSQC (the full spectrum, is shown in SI Fig. S1).

3.1 Recording and analysis of the 4D NOESY-NOESY-TROSY spectrum

As highlighted by the work of Kaptein and co-workers, 3D NOE-NOE experiments provided an effective method for studying the ^{1}H-^{1}H cross relaxation network in proteins in more detail. Here, we extend this powerful experiment to four dimensions, making it more straightforward to analyze such a spectrum, while limiting the relaxation pathways by perdeuteration of the protein.

The pulse scheme of this 4D NOESY-NOESY-TROSY is shown in Figure 1b. It represents a straightforward extension of the original NOE-NOE 3D experiment (Boelens et al., 1989), but with the detection period substituted by the gradient-enhanced 2D ^{1}H-^{15}N TROSY scheme (Pervushin et al., 1998). The latter enhances the attainable spectral resolution in the t_3 and t_4 dimensions, while dispersing the detected $^{1}H^{N}$ resonances in the ^{15}N dimension. A number of minor technical considerations are also relevant in this respect. (1) First, in order to maximize the number of (t_1, t_2, t_3) data points sampled, the phase cycling of the 4D experiment was reduced to 4 steps, and the observed spectral window was restricted to the region downfield of the H_2O resonance. To prevent bleeding in of several weaker imperfectly deuterated aliphatic or exchangeable resonances present in the upfield spectral region, selective EBURP2 and reverse-EBURP2 pulses (Geen and Freeman, 1991) were used to also restrict the regions where ^{1}H resonances were excited to those resonating downfield from the water resonance. As a result, no NOE peaks from a few amide protons resonating near water or upfield from water were observed. (2) Recording of a 4D NMR spectrum at adequate resolution requires the use of non-uniform sampling. High quality NUS reconstruction of a 4D NMR spectrum can be accomplished by the SMILE program (Ying et al., 2017) but this as well as most other NUS reconstruction software performs better if the various time domains are acquired in a manner that results in either a 0˚ or a 180˚ linear phase correction across the spectrum. For this purpose, and to ensure that the non-suppressed axial peaks can be optimally reconstructed, which requires 0˚ linear phase correction, it was preferable to insert a non-selective $90°_x207°_y90°_x$ composite ^{1}H inversion pulse (highlighted as the green open bar in Fig. 1b), followed by a second such pulse that reverses any phase imperfections introduced by the first composite pulse (Hwang et al., 1997). Specifically, the ϕ_1 phase cycling serves to eliminate axial peaks in the t_1 dimension caused by pulse imperfection as well as T_1 relaxation and amide exchange with solvent during T_{m1}, while also suppressing axial peaks in the t_2 dimension resulting from T_1 relaxation and water exchange during T_{m2}. To minimize the number of phase cycling steps, ϕ_2 was not phase cycled.

However, this resulted in small residual axial peaks along the F_2 dimension caused by pulse imperfections. To ensure that these residual axial peaks were absorptive in the final spectrum, thereby simplifying SMILE NUS reconstruction, an echo is generated by the application of two composite ^{1}H 180° pulses in order to suppress initial chemical shift evolution at t_2=0, thereby eliminating the need for a linear phase correction. Considering that the real and imaginary components of the residual axial signals have the same amplitude, they result in a 45º phase error for the axial peaks in the F_2 dimension. Shifting the ϕ_2 phase by -45° ensures that the NOE and axial peaks both can be phased absorptive using the same phase correction, thus facilitating NUS processing.

Compared to the 4D TROSY-NOESY-TROSY pulse scheme, the 4D NOESY-NOESY-TROSY experiment avoids the lossy magnetization transfer step from ^{1}H to ^{15}N and back (leading to a slightly larger number of 241 diagonal peaks on the ^{15}N-^{1}H (F_3,F_4) projected plane, compared to 231 for TROSY-NOESY-TROSY). Instead, its magnetization is simply transferred, in part, to its nearest neighbors by cross relaxation during the first NOE mixing period of duration T_{m1} = 50 ms. There is virtually no loss in total spin polarization summed over the initial "starting spin", whose t_1 evolution is monitored, and those of its immediate neighbors that are within cross-relaxation contact. As a result, the intrinsic sensitivity of such NOESY-NOESY-TROSY measurements is quite high, allowing the choice of a long duration of 300 ms for the second NOE mixing time, T_{m2}. During this second, much longer mixing time, the *z*-magnetization distributes over considerable distances due to indirect transfers (Fig. 3). Even in this extensively perdeuterated protein, NOEs to nearly a dozen neighboring protons are observed on the diagonals of the (F_1,F_2) cross sections, taken at the same (^{15}N, ^{1}H) frequencies used for illustrating the utility of the 4D TROSY-NOESY-TROSY spectrum of Figure 2. However, as pointed out by Boelens *et al.* and Breg *et al.* (Boelens et al., 1989;Breg et al., 1990), the NOE-NOE combination offers a wealth of new information on the cross-relaxation pathways that led to the long-distance NOEs, substantially aiding both the assignment and analysis of distance information. Below, we briefly highlight a few examples.

As expected, α-helical residue L232 shows intense cross peaks to both of its sequential neighbors, N231 and V233, as well as a weaker cross peak to F230. Despite the relatively short mixing time of only 50 ms that separates t_1 and t_2 evolution, the latter must result mostly from indirect

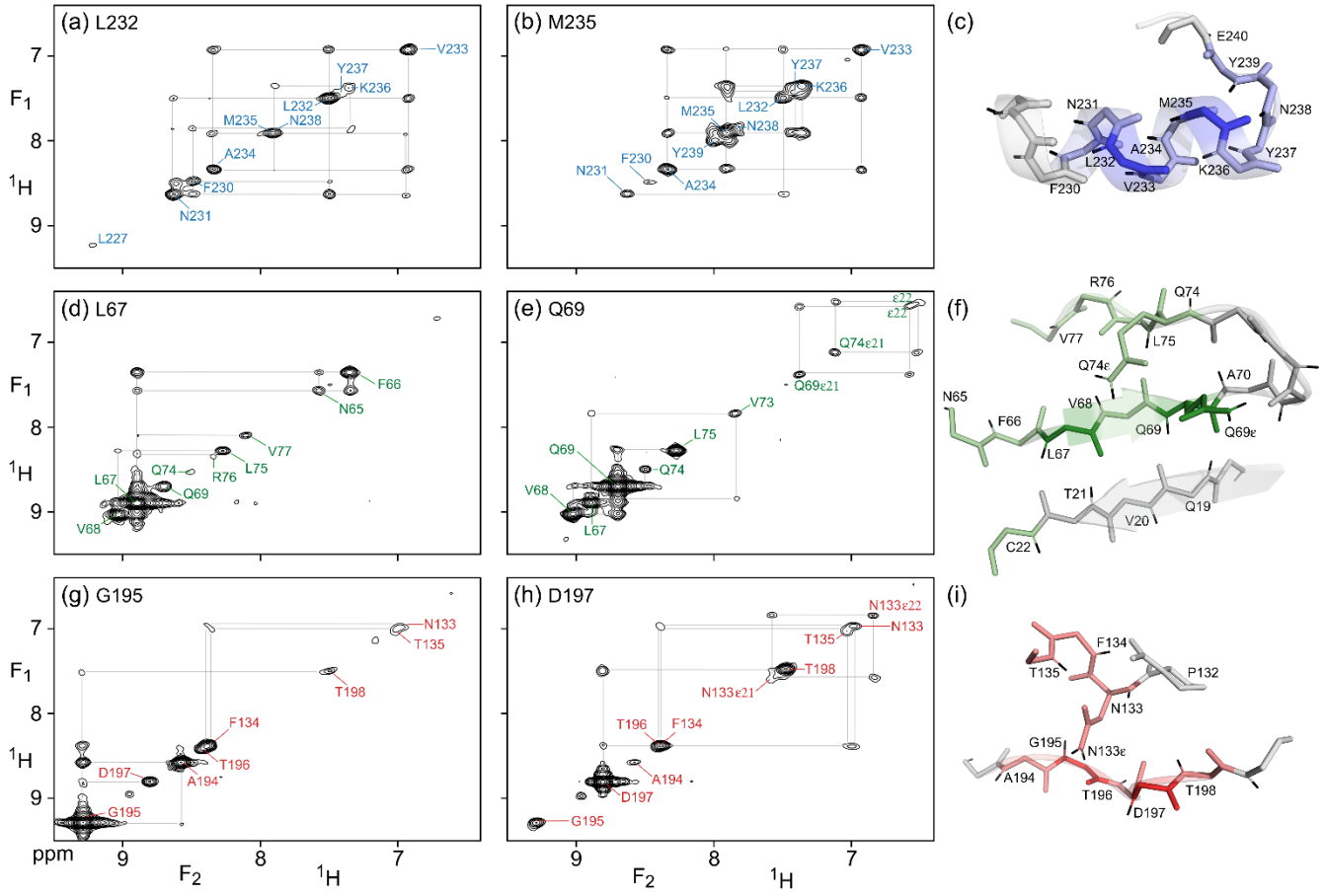

Figure 3. (F_1,F_2) cross sections through the 4D NOESY-NOESY-TROSY spectrum of M^{pro}_{C145A}, taken orthogonal to the (F_3,F_4) TROSY-HSQC plane at the ^{15}N,^{1}H frequencies of (a) L232, (b) M235, (d) L67, (e) Q69), (g) G195, and (h) D197. These cross sections show diagonal resonances for amide protons that are within long-range contact of the selected amide, either through direct or indirect NOE transfer during the two mixing periods that have a total duration of 50+300 ms. Off-diagonal resonances in these cross sections correspond to NOE magnetization transfer during the 50-ms mixing period that separates the t_1 and t_2 evolution periods. Colors match those of the corresponding residues in Figure 2. Expanded views of the structural elements (PDB: 6LU7) that gave rise to the observed NOEs are shown in panels (c), (f), and (i).

transfer through N231, because N231 and F230 share an intense cross peak. So in effect, each cross section through the 4D spectrum shown in Figure 3 corresponds to a 2D NOESY spectrum of a small, localized, region within the protein structure - making its analysis far simpler. For residues with few neighbors, direct NOE contacts between neighbors separated by as much as 4.5 Å give rise to quite intense cross peaks after 50 ms NOE mixing, as exemplified by the contacts between G195 and its A194 and T196 neighbors (Fig. 3g). A weaker cross peak between G195 and D197, at an interproton distance of 6.4 Å, appears not to be mediated by spin diffusion because

386 the G195 and D197 panels (Fig. 3g) show no common strong NOE to any visible resonance.
387 However, the possibility that the hydroxyl proton of T196 serves as a relay partner cannot be
388 excluded.

389 The NOESY-NOESY-TROSY spectrum also shows multiple NOEs to sidechain amide protons
390 that are not visible in the TROSY-NOESY-TROSY spectrum because the TROSY element does
391 not select magnetization transfer for NH_2 groups. For example, D197 shows long-range NOEs to
392 the N133 carboxamide protons, whereas Q69 shows NOEs to both its own carboxamide protons
393 and to those of Q74. The non-equivalent NH_2 pairs are readily recognized by cross-peak to
394 diagonal peak intensity ratios that are close to one, owing to their short interproton distance.

395

396 **4. Concluding remarks**

397 The spectra shown in this study were recorded during the summer of 2020, when access to campus
398 facilities was strongly restricted due to COVID-19 pandemic mitigation efforts. These restrictions
399 allowed for much lengthier acquisition of spectra than commonly used, for a total of eight days for
400 the two 4D spectra. As a benefit of NUS reconstruction, it is possible to generate spectra of the
401 same resolution recorded in any fraction of that time. Alternatively, we can discard the data
402 recorded at the longest values of t_1, t_2, and t_3. Indeed, processing the same time domain data sets
403 but shortening the time domains using a previously described protocol that considers the total
404 normalized length of the 3D (t_1,t_2,t_3) time domain vector (Ying et al., 2019), using only one third
405 of the acquired time domain data yields spectra that are very similar to the ones shown in Figures
406 2 and 3, albeit at slightly lower resolution and signal-to-noise, due to the use of three-fold less time
407 domain data. Nevertheless, the quality of the resulting spectra remains excellent, with near-
408 identical information content (Supporting Information Figures S2 and S3).

409 Use of the lengthy data acquisition times needed to collect the 4D spectra requires a high stability
410 sample, which in our case benefited from the C145A active site mutation, protecting the sample
411 from auto-proteolysis. As with all NMR experiments, S/N is approximately proportional to sample
412 concentration. Therefore, working at high concentrations benefits S/N of these experiments that
413 involve multiple magnetization transfer steps, an issue that is particularly important for NOE
414 experiments where magnetization from a single nucleus is distributed over many neighbors.

415 We note that the TROSY-NOESY-TROSY experiment used a long NOE mixing time of 200 ms,

such as to increase the number of observed connectivities by adding indirect NOE effects, including spin diffusion through hydroxyl protons (Koharudin et al., 2003), thereby aiding the assignment process. The use of a 50-ms NOE mixing period in the subsequent 4D NOESY-NOESY-TROSY experiment then provided a semi-quantitative measure of distance between these protons and their neighbors. Indeed, as pointed out by Kaptein and co-workers, recording of NOE-NOE spectra provides important experimental data on the pathway of magnetization transfer during NOE mixing. Such information could be used to convert this data into more quantitative distance information than the typical qualitative analysis of NOE intensities, potentially leading to the generation of higher resolution structures (Vogeli et al., 2009;Vogeli et al., 2012). Quantitative NOE interpretation traditionally relied on the recording of a series of NOE buildup data, which can become comparably time-consuming as the recording of 4D NMR spectra if resonance overlap is a limiting factor, as typically is the case for NOE spectra. This problem is further exacerbated by the spectral crowding of large proteins, particularly in the ^{1}H dimension, and while 3D spectra may give higher signal to noise ratios than 4D spectra, downstream analysis frequently requires extensive disambiguation of overlapped peaks. Our study of M^{pro}_{C145A} shows that a large number of semi-quantitative NOE distances become accessible by recording of 4D NMR spectra on a perdeuterated larger protein with little or no ambiguity about the nuclei involved.

While the high signal to noise and spectral simplicity of working with perdeuterated proteins has long been recognized (Torchia et al., 1988;Lemaster and Richards, 1988;Grzesiek et al., 1993;Tugarinov et al., 2004) the number of structural restraints accessible used to be small. Our present study demonstrates that a much larger number of NOE interactions becomes available by the recording of 4D NOE spectra. Moreover, it highlights the exquisite detail and value of NOE-NOE interaction analysis explored by the Kaptein group and it demonstrates that this approach is highly suitable for the larger biomolecules and biomolecular complexes being explored today, in particular when using extensive perdeuteration. Therefore, we believe that the recording of high quality 4D NMR spectra of the type presented in this study is entirely practical and invaluable for the structural and functional analysis of large proteins and their complexes, with possible extension to the study of nucleic acids. We note, however, that in the absence of extensive deuteration the dilution of nuclear magnetization over sidechain resonances will strongly lower the sensitivity of the experiment, which is further exacerbated by decreased effectiveness of TROSY-based line narrowing in such samples. On the other hand, adaptations of the NOESY-NOESY-TROSY

experiment to methyl-protonated but otherwise perdeuterated proteins (Tugarinov et al., 2005b) are expected to be readily feasible.

5. Data availability

The raw Bruker NMR data sets including the acquisition parameters and NUS sampling lists, pulse programs, include file, and NMRPipe processing scripts are available for download from zenodo.org: https://zenodo.org/record/4625615

6. Special issue statement

This article is part of the special issue "Robert Kaptein Festschrift". It is not associated with a conference.

7. Supplement

Supporting information related to this article is available online at: https://doi.org/xxxxxxxxxxx.

8. Author contributions

AJR expressed and purified protein samples, collected and analyzed the data, and edited the manuscript; JY optimized pulse sequence parameterization and processing and edited the manuscript; AB supervised the project and wrote the manuscript.

9. Competing interest

The authors declare that they have no conflict of interest.

10. Acknowledgements

We thank John M. Louis, Joseph Courtney, Yang Shen, James L. Baber and Dennis A. Torchia for helpful discussions. This work was supported by the Intramural Research Program of the NIDDK and by the Intramural Antiviral Target Program of the Office of the Director, NIH.

11. References

Altincekic, N.: Large-scale recombinant production of the SARS-CoV-2 proteome for high-throughput and structural biology applications, Frontiers in Molecular Biosciences, doi: 10.3389/fmolb.2021.653148 2021, 2021.

Barnes, C. A., Shen, Y., Ying, J. F., Takagi, Y., Torchia, D. A., Sellers, J. R., and Bax, A.: Remarkable Rigidity of the Single alpha-Helical Domain of Myosin-VI As Revealed by NMR Spectroscopy, J. Am. Chem. Soc., 141, 9004-9017, 10.1021/jacs.9b03116, 2019.

Boelens, R., Koning, T. M. G., and Kaptein, R.: Determination of Biomolecular Structures From Proton-Proton Noes Using a Relaxation Matrix Approach, J. Mol. Struct., 173, 299-311, 1988.

Boelens, R., Vuister, G. W., Koning, T. M. G., and Kaptein, R.: Observation of spin diffusion in biomolecules by 3-dimensional NOE-NOE spectroscopy, J. Am. Chem. Soc., 111, 8525-8526, 10.1021/ja00204a039, 1989.

Boras, B., Jones, R. M., Anson, B. J., Arenson, D., Aschenbrenner, L., Bakowski, M. A., Beutler, N., Binder, J., Chen, E., Eng, H., Hammond, H., Hammond, J., Haupt, R. E., Hoffman, R., Kadar, E. P., Kania, R., Kimoto, E., Kirkpatrick, M. G., Lanyon, L., Lendy, E. K., Lillis, J. R., Logue, J., Luthra, S. A., Ma, C., Mason, S. W., McGrath, M. E., Noell, S., Obach, R. S., O'Brien, M. N., O'Connor, R., Ogilvie, K., Owen, D., Pettersson, M., Reese, M. R., Rogers, T. F., Rossulek, M. I., Sathish, J. G., Shirai, N., Steppan, C., Ticehurst, M., Updyke, L. W., Weston, S., Zhu, Y., Wang, J., Chatterjee, A. K., Mesecar, A. D., Frieman, M. B., Anderson, A. S., and Allerton, C.: Discovery of a Novel Inhibitor of Coronavirus 3CL Protease for the Potential Treatment of COVID-19, bioRxiv, 2020.2009.2012.293498, 10.1101/2020.09.12.293498, 2021.

Breg, J. N., Boelens, R., Vuister, G. W., and Kaptein, R.: 3D NOE-NOE spectroscopy of proteins - Observation of sequential 3D NOE cross peaks in Arc repressor, J. Magn. Reson., 87, 646-651, 10.1016/0022-2364(90)90324-3, 1990.

Clore, G. M., Kay, L. E., Bax, A., and Gronenborn, A. M.: Four-dimensional 13C/13C-edited nuclear Overhausesr enhancement spectroscopy of a protein in solution: Application to interleukin 1β., Biochemistry, 30, 12-18, 1991.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A.: NMRpipe - a multidimensional spectral processing system based on Unix pipes, J. Biomol. NMR, 6, 277-293, 1995.

Diercks, T., Truffault, V., Coles, M., and Millett, O.: Diagonal-Free 3D/4D HN,HN-TROSY-NOESY-TROSY, J. Am. Chem. Soc., 132, 2138-+, 10.1021/ja910523q, 2010.

Douangamath, A., Fearon, D., Gehrtz, P., Krojer, T., Lukacik, P., Owen, C. D., Resnick, E., Strain-Damerell, C., Aimon, A., Abranyi-Balogh, P., Brandao-Neto, J., Carbery, A., Davison, G., Dias, A., Downes, T. D., Dunnett, L., Fairhead, M., Firth, J. D., Jones, S. P., Keeley, A., Keseru, G. M., Klein, H. F., Martin, M. P., Noble, M. E. M., O'Brien, P., Powell, A., Reddi, R. N., Skyner, R., Snee, M., Waring, M. J., Wild, C., London, N., von Delft, F., and Walsh, M. A.: Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease, Nature Communications, 11, 10.1038/s41467-020-18709-w, 2020.

Geen, H., and Freeman, R.: Band-selective radiofrequency pulses, J. Magn. Reson., 93, 93-141, 1991.

Grzesiek, S., Anglister, J., Ren, H., and Bax, A.: ^{13}C line narrowing by ^{2}H decoupling in $^{2}H/^{13}C/^{15}N$-enriched Proteins. Application to triple resonance 4D connectivity of sequential amides., Journal of American Chemical Society, 115, 4369-4370, 1993.

Grzesiek, S., and Bax, A.: The Importance of Not Saturating H_2O in Protein NMR. Application to Sensitivity Enhancement and NOE Measurement, J. Am. Chem. Soc., 115, 12593-12594, 1993.

Grzesiek, S., Wingfield, P., Stahl, S., Kaufman, J. D., and Bax, A.: Four-dimensional 15N-separated NOESY of slowly tumbling perdeuterated 15N-enriched proteins. Applications to HIV-1 Nef., J. Am. Chem. Soc., 117, 9594-9595, 1995.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Rio, J. F., Wiebe, M., Peterson, P., Gerard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E.: Array programming with NumPy, Nature, 585, 357-362, 10.1038/s41586-020-2649-2, 2020.

Helmus, J. J., and Jaroniec, C. P.: Nmrglue: an open source Python package for the analysis of multidimensional NMR data, J. Biomol. NMR, 55, 355-367, 10.1007/s10858-013-9718-x, 2013.

Hunter, J. D.: Matplotlib: A 2D graphics environment, Computing in Science & Engineering, 9, 90-95, 10.1109/mcse.2007.55, 2007.

Hwang, T. L., van Zijl, P. C. M., and Garwood, M.: Broadband adiabatic refocusing without phase distortion, J. Magn. Reson., 124, 250-254, 1997.

Ikura, M., Kay, L. E., and Bax, A.: A novel approach for sequential assignment of ^{1}H, ^{13}C, and ^{15}N spectra of larger proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. application to calmodulin, Biochemistry, 29, 4659-4667, 1990.

Kay, L. E., Clore, G. M., Bax, A., and Gronenborn, A. M.: Four-dimensional heteronuclear triple-resonance NMR spectroscopy of interleukin-1B in solution, Science, 249, 411-414, 1990.

Kay, L. E., Keifer, P., and Saarinen, T.: Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity, J. Am. Chem. Soc., 114, 10663-10665, 1992.

Koharudin, L. M. I., Bonvin, A., Kaptein, R., and Boelens, R.: Use of very long-distance NOEs in a fully deuterated protein: an approach for rapid protein fold determination, J. Magn. Reson., 163, 228-235, 10.1016/s1090-7807(03)00149-6, 2003.

Lee, W., Tonelli, M., and Markley, J. L.: NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy., Bioinformatics (Oxford, England), 31, 1325-1327, 10.1093/bioinformatics/btu830, 2015.

Lemaster, D. M., and Richards, F. M.: NMR sequential assignment of Escherichia-coli thioredoxin utilizing random fractional deuteration Biochemistry, 27, 142-150, 10.1021/bi00401a022, 1988.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M.: Overcoming the overlap problem in the assignment of ^{1}H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^{1}H-^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1β, Biochemistry, 28, 6150-6156, 1989a.

Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A.: Three-dimensional heteronuclear NMR of ^{15}N-labeled proteins, J. Am. Chem. Soc., 111, 1515-1517, 1989b.

Oschkinat, H., Griesinger, C., Kraulis, P. J., Sorensen, O. W., Ernst, R. R., Gronenborn, A. M., and Clore, G. M.: 3-Dimensional NMR spectroscopy of a protein in solution, Nature, 332, 374-376, 1988.

Pervushin, K., Riek, R., Wider, G., and Wuthrich, K.: Attenuated T2 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. USA, 94, 12366-12371, 1997.

569 Pervushin, K. V., Wider, G., and Wuthrich, K.: Single transition-to-single transition polarization
570 transfer (ST2-PT) in [N15, H1]-TROSY, J. Biomol. NMR, 12, 345-348, 1998.
571 Rovnyak, D., Frueh, D. P., Sastry, M., Sun, Z. Y. J., Stern, A. S., Hoch, J. C., and Wagner, G.:
572 Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling
573 and maximum entropy reconstruction, J. Magn. Reson., 170, 15-21, 10.1016/j.jmr.2004.05.016,
574 2004.
575 Torchia, D. A., Sparks, S. W., and Bax, A.: Delineation of Alpha-Helical Domains in
576 Deuteriated Staphylococcal Nuclease By 2d Noe Nmr-Spectroscopy, J. Am. Chem. Soc., 110,
577 2320-2321, 1988.
578 Tugarinov, V., Muhandiram, R., Ayed, A., and Kay, L. E.: Four-dimensional NMR spectroscopy
579 of a 723-residue protein: Chemical shift assignments and secondary structure of malate synthase
580 G, J. Am. Chem. Soc., 124, 10025-10035, 2002.
581 Tugarinov, V., Hwang, P. M., and Kay, L. E.: Nuclear magnetic resonance spectroscopy of high-
582 molecular-weight proteins, Annu. Rev. Biochem., 73, 107-146, 2004.
583 Tugarinov, V., Choy, W. Y., Orekhov, V. Y., and Kay, L. E.: Solution NMR-derived global fold
584 of a monomeric 82-kDa enzyme, Proc. Natl. Acad. Sci. U. S. A., 102, 622-627, 2005a.
585 Tugarinov, V., Kay, L. E., Ibraghimov, I., and Orekhov, V. Y.: High-resolution four-dimensional
586 H-1-C-13 NOE spectroscopy using methyl-TROSY, sparse data acquisition, and
587 multidimensional decomposition, J. Am. Chem. Soc., 127, 2767-2775, 10.1021/ja04032o,
588 2005b.
589 Vogeli, B., Segawa, T. F., Leitz, D., Sobol, A., Choutko, A., Trzesniak, D., van Gunsteren, W.,
590 and Riek, R.: Exact Distances and Internal Dynamics of Perdeuterated Ubiquitin from NOE
591 Buildups, J. Am. Chem. Soc., 131, 17215-17225, 10.1021/ja905366h, 2009.
592 Vogeli, B., Kazemi, S., Guntert, P., and Riek, R.: Spatial elucidation of motion in proteins by
593 ensemble-based structure calculation using exact NOEs, Nat. Struct. Mol. Biol., 19, 1053-1058,
594 10.1038/nsmb.2355, 2012.
595 Vuister, G. W., Boelens, R., and Kaptein, R.: Non-selective 3-dimensional NMR spectroscopy -
596 The 3D NOE-HOHAHA experiment, J. Magn. Reson., 80, 176-185, 1988.
597 Wagner, G.: Prospects for NMR of large proteins, J. Biomol. NMR, 3, 375-385, 1993.
598 Wüthrich, K.: NMR of Proteins and Nucleic Acids, John Wiley & Sons, New York, 1986.

Xia, Y. L., Sze, K. H., and Zhu, G.: Transverse relaxation optimized 3D and 4D N-15/N-15 separated NOESY experiments of N-15 labeled proteins, J. Biomol. NMR, 18, 261-268, 10.1023/a:1026590201757, 2000.

Ying, J., Delaglio, F., Torchia, D. A., and Bax, A.: Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data, J. Biomol. NMR, 68, 101-118, 2017.

Ying, J. F., Barnes, C. A., Louis, J. M., and Bax, A.: Importance of time-ordered non-uniform sampling of multidimensional NMR spectra of A beta(1-42) peptide under aggregating conditions, J. Biomol. NMR, 73, 429-441, 10.1007/s10858-019-00235-7, 2019.

Zuiderweg, E. R. P., and Fesik, S. W.: Heteronuclear Three-Dimensional NMR Spectroscopy of the Inflammatory Protein $C5_a$, Biochemistry, 28, 2387-2391, 1989.

Zuiderweg, E. R. P., Petros, A. M., Fesik, S. W., and Olejniczak, E. T.: 4-Dimensional [C-13, H-1, C-13, H-1] Hmqc-Noe-Hmqc Nmr-Spectroscopy - Resolving Tertiary Noe Distance Constraints in the Spectra of Larger Proteins, J. Am. Chem. Soc., 113, 370-372, 10.1021/ja00001a060, 1991.