ACS PHYSICAL CHEMISTRY Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Editors' Choice
Open Access

pubs.acs.org/physchemau
Article

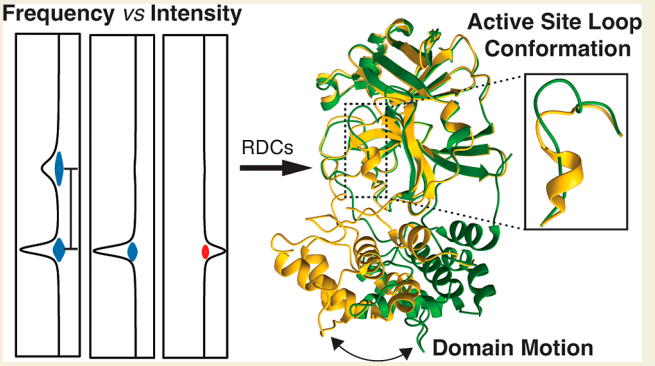
Solution Domain Dynamics of Monomeric SARS-CoV-2 Main Protease Revealed by Optimized NMR Residual Dipolar Coupling Measurements

Marshall J. Smith, Jinfa Ying, Yang Shen, John M. Louis, and Ad Bax*

Cite This: ACS Phys. Chem Au 2026, 6, 81−94
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

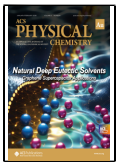
ABSTRACT: The main protease (MPro) of SARS-CoV-2, released from a monomeric precursor, is essential for viral replication, and multiple antiviral compounds target the active site of the mature dimer. However, the solution structure and dynamics of the monomer remain poorly understood. Here, we utilize residual dipolar couplings (RDCs) to characterize conformational and dynamical differences between monomeric and dimeric MPro. While most protein NMR studies rely primarily on $^{1}H-^{15}N$ RDCs, we demonstrate that at least four backbone RDCs can be measured at high accuracy from a single aligned sample. We compare frequency-based and intensity-based methods and introduce a mixed-time evolution scheme that improves resolution in 2D N−C′ RDC measurements. These methods are applied to a 9-residue N-terminal deletion mutant of the SARS-CoV-2 MPro, revealing that the monomeric active site loop conformation closely resembles that of the SARS-CoV monomer X-ray structure, which differs substantially from the dimer. The C-terminal helical domain also undergoes large amplitude motions relative to the catalytic domain. In contrast, AlphaFold2 models of the deletion mutant predict structures adopting only the dimer-like conformation. Refinement of the monomeric X-ray structure (PDB: 2QCY) against our experimental RDCs resulted in backbone rearrangements up to ca. 1 Å, improved MolProbity statistics, and better agreement with independent $^{2}D_{C'H}$ RDCs not included in the refinement. These findings highlight the power of RDCs for probing conformational states and dynamics and may aid future identification and characterization of compounds targeting the precursor monomer active site.



KEYWORDS: *domain dynamics, liquid crystal, protein NMR, structure refinement, 3CLpro, nsp5, Mpro monomer*

The Main Protease (MPro) of SARS-CoV-2 is critical for polyprotein processing and viral replication. Hence, MPro is of high interest for antiviral drug development, exemplified by the very successful covalent drug nirmatrelvir targeting this protease,[1] and by the more recently introduced noncovalent ensitrelvir.[2] However, the potential for drug-resistant mutants necessitates the exploration of alternative therapeutic strategies for targeting this key enzyme. Mature MPro exists as a homodimer, with each monomer containing three domains: domains 1 (residues 10−100) and 2 (residues 101−184) together form the catalytic domain with a chymotrypsin-like fold, while domain 3 (residues 201−306) is α-helical and is critical for dimerization.[3−5] Each monomer of MPro contains a complete catalytic dyad (H41 and C145) that is located in the cleft between domains 1 and 2. However, catalytic activity is very low in the monomeric state and the protein requires dimerization for full enzymatic activity.[6,7] Inactivity of the monomer is attributed to structural differences adjacent to the active site between the monomeric and dimeric forms of MPro, with the oxyanion loop (G138-C145) forming a 3_{10} helix in the monomeric state that is unwound in the homodimer.[5,6] Hence, the active site conformation of the monomer as well as the dimerization process are potential targets for future antiviral drug development, but designing such compounds requires detailed information on the solution structure and dynamics of this protease.[8,9] While monomerization of oligomeric proteins often can be obtained by the application of hydrostatic pressure,[10,11] for MPro we observed no shift to the monomeric state even at 3 kbar, presumably because of the absence of significant void volume at the dimer interface. To gather information on the monomeric state, we therefore perform RDC analyses on a monomeric mutant of MPro lacking its N-

Received: August 14, 2025
Revised: September 29, 2025
Accepted: September 30, 2025
Published: November 27, 2025



https://doi.org/10.1021/acsphyschemau.5c00081
ACS Phys. Chem Au 2026, 6, 81−94

terminal finger residues, S1−P9, that are required for dimerization, with further inactivation by the active site H41Q mutation (MPro$^{10-306,H41Q}$). The latter mutation restricts autoproteolysis, thereby prolonging the longevity of the protein. Attempts to crystalize this protein, like many other monomeric constructs of MPro, were unsuccessful. Hence, alternative strategies to obtain critical information on structure and dynamics were required.

Solution NMR spectroscopy can provide valuable insights into the structure and dynamics of biomolecules at the atomic level.[12] However, NMR relaxation measurements and traditional NMR restraints such as NOE distances, chemical shifts, and scalar couplings, typically only provide information on the local environment of a nucleus. In contrast, RDCs offer complementary global orientational information, making them a powerful complement for studying macromolecular structure and dynamics.[13−15]

Dipolar couplings arise from the interactions between magnetically active nuclei in an external magnetic field, $\boldsymbol{B}_0$. If the internuclear vector is parallel to the external magnetic field, the strength of the interaction is at a maximum, with a magnitude (D_{max}) given by

$$D_{max} = \frac{-\mu_0 h \gamma_i \gamma_j}{8\pi^3 r_{ij}^{\ 3}} \tag{1}$$

where μ is the vacuum permeability constant, h is Planck's constant, γ_x is the gyromagnetic ratio of nucleus x and, and r_{ij} is the distance between nuclei i and j.[13] In solution, averaging over the orientations of the internuclear vector with respect to the external magnetic field results in a reduction of the observed dipolar coupling (D)

$$D = D_{max}\left\langle \frac{3\cos^2(\Theta) - 1}{2} \right\rangle \tag{2}$$

where the angular brackets denote the time or population average and Θ is the angle of the internuclear vector with respect to the external magnetic field. Equation 2 can be rewritten as

$$D = D_a\left[(3\cos^2(\theta) - 1) + \frac{3}{2}Rh^*\cos(2\phi)\sin^2(\theta)\right] \tag{3}$$

where D_a describes the alignment order, Rh the rhombicity (asymmetry) of the diagonalized alignment tensor and θ and ϕ represent the spherical coordinates of the internuclear vector in the alignment frame.[13,16]

By inducing weak protein alignment, either by exploiting the magnetic susceptibility anisotropy of a protein,[16] by using a dilute aqueous liquid crystalline alignment medium[13] such as bacteriophage[17,18] or by immersing the protein in highly hydrated stretched polyacrylamide gels,[19,20] an anisotropic condition is imposed on the macromolecule and therefore the dipolar couplings no longer average to zero.[21] Typically, the deviation from a uniform orientational distribution is kept relatively low, corresponding to ca 0.1% alignment, such that only the large one-bond interactions remain visible without causing major line broadening from more distant ^{1}H−^{1}H interactions. For example, D_{max} for a protein backbone NH spin pair is ∼21.7 kHz, which by using weak alignment is reduced to fall in the ca ± 20 Hz range.[22] Achieving this weak protein alignment is often the most challenging step in an RDC study, as the medium must be compatible with the biomolecule under investigation, stable over a wide range of experimental conditions, and allow for a tunable degree of alignment. Once a suitable sample has been prepared, a host of NMR methods can be applied to measure different types of internuclear RDCs.[23]

Acquiring RDCs for multiple types of bond vectors is important when used for protein structure refinement. For example, if only amide ^{15}N−^{1}H RDC data are used, an incorrect structure can be forced to become compatible with those RDCs.[24] However, when measuring three types of RDCs, $^1D_{NH}$, $^1D_{C'N}$, and $^1D_{C'C\alpha}$, these couplings contain nearly independent information[25] and define the peptide plane orientation, albeit with 16-fold degeneracy.[26]

Unfortunately, only few studies have used three or more different types of RDCs,[21] despite their measurement at high precision being demonstrated for relatively large proteins such as maltose binding protein (MBP).[27,28] Reasons for this underutilization include the limited availability of pulse sequences in the spectrometer manufacturer's library, lack of dedicated software for extracting RDCs, and uncertainty regarding the best pulse sequences for specific cases.

RDCs manifest as an additional contribution to the scalar coupling, J, altering the frequency separation from J to $J + D$, where D is the RDC. Generally, methods for measuring RDCs can be divided into two categories: methods where the coupling is obtained from the frequency difference between multiplet components, and intensity-based methods where the coupling is encoded in the signal intensity. For small proteins, amide RDCs can be obtained directly from the frequency separation between the doublet components in a ^{1}H−^{15}N HSQC spectrum when acquired without ^{1}H decoupling during ^{15}N evolution. While the doubling of the number of resonances can result in increased signal overlap, recording such spectra in an interleaved in-phase and anti-phase (IPAP) manner[29] solves this problem by generating separate spectra for the upfield and downfield components of the amide doublet. However, for larger, slower tumbling proteins the upfield ^{15}N-{^{1}H} doublet relaxes rapidly, causing this component to be broad and weak, thereby reducing the accuracy of the extracted couplings. For such proteins, intensity-based methods often are preferable.

Intensity-based methods can be divided into two categories: quantitative J-correlation (QJ) and J-modulation.[30] J-modulation methods require multiple spectra with different dephasing delays, and the modulation pattern is best-fitted with a J-dependent analytical function.[30−32] QJ is essentially a truncated version of J-modulated methods, applicable to doublets, where only two spectra are acquired: one where J-dephasing during an interval Δ is adjusted to give maximum signal intensity, and a second spectrum where J-dephasing during the same Δ duration is altered to yield zero intensity for $\Delta = 1/J$, or analogous variations on this theme. QJ methods require less time than J-modulation methods, especially when applied to 3D triple resonance experiments, where J-modulation methods can result in impractical experiment times when many dephasing delays are required. On the other hand, QJ methods can yield systematic errors due to pulse imperfections or incomplete isotopic enrichment. Such effects are evident in deviations from pure sinusoidal intensity variation in J-modulated data sets[33] but are less easily identified in QJ methods. Here, we compare the merits and limitations of frequency- and intensity-based methods for measuring four types of protein backbone RDCs for a monomeric construct of the SARS-CoV-2 main protease, utilizing 2-dimensional (2D)

https://doi.org/10.1021/acsphyschemau.5c00081
ACS Phys. Chem Au 2026, 6, 81−94

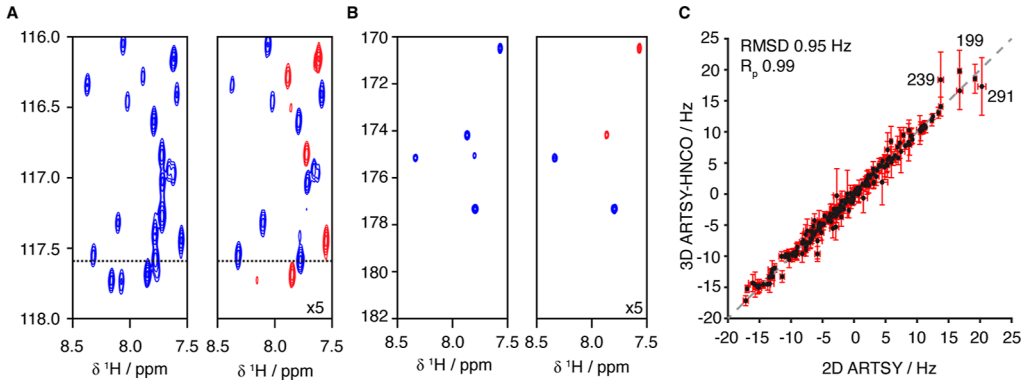


Figure 1. Comparison of 2D and 3D ARTSY $^1D_{NH}$ RDCs for MPro$^{10-306,H41Q}$. (A) Small regions of the reference (left) and attenuated (right) 2D $^1H-^{15}N$ ARTSY spectra, recorded at 800 MHz with a total measurement time of 5 h to acquire both. (B) $^1H/^{13}C'$ cross sections through the reference and attenuated 3D ARTSY-HNCO spectra (17 h), taken at $\delta(^{15}N)$ = 117.6 ppm (black dotted line in panel A). In A and B, the positive contours are shown in blue, negative contours in red. In the attenuated spectra, negative peaks correspond to amides with $J_{NH} + D_{NH} > 93$ Hz. Contours for the attenuated spectra in (A,B) are at 5-fold lower levels than for the reference spectra. (C) Correlation between RDCs that were measured with the 2D and 3D ARTSY methods. Outliers (e.g., T199, Y239 and F291) result from low S/N in the 3D ARTSY-HNCO experiment, as evidenced by the large error bars, derived from the S/N in the reference spectrum.[37]

and 3-dimensional (3D) techniques. We also provide tools for implementing these experiments.

METHODS

Sample Preparation

MPro$^{10-306,H41Q}$ was expressed analogous to the protocol of Aniana et al.[34] and outlined in Figure S1. $^{15}N/^{13}C/^2H$-enriched MPro$^{H41Q,10-306}$ was expressed in M9 medium containing U$-^{13}C/^2H$ glucose (2 g/L), $^{15}NH_4Cl$ (1 g/L), and 99% D_2O.

Recording and Analysis of NMR Data

Resonance assignments (Figure S14) were completed using standard TROSY-based triple resonance experiments,[35] in conjunction with a 4D $^1H-^{15}N$ TROSY-NOESY-TROSY[36] experiment (Figure S15). All assignment data were acquired at 800 MHz on a perdeuterated 0.5 mM sample in 25 mM HEPES buffer at pH 6.9 containing 20 mM NaCl, 1 mM TCEP, 3% D_2O and 0.02% NaN_3 at 15 °C. RDC data were collected at 25 °C using a 0.43 mM sample with 13 mg/mL Pf1 used as the alignment medium. Resonance assignments made at 15 °C were transferred to spectra at 25 °C using a series of 2D $^1H-^{15}N$ TROSY spectra acquired at temperatures over this range.

The $^1D_{NH}$ couplings were measured at 800 MHz using the 2D $^1H-^{15}N$ ARTSY[37] and 3D ARTSY-HNCO[38] methods, Figures S2 and S3. Recording of the 2D data required 5 h, whereas recording of the 3D ARTSY-HNCO data took 17.5 h. $^1D_{NC'}$ couplings were measured with QJ (constant-time and mixed-time, Figure S4) and with E. COSY type spectra, all at 800 MHz. Each had a total experiment time of 4 h, with the CT-QJ measurement having four times more scans for the attenuated spectrum compared to the reference spectrum. $^2D_{C'H}$ data were collected at 600 MHz utilizing the TROSY Anti-TROSY Encoded RDC (TATER) method[38] (Figure S5) with a total experiment time of 28 h for acquiring both the upfield and downfield doublet data.

$^1D_{C'C\alpha}$ data were acquired at 600 MHz using a 3D HNCO-based QJ experiment[39] (Figure S6) and by a 3D HNCO experiment without $^{13}C^\alpha$ decoupling during the $^{13}C'$ chemical shift evolution. The QJ experiment, including the reference and attenuated spectra, and the HNCO experiment without $^{13}C^\alpha$ decoupling, each required a total duration of 17.5 h. Each experiment utilized 22% nonuniform sampling.

All data were processed using NMRPipe,[40] while CCPN[41] v3.1 was used for resonance assignment and MATLAB scripts were used to extract scalar couplings and RDCs. To make RDC analysis more user-friendly and efficient, a CCPN macro was developed for extracting $^1D_{NH}$, $^1D_{NC'}$ and $^1D_{C'C\alpha}$ from QJ spectra and generating input files for the DC server.[40] The DC server was used for fitting the RDCs to various MPro X-ray structures. Further experimental details, data, and pulse sequence code are provided in the Supporting Information.

Structure Refinement Using RDCs

The X-ray crystal structure of monomeric MPro (PDB ID: 2QCY), an R298A mutant of the wild-type SARS-CoV MPro sequence,[42] was used as the starting model for structure refinement. Missing hydrogen atoms were added using the Dynamo package,[40] and the residues differing between SARS-CoV-1 and -2 were replaced using Rosetta.[43] Refinement by including the RDC information was carried out by simulated annealing and energy minimization calculations, performed in Cartesian space using the internal variable dynamics module (IVM)[44] of the XPLOR-NIH structure determination package.[45] Refinement closely followed the protocol described by Chou et al.[46] An ensemble of ten structures was generated for each set of calculations, with each having different initial random velocities at the start of the protocol. The target function included terms for experimental NMR restraints, covalent geometry, and nonbonded interactions.

The NMR restraints comprised 830 RDCs (496 for the catalytic domain and 334 for the helical domain), with final reference force constants of 1, 25, 9.18, and 64 kcal·mol^{-1}·Hz^{-2} for $^1D_{NH}$, $^1D_{NC'}$, $^1D_{C'C\alpha}$ and $^2D_{C'H}$ RDCs, respectively. Force constants for bond lengths, bond angles, and improper torsions were set to 1000 kcal·mol^{-1}·Å^{-2}, 500 kcal·mol^{-1}·rad^{-2}, and 500 kcal·mol^{-1}·rad^{-2}, respectively. Nonbonded interactions included a quartic van der Waals repulsion term (final force constant: 4 kcal·mol^{-1}·Å^{-2}), a torsion-angle database-derived potential of mean force (weighting factor: 1), a gyration volume restraint (final energy scale: 50 kcal·mol^{-1}·Å^{-2}), and a hydrogen bond database potential of mean force (HBPot; weighting factor: 2.5).[47] Finally, a PosDiffPot term with a small force constant of 1.5 kcal·mol^{-1}·Å^{-2} was included throughout to prevent large excursions from initial coordinates.

As the catalytic domain (domains 1 and 2) and the helical domain are flexibly linked and have different alignment tensors, separate refinements were performed. For the catalytic domain, residues in the oxyanion loop were excluded from the alignment tensor calculation. The high temperature simulated annealing protocol began at 3500 K and cooled to 25 K in 10 K decrements with relatively strong harmonic ϕ and ψ dihedral restraints applied, 300 kcal mol^{-1}·rad^{-2}, to retain the backbone torsion angles close to those of the starting structure. Such a protocol limits local structural changes but allows more distant rearrangements to occur through the accumulation of small torsion angle adjustments, if these changes lead to improved RDC agreement.[46] At this point the target dihedral restraints were reset to the current values, and a second low temperature refinement

https://doi.org/10.1021/acsphyschemau.5c00081
ACS Phys. Chem Au 2026, 6, 81−94

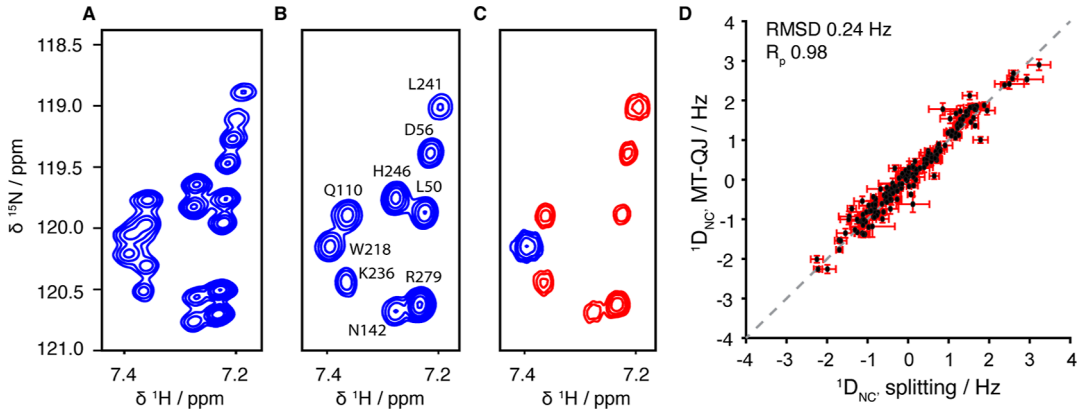


Figure 2. Small regions of the 800 MHz ^{1}H−^{15}N TROSY spectra of U−^{2}H/^{15}N/^{13}C MPro$^{10-306,H41Q}$. (A) 2D ^{1}H−^{15}N TROSY-HSQC spectrum, recorded in the absence of ^{13}C′ decoupling during t_1 evolution and ^{1}H^{N} detection, giving rise to E. COSY type doublets.[58] ^{13}C$^{\alpha}$ decoupling was used in both dimensions. (B,C) Reference and attenuated spectra, respectively, obtained with the MT-QJ $^{1}D_{NC'}$ pulse scheme. Contours in the attenuated spectrum are at a 5-fold lower level than the reference spectrum. $J_{NC'}$ dephasing was either absent during t_1 evolution (B), or active for 33.3 ms (C). (D) Comparison of the $^{1}D_{NC'}$ couplings extracted from the E. COSY and MT-QJ methods.

was performed, where the temperature was reduced from 25 to 1 K and the dihedral force constant was ramped down from 300 kcal mol^{-1} rad^{-2} to 50 kcal mol^{-1} rad^{-2}. All other force constants were the same as during the first round of simulated annealing. The final structure was generated by taking the regularized average of the ten ensemble members, utilizing an additional round of refinement with the same potential terms as the low temperature annealing step. The full refinement protocol is available at 10.5281/zenodo.16328473.

The RDC-refined structures for the catalytic and helical domains have been deposited in the PDB under the accession numbers 9PYT and 9PYS, respectively.

RESULTS AND DISCUSSION

Measurement of Amide ^{1}H−^{15}N RDCs

The Amide RDC by TROSY (ARTSY) method,[37] Figure S2, encodes RDCs into the signal intensity of ^{1}H−^{15}N TROSY-HSQC spectra (Figure 1A). Two interleaved data sets are recorded that yield reference and attenuated spectra, corresponding to the ^{1}H J_{NH} dephasing being active for half or throughout the full initial INEPT[48] period (extended to $1/J_{NH}$), respectively. The value of the doublet splitting ($\lambda_{NH} = J_{NH} + D_{NH}$) is then obtained from the ratio, $q = I_a/I_r$, between signal intensity in the attenuated (I_a) and reference (I_r) spectra

$$\lambda_{NH} = \frac{2}{\pi T}\cos^{-1}(q/2) \tag{4}$$

The line-narrowing effect provided by the TROSY component[49] results in excellent resolution and high sensitivity in ARTSY experiments. However, even at high magnetic fields, 2D ARTSY spectra for medium to large proteins can suffer from signal overlap, thereby decreasing the number of accurately measurable RDCs. The 3D ARTSY-HNCO method,[38] Figure S3, resolves this issue by spreading the signals across an additional ^{13}C dimension (Figure 1B), significantly improving resolution and providing additional RDCs. This comes at the cost of measurement time, with up to several days being required to achieve high-quality fully sampled 3D data sets. Fortunately, advances in nonuniform sampling (NUS)[50−54] have made these experiments more feasible, reducing experiment times to less than a day with little to no impact on the accuracy of the extracted RDCs, other than the reduced signal-to-noise (S/N) ratio associated with the shorter acquisition.[38,53,55] For example, Figure 1C shows that RDCs derived from 2D-ARTSY and from 3D ARTSY-HNCO data recorded with 21% NUS are in good agreement, with a root-mean-square difference (RMSD) of 0.95 Hz and a Pearson correlation coefficient (R_P) of 0.99. This RMSD is dominated by several residues that yield weak correlations in the HNCO spectrum, e.g. T199, Y239, and F291. Assuming the measurement errors are uncorrelated in the two data sets, the uncertainty in the individual measurements is ca 0.7 Hz which is far lower than the RMSD between observed RDCs and those predicted by a best-fit to any of the multitude of X-ray structures (vide infra). In contrast to the 2D ^{1}H−^{15}N ARTSY experiments where minimizing resonance overlap requires high magnetic field strength, 3D HNCO-based experiments benefit from medium magnetic field strengths (500−600 MHz) that offer a good compromise between resolution and sensitivity because ^{13}C′ transverse relaxation is dominated by chemical shift anisotropy and therefore scales with the square of the magnetic field.

Normalized for data collection time, the S/N·h$^{-1/2}$ of the 2D ARTSY was ca 6-fold higher than for 3D ARTSY-HNCO. This lower S/N results from several factors: (1) adding the amplitude-modulated ^{13}C′ dimension lowers S/N by $\sqrt{2}$; (2) relaxation losses during the ^{15}N-{^{13}C′} de- and rephasing delays of the HNCO pulse sequence are significant; (3) transverse relaxation of ^{13}C′ and apodization of this dimension to avoid truncation further decrease S/N. However, the 3D ARTSY-HNCO yielded 229 RDCs, vs 219 for the 2D ARTSY. These numbers are both substantially lower than the 283 values expected for this 297-residue protein (incl. twelve Pro residues and two N-terminal amides, S10 and G11, that are not observed due to fast hydrogen exchange), which is due to difficulties in protonating amide groups after expression in D_2O.[56]

Measurement of One-Bond ^{15}N−^{13}C′ RDCs

Due to the smaller product of the gyromagnetic ratios and longer internuclear distances, $^{1}D_{NC'}$ RDCs are approximately eight times smaller than $^{1}D_{NH}$ values, requiring greater precision in their measurement. A straightforward approach for measuring these couplings involves acquiring the ^{1}H−^{15}N TROSY-HSQC spectrum without ^{13}C′ decoupling during the ^{15}N evolution and ^{1}H detection periods. Correlations in the resulting spectrum exhibit an E. COSY pattern,[57] (Figure 2A),

https://doi.org/10.1021/acsphyschemau.5c00081
ACS Phys. Chem Au 2026, 6, 81−94

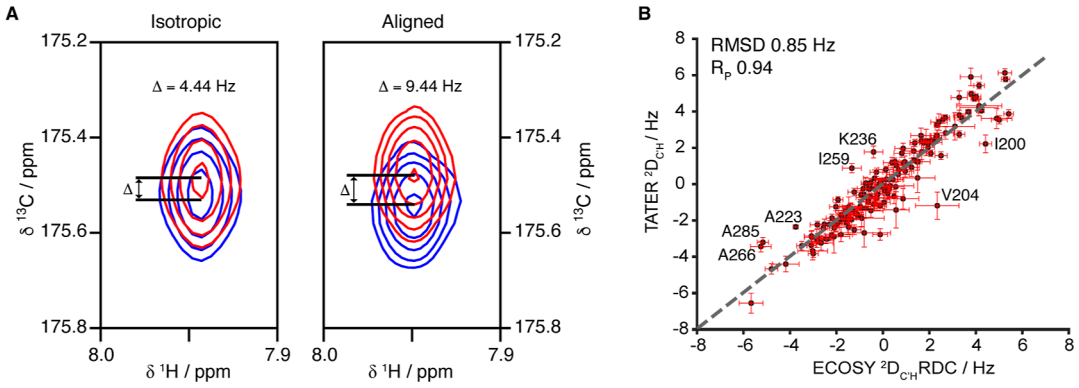


Figure 3. Measurement of $^{2}D_{C'H}$ RDCs. (A) Overlay of small, expanded regions of $^{1}H^{N}-^{13}C'$ cross sections taken at the ^{15}N shift of residue L286 through the TATER spectra. Overlayed signals from the TATER spectra are shown in blue and red, respectively. The frequency separation, Δ, in the $^{13}C'$ dimension provides the coupling. (B) Comparison of the $^{2}D_{C'H}$ couplings extracted from the 2D E. COSY spectrum shown in Figure 2A (4 h recording time) and those extracted from the 3D TATER spectra (18 h recording time), with S/N-based error bars in red. The gray dashed diagonal has unit slope.

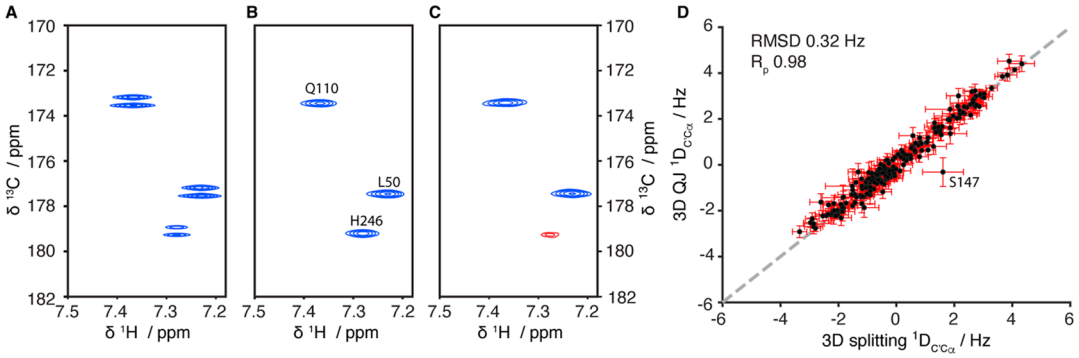


Figure 4. Measurement of $^{1}D_{C'C\alpha}$ RDCs in MPro$^{10-306,H41Q}$. (A) Small region of a cross-section through the 600 MHz HNCO spectrum, recorded without $^{13}C^{\alpha}$ decoupling during $^{13}C'$ chemical shift evolution. (B, C) QJ-HNCO spectra for the same region as in panel A, showing both the reference (B) and attenuated (C) spectra. Contours for the attenuated spectrum (C) are at 5-fold lower level than for the reference spectrum (B). Measurement time for the QJ and coupled HNCO experiments was 17 h each. (D) Comparison of the $^{1}D_{C'C\alpha}$ values extracted using the frequency-based and QJ methods. The gray dashed diagonal has unit slope. Red error bars represent the estimated uncertainties (see error analysis).

where the frequency separation in the ^{15}N dimension reveals the $^{15}N-^{13}C'$ coupling, and the frequency difference in the ^{1}H dimension corresponds to the two-bond $^{13}C'$-H^{N} coupling. This method allows both types of couplings to be extracted from a single spectrum, but for aligned proteins it typically requires perdeuteration of nonexchangeable hydrogens to reduce line broadening from unresolved RDCs to aliphatic protons.[58]

The 2-fold increase in the number of signals associated with the E. COSY doublets, and the difficulty in fully separating the two components in larger proteins, results in resonance overlap. Even partial resonance overlap raises the uncertainty in measured peak positions, thereby reducing the accuracy of extracted couplings. Intensity-based methods address this issue by encoding the $^{15}N-^{13}C'$ coupling in the resonance intensity, yielding a single cross peak per amide group. This results in reduced spectral overlap and simplifies analysis as peak positions are identical to those of the regular $^{1}H^{N}-^{15}N$ TROSY spectrum and are not impacted by the size of J couplings, allowing protein assignments to be directly transferred from TROSY spectra used during the resonance assignment process. The $^{1}J_{NC'}$ + $^{1}D_{NC'}$ couplings then are simply determined from corresponding ratios of peak intensities (Figure 2B,C).[33]

When carried out in a 2D fashion, the resolution of the QJ $^{15}N-^{13}C'$ pulse scheme is limited by the 65 ms duration of the constant-time (CT) ^{15}N evolution period, causing increased signal overlap (Figure S7) which limits spectral resolution and thereby the number and accuracy of extracted RDCs. Here, we introduce mixed-time (MT) chemical shift evolution (Figure S4),[59] where the CT evolution is proceeded by conventional chemical shift evolution once the CT evolution has been exhausted. The MT approach offers significantly improved resolution over the CT method, but at the expense of a ca. 40% decrease in $S/N{\cdot}h^{-1/2}$ due to the ^{15}N signal decay that follows CT ^{15}N evolution. The MT-QJ method yielded less resonance overlap than the E. COSY spectrum, resulting in 219 vs 207 RDCs. Comparison of these 207 RDCs with those of the MT QJ $^{15}N-^{13}C'$ pulse scheme shows excellent agreement (Figure 2D), even though the pairwise RMSD is dominated by the ca. 2-fold higher uncertainty of the E. COSY RDCs.

Measurement of Two-Bond $^{1}H^{N}-^{13}C'$ RDCs

Since, in contrast to the E. COSY spectrum of Figure 2A, only a single RDC is encoded in the QJ NC′ spectra, an additional experiment is needed to obtain the $^{2}D_{C'H}$ RDCs. These RDCs are only ca. 3-fold smaller than $^{1}D_{NH}$ interactions, with the higher ^{13}C magnetogyric ratio compared to ^{15}N partially compensating for the ca 2-fold larger internuclear distance. For measurement of two-bond $^{1}H^{N}-^{13}C'$ RDCs, the TROSY and

https://doi.org/10.1021/acsphyschemau.5c00081
ACS Phys. Chem Au 2026, 6, 81−94

ANTI-TROSY Encoded RDC (TATER) method was previously found to be very effective,[38] Figure S5. TATER is a frequency discrimination technique that records two interleaved 3D HNCO spectra where in one of the two experiments the $^{1}H^{N}$ spin state is temporarily flipped from $|\beta\rangle$ to $|\alpha\rangle$ during $^{13}C'$ evolution. This yields two TROSY-HNCO spectra where the resonances are displaced by $^{2}J_{C'H} + {}^{2}D_{C'H}$ in the $^{13}C'$ dimension. The $^{13}C'-^{1}H^{N}$ couplings are then determined from the frequency displacement in the $^{13}C'$ dimension (Figure 3A). The small ~4 Hz $^{2}J_{C'H}$ coupling limits the suitability of QJ methods because $^{2}J_{C'H}$ and $^{2}D_{C'H}$ are of comparable magnitude and QJ methods cannot determine the sign of $^{2}J_{C'H} + {}^{2}D_{C'H}$. For signals that are well resolved in both the 2D E. COSY and the 3D TATER spectra, the $^{2}D_{C'H}$ RDCs show good agreement, with $R_P \approx 0.94$ (Figure 3B). The uncertainties associated with $^{2}D_{C'H}$ RDCs measured from the E. COSY and TATER methods are comparable. However, the TATER method yielded 229 RDCs while only 207 reliable RDCs were obtained from the E. COSY method. While, in contrast to the E. COSY measurement, TATER-derived RDCs are not impacted by incomplete ^{13}C enrichment, imperfect ^{1}H inversion by the 180° pulses that flip the $^{1}H^{N}$ state from $|\beta\rangle$ to $|\alpha\rangle$ and back can result in smaller values for the measured couplings. Adiabatic hyperbolic tangent pulses[60] were therefore used for this purpose.

Measurement of $^{1}D_{C'C\alpha}$ RDCs

$^{1}D_{C'C\alpha}$ RDCs are inherently about five times weaker than $^{15}N-^{1}H^{N}$ RDCs, and they therefore only become useful if measured at high precision. $^{1}D_{C'C\alpha}$ RDCs couplings can be obtained by recording a 3D HNCO spectrum without $^{13}C^{\alpha}$ decoupling during $^{13}C'$ chemical shift evolution. The couplings are then extracted from the frequency separation between the two signals (Figure 4A). Although this method doubles the number of signals in the spectrum, signal overlap is less problematic due to the inherent high resolution of the 3D HNCO spectrum.

$^{1}D_{C'C\alpha}$ couplings can also be measured using intensity-based methods.[39] One approach is to encode the couplings into the signal intensity of a 2D $^{1}H-^{15}N$ TROSY spectrum by incorporating a $^{13}C'$ heteronuclear spin echo and varying the time point where the $^{13}C^{\alpha}$ 180° pulse is applied (Figure S6). Although these experiments offer a significant $S/N{\cdot}h^{-1/2}$ advantage over their 3D counterparts, the lower spectral resolution can result in extensive resonance overlap (Figure S8). Signal overlap is further exacerbated for $^{1}D_{C'C\alpha}$ measurements because these experiments are best suited for acquisition at medium field strengths due to the dominating effect of $^{13}C'$ chemical shift anisotropy (CSA) on transverse relaxation during the $3/(2J_{C'C\alpha}) \approx 28$ ms dephasing period, thereby reducing resolution of the 2D $^{1}H-^{15}N$ TROSY spectrum. Encoding the couplings into the signal intensity of a 3D HNCO experiment (Figure S6) reduces overlap and yields many more RDCs, but with a significant decrease in sensitivity (Figure 4). Nevertheless, with a RMSD of 0.32 Hz and a R_P = 0.98, RDCs obtained from the 3D frequency- and intensity-based methods are in excellent agreement (Figure 4D). The frequency-based method has errors similar to those in the intensity-based method, as shown by the error bars in Figure 4D. The S/N of the frequency-based experiment is higher than for the QJ method (294 vs 84), due to relaxation losses during the constant-time chemical shift evolution of the QJ method. However, the jackknifed Q_{JK} values[61] obtained from best fitting the RDCs to the X-ray structure are similar for the two methods (vide infra).

Precision of RDCs

The precision of a coupling derived from QJ measurements is proportional to the experimental uncertainty, σ, in the intensity ratio

$$\sigma = q^{*}\sqrt{\left(\frac{N}{S}\right)_{R}^{2} + \left(\frac{N}{S}\right)_{A}^{2}} \tag{5}$$

where q is the signal attenuation and $(N/S)_R$ and $(N/S)_A$ denote the inverse signal-to-noise ratio of the reference and attenuated spectra, respectively. Standard propagation of errors yields the error in the measured RDC

$$\sigma_{\lambda} = \sigma\frac{\partial}{\partial q}\lambda(q) \tag{6}$$

where $\lambda(q)$ is the function used to extract the measured couplings. For instance, the $^{1}J_{C'C\alpha}$ and $^{1}D_{C'C\alpha}$ values in this work (Table S5) were obtained by using eq 7, where $2T = 28.4$ ms $\approx 3/(2^{1}J_{C'C\alpha})$ is the $^{1}J_{C'C\alpha} + {}^{1}D_{C'C\alpha}$ dephasing delay

$$\lambda = \frac{\cos^{-1}(q)}{2\pi T} + \frac{1}{2T} \tag{7}$$

and the error is approximated by

$$\sigma_{\lambda} = \left(\frac{q}{2\pi T\sqrt{1-q^{2}}}\right)\sqrt{\left(\frac{N}{S}\right)_{R}^{2} + \left(\frac{N}{S}\right)_{A}^{2}} \tag{8}$$

To minimize σ per unit of measurement time, we propose to acquire the attenuated spectrum with more scans than the corresponding reference spectrum. Since the attenuated spectrum is near a null in QJ experiments, the error in the intensity ratio and consequently the error in the total coupling, $^{1}J_{C'C\alpha} + {}^{1}D_{C'C\alpha}$, is dominated by $(S/N)_A$. By acquiring the attenuated spectrum with more scans, the error in the extracted coupling decreases with the inverse square root of the number of scans, until the S/N of the reference spectrum becomes a limiting factor. For a ratio, β, between the number of transients for the attenuated and reference spectra, improvements that approach $\sqrt{[(2\beta)/(\beta+1)]}$ per unit of total measurement time can be obtained for $q = 0$. For $q \neq 0$, $(S/N)_R$ can become the limiting factor and in practice a value of $\beta \approx 4$ is a good compromise, yielding a ca. 1.25-fold improvement in the precision of $^{1}D_{C'C\alpha}$ per unit of measuring time (Figure S9).

As highlighted by Liu and Prestegard,[33] the amplitude of the sinusoidal J modulation, and thereby the extracted coupling, is impacted by multiple factors that can lead to systematic errors in couplings when just using two dephasing time points, 0 and $2T$. These sources of error include incomplete isotope labeling, relaxation-induced errors, pulse imperfections, and the effects of passive couplings. Fortunately, when the proportion of incomplete labeling is relatively small, these factors to a good approximation impact all couplings measured in a given experiment by the same factor. Therefore, scaling factors which relate different types of measured couplings to their true value can be extracted from a fit of the RDCs to the coordinates of the corresponding bond vectors. This consideration further emphasizes the advantage of obtaining different types of RDCs, because systematic scaling errors are readily identified when all RDCs are simultaneously fit to an X-ray structure. A scaling

https://doi.org/10.1021/acsphyschemau.5c00081
ACS Phys. Chem Au 2026, 6, 81−94

Table 1. Jackknifed Q_{JK} Values[a] for Fits of Various RDC Types to the Invariant Residues of MPro's Three Domains[b]

	domain 1		domain 2		domain 3	
	quantitative *J*-correlation	frequency based	quantitative *J*-correlation	frequency based	quantitative *J*-correlation	frequency based
$^1D_{NC'}$	0.27	0.30	0.33	0.32	0.24	0.21
$^1D_{C'Ca}$	0.18	0.22	0.22	0.27	0.21	0.18
$^1D_{NH}$	0.21[c]		0.30[c]		0.20[c]	
$^2D_{C'H}$		0.13[d]		0.35[d]		0.21[d]

[a]Uncertainties in Q_{JK} values are ±0.03. [b]Fits to vector orientations averaged over those in X-ray structures of monomeric MPro (PDB entry 2QCY), a monomer subunit of the active homodimer (7K3T), and the average vector orientation of five AlphFold2 predicted structure of MPro$^{10-306,H41Q}$ with structures of >30% sequence identity removed from the AlphaFold database. [c]Obtained from 3D ARTSY-HNCO. [d]Obtained from 3D TATER-HNCO.

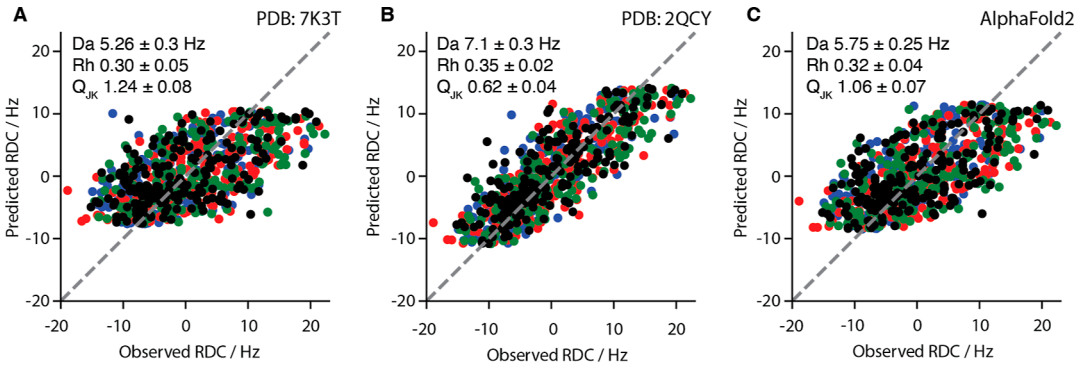


Figure 5. Simultaneous fit of the normalized $^1D_{NH}$, $^1D_{C'C\alpha}$, $^2D_{C'H}$ and $^1D_{NC'}$, represented by blue, red, green and black data points, respectively, to (A) the monomeric subunit of the homodimer (PDB: 7K3T), (B) the MPro^{R298A} SARS-CoV monomeric mutant (PDB: 2QCY) and (C) an ensemble of five AlphaFold2 predicted structures with proteins with >30% sequence identity removed from the AlphaFold database. $^1D_{NH}$, $^1D_{NC'}$, $^1D_{C'C\alpha}$ and $^2D_{C'H}$ couplings were obtained from ARTSY-HNCO, MT-$QJ_{NC'}$, $QJ_{C'C\alpha}$ and TATER experiments, respectively.

factor that differs substantially from unity may point to a problem with incomplete isotopic enrichment or unusually large pulse imperfections.

Frequency based methods, to their credit, often provide accurate couplings even in cases of incomplete labeling. However, when resolution is limited, the estimated RDC obtained from the frequency separation is often underestimated.[62−64] In the absence of significant overlap between doublet components, the uncertainty in the RDCs obtained by frequency based methods is estimated by[65]

$$\sigma \approx \frac{\mathrm{LW}}{\mathrm{S/N}} \tag{9}$$

where LW is the line width of the broadest doublet component, and S/N refers to the lowest signal-to-noise ratio observed for any of the doublet components in either the isotropic or aligned sample.

RDC Fits to Invariant Domains

The quality of RDC fits to atomic coordinates is adversely impacted by dynamics that are invisible in a single, static structure representation.[66] As previously demonstrated for MPro, when the angular excursions associated with such dynamics are small, this decrease in fit quality can be mitigated by fitting the RDCs to an ensemble of X-ray structures as shown in Figure S10.[61] Alternatively, the RDCs can be used to generate dynamic ensemble representations of structures that then reflect macromolecular motions.[15,67,68]

While the above averaging of bond vector orientations over multiple structures improves RDC fit quality in well-ordered regions of a protein, this approach becomes inadequate when structural differences are large as the nonlinearity of RDC dependence on orientation becomes significant with increasing fluctuations. Moreover, a limited number of structural models generally will represent a skewed representation of the solution state. Therefore, for the purpose of evaluating which RDC measurement method provides the most accurate results, residues that undergo large amplitude motions must be excluded. Furthermore, previous monomeric MPro RDC measurements indicated that the alignment strength observed for α-helices in its C-terminal domain were uniformly stronger than for the single α-helix in domain 1, indicative of interdomain motions.[69]

Therefore, we separate the RDC analysis into the three structural domains of MPro. Next, "well-defined peptide planes" are identified based on two criteria: (1) the backbone heavy atom positions differ by less than 0.6 Å across a best-fit superposition of a three-member structural ensemble comprising of an AlphaFold2 model of MPro$^{10-306,H41Q}$ (AF2-MPro), and the X-ray structures of monomeric (PDB 2QCY)[42] and homodimeric (PDB 7K3T)[70] MPro, and (2) the average angular differences in the N−H, N−C′ and C′-C$^{\alpha}$ vectors are less than 4°. Based on these criteria, we identified 60 well-defined peptide planes in domain 1, 53 in domain 2, and 83 in domain 3 (Figure S11). Subsequently fitting the RDC data for the identified residues to the average vector orientation of the three-member ensemble, resulted in a better fit than any individual X-ray structure or pairwise combination thereof (Table S1). Interestingly, the AF2-MPro structure yielded Q_{JK} values lower than those of any individual X-ray structure and matched the fit quality of the three-member ensemble.

RDC fit quality is represented by a jackknifed quality factor, Q_{JK}. The uncertainty in Q_{JK} is given by $Q_{JK}/\sqrt{N}$, where N is the number of RDCs. The value of Q_{JK} increases with both

https://doi.org/10.1021/acsphyschemau.5c00081
ACS Phys. Chem Au 2026, 6, 81−94

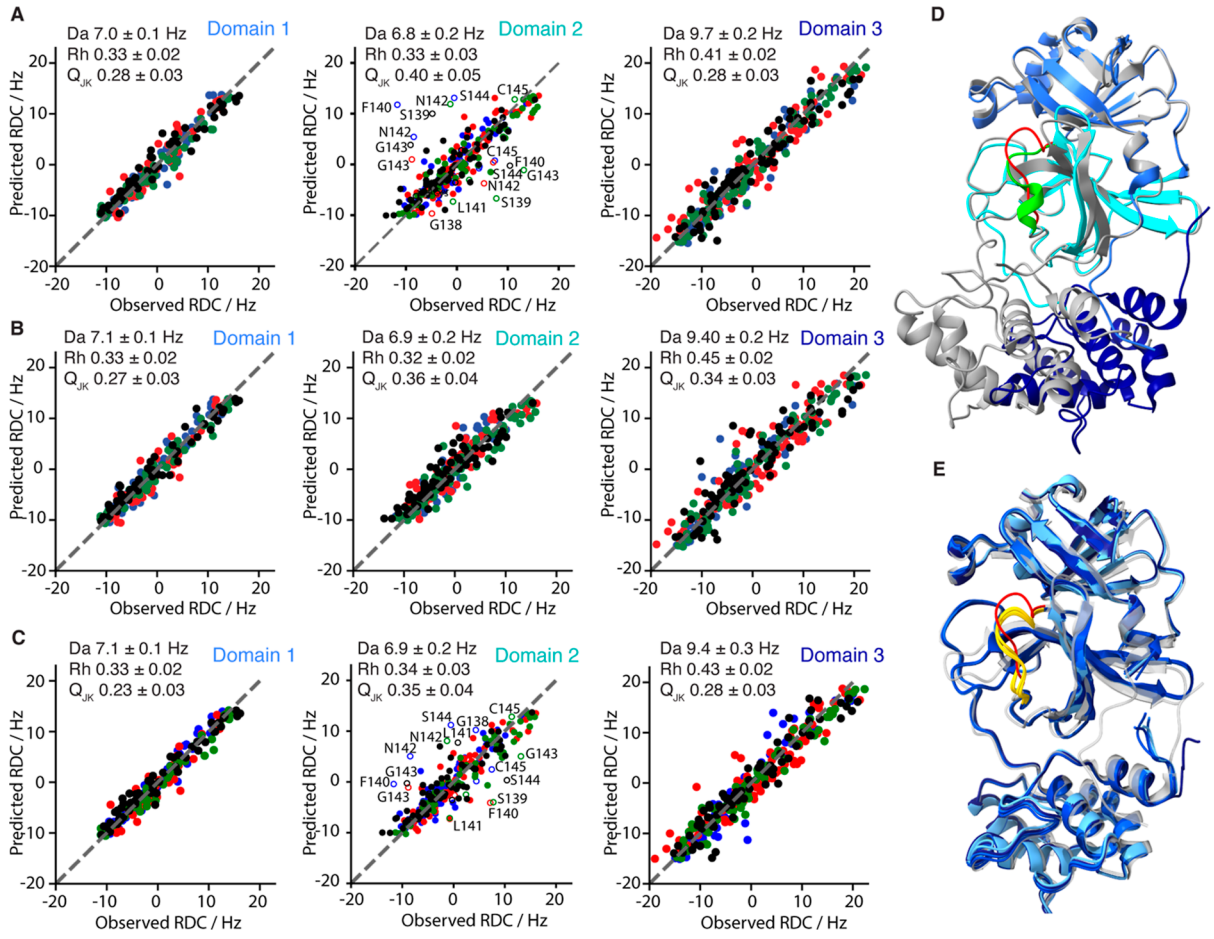


Figure 6. Simultaneous SVD fits of $^{1}D_{NH}$ (blue), $^{1}D_{NC'}$ (red), $^{1}D_{C'C\alpha}$ (green) and $^{2}D_{C'H}$ (black) to domain 1 (S10−K100), domain 2 (Y101−P184), and domain 3 (T201-C300). RDC values were normalized to the $^{1}D_{NH}$ interaction strength, and all were given the same weight in the SVD fit. (A) Fit to domains of the native homodimeric X-ray structure (PDB: 7K3T). (B) Corresponding fits to the X-ray structure of the R298A monomer (2QCY) and (C) the average vector orientation of five AlphaFold predicted models where structures with >30% sequence similarity were removed from the training set. Panels include alignment strength (D_a), rhombicity (Rh), and jackknifed quality factor (Q_{JK}), obtained for each fit, after exclusion of labeled oxyanion loop RDCs. (D) Monomeric subunit taken from the homodimeric X-ray structure (with domains 1, 2, and 3 shown in cornflower blue, cyan and navy, respectively) superimposed on the R298A monomer structure (gray) to yield best alignment for domains 1 and 2 (top half of figure). Residues of the homodimer oxyanion loop (G138-C145, red) differ from those in the monomer (green). (E) Overlay of five AlphaFold2 models (blue) and the 7K3T (gray outline) X-ray structure, with the oxyanion loop residues highlighted in yellow and red, respectively. The AlphaFold predicted models are available at 10.5281/zenodo.16328473. The $^{1}D_{NH}$, $^{1}D_{NC'}$, and $^{1}D_{C'C\alpha}$ RDCs were extracted from 3D ARTSY-HNCO, MT-QJ$_{NC'}$ and 3D QJ-HNCO spectra, respectively, whereas $^{2}D_{C'H}$ RDCs were obtained with the 3D TATER pulse sequence.

measurement error in the RDCs and uncertainty in the atomic coordinates. The latter often dominates but can be reduced by identifying well-defined peptide planes and averaging their bond vector orientations over the available experimental structures. For MPro$^{10-306,H41Q}$, fits of the different types of RDCs indeed show up to a 2-fold decrease in Q_{JK} when comparing well-defined peptide planes averaged over AF2-MPro, 2QCY, and 7K3T with fits of all RDCs to domains of the best-fit single X-ray structure (see Table 1 and Table S2).

For $^{1}D_{NC'}$ and $^{1}D_{C'C\alpha}$, comparisons of the Q_{JK} values obtained with QJ and frequency-based methods shows comparable quality fits, with best fits observed for the structurally most conserved C-terminal domain (Table 1). For all four types of RDCs, however, the RMSD between observed and best-fitted RDCs exceeds the estimated precision of the measured values. This result indicates that the fit qualities are dominated by uncertainties in the coordinates of the fitted structures, a conclusion supported by the improved fit quality when vector orientations of different models are averaged prior to fitting. True differences between the average structure in solution and in the crystalline state also are expected to decrease the quality of the RDC fit but these differences are likely to be reduced by our selection of "well-defined peptide planes".

Insights into MPro10−306, H41Q Structure and Dynamics

Fits of all $^{1}D_{NH}$, $^{1}D_{NC'}$, $^{2}D_{C'H}$ and $^{1}D_{C'C\alpha}$ couplings derived from the QJ and TATER methods (Tables S3−S6) to the full structure of either the monomeric subunit of the native dimer (PDB entry 7K3T)[70] (Figure 5A) or the monomeric R298A mutant of SARS-CoV MPro (2QCY)[42] (Figure 5B) both yield very poor fits, confirming our earlier analysis that the solution structure of the monomer differs from these X-ray structures.[69] Perhaps not surprising, predictions of the MPro$^{10-306,H41Q}$ mutant made by AlphaFold2[71] that include X-ray structures of SARS-CoV-2 MPro and its homologues from the SARS and MERS viruses predicted a structure that is close to 7K3T and that fit as poorly as 7K3T. When removing any proteins with >30% sequence identity to MPro from the AlphaFold2 training

https://doi.org/10.1021/acsphyschemau.5c00081
ACS Phys. Chem Au 2026, 6, 81−94

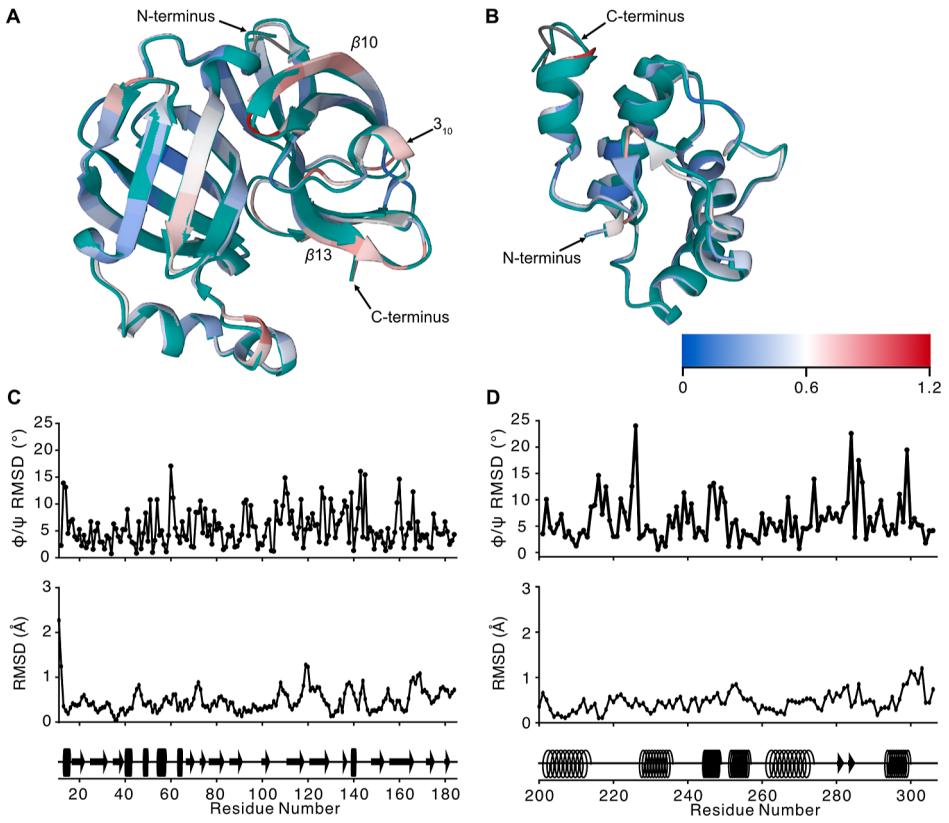


Figure 7. XPLOR-NIH RDC-refined structures of MPro$^{10-306,H41Q}$. (A) Superposition of the catalytic domain from the RDC-refined structure (blue-red) with the reference crystal structure (2QCY, green). (B) Equivalent comparison for the helical domain. The color scale indicates the per-residue backbone heavy atom RMSD between the initial and refined structures, ranging from 0 Å (blue) to 1.2 Å (red). (C,D) Per-residue RMSD in ϕ/ϕ backbone torsion angles ($\sqrt{[(\Delta\phi^2 + \Delta\psi^2)/2]}$) (top) and backbone heavy atom coordinates (middle), and secondary structure annotation (bottom) for the catalytic (C) and helical (D) domains. The secondary structure key uses arrows to represent β-sheets and coils to depict α-helices. The disordered C-terminal tail residues were excluded from the structure refinement.

set, the five models it generated were more heterogeneous (vide infra) but still resembled the dimeric crystal structure with a backbone heavy atom RMSD of 1.45 Å compared to 7K3T and, consequently, a poor fit to the RDC data (Figure 5C).

With backbone heavy atom coordinate RMSD values <1 Å, structures of domains 1 and 3 of the R298A monomeric mutant (PDB: 2QCY) are very similar to those of the native dimer (PDB: 7K3T) and AF2-MPro (Figure 6A−C). Simultaneous RDC fits for all four types of couplings, normalized to the $^{15}N-^{1}H$ dipolar interaction strength, show excellent agreement for Domain 1 (S10−K100) and Domain 3 (T201-C300) across all three structures. In contrast, RDCs of Domain 2 agree best with the X-ray structure of the R298A monomer, which exhibits a conformational rearrangement of the oxyanion loop residues (G138-C145 shown in green and red for the monomer and dimer, respectively, in Figure 6D) from a loop conformation in the active homodimer to a 3_{10} helix in the monomer. A similar rearrangement was reported for SARS-CoV-2 MPro lacking the helical domain (MPro$^{1-199}$; PDB entry 7UJ9), Figure S12.

Indeed, outliers in the RDC fits to domain 2 of the 7K3T homodimer correspond to residues in this oxyanion loop (Figure 6A). Removal of these RDCs results in a substantial Q_{JK} decrease from 0.69 to 0.40, confirming that the conformation of the oxyanion loop is largely responsible for the poor fit.

To within their jackknifed uncertainties, the alignment tensors for domains 1 and 2 are essentially the same, both in terms of D_a and Rh values, whereas the tensor for domain 3 is markedly different. This indicates that domains 1 and 2 undergo large-amplitude concerted motions relative to domain 3, which explains the poor RDC fits to any static X-ray structure when simultaneously fitting to all three domains (Figure 5). Comparison of the monomer and dimer X-ray structures shows that the helical domain in the R298A monomeric mutant has an orientation relative to the catalytic domain that differs significantly from that in the homodimer (Figure 6D), but neither is compatible with the RDCs (Figure 5A,B).

AlphaFold2 predictions on MPro$^{10-306,\ H41Q}$, generated after excluding training structures with >30% sequence similarity, showed slight deviations in the oxyanion loop residues compared to 7K3T but fail to reproduce the 3_{10}-helix conformation (Figure 6E). RDC fits to the average of five AlphaFold-predicted structures again yield excellent agreement for domains 1 and 3 with Q_{JK} values of 0.23 and 0.28, respectively (Figure 6C). However, domain 2 showed a poor fit with a Q_{JK} value of 0.49 that decreases to 0.35 upon exclusion of the oxyanion loop RDCs. Furthermore, AlphaFold2 does not predict any change in the orientation of the C-terminal helical domain relative to the catalytic domain, with all five predicted models having the same orientation as the dimer (Figure 6E).

https://doi.org/10.1021/acsphyschemau.5c00081
ACS Phys. Chem Au 2026, 6, 81−94

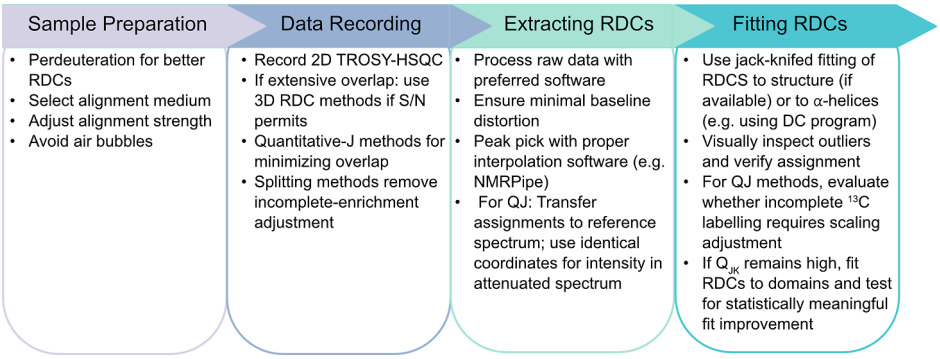


Figure 8. Roadmap summarizing key considerations for utilizing RDCs to investigate the structure and dynamics of medium sized proteins.

RDC Refinement of Monomeric X-ray Structure

To further demonstrate the benefits of acquiring multiple sets of RDCs and to overcome the challenge posed by the elusive crystallization of MPro$^{10-306,H41Q}$, the X-ray structure of its monomeric R298A mutant (2QCY)[42] was refined utilizing RDC restraints. Due to domains 1 and 2 (together forming the catalytic domain) having a different alignment tensor from the helical domain, the refinement procedure was performed separately for the two sections of the protein. The refined structures are shown in Figure 7A,B, colored according to the backbone heavy atom RMSD relative to the 2QCY crystal structure (shown in green).

Deviations that are less than 0.6 Å are colored in blue and those greater than 0.6 Å are shown in red. With average RMSDs of 0.49 Å and 0.57 Å across all residues, the refined structures remain close to those determined by X-ray crystallography. Excellent agreement with the four types of RDCs is achieved following refinement, with Q_{JK} values of 0.17 and 0.16 for the catalytic and helical domains, respectively. Structural quality as exemplified by MolProbity scores[72] improved from 1.99 to 1.26 (76th to 99th percentile) for the catalytic domain, whereas remaining essentially unchanged (1.21 vs 1.25; 99th vs 99th percentile) for the C-terminal domain. To confirm the validity of the refined models, structural refinement was repeated with the $^2D_{C'H}$ RDCs excluded. The resulting structures for the catalytic and helical domains had backbone RMSDs of 0.32 Å and 0.29 Å, respectively, relative to the fully refined domains, indicating high structural consistency. The Q_{JK} values for the excluded $^2D_{C'H}$ RDCs were 0.20 and 0.25 (versus 0.35 and 0.38 for 2QCY), demonstrating good cross-validation of the refined solution state structures.

The per residue RMSD of the backbone torsion angles and heavy atom coordinates of the refined structures relative to the starting structure show that changes are fairly small (Figure 7C,D) for both the catalytic and the helical domain. Unsurprisingly, the N- and C-terminal residues of the protein show the largest deviations. For clarity, these regions, colored in gray, are not included in the heat map (Figure 7A,B). Residues corresponding to the oxyanion loop show ∼0.5 Å RMSD between the initial and refined structures and up to 15° angular RMSD, which may more accurately reflect the solution state structure of the oxyanion loop. Interestingly, β10, a proline containing β-strand spanning residues S121 to C129, appears kinked with a relatively large RMSD at G124. Comparing the monomeric 2QCY and dimeric 7K3T structures (Figure S13), also shows substantial distance and angular RMSDs for this region. The helical domain remains very well conserved with the largest RMSDs observed in loop regions, Figure 7B,D. Only the C-terminal helix, which contains the R298A mutation in the 2QCY reference structure, shows a relatively large distance and angular RMSD; moderate distance and angular RMSDs are observed in the linker connecting the two C-terminal helices.

CONCLUDING REMARKS

Collection and analysis of RDC data for the investigation of the structure and motions of proteins involves multiple considerations during the four stages of such a study: (1) preparation of a suitably aligned sample; (2) selection of the most suitable pulse sequences; (3) extracting the RDCs; and (4) fitting of RDCs to coordinates (Figure 8). Selection of a suitable alignment medium is a key consideration, with commercially available filamentous Pf1 being the most widely used,[17] but anisotropically compressed acrylamide gel[19,20,73] and polyethylene glycol nematic liquid crystal[74] also being widely applicable. In our study, we perdeuterated the protein, which enables the use of ca 2-fold stronger alignment, owing to the reduced number and magnitude of $^1H^N-^1H^N$ RDCs between exchangeable protons compared to $^1H^N-^1H$ RDCs in aligned protonated proteins. The latter can give rise to large line broadening of the $^1H^N$ resonances and therefore limits the degree to which proteins can be aligned while still yielding high quality spectra. We note that this limitation is largely removed when using ^{13}C direct detection for measurement of RDCs.[75]

The introduction of a lanthanide-chelating paramagnetic tag provides another powerful method for introducing weak alignment on a protein, which allows both structural and domain dynamics information to be extracted from RDCs and paramagnetic shifts.[76−79] Such lanthanide tags are commonly attached to a Cys side chain, but this approach proved inapplicable to MPro due to its 12 Cys residues, many of which are highly conserved and solvent accessible, which prevented attachment of the tag at a unique site.

Our study demonstrates that once a sample is adequately aligned, multiple sets of backbone RDCs can be easily obtained in addition to the most widely used amide $^{15}N-^1H$ couplings, providing additional structural restraints and improved cross-validation,[80,81] thereby potentially increasing the accuracy of structures obtainable from NMR data.

A direct comparison of frequency-based and intensity-based methods for measuring RDCs revealed that both approaches resulted in similar quality fits to the X-ray structure. Although 2D experiments have a significant time advantage over 3D and yield accurate RDCs for well resolved signals, for medium-to large-sized proteins signal overlap often reduces the number of

https://doi.org/10.1021/acsphyschemau.5c00081
ACS Phys. Chem Au 2026, 6, 81−94

RDCs that can be reliably extracted, even at high magnetic fields. For instance, the 2D $^{1}H-^{15}N$ ARTSY method allowed the measurement of 219 reliable $^{1}D_{NH}$ RDCs at 800 MHz, whereas the 3D TROSY−HNCO−ARTSY at 600 MHz provided 229 $^{1}D_{NH}$ RDCs.

Experiments that provide multiple types of RDCs in a single experiment, such as $^{1}D_{NC'}$ and $^{2}D_{C'H}$ RDCs in E. COSY spectra, are intrinsically advantageous. However, the obtained RDCs can suffer from lower precision and increased signal overlap, caused by the doubling of the number of resonances, which reduces the number of accessible RDCs. For the 2D CT-$QJ_{NC'}$ method, which also is affected by resonance overlap, spectral resolution was strongly improved by utilizing mixed-time evolution while retaining the higher sensitivity of 2D detection over the analogous 3D experiment.[82]

The frequency discrimination and intensity-based $^{1}D_{C'C\alpha}$ 3D HNCO methods both provided highly precise RDCs. However, the effect of incomplete ^{13}C enrichment and imperfect pulses in $^{1}D_{C'C\alpha}$ and $^{1}D_{NC'}$ QJ experiments can result in systematically smaller RDCs. These effects typically manifest as slightly different D_a values for the various types of couplings relative to those expected based on internuclear distances and magnetogyric ratios. Collecting multiple RDCs and fitting them to an X-ray structure clearly identifies such systematic deviations and allows for appropriate scaling. After such scaling, the QJ RDCs provided superior fits to the structure compared to frequency-based methods. While such systematic underestimates are absent in splitting-based measurements of $^{1}D_{C'C\alpha}$, systematic errors can be present in E. COSY measurements of $^{1}D_{NC'}$ and $^{2}D_{C'H}$ RDCs when the downfield ^{15}N-{$^{13}C'$} doublet component partially overlaps with the resonance of the ^{15}N-{$^{12}C'$} isotopomer, which resonates at lower field due to the ^{13}C isotope shift.[83]

RDC analysis of the monomeric mutant of the SARS-CoV-2 main protease, $MPro^{H41Q,10-306}$, revealed that domains 1 and 3 adopt structures that are close to those of the active dimer. On the other hand, domain 2 which encompasses the active site loop residues yielded a considerably poorer fit to the RDC data, indicative of differences in structure. Interestingly, AlphaFold2 models[71] of the monomeric mutant did not accurately predict the conformational change in the active site of the protease relative to the dimer, highlighting the unique importance of NMR in structural biology. Furthermore, whereas domains 1 and 2 have alignment tensors that are very similar in D_a and *Rh* values, suggestive of a well-ordered relative domain orientation, the alignment tensor of domain 3 is about 40% stronger. This result indicates that the C-terminal helical domain undergoes large amplitude motions relative to domains 1 and 2, consistent with the inability to simultaneously fit RDCs of all three domains to any X-ray structure (Figure 5). A similarly poor fit is obtained for AlphaFold models, with the orientation of the helical domain relative to the catalytic domain being the same as the dimeric X-ray crystal structure.

Our protocol used for refining the MPro monomer X-ray crystal structure by experimental RDCs showed that structural rearrangements were small, up to ca 1.2 Å, with the largest changes confined to loop regions (Figure 7A). Previous work to integrate RDCs and X-ray diffraction data aimed to generate structures that satisfied both sets of experimental data simultaneously.[84,85] Such refinement requires minimal atomic displacements relative to the X-ray data, but can highlight areas where the RDC and X-ray data are inconsistent with one another, i.e. where true structural differences between the solution and crystalline state may be present.[86] The protocol used by us allows structural deviations from the starting X-ray model but prevents large changes in local backbone torsion angles. Coordinates of the RDC-refined solution structure then can deviate substantially from the X-ray structure, and as demonstrated by us for $MPro^{H41Q,10-306}$ it can lead to improved agreement with RDCs not used in the refinement, as well as improved MolProbity scores, a comprehensive metric for evaluating the quality of structural models.[72] These results indicate that the RDC-refined X-ray model is a better representation of the protein structure in solution than the X-ray derived starting model.

We hope that our results will aid in future research decisions regarding RDC collection and analysis while promoting wider adoption of these methods (summarized in Figure 8). Such analyses will ultimately lead to better refined structural models of biomolecules and improved understanding of their dynamic behavior. To facilitate rapid adoption of these methods, we developed a streamlined CCPN[41] workflow for analyzing RDC data sets, which is available in the Supporting Information (section A3.3). Importantly, this study also highlights the application of these methods for unveiling the solution structure of the MPro monomer. NMR characterization of the monomer may facilitate in the discovery of compounds, identified through virtual or fragment-based screening, that restrict the active site from transitioning between the unwound (inactive) and wound (active) states, or that perturb dimerization following MPro excision from the polyprotein.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsphyschemau.5c00081.

Additional experimental details; pulse sequence diagrams and code; tables with scalar and dipolar couplings; annotated multidimensional NMR spectra; programs and scripts for processing and analysis of the NMR data. Acquisition parameters, pulse sequence code, data, XPLOR-NIH refinement scripts and the AlphaFold2 predicted models are also available at 10.5281/zenodo.16328473 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Ad Bax − *Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States;* orcid.org/0000-0002-9809-5700; Email: bax@nih.gov

Authors

Marshall J. Smith − *Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States*

Jinfa Ying − *Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States*

Yang Shen − *Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases,*

National Institutes of Health, Bethesda, Maryland 20892-0520, United States

John M. Louis − *Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States;* orcid.org/0000-0002-0052-1899

Complete contact information is available at: https://pubs.acs.org/10.1021/acsphyschemau.5c00081

Author Contributions

CRediT: **Marshall J Smith** conceptualization, data curation, formal analysis, investigation, methodology, software, validation, writing - original draft, writing - review & editing; **Jinfa Ying** investigation, software, writing - review & editing; **Yang Shen** resources, software; **John M. Louis** investigation, resources, writing - review & editing; **Ad Bax** conceptualization, methodology, supervision, writing - original draft, writing - review & editing.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Ivan Gagarinov for preliminary assignments of $MPro^{H41Q,10-306}$, Charles Schwieters for guidance in utilizing XPLOR-NIH, James Baber for technical assistance and Annie Aniana for help with protein expression. This research was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK; Grant DK029046-19) within the National Institutes of Health (NIH). The contributions of the authors were made as part of their official duties as NIH federal employees, are in compliance with agency policy requirements, and are considered Works of the United States Government. However, the findings and conclusions presented in this paper are those of the authors and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services.

■ REFERENCES

(1) Owen, D. R.; Allerton, C. M. N.; Anderson, A. S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin, R. D.; Carlo, A.; Coffman, K. J.; Dantonio, A.; Di, L.; Eng, H.; Ferre, R.; Gajiwala, K. S.; Gibson, S. A.; Greasley, S. E.; Hurst, B. L.; Kadar, E. P.; Kalgutkar, A. S.; Lee, J. C.; Lee, J.; Liu, W.; Mason, S. W.; Noell, S.; Novak, J. J.; Obach, R. S.; Ogilvie, K.; Patel, N. C.; Pettersson, M.; Rai, D. K.; Reese, M. R.; Sammons, M. F.; Sathish, J. G.; Singh, R. S. P.; Steppan, C. M.; Stewart, A.; Tuttle, J. B.; Updyke, L.; Verhoest, P. R.; Wei, L. Q.; Yang, Q. Y.; Zhu, Y. A. An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19. *Science* **2021**, *374*, 1586−1593.

(2) Unoh, Y.; Uehara, S.; Nakahara, K.; Nobori, H.; Yamatsu, Y.; Yamamoto, S.; Maruyama, Y.; Taoda, Y.; Kasamatsu, K.; Suto, T.; Kouki, K.; Nakahashi, A.; Kawashima, S.; Sanaki, T.; Toba, S.; Uemura, K.; Mizutare, T.; Ando, S.; Sasaki, M.; Orba, Y.; Sawa, H.; Sato, A.; Sato, T.; Kato, T.; Tachibana, Y. Discovery of S-217622, a Noncovalent Oral SARS-CoV-2 3CLProtease Inhibitor Clinical Candidate for Treating COVID-19. *J. Med. Chem.* **2022**, *65*, 6499−6512.

(3) Anand, K.; Palm, G. J.; Mesters, J. R.; Siddell, S. G.; Ziebuhr, J.; Hilgenfeld, R. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra α-helical domain. *EMBO J.* **2002**, *21*, 3213−3224.

(4) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020**, *582*, 289−293.

(5) Nashed, N. T.; Kneller, D. W.; Coates, L.; Ghirlando, R.; Aniana, A.; Kovalevsky, A.; Louis, J. M. Autoprocessing and oxyanion loop reorganization upon GC373 and nirmatrelvir binding of monomeric SARS-CoV-2 main protease catalytic domain. *Commun. Biol.* **2022**, *5*, 976.

(6) Chen, S.; Jonas, F.; Shen, C.; Higenfeld, R. Liberation of SARS-CoV main protease from the viral polyprotein: N-terminal autocleavage does not depend on the mature dimerization mode. *Protein Cell* **2010**, *1*, 59−74.

(7) Nashed, N. T.; Aniana, A.; Ghirlando, R.; Chiliveri, S. C.; Louis, J. M. Modulation of the monomer-dimer equilibrium and catalytic activity of SARS-CoV-2 main protease by a transition-state analog inhibitor. *Commun. Biol.* **2022**, *5*, 160.

(8) Cantrelle, F. X.; Boll, E.; Brier, L.; Moschidi, D.; Belouzard, S.; Landry, V.; Leroux, F.; Dewitte, F.; Landrieu, I.; Dubuisson, J.; Deprez, B.; Charton, J.; Hanoulle, X. NMR Spectroscopy of the Main Protease of SARS-CoV-2 and Fragment-Based Screening Identify Three Protein Hotspots and an Antiviral Fragment. *Angew. Chem., Int. Ed.* **2021**, *60*, 25428−25435.

(9) Altincekic, N.; Jores, N.; Löhr, F.; Richter, C.; Ehrhardt, C.; Blommers, M. J. J.; Berg, H.; Öztürk, S.; Gande, S. L.; Linhard, V.; Orts, J.; Abi Saad, M. J.; Bütikofer, M.; Kaderli, J.; Karlsson, B. G.; Brath, U.; Hedenström, M.; Gröbner, G.; Sauer, U. H.; Perrakis, A.; Langer, J.; Banci, L.; Cantini, F.; Fragai, M.; Grifagni, D.; Barthel, T.; Wollenhaupt, J.; Weiss, M. S.; Robertson, A.; Bax, A.; Sreeramulu, S.; Schwalbe, H. Targeting the Main Protease (MPro, nsp5) by Growth of Fragment Scaffolds Exploiting Structure-Based Methodologies. *ACS Chem. Biol.* **2024**, *19*, 563−574.

(10) Silva, J.; Foguel, D.; DaPoian, A.; Prevelige, P. The use of hydrostatic pressure as a tool to study viruses and other macromolecular assemblages. *Curr. Opin. Struct. Biol.* **1996**, *6*, 166−175.

(11) Nguyen, T.; Ghirlando, R.; Roche, J.; Venditti, V. Structure elucidation of the elusive Enzyme I monomer reveals the molecular mechanisms linking oligomerization and enzymatic activity. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, No. e2100298118.

(12) Arthanari, H.; Takeuchi, K.; Dubey, A.; Wagner, G. Emerging solution NMR methods to illuminate the structural and dynamic properties of proteins. *Curr. Opin. Struct. Biol.* **2019**, *58*, 294−304.

(13) Tjandra, N.; Bax, A. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **1997**, *278*, 1111−1114.

(14) Tolman, J. R.; Al-Hashimi, H. M.; Kay, L. E.; Prestegard, J. H. Structural and Dynamic Analysis of Residual Dipolar Coupling Data for Proteins. *J. Am. Chem. Soc.* **2001**, *123*, 1416−1424.

(15) Lange, O. F.; Lakomek, N. A.; Fares, C.; Schroder, G. F.; Walter, K. F. A.; Becker, S.; Meiler, J.; Grubmuller, H.; Griesinger, C.; de Groot, B. L. Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* **2008**, *320*, 1471−1475.

(16) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. Nuclear magnetic dipole interactions in field-oriented proteins - Information for structure determination in solution. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9279−9283.

(17) Hansen, M. R.; Mueller, L.; Pardi, A. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Mol. Biol.* **1998**, *5*, 1065−1074.

(18) Clore, G. M.; Starich, M. R.; Gronenborn, A. M. Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses. *J. Am. Chem. Soc.* **1998**, *120*, 10571−10572.

(19) Sass, H.-J.; Musco, G.; Stahl, S. J.; Wingfield, P. T.; Grzesiek, S. Solution NMR of proteins within polyacrylamide gels: Diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. *J. Biomol. NMR* **2000**, *18*, 303−309.

(20) Ishii, Y.; Markus, M. A.; Tycko, R. Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel. *J. Biomol. NMR* **2001**, *21*, 141−151.
(21) Prestegard, J. H.; Bougault, C. M.; Kishore, A. I. Residual dipolar couplings in structure determination of biomolecules. *Chem. Rev.* **2004**, *104*, 3519−3540.
(22) Lipsitz, R. S.; Tjandra, N. Residual dipolar couplings in NMR structure analysis. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 387−413.
(23) Chiliveri, S. C.; Robertson, A. J.; Shen, Y.; Torchia, D. A.; Bax, A. Advances in NMR Spectroscopy of Weakly Aligned Biomolecular Systems. *Chem. Rev.* **2022**, *122*, 9307−9330.
(24) Bax, A.; Grishaev, A. Weak alignment NMR: a hawk-eyed view of biomolecular structure. *Curr. Opin. Struct. Biol.* **2005**, *15*, 563−570.
(25) Bryce, D. L.; Bax, A. Application of correlated residual dipolar couplings to the determination of the molecular alignment tensor magnitude of oriented proteins and nucleic acids. *J. Biomol. NMR* **2004**, *28*, 273−287.
(26) Hus, J.-C.; Salmon, L.; Bouvignies, G.; Lotze, J.; Blackledge, M.; Brueschweiler, R. 16-Fold Degeneracy of Peptide Plane Orientations from Residual Dipolar Couplings: Analytical Treatment and Implications for Protein Structure Determination. *J. Am. Chem. Soc.* **2008**, *130*, 15927−15937.
(27) Hwang, P. M.; Skrynnikov, N. R.; Kay, L. E. Domain orientation in β-cyclodextrin-loaded maltose binding protein:: Diffusion anisotropy measurements confirm the results of a dipolar coupling study. *J. Biomol. NMR* **2001**, *20*, 83−88.
(28) Evenäs, J.; Tugarinov, V.; Skrynnikov, N. R.; Goto, N. K.; Muhandiram, R.; Kay, L. E. Ligand-induced structural changes to maltodextrin-binding protein as studied by solution NMR spectroscopy. *J. Mol. Biol.* **2001**, *309*, 961−974.
(29) Ottiger, M.; Delaglio, F.; Bax, A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.* **1998**, *131*, 373−378.
(30) Tolman, J. R.; Prestegard, J. H. A quantitative J-correlation experiment for the accurate measurement of one-bond amide N-15-H-1 couplings in proteins. *J. Magn. Reson. B* **1996**, *112*, 245−252.
(31) Ottiger, M.; Delaglio, F.; Marquardt, J. L.; Tjandra, N.; Bax, A. Measurement of dipolar couplings for methylene and methyl sites in weakly oriented macromolecules and their use in structure determination. *J. Magn. Reson.* **1998**, *134*, 365−369.
(32) Evenas, J.; Mittermaier, A.; Yang, D. W.; Kay, L. Measurement of 13Ca-13Cb dipolar couplings in 15N,13C,2H-labeled proteins: application to domain orientation in maltose binding protein. *J. Am. Chem. Soc.* **2001**, *123*, 2858−2864.
(33) Liu, Y. Z.; Prestegard, J. H. Measurement of one and two bond N-C couplings in large proteins by TROSY-based J-modulation experiments. *J. Magn. Reson.* **2009**, *200*, 109−118.
(34) Aniana, A.; Nashed, N. T.; Ghirlando, R.; Drago, V. N.; Kovalevsky, A.; Louis, J. M. Characterization of alternate encounter assemblies of SARS-CoV-2 main protease. *J. Biol. Chem.* **2024**, *300*, 107675.
(35) Salzmann, M.; Wider, G.; Pervushin, K.; Senn, H.; Wuthrich, K. TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. *J. Am. Chem. Soc.* **1999**, *121*, 844−848.
(36) Robertson, A. J.; Ying, J.; Bax, A. Four-dimensional NOE-NOE spectroscopy of SARS-CoV-2 Main Protease to facilitate resonance assignment and structural analysis. *Magn. Reson.* **2021**, *2*, 129−138.
(37) Fitzkee, N. C.; Bax, A. Facile measurement of H-1-N-15 residual dipolar couplings in larger perdeuterated proteins. *J. Biomol. NMR* **2010**, *48*, 65−70.
(38) Robertson, A. J.; Courtney, J. M.; Shen, Y.; Ying, J. F.; Bax, A. Concordance of X-ray and AlphaFold2Models of SARS-CoV-2 Main Protease with Residual Dipolar Couplings Measured in Solution. *J. Am. Chem. Soc.* **2021**, *143*, 19306−19310.
(39) Jaroniec, C. P.; Ulmer, T. S.; Bax, A. Quantitative J correlation methods for the accurate measurement of 13C′-13Cαdipolar couplings in proteins. *J. Biomol. NMR* **2004**, *30*, 181−194.
(40) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRpipe - a multidimensional spectral processing system based on Unix pipes. *J. Biomol. NMR* **1995**, *6*, 277−293.
(41) Skinner, S. P.; Fogh, R. H.; Boucher, W.; Ragan, T. J.; Mureddu, L. G.; Vuister, G. W. CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J. Biomol. NMR* **2016**, *66*, 111−124.
(42) Shi, J. H.; Sivaraman, J.; Song, J. X. Mechanism for controlling the dimer-monomer switch and coupling dimerization to catalysis of the severe acute respiratory syndrome coronavirus 3C-like protease. *J. Virol.* **2008**, *82*, 4620−4629.
(43) Leman, J. K.; Weitzner, B. D.; Lewis, S. M.; Adolf-Bryfogle, J.; Alam, N.; Alford, R. F.; Aprahamian, M.; Baker, D.; Barlow, K. A.; Barth, P.; Basanta, B.; Bender, B. J.; Blacklock, K.; Bonet, J.; Boyken, S. E.; Bradley, P.; Bystroff, C.; Conway, P.; Cooper, S.; Correia, B. E.; Coventry, B.; Das, R.; De Jong, R. M.; DiMaio, F.; Dsilva, L.; Dunbrack, R.; Ford, A. S.; Frenz, B.; Fu, D. Y.; Geniesse, C.; Goldschmidt, L.; Gowthaman, R.; Gray, J. J.; Gront, D.; Guffy, S.; Horowitz, S.; Huang, P.-S.; Huber, T.; Jacobs, T. M.; Jeliazkov, J. R.; Johnson, D. K.; Kappel, K.; Karanicolas, J.; Khakzad, H.; Khar, K. R.; Khare, S. D.; Khatib, F.; Khramushin, A.; King, I. C.; Kleffner, R.; Koepnick, B.; Kortemme, T.; Kuenze, G.; Kuhlman, B.; Kuroda, D.; Labonte, J. W.; Lai, J. K.; Lapidoth, G.; Leaver-Fay, A.; Lindert, S.; Linsky, T.; London, N.; Lubin, J. H.; Lyskov, S.; Maguire, J.; Malmström, L.; Marcos, E.; Marcu, O.; Marze, N. A.; Meiler, J.; Moretti, R.; Mulligan, V. K.; Nerli, S.; Norn, C.; Ó'Conchúir, S.; Ollikainen, N.; Ovchinnikov, S.; Pacella, M. S.; Pan, X.; Park, H.; Pavlovicz, R. E.; Pethe, M.; Pierce, B. G.; Pilla, K. B.; Raveh, B.; Renfrew, P. D.; Burman, S. S. R.; Rubenstein, A.; Sauer, M. F.; Scheck, A.; Schief, W.; Schueler-Furman, O.; Sedan, Y.; Sevy, A. M.; Sgourakis, N. G.; Shi, L.; Siegel, J. B.; Silva, D.-A.; Smith, S.; Song, Y.; et al. Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nat. Methods* **2020**, *17*, 665−680.
(44) Schwieters, C. D.; Clore, G. M. Internal coordinates for molecular dynamics and minimization in structure determination and refinement. *J. Magn. Reson.* **2001**, *152*, 288−302.
(45) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Marius Clore, G. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **2003**, *160*, 65−73.
(46) Chou, J. J.; Li, S.; Bax, A. Study of conformational rearrangement and refinement of structural homology models by the use of heteronuclear dipolar couplings. *J. Biomol. NMR* **2000**, *18*, 217−227.
(47) Schwieters, C. D.; Bermejo, G. A.; Clore, G. M. A three-dimensional potential of mean force to improve backbone and sidechain hydrogen bond geometry in Xplor-NIH protein structure determination. *Protein Sci.* **2020**, *29*, 100−110.
(48) Morris, G. A.; Freeman, R. Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc.* **1979**, *101*, 760−762.
(49) Pervushin, K.; Riek, R.; Wider, G.; 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. U.S.A.* **1997**, *94*, 12366−12371.
(50) Hyberts, S. G.; Frueh, D. P.; Arthanari, H.; Wagner, G. FM reconstruction of non-uniformly sampled protein NMR data at higher dimensions and optimization by distillation. *J. Biomol. NMR* **2009**, *45*, 283−294.
(51) Hoch, J. C.; Maciejewski, M. W.; Mobli, M.; Schuyler, A. D.; Stern, A. S. Nonuniform Sampling and Maximum Entropy Reconstruction in Multidimensional NMR. *Acc. Chem. Res.* **2014**, *47*, 708−717.
(52) Mayzel, M.; Rosenlow, J.; Isaksson, L.; Orekhov, V. Y. Time-resolved multidimensional NMR with non-uniform sampling. *J. Biomol. NMR* **2014**, *58*, 129−139.
(53) Ying, J.; Delaglio, F.; Torchia, D. A.; Bax, A. Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruc-

https://doi.org/10.1021/acsphyschemau.5c00081
ACS Phys. Chem Au 2026, 6, 81−94

tion of both non-uniformly sampled and conventional NMR data. *J. Biomol. NMR* **2017**, *68*, 101−118.
(54) Kazimierczuk, K.; Orekhov, V. Y. Accelerated NMR Spectroscopy by Using Compressed Sensing. *Angew. Chem., Int. Ed.* **2011**, *50*, 5556−5559.
(55) Kubat, J. A.; Chou, J. J.; Rovnyak, D. Nonuniform sampling and maximum entropy reconstruction applied to the accurate measurement of residual dipolar couplings. *J. Magn. Reson.* **2007**, *186*, 201−211.
(56) Robertson, A. J.; Ying, J. F.; Bax, A. NMR Observation of Sulfhydryl Signals in SARS-CoV-2 Main Protease Aids Structural Studies. *Chembiochem* **2022**, *23*, No. e202200471.
(57) Griesinger, C.; Sørensen, O. W.; Ernst, R. R. Practical aspects of the E.COSY technique. Measurement of scalar spin-spin coupling constants in peptides. *J. Magn. Reson.* **1987**, *75*, 474−492.
(58) Wang, Y. X.; Marquardt, J. L.; Wingfield, P.; Stahl, S. J.; Lee-Huang, S.; Torchia, D.; Bax, A. Simultaneous measurement of 1H-15N, 1H-13C', and 15N-13C' dipolar couplings in a perdeuterated 30 kDa protein dissolved in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* **1998**, *120*, 7385−7386.
(59) Ying, J. F.; Chill, J. H.; Louis, J. M.; Bax, A. Mixed-time parallel evolution in multiple quantum NMR experiments: sensitivity and resolution enhancement in heteronuclear NMR. *J. Biomol. NMR* **2007**, *37*, 195−204.
(60) Hwang, T. L.; van Zijl, P. C. M.; Garwood, M. Fast broadband inversion by adiabatic pulses. *J. Magn. Reson.* **1998**, *133*, 200−203.
(61) Shen, Y.; Robertson, A.; Bax, A. Validation of X-ray Crystal Structure Ensemble Representations of SARS-CoV-2 Main Protease by Solution NMR Residual Dipolar Couplings. *J. Mol. Biol.* **2023**, *435*, 168067.
(62) Neuhaus, D.; Wagner, G.; Vasak, M.; Kagi, J. H. R.; Wuthrich, K. Systematic Application of High-Resolution, Phase-Sensitive Two-Dimensional H-1-Nmr Techniques For the Identification of the Amino-Acid-Proton Spin Systems in Proteins - Rabbit Metallothionein-2. *Eur. J. Biochem.* **1985**, *151*, 257−273.
(63) Montelione, G. T.; Emerson, S. D.; Lyons, B. A. A General-Approach For Determining Scalar Coupling-Constants in Polypeptides and Proteins. *Biopolymers* **1992**, *32*, 327−334.
(64) Harbison, G. S. Interference Between J-Couplings and Cross-Relaxation in Solution NMR-Spectroscopy - Consequences For Macromolecular Structure Determination. *J. Am. Chem. Soc.* **1993**, *115*, 3026−3027.
(65) Kontaxis, G.; Clore, G. M.; Bax, A. Evaluation of cross-correlation effects and measurement of one- bond couplings in proteins with short transverse relaxation times. *J. Magn. Reson.* **2000**, *143*, 184−196.
(66) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. NMR evidence for slow collective motions in cyanometmyoglobin. *Nat. Struct. Biol.* **1997**, *4*, 292−297.
(67) Zhang, Q.; Sun, X. Y.; Watt, E. D.; Al-Hashimi, H. M. Resolving the motional modes that code for RNA adaptation. *Science* **2006**, *311*, 653−656.
(68) Tang, C.; Schwieters, C. D.; Clore, G. M. Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* **2007**, *449*, 1078−1082.
(69) Shen, Y.; Smith, M. J.; Louis, J. M.; Bax, A. Alpha-helices as alignment reporters in residual dipolar coupling analysis of proteins. *J. Biomol. NMR* **2025**, *79*, 47−57.
(70) Andi, B.; Kumaran, D.; Kreitler, D. F.; Soares, A. S.; Keereetaweep, J.; Jakoncic, J.; Lazo, E. O.; Shi, W. X.; Fuchs, M. R.; Sweet, R. M.; Shanklin, J.; Adams, P. D.; Schmidt, J. G.; Head, M. S.; McSweeney, S. Hepatitis C virus NS3/4A inhibitors and other drug-like compounds as covalent binders of SARS-CoV-2 main protease. *Sci. Rep.* **2022**, *12*, 12197.
(71) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583−589.
(72) Williams, C. J.; Headd, J. J.; Moriarty, N. W.; Prisant, M. G.; Videau, L. L.; Deis, L. N.; Verma, V.; Keedy, D. A.; Hintze, B. J.; Chen, V. B.; Jain, S.; Lewis, S. M.; Arendall, W. B.; Snoeyink, J.; Adams, P. D.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **2018**, *27*, 293−315.
(73) Tycko, R.; Blanco, F. J.; Ishii, Y. Alignment of biopolymers in strained gels: A new way to create detectable dipole-dipole couplings in high-resolution biomolecular NMR. *J. Am. Chem. Soc.* **2000**, *122*, 9340−9341.
(74) Ruckert, M.; Otting, G. Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J. Am. Chem. Soc.* **2000**, *122*, 7793−7797.
(75) Felli, I. C.; Pierattelli, R. 13C Direct Detected NMR for Challenging Systems. *Chem. Rev.* **2022**, *122*, 9468−9496.
(76) Bertini, I.; Del Bianco, C.; Gelis, I.; Katsaros, N.; Luchinat, C.; Parigi, G.; Peana, M.; Provenzani, A.; Zoroddu, M. A. Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6841−6846.
(77) Haussinger, D.; Huang, J. R.; Grzesiek, S. DOTA-M8: an extremely rigid, high-affinity lanthanide chelating tag for PCS NMR spectroscopy. *J. Am. Chem. Soc.* **2009**, *131*, 14761−14767.
(78) Karschin, N.; Becker, S.; Griesinger, C. Interdomain Dynamics via Paramagnetic NMR on the Highly Flexible Complex Calmodulin/Munc13−1. *J. Am. Chem. Soc.* **2022**, *144*, 17041−17053.
(79) Wang, Y.; An, L.; Yang, Y.; Yao, L. Generating five independent molecular alignments for simultaneous protein structure and dynamics determination using nuclear magnetic resonance spectroscopy. *Anal. Chem.* **2020**, *92*, 15263−15269.
(80) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* **1998**, *120*, 6836−6837.
(81) Clore, G. M.; Garrett, D. S. R-factor, free R, and complete cross-validation for dipolar coupling refinement of NMR structures. *J. Am. Chem. Soc.* **1999**, *121*, 9008−9012.
(82) Chou, J. J.; Delaglio, F.; Bax, A. Measurement of 15N-13C' dipolar couplings in medium sized proteins. *J. Biomol. NMR* **2000**, *18*, 101−105.
(83) Ikura, M.; Krinks, M.; Torchia, D. A.; Bax, A. An Efficient Nmr Approach For Obtaining Sequence-Specific Resonance Assignments of Larger Proteins Based On Multiple Isotopic Labeling. *FEBS Lett.* **1990**, *266*, 155−158.
(84) Carlon, A.; Ravera, E.; Hennig, J.; Parigi, G.; Sattler, M.; Luchinat, C. Improved accuracy from joint X-ray and NMR refinement of a protein-RNA complex structure. *J. Am. Chem. Soc.* **2016**, *138*, 1601−1610.
(85) Rinaldelli, M.; Ravera, E.; Calderone, V.; Parigi, G.; Murshudov, G. N.; Luchinat, C. Simultaneous use of solution NMR and X-ray data in REFMAC5 for joint refinement/detection of structural differences. *Acta Crystallogr., Sect. D:Biol. Crystallogr.* **2014**, *70*, 958−967.
(86) Carlon, A.; Ravera, E.; Parigi, G.; Murshudov, G. N.; Luchinat, C. Joint X-ray/NMR structure refinement of multidomain/multi-subunit systems. *J. Biomol. NMR* **2019**, *73*, 265−278.

https://doi.org/10.1021/acsphyschemau.5c00081
ACS Phys. Chem Au 2026, 6, 81−94