ARTICLE



Alpha-helices as alignment reporters in residual dipolar coupling analysis of proteins

Yang Shen¹ · Marshall J. Smith¹ · John M. Louis¹ · Ad Bax¹

Received: 26 September 2024 / Accepted: 29 November 2024 This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2024

Abstract

Inclusion of residual dipolar couplings (RDCs) during the early rounds of protein structure determination requires use of a floating alignment tensor or knowledge of the alignment tensor strength and rhombicity. For proteins with interdomain motion, such analysis can falsely hide the presence of domain dynamics. We demonstrate for three proteins, maltotriose-ligated maltose binding protein (MBP), Ca²⁺-ligated calmodulin, and a monomeric N-terminal deletion mutant of the SARS-CoV-2 Main Protease, MPro, that good alignment tensor estimates of their domains can be obtained from RDCs measured for residues that are identified as α -helical based on their chemical shifts. The program, Helix-Fit, fits the RDCs to idealized α -helical coordinates, often yielding a comparable or better alignment tensor estimate than fitting to the actual high-resolution X-ray helix coordinates. The 13 helices of ligated MBP all show very similar alignment tensors, indicative of a high degree of order relative to one another. By contrast, while for monomeric MPro the alignment strengths of the five helices in the C-terminal helical domain (residues 200–306) are very similar, pointing to a well-ordered domain, the single α -helix Y54-I59 in the N-terminal catalytic domain (residues 10–185) aligns considerably weaker. This result indicates the presence of large amplitude motions of either Y54-I59 or of the entire N-terminal domain relative to the C-terminal domain, contrasting with the high degree of order seen in the native homodimeric structure.

Keywords RDC · Protein dynamics · Protein NMR · Domain motion · SARS-CoV-2 MPro

Introduction

Residual dipolar couplings (RDCs) in biological macromolecules can be observed by solution NMR when there exists a net, time-averaged orientation of the molecule relative to an external magnetic field (Tolman et al. 1995; Tjandra and Bax 1997). These RDCs can yield very precise information on the time-averaged orientation of bonds relative to the global reference frame of the molecule's alignment tensor, thereby encoding both structural and dynamic information (Bertini et al. 2004; Lindorff-Larsen et al. 2005; Lange et al. 2008).

Weak molecular alignment relative to the magnetic field can be obtained by multiple mechanisms, including paramagnetism (Tolman et al. 1995; Wang et al. 2020), dilute liquid crystalline phases (Tjandra and Bax 1997; Clore et al.

1998a, b; Hansen et al. 1998; Ruckert and Otting 2000), and anisotropically compressed hydrogels (Sass et al. 2000; Tycko et al. 2000). However, such measurements are often restricted to a single mode of alignment because optimizing each alignment method can be labor intensive. However, once a suitably aligned sample for RDC measurements has been generated, it is relatively straightforward to measure not only the most commonly used $^1\text{D}_{\rm NH}$ RDCs, but also many additional ¹H-¹³C, ¹³C-¹⁵N, and ¹³C-¹³C dipolar couplings (Tjandra and Bax 1997; Yang et al. 1999; Evenas et al. 2001b; Chiliveri et al. 2022). The latter can agree better with crystallographically determined reference structures because they are not impacted by errors in the precise positions of the amide hydrogen atoms, which typically are added to the X-ray crystal structure by assuming standard, in-plane geometry (Ottiger and Bax 1998).

The strong dependence of RDCs on bond vector orientation can make it difficult to effectively use them during early rounds of structure calculations because it often causes very large scatter between experimental RDCs and values predicted for imperfect structural models. Either a

Ad Bax bax@nih.gov

¹ Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, USA

floating alignment tensor may be used during the simulated annealing process (Sass et al. 2001), or a rough estimate of the alignment tensor's magnitude and rhombicity can be obtained from a histogram analysis of the distribution of RDCs (Clore et al. 1998a, b). Even when an approximate protein structure is available it may be difficult to accurately define its alignment tensor or to evaluate whether the protein exists in solution as a single globular unit, or as a set of flexibly linked domains. Poor fits of the RDCs to molecular coordinates also can be caused by experimental errors, e.g. assignment errors or RDCs derived from partially overlapped resonances. Alternatively, a poor fit can result from dynamics in solution that is not reflected in the reference structure, or from true differences between solution and crystalline states. It therefore is desirable to identify self-consistent RDCs that correspond to a well-defined local alignment tensor, or multiple well-defined tensors for different regions in a protein that contains flexibly linked domains.

Here, we demonstrate that RDCs in α -helices, which are identified from backbone chemical shifts by the TALOS-N program (Shen and Bax 2013), can yield high quality alignment tensor information, provided that multiple types of backbone RDCs are available per residue. We introduce the program Helix-Fit, which relies on the assumption of idealized helical coordinates, and often yields a quality for the RDC fit that is comparable to using the high-resolution X-ray coordinates of such helices. Use of Helix-Fit is demonstrated for three proteins, maltose binding protein (MBP), calmodulin, and a monomeric form of the main Protease (MPro) of SARS-CoV-2.

The accuracy of alignment tensors depends on how well orientations are sampled by the internuclear vectors for which RDCs are measured. Fushman et al. introduced a "generalized sampling parameter", Ξ , that quantifies the distribution of vector orientations on a scale from 0 (optimal sampling) to 1 (poor sampling) (Fushman et al. 2000). Whereas for N-H vectors in α -helices this sampling is poor ($\Xi \approx 0.83$), excellent sampling ($\Xi \le 0.17$) applies for other vectors (C^{α} -C'; N-C'; C^{α} -H^{α}; C'-H^N; and C'-H^{α}) whose RDCs are readily measured. The favorable orientational sampling by these vectors benefits the robustness at which alignment tensors of α -helices can be determined (Fushman et al. 2000).

Results

Identification of α-helices

When atomic coordinates are available, α -helices are characterized by backbone torsion angles, typically falling in a relatively narrow range ($\phi = -63 \pm 6^\circ$; $\psi = -42 \pm 6^\circ$), and their H-bonding between the C=O of residue *i* and the H^N of i+4 (Kabsch and Sander 1983). In the absence of structural information, α -helices are readily identified by their characteristic deviations from random coil chemical shifts (Spera and Bax 1991; Wishart et al. 1991) or, if available, sequential NOE patterns (Wüthrich 1986). The H-bonding pattern in α -helices results in highly regular structures that often closely fit to an idealized α -helix (Table S1 for ideal helical parameters used in this study). As a consequence, RDCs for backbone N-H pairs, whose internuclear vectors differ by 15.8° in orientation from the helix axis, show a characteristic sinusoidal pattern as a function of residue number with a periodicity of 3.6 residues, known as the "dipolar wave" (Mesleh et al. 2003). Below, we used the neural-network based TALOS-N program (Shen and Bax 2013) for identifying α -helices of ≥ 5 residues in length. For our current purpose of finding close-to-ideal α -helices, we found it necessary to terminate helices at the last residue preceding a Pro residue, or to start them at the first residue following a Pro when such a residue is embedded in a longer, kinked α -helix.

RDC fitting of α-helices

Agreement between RDCs and atomic coordinates is limited by the precision at which RDCs can be measured as well as by uncertainties in the atomic coordinates. The latter often dominate the scatter between observed and best-fitted RDCs, obtained from a singular value decomposition (SVD) fit (Losonczi et al. 1999) of the RDCs to the internuclear vector orientations (Shen et al. 2023). The SVD fit can include weighting to account for the experimental error estimates, but in our examples below, the experimental errors are assumed to scale with the inverse of the dipolar interaction constants that are determined by the magnetogyric ratios and internuclear distances of the pairs of atoms. For example, the errors in ${}^{1}D_{C'C\alpha}$ are assumed to be five times smaller than for ${}^{1}D_{NH}$, effectively giving ${}^{1}D_{C'C\alpha}$ and ${}^{1}D_{NH}$ equal weight in the analysis. The use of equal weights for fitting the various types of normalized RDCs, available for MBP, calmodulin, and monomeric MPro, is justified by comparable qualities of their SVD fits, indicating that the errors in the fit are dominated by coordinate uncertainty, not by measurement error of the RDCs. However, in the program Helix-Fit, as well as all our other RDC analysis software, the estimated error in the measured RDC will be used if specified in the input table.

For a well-ordered globular domain that contains multiple helices, separate fits of the RDCs to its various helices are expected to yield very similar alignment tensors, with the uncertainty in each alignment tensor determined by jackknifing. The jackknife procedure cyclically omits one RDC from the total set of N RDCs, and then carries out N SVD fits on the remaining N-1 RDCs, resulting in N alignment tensors. Similarly, jackknifing at the residue level can be carried out by cyclically leaving out all RDCs (e.g. ${}^{1}D_{NH}$, ${}^{1}D_{CN}$, ${}^{1}D_{C\alpha C'}$, ${}^{2}D_{HC'}$) measured for a given amide, a procedure useful for identifying whether a residue identified by TALOS-N as the terminal helix residue indeed is consistent with helical geometry.

We represent the uncertainty in the orientation and rhombicity of the final, averaged alignment tensor, $\langle S \rangle$, by $\varepsilon(S)$, which is related to the spread in their normalized scalar products (Sass et al. 1999):

$$\varepsilon(S) = (N)^{1/2} RMS \left[\left\{ 1 - P(S_i, \langle S \rangle) \right\} \right]$$
(1a)

where RMS{1-P($S_i < S >$)} = $\sqrt{\sum_{i=1..N} (1 - P(S_i, \langle S \rangle))^2/N}$ is the root-mean-square (RMS) deviation from unity for the normalized scalar products P($S_i, < S >$) of the N jackknifed five-dimensional Saupe matrix vectors, $S_i = [S_{zz}, (1/\sqrt{3})(S_{xx}-S_{yy}), (2/\sqrt{3})(S_{xy}), (2/\sqrt{3})(S_{xz}), (2/\sqrt{3})(S_{yz})]_{i=1..N}$, relative to the averaged Saupe matrix vector < S >, or P($S_i, <S >$) = $S_i \cdot <S >/(|S_i| \times |<S >|)$.

Analogously, the fractional uncertainty in the alignment strength is

$$\epsilon(G) = (\mathbf{N})^{1/2} \mathbf{RMS} \left(G_i - \langle G \rangle \right) / \langle G \rangle$$
(1b)

with G_i given by (Clore and Garrett 1999):

$$G_i = \left\{ D_{a,i}^2 \left[4 + 3Rh_i^2 \right] / 5 \right\}^{1/2}$$
(1c)

where $D_{a,i}$ and Rh_i are the magnitude and rhombicity of the alignment tensor, obtained when RDC *i* is removed from the fit. The fit is repeated *N* times and each SVD fit omits a different coupling from the *N* available RDCs. $\langle G \rangle$ represents the average over the *N* G_i values obtained by this jackknife procedure, with reported values normalized to the interaction strength of the ¹⁵N-¹H backbone amide pair.

A jackknifed Q_{jk} factor (Shen et al. 2023) is used to evaluate the SVD fit:

$$Q_{jk} = \sqrt{\sum_{i=1..N} \frac{(W_{pq}D_i^{pred} - W_{pq}D_i^{meas})^2}{N\left\{D_{a,i}^2(4 + 3Rh_i^2)/5\right\}}}$$
(2)

where D_i^{pred} and D_i^{meas} refer to the predicted and the measured values for coupling *i* when the SVD is carried out for all other N - 1 couplings, excepting *i*. When fitting a large number (N>>5) of RDCs, Q_{jk} approaches the standard *Q* value, derived when including all RDCs in the SVD fit. However, because the SVD fit includes five adjustable parameters, the standard method for deriving *Q* strongly overestimates the goodness of the fit when the number of RDCs is small (N < ~25). This problem is solved by the computationally more burdensome jackknifing procedure. With five adjustable parameters in the SVD fit, the jackknifing procedure requires a minimum of six RDCs. For more robust results, in practice we only consider helices with at least eight RDCs. Therefore, if only ${}^{1}D_{\rm NH}$ RDCs are measured, helices are required to be relatively long for such an analysis. However, since several other backbone RDCs can be obtained at high relative precision, we limit the minimum length of α -helical elements evaluated by Helix-Fit to five residues, *i.e.* including at least one *i to i* – 4 amide-tocarbonyl H-bond.

Helix-Fit can use either the atomic coordinates of the reference structure for SVD fitting or, controlled by a flag, uses the coordinates of an idealized helix that is best-fit superimposed on the heavy backbone atoms (N, C^{α} , and C') of the residues selected.

Maltose binding protein

For ligated maltose binding protein (MBP), there is close agreement between α -helices identified from crystallographic atomic coordinates (PDB entry 3MBP) (Quiocho et al. 1997) and from NMR backbone chemical shifts (BMRB entry 4354) (Gardner et al. 1998) (Fig. 1A). Small differences are confined to the N- and C-termini which may be recognized as helical by TALOS-N, even though their respective ϕ and ψ angles sometimes deviate by more than 30° from canonical helix values (Fig. 1A). For RDC-fitting purposes, we therefore evaluate whether removal of RDCs that report on the orientation of the N-terminal and C-terminal amide planes, identified as helical by TALOS-N, can greatly improve the quality of the RDC fit to the coordinates of an idealized helix (see Methods). For example, the Q_{ik} value obtained from fitting helix 12 decreases from 0.53 to 0.33 after removal of the RDCs arising from the N-terminal amide plane of Q335; a similar improvement of Q_{ik} from 0.30 to 0.19 is obtained for helix 8 after removing the RDCs related to the C-terminal amide plane of E281. The G values and their standard deviations can be compared for fits of the RDCs to the ideal helical coordinates and those in the actual X-ray structures (Fig. 1B). As can be seen, for MBP the spread in G values is small, and the normalized scalar products between the global alignment tensor of the entire protein, S_{Global} , and fits to tensors obtained for the individual X-ray ($S_{Xray-helix}$) or idealized ($S_{ideal-helix}$) α -helices are very large (Fig. 1C), confirming that the helices in ligated MBP are highly ordered.

Calmodulin

Calmodulin is a two-domain protein, regulating the activity of more than one hundred targets, many of them kinases, in a Ca^{2+} -dependent manner (Crivici and Ikura 1995). It contains two globular domains that are homologous in sequence, and



Fig. 1 ¹D_{HN}, ¹D_{C'N} and ¹D_{C\alphaC'} RDC analysis of α -helices in MBP. The generalized sampling parameter for these couplings is $\Xi = 0.066$. **A** Colored bars along the amino acid sequence refer to residues defined as helical by TALOS-N. Only helices with length \geq 5 residues are used for RDC analysis, with some terminal residues (red) culled prior to final fitting due to their RDC incompatibility with α -helical structure. Helices < 5 residues are marked in blue and were not considered in the Helix-Fit analysis. **B** *G* values and their jack-knifed standard deviations for the 13 helices in maltotriose-ligated

MBP obtained when using idealized α -helical coordinates (orange) and X-ray atomic coordinates (PDB entry 3MBP, gray); the *G* value obtained when fitting all RDCs is plotted as the horizontal line, and the corresponding Q_{jk} values (open circles) correspond to the scale at the right side of the panel. **C** Correlation of normalized scalar products between the alignment tensors obtained for the individual α -helices (labeled by helix number) and the tensor obtained from the full set of RDCs (S_{global}), covering the entire protein

each consists of two EF-hand Ca^{2+} -binding motifs. The two domains are connected by a linker that is entirely α -helical in the crystalline state (Babu et al. 1988; Wilson and Brunger 2000). However, based on solution ¹⁵N relaxation analysis that showed independent, near-isotropic rotational diffusion of its two domains, this so-called "central helix" is highly disordered near its midpoint (Barbato et al. 1992). As expected, binding of a paramagnetic lanthanide in one of the two N-terminal Ca²⁺-binding sites imposes substantial magnetic field alignment for the N-terminal domain, but much weaker alignment on the flexibly linked C-terminal domain (Bertini et al. 2004).

A large set of ¹H-¹⁵N, ¹H^{α -13}C^{α}, ¹³C'_{i-1}-N_i and ¹³C^{α -13}C', as well as two-bond ¹H^{α -13}C' RDCs were previously reported for Ca²⁺-ligated calmodulin (Chou et al. 2001). For each calmodulin domain, the RDC data pointed to an, on average, narrower target-binding groove in solution than seen in the X-ray structure. Indeed, a best-fit of the RDCs in calmodulin's four N-terminal domain α -helices to the 1 Å X-ray structure shows a rather poor fit (Q_{jk}=0.44) (Fig. 2A). Fitting the same RDCs to the experimental X-ray structure

but with coordinates of each of the four helices replaced by best-fit superimposed ideal helices shows a comparably poor fit (Q_{jk} =0.43). However, separately fitting the RDCs to coordinates of individual idealized helices that are best-fit superimposed on the X-ray helices are considerably better (Fig. 2B–E), comparable to results obtained by fitting to the 1.0 Å X-ray coordinates of these four helices (SI Fig. S1). This result is consistent with the earlier observation that the average EF-hand interhelical angles in solution differ from those seen in the X-ray structure by *ca* 25° (Chou et al. 2001).

RDC fits of the four individual N-terminal domain helices all exhibit G values well above the G value obtained when fitting RDCs of the entire N-terminal domain, regardless of using idealized coordinates or the X-ray coordinates for the SVD fits (Fig. 3A). This discrepancy in G values is primarily caused by a difference between the average relative helix orientations seen in the X-ray structure and those present in solution (Chou et al. 2001). These differences in average helix orientation are reflected in the well below unity value of the normalized scalar products of the alignment tensors



Fig.2 RDC analysis of α -helices in Ca²⁺-calmodulin's N-terminal domain. **A** Fit of the normalized ${}^{1}D_{NH}$ (red ball), ${}^{1}D_{H\alphaC\alpha}$ (yellow ball), ${}^{1}D_{C'N}$ (green ball), ${}^{1}D_{C\alphaC'}$ (blue ball), and ${}^{2}D_{H\alphaC'}$ (pink ball) RDCs, reported by Chou et al. (2001) against values predicted by an SVD fit to the 1 Å X-ray structure (PDB entry 1EXR) (Wilson and Brunger 2000). **B–E** Individual SVD fits of normalized experimental

RDCs to coordinates of idealized α -helices, best-fit superimposed on the corresponding backbone atoms (N, C^{α}, C') of the X-ray coordinates. **B** Helix 1 (E6-F19); **C** Helix 2 (T29 to S38); **D** Helix 3 (E45– E54); and **E** Helix 4 (F65-R74). *N* refers to the number of backbone RDCs available for the SVD fits, and Q_{ik} to the jackknifed Q values



Fig.3 Alignment of helices in Ca²⁺-ligated calmodulin, oriented in 15 mg/ml Pf1 (Chou et al. 2001). **A** Generalized alignment strengths, *G*, obtained by eight separate SVD fits of its α-helices (E6-F19; T29-S38; E45-E54; F65-R74; E82-F92; A102-N111; D118-E127; Y138-M145) when using the X-ray structure (1EXR; grey) or coordinates of idealized helices that are best-fitted to the X-ray structure (orange). ¹D_{NH}, ¹D_{HαCα}, ¹D_{CN}, ¹D_{CαC}, and.²D_{HαC} couplings were used (Ξ =0.00). Error bars correspond to the $\varepsilon(G)$ values derived by jackknifing (Eq. 1b). The left horizontal line corresponds to the <*G*> value obtained by simultaneously fitting RDCs of all four N-terminal domain helices to the X-ray structure (D_a =9.97±0.17 Hz; Rh=0.41±0.03; G=9.46±0.15); the right

horizontal line corresponds to the $\langle G \rangle$ value obtained by fitting RDCs of all four C-terminal domain helices to the X-ray structure $(D_a = 9.27 \pm 0.17 \text{ Hz}; Rh = 0.65 \pm 0.02; G = 9.52 \pm 0.19)$. **B** Q_{jk} values obtained by fitting RDCs to the X-ray (grey) and idealized (orange) helical coordinates. **C** Normalized scalar product values $P(S_i, S_j)$ for the alignment tensors *S* obtained for N-terminal domain α -helices relative to one another, and relative to the alignment tensor obtained for a global RDC fit of the entire domain (i,j = 'H1', 'H2', 'H3', 'H4', and 'Nterm'). The lower right half corresponds to use of idealized helical coordinates; the top left half corresponds to using X-ray coordinates. **D** Same as **C** but for the C-terminal domain (see also SI Fig. S2)

obtained for the individual helices, both relative to one another and relative to a fit of each entire domain (Fig. 3B, C). Indeed, the variability in the interhelical EF-hand angles seen in high-resolution X-ray structures of complexes with different target sites highlights the importance of the flexible nature of these EF-hands, which allows calmodulin to finetune the interactions with its wide range of targets (Akke and Chazin 2001).

As previously reported (Chou et al. 2001), RDCs measured for calmodulin's C-terminal domain are more consistent with the X-ray structure than those of the N-terminal domain, simply reflecting a smaller difference in the average EF-hand interhelical angles in solution relative to the crystalline state. This conclusion is also reflected in $P(S_i, S_j)$ values for the C-terminal domain helices (i,j = 'H5' to 'H8') (Fig. 3D) that are closer to unity than for the N-terminal domain (i,j = 'H1' to 'H4') (Fig. 3C). Remarkably, the G values of the N- and C-terminal domains are very similar to one another, despite the fact that these domains are flexibly linked. While the $P(S_{Nterm}, S_{Cterm})$ value for the alignment tensors of the N- and C-terminal domains, S_{Nterm} and S_{Cterm} , is low (0.08) when using the X-ray structure as a reference, reorienting the C-terminal domain such that its principal axis system coincides with that of the N-terminal domain raises this $P(S_{Nterm}, S_{Cterm})$ value to 0.98 (Fig. S3). It could be argued that the latter result is consistent with a static structure that adopts this alternate relative domain orientation. This incorrect conclusion highlights the fact that a $P(S_{Nterm}, S_{Cterm})$ value near unity does not prove the absence of large angular motions; however, a low $P(S_{Nterm}, S_{Cterm})$ value reflects a difference in static structure and/or relative domain motions. A statistically meaningful difference in Gvalue for different domains requires interdomain motion, typically of substantial angular amplitude. For randomly

oriented pairs of alignment tensors with random rhombicity, $[S_i, S_j]$, the P(S_i, S_j) distribution spans the range of 0 to 1, with the relative probability of a P(S_i, S_j) = α value approximately decreasing with cos($2\alpha/\pi$) (SI Fig. S4).

Monomeric SARS-CoV-2 MPro

A SARS-CoV-2 MPro construct, lacking the N-terminal residues S1-P9 that stabilize the dimerization of the native form of the enzyme, and further inactivated by the active site H410 mutation (MPro^{10-306,H41Q}), was expressed in D₂O medium. For multiple β -sheet residues in the N-terminal domain and α -helical residues in the C-terminal domain, back exchange of the amide protons with the protonated solvent was incomplete, even after several days at neutral pH and room temperature, causing multiple backbone amide signals to be very weak or absent. Nevertheless, a sufficiently large set of ${}^{1}D_{NH}$, ${}^{1}D_{C'N}$, ${}^{1}D_{C\alpha C'}$, and ${}^{2}D_{C'H}$ RDCs was obtained (Table S2) that yielded well-defined alignment tensors for each of its six α -helices: one in the N-terminal domain (Y54-I59) and five in the C-terminal domain (R201-I213; L227-Y237; D245-L250; V261-O273; and F294-Q299 (Fig. 4). There are several X-ray structures for monomeric MPro (2QCY (Shi et al. 2008); 2PWX (Chen et al. 2008); and 3F9E (Hu et al. 2009). These all pertain to the highly homologous (~96% identity, Zhang et al. 2011) previous SARS coronavirus isolate and display a different orientation of the C-terminal domain (I200-Q306) relative to the N-terminal catalytic domain from that seen in the native dimeric state (Fig. 4). As was observed for calmodulin, good SVD fits were obtained for RDCS of the individual helices measured in the monomeric state to coordinates of helices in the 1.2-A X-ray structure of the native dimer, 7K3T (Andi et al. 2022) (Fig. 5A-F). On average, slightly lower quality fits and less well-defined alignment tensors, reflected in higher Q_{ik} and $\varepsilon(G)$ values, were obtained for fits to idealized helices (Fig. 5G).

Remarkably, the generalized alignment magnitude, G, is considerably larger for the C-terminal domain than for the Y54-I59 helix in the N-terminal domain, indicative of large amplitude motions of this helix or of the entire N-terminal domain relative to the C-terminal domain (Fig. 5G). Although assignments for the N-terminal domain remain incomplete due to the above mentioned slow back exchange as well as conformational exchange broadening, a large number of yet to be fully identified RDCs was measured for the N-terminal domain. The distributions of normalized RDCs seen in the N-terminal domain appears considerably narrower than for the C-terminal domain (Fig. 6), pointing to a lower alignment strength (Clore et al. 1998a, b). However, the histogram RDC distribution expected for the N-terminal domain when using the alignment tensor obtained for its Y54-I59 helix is somewhat narrower than the observed



Fig. 4 Superposition of MPro subunit taken from the native homodimer X-ray structure, 7K3T (pink), and its R298A monomeric mutant, 2QCY (grey). The six 7K3T helices used for RDC analysis are shown in green

distribution (Fig. 6A), suggesting that the helix orientation undergoes dynamic fluctuations relative to the N-terminal domain.

For the C-terminal domain, its helices show very similar alignment strengths, indicative of a well-ordered domain which is consistent with the slow hydrogen exchange rates seen for many of its backbone amides. Mobility of helix Y54-I59 relative to the C-terminal domain differs strongly from what is observed when evaluating the RDCs previously measured for the homodimeric state (Robertson et al. 2021), which show very similar *G* values and high $P(S_i, S_j)$ values (i,j = 'H1' to 'H6') (Fig. S5).

For RDCs measured here in the monomeric state and for all of the X-ray structures evaluated (2QCY; 2PWX; 3F9E; and dimeric 7K3T), we find a low value of the normalized scalar products, $P(S_i, S_j)$, of the alignment tensor of the Y54-I59 helix (S_i , i = 'H1') relative to those of the C-terminal domain helices (S_j , j = 'H2' to 'H6') (Fig. 5H; SI Fig. S6), confirming that in solution the relative domain orientation differs from that in any of the X-ray structures and presumably is subject to large amplitude dynamics.





Fig. 5 RDC analysis of α-helices in monomeric SARS-CoV-2 MPro^{10–306,H41Q}. Fits of the normalized ¹D_{NH} (red ball), ¹D_{C'N} (green ball), ¹D_{CαC'} (blue ball), and.²D_{HNC'} (pink ball) RDCs to the X-ray coordinates of homodimeric MPro (PDB entry 7K3T) are shown for the following helices: **A** Y54-I59; **B** T201-I213; **CL**227-Y237; **D** D245-L250; **E** V261-Q273; and **F** F294-Q299. The generalized sampling parameter for these couplings, Ξ, is 0.062. **G** *G* values obtained for the six helices when using 7K3T (grey) or idealized helices (orange) as reference structures; the reference <*G*> value of 7.83 obtained for the five C-terminal helices (Helix-2 to Helix-6),

and $\langle G \rangle = 5.34$ obtained for Helix-1 in the N-terminal domain, are plotted as horizontal lines. Q_{jk} values (open circles) correspond to the scale at the right side of the panel. **H** Normalized scalar products $P(S_i, S_j)$ for all pairs of the alignment tensors of six helices, $[S_i, S_j]$ (i,j = "H1" to "H6"), obtained when fitting the helical RDCs to the coordinates of the homodimeric X-ray structure (7K3T) (upper-left half), and ideal helices (lower-right half). For the corresponding $P(S_i, S_j)$ values when using the monomeric MPro X-ray structures (2QCY; 2PWX; 3F9E), see SI Fig. S6

Fig. 6 Histogram distributions of backbone RDCs in monomeric MPro, normalized to ¹D_{NH}. A Experimental RDC distribution for the N-terminal domain in blue, and in red the histogram expected for uniformly distributed vector orientations and the alignment tensor obtained for helix Y54-I59. B Analogous plot of C-terminal domain RDCs, and the expected histogram distribution if vectors were uniformly distributed with an alignment strength and rhombicity obtained from simultaneously fitting all helical RDCs in this domain



For the C-terminal domain helices, the $P(S_i, S_j)$ values (i,j = 'H2' to 'H6') are close to unity when using the 1.2 Å X-ray structure of the homodimer as a reference but, on average, somewhat lower when using the monomer X-ray structures as references (Fig. S6). However, superposition of the C-terminal domain of monomeric MPro X-ray

structures on those of the homodimer shows close correspondence, with a C^{α} RMSD \leq 0.64 Å. Therefore, these somewhat lower P(S_i , S_j) values likely result from slightly less accurate atomic coordinate positions in these monomeric X-ray structures.

Discussion

Our RDC analysis of α -helical regions in proteins takes advantage of the highly regular backbone geometry imposed by the backbone H-bonding pattern in such helices. Consequently, RDCs fit about equally well to idealized α -helices and to the corresponding high-resolution X-ray coordinates, as reflected in comparable Q_{jk} values (Figs. 1B, 3B, 5G). If perturbations in the α -helical H-bonding pattern are present, they often are apparent in strong deviations from the expected secondary chemical shifts, and RDC fitting to idealized α -helical coordinates must be restricted to helices that display canonical ¹³C^{α}, ¹³C^{β}, and ¹³C' secondary chemical shift values, and only to regions that do not include a Pro residue, which invariably kinks the helix axis.

RDC analysis of α -helices can provide insights into the dynamics of a system, even before either a full-fledged structure determination or a ¹⁵N NMR relaxation analysis has been completed. In practice, such an analysis requires that multiple RDCs per peptide group are measured because the uncertainty in alignment tensor parameters obtained for any given helix scales roughly with $1/\sqrt{(N-1)^2}$ 5), where N is the number of measured RDCs per helix and 5 corresponds to the degrees of freedom in the SVD fit. Moreover, the orientational sampling of just N-H vectors in α -helices ($\Xi = 0.83$) typically will be insufficient for deriving accurate alignment tensors (Fushman et al. 2000). Even for the MPro dimer, two RDCs per amide $({}^{1}D_{NH} \text{ and } {}^{2}D_{CH})$ were measured at high precision, despite the large size and slow tumbling of this 68 kD system (Robertson et al. 2021). For the monomeric MPro, which exhibits more favorable relaxation properties, four RDCs were measured per peptide plane $({}^{1}D_{NH}, {}^{1}D_{C'N}, {}^{1}D_{C\alpha C'}, and$ $^{2}D_{HC'}$). Even though, due to the planar geometry, these four couplings are partially redundant (Bryce and Bax 2004), their measurement errors are uncorrelated and the availability of all four couplings improves the reliability of the fits.

Our RDC analysis of maltotriose-ligated MBP showed that within experimental uncertainty, all its helices are subject to the same alignment tensor, confirming that in the ligated state this protein behaves as one globular, well-ordered system. RDC analysis of Ca^{2+} -calmodulin revealed that the normalized scalar products between each of its four individual helices and the alignment tensor obtained for the entire N terminal domain falls well below unity, indicative of the previously identified substantial structural differences between the relative helix orientations of this domain in solution, (Chou et al. 2001) and in the X-ray structure. The calmodulin example also shows that the presence of essentially indistinguishable alignment tensor magnitudes, G, of its N- and C-terminal domains should not be misinterpreted as evidence for the absence of interdomain motion. Presumably as a result of the two domains' very similar shape and charge distribution, their alignments in the Pf1 liquid crystal simply happen to be equally strong even while their relative orientation is highly dynamic (Barbato et al. 1992; Bertini et al. 2004).

Great interest exists in the monomeric structure of MPro because this protein is initially expressed as a polyprotein, and it needs to excise itself from this precursor before it can adopt the catalytically active homodimeric state (Tan et al. 2005; Shi et al. 2008; Nashed et al. 2022). Therapeutic strategies that target its ability to dimerize therefore may be powerful complements to traditional active site inhibitors (Goyal and Goyal 2020; Cantrelle et al. 2021; Berg et al. 2022). Even before assignments and NMR relaxation analysis of the monomeric MPro are complete, RDC analysis points to considerably weaker alignment of the active-site containing N-terminal domain than for the mostly α-helical C-terminal domain. Relative helix orientations in the C-terminal domain are largely consistent with those seen in both monomeric and homodimeric X-ray structures, but the relative orientation of the N- and C-terminal domains clearly does not match any of the X-ray structures, again pointing to interdomain dynamics. The presence of this interdomain motion is perhaps not surprising considering that the globular N-terminal domain retains much of its structure, with largely unchanged chemical shifts when expressed in isolation compared to values seen in the homodimeric state (Zhang et al. 2011; Robertson et al. 2022). Moreover, the linker between the N-and C-terminal domains in the dimeric protein is subject to considerable dynamic disorder, both in solution and in the crystalline state (Ebrahim et al. 2022; Shen et al. 2023). On the other hand, a 50-ns molecular dynamics trajectory revealed no significant changes in relative domain orientation for the monomeric state (Parmar et al. 2023), and a full RDC and ¹⁵N NMR relaxation analysis of monomeric MPro therefore has been initiated in our laboratory.

Methods

RDC data for MBP are from (Evenäs et al. 2001a) and kindly provided to us by Dr. V. Tugarinov. Ca²⁺-calmodulin RDC data, recorded in 15 mg/mL Pf1 medium, were taken from (Chou et al. 2001).

Resonance assignments of the monomeric mutant MPro^{10–306,H41Q} (Nashed et al. 2022; Kovalevsky et al. 2024) were carried out using standard TROSY-based 3D triple resonance methods (Salzmann et al. 1999) on a perdeuterated, partially back-exchanged construct. Data were collected at 800 MHz for a 0.5 mM sample in 25 mM HEPES buffer

at pH 6.9 with 20 mM NaCl, 1 mM TCEP, 3% D_2O and 0.02% NaN₃ at 15 °C. RDC data were collected at 25 °C on a 0.43 mM sample, with 11 mg/mL Pf1 used as the alignment medium. The resonance assignments were transferred from spectra acquired at 15 °C to those at 25 °C using multiple TROSY spectra collected over this temperature range.

 ${}^{1}D_{NH}$ couplings were measured at 800 MHz, using the ARTSY method (Fitzkee and Bax 2010), with RDCs calculated from the cross-peak intensity ratios of two interleaved 2D ¹H-¹⁵N TROSY spectra (Fitzkee and Bax 2010) for a total measurement time of 5 h for isotropic data and the same for the aligned sample. Additionally, two interleaved 3D ¹H-¹⁵N TROSY-HNCO ARTSY spectra (17 h each) were collected at 600 MHz (Chiliveri et al. 2022). ${}^{1}D_{NC'}$ and ${}^{1}D_{C\alpha C'}$ RDCs were measured at 800 and 600 MHz ¹H frequencies, respectively, using 2D and 3D quantitative-J experiments (Chou et al. 2000; Jaroniec et al. 2004). Data recording time was 11 h for ${}^{1}D_{NC'}$; 19 h for ${}^{1}D_{C\alpha C'}$, and similar for the measurement of ${}^{1}J_{NC'}$ and ${}^{1}J_{C\alpha C'}$ under isotropic conditions. For improved precision of the extracted ${}^{1}D_{NC'}$ and ${}^{1}D_{C\alpha C'}$ RDCs, data for the attenuated spectra were collected with six times more scans than the reference spectrum. ${}^{2}D_{C'H}$ RDCs were measured at 600 MHz, using the 3D TROSY-AntiTROSY-Encoded RDC (TATER) method (Robertson et al. 2021). The total recording time was 18 h for each of the isotropic and aligned samples. Data processing was carried out using NMRPipe (Delaglio et al. 1995); assignments were carried out using CCPN-v3.1 software (Skinner et al. 2016), and RDC values were extracted using MATLAB scripts. RDC data used for analysis of the helices are included in SI Table S2 and RDC data for the full protein will be deposited in the BMRB once analysis has been completed.

Further details of the MPro^{10–306,H41Q} assignments and RDC data collection, carried out as part of a search to find the optimal RDC measurement methods for proteins of this intermediate size class, will be presented elsewhere.

Software availability

The Helix-Fit program is available in server format at https:// spin.niddk.nih.gov/bax/nmrserver/dc/helix_fit.html (requiring a generic google login) and will also be made available at NMRBox: https://nmrbox.nmrhub.org/.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10858-024-00456-5.

Acknowledgements This work was supported by the Intramural Research Program of the NIDDK at the National Institutes of Health (Grant DK029046). We thank Dr. Ivan Gagarinov for preliminary assignments of the MPro monomer, Drs. James L. Baber and Jinfa Ying for assistance with instrumentation and software, Dr. Dennis Torchia for helpful discussions, Annie Aniana for assistance with preparation of the MPro^{10–306,H41Q} sample, and Vitali Tugarinov and Lewis Kay for providing the RDCs of MBP.

Author contributions Y.S. developed code and prepared all figures; M.J.S. collected and analyzed MPro RDC data; J.M.L. prepared isotopically enriched MPro samples; A.B. wrote the main manuscript text. All authors reviewed and edited the manuscript.

Funding Open access funding provided by the National Institutes of Health. Funding was provided by National Institute of Diabetes and Digestive and Kidney Diseases (Grant No. DK029046).

Data availability All data is included in the Supporting Information.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

- Akke M, Chazin WJ (2001) An open and shut case. Nat Struct Mol Biol 8:910–912
- Andi B, Kumaran D, Kreitler DF, Soares AS, Keereetaweep J, Jakoncic J, Lazo EO, Shi WX, Fuchs MR, Sweet RM, Shanklin J, Adams PD, Schmidt JG, Head MS, McSweeney S (2022) Hepatitis C virus NS3/4A inhibitors and other drug-like compounds as covalent binders of SARS-CoV-2 main protease. Sci Rep 12:12197
- Babu YS, Bugg CE, Cook WJ (1988) Structure of calmodulin refined at 2.2 a resolution. J Mol Biol 204:191–204
- Barbato G, Ikura M, Kay LE, Pastor RW, Bax A (1992) Backbone dynamics of calmodulin studied by 15N relaxation using inverse detected 2-dimensional NMR spectroscopy. The central helix is flexible. Biochemistry 31:5269–5278
- Berg H, Martin MAW, Altincekic N, Alshamleh I, Bains JK, Blechar J, Ceylan B, de Jesus V, Dhamotharan K, Fuks C, Gande SL, Hargittay B, Hohmann KF, Hutchison MT, Korn SM, Krishnathas R, Kutz F, Linhard V, Matzel T, Meiser N, Niesteruk A, Pyper DJ, Schulte L, Trucks S, Azzaoui K, Blommers MJJ, Gadiya Y, Karki R, Zaliani A, Gribbon P, Almeida MD, Anobom CD, Bula AL, Bütikofer M, Caruso IP, Felli IC, Da Poian AT, de Amorim GC, Fourkiotis NK, Gallo A, Ghosh D, Gomes-Neto F, Gorbatyuk O, Hao B, Kurauskas V, Lecoq L, Li YF, Mebus-Antunes NC, Mompeán M, Neves-Martins TC, Ninot-Pedrosa M, Pinheiro AS, Pontoriero L, Pustovalova Y, Riek R, Robertson AJ, Saad MJA, Treviño MA, Tsika AC, Almeida FCL, Bax A, Henzler-Wildman K, Hoch JC, Jaudzems K, Laurents DV, Orts J, Pierattelli R, Spyroulias GA, Duchardt-Ferner E, Ferner J, Fürtig B, Hengesbach M, Löhr F, Qureshi N, Richter C, Saxena K, Schlundt A, Sreeramulu S, Wacker A, Weigand JE, Wirmer-Bartoschek J, Wöhnert J, Schwalbe H (2022) Comprehensive fragment screening of the SARS-CoV-2 proteome explores novel chemical space for drug development. Angew Chem Int Ed 61:e202205858

- Bertini I, Del Bianco C, Gelis I, Katsaros N, Luchinat C, Parigi G, Peana M, Provenzani A, Zoroddu MA (2004) Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. Proc Natl Acad Sci USA 101:6841–6846
- Bryce DL, Bax A (2004) Application of correlated residual dipolar couplings to the determination of the molecular alignment tensor magnitude of oriented proteins and nucleic acids. J Biomol NMR 28:273–287
- Cantrelle FX, Boll E, Brier L, Moschidi D, Belouzard S, Landry V, Leroux F, Dewitte F, Landrieu I, Dubuisson J, Deprez B, Charton J, Hanoulle X (2021) 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 60:25428–25435
- Chen S, Hu T, Zhang J, Chen J, Chen K, Ding J, Jiang H, Shen X (2008) Mutation of Gly-11 on the dimer interface results in the complete crystallographic dimer dissociation of severe acute respiratory syndrome coronavirus 3C-like protease—crystal structure with molecular dynamics simulations. J Biol Chem 283:554–564
- Chiliveri SC, Robertson AJ, Shen Y, Torchia DA, Bax A (2022) Advances in NMR spectroscopy of weakly aligned biomolecular systems. Chem Rev 122:9307–9330
- Chou JJ, Delaglio F, Bax A (2000) Measurement of 15N-13C' dipolar couplings in medium sized proteins. J Biomol NMR 18:101-105
- Chou JJ, Li S, Klee CB, Bax A (2001) Solution structure of Ca2+calmodulin reveals flexible hand-like properties of its domains. Nat Struct Biol 8:990–997
- Clore GM, Garrett DS (1999) R-factor, free R, and complete crossvalidation for dipolar coupling refinement of NMR structures. J Am Chem Soc 121:9008–9012
- Clore GM, Gronenborn AM, Bax A (1998a) A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. J Magn Reson 133:216–221
- Clore GM, Starich MR, Gronenborn AM (1998b) Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses. J Am Chem Soc 120:10571–10572
- Crivici A, Ikura M (1995) Molecular and structural basis of target recognition by calmodulin. Annu Rev Biophys Biomol Struct 24:85–116
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRpipe—a multidimensional spectral processing system based on Unix pipes. J Biomol NMR 6:277–293
- Ebrahim A, Riley BT, Kumaran D, Andi B, Fuchs MR, McSweeney S, Keedy DA (2022) The temperature-dependent conformational ensemble of SARS-CoV-2 main protease (M-pro). Iucrj 9:682–694
- Evenäs J, Tugarinov V, Skrynnikov NR, Goto NK, Muhandiram R, Kay LE (2001a) Ligand-induced structural changes to maltodextrinbinding protein as studied by solution NMR spectroscopy. J Mol Biol 309:961–974
- Evenäs J, Mittermaier A, Yang DW, Kay L (2001b) Measurement of 13Ca-13Cb dipolar couplings in 15N,13C,2H-labeled proteins: application to domain orientation in maltose binding protein. J Am Chem Soc 123:2858–2864
- Fitzkee NC, Bax A (2010) Facile measurement of H-1-N-15 residual dipolar couplings in larger perdeuterated proteins. J Biomol NMR 48:65–70
- Fushman D, Ghose R, Cowburn D (2000) The effect of finite sampling on the determination of orientational properties: a theoretical treatment with application to interatomic vectors in proteins. J Am Chem Soc 122:10640–10649
- Gardner KH, Zhang XC, Gehring K, Kay LE (1998) Solution NMR studies of a 42 KDa *Escherichia coli* maltose binding protein

 β -cyclodextrin complex: chemical shift assignments and analysis. J Am Chem Soc 120:11738–11748

- Goyal B, Goyal D (2020) Targeting the dimerization of the main protease of coronaviruses: a potential broad-spectrum therapeutic strategy. ACS Comb Sci 22:297–305
- Hansen MR, Mueller L, Pardi A (1998) Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. Nat Struct Mol Biol 5:1065–1074
- Hu TC, Zhang Y, Li LW, Wang KF, Chen S, Chen J, Ding JP, Jiang HL, Shen X (2009) Two adjacent mutations on the dimer interface of SARS coronavirus 3C-like protease cause different conformational changes in crystal structure. Virology 388:324–334
- Jaroniec CP, Ulmer TS, Bax A (2004) Quantitative J correlation methods for the accurate measurement of 13C'-13Cαdipolar couplings in proteins. J Biomol NMR 30:181–194
- Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637
- Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KFA, Becker S, Meiler J, Grubmuller H, Griesinger C, de Groot BL (2008) Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. Science 320:1471–1475
- Lindorff-Larsen K, Best RB, DePristo MA, Dobson CM, Vendruscolo M (2005) Simultaneous determination of protein structure and dynamics. Nature 433:128–132
- Losonczi JA, Andrec M, Fischer MWF, Prestegard JH (1999) Order matrix analysis of residual dipolar couplings using singular value decomposition. J Magn Reson 138:334–342
- Mesleh MF, Lee S, Veglia G, Thiriot DS, Marassi FM, Opella SJ (2003) Dipolar waves map the structure and topology of helices in membrane proteins. J Am Chem Soc 125:8928–8935
- Nashed NT, Aniana A, Ghirlando R, Chiliveri SC, Louis JM (2022) Modulation of the monomer-dimer equilibrium and catalytic activity of SARS-CoV-2 main protease by a transition-state analog inhibitor. Commun Biol 5:1
- Ottiger M, Bax A (1998) Determination of relative N-H-N N-C', C-alpha-C', andC(alpha)-H-alpha effective bond lengths in a protein by NMR in a dilute liquid crystalline phase. J Am Chem Soc 120:12334–12341
- Parmar M, Thumar R, Patel B, Athar M, Jha PC, Patel D (2023) Structural differences in 3C-like protease (Mpro) from SARS-CoV and SARS-CoV-2: molecular insights revealed by molecular dynamics simulations. Struct Chem 34:1309–1326
- Quiocho FA, Spurlino JC, Rodseth LE (1997) Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport chemosensory receptor. Structure 5:997–1015
- Robertson AJ, Courtney JM, Shen Y, Ying JF, Bax A (2021) Concordance of X-ray and AlphaFold2 models of SARS-CoV-2 main protease with residual dipolar couplings measured in solution. J Am Chem Soc 143:19306–19310
- Robertson AJ, Ying JF, Bax A (2022) NMR observation of sulfhydryl signals in SARS-CoV-2 main protease aids structural studies. ChemBioChem 23:e202200471
- Ruckert M, Otting G (2000) Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. J Am Chem Soc 122:7793–7797
- Salzmann M, Wider G, Pervushin K, Senn H, Wuthrich K (1999) TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. J Am Chem Soc 121:844–848
- Sass J, Cordier F, Hoffmann A, Rogowski M, Cousin A, Omichinski JG, Lowen H, Grzesiek S (1999) Purple membrane induced alignment of biological macromolecules in the magnetic field. J Am Chem Soc 121:2047–2055

- Sass H-J, Musco G, Stahl SJ, Wingfield PT, Grzesiek S (2000) Solution NMR of proteins within polyacrylamide gels: diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. J Biomol NMR 18:303–309
- Sass HJ, Musco G, Stahl SJ, Wingfield PT, Grzesiek S (2001) An easy way to include weak alignment constraints into NMR structure calculations. J Biomol NMR 21:275–280
- Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. J Biomol NMR 56:227–241
- Shen Y, Robertson A, Bax A (2023) Validation of ensemble and static X-ray crystal structure representations of SARS-CoV-2 main protease by solution NMR residual dipolar couplings. J Mol Biol 435:168067
- Shi JH, Sivaraman J, Song JX (2008) Mechanism for controlling the dimer-monomer switch and coupling dimerization to catalysis of the severe acute respiratory syndrome coronavirus 3C-like protease. J Virol 82:4620–4629
- Skinner SP, Fogh RH, Boucher W, Ragan TJ, Mureddu LG, Vuister GW (2016) CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. J Biomol NMR 66:111–124
- Spera S, Bax A (1991) Empirical correlation between protein backbone conformation and Ca and Cb 13C nuclear magnetic resonance chemical shifts. J Am Chem Soc 113:5490–5492
- Tan JZ, Verschueren KHG, Anand K, Shen JH, Yang MJ, Xu YC, Rao ZH, Bigalke J, Heisen B, Mesters JR, Chen KX, Shen X, Jiang HL, Hilgenfeld R (2005) pH-dependent conformational flexibility of the SARS-CoV main proteinase (Mpro) dimer: molecular dynamics simulations and multiple X-ray structure analyses. J Mol Biol 354:25–40
- Tjandra N, Bax A (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science 278:1111–1114
- Tolman JR, Flanagan JM, Kennedy MA, Prestegard JH (1995) Nuclear magnetic dipole interactions in field-oriented

proteins—information for structure determination in solution. Proc Natl Acad Sci USA 92:9279–9283

- Tycko R, Blanco FJ, Ishii Y (2000) Alignment of biopolymers in strained gels: a new way to create detectable dipole-dipole couplings in high-resolution biomolecular NMR. J Am Chem Soc 122:9340–9341
- Wang Y, An L, Yang Y, Yao L (2020) Generating five independent molecular alignments for simultaneous protein structure and dynamics determination using nuclear magnetic resonance spectroscopy. Anal Chem 92:15263–15269
- Wilson MA, Brunger AT (2000) The 1.0 A crystal structure of Cabound calmodulin: an analysis of disorder and implications for functionally relevant plasticity. J Mol Biol 301:1237–1256
- Wishart DS, Sykes BD, Richards FM (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. J Mol Biol 222:311–333
- Wüthrich K (1986) NMR of proteins and nucleic acids. Wiley, New York
- Yang DW, Venters RA, Mueller GA, Choy WY, Kay LE (1999) TROSY-based HNCO pulse sequences for the measurement of (HN)-H- 1-N-15, N-15-(CO)-C-13, (HN)-H-1-(CO)-C-13, (CO)-C-13-C- 13(alpha) and (HN)-H-1-C-13(alpha) dipolar couplings in N-15, C-13, H-2-labeled proteins. J Biomol NMR 14:333–343
- Zhang SN, Zhong N, Ren XB, Jin CW, Xia B (2011) 1H, 13C and 15N resonance assignments of SARS-CoV main protease N-terminal domain. Biomol NMR Assign 5:143–145

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.