


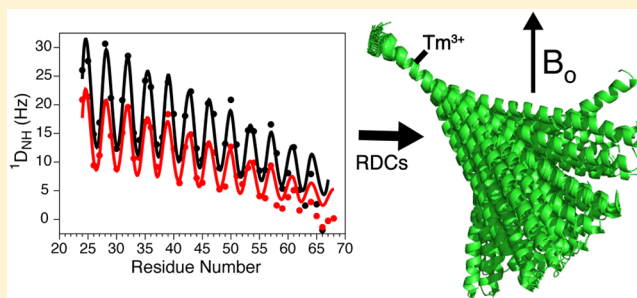
# Remarkable Rigidity of the Single $\alpha$ -Helical Domain of Myosin-VI As Revealed by NMR Spectroscopy

C. Ashley Barnes,<sup>†</sup> Yang Shen,<sup>†</sup> Jinfa Ying,<sup>†</sup> Yasuharu Takagi,<sup>‡</sup> Dennis A. Torchia,<sup>†</sup> James R. Sellers,<sup>‡</sup> and Ad Bax<sup>\*†</sup> 

<sup>†</sup>Laboratory of Chemical Physics, NIDDK, and <sup>‡</sup>Laboratory of Molecular Physiology, NHLBI, National Institutes of Health, Bethesda, Maryland 20892-0520, United States

## Supporting Information

**ABSTRACT:** Although the  $\alpha$ -helix has long been recognized as an all-important element of secondary structure, it generally requires stabilization by tertiary interactions with other parts of a protein's structure. Highly charged single  $\alpha$ -helical (SAH) domains, consisting of a high percentage (>75%) of Arg, Lys, and Glu residues, are exceptions to this rule but have been difficult to characterize structurally. Our study focuses on the 68-residue medial tail domain of myosin-VI, which is found to contain a highly ordered  $\alpha$ -helical structure extending from Glu-6 to Lys-63. High hydrogen exchange protection factors (15–150), small (ca. 4 Hz)  $^3J_{\text{HNH}\alpha}$  couplings, and a near-perfect fit to an ideal model  $\alpha$ -helix for its residual dipolar couplings (RDCs), measured in a filamentous phage medium, support the high regularity of this helix. Remarkably, the hydrogen exchange rates are far more homogeneous than the protection factors derived from them, suggesting that for these transiently broken helices the intrinsic exchange rates derived from the amino acid sequence are not appropriate reference values.  $^{15}\text{N}$  relaxation data indicate a very high degree of rotational diffusion anisotropy ( $D_{\parallel}/D_{\perp} \approx 7.6$ ), consistent with the hydrodynamic behavior predicted for such a long, nearly straight  $\alpha$ -helix. Alignment of the helix by a paramagnetic lanthanide ion attached to its N-terminal region shows a decrease in alignment as the distance from the tagging site increases. This decrease yields a precise measure for the persistence length of  $224 \pm 10 \text{ \AA}$  at 20 °C, supporting the idea that the role of the SAH helix is to act as an extension of the myosin-VI lever arm.



## ■ INTRODUCTION

Highly charged single  $\alpha$ -helix (SAH) domains are found in a significant fraction (0.2–0.5%) of the human protein database.<sup>1</sup> As previously pointed out,<sup>2–4</sup> due to their high contents of charged and polar residues and the scarcity of X-ray structure data for such motifs, they originally were believed to form coiled-coil structures based upon predictions of secondary structure programs. Unlike normal  $\alpha$ -helices, which require tertiary interactions (hydrogen bonding, hydrophobic, disulfide, and/or van der Waals) to maintain their helical structure, SAH domains are stabilized by a repeating pattern of four negatively (glutamate, E) and four positively (arginine, R, and/or lysine, K) charged residues that form dynamic intrahelical salt bridges, while the relatively long alkyl side chains of these residues provide stabilizing hydrophobic interactions. This unique,  $E_4(R/K)_4$  feature of the peptide sequence can confer a much higher degree of stability across a wide range of temperatures, pH, and ionic strength than seen in, for example, poly-Ala helices.<sup>5–7</sup>

Although dispersed throughout the human genome, there are just a few examples of SAH domains isolated in solution, including those of caldesmon<sup>8</sup> (~160 residues) and those from the mechano-enzymes, myosin-VI,<sup>9</sup> myosin-X,<sup>2</sup> and myosin-VIIA.<sup>10,11</sup> Functionally, within these proteins, the SAH

domains are believed to either serve as rigid spacers between protein domains, such as seen in caldesmon, or act as an extension of the lever arms in the myosin family of proteins. The extension of the lever arm is critical in some myosins.<sup>12</sup> The lever arm of the well-studied myosin-V has six calmodulin bound, which allows the two motor domains to simultaneously bind to actin monomers separated by 36 nm that corresponds to the half helical repeat of the actin filament. The lever arm of myosin-VI only has one calmodulin bound, but single molecule stepping studies demonstrate that it can also take steps consistent with a motor domain separation of 36 nm. This observation suggests that the extended SAH domain helps to provide the additional lever arm length in myosin-VI. Charged SAH domains are also encountered as structural elements in otherwise intrinsically disordered proteins such as MFAP1, where they appear to regulate the repeated conformational remodeling of the spliceosome, required for activity of this molecular machine.<sup>13</sup> They can also function as interprotein “interaction motifs”, where the distinct electrostatic SAH surface can play an important role, while the degree of intrahelical ER/K stabilization can fine-tune the strength of the

Received: March 21, 2019

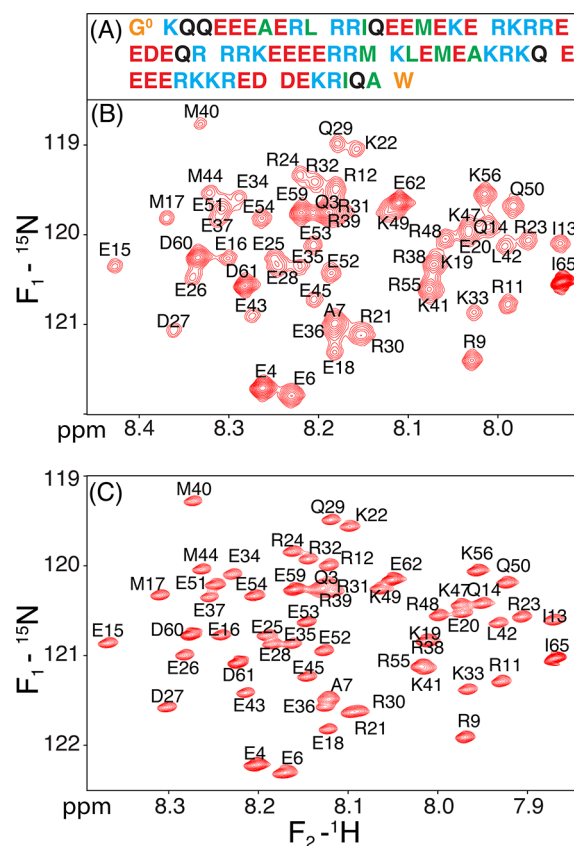
Published: May 22, 2019

intermolecular interaction.<sup>13</sup> The remarkable spring-like properties of charged SAH rods also hold strong potential for applications in protein design and engineering.<sup>14–18</sup>

Despite extensive biochemical and biophysical studies of SAH domains, atomic level structural characterization of such helices is very sparse. To date, there is only one X-ray structure of an SAH helix in isolation,<sup>10</sup> and very few in the context of larger protein structures or structural complexes. The latter invariably raise the question to what extent the observed helical structures are impacted or stabilized by the lattice environment in which they are observed. For example, the (non-SAH) central helix of calmodulin contains two fully solvent-exposed turns of  $\alpha$ -helix in its X-ray structure,<sup>19</sup> whereas in solution this region of the central helix is fully disordered.<sup>20</sup> Solution X-ray scattering (SAXS) experiments on the medial tail domain (MT) of myosin-VI reveal a length of  $\sim 10$ – $15$  nm, fully consistent with an SAH structure.<sup>9,21</sup> Two distinct kinks in this helix, deduced from the SAXS data, correlate with the positions of small hydrophobic patches interrupting the pattern of  $E_4(R/K)_4$  residues. However, whereas the SAXS data suggested these kinks to be static, molecular dynamics calculations pointed to transient unfolding of the helical turns comprising these patches.<sup>22</sup>

Studies of SAH domains by NMR spectroscopy have proven challenging. The very high contents of Glu, Lys, and Arg residues, all in the  $\alpha$ -helical region of Ramachandran space, lead to very poor resonance dispersion and cause many ambiguities in the conventional triple resonance assignment procedure. More than a decade ago, the 35-residue proximal tail region of myosin 10, containing a 26-residue SAH segment, was studied by NMR.<sup>2</sup> However, although the NOEs and chemical shifts clearly indicated a strong  $\alpha$ -helical signature for the SAH segment, resonance overlap precluded a residue-specific analysis. Very recently, however, an in-depth study of the myosin-VIIA SAH domain was reported by the same research group,<sup>11</sup> providing strong evidence for the  $\alpha$ -helical structure of this element, and, most importantly, providing quantitative information on the dynamic nature of individual salt bridges that are important for its stabilization. Remarkably, the  $\alpha$ -helical structure appears not to be disrupted by a five-residue insert near its midpoint that deviates from the canonical ER/K sequence pattern.<sup>11</sup>

Here, we present a detailed multinuclear NMR characterization of the structure and dynamics of the 68-residue MT domain of *S. scrofa* (pig) myosin-VI (Figure 1A), comprising residues 918–985, which have been shown to adopt an SAH structure.<sup>9,21</sup> Resonance dispersion of the  $^{13}C^\alpha$  resonances is found to be particularly poor, as expected on the basis of the sequence composition and high  $\alpha$ -helicity of the structure.  $^{15}N$  and  $^1H^N$  resonance line widths are much larger than usual for a 68-residue construct because nearly all amide vectors are approximately parallel to the axis of the  $\sim 100$ -Å helical rod, causing their effective rotational correlation times to be long. However, fairly well-resolved  $^1H$ – $^{15}N$  maps can be obtained through full perdeuteration of the nonexchangeable protons and utilization of the TROSY line-narrowing effect.<sup>23</sup> Use of four-dimensional (4D)  $^1H$ – $^{15}N$ – $^{15}N$ – $^1H$  NOESY spectra at 900 MHz  $^1H$  frequency then enables full backbone amide resonance assignments. This strategy takes advantage of the  $\alpha$ -helical nature of the structure, which gives rise to strong sequential  $^1H^N$ – $^1H^N$  NOEs, and allows a straightforward link to the corresponding  $^{13}C^\alpha$  chemical shifts through standard TROSY-based triple resonance NMR,<sup>24</sup> thereby providing



**Figure 1.** NMR spectra of myosin-VI E68W MT domain. (A) Sequence of the construct used, with residue types color coded as follows: Arg and Lys in blue; Glu and Asp in red; polar, neutral Gln in black; hydrophobic Ala, Leu, Ile and Met in green; and non-native N- and C-terminal residues in orange. (B) Most crowded region of the regular Rance-Kay HSQC spectrum of perdeuterated, amide-protonated  $^{15}N$ -enriched MT,<sup>72</sup> recorded at 900 MHz  $^1H$  frequency, 20 °C, pH 6.3, 2 mM EDTA, yielding broad resonances. (C) TROSY-HSQC spectrum,<sup>23</sup> recorded under identical conditions.

important residue-specific secondary structure information that complements the NOE distance information.

With full amide resonance assignments in hand, characterization of individual H-bonds by measurement of their protection against amide proton exchange (HX) with solvent provides direct access to the fraction of time such bonds are present.<sup>25</sup> Although chemical shift dispersion of the aliphatic  $^{13}C$  and  $^1H$  resonances was far too low to permit measurement of sufficient experimental NOE restraints to define the SAH structure by conventional NOE-based methods, a full set of  $^1H^N$ – $^1H^N$  NOEs complemented by  $^{15}N$ – $^1H^N$ ,  $^{15}N$ – $^{13}C'$ , and  $^{13}C'$ – $^1H^N$  RDCs could be measured in Pf1 alignment medium and sufficed to define the structure. Moreover, a nearly complete set of  $^3J_{HNH\alpha}$  couplings provided restraints for the  $\varphi$  backbone torsion angles.

Classical  $^{15}N$  relaxation analysis is adversely impacted by the near-parallel alignment of the  $^{15}N$ – $^1H$  bond vectors. However, when supplemented by  $^{13}C$ -detected  $^{13}C^\alpha$  relaxation rates, the expected high degree of rotational diffusion anisotropy could be quantified. Importantly, chelation of a lanthanide tag near the N-terminal end of the SAH helix<sup>26</sup> created strong local alignment of the helix relative to the magnetic field. Analogous to the pioneering work of Bertini et al.,<sup>27</sup> the decrease in alignment as a function of distance from the tagging site then

provides direct access to the persistence length of the helix. Together, these data provide a detailed characterization of both the structure and the backbone dynamics of the SAH domain embedded in the myosin-VI MT, revealing a highly regular, kink-free helical structure that extends from residue E6 to E62, with only a moderate degree of increased dynamics near the ends of the helix.

## METHODS

**Expression and Purification.** The cDNA of the native MT domain of *S. scrofa* (pig) myosin-VI, containing the SAH, was fused to the IgG domain B1 of Protein G (GB1) and cloned into the pET24a vector by Genscript (Piscataway, NJ), which subsequently performed site-directed mutagenesis to create MT variants (see below). Specifically, the N-terminus of SAH was spliced to a TEV cleavage site (ENLYFQG) added to the C-terminus of GB1 (GB1-TEV-SAH). To facilitate Ni:NTA purification, the GB1 domain included a 6xHis site at its N-terminus. The GB1-TEV-MT containing vector was transformed into *E. coli* BL21(DE3) cells and grown to an OD of  $\sim 0.7$  at 37 °C in M9 medium before induction with 1 mM of IPTG. Expression of the GB1-TEV-MT construct (3 h at 37 °C) was followed by centrifugation at 4000 rpm for 20 min at 4 °C. The cells were then resuspended in 50 mM Tris-HCl, 500 mM NaCl, 20 mM imidazole, pH 7.4, including cComplete Roche protease inhibitor cocktail (Buffer A), and passed once through an Avestin Emulsiflex C3 homogenizer (ATA Scientific; Taren Point, Australia). The resulting lysate was then centrifuged for 1 h at 20 000 rpm and 4 °C. Next, the supernatant was filtered through a 0.22 mm filter before passage through a 5 mL GE His-trap column, equilibrated with buffer A, and eluted using a step gradient of 3% and 30% buffer B (buffer A plus 1 M imidazole), respectively. Fractions containing GB1-TEV-MT were collected and, if needed, concentrated to a volume of  $\sim 5$ –7 mL before injection onto a Superdex 26/60 G75 column equilibrated with 25 mM Tris-HCl, 100 mM NaCl, 0.5 mM EDTA/TCEP, pH 8. To isolate the MT domain, TEV protease was added to the combined GB1-TEV-SAH fractions at a TEV:protein ratio of 1:50. After the reaction mixture was left for  $\sim 16$ –18 h at room temperature, it was passed again through a 5 mL GE His-trap column using the aforementioned protocol. Purified MT was obtained in the flow-through. For NMR experiments, the MT domain was concentrated and buffer-exchanged into the buffers noted in the text. The construct used for our NMR studies has the following sequence, G K<sup>1</sup>QQEEEAERL RRIQEEMEKE RKRREDEQQR RRKEEEERRM KLEMEAKRKQ EEEERKKRED DEKRIQAE/W, where K<sup>1</sup> corresponds to K<sup>917</sup> in *S. scrofa* (pig) myosin-VI, which differs from *H. sapiens* by a K945Q substitution. The non-native N-terminal Gly is a remnant of the TEV cleavage sequence, and the C-terminal Glu was substituted by Trp during later stages of the project to facilitate quantification by optical density. The total mass of the polypeptide, assuming natural isotopic abundance, is 8972 Da (9029 Da for E68W).

For generating the requisite <sup>15</sup>N-, <sup>13</sup>C-, and <sup>2</sup>H-isotopically enriched samples, <sup>15</sup>N-ammonium chloride, <sup>2</sup>H/<sup>13</sup>C D-glucose, or <sup>2</sup>H D-glucose were used as the sole sources of nitrogen, carbon, and hydrogen. Additionally, for generating perdeuterated sample, deuterium oxide (99% D<sub>2</sub>O, Cambridge Isotope Laboratories, Boston, MA) was used in place of H<sub>2</sub>O in the expression medium and supplemented with 1 g/L of ISOGRO of the desired isotopic composition (Sigma-Aldrich).

For tagging MT with 4R,4S-DOTA-M8, loaded with either thulium (Tm) or lutetium (Lu), purified I13C MT domain was buffer-exchanged into 100 mM Tris-HCl pH 7.8 using a PD10 G25 column (GE Healthcare). The concentration of MT, post-PD10, was  $\sim 20$   $\mu$ M. Immediately after buffer exchange,  $\sim 20$ -fold molar excess of 4R,4S-DOTA-M8-Lu/Tm was added to the MT domain solution. The reaction mixture was then incubated for  $\sim 18$ –20 h at room temperature before it was passed twice through a PD10 column to remove excess tag and exchange to NMR buffer. Liquid chromatography–electrospray ionization–mass spectrometry (LC–ESI–MS)

was used to verify the identify and purity of the I13C MT-DOTA-Lu/Tm. No impurities or unlabeled I13C MT were observed in this LC–ESI–MS analysis.

**Circular Dichroism Measurements.** CD spectra were recorded in 20 mM sodium phosphate buffer, 2 mM EDTA, pH 6.3, on a JASCO J-810 spectropolarimeter using a 0.1 cm path length quartz cuvette. Thermal melt experiments were carried out by recording the CD signal at 222 nm as a function of increasing temperature from 0 to 80 °C at 0.5 °C/min. Using the CD spectra from 190–260 nm, the  $\alpha$ -helical content was derived using the CDNN program.<sup>28</sup>

**Backbone Assignment Experiments.** All experiments used for backbone assignments were in 20 mM sodium phosphate buffer, 2 mM EDTA, pH = 6.3, 20 °C. The 4D <sup>1</sup>H–<sup>15</sup>N–<sup>15</sup>N–<sup>1</sup>H HMQC-NOESY-TROSY-HSQC spectrum was pivotal in the assignment procedure and was carried out on a 900-MHz Bruker Avance-III spectrometer equipped with a z-gradient TCI cryogenic probe, on a 1 mM solution of <sup>2</sup>H/<sup>15</sup>N MT in 92% H<sub>2</sub>O, 8% D<sub>2</sub>O, and assignments of the TROSY-HSQC spectrum subsequently were transferred to different conditions mostly by stepwise titration. To generate sufficient spectral resolution in the available measurement time (2.5 days), nonuniform sampling was employed,<sup>29</sup> with acquisition times of 39.7 ms (<sup>1</sup>H,<sub>t<sub>1</sub></sub>), 38.4 ms (<sup>15</sup>N,<sub>t<sub>2</sub></sub>), 128 ms (<sup>15</sup>N,<sub>t<sub>3</sub></sub>), and 170 ms (<sup>1</sup>H,<sub>t<sub>4</sub></sub>), with a sparsity of 0.49% prior to extension of the time domain data. A total of four scans per FID was used to permit phase cycling for axial peak suppression. Spectral reconstruction was carried out with the SMILE program,<sup>30</sup> yielding a final digital resolution of 6.7 Hz (<sup>1</sup>H,<sub>F<sub>1</sub></sub>), 7.1 Hz (<sup>15</sup>N,<sub>F<sub>2</sub></sub>), 1.6 Hz (<sup>15</sup>N,<sub>F<sub>3</sub></sub>), and 2.2 Hz (<sup>1</sup>H,<sub>F<sub>4</sub></sub>). Additionally, a 3D TROSY-HNCACB spectrum was recorded at 700 MHz <sup>1</sup>H frequency on a 1 mM sample of <sup>2</sup>H/<sup>15</sup>N/<sup>13</sup>C MT. All data were processed using NMRPipe,<sup>31</sup> and assignments were made by means of analysis software included with NMRPipe, principally scroll.tcl and NMRDraw.

**Backbone Dynamics.** Heteronuclear TROSY-based relaxation experiments using previously described pulse sequences<sup>32</sup> were carried at 600 and 800 MHz <sup>1</sup>H frequency on a 1 mM solution of <sup>2</sup>H/<sup>15</sup>N MT at 20 °C, to obtain <sup>15</sup>N longitudinal ( $R_1$ ), transverse ( $R_{1\rho}$ ), and heteronuclear <sup>15</sup>N{<sup>1</sup>H} NOE relaxation data.  $R_{1\rho}$  measurements were carried out using a spin-lock field strength of 2.5 kHz (calibrated according to the power level needed to obtain a 100  $\mu$ s 90° pulse width), and reported values were corrected for off-resonance effects using  $R_{1\rho} = (R_{1\rho}' - \sin^2 \varphi R_1) / \cos^2 \varphi$ , where  $R_{1\rho}'$  is the experimentally measured value,  $R_{1\rho}$  is the on-resonance value, and  $\varphi = \tan^{-1}(\delta/\nu_{RF})$ , with  $\delta$  being the resonance offset and  $\nu_{RF}$  the strength of the spin lock field, both in units of Hz.<sup>33</sup> Analysis of the relaxation data was carried out using the Modelfree4 program,<sup>34</sup> downloaded from the NYSBC Web site ([www.NYSBC.org](http://www.NYSBC.org)).

**<sup>3</sup>J<sub>HNH $\alpha$</sub>  Measurement.** Residue-specific values for <sup>3</sup>J<sub>HNH $\alpha$</sub>  were obtained at 35 °C and 900 MHz from a fully protonated, <sup>15</sup>N-enriched sample (1 mM) of MT. Two ARTSY-J experiments<sup>35</sup> were carried out, using <sup>3</sup>J<sub>HNH $\alpha$</sub>  dephasing durations,  $\tau_d$ , of 30 and 50 ms. Using the relation  $^3J_{\text{HNH}\alpha} = (1 + 0.206R_{1\alpha}\tau_d)J_{\text{apparent}}$ <sup>35</sup> comparison of the two sets of  $J_{\text{apparent}}$  values (Figure S1) yielded  $R_{1\alpha} = 5.7 \text{ s}^{-1}$ . Values reported here (Table S1) include the correction relative to the values measured for  $\tau_d = 30$  ms, and are 3.5% larger than  $J_{\text{apparent}}$  at  $\tau_d = 30$  ms.

**Hydrogen Exchange.** HX rates were measured at 900 MHz <sup>1</sup>H frequency at both 20 and 30 °C, using a 1 mM sample of perdeuterated, <sup>15</sup>N-enriched MT sample in 95% H<sub>2</sub>O, 5% D<sub>2</sub>O, 25 mM sodium phosphate buffer, pH 7.8. Rates were measured using the WEX-III TROSY experiment,<sup>36</sup> using acquisition durations of 155 ms ( $t_1$ , <sup>15</sup>N) and 122 ms ( $t_2$ , <sup>1</sup>H<sup>N</sup>), and seven durations for the water inversion interval, ranging from 5 ms to 1 s.

**Residual Dipolar Couplings.** RDCs were measured under various different alignment conditions. Two sets were collected where the MT alignment was induced by immersing it in a dilute liquid crystalline suspension of Pfl, purchased from ASLA Biotech, Riga, Latvia. The first sample contained 13 mg/mL Pfl in 25 mM sodium phosphate buffer, pH 6.3, 2 mM EDTA; the second sample contained an additional 100 mM NaCl. Pfl was purchased as a concentrated stock solution, which due to its high viscosity made

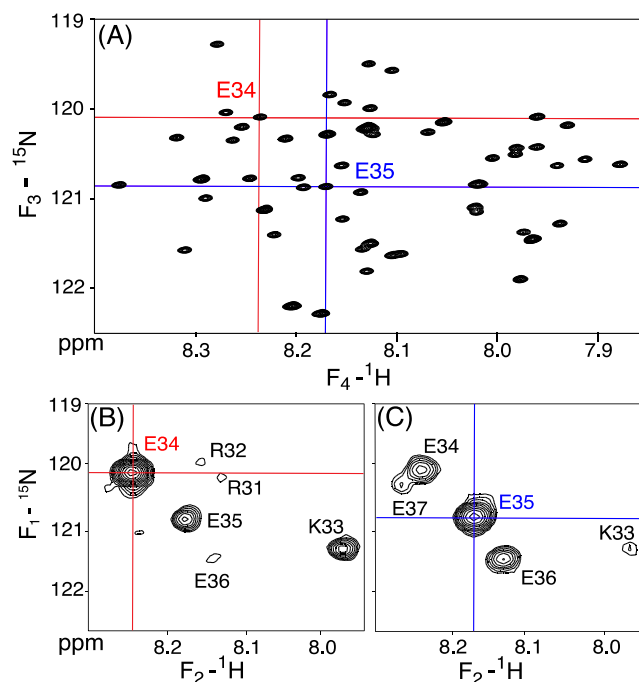
precise pipetting challenging. Reported Pfl sample concentrations therefore were derived from the observed  $^2\text{H}$  lock solvent splitting, using the approximation that a lock splitting of  $p$  Hz corresponds to  $p$  mg/mL Pfl.<sup>37</sup> A third set of RDCs was acquired using paramagnetic alignment, in a solution containing 25 mM sodium phosphate buffer, pH 6.3, 2 mM EDTA. This latter set required generation of two samples, with either a paramagnetic (Tm) or diamagnetic (Lu) lanthanide chelated to a DOTA-M8 metal-binding tag, engineered and kindly provided to us by Dr. D. Haussinger.<sup>26</sup> The DOTA-M8 was attached to Cys-13 of a I13C mutant of the MT construct. For all samples,  $^1\text{H}$ - $^{15}\text{N}$  RDCs were obtained at 900 MHz  $^1\text{H}$  frequency using the ARTSY experiment,<sup>38</sup> which derives the RDCs from intensity ratios in two TROSY-HSQC spectra, recorded in an interleaved manner. Additionally, for the sample with added salt, one-bond  $^{13}\text{C}'$ - $^{15}\text{N}$  and two-bond  $^{13}\text{C}'$ - $^1\text{H}^{\text{N}}$  RDCs were derived from the difference in splitting observed between two 900 MHz E.COSY-HSQC-TROSY spectra,<sup>39</sup> recorded with and without Pfl on a uniformly  $^{15}\text{N}/^{13}\text{C}/^2\text{H}$ -MT sample.

## RESULTS

### Resonance Assignment and Secondary Structure.

The highly repetitive MT sequence (Figure 1A) and its very high fraction (72%) of Arg, Glu, and Lys residues, which have nearly the same random coil chemical shifts, and therefore very similar  $^{13}\text{C}^{\alpha}$  chemical shifts when embedded in  $\alpha$ -helical secondary structure, caused numerous ambiguities when following the conventional triple resonance assignment protocol. Moreover, we observed rapid transverse relaxation of both the amide  $^1\text{H}$  and the  $^{15}\text{N}$  resonances, much faster than anticipated for the small size of the MT domain studied here (69 residues), resulting in broad, weak resonances (Figure 1B). The rapid transverse relaxation results from the high anisotropy of MT's rotational diffusion tensor, where the amide vectors are aligned approximately parallel to the helical axis and thereby experience the slowest rotational diffusion.<sup>11</sup> However, TROSY-HSQC measurements carried out at 900 MHz dramatically improve the attainable spectral resolution (Figure 1C) and remove the large variation in intensities that results from modest fractional differences in transverse relaxation rates. The anticipated high helicity of the MT domain<sup>22</sup> (>90% based on CD measurements; Figure S2) resulted in relatively strong and quite uniform sequential amide–amide NOE connectivities,  $d_{\text{NN}}$ . This feature, combined with improved resolution attained through TROSY detection and the use of nonuniform sampling (NUS), enabled the recording of highly resolved 4D amide–amide NOESY spectra on a perdeuterated, amide-protonated sample (Figure 2), which in addition to strong sequential,  $d_{\text{NN}}(i, i \pm 1)$ , connectivities also showed numerous weaker  $d_{\text{NN}}(i, i \pm 2)$  and even  $d_{\text{NN}}(i, i \pm 3)$  NOEs, permitting straightforward and unambiguous linking of all amides throughout the entire helical section of the MT construct. Assignments of the most N- and C-terminal residues were confirmed by recording a TROSY-HNCACB spectrum, which also yielded  $^{13}\text{C}^{\alpha}$  and some  $^{13}\text{C}^{\beta}$  chemical shifts. Large secondary  $^{13}\text{C}^{\alpha}$  chemical shifts confirm the uniformity of the helical structure of the MT domain, with a gradual reduction in secondary chemical shifts toward the termini of the MT domain, indicative of increased helical fraying (Figure 3A).

Despite the rather fast transverse relaxation of the  $^1\text{H}^{\text{N}}$  resonances, even in TROSY-type spectra, and the small values of  $^3J_{\text{HNH}\alpha}$  couplings in helical structure, reliable values for this coupling were measured at 35 °C using the recently introduced ARTSY-J experiment.<sup>35</sup> This experiment relies on the

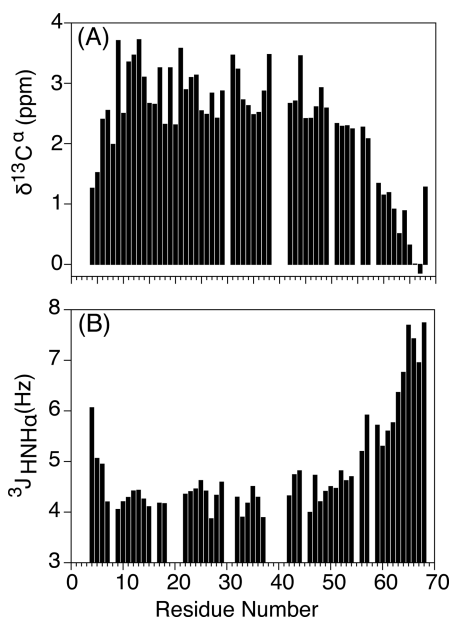


**Figure 2.** Depiction of the residue assignment protocol for the MT domain using a 4D amide–amide HMQC-NOESY-TROSY-HSQC spectrum (900 MHz; 250 ms NOE mixing time; 20 °C), recorded with nonuniform (0.49% sparsity) sampling. (A) Projection of the 4D spectrum on the  $(F_3, F_4)$  TROSY-HSQC plane. The chemical shift scales have been adjusted to make them consistent with the HSQC cross sections by adjustment for the  $^1J_{\text{NH}}/2$  displacement of cross peak positions in TROSY-HSQC spectra. (B,C) Two-dimensional  $(F_1, F_2)$  cross sections, orthogonal to the  $(F_3, F_4)$  planes, at the chemical shift positions of the amide resonances of (B) E34, marked by the red cross hairs, and (C) E35, marked by the blue cross hairs. Cross peaks to  $i \pm 1$  and  $i \pm 2$  are marked by residue type and number. E34 in (B) also displays a weak cross peak with  $i - 3$  residue R31.

comparison of the amide signal TROSY intensity in the presence and absence of  $^3J_{\text{HNH}\alpha}$  dephasing. The measurement was carried out for dephasing durations of 30 and 50 ms, yielding very close agreement between the two sets of data (Figure S1). These  $^3J_{\text{HNH}\alpha}$  values (Figure 3B) are highly consistent with the corresponding secondary  $^{13}\text{C}^{\alpha}$  chemical shifts and confirm the uninterrupted helical structure of the MT domain, explaining the exceptionally high mean residue ellipticity at 222 nm of ca.  $-3500 \text{ deg}\cdot\text{cm}^2/\text{dmol}$  at 20 °C (Figure S2).

**Structure of MT.** While initially the MT domain of myosin-VI was believed to be a dimeric coiled coil,<sup>40,41</sup> its monomeric nature was subsequently established by multiple-angle light scattering.<sup>9</sup> Indeed, indistinguishable resonance positions and line widths in our TROSY-HSQC spectra recorded at concentrations of 10  $\mu\text{M}$  to 1 mM do not show any sign of intermolecular interactions (Figure S3A). With the uninterrupted  $\alpha$ -helical structure of MT firmly established from  $^{13}\text{C}^{\alpha}$  chemical shifts and  $^3J_{\text{HNH}\alpha}$  couplings, RDCs provide a unique and unambiguous route to study the global structure of the MT domain and probe for the presence of potential bends or kinks.

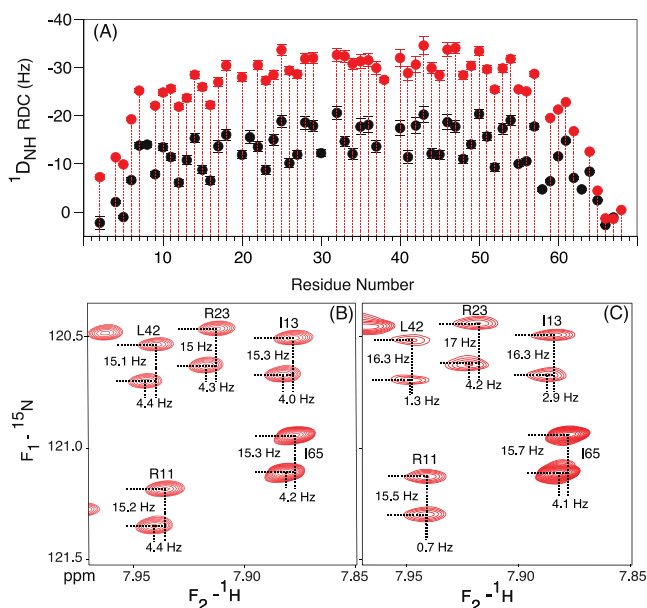
As anticipated, alignment of the highly charged MT domain in a liquid crystalline suspension of filamentous bacteriophage, Pfl, is impacted by its extreme deviation from spherical shape,



**Figure 3.** Residue-specific NMR parameters reporting on secondary structure of MT. (A)  $^{13}\text{C}^\alpha$  secondary chemical shifts. Secondary chemical shifts are relative to the Poulsen random coil values,<sup>73</sup> corrected for the effect of deuteration.<sup>74</sup> (B)  $^3J_{\text{HNH}\alpha}$  couplings measured at 35 °C. Values were measured at 900 MHz  $^1\text{H}$  frequency using the ARTSY-J experiment, and corrected for the effect of  $R_1(\text{H}^\alpha)$ .<sup>35</sup> The secondary chemical shifts and  $^3J_{\text{HNH}\alpha}$  were measured at 35 °C, for a construct that includes E68.

which results in strong steric alignment. Additionally, the alignment is impacted by electrostatic interaction with the highly negatively charged Pf1 phage. This latter interaction can be reduced by increasing the ionic strength of the buffer, and measurements therefore were carried out at both our standard condition (20 mM sodium phosphate, 2 mM EDTA, pH 6.3) and the same buffer with addition of 100 mM NaCl. SAH domains are known to be little impacted by ionic strength at salt concentrations below ca. 1 M.<sup>4</sup> Indeed, the NMR spectrum is minimally affected by the addition of 100 mM NaCl (Figure S3B), confirming that the structure is not altered within our detection limits, while making the spectrum more amenable to measuring RDCs at high precision.

Measurement of  $^1D_{\text{NH}}$  RDCs in 13 mg/mL Pf1 shows considerably larger couplings in the sample containing the additional 100 mM NaCl than the measurement in standard buffer, confirming that the MT alignment is impacted by both steric and electrostatic interactions with the filamentous phage. In the sample containing 100 mM NaCl, the sinusoidal oscillation in the size of the RDC, commonly referred to as a dipolar wave pattern,<sup>42</sup> is of small amplitude and nearly vanishes for the center region (E28-E37; Figure 4A), as expected when the helical axis becomes nearly coincident with the symmetry axis of a (nearly) axially symmetric alignment tensor (Supporting Information text). In fact, the small amplitudes of the dipolar wave oscillations on both the N- and the C-terminal sides of this center region already foreshadow that the helix axis does not contain any substantial kinks or bends. In the absence of the additional salt, the alignment of the MT domain is somewhat weaker, and the center of the helix orients at a small angle (of ca. 18°) relative to the z-axis of the now slightly more rhombic alignment tensor (Table S8). In the presence of salt, alignment was



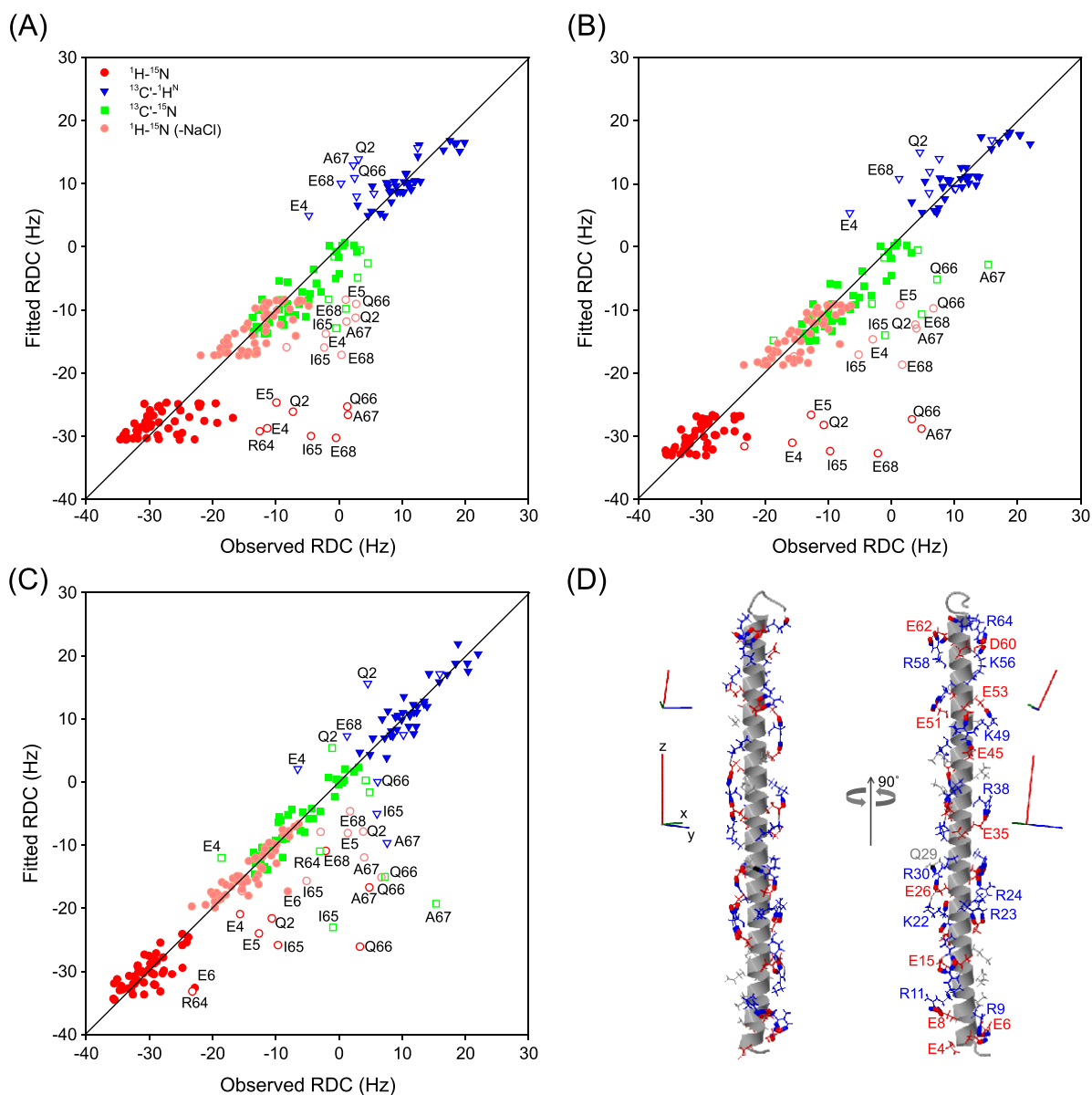
**Figure 4.** Measurement of RDCs at 900 MHz, 20 °C. (A)  $^1D_{\text{NH}}$  values measured in the presence of 13 mg/mL Pf1 in 20 mM sodium phosphate, 2 mM EDTA, pH 6.3 (black), and upon addition of 100 mM NaCl (red) on a 1 mM sample of perdeuterated  $^{15}\text{N}/^{13}\text{C}$  enriched MT. Plotted values ignore the negative sign of  $\gamma(^{15}\text{N})$ ; i.e., the plotted negative values correspond to  $^1J_{\text{NH}} + ^1D_{\text{NH}}$  splittings smaller than 90 Hz. (B,C) Small regions of the TROSY-HSQC-E-COSY spectra on (B) the isotropic and (C) the aligned sample, illustrating the measurement of  $^1D_{\text{NC}'}$  and  $^2D_{\text{CH}}$  couplings in the sample containing 100 mM NaCl. RDCs are listed in Table S2.

sufficiently strong to permit measurement of the smaller  $^1D_{\text{C}'\text{N}}$  and  $^2D_{\text{C}'\text{H}}$  RDCs, in addition to  $^1D_{\text{NH}}$ . The relative precision at which these smaller couplings can be measured from a HSQC-TROSY-E-COSY spectrum<sup>39</sup> is somewhat lower than that for  $^1D_{\text{NH}}$  RDCs (Figure 4B,C), but nevertheless these couplings serve as important complements to  $^1D_{\text{NH}}$  as they strongly increase the width of the orientational distribution of the sampled RDC vectors.

With a  $Q$ -factor of 21%, the RDCs of residues E6-K63 fit quite well to an idealized  $\alpha$ -helical structure (Figure 5A), but deviations exceed the errors in the RDC measurements by at least 2-fold, meaning that some deviations from an ideal  $\alpha$ -helix are present.

It is well recognized that RDCs are impacted by internal motions,<sup>43,44</sup> a fact that has been exploited extensively to quantify the amplitude and direction of internal motions, integrated over time scales ranging from pico- to micro-seconds.<sup>45</sup> For moderate amplitude isotropic internal motion, the experimentally detected RDCs, to first order, scale with the size of the generalized local order parameter  $S$ . In the absence of slower, micro- or millisecond internal motions, as applies for MT (vide infra),  $S$  is equivalent to the square root of the Lipari–Szabo<sup>46</sup> generalized order parameter,  $S^2$ , commonly extracted from  $^{15}\text{N}$  relaxation analysis.

Although for globular proteins, the amplitude of internal motions across residues involved in secondary structure is relatively uniform, and correction for the amplitude of motions in the calculation of an accurate, time-averaged structure is not needed, for the MT domain the situation is somewhat more complex. As discussed below,  $^{15}\text{N}$   $R_2$  rates monotonically decrease when approaching the ends of the helical domain,



**Figure 5.** Results of SVD fits of experimental RDCs to an idealized  $\alpha$ -helical structure with backbone torsion angles of  $\varphi = -62.5^\circ$ ,  $\psi = -42.5^\circ$ ,  $\omega = 180^\circ$ . Only the RDCs for residues E6–R63 were used to carry out the fit, but the correlation is shown for all observed couplings.  $^1D_{\text{NC}'}^{\text{H}}$  and  $^2D_{\text{CH}}^{\text{N}}$  were upscaled by multiplication with the inverse of the respective dipolar interaction constants (2609 and 6962 Hz)<sup>75</sup> relative to  $^1D_{\text{NH}}$  (21 585 Hz), ignoring the effect of the sign of the  $^{15}\text{N}$  gyromagnetic ratio on the RDC, thereby ensuring that normalized RDCs of the same value correspond to the same orientational restraint. Thus, the uniformly negative values of  $^1D_{\text{NH}}$  in the plot correspond to  $|^1J_{\text{NH}} + ^1D_{\text{NH}}| < 92$  Hz. RDCs were measured in the presence of 13 mg/mL Pf1, 100 mM NaCl, except for the light red  $^1D_{\text{NH}}$  values (no NaCl). (A) SVD fit of the RDCs (scaled for the interaction constants). (B) SVD fit of the RDCs after correction for the variation in  $^{15}\text{N}$   $R_2$  value, as discussed in the text, to (B) the idealized  $\alpha$ -helical structure and (C) the XPLOR-NIH structure, refined with the  $R_2$ -corrected RDCs. (D) Two orthogonal views of the refined structure, with the backbone displayed as a ribbon diagram, and the corresponding alignment frames depicted next to them (top, no NaCl; bottom, with 100 mM NaCl); side chains are shown in full atom representation, but are restrained solely by the empirical force field terms (tDB and eefxpot) in the XPLOR-NIH<sup>48</sup> structure calculation. Glu and Asp are in red; Lys and Arg in blue; and gray for all other residues. Salt bridges and H-bonds between the side chains are drawn as thick lines.

indicative of increased amplitudes of motion. To first order,  $^{15}\text{N}$   $R_2$  equals  $S^2 \times J(0)$ , where  $J(0)$  is the spectral density at zero frequency. The possibility that this decrease could be caused by a gradual bending of the helix relative to an approximately axially symmetric rotational diffusion tensor is excluded by the consideration that  $^{15}\text{N}-^1\text{H}$  vectors are pointing away from the helical axis by about  $14^\circ$ . Therefore, a bending of the helix axis would cause a modulation of the  $^{15}\text{N}-^1\text{H}$  angle, as a function of residue number, relative to the

diffusion tensor, and thereby introduce a periodic oscillation of  $J(0)$ , which is contrary to observation (vide infra).

We therefore make a first-order correction by scaling all RDCs observed for residue  $i$  by the square root of  $R_2(\text{center})/R_2(i)$ , where  $R_2(\text{center})$  equals the average  $^{15}\text{N}$   $R_2$  rate for the center 15 residues (residues E28–L42). Importantly, this correction of the RDCs for the effect of internal dynamics results in improved fits of these scaled RDCs to an idealized  $\alpha$ -helix (Figure 5B), as well as improved cross validation statistics

for a structure derived de novo from RDCs, in particular for residues near the ends of the SAH domain (Table S3).

Together with the information derived from hydrogen exchange measurements (vide infra) that the amides of A7–I65 are engaged in  $\alpha$ -helical H-bonds, allowing the use of an empirical potential-of-mean-force (PMF),<sup>47</sup> these scaled RDCs then serve as the primary restraints in a calculation of the structure using the XPLOR-NIH program.<sup>48</sup> As expected, the resulting structure is highly regular, lacks any significant kinks, and only shows a very minor degree of bending along its ca. 90 Å long axis (Figure 5D). With a  $Q_{\text{free}}$  of 12.9%, agreement between the observed and best-fitted RDCs to this “static” structure (Figure 5C) is remarkably good (Table S3). It is worth pointing out that this  $Q_{\text{free}}$  factor is a true cross validation measure: Four sets of structures were calculated, leaving out all RDCs for residues  $m + n \times 4$ , with  $m = 5, 6, 7$ , or 8, and  $n = 0, \dots, 14$ .  $Q_{\text{free}}$  values were then calculated on the basis of the differences between observed RDCs and those not used in each of the four sets of structures, while using the alignment tensor parameters obtained for a structure where all RDCs were included. Slightly lower (i.e., better)  $Q_{\text{free}}$  values can be obtained by randomly, cyclically omitting a smaller fraction of the RDCs.<sup>49</sup> However, with four RDCs per peptide group of fixed, planar covalent geometry, the issue of whether the obtained  $Q$  value then represents a true cross validation statistic can be debated, and we therefore opted against this latter approach and chose to use the more challenging statistic.

As mentioned above and discussed in more detail below, our finding that, on average, the helical axis is nearly straight does not mean that it does not undergo substantial time-dependent deviations from this average. Below, we first discuss the amplitude of local internal backbone dynamics, followed by a discussion of the impact of such motions on the persistence length of the helix.

**Relaxation-Derived Backbone Dynamics of MT.** The amplitude and time scale of internal backbone motions are first evaluated by standard analysis of  $^{15}\text{N}$   $R_1$ ,  $R_2$ , and  $^{15}\text{N}\{^1\text{H}\}$  NOE relaxation rates measured at 800 and 600 MHz  $^1\text{H}$  frequency (18.8 and 14.1 T, respectively) while accounting for anisotropic rotational diffusion.<sup>50,51</sup> Near the midpoint of the MT domain, its  $R_1$  rates at 18.8 T are low and approach  $0.5 \text{ s}^{-1}$ , while, at a value of ca.  $40 \text{ s}^{-1}$ , the  $R_2$  rates are exceptionally high for a construct of only 69 residues. A plot of the  $R_{1\rho}$  rates measured at 18.8 T against the 14.1 T values yields a straight line with a slope of 1.19 and no significant outliers (Figure S4). This slope is very close to the ratio expected for a process dominated by  $J(0)$  spectral density, a field-independent dipolar interaction constant of 21.6 kHz, and a  $^{15}\text{N}$  chemical shift anisotropy of 170 ppm. This result therefore excludes the presence of significant exchange contributions,  $R_{\text{ex}}$  to the observed transverse relaxation rates. We note that a spherically shaped protein with a volume equal to that of the MT domain (ca.  $10 \text{ nm}^3$ ) with an additional 0.3 nm layer to account for surface hydrogens and the hydration layer<sup>52,53</sup> has a calculated overall correlation time of about 6 ns at 20 °C, and  $R_2/R_1$  ratios of 4.4 and 6.9 at 600 and 800 MHz, respectively. The  $R_2/R_1$  ratios of MT are more than 10 times larger, foreshadowing that its overall diffusion is highly anisotropic.<sup>50,51</sup>

Unfortunately, the very narrow orientational distribution of N–H vectors within a single, nearly straight helix is insufficient to determine the anisotropy of a rotational diffusion tensor,<sup>54</sup> and we therefore first aim to measure the  $^{13}\text{C}^\alpha$  relaxation rates

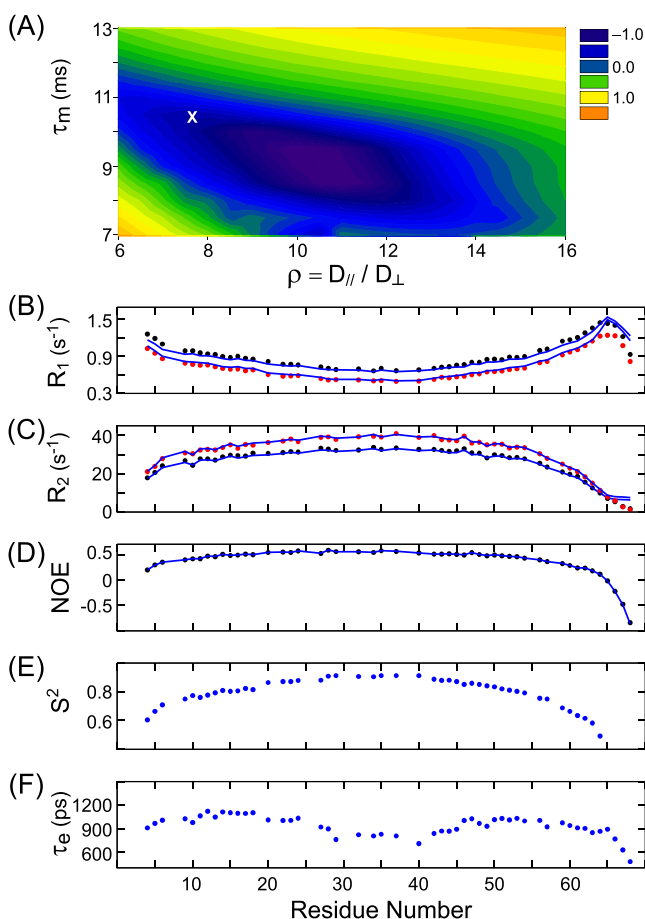
and  $^{13}\text{C}^\alpha\text{--}\{^1\text{H}^\alpha\}$  NOE values in a fully protonated,  $^{13}\text{C}/^{15}\text{N}$ -labeled MT sample. The  $^{13}\text{C}^\alpha\text{--}^1\text{H}^\alpha$  vectors make an angle of ca.  $65^\circ$  with the helix axis, and therefore  $^{13}\text{C}^\alpha$  relaxation is dominated by rotation of MT around its long axis, that is, by the  $D_{\parallel}$  component of the diffusion tensor. This contrasts with  $^{15}\text{N}$   $R_2$  relaxation, which is dominated by  $D_{\perp}$ .

Site-specific measurements of  $^{13}\text{C}^\alpha$  relaxation rates, which, at least in principle, could be carried out by transfer of  $^{13}\text{C}^\alpha$  magnetization via  $^{13}\text{C}'$  to  $^{15}\text{N}$ , followed by amide  $^1\text{H}$  detection, are challenging for the MT domain due to the fast transverse relaxation of  $^{13}\text{C}'$ , yielding low sensitivity. Instead, we resorted to direct  $^{13}\text{C}$  observation, where the  $^{13}\text{C}^\alpha$  resonances of the helical Glu, Lys, and Arg residues are completely overlapped (Figure S5A). The intense, downfield shoulder of this unresolved cluster of resonances has the largest secondary  $^{13}\text{C}^\alpha$  chemical shifts and therefore is representative of the most ordered residues in the MT helix. Indeed, its  $R_1$  ( $1.4 \text{ s}^{-1}$ ) and  $R_2$  ( $38 \text{ s}^{-1}$ ) values (Figure S5) are consistent with rapid reorientation of the  $^{13}\text{C}^\alpha\text{--}^1\text{H}^\alpha$  vector. Considering that  $R_2$  rates in proteins are dominated by  $J(0)$  spectral density and that the prefactor from the one-bond dipolar interaction combined with the effect of chemical shift anisotropy is about 3.5-fold larger for  $^{13}\text{C}^\alpha$  than for  $^{15}\text{N}$ , for isotropic diffusion one would expect  $R_2(^{13}\text{C}^\alpha) \approx 3.5 \times R_2(^{15}\text{N}) \approx 120 \text{ s}^{-1}$  instead of the observed, more than 3-fold lower value, pointing to extreme rotational diffusion anisotropy.

To quantify the rotational diffusion anisotropy, we first carried out a model-free<sup>46</sup> analysis of the center section of the MT domain, whose RDCs were least impacted by internal dynamics and whose structure is precisely known from the above RDC analysis.

Relaxation data analysis was carried out using the Model-free4 software package,<sup>34</sup> which permits restriction of diffusion tensor anisotropy values,  $\rho = D_{\parallel}/D_{\perp}$ , or fixation of the mean correlation time,  $\tau_m = 1/(2D_{\parallel} + 4D_{\perp})$ . A systematic search for the applicable rotational diffusion tensor was carried out using the relaxation rates of the most ordered, center part of the MT domain (residues E28–L42). Perhaps surprisingly, this search reveals that there is a substantial range of  $\rho$  and  $\tau_m$  combinations that can fit the  $^{15}\text{N}$  and  $^{13}\text{C}^\alpha$  relaxation rates to within experimental error (Figure 6A), with a shallow minimum centered near  $\rho = 10.5$ ,  $\tau_m = 9 \text{ ns}$ , but the entire blue region of the plot satisfies the relaxation data well within experimental error. These  $\rho$  and  $\tau_m$  values are in agreement with results of hydrodynamic calculations for a straight  $\alpha$ -helix of 60 residues (90 Å) in length and a radius of 6 Å, where the radius has been derived assuming a protein density of  $1.4 \text{ g/cm}^3$ . Adding the standard 3 Å surface layer to account for the sum of the “bare-atom” correction and the hydration surface layer,<sup>52</sup> this yields a hydrated cylinder of 96 Å in length and 18 Å in diameter (Supporting Information text). Hydrodynamic modeling calculations by Garcia de la Torre and Bloomfield<sup>53,55</sup> then yield a  $\tau_m$  value of 10.3 ns, and a diffusion anisotropy  $\rho = 7.6$ , fully consistent with those obtained by the Model-free4 fit<sup>34</sup> to our experimental data.

With the global diffusion tensor established above, the full set of relaxation rates derived for the entire MT construct was used to evaluate the amplitude and time scales of motions within the frame of this diffusion tensor. We found that the “simple model-free” fitting<sup>46</sup> that is, ascribing a single time scale to the internal motions, can adequately fit the experimental data (Figure 6B–D). Near the middle of the MT domain, the fitting results yielded  $S^2$  values that exceed 0.9



**Figure 6.** Analysis of MT backbone dynamics from  $^{15}\text{N}$  relaxation data using the Modelfree4 program.<sup>34</sup> (A) Grid search over the mean rotational correlation time  $\tau_m$  and the diffusion anisotropy  $\rho = D_{\parallel}/D_{\perp}$  for the minimum normalized  $\chi^2 = \{\sum_{n=28, \dots, 42} [(R_{1,n}^{800, \text{obs}} - R_{1,n}^{800, \text{fit}})^2 / \epsilon_1^2 + (R_{1,n}^{600, \text{obs}} - R_{1,n}^{600, \text{fit}})^2 / \epsilon_1^2 + (R_{2,n}^{800, \text{obs}} - R_{2,n}^{800, \text{fit}})^2 / \epsilon_2^2 + (R_{2,n}^{600, \text{obs}} - R_{2,n}^{600, \text{fit}})^2 / \epsilon_2^2 + (\text{NOE}_n^{600, \text{obs}} - \text{NOE}_n^{600, \text{fit}})^2 / \epsilon_3^2 + (R_{1,C}^{600, \text{obs}} - R_{1,C}^{600, \text{fit}})^2 / \epsilon_4^2 + (R_{2,C}^{600, \text{obs}} - R_{2,C}^{600, \text{fit}})^2 / \epsilon_5^2 + (\text{NOE}_C^{600, \text{obs}} - \text{NOE}_C^{600, \text{fit}})^2 / \epsilon_6^2\} / N$  function when fitting the relaxation rates of residues  $n = 28-42$  to the simple modelfree formalism with axially symmetric overall diffusion, using the coordinates of the RDC-refined MT domain.  $R_{1,n}$ ,  $R_{2,n}$  and  $\text{NOE}_n$  refer to the longitudinal and transverse  $^{15}\text{N}$  relaxation rates of residue  $n$ , and its  $^{15}\text{N}\{^1\text{H}\}$  NOE, whereas  $R_{1,C}$ ,  $R_{2,C}$  and  $\text{NOE}_C$  are the corresponding average rates measured for the overlapping group of E/R/K/Q residues in Figure S5 and include the entire helix.  $N$  corresponds to the total number of fitted experimental constraints, and the colors correspond to  $\log(\chi^2)$ . Uncertainties in the experimental  $R_1$  rates ( $\epsilon_1$ ) and  $R_2$  rates ( $\epsilon_2$ ) were estimated to be 3% of their measured values; the error in the NOE was  $\epsilon_3 = 0.03$  based on the signal-to-noise in the acquired spectra, and  $^{13}\text{C}$  errors were estimated at  $\epsilon_4 = 0.03 \text{ s}^{-1}$ ,  $\epsilon_5 = 1 \text{ s}^{-1}$ , and  $\epsilon_6 = 0.01$ . The blue region of the plot marks the range of  $\tau_m$  and  $\rho$  where the data can be fit to within their experimental uncertainty ( $\log(\chi^2) \leq 0$ ). (B–D) Agreement between experimental (B)  $R_1$ , (C)  $R_2$ , and (D)  $^{15}\text{N}\{^1\text{H}\}$  NOE data (black symbols, 600 MHz; red symbols, 800 MHz) and corresponding values predicted by Modelfree (solid blue lines) when using  $\tau_m = 10.3 \text{ ns}$ ;  $\rho = 7.6$  (position X in panel A), and the fitted residue-specific order parameters  $S^2$  (E) and internal correlation times  $\tau_e$  (F) obtained by Modelfree4.

(Figure 6E), and therefore are comparable to what is found in well-structured globular proteins. When moving closer to the N- and C-termini of the helix, a gradual decrease in  $S^2$  is observed, consistent with somewhat smaller RDCs, discussed

above. The effective correlation times for the internal motions derived by Modelfree4, ca. 1 ns (Figure 6F), are considerably longer than the typical values of 50–100 ps seen in globular proteins. We interpret such motions as undulations in the straightness of the helical rod, with the amplitude relative to the inertia frame naturally increasing (i.e., decreasing  $S^2$ ) when approaching the ends of the helix. Such motions clearly are expected to span a range of time scales, but as stipulated in the modelfree approach,<sup>46</sup> a distribution of these times cannot be extracted uniquely from the available experimental data and therefore is substituted by a single effective value. The amplitudes of the fluctuations of the N–H vector orientations, reflected in  $S^2$ , in this analysis can be interpreted as resulting from local fluctuations in the orientation of the helix axis relative to its average, nearly straight structure. If interpreting this motion as diffusion in a cone of half angle  $\theta_0$ ,  $S^2 \approx 0.6$  values near the ends of the helix correspond to  $S \approx 0.775$ , or  $\theta_0 \approx 33^\circ$ .

**Rotational Diffusion from  $^1\text{H}$ – $^1\text{H}$  NOE.** The rotational diffusion of MT is also encoded in the  $^1\text{H}^{\text{N}}$ – $^1\text{H}^{\text{N}}$  NOESY spectrum of fully perdeuterated, amide-protonated protein. In particular, the ratios of the sequential cross-peak to diagonal peak intensity,  $I_{i\pm 1}/I_d$ , prove useful for this purpose. Growth of NOE cross peaks decreases with mixing time, making the buildup of NOE intensity nonlinear in the region where high intensity cross peaks are obtained. On the other hand, diagonal intensity decreases with mixing time. As a consequence, the ratio of diagonal to cross peak intensity remains close to linear over a much wider range of mixing times than the cross-peak buildup itself (Supporting Information text; Figure S6A). Importantly, our 4D NOESY spectrum of this sparse  $^1\text{H}$  spin system provides well-resolved access to both the diagonal and the cross peaks and provides a simple alternative to recent work aimed at making the relation between NOE and interproton distance more quantitative.<sup>56</sup>

The buildup of NOE cross peak intensities between protons  $i$  and  $j$  is proportional to the  $^1\text{H}$ – $^1\text{H}$  cross-relaxation rate,  $\sigma_{ij}$ . This rate,  $\sigma_{ij}$ , is a function of (a) the overall rotational diffusion tensor of the helix, (b) the internuclear  $\text{H}_i$ – $\text{H}_j$  distance, and (c) the angle of the  $\text{H}_i$ – $\text{H}_j$  internuclear vector relative to the long axis of the helix. In addition, an order parameter  $S_{\text{HH}}^2$  that accounts for internal motion of the  $\text{H}_i$ – $\text{H}_j$  internuclear vector scales  $\sigma_{ij}$ . The NOESY spectrum was recorded using a long mixing time (0.25 s) to ensure adequate signal/noise of the cross-peak intensities. Its cross-peak intensities then can be readily compared to those calculated for an idealized  $\alpha$ -helix (Supporting Information text). Spin-diffusion effects were accounted for by including cross-relaxation involving seven sequential spins in the simulations of the cross-peak intensities. For a mixing time of 250 ms, the calculation using the above derived diffusion tensor yields a value of  $I_{i\pm 1}/I_d = 0.33$ , in fair agreement with experimentally observed ratios of  $0.26 \pm 0.03$  for residues E6–K63 (Figure S6). The result that the calculated value is somewhat larger than that measured is attributed to internal motions, that is,  $S_{\text{HH}}^2 < 1$ , similar to what is found for the N–H order parameter. The experimentally observed decrease in  $I_{i\pm 1}/I_d$  ratios as residue positions recede from the center of the helix (Figure S6B) closely parallels the behavior of  $^{15}\text{N}$   $R_2$  and  $S^2$  (Figure 6C,E). Although for MT, the decrease in sequential  $^1\text{H}^{\text{N}}$ – $^1\text{H}^{\text{N}}$  NOE when approaching the terminal regions is foreshadowed by the  $^{15}\text{N}$  relaxation analysis, ignoring that the increased amplitude of backbone dynamics results in only very small distance errors: a 20%



decrease in  $^1\text{H}^{\text{N}}-^1\text{H}^{\text{N}}$  NOE leads to an overestimate of  $r_{\text{HH}}$  by as little as 0.1 Å. The low degree of sensitivity of interproton distance to internal dynamics can be seen as a strength but also poses challenges when aiming to derive quantitative dynamic information from a joint structural and dynamical analysis of NMR data,<sup>57</sup> an issue that also has resulted in much debate regarding the quantitative analysis of RDCs in terms of dynamics.<sup>45,58</sup>

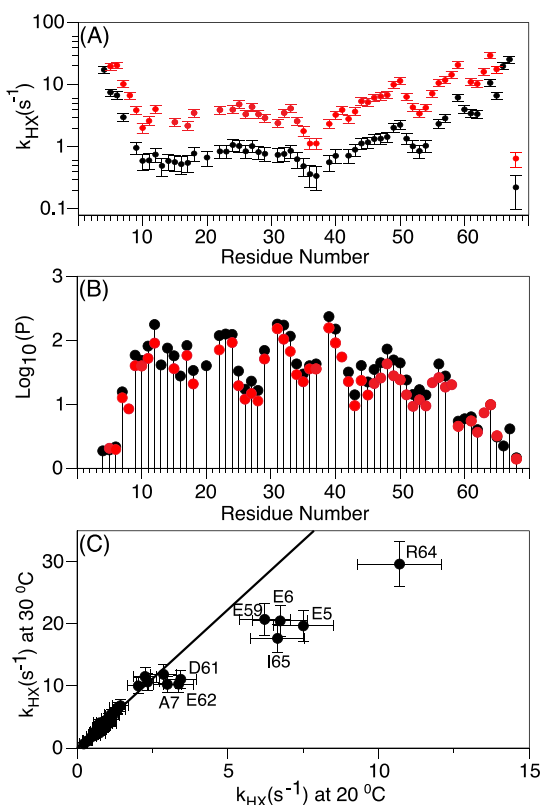
**Backbone Amide Hydrogen Exchange.** Exchange of a protein backbone amide hydrogen with solvent (HX) requires that it is not involved in a H-bond. For a protein, where H-bonds transiently break due to dynamic fluctuations, the ratio of the rate at which an amide exchanges with solvent and that rate in a disordered short peptide of the same linear sequence (intrinsic HX rate) is commonly interpreted as quantitative information about the fraction of time the H-bond is broken.<sup>59</sup> The inverse of this ratio is often referred to as the protection factor,  $P$ . We have measured the hydrogen exchange rates at two temperatures, at a pH value of 7.8 where most of the exchange rates fall in the 0.5–20  $\text{s}^{-1}$  range, which allows for their convenient and accurate measurement using the WEX-III TROSY hydrogen exchange experiment (Figure 7A).<sup>36</sup> Measured  $P$  values fall in the range of 15–150 for residues A7–I65, and somewhat lower values closer to the ends of the SAH domain (Figure 7B). Consistent with the  $^{15}\text{N}$  relaxation

data, the RDCs, and the  $^3J_{\text{HNH}\alpha}$  couplings, protection factors less than two for residues E4–E6 and Q66–E68 indicate that the amide hydrogens of these residues are less than 50% of the time engaged in  $\alpha$ -helical H-bonds. The amide of A7 ( $P \approx 15$ ) is expected to make an  $\alpha$ -helical H-bond to the carbonyl oxygen of Q3, and the HX measurements therefore indicate that the helix initiates at Q3, and terminates at I65 ( $P \approx 3$ ), with increased internal dynamics near the termini of this SAH region.

A caveat regarding the precise interpretation of  $P$  values is that they are referenced to random coil values, which only account for the residue type of the amide considered and that of its preceding neighbor while ignoring the effect of more distant residues. Considering the substantial effects of nearby charged groups on the local  $\text{OH}^-$  concentration, which dominates the HX rates at the near-neutral pH used in our study, uncertainty in the actual intrinsic HX rates may impact the true  $P$  values by non-negligible amounts. A second, and potentially even larger problem with using intrinsic, random coil HX rates as the reference for the not H-bonded state could result from the fact that the newly solvent-exposed amide at a transient break in an  $\alpha$ -helix then is situated at the N-terminus of a newly formed  $\alpha$ -helical fragment. The large electric dipole moment of this fragment elevates the electrostatic potential at the position of its solvent-exposed N-terminal amide, thereby increasing the  $\text{OH}^-$  concentration and the associated rate of base-catalyzed HX. Thus, it is possible that the solvent exchange rate of transiently exposed amides is considerably higher than the intrinsic HX rates in short, unstructured peptides. Our experimentally measured rates provide some support for this hypothesis as the residue-by-residue variation along the sequence is far smaller than the variation in the protection factor,  $P$  (Figure 7).

The temperature dependence of the HX rates has been widely used to probe the nature of the unfolding process in globular proteins.<sup>60</sup> For MT, a plot of HX rates measured at 20 °C versus those at 30 °C (Figure 7C) shows a remarkably linear correlation for all but the terminal residues, with a slope of 4.5, considerably larger than the factor of 2.5 that applies for the intrinsic HX rate.<sup>59</sup> Only the residues near the ends of the SAH domain of MT, specifically E5–A7 and D61–I65, show a ratio intermediate between the random coil value of 2.5 and the factor of ca. 4.5 observed for the other SAH residues.

For a global unfolding process, an increase in temperature or denaturant concentration is expected to result in the same fractional change of the measured HX rates,<sup>60</sup> consistent with what we observe for MT. At first sight, this result therefore suggests that the unfolding of the MT domain is highly cooperative. Frequently,  $P$  is interpreted as an apparent difference in free energy,  $\Delta G_i$ , between the H-bonded and open, HX-prone state of H-bond  $i$ :  $\Delta G_i = RT \ln(P)$ . The apparent activation energy for  $\text{OH}^-$  catalyzed HX, as applies at the near neutral pH used in our study, corresponds to 71 kJ/mol,<sup>59</sup> and is responsible for the ratio of 2.5 in intrinsic HX rate, that is, those not involved in intramolecular H-bonds. The  $4.5/2.5 = 1.8$ -fold larger ratio yields a total activation energy of 113 kJ/mol, which represents the sum of the activation energies for breaking the H-bonds plus intrinsic HX, or an activation energy of ca. 42 kJ/mol for generating the exchangeable, non-H-bonded state. This activation energy for unfolding is comparable to what is seen for several small proteins, and together with the observation that the temperature dependence of the H-bond opening is highly uniform



**Figure 7.** Backbone amide hydrogen exchange of the MT domain in 20 mM sodium phosphate, 2 mM EDTA, pH 7.8. (A) HX rates as a function of residue number at 20 °C (black) and 30 °C (red). (B) HX rates converted to protection factors,  $P$ , by dividing them by the corresponding intrinsic exchange rate number at 20 °C (black) and 30 °C (red). (C) Correlation graph between the hydrogen exchange rates at 20 and 30 °C. The slope of the correlation, when excluding the outlying, labeled terminal residues, is 4.5. For amides that are not H-bonded, the expected slope is 2.5.

indeed could point to a high degree of cooperativity in the folding of the SAH domain.

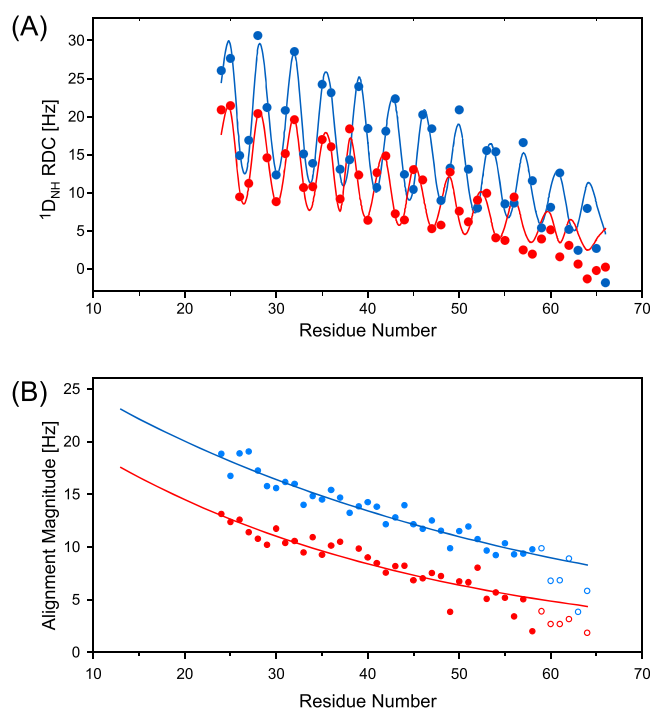
However, the observation that the HX rates scale with temperature in a uniform manner across nearly the entire SAH region of the MT domain does not mean that opposite ends of the SAH region are directly coupled in the transient unfolding process that enables HX. Considering that the sequence is relatively uniform in its composition, one might expect relatively homogeneous behavior of the temperature dependence of backbone amide exchange. On the other hand, the absence of significant variations on a smaller scale, where sequence composition differences are much more pronounced (Figure 1A), suggests that, at a minimum on this smaller length scale, there is a substantial degree of cooperativity in the unfolding of such regions.

**Persistence Length from Paramagnetically Induced RDCs.** A key question in the functioning of the myosin-VI MT focuses on a quantitative understanding of whether this domain has sufficient stiffness to function as a lever arm extension. For this purpose, it is important to have an unambiguous measurement of its persistence length,  $L_p$ . By treating an SAH domain as a worm-like chain, comparisons between observed and predicted small-angle X-ray scattering (SAXS) patterns have been used to provide experimental bounds on  $L_p$ . However, such measurements are also impacted by the structural model used for the SAH domain, which in some cases included significant kinks.<sup>9</sup> Interpretation of SAXS data in terms of a  $L_p$  value becomes more unambiguous once the contour length of the helix (ca.  $1.5 \times N \text{ \AA}$ , where  $N$  is the number of residues in the helix) is larger than  $L_p$ , a condition that does not apply for myosin-VI MT. SAXS measurements on the much longer Kelch SAH domain (196 residues) yielded a definitive  $L_p = 150 \text{ \AA}$  for this domain, a value consistent with the conclusion of a relatively rigid MT domain in myosin-VI, drawn from optical trapping experiments.<sup>21</sup> However, we note that there are considerable deviations from the canonical ER/K sequence pattern near the midpoint of the Kelch SAH domain, which could reduce its stiffness relative to myosin-VI MT. We therefore set out to provide a more direct measurement of the  $L_p$  of MT by measurement of its RDCs under paramagnetically induced alignment.

As first demonstrated by Bertini and co-workers for calmodulin,<sup>27</sup> very weak magnetic alignment of one part of a molecule, introduced by a tightly bound paramagnetic lanthanide ion with a substantial magnetic susceptibility anisotropy, will transfer to other parts of a molecule if rigidly linked, but in an attenuated manner if the linker is flexible. This approach represents a uniquely quantitative, nonperturbing method for measuring both the average orientation and the degree to which a remote domain is ordered relative to the domain that is directly aligned by the paramagnetic ion.

In our study, we used the DOTA-M8 tag introduced by Haussinger and co-workers,<sup>26</sup> which very tightly chelates lanthanide ions and can be conveniently attached to any surface-exposed cysteine. For MT, which lacks native cysteines, we chose residue I13 for mutation to cysteine, thereby minimally impacting the salt-bridges that stabilize the SAH domain, while causing the strongest alignment to be close to one end of the helix. Importantly, the  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts for residues more than two turns of helix away from the I13C mutation site are not affected by the sequence change (Figure S7). Even the chemical shift perturbations for residues directly contacting the mutation site ( $i \pm 1, i - 3, i + 4$ ) are

modest in size, indicating that the mutation leaves the helical structure intact. When tagging with the paramagnetic Tm lanthanide ion, strong paramagnetic relaxation broadening of protons within ca.  $15 \text{ \AA}$  of the lanthanide ion makes them unobservable in TROSY-HSQC spectra. However, residues starting from R24 and beyond are sufficiently intense that  $^1\text{H}$ - $^{15}\text{N}$  RDCs can be reliably measured (Figure 8A).



**Figure 8.** Analysis of  $^1\text{D}_{\text{NH}}$  RDCs, collected at 900 MHz for I13C MT, aligned by the DOTA-M8-Tm tag,<sup>26</sup> attached to Cys-13. (A) RDCs collected at 20 (blue) and 35 °C (red). Amides of residues Q3–R23 could not be observed due to paramagnetic broadening. The superimposed sinusoids correspond to the dipolar wave pattern<sup>42</sup> expected for an idealized  $\alpha$ -helix, multiplied by a decaying exponential function with a decay constant of 49.75 (blue) or 36.63 (red) residues. (B) Local alignment strength ( $D_a$ ) as a function of residue number,  $N$ . Alignment tensor values, extrapolated to Cys-13, are listed in Table S8. Exponential fits to the decreasing  $D_a$  values (solid lines) yield persistence lengths of 149 and 110 residues at 20 and 35 °C, respectively.  $D_a$  values that are not used for the exponential fits are plotted as open circles; fitted exponential curves are  $23.06 \times \exp[-(i - 13)/49.75] \text{ Hz}$  and  $17.54 \times \exp[-(i - 13)/36.63] \text{ Hz}$ , where  $i$  is the residue number, and fitted rmsd errors are 0.8 and 0.9 Hz at 20 and 35 °C, respectively.

For calculating the relation between persistence length and the degree of alignment, as reflected in the magnitude of the local alignment tensor, we follow Spudich and co-workers and consider the SAH domain to be a worm-like chain.<sup>21</sup> In this model, the angle  $\beta_n$  of the helical axis at the position of residue  $n$  randomly deviates from its orientation at  $n = 0$ , and follows the equation:

$$\langle \cos(\beta_n) \rangle = \exp(-n/L_p) \quad (1)$$

where  $L_p$  is the persistence length in units of helical rise per residue ( $\sim 1.5 \text{ \AA}$ ). RDCs between nuclei  $i$  and  $j$  in a rigid molecule are described by<sup>61</sup>

$$D_{ij} = (D_a/r_{ij}^3)[(3 \cos^2 \theta - 1) + \frac{3}{2}R \sin^2 \theta \cos(2\phi)] \quad (2)$$

where  $r_{ij}$  is the distance between  $i$  and  $j$ ,  $D_a$  is the strength of the alignment,  $R$  is its rhombicity, and  $\theta$  and  $\phi$  are the polar angles of the  $i$ - $j$  vector in a frame where the alignment tensor is diagonalized. As shown in the [Supporting Information](#), dynamic disorder such as found in a worm-like chain model causes scaling of the RDCs at position  $n$  with  $\langle P_2(\cos(\beta_n)) \rangle$ , where  $P_2(x)$  is the second-order Legendre polynomial:

$$\langle P_2(\cos(\beta_n)) \rangle = \exp(-3n/L_p) \quad (3)$$

For diffusive motion around a symmetry axis, such as applies to the worm-like chain, only the  $D_a$  term in [eq 2](#) is impacted and scales with  $\langle P_2(\cos(\beta_n)) \rangle$ . The decrease in apparent  $D_{a,n}$  as a function of residue position,  $n$ , in the MT domain ([Figure 8B](#)) is then derived from the paramagnetic  $^1D_{NH}(n)$  of residue  $n$ , by the scale factor needed to match it to the value predicted by an alignment tensor whose  $D_a$  and  $R$  values are obtained from a singular value decomposition (SVD) fit<sup>61</sup> to the paramagnetic  $^1D_{NH}$  RDCs measured for residues R24–E45 of the MT structure. As described above, this MT structure was determined from RDCs measured in Pf1 alignment medium. The result of the SVD fit of the paramagnetic RDCs then represents a first-order approximation for the average paramagnetic alignment of R24–E45, neglecting the effect of differential motional averaging of the RDCs across the segment. The paramagnetic RDCs deviate substantially from those measured in Pf1, as exemplified by the large dipolar wave amplitude ([Figure 8A](#)), but fit fairly well to the structure derived from RDCs measured in Pf1 ([Figure 9A](#)), even while ignoring the increased orientational spread when moving away from the paramagnetic tagging site.

The decrease in alignment strength when moving away from the paramagnetic tagging site corresponds to  $\langle P_2(\cos(\beta_n)) \rangle$ , and a fit of these data to a monoexponential function then yields the persistence length,  $L_p = 150 \pm 7$  residues, or  $224 \pm 10$  Å, at 20 °C, and  $L_p = 110 \pm 6$  residues, or  $165 \pm 9$  Å, at 35 °C ([Figure 8B](#)). When upscaling the experimental paramagnetic RDCs by the inverse of this exponentially decaying function, their SVD fit to the MT structure is remarkably good, all the way to residue R64 ([Figure S8B](#)). When viewing the time averaged structure, R64 therefore may be considered the last residue of the SAH domain, even while measurements of HX rates and  $^{15}\text{N}$  relaxation indicate an increased degree of transient H-bond opening and rapid internal motion when approaching this end of the helix.

## CONCLUDING REMARKS

Prior work has provided compelling evidence for the pivotal role of ionic interactions between the side chains of Glu residues with those of Arg and Lys, which group in patterns of roughly four like-charged residues in SAH domains.<sup>1,62</sup> A very recent NMR study of the SAH domain of murine myosin-VIIA highlights the dynamic character of such ion pair interactions, adding a substantial favorable entropic component to the remarkable stability of SAH domains, and explaining the difficulty in obtaining high-resolution X-ray structural information, in particular for their side chains.<sup>11</sup> Our study extends and complements that work by providing the high-resolution backbone structure of an SAH domain, unconstrained by lattice packing forces or long-range intramolecular

interactions, and globally defined by RDCs. The latter is critical as the absence of long-range NOEs and the strictly local information contained in chemical shifts result in the accumulation of uncertainty regarding the global structure of an SAH domain. Our results differ somewhat from the previous analysis of SAXS data of the myosin-VI MT domain, which suggested the presence of two oblique but distinct kinks in its SAH domain.<sup>22</sup> However, it has also been noted that the presence of such kinks is difficult to distinguish from random deviations in a straight helix by SAXS data, and that SAXS data can only become an unambiguous reporter for the persistence length when the helix length becomes comparable to or larger than  $L_p$ , a situation where dynamic bending significantly impacts the end-to-end distance of such helices.<sup>21</sup> By collecting three different types of backbone RDCs, we were able to unambiguously evaluate the structural features of MT. Remarkably, after correcting the RDCs for the increased amplitude of backbone dynamics by amounts derived from  $^{15}\text{N}$  relaxation, the data fit remarkably well to an idealized, perfectly straight  $\alpha$ -helix. The  $Q$ -factor, which provides a quantitative measure for the quality of the fit, then equals 15.6%, that compares favorably to what is typically seen when fitting RDCs to high-resolution X-ray structures. For example, fits of the same three types of backbone couplings to the 1.8 Å X-ray structure of ubiquitin yielded  $Q \approx 18\%$ , which dropped to 12% when fitting the RDCs to a set of 15 independently determined ubiquitin X-ray structures, each solved at a resolution  $\leq 1.8$  Å.<sup>58</sup> Refinement of the MT domain by using the experimental RDCs, corrected for  $^{15}\text{N}$  dynamics, reduces the cross-validated free  $Q$  factor to 12.9% ([Table S3](#)). This observation confirms that, although deviations from a straight, idealized  $\alpha$ -helix are detectable in the MT SAH domain, the magnitude of these deviations is remarkably small ([Figure 5D](#)).

In our study, we adopted the traditional, sequential mode of studying the structure and dynamics of MT. In principle, carrying out the two types of analyses simultaneously could be advantageous.<sup>57,63–65</sup> However, it then becomes difficult to separate uncertainty in structure from uncertainty in dynamics, which in some cases has led to widely divergent conclusions, even for a simple model system.<sup>45,58</sup> For MT, the exceptional high quality of the RDC fit to the coordinates of an idealized helix validated the accuracy of its local structure, then enabling the study of its motional characteristics in a subsequent step.

As implied by the virtual absence of chemical shift changes between pH 6.3 and 7.8 ([Figure S3C](#)), the structure is essentially independent of pH over this range. Base-catalyzed HX at pH 8.9 is found to scale relative to pH 7.8 approximately linearly with  $[\text{OH}^-]$  ([Table S6](#)), and chemical shifts remain minimally impacted, suggesting that the structure and stability of MT extends well into this basic range. The RDCs measured for the paramagnetically aligned MT domain at pH 6.3 and 7.0 are indistinguishable within experimental uncertainty ([Figure S9](#)), indicating that the persistence length is also not measurably impacted by pH at near physiological values.

Previously, the gradual nature of the melting of an SAH domain upon increasing temperature, as observed by CD spectroscopy, has been interpreted as evidence for non-cooperative unfolding of SAH domains,<sup>9</sup> consistent with the spring-like behavior seen in single-molecule pulling experiments.<sup>6</sup> While the highly uniform temperature dependence of our hydrogen exchange data at first sight seems to be at odds with these prior interpretations, we note that for the SAH domain of myosin-VI, the hydrogen exchange data can only be

interpreted as indicative of local cooperativity: The average sequence composition varies very little along the MT sequence, and the absence of a difference in temperature dependence of the HX-rates of, for example, the N-terminal quarter of the SAH domain relative to the C-terminal quarter does not mean that these well-separated regions unfold cooperatively. However, the virtual absence of local variation in the temperature dependence of the HX rates, even at locations where the ER/K pattern is disrupted by alternate residue types, implies that, locally, the backbone H-bonding is highly cooperative, even more so than is seen in regular non-SAH  $\alpha$ -helices.<sup>66,67</sup> This result appears consistent with a minimum-length requirement of ca. 16 residues for the formation of SAH helices seen in the early synthetic peptide work of Kallenbach and co-workers.<sup>7</sup> In this respect, it is interesting to note that the local parameters for secondary structure, including  $^3J_{\text{HNH}\alpha}$ ,  $\Delta\delta^{13}\text{C}^\alpha$ , and the HX rates, show evidence of decreased helicity starting at residue R58, whereas the RDC-refined structure indicates that, on average, the helix continues for nearly two full turns beyond this point. This result suggests that the absence of helical residues beyond E64 impacts the H-bond stability several turns earlier in the helix.

Use of the paramagnetic tagging to magnetically align one end of the MT domain while observing the decrease in alignment when moving away from the tagging site provides a uniquely sensitive method for measuring the persistence length of the SAH domain in a very benign manner. The magnetic alignment forces, exerted by the paramagnetic tag, are 3 orders of magnitude weaker than  $kT$  and therefore can be considered nonperturbing. The high precision at which the strength of local alignment can be measured from RDCs provides a unique and highly sensitive probe for detection of the local ordering as a function of distance from the tagging site. Our measurements for MT yielded a persistence length,  $L_p = 224 \pm 10 \text{ \AA}$ , that is ca. 50% longer than that reported for the 196-residue Kelch SAH domain.<sup>21</sup> However, considering that this latter domain contains substantial deviations from the canonical ER/K motif, this moderate difference in persistence length is perhaps not surprising.

As noted by Spudich and others,<sup>1,3,15,21,68</sup> SAH domains are ideally suited to function as rods for designing proteins with optimized interdomain spacing. An alternate application could be their use as flexible springs, linking two sites of a molecule or molecular complex and thereby creating a physical force.<sup>69,70</sup> If the two SAH termini were connected by intrinsically unfolded peptide sequences to two sites on a target protein, the entropic component of partially ordering these linkers also would serve as a mechanical force.<sup>71</sup> The size of SAH helices is sufficiently small that adding such elements to a protein or protein complex will not be prohibitive for recording high-quality NMR spectra, thereby enabling the study of mechanical strain induced in a system at atomic resolution by isotropic solution NMR spectroscopy. Such experiments can provide an atomic resolution complement to single molecule pulling experiments and are currently being explored in our laboratory.

## ■ ASSOCIATED CONTENT

### ● Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.9b03116.

Relaxation analysis, structure calculations, and additional data (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*bax@nih.gov

### ORCID

Ad Bax: 0000-0002-9809-5700

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank J. L. Baber, for technical support, John Lloyd and the NIDDK Core Facility as well as Duck-Yeon Lee and the NHLBI Biochemistry Core Facility for mass spectrometry, Dr. Daniel Haussinger for his generous gift of the DOTA-M8-Spy lanthanide tags, and G. Marius Clore, Attila Szabo, Charles Schwieters, and Robert Best for valuable discussions. This work was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases.

## ■ REFERENCES

- (1) Swanson, C. J.; Sivaramakrishnan, S. Harnessing the Unique Structural Properties of Isolated alpha-Helices. *J. Biol. Chem.* **2014**, *289* (37), 25460–25467.
- (2) Knight, P. J.; Thirumurugan, K.; Xu, Y. H.; Wang, F.; Kalverda, A. P.; Stafford, W. F.; Sellers, J. R.; Peckham, M. The predicted coiled-coil domain of myosin 10 forms a novel elongated domain that lengthens the head. *J. Biol. Chem.* **2005**, *280* (41), 34702–34708.
- (3) Peckham, M.; Knight, P. J. When a predicted coiled coil is really a single alpha-helix, in myosins and other proteins. *Soft Matter* **2009**, *5*, 2493–2503.
- (4) Suveges, D.; Gaspari, Z.; Toth, G.; Nyitray, L. Charged single alpha-helix: A versatile protein structural motif. *Proteins: Struct., Funct., Genet.* **2009**, *74* (4), 905–916.
- (5) Zagrovic, B.; Jayachandran, G.; Millett, I. S.; Doniach, S.; Pande, V. S. How large is an alpha-helix? Studies of the radii of gyration of helical peptides by small-angle X-ray scattering and molecular dynamics. *J. Mol. Biol.* **2005**, *353* (2), 232–241.
- (6) Wolny, M.; Batchelor, M.; Knight, P. J.; Paci, E.; Dougan, L.; Peckham, M. Stable Single alpha-Helices Are Constant Force Springs in Proteins. *J. Biol. Chem.* **2014**, *289* (40), 27825–27835.
- (7) Lyu, P. C. C.; Gans, P. J.; Kallenbach, N. R. Energetic contribution of solvent-exposed ion-pairs to alpha-helix structure. *J. Mol. Biol.* **1992**, *223* (1), 343–350.
- (8) Wang, C. L. A.; Chalovich, J. M.; Graceffa, P.; Lu, R. C.; Mabuchi, K.; Stafford, W. F. A long helix from the central region of smooth-muscle caldesmon. *J. Biol. Chem.* **1991**, *266*, 13958–13963.
- (9) Spink, B. J.; Sivaramakrishnan, S.; Lipfert, J.; Doniach, S.; Spudich, J. A. Long single alpha-helical tail domains bridge the gap between structure and function of myosin VI. *Nat. Struct. Mol. Biol.* **2008**, *15* (6), 591–597.
- (10) Li, J. C.; Chen, Y. Y.; Deng, Y. S.; Unarta, I. C.; Lu, Q.; Huang, X. H.; Zhang, M. J. Ca<sup>2+</sup>-Induced Rigidity Change of the Myosin VIIa IQ Motif-Single alpha Helix Lever Arm Extension. *Structure* **2017**, *25* (4), 579–591.
- (11) Batchelor, M.; Wolny, M.; Baker, E. G.; Paci, E.; Kalverda, A. P.; Peckham, M. Dynamic ion pair behavior stabilizes single alpha helices in proteins. *J. Biol. Chem.* **2019**, *294* (9), 3219–3234.
- (12) Sun, Y. J.; Goldman, Y. E. Lever-Arm Mechanics of Processive Myosins. *Biophys. J.* **2011**, *101* (1), 1–11.
- (13) Ulrich, A. K. C.; Seeger, M.; Schutze, T.; Bartlick, N.; Wahl, M. C. Scaffolding in the Spliceosome via Single alpha Helices. *Structure* **2016**, *24* (11), 1972–1983.

- (14) Baboolal, T. G.; Sakamoto, T.; Forgacs, E.; White, H. D.; Jackson, S. M.; Takagi, Y.; Farrow, R. E.; Molloy, J. E.; Knight, P. J.; Sellers, J. R.; Peckham, M. The SAH domain extends the functional length of the myosin lever. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (52), 22193–22198.
- (15) Sivaramakrishnan, S.; Spudich, J. A. Systematic control of protein interaction using a modular ER/K alpha-helix linker. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (51), 20467–20472.
- (16) Ritt, M.; Guan, J. L.; Sivaramakrishnan, S. Visualizing and Manipulating Focal Adhesion Kinase Regulation in Live Cells. *J. Biol. Chem.* **2013**, *288* (13), 8875–8886.
- (17) Swanson, C. J.; Ritt, M.; Wang, W.; Lang, M. J.; Narayan, A.; Tesmer, J. J.; Westfall, M.; Sivaramakrishnan, S. Conserved Modular Domains Team up to Latch-open Active Protein Kinase C alpha. *J. Biol. Chem.* **2014**, *289* (25), 17812–17829.
- (18) Zhou, X. X.; Chung, H. K.; Lam, A. J.; Lin, M. Z. Optical Control of Protein Activity by Fluorescent Protein Domains. *Science* **2012**, *338* (6108), 810–814.
- (19) Kretsinger, R. H.; Rudnick, S. E.; Weissman, L. J. Crystal structure of calmodulin. *J. Inorg. Biochem.* **1986**, *28* (2–3), 289–302.
- (20) Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A. Backbone dynamics of calmodulin studied by <sup>15</sup>N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry* **1992**, *31* (23), 5269–5278.
- (21) Sivaramakrishnan, S.; Sung, J.; Ali, M.; Doniach, S.; Flyvbjerg, H.; Spudich, J. A. Combining Single-Molecule Optical Trapping and Small-Angle X-Ray Scattering Measurements to Compute the Persistence Length of a Protein ER/K alpha-Helix. *Biophys. J.* **2009**, *97* (11), 2993–2999.
- (22) Sivaramakrishnan, S.; Spink, B. J.; Sim, A. Y. L.; Doniach, S.; Spudich, J. A. Dynamic charge interactions create surprising rigidity in the ER/K alpha-helical protein motif. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (36), 13356–13361.
- (23) 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* (23), 12366–12371.
- (24) 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* (4), 844–848.
- (25) Englander, S. W.; Downer, N. W.; Teitelbaum, H. Hydrogen-exchange. *Annu. Rev. Biochem.* **1972**, *41*, 903–924.
- (26) 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* (41), 14761–14767.
- (27) 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* (18), 6841–6846.
- (28) Bohm, G.; Muhr, R.; Jaenicke, R. Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng., Des. Sel.* **1992**, *5* (3), 191–195.
- (29) Mobli, M.; Hoch, J. C. Nonuniform sampling and non-Fourier signal processing methods in multidimensional NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *83*, 21–41.
- (30) Ying, J.; Delaglio, F.; Torchia, D. A.; Bax, A. Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. *J. Biomol. NMR* **2017**, *68*, 101–118.
- (31) 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* (3), 277–293.
- (32) Lakomek, N. A.; Ying, J. F.; Bax, A. Measurement of 15N relaxation rates in perdeuterated proteins by TROSY-based methods. *J. Biomol. NMR* **2012**, *53*, 209–221.
- (33) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G.; Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice*; Academic Press: San Diego, CA, 1996; p 587.
- (34) Mandel, A. M.; Akke, M.; Palmer, A. G. Backbone Dynamics of Escherichia coli Ribonuclease HI: correlations with Structure and Function in an Active enzyme. *J. Mol. Biol.* **1995**, *246*, 144–163.
- (35) Roche, J.; Ying, J. F.; Shen, Y.; Torchia, D. A.; Bax, A. ARTSY-J: Convenient and precise measurement of (3)J(HNH alpha) couplings in medium-size proteins from TROSY-HSQC spectra. *J. Magn. Reson.* **2016**, *268*, 73–81.
- (36) Fitzkee, N. C.; Torchia, D. A.; Bax, A. Measuring rapid hydrogen exchange in the homodimeric 36 kDa HIV-1 integrase catalytic core domain. *Protein Sci.* **2011**, *20* (3), 500–512.
- (37) Hansen, M. R.; Mueller, L.; Pardi, A. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.* **1998**, *5* (12), 1065–1074.
- (38) Fitzkee, N. C.; Bax, A. Facile measurement of H-1-N-15 residual dipolar couplings in larger perdeuterated proteins. *J. Biomol. NMR* **2010**, *48* (2), 65–70.
- (39) Wang, Y. X.; Marquardt, J. L.; Wingfield, P.; Stahl, S. J.; Lee-Huang, S.; Torchia, D.; Bax, A. Simultaneous measurement of H-1-N-15, H-1-C-13", and N-15-C-13" dipolar couplings in a perdeuterated 30 kDa protein dissolved in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* **1998**, *120* (29), 7385–7386.
- (40) Rock, R. S.; Ramamurthy, B.; Dunn, A. R.; Beccafico, S.; Rami, B. R.; Morris, C.; Spink, B. J.; Franzini-Armstrong, C.; Spudich, J. A.; Sweeney, H. L. A flexible domain is essential for the large step size and processivity of myosin VI. *Mol. Cell* **2005**, *17* (4), 603–609.
- (41) Okten, Z.; Churchman, L. S.; Rock, R. S.; Spudich, J. A. Myosin VI walks hand-over-hand along actin. *Nat. Struct. Mol. Biol.* **2004**, *11* (9), 884–887.
- (42) Mesleh, M. F.; Opella, S. J. Dipolar Waves as NMR maps of helices in proteins. *J. Magn. Reson.* **2003**, *163* (2), 288–299.
- (43) Meiler, J.; Prompers, J. J.; Peti, W.; Griesinger, C.; Bruschweiler, R. Model-free approach to the dynamic interpretation of residual dipolar couplings in globular proteins. *J. Am. Chem. Soc.* **2001**, *123* (25), 6098–6107.
- (44) 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.
- (45) 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* (5882), 1471–1475.
- (46) Lipari, G.; Szabo, A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **1982**, *104*, 4546–4559.
- (47) Grishaev, A.; Bax, A. An empirical backbone-backbone hydrogen-bonding potential in proteins and its applications to NMR structure refinement and validation. *J. Am. Chem. Soc.* **2004**, *126* (23), 7281–7292.
- (48) Schwieters, C. D.; Kuszewski, J. J.; Clore, G. M. Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48* (1), 47–62.
- (49) 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* (39), 9008–9012.
- (50) Schurr, J. M.; Babcock, H. P.; Fujimoto, B. S. A Test of the Model-Free Formulas - Effects of Anisotropic Rotational Diffusion and Dimerization. *J. Magn. Reson., Ser. B* **1994**, *105* (3), 211–224.
- (51) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. Rotational diffusion anisotropy of human ubiquitin from N-15 NMR relaxation. *J. Am. Chem. Soc.* **1995**, *117* (50), 12562–12566.
- (52) Halle, B.; Davidovic, M. Biomolecular hydration: From water dynamics to hydrodynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (21), 12135–12140.

- (53) Garcia de la Torre, P.; Bloomfield, V. A. Hydrodynamic properties of complex, rigid biological macromolecules. Theory and applications. *Q. Rev. Biophys.* **1981**, *14*, 81–139.
- (54) Fushman, D.; Ghose, R.; Cowburn, D. 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.* **2000**, *122* (43), 10640–10649.
- (55) Tirado, M. M.; Garcíadelatorre, J. Rotational-Dynamics of Rigid, Symmetric Top Macromolecules - Application to Circular-Cylinders. *J. Chem. Phys.* **1980**, *73* (4), 1986–1993.
- (56) Vogeli, B. The nuclear Overhauser effect from a quantitative perspective. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *78*, 1–46.
- (57) Vogeli, B.; Kazemi, S.; Guntert, P.; Riek, R. Spatial elucidation of motion in proteins by ensemble-based structure calculation using exact NOEs. *Nat. Struct. Mol. Biol.* **2012**, *19* (10), 1053–1058.
- (58) Maltsev, A. S.; Grishaev, A.; Roche, J.; Zasloff, M.; Bax, A. Improved Cross Validation of a Static Ubiquitin Structure Derived from High Precision Residual Dipolar Couplings Measured in a Drug-Based Liquid Crystalline Phase. *J. Am. Chem. Soc.* **2014**, *136* (10), 3752–3755.
- (59) Bai, Y.; Milne, J. S.; Englander, S. W. Primary structure effects on peptide group exchange. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 75–86.
- (60) Bai, Y. W.; Englander, S. W. Future directions in folding: The multi-state nature of protein structure. *Proteins: Struct., Funct., Genet.* **1996**, *24* (2), 145–151.
- (61) Prestegard, J. H.; Bougault, C. M.; Kishore, A. I. Residual dipolar couplings in structure determination of biomolecules. *Chem. Rev.* **2004**, *104* (8), 3519–3540.
- (62) Wolny, M.; Batchelor, M.; Bartlett, G. J.; Baker, E. G.; Kurzawa, M.; Knight, P. J.; Dougan, L.; Woolfson, D. N.; Paci, E.; Peckham, M. Characterization of long and stable de novo single alpha-helix domains provides novel insight into their stability. *Sci. Rep.* **2017**, *7*, 1 DOI: [10.1038/srep44341](https://doi.org/10.1038/srep44341).
- (63) Lindorff-Larsen, K.; Best, R. B.; DePristo, M. A.; Dobson, C. M.; Vendruscolo, M. Simultaneous determination of protein structure and dynamics. *Nature* **2005**, *433* (7022), 128–132.
- (64) Clore, G. M.; Schwieters, C. D. Concordance of residual dipolar couplings, backbone order parameters and crystallographic B-factors for a small alpha/beta protein: A unified picture of high probability, fast atomic motions in proteins. *J. Mol. Biol.* **2006**, *355* (5), 879–886.
- (65) Bouvignies, G.; Markwick, P. R. L.; Blackledge, M. Simultaneous definition of high resolution protein structure and backbone conformational dynamics using NMR residual dipolar couplings. *ChemPhysChem* **2007**, *8* (13), 1901–1909.
- (66) Wieczorek, R.; Dannenberg, J. J. H-bonding cooperativity and energetics of alpha-helix formation of five 17-amino acid peptides. *J. Am. Chem. Soc.* **2003**, *125* (27), 8124–8129.
- (67) Li, J.; Wang, Y.; Chen, J.; Liu, Z.; Bax, A.; Yao, L. Observation of alpha-Helical Hydrogen-Bond Cooperativity in an Intact Protein. *J. Am. Chem. Soc.* **2016**, *138* (6), 1824–1827.
- (68) Arai, R.; Ueda, H.; Kitayama, A.; Kamiya, N.; Nagamune, T. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng., Des. Sel.* **2001**, *14* (8), 529–532.
- (69) Zocchi, G. Controlling proteins through molecular springs. *Annu. Rev. Biophys.* **2009**, *38*, 75–88.
- (70) Choi, B.; Zocchi, G.; Wu, Y.; Chan, S.; Perry, L. J. Allosteric control through mechanical tension. *Phys. Rev. Lett.* **2005**, *95* (7), 1 DOI: [10.1103/PhysRevLett.95.078102](https://doi.org/10.1103/PhysRevLett.95.078102).
- (71) Keul, N. D.; Oruganty, K.; Bergman, E. T. S.; Beattie, N. R.; McDonald, W. E.; Kadirvelraj, R.; Gross, M. L.; Phillips, R. S.; Harvey, S. C.; Wood, Z. A. The entropic force generated by intrinsically disordered segments tunes protein function. *Nature* **2018**, *563* (7732), 584–588.
- (72) Kay, L. E.; Keifer, P.; Saarinen, T. Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J. Am. Chem. Soc.* **1992**, *114*, 10663–10665.
- (73) Kjaergaard, M.; Poulsen, F. M. Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *J. Biomol. NMR* **2011**, *50* (2), 157–165.
- (74) Maltsev, A. S.; Ying, J. F.; Bax, A. Deuterium isotope shifts for backbone <sup>1</sup>H, <sup>15</sup>N and <sup>13</sup>C nuclei in intrinsically disordered protein alpha-synuclein. *J. Biomol. NMR* **2012**, *54*, 181–191.
- (75) Ottiger, M.; Bax, A. Determination of relative N-H-N N-C", C-alpha-C", and C(alpha)-H-alpha effective bond lengths in a protein by NMR in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* **1998**, *120* (47), 12334–12341.