NMR Determines Transient Structure and Dynamics in the Disordered C-Terminal Domain of WASp Interacting Protein

Noam Y. Haba,[†△] Renana Gross,[†△] Jiri Novacek,[§] Hadassa Shaked,[†] Lukas Zidek,[§] Mira Barda-Saad,[‡] and Jordan H. Chill[†*]

[†]Department of Chemistry and [‡]The Mina and Everard Goodman Faculty of Life Sciences, Bar Ilan University, Ramat Gan, Israel; and [§]Faculty of Science, NCBR and CEITEC, Masaryk University, Brno, Czech Republic

ABSTRACT WASp-interacting protein (WIP) is a 503-residue proline-rich polypeptide expressed in human T cells. The WIP C-terminal domain binds to Wiskott-Aldrich syndrome protein (WASp) and regulates its activation and degradation, and the WIP-WASp interaction has been shown to be critical for actin polymerization and implicated in the onset of WAS and X-linked thrombocytopenia. WIP is predicted to be an intrinsically disordered protein, a class of polypeptides that are of great interest because they violate the traditional structure-function paradigm. In this first (to our knowledge) study of WIP in its unbound state, we used NMR to investigate the biophysical behavior of WIP^C, a C-terminal domain fragment of WIP that includes residues 407–503 and contains the WASp-binding site. In light of the poor spectral dispersion exhibited by WIP^C and the high occurrence (25%) of proline residues, we employed 5D-NMR ^{13}C-detected NMR experiments with nonuniform sampling to accomplish full resonance assignment. Secondary chemical-shift analysis, ^{15}N relaxation rates, and protection from solvent exchange all concurred in detecting transient structure located in motifs that span the WASp-binding site. Residues 446–456 exhibited a propensity for helical conformation, and an extended conformation followed by a short, capped helix was observed for residues 468–478. The ^{13}C-detected approach allows chemical-shift assignment in the WIP^C polyproline stretches and thus sheds light on their conformation and dynamics. The effects of temperature on chemical shifts referenced to a denatured sample of the polypeptide demonstrate that heating reduces the structural character of WIP^C. Thus, we conclude that the disordered WIP^C fragment is comprised of regions with latent structure connected by flexible loops, an architecture with implications for binding affinity and function.

INTRODUCTION

Wiskott-Aldrich syndrome protein (WASp) is expressed predominantly in hematopoietic cells and is a key participant in actin polymerization, which accompanies activation of T cells. Its cellular levels are closely controlled by a combination of activation and degradation mechanisms. Mutations in the WASp-encoding gene that disturb this delicate balance have been implicated in Wiskott-Aldrich syndrome (WAS) and X-linked thrombocytopenia (XLT) (1), both of which are hereditary immunodeficiency conditions characterized by impaired cytoskeleton formation and an increased incidence of autoimmune diseases and malignancies (2). WASp-interacting protein (WIP) acts as the cellular chaperone of WASp, conveys WASp to areas of active actin assembly after antigen-receptor and chemokine receptor signaling (3,4), and stabilizes WASp by shielding it from cellular degradation systems (5). The pathological effects of WAS-inducing mutations located in the N-terminal EVH1 domain of WASp result from a reduced affinity to WIP, suggesting that WIP's chaperone function is essential for WASp stability. Supporting this notion is the fact that in one study (6), WIP-deficient mice had normal levels of WASp messenger RNA (mRNA) levels but suffered from decreased WASP activity, which could be rescued by exogenous addition of WIP. In resting lymphocytes, >95% of WASp is complexed with WIP, which inhibits Cdc42-induced activation of WASp (7).

As a member of the mammalian verprolin family, the 503-residue WIP is comprised of an N-terminal actin-binding WAS-homology (WH2) domain, proline-rich domains (PRDs) that bind profilin and SH3 domains, and a C-terminal segment capable of binding the N-terminal EVH1 domain of WASp (8). Disruption of the latter interface is the cause of WAS. The C-terminal region contains a consensus phosphorylation motif (9) that is recognized by the θ isoform of protein kinase C (PKCθ), and this phosphorylation leads to activation of WASp (7,10). A 35-mer peptide consisting of WIP residues 451–485 fused to a neuronal WASp homolog (N-WASp) was demonstrated to be an N-WASp-binding determinant comprised of three well-conserved epitopes, spanning residues 454–478. The WIP peptide, including a central polyproline motif, was shown to wrap around and bind to an elongated interaction surface on N-WASp that is best described as a concave spool (11,12). To date, the unbound state of WIP has not been structurally investigated.

WIP (particularly its C-terminal domain) is predicted to be an intrinsically disordered protein (IDP). IDPs fail to adopt a well-defined structure under biologically native conditions, and are best described as an ensemble of unfolded

Submitted January 2, 2013, and accepted for publication May 20, 2013.

△Noam Y. Haba and Renana Gross contributed equally to this work.

*Correspondence: jordan.chill@biu.ac.il

Editor: Josh Wand.

© 2013 by the Biophysical Society

0006-3495/13/07/0481/13 $2.00

http://dx.doi.org/10.1016/j.bpj.2013.05.046

and partially folded states that contribute to their overall behavior in solution (13,14). IDPs are involved in protein aggregation, translocation, and degradation, are quite common in eukaryotic systems, and are closely related to human disease (15,16), accounting for the desire to understand their conformation and dynamics. It is now clear that transient structural elements observed in these polypeptides are associated with their biological function, often reflecting a latent structural propensity that may be evoked and stabilized by formation of a complex with a binding partner (17,18). The multiconformational nature of IDPs makes them flexible biological entities that are capable of interacting with—and thus being regulated by—a variety of biological effectors in response to changing cellular conditions. Located at the crossroads of T cell activation and regulated by its interaction with several cellular factors, WIP appears to fit this paradigm well.

Crystallographic studies of IDPs have been frustrated by the lack of permanent structure, since crystals are difficult to obtain, and any information extracted cannot correctly represent the flexible nature of this class of proteins. In contrast, it is precisely this attribute of IDPs that emphasizes the inherent advantages of high-resolution NMR over other structural methods, explaining why NMR has emerged as the preferred approach for investigating nascent structure in IDPs. Chemical shifts measured by NMR are highly sensitive probes of polypeptide secondary structure, representing an average over multiple populations and allowing the detection of minor contributions (19,20). NMR methods are also capable of detecting molecular motions over a wide range of timescales (21,22), as well as identifying long-range intramolecular contacts and intermolecular interaction surfaces that exhibit intermediate and even low affinity (23,24). The backbone of IDPs exhibits considerable motion on the picosecond-to-nanosecond timescale, resulting in favorable nuclear relaxation properties for NMR measurements, and recent methodological advances have alleviated the difficulties associated with the low spectral dispersion characteristic of IDPs (25–29). These two factors enable resonance assignment and measurement of structural parameters for this class of proteins possible.

Despite its central role in activation-induced actin polymerization in T cells, only limited structural information is available for WIP, and its biological activity is poorly understood on the molecular level. In particular, WIP has not been studied in the unbound state, which may be biologically significant once activation has occurred. In this work, we used NMR methods to determine the structural propensity and dynamics of the C-terminal domain of WIP. Secondary chemical-shift analysis, ^{15}N relaxation measurements, solvent exchange of amide protons, and temperature-induced chemical-shift effects all identified within the WIP C-terminal domain short segments with transient structure connected by unstructured loops. The ^{13}C′-based assignment provided new (to our knowledge) insight into the behavior of polyproline segments, which are usually inaccessible to HN-based assignment experiments. The alternating pattern of structured and unstructured regions may give WIP the flexibility to bind multiple cellular factors in its native biological setting.

MATERIALS AND METHODS

Acquisition of NMR data

The construct GSSHHHHHH-WIP(407–503)-LEHHHHHH (WIP^{C}) was expressed in M9 minimal medium containing $^{15}NH_4Cl$ (1 g/L), $^{13}C_6$-D-glucose (2.5 g/L). Detailed expression and purification protocols are described in full in the Supporting Material. All NMR samples (unless stated otherwise) were prepared in 20 mM phosphate buffer, pH 5.0–5.2, 20 mM NaCl, 10 mM βME and 7% $^{2}H_2O$, and placed in a Shigemi tube (Shigemi, Allison, PA). All 2D- and 3D-NMR measurements were conducted on a DRX700 Bruker spectrometer using a cryogenic triple-resonance TCI probe head equipped with z-axis pulsed field gradients. 4D and 5D experiments were performed on a DRX600 Bruker spectrometer using a cryogenic triple-resonance TCI probe head equipped with z-axis pulsed field gradients.

For backbone assignment we utilized a ^{13}C′-detected strategy (CON spectrum as readout) based on the 3D experiments CANCO, CBCACON, CBCANCO, and C-(CC-TOCSY)-CON (25,30), and the 5D experiments CACONCACO and NCOCANCO (27,28). Details of the NMR experiments performed for purposes of resonance assignment appear in the Supporting Material. All chemical shifts were referenced directly to an external 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) reference measured under sample conditions identical to those used for WIP^{C}. Assigned chemical shifts were deposited in the BioMagResBank (BMRB entry 18265).

^{15}N relaxation data were acquired for a 0.9 mM sample of WIP^{C} on the 700 MHz spectrometer and at 283 K. Previously reported heteronuclear single quantum coherence (HSQC)-based experiments were used to acquire R_1, R_2, and ^{15}N-{^{1}H}-nuclear Overhauser effect (NOE) parameters for the backbone ^{15}N nuclei (31). Cross-correlated (CC) relaxation η_{xy} was estimated by acquiring two tr-HSQC spectra in interleaved fashion, differing by a time period of 2Δ during which the $N^{+}H^{\beta}$ and $N^{+}H^{\alpha}$ operators were interchanged, resulting in a intensity ratio of $\exp(-4\Delta\eta_{xy})$ between peaks in the two spectra (32). Details of the experiments conducted to measure WIP^{C} relaxation rates appear in the Supporting Material. The exchange rates of amide protons with solvent at 298 K were estimated using a CLEANEX-PM-based experiment (33) modified to allow for a CON-type readout. Sample pH was raised to 6.0 for purposes of these measurements, with negligible effects on the spectrum, and amide recovery rates after 10 and 20 ms of mixing time were measured. A ^{13}C-detected HNCO experiment served as the reference spectrum.

Analysis of multidimensional spectra and resonance assignment

We processed and visualized 2D- and 3D-NMR spectra using TopSpin2.1 (Bruker, Karlsruhe, Germany). In the 5D-NMR experiments, we processed preliminary 3D (CACO)NCACO and (NCO)CANCO hyperplanes using spectral processing and the analysis system NMRPipe/NMRDraw 3.0 (34), and processed the full 5D experiment using the sparse multidimensional Fourier transform (SMFT) algorithm (35). The procedure for extracting spectral information from these spectra was previously described (28). All of the spectra were analyzed with the use of the NMR assignment and integration software Sparky 3.115 (T.D. Goddard and D.G. Kneller, University of California, San Francisco, CA). Relaxation rates were extracted by least-square fitting of the decay curve using in-house algorithms, and errors in relaxation rates were estimated from the signal/noise ratio in the relevant experiment.

Sedimentation equilibrium experiments

Sedimentation equilibrium (SE) experiments on ^{15}N,^{13}C-WIP^{C} (molecular mass 13,770 g/mol) in 20 mM phosphate buffer (pH 5.2), 20 mM NaCl, 10 mM βME were performed using an XL-I analytical ultracentrifuge equipped with An-60Ti rotor and absorbance optics (Beckman-Coulter, Brea, CA). Data (absorption at 280 nm) were collected at 298 K in double-sector cells of 12 mm thickness. For three different WIP^{C} concentrations (22, 44, and 66 μM), sedimentation curves were recorded after spinning at 30,000 rpm for 10, 12, and 14 h, providing triplicate data with typical standard deviations (SDs) of 0.004 A280 units, or 0.5–2%. We analyzed the average curve using in-house-written scripts based on MATLAB (The MathWorks, Natick, MA) with a nonlinear least-squares approach to extract molecular weight information (36), using a value of 0.68 cm^3/g for protein specific volume (37). Selection among the models tested (monomer, monomer-dimer equilibrium, and monomer with high MW aggregate) was based on an F statistic obtained from comparison of residuals.

Circular dichroism experiments

Circular dichroism (CD) experiments were performed on a 6.7 μM sample of ^{13}C,^{15}N-labeled WIP^{C} in the NMR buffer (vide infra) on a Chirascan polarimeter (Applied Photophysics, Surrey, United Kingdom). Experiments were conducted at 278–308 K at 5 K intervals, with each measurement repeated three times and subtracted from a measurement of an identical buffer sample. Results were analyzed using the CDSSTR module of the DichroWeb platform for the 190–240 nm range (38) (http://dichroweb.cryst.bbk.ac.uk).

RESULTS

The C-terminal domain of WIP is a disordered protein

Heterologous expression of large IDPs is challenging to achieve due to their sensitivity to proteolytic degradation in the host cell. We therefore focused on the WASp-binding C-terminal domain of WIP as our structural target, and explored the possibility of expressing C-terminal fragments at sufficient levels. The best results were obtained for a polypeptide corresponding to residues 407–503 of WIP (WIP^{C}; Fig. 1 *A*), which yielded a sample amenable to NMR study. This segment includes the central verprolin conserved region (VCR; residues 444–478), which is a WASp-binding

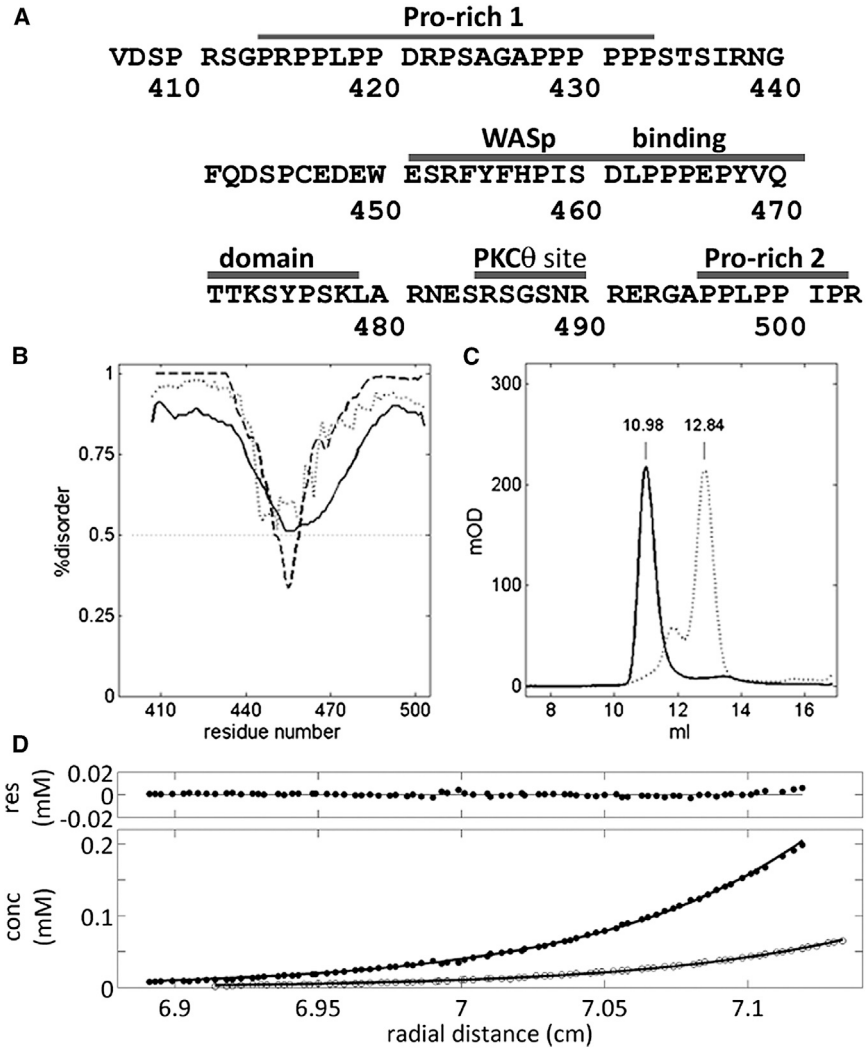

FIGURE 1 WIP^{C}. (*A*) Sequence of WIP^{C} showing the WASp-binding epitopes (as defined in Volkman et al. (11)), polyproline stretches, and PKCθ phosphorylation site. (*B*) Predictions of disorder along WIP^{C} by IUPRED (*solid line*), RONN (*dotted line*), and ISUnstruct2.0 (*dashed line*). (*C*) Size-exclusion chromatograms for WIP^{C} (*solid line*) and ribonuclease A (ribA, *dotted line*), a 13.7 kDa protein. The smaller ribA peak at 11.9 ml behaves as a dimer of ribA (27 kDa) in SDS-PAGE. (*D*) SE curves for uniformly ^{15}N, ^{13}C-labeled WIP^{C} under conditions identical to those used for the NMR samples at 298 K and 30,000 rpm. Results for WIPC concentrations of 22 and 66 μM are shown in open and solid circles, respectively, with fitted curves corresponding to a monomeric species of 15.2 kDa shown as solid lines. Residuals are shown in the upper panel for the 66 μM concentration.

domain that is common to the veprolin family, exhibiting 60–65% homology among WIP, WIP-related protein, and CR16). The VCR is flanked by two PRDs, the N-terminal segment of residues 413–433 (PRD1), and the C-terminal segment of residues 495–503 (PRD2). The PKCθ consensus sequence R^{485}SGSNR^{490} is located immediately before the PRD2 sequence (39).

Sequence analysis of WIP^C using three prediction algorithms predicts this 12.9 kDa fragment (13,770 g/mol in its ^{13}C,^{15}N-labeled form) to be generally disordered, with an indication of a more structured region in the VCR segment (Fig. 1 *B*). This predicted structural propensity was strongest for the aromatic-rich sequence FYFH (residues 454–457), which coincides with one of the conserved WASp-binding epitopes. Indeed, WIP^C migrated on a native polyacrylamide gel as a larger protein than expected (data not shown), and in size-exclusion chromatography (SEC) the protein preceded its expected elution volume on an analytical Superdex 75 column of 0.55 column volumes, appearing instead at 0.45 column volumes, corresponding to a protein of ~45–50 kDa (Fig. 1 *C*). To ensure that these findings were not biased by aggregation of WIP^C, we employed SE methods, which offer a quantitative estimation of protein size. SE curves of ^{13}C,^{15}N-WIP^C suggested a mass of ~15,200 g/mol, close to that expected for a monomer. This deviation from the expected mass may be attributed to the difficulty of accurately determining the specific volume of an IDP, which would cause an overestimation of the molecular mass (37). In any event, the curves were incompatible with the formation of a dimer above the detection limit of 1%, and, notably, the fitted values were independent of sample concentration (Fig. 1 *D*). Together, these findings suggest a larger-than-expected radius of gyration for WIP^C and confirm its disordered character.

Resonance assignment of WIP^C

A comparison of the 2D ^{1}H,^{15}N-HSQC, and 2D-IPAP $^{13}C'$-^{15}N (CON) spectra of WIP^C (Fig. 2) reveals a dual justification for employing $^{13}C'$-detected experiments in the case of WIP^C: 1), they detect the otherwise invisible proline residues, which comprise 26% of the WIP^C sequence; and 2), they display significantly advantageous spectral dispersion. Despite this spread of resonances, a set of 3D-NMR experiments, including CANCO, CBCANCO, and CBCACON (25), could achieve no more than 50% of the assignment. We attributed this to the highly repetitive WIP^C sequence, which contains (in its 97 WIP-derived residues) 26% Pro, 15% Ser, and 17% Arg/Glu (with a highly similar $^{13}C^{\alpha}/^{13}C^{\beta}$ pair of shifts). To overcome this difficulty, we employed a recently proposed method for assigning unstructured, repetitive sequences based on a pair of 5D-NMR experiments, (H)NCOCANCO and (H)CACONCACO, which utilize nonuniform sampling (NUS) and T_1-optimized excitation to correlate five different frequencies within a

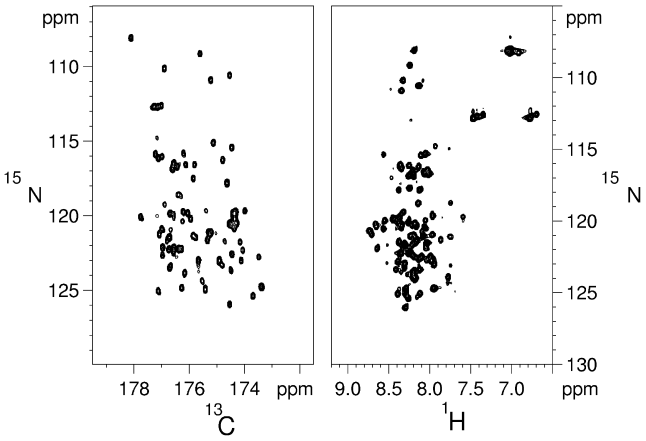

FIGURE 2 Comparison of fingerprint spectra of WIP^C. ^{13}C-^{15}N CON-IPAP (*left*) and ^{1}H-^{15}N HSQC (*right*) spectra of WIP^C, both acquired at 16.4 T and 298 K. The spectral width shown in the F_2 dimension was chosen so that typical crosspeak line widths in this dimension (6 Hz and 23 Hz for the CON and HSQC, respectively) appear equal, demonstrating the advantageous spectral dispersion of the CON spectrum. Peaks emanating from proline residues (in the 134–141 ^{15}N ppm range) are not shown, but are clearly another advantageous feature of the CON experiment.

reasonable experimental time (27,28). This approach is demonstrated in Fig. 3 *A*. All WIP^C residues, with the exception of the 450–456 segment, were successfully assigned using this approach. The absence of this missing region is attributable to its more structured nature (vide infra), with faster relaxation rates limiting the obtainable signal/noise ratio in the relaxation-sensitive 5D-NMR experiments. Supporting this explanation is the fact that these residues afforded significantly weaker peaks in the CON experiment, and signal intensity showed a gradual decrease for residues flanking this segment, pointing to a systematic structural effect rather than an effect of random spectral overlap. By the process of elimination, reconsideration of the above 3D spectra and proton-detected HNCACB spectrum identified a set of weaker peaks corresponding to the missing residues. Using this combined assignment strategy, we were able to fully assign the 97 WIP^C segment and flanking residues (Fig. 3 *B*). In addition, H^{α} and $^{13}C^{\beta}$ resonances were obtained from a $^{1}J(NC^{\alpha})$-selective HCBCANCO experiment (40). The complete set of backbone and $^{13}C^{\beta}$ frequencies lays the necessary groundwork for further characterization of the structure and dynamics of WIP^C.

Structural features of WIP^C

Backbone assignment immediately provides a view of secondary structure, even if it is transient, because chemical shifts are very sensitive to dihedral angles along the polypeptide backbone. Secondary chemical shifts, or the change in resonance frequency compared with those measured for a random coil conformation, have traditionally been relied upon as an indicator of secondary structure (41–44). This general relationship was recently refined for unstructured

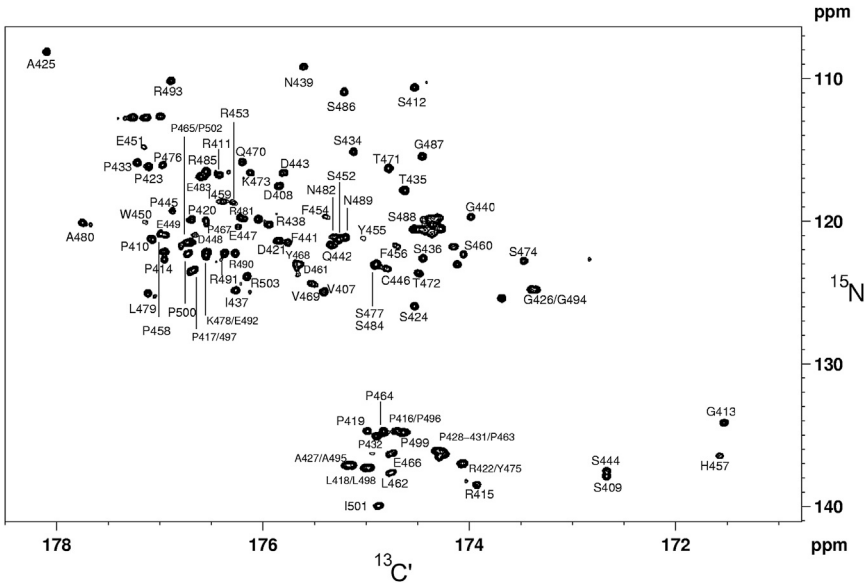

FIGURE 3 5D-NMR-based assignment of WIP^{C} resonance frequencies. Full assignment of WIP^{C}. Peaks are identified by their ^{13}C′ residue. Unmarked peaks arise from side-chain, His$_6$-tag, or linker-residue ^{13}C′-^{15}N pairs.

proteins based on a representative database of 14 IDPs (neighbor-corrected IDP (ncIDP) (45)) and an ensemble approach based on chemical shifts in folded proteins (δ2D (46)). Fig. 4 shows the results of these analyses for WIP^{C}; secondary chemical shifts for WIP^{C} ^{13}C′, ^{13}C$^{\alpha}$, and ^{13}C$^{\beta}$ nuclei ($\Delta\delta$ values) obtained from the ncIDP prediction; and secondary structure predictions from the δ2D database, revealing several intriguing features.

Most striking is the propensity for α-helical structure in the central core of WIP^{C} at residues C^{446}EDEWESRFYF^{456}, for which all chemical-shift parameters are in agreement. The magnitude of the shifts observed is consistent with a 15–20% population of such a helical conformation, consistent with the value obtained in the secondary structure prediction. Supporting this conclusion is the fact that CON peaks emanating from these residues were relatively weak, typically exhibiting a 10-fold reduction in the signal/noise ratio. This behavior cannot be attributed to a dimerization of WIP^{C} via the side chains of C^{446} residues, since spectra were recorded in the presence of 10 mM βME and were unaffected by addition of 5 mM dithiothreitol, and the C^{446} ^{13}C$^{\beta}$ chemical shift (28.2 ppm) is typical for a reduced thiol group. Another region with structural tendency is apparent at residues K^{473}SYPS^{477}, which exhibits the characteristics of a short N-capped helix, with a typical negative-positive motif for $\Delta\delta$ of ^{13}C′ and ^{13}C$^{\alpha}$ and an opposite effect for ^{13}C$^{\beta}$. Less pronounced is a marginal propensity for helical conformation in the S^{484}RSGSNRRE^{492} segment.

Secondary shifts and structural propensity in polyproline segments

Of particular interest is the residual structure assumed by polyproline stretches. Proline-rich WIP^{C} has a six-residue polyproline motif (P^{428}PPPPP^{433}) and three broken stretches (the N-terminal P^{414}RPPLPP^{420}, the central (WASp-binding) P^{463}PPEP^{467}, and the C-terminal P^{496}PLPPIP^{502}). The chemical shifts of these sequences are inaccessible to assignment approaches using amide ^{1}H-^{15}N moieties as the spectroscopic link between two adjacent amino acids. Consequently, these chemical shifts are poorly represented in resonance frequency databases, rendering secondary structure prediction for these regions only marginally accurate. In contrast, the ^{13}C-detected approach used in this study is well suited for measuring chemical shifts in such regions. Internal proline residues within the hexa-proline segment exhibited shifts of 174.3, 61.2, and 30.5 ppm for the ^{13}C′, ^{13}C$^{\alpha}$, and ^{13}C$^{\beta}$ nuclei, respectively. These values represent $\Delta\delta$(^{13}C′,^{13}Cα,^{13}Cβ) secondary shifts (−0.5, +0.5, −0.8) when using the ncIDP database, and (−1.2, +0.2, −0.4) when compared with neighbor-corrected random-coil values calculated by SPARTA+ (47). The unique imino acid proline is known to exert significant effects on polypeptide backbone conformation (48–50). A closer look at ncIDP-derived secondary chemical shifts revealed that this characteristic down-up-down secondary shift motif is observed in Pro(i)-Pro($i+1$) dyads throughout the WIP^{C} sequence for the ^{13}Cα/^{13}Cβ nuclei in residue Pro(i) and the ^{13}C′ nucleus of the preceding ($i-1$) residue.

In view of the shortage of polyproline segments with well-documented chemical shifts, the structural interpretation of these results is not straightforward. Of the 15 BioMagResBank entries containing polyproline segments, only seven provided chemical-shift data for the polyproline region, and only two provided ^{13}Cα and ^{13}Cβ chemical shifts; ^{13}C′ data were unavailable. The pentaproline segment of the SPCp41 protein (BMRB #15013; PDB accession code 2JMC) (51) and the PPLPP motif of a tyrosine kinase-interacting peptide (BMRB #6456; PDB accession code 1WA7) (52) exhibit shifts that are highly similar to those obtained

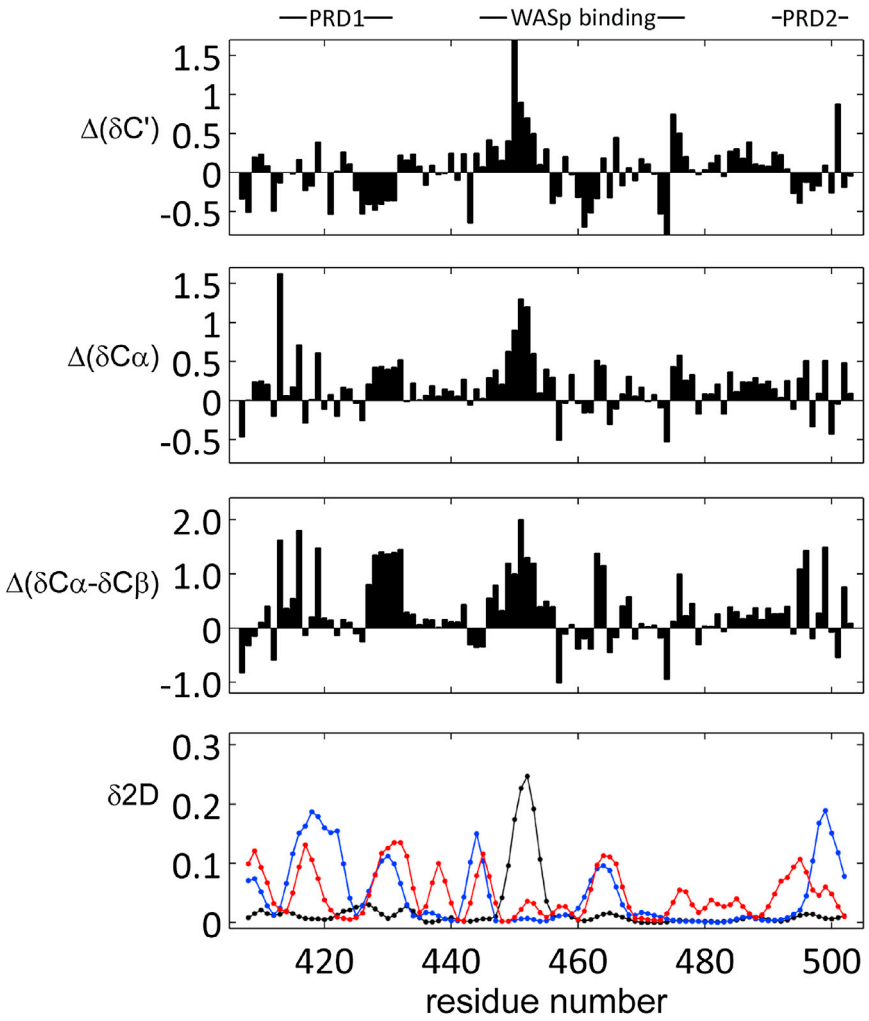

FIGURE 4 Secondary chemical-shift analysis for WIP^C. Chemical-shift analysis using the ncIDP and δ2D databases, comparing WIP^C chemical shifts with reference values in IDPs corrected for neighbor effects. Data are shown for $^{13}C'$ (first panel), $^{13}C\alpha$ (second panel), and ($^{13}C\alpha - {}^{13}C\beta$) (third panel). Fourth panel: δ2D analysis of WIP^C chemical shifts. Shown are the probabilities for α-helix (*black*), β-sheet/extended (*dark grey, blue* online), and polyproline II helix (*light grey, red* online). The remaining probabilities (not plotted) are for coil structure.

for residues 428–433 and 416–420/496–500, respectively, of WIP^C. They also both adopt helical polyproline-II (P_{II}) structures, characterized by a *trans* configuration of the peptide bonds and a left-handed helical arrangement, creating dihedral angles of $(\phi, \psi) = (-75, 140)$ and a near alignment of each third residue in an axial view of the helix. In light of this similarity, we conclude that polyproline stretches in WIP^C do contain a significant measure of P_{II} structure.

Backbone dynamics in WIP^C

We investigated the dynamics of WIP^C by measuring ^{15}N relaxation rates, which are well-established reporters of backbone motions on the picosecond-to-nanosecond timescale and a useful probe of transient structure in IDPs. A fully unstructured polypeptide conforming to the segmental motion model (53) should exhibit a flat, bell-shaped profile with a central plateau; however, this is not the case for relaxation rates in WIP^C measured at 283 K (Fig. 5, *top four panels*). Relaxation rates are relatively uniform throughout WIP^C, with R_1 values of 1.5–1.6 s^{-1}, R_2 values of 6–8 s^{-1}, and ^1H-^{15}N NOE (hetNOE) values of 0.15–0.30. Notably, nonproline residues within the PRDs and the short helical motif at residues 473–477 exhibit such typical relaxation rates as well. The exception is the VCR region, and particularly residues 445–460, where relaxation rates change to R_1 values of 1.3–1.4 s^{-1}, R_2 values of 15–20 s^{-1}, and hetNOE values of 0.40–0.60. Such changes are consistent with an increase in structural character in this region. Generally, in the WIP^C VCR domain, R_1 and R_2 values are anticorrelated, a behavior that is typical of polypeptides in the slow-tumbling regime and is notably absent in the less-structured terminal regions. Cross-correlated relaxation rates, which are insensitive to exchange processes, exhibit a similar pattern, with peak values of $\eta_{xy} =$ 9–14 s^{-1} in the core region, as opposed to $\eta_{xy} =$ 2–5 s^{-1} in the unstructured regions. This clearly indicates that it is the more-structural nature of the VCR segment that causes elevated R_2 values, rather than exchange on the microsecond-to-millisecond timescale. Although IDPs are incompatible with the standard Lipari-Szabo model-free approach (54), the observed patterns are consistent with the structural

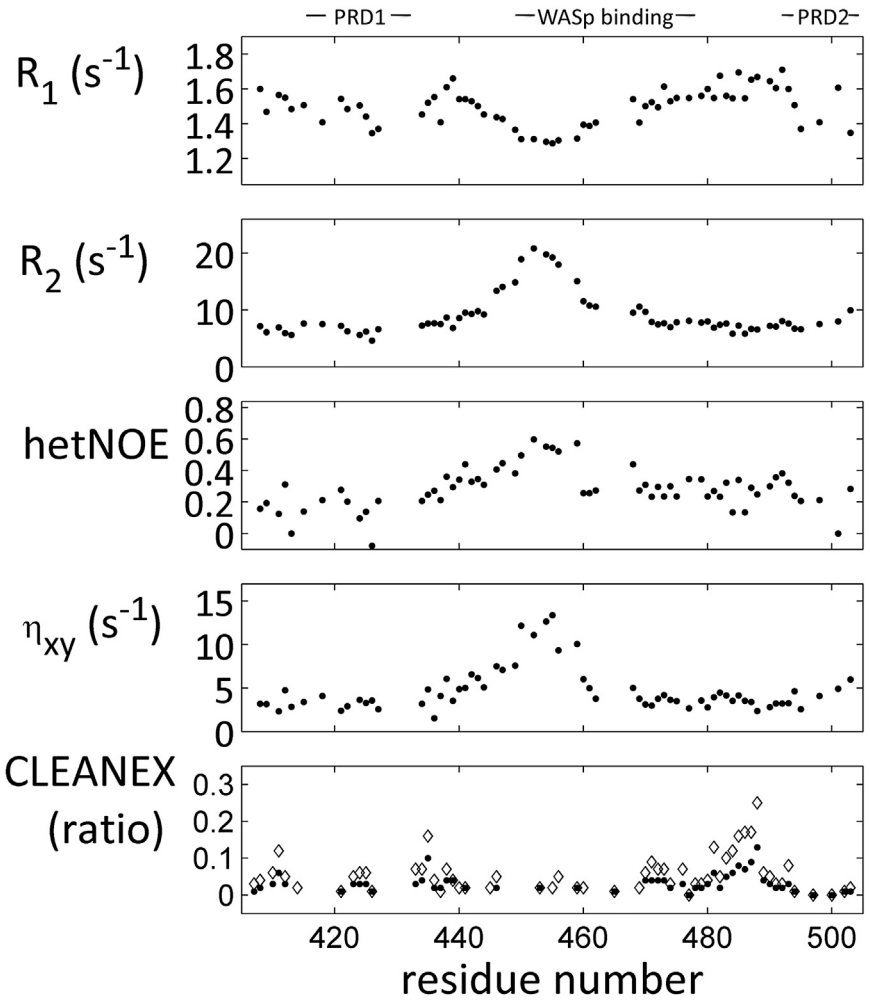

FIGURE 5 Backbone dynamics along the WIP^{C} polypeptide. Relaxation rates were measured for WIP^{C} backbone ^{15}N nuclei at 16.4 T and 283 K. Top to bottom: ^{15}N R_1, ^{15}N R_2, ^{1}H-^{15}N NOE, ^{15}N cross-correlated relaxation (η_{xy}), and rate of amide proton exchange with bulk solvent as measured in a CLEANEX-PM experiment with 10 (*solid circles*) and 20 (*open diamonds*) ms of exchange time. Residue numbers refer to the ^{13}C′ nucleus of the residue preceding the amide proton for which the rate was measured.

view offered by secondary chemical shifts, which show the central region containing the WASp binding epitopes to be more structured than the generally flexible terminal segments of WIP^{C}.

Solvent accessibility along the WIP^{C} backbone

To further characterize the degree of structure along the WIP^{C}, we estimated the exchange rates of amide protons with the solvent using a CON-based CLEANEX-PM experiment (33) in which exchange-mediated recovery of magnetization was measured for WIP^{C} at 298 K and pH 6 after 10 and 20 ms. Reduced WIP^{C} solubility at pH values in the 6.5–7.5 range required these CLEANEX experiments to be conducted under conditions in which solvent exchange is inherently lower. Nevertheless, data acquired after $t = 20$ ms, which still met the necessary $t \ll 1/R_1$ condition, allowed us to discriminate among different protection factors along the WIP^{C} backbone (Fig. 5, *bottom panel*). Regions that were most protected from solvent exchange, exhibiting negligible recovery after 20 ms, included residues 440–468, spanning the rigid VCR domain, and residues 475–479, both consistent with structural elements delineated by chemical-shift data. Notably, nonproline residues nested in proline-rich segments, such as R^{415}, L^{418}, E^{466}, L^{498}, and I^{501}, exhibited low recovery rates as well. This may be due to the relatively hydrophobic character of the polyproline backbone environment. Low protection factors identified increased flexibility in two regions of WIP^{C}, residues 470–474 and 481–490, which exhibited recovery rates of 5–10% and 10–25%, respectively. The former is the segment that links two structured regions, the WASp-binding poly-Pro sequence and the N-capped short helix, and the latter corresponds to the PKCθ phosphorylation sequence. Both relaxation and solvent-exchange results concur with the secondary structure tendencies identified by the chemical-shift data, together demarcating regions of relative rigidity, reflecting a structural propensity, and more flexible regions that are fully unfolded.

Effects of temperature on WIP^{C} structure

As a disordered protein, WIP^{C} exists in rapidly interchanging multiple conformations separated by low-activation

energy barriers, and the effect of temperature on this conformational distribution is a reliable indicator of their relative stability. Increasing temperature caused a reduction and bathochromic shift of the polyproline-II helix (P_{II}) signal and an apparent increase in the α-helical signal in the far-UV CD curve of WIP^C, creating a characteristic isodichroic intersection of CD curves at 209 nm (Fig. 6 *A*). However, it was previously noted that a shift of residues from the P_{II} to the extended region of the Ramachandran plot, rather than an increase in helical content, may account for this change (55). Whereas CD provides a macroscopic view of protein structure, chemical shifts offer more quantitative information with per-residue resolution, motivating us to use them to investigate temperature-dependent structure changes.

Reference statistical coil chemical shifts are also affected by temperature and must be correctly estimated for accurate measurement of temperature-dependent secondary chemical shifts, particularly since IDPs with transient structure will typically exhibit only minor temperature effects. We accomplished this using intrinsic random coil referencing (56), comparing $^{13}C'$ and $^{13}C^{\alpha}$ resonances under normal and denaturing conditions at various temperatures. We acquired the fingerprint CON spectrum of WIP^C at temperatures ranging from 278 to 308 K at 10 K intervals, and the 3D CACON spectrum, providing $\{^{13}C^{\alpha}(i),^{13}C'(i),^{15}N(i+1)\}$ correlations, at 288 and 308 K, repeating all measurements for a identical sample denatured by addition of 7.5 M urea. All spectra could be assigned by comparison with the data obtained for WIP^C under normal conditions at 298 K; to facilitate this process, spectra were recorded for WIP^C in 2.5 and 5 M urea as well. Intrinsically referenced chemical shifts (IRCSs) at each temperature were calculated as the

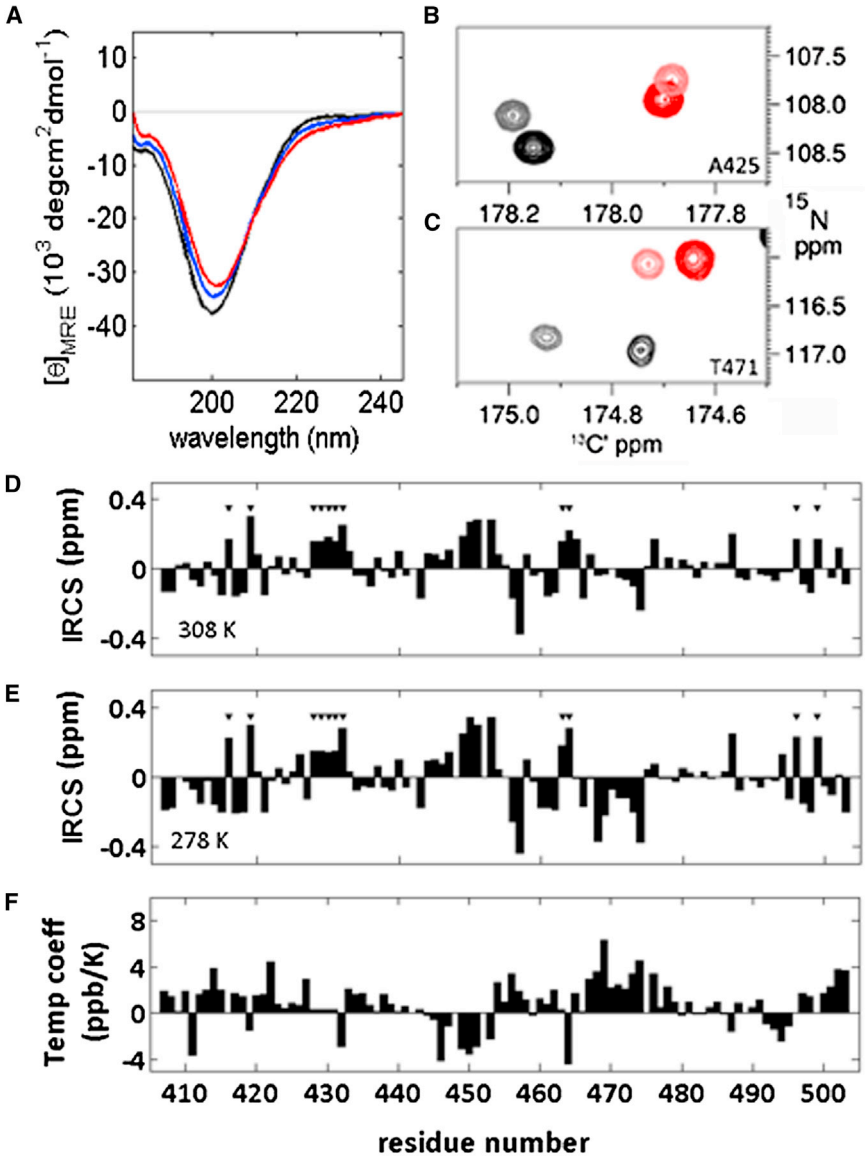

FIGURE 6 Effects of temperature on structural motifs in WIP^C. (*A*) CD curves for a 6.7 μM WIP^C sample at 278 (*black*), 293 (*blue*), and 308 (*red*) K. An increase (decrease) in signal is observed for the 215–230 (195–205) range. (*B* and *C*) Temperature effects on the intrinsic $^{13}C'$ shifts of WIP^C as observed in CON-IPAP spectra. Black (*red*) peaks were obtained at 278 (308) K; dark (light) peaks were obtained for samples under normal (denaturing) conditions. Plots focus on residues A425 (*B*) and T471 (*C*). (*D–F*) $^{13}C'$ IRCSs at 308 K (*D*) and 278 K (*E*), and temperature factors (*F*) plotted against the WIP^C sequence. The intrinsic temperature coefficient was calculated from the slope of IRCS versus temperature in the 278–308 range. The first residues of Pro-Pro dyads are designated by black triangles. Typical errors for temperature factors were 0.2 and 0.5 ppb/K for the unstructured and structured regions of WIP^C, respectively.

difference between shifts in samples under normal and denaturing conditions. Generally, $^{13}C^{\alpha}$ IRCS values were less informative than their $^{13}C'$ counterparts, particularly since several shifts in more structured regions were missing due to increased relaxation losses from the low-temperature measurement. Therefore, the following analysis focuses on temperature effects on $^{13}C'$ shifts.

The derivation of $^{13}C'$ IRCS values at 308 K and 278 K for residues in random coil and transiently structured domains (A425 and T471, respectively) is shown in Fig. 6 *B*, and the data for the entire WIP^C domain are presented in Fig. 6 *C*. IRCS values are residue specific and clearly distinguish between regions with undetectable transient structure, such as residues 422–427, 434–444, and 479–494, with values close to zero, and more structured segments as determined by previous experiments. Throughout the sequence and at all temperatures, Pro(i)-Pro($i + 1$) dyads exhibit a pattern of negative $^{13}C'(i-1)$ and positive $^{13}C^{\alpha}(i-1)/^{13}C'(i)$ shifts. This observation highlights the effect of the constrained Xxx-Pro peptide bond on chemical shifts. The more-structured region of WIP^C corresponding to the WASp-binding site (residues 446–478) exhibits the most significant IRCS values. Positive values for residues 446–453 and negative values for residues 460–475 (with the exception of the poly-Pro segment) are indicative of helical and extended conformations, respectively, in these regions.

Intrinsic temperature coefficients, defined as a slope of IRCS versus temperature, were calculated to assess the effects of temperature on WIP^C transient structure. Cooling WIP^C to 278 K caused an enhancement of the structural pattern, as shown by an analysis of $^{13}C'$ intrinsic temperature coefficients (Fig. 6 C, *bottom panel*). Negative values for the helical region (residues 446–453) and positive values in the extended region (residues 468–475) are consistent with loss of structure at higher temperatures. Similarly, positive values in the PRDs indicate that a decrease in temperature stabilized the proline-imposed rigidity in these regions. The small (<1.5 ppb/K) positive baseline in unstructured regions is consistent with the previously reported behavior of $^{13}C'$ IRCS values (55). We conclude that an increase in temperature induces an increase in flexibility and a loss of structural content, and that the CD results must be interpreted as a reduction in helical and P_{II} conformations on the background of a redistribution of the statistical random coil ensemble.

DISCUSSION

The flexibility and multiconformational nature of IDPs are inherent contributors to their functionality in biological systems. The ability of NMR to study structure and motions on a per-residue basis along the polypeptide backbone of IDPs is unique and accounts for its increasing utility in investigations of this class of proteins (20). The introduction of two recent breakthrough methodologies, ^{13}C-detected backbone assignment experiments (25) and higher-dimensionality NMR experiments (57–59), supported by nonuniform data sampling techniques (28,35,60,61), has enhanced this capability by effectively addressing the difficulties presented by the low-spectral-dispersion characteristic of IDPs. In this study, these technologies were instrumental in completing backbone assignment of WIP^C, allowing for structural interpretation of other experiments. The assignment process utilized a combination of two $^{13}C'$-detected approaches: 3D-NMR experiments correlating the backbone $^{13}C'$-^{15}N group with the four adjacent $^{13}C\alpha/^{13}C\beta$ nuclei, and 5D-NMR with nonuniform sampling correlating among five backbone nuclei spanning two to three residues. As expected, the 5D experiments that required more magnetization-transferred steps were most challenged by the more structured regions of WIP^C, where relaxation losses are most pronounced. Still, whereas the 3D-NMR experiments were unable to provide >50% of the assignment, >90% could be assigned by 5D-NMR, with only the most structured region defying assignment. Once the majority of residues were assigned, the more sensitive 3D-NMR experiments, both proton- and $^{13}C'$-detected, were successful in identifying the missing frequencies. Thus, in tandem the two methods performed in complementary fashion. In another methodological study (40), it was established that two $^{13}C'$-detected 4D experiments, HabCabCON and intra-HabCabNCO, could also provide sequential connectivities and allow assignment, although this approach was considerably more tedious than the 5D approach described above.

In this work we utilized chemical-shift information, ^{15}N relaxation rates, and exposure to solvent exchange to characterize the unbound state of a WIP^C fragment. The results of these measurements are consistent in identifying regions in which a higher contribution of structured conformations is observed. Structure is most tangible in the helical WASp-binding core domain including residues 446–456, which represents the main structural feature of WIP^C. This conclusion is supported by secondary chemical shifts as well as relaxation rates in this region. It is unlikely that the helical conformation is affected or induced by intermolecular interactions, since 1), SE results did not reflect WIP oligomerization; and 2), chemical shifts exhibited only negligible concentration dependence. In addition, the similarity of η_{xy}/R_2 ratios in this region to those predicted by theory rules out significant exchange contributions to relaxation. Additional regions with an appreciable population of structured conformation are the polyproline stretches, residues 414–423, 428–433, 463–467 (in the WASp-binding domain) and 496–502, and the 473–478 segment, which adopts a short N-capped helical conformation. A comparison of WIP^C with other recently studied IDPs demonstrates a similarity between levels of transient structure observed for these proteins. The secondary shifts measured for WIP^C were comparable to those observed for securin (62), tau protein (63), and others (64,65). The transiently

structured regions are connected by unstructured linker segments, as could be deduced from minimal secondary structure, low transverse relaxation and hetNOE rates, minimal IRCS values, and exposure to solvent exchange. One such region of interest is the phosphorylation site of residues 484–492, whose flexibility is consistent with its need to be sufficiently exposed to allow access to the phosphorylating enzyme PKCθ. These findings are in agreement with the paradigm that IDPs, rather than being fully disordered, are actually comprised of flexible segments connecting regions of transient secondary structure. The latter sample a more restricted conformational space and may represent nuclei of emerging structure that can be evoked by interaction with binding partners.

A unique aspect of WIP^{C} is the prevalence of proline residues. Proline-rich segments have proven roles in mediating protein-protein interactions (66–68), suggesting a biological role for such segments along WIP^{C}. The use of $^{13}C'$-detected experiments is an attractive alternative to other approaches (69,70) that render such segments NMR visible, here allowing the first (to our knowledge) assignment of extensive poly-proline segments in the context of a larger protein, whereas previously this could be achieved only for small peptides. Chemical shifts observed for the $P^{428}PPPPP^{433}$ and $P^{416}PLPP^{420}$/ $P^{496}PLPP^{500}$ segments indicated that they adopted P_{II} structures, which appeared to be relatively rigid as suggested by the fact that these segments were protected from solvent exchange. Pro-Pro dyads induced characteristic deviations from random coil chemical shifts, particularly the down-up-down motif for chemical shifts $\{^{13}C'(i-1),^{13}C^{\alpha}(i),^{13}C^{\beta}(i)\}$ and the down-up-up motif for IRCS values of $\{^{13}C'(i-1),^{13}C^{\alpha}(i-1),^{13}C'(i)\}$. Regarding the former, however, we note that the ncIDP database (which lacks poly-Pro segments) does not consider two-residue neighbor effects, which are nonnegligible for proline residues (50). Adjusting for this additional effect would mitigate the observed chemical-shift deviations.

Characteristic temperature-dependent changes have been observed in several IDPs (55,71,72). Here, temperature-dependent chemical shifts, compared between normal and denaturing conditions to obtain the intrinsic structure-induced chemical shift, were acquired to follow the effects of temperature on the transient structure in WIP^{C}. Generally, IRCS values in more structured regions were low compared with those reported in a previous study (55), indicating lower levels of transient structure, and temperature effects were commensurately reduced. However, a pattern of increasing structural content at lower temperatures was evident, in agreement with a previous study showing that IDPs did not undergo cold denaturation (73). Analysis of CD curves (in the 278–308 K range) and δ2D analysis of chemical shifts (in the 278–298 K range; data not shown) was consistent with these findings, with both failing to identify temperature-induced structural changes in excess of 5–10%. Nevertheless, the correlation between significant IRCS values and regions with transient structure as determined by other NMR parameters suggests that intrinsic shifts are useful for identifying structural traces in IDPs. However, such an analysis assumes that all structural elements are fully denatured in 7.5 M urea, which may be an overgeneralization. Thus, positive temperature coefficients for the PRDs (residues 407–423, 496–502) may result from a stabilization of the P_{II} conformation (characterized by negative $^{13}C'$ shifts) in 7.5 M urea, as suggested in previous studies (74,75). Residues 428–431 from the hexa-proline segment exhibited negligible temperature effects, reflecting the fact that poly-Pro structures are stabilized by steric clashes rather than enthalpically favored interactions. These observations underline the fact that it is extremely difficult to establish a truly random peptide chain conformation. Whereas in folded proteins effects of this magnitude are usually negligible on the backdrop of secondary shifts of 2–3 ppm, they must be considered in the context of IDPs, where secondary chemical shifts are far less pronounced.

A schematic picture of structural propensity along the WIP^{C} backbone summarizing our findings is shown in Fig. 7. It is instructive to compare the structural tendencies of the core region of WIP^{C} in its unbound state with the conformation of a WIP-derived peptide (residues 451–485) when bound to a neuronal homolog of WASp, N-WASp (12). With the WIP peptide fused before the N-terminus of N-WASp, the three salient structural features observed were 1), a *cis*-peptide H^{457}-P^{458} bond connecting the aromatic stretch $F^{454}YF^{456}$ to a tight turn at residues $I^{459}SD^{461}$; 2), a P_{II} conformation for the polyproline segment of residues $L^{462}PPP^{465}$; and 3), a helical turn at residues $S^{474}YPSK^{478}$. The last two elements were connected by a flexible linker in which T^{471} was the least structured residue (12). Structural motifs 2 and 3 are paralleled by the findings

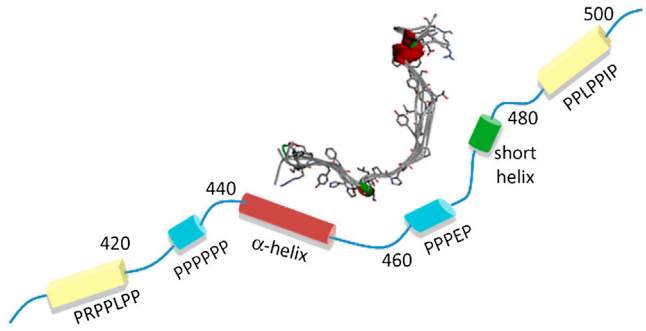

FIGURE 7 Comparison of WIP^{C} in the WASp-bound and unbound states. In ribbons/stick style, the conformation of a WIP peptide (residues 451–485) bound to neuro-WASp (12) is shown as a superposition of five low-energy structures. For clarity, side chains are shown for one conformer only. Below is a schematic representation of unbound WIP^{C} obtained in this study, showing unstructured linkers (*thin lines*) and regions of structural tendency (*cylinders*).

presented here, and the 468–473 linker segment exhibits high-temperature coefficients although it lacks structure, indicating it may adopt a specific biologically relevant conformation in the bound form. In contrast, free WIP^{C} does not foreshadow the *cis* H^{457}-P^{458} bond or the turn at residues 459–461. Thus, contributing to the formation of the WIP/WASp complex are preformed motifs at residues 462–465 and 474–478, and a third motif (residues 457–461) whose structure is formed only in the presence of WASp. It is possible that the flexible segment connecting motifs 2 and 3 makes different interaction modes possible, as would be required in various cellular situations.

It is tempting to speculate about the contribution of helical WIP^{C} residues 446–452 to the affinity of the WIP/WASp complex, since sequences of previously determined WIP/WASp structures (11,12) failed to include this structural element. Binding to WASp can be expected to stabilize this helical conformation because the structural propensities of long IDP segments (six to seven residues) are good predictors of the bound conformation (76). Using the known position of the shorter WIP^{C} fragment in relation to N-WASp, it is possible to suggest potential intermolecular interactions contributed by the predicted helix. Specifically, N-WASp presents a positively charged surface consisting of residues K56, K57, K81, and K84, which could form contacts with the aforementioned putative helix, including four negatively charged residues. Notably, only residues K56 and K84 are conserved (replaced by Arg residues) in the homologous T-cell WASp, allowing the ubiquitous WIP to interact differently with its binding partners in different biological environments. Further studies will be necessary to determine the role of the helical segment at residues 446–452 in the context of WIP/WASp complex formation.

CONCLUSIONS

In this work, we used advanced NMR methods to fully assign the backbone resonances of the 97-residue intrinsically disordered fragment of WIP, which allowed us to study its structure and dynamics in its unbound state. The emerging picture is one of a pattern of alternating partially structured and fully unstructured regions, with the established WASp-binding epitope at the core of the more-structured domain. The WIP-WASp binding interface is a combination of preformed and WASp-induced structural motifs, and it is likely that this pattern is closely related to its function. This study also lays the foundation for investigating the interaction between the extended WIP^{C} fragment and WASp, as well as other cellular binding partners, which has important biological and pharmaceutical implications.

SUPPORTING MATERIAL

Supporting analysis, one figure, and references (77,78) are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)00633-4.

We thank Drs. Hugo Gottlieb and Keren Keinan-Adamsky for spectrometer assistance, Prof. Isabella Felli (CERM, University of Florence) for the CON-based CLEANEX-PM experiment, members of the Barda-Saad laboratory (Bar Ilan University) for assistance in cloning of WIP^{C}, and Dvir Doron (Bar Ilan University) for support in modeling the bound WIP conformation.

This work was supported by the Access to Research Infrastructures activity in the 7th Framework Programme of the European Commission (contract 228461, EAST-NMR), the Czech Science Foundation (grant P206/11/0758), and the Heritage Legacy fund (award 491/10). The 700 MHz spectrometer was purchased with the assistance of a Converging Technologies award and a generous donation by Fundacion Adar.

REFERENCES

1. Derry, J. M., J. A. Kerns, …, U. Francke. 1995. WASP gene mutations in Wiskott-Aldrich syndrome and X-linked thrombocytopenia. *Hum. Mol. Genet.* 4:1127–1135.
2. Thrasher, A. J., and S. O. Burns. 2010. WASP: a key immunological multitasker. *Nat. Rev. Immunol.* 10:182–192.
3. Antón, I. M., and G. E. Jones. 2006. WIP: a multifunctional protein involved in actin cytoskeleton regulation. *Eur. J. Cell Biol.* 85:295–304.
4. Martinez-Quiles, N., R. Rohatgi, …, N. Ramesh. 2001. WIP regulates N-WASP-mediated actin polymerization and filopodium formation. *Nat. Cell Biol.* 3:484–491.
5. de la Fuente, M. A., Y. Sasahara, …, N. Ramesh. 2007. WIP is a chaperone for Wiskott-Aldrich syndrome protein (WASP). *Proc. Natl. Acad. Sci. USA.* 104:926–931.
6. Konno, A., M. Kirby, …, F. Candotti. 2007. The expression of Wiskott-Aldrich syndrome protein (WASP) is dependent on WASP-interacting protein (WIP). *Int. Immunol.* 19:185–192.
7. Sasahara, Y., R. Rachid, …, R. S. Geha. 2002. Mechanism of recruitment of WASP to the immunological synapse and of its activation following TCR ligation. *Mol. Cell.* 10:1269–1281.
8. Ramesh, N., I. M. Antón, …, R. S. Geha. 1997. WIP, a protein associated with wiskott-aldrich syndrome protein, induces actin polymerization and redistribution in lymphoid cells. *Proc. Natl. Acad. Sci. USA.* 94:14671–14676.
9. Bi, K., and A. Altman. 2001. Membrane lipid microdomains and the role of PKCθ in T cell activation. *Semin. Immunol.* 13:139–146.
10. Dong, X., G. Patino-Lopez, …, S. Shaw. 2007. Structure-function analysis of the WIP role in T cell receptor-stimulated NFAT activation: evidence that WIP-WASP dissociation is not required and that the WIP NH2 terminus is inhibitory. *J. Biol. Chem.* 282:30303–30310.
11. Volkman, B. F., K. E. Prehoda, …, W. A. Lim. 2002. Structure of the N-WASP EVH1 domain-WIP complex: insight into the molecular basis of Wiskott-Aldrich syndrome. *Cell.* 111:565–576.
12. Peterson, F. C., Q. Deng, …, B. F. Volkman. 2007. Multiple WASP-interacting protein recognition motifs are required for a functional interaction with N-WASP. *J. Biol. Chem.* 282:8446–8453.
13. Dunker, A. K., I. Silman, …, J. L. Sussman. 2008. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* 18:756–764.
14. Rezaei-Ghaleh, N., M. Blackledge, and M. Zweckstetter. 2012. Intrinsically disordered proteins: from sequence and conformational properties toward drug discovery. *ChemBioChem.* 13:930–950.
15. Uversky, V. N. 2009. Intrinsic disorder in proteins associated with neurodegenerative diseases. *Front. Biosci.* 14:5188–5238.
16. Babu, M. M., R. van der Lee, …, J. Gsponer. 2011. Intrinsically disordered proteins: regulation and disease. *Curr. Opin. Struct. Biol.* 21:432–440.
17. Ward, J. J., J. S. Sodhi, …, D. T. Jones. 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* 337:635–645.

18. Dyson, H. J., and P. E. Wright. 2005. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6:197–208.
19. Mittag, T., and J. D. Forman-Kay. 2007. Atomic-level characterization of disordered protein ensembles. *Curr. Opin. Struct. Biol.* 17:3–14.
20. Eliezer, D. 2009. Biophysical characterization of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* 19:23–30.
21. Torchia, D. A. 2011. Dynamics of biomolecules from picoseconds to seconds at atomic resolution. *J. Magn. Reson.* 212:1–10.
22. Mittermaier, A. K., and L. E. Kay. 2009. Observing biological dynamics at atomic resolution using NMR. *Trends Biochem. Sci.* 34:601–611.
23. Zuiderweg, E. R. P. 2002. Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry.* 41:1–7.
24. Clore, G. M., C. Tang, and J. Iwahara. 2007. Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Curr. Opin. Struct. Biol.* 17:603–616.
25. Bermel, W., I. Bertini, …, R. Pierattelli. 2006. Protonless NMR experiments for sequence-specific assignment of backbone nuclei in unfolded proteins. *J. Am. Chem. Soc.* 128:3918–3919.
26. Mäntylahti, S., M. Hellman, and P. Permi. 2011. Extension of the HA-detection based approach: (HCA)CON(CA)H and (HCA) NCO(CA)H experiments for the main-chain assignment of intrinsically disordered proteins. *J. Biomol. NMR.* 49:99–109.
27. Motáčková, V., J. Nováček, …, V. Sklenář. 2010. Strategy for complete NMR assignment of disordered proteins with highly repetitive sequences based on resolution-enhanced 5D experiments. *J. Biomol. NMR.* 48:169–177.
28. Nováček, J., A. Zawadzka-Kazimierczuk, …, V. Sklenář. 2011. 5D 13C-detected experiments for backbone assignment of unstructured proteins with a very low signal dispersion. *J. Biomol. NMR.* 50:1–11.
29. Bermel, W., I. Bertini, …, J. Stanek. 2012. Speeding up sequence specific assignment of IDPs. *J. Biomol. NMR.* 53:293–301.
30. Bermel, W., I. Bertini, …, R. Pierattelli. 2010. Exclusively heteronuclear NMR experiments to obtain structural and dynamic information on proteins. *ChemPhysChem.* 11:689–695.
31. Farrow, N. A., R. Muhandiram, …, L. E. Kay. 1994. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry.* 33:5984–6003.
32. Chill, J. H., J. M. Louis, …, A. Bax. 2006. Measurement of ^{15}N relaxation in the detergent-solubilized tetrameric KcsA potassium channel. *J. Biomol. NMR.* 36:123–136.
33. Mori, S., J. M. Berg, and P. C. van Zijl. 1996. Separation of intramolecular NOE and exchange peaks in water exchange spectroscopy using spin-echo filters. *J. Biomol. NMR.* 7:77–82.
34. Delaglio, F., S. Grzesiek, …, A. Bax. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* 6:277–293.
35. Kazimierczuk, K., A. Zawadzka-Kazimierczuk, and W. Koźmiński. 2010. Non-uniform frequency domain for optimal exploitation of non-uniform sampling. *J. Magn. Reson.* 205:286–292.
36. Cole, J. L., and J. C. Hansen. 1999. Analytical ultracentrifugation as a contemporary biomolecular research tool. *J. Biomol. Tech.* 10:163–176.
37. Fischer, H., I. Polikarpov, and A. F. Craievich. 2004. Average protein density is a molecular-weight-dependent function. *Protein Sci.* 13:2825–2828.
38. Whitmore, L., and B. A. Wallace. 2008. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers.* 89:392–400.
39. Ramesh, N., and R. S. Geha. 2009. Recent advances in the biology of WASP and WIP. *Immunol. Res.* 44:99–111.
40. Novacek, J., N. Y. Haba, …, L. Zidek. 2012. 4D non-uniformly sampled HABCABCON and ^{1}J(NCα)-selective HCBCANCO experiments for the sequential assignment and chemical shift analysis of intrinsically disordered proteins. *J. Biomol. NMR.* 53:139–148.
41. Wishart, D. S., and D. A. Case. 2001. Use of chemical shifts in macromolecular structure determination. *Methods Enzymol.* 338:3–34.
42. Cavalli, A., X. Salvatella, …, M. Vendruscolo. 2007. Protein structure determination from NMR chemical shifts. *Proc. Natl. Acad. Sci. USA.* 104:9615–9620.
43. Shen, Y., O. Lange, …, A. Bax. 2008. Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. USA.* 105:4685–4690.
44. Raman, S., O. F. Lange, …, D. Baker. 2010. NMR structure determination for larger proteins using backbone-only data. *Science.* 327:1014–1018.
45. Tamiola, K., B. Acar, and F. A. A. Mulder. 2010. Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J. Am. Chem. Soc.* 132:18000–18003.
46. Camilloni, C., A. De Simone, …, M. Vendruscolo. 2012. Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry.* 51:2224–2231.
47. Shen, Y., and A. Bax. 2010. SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *J. Biomol. NMR.* 48:13–22.
48. Ho, B. K., and R. Brasseur. 2005. The Ramachandran plots of glycine and pre-proline. *BMC Struct. Biol.* 5:14.
49. MacArthur, M. W., and J. M. Thornton. 1991. Influence of proline residues on protein conformation. *J. Mol. Biol.* 218:397–412.
50. Schwarzinger, S., G. J. A. Kroon, …, H. J. Dyson. 2001. Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc.* 123:2970–2978.
51. Candel, A. M., F. Conejero-Lara, …, M. Bruix. 2007. The high-resolution NMR structure of a single-chain chimeric protein mimicking a SH3-peptide complex. *FEBS Lett.* 581:687–692.
52. Bauer, F., K. Schweimer, …, H. Sticht. 2005. Structural characterization of Lyn-SH3 domain in complex with a herpesviral protein reveals an extended recognition motif that enhances binding affinity. *Protein Sci.* 14:2487–2498.
53. Schwalbe, H., K. M. Fiebig, …, C. M. Dobson. 1997. Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry.* 36:8977–8991.
54. Lipari, G., and A. Szabo. 1982. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. 2. Analysis of experimental results. *J. Am. Chem. Soc.* 104:4546–4570.
55. Kjaergaard, M., A.-B. Nørholm, …, B. B. Kragelund. 2010. Temperature-dependent structural changes in intrinsically disordered proteins: formation of α-helices or loss of polyproline II? *Protein Sci.* 19:1555–1564.
56. Modig, K., V. W. Jürgensen, …, F. M. Poulsen. 2007. Detection of initiation sites in protein folding of the four helix bundle ACBP by chemical shift analysis. *FEBS Lett.* 581:4965–4971.
57. Atreya, H. S., A. Eletsky, and T. Szyperski. 2005. Resonance assignment of proteins with high shift degeneracy based on 5D spectral information encoded in G2FT NMR experiments. *J. Am. Chem. Soc.* 127:4554–4555.
58. Hiller, S., C. Wasmer, …, K. Wüthrich. 2007. Sequence-specific resonance assignment of soluble nonglobular proteins by 7D APSY-NMR spectroscopy. *J. Am. Chem. Soc.* 129:10823–10828.
59. Wen, J., J. Wu, and P. Zhou. 2011. Sparsely sampled high-resolution 4-D experiments for efficient backbone resonance assignment of disordered proteins. *J. Magn. Reson.* 209:94–100.
60. Kupce, E., and R. Freeman. 2003. Projection-reconstruction of three-dimensional NMR spectra. *J. Am. Chem. Soc.* 125:13958–13959.
61. Coggins, B. E., and P. Zhou. 2008. High resolution 4-D spectroscopy with sparse concentric shell sampling and FFT-CLEAN. *J. Biomol. NMR.* 42:225–239.

62. Csizmok, V., I. C. Felli, …, I. Bertini. 2008. Structural and dynamic characterization of intrinsically disordered human securin by NMR spectroscopy. *J. Am. Chem. Soc.* 130:16873–16879.

63. Mukrasch, M. D., S. Bibow, …, M. Zweckstetter. 2009. Structural polymorphism of 441-residue tau at single residue resolution. *PLoS Biol.* 7:e34.

64. Rao, J. N., Y. E. Kim, …, T. S. Ulmer. 2009. Effect of pseudorepeat rearrangement on α-synuclein misfolding, vesicle binding, and micelle binding. *J. Mol. Biol.* 390:516–529.

65. Marsh, J. A., V. K. Singh, …, J. D. Forman-Kay. 2006. Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: implications for fibrillation. *Protein Sci.* 15:2795–2804.

66. Williamson, M. P. 1994. The structure and function of proline-rich regions in proteins. *Biochem. J.* 297:249–260.

67. Ball, L. J., R. Kühne, …, H. Oschkinat. 2005. Recognition of proline-rich motifs by protein-protein-interaction domains. *Angew. Chem. Int. Ed. Engl.* 44:2852–2869.

68. Ingham, R. J., K. Colwill, …, T. Pawson. 2005. WW domains provide a platform for the assembly of multiprotein networks. *Mol. Cell. Biol.* 25:7092–7106.

69. Bottomley, M. J., M. J. Macias, …, M. Sattler. 1999. A novel NMR experiment for the sequential assignment of proline residues and proline stretches in $^{13}C/^{15}N$-labeled proteins. *J. Biomol. NMR.* 13:381–385.

70. Kanelis, V., L. Donaldson, …, L. E. Kay. 2000. Sequential assignment of proline-rich regions in proteins: application to modular binding domain complexes. *J. Biomol. NMR.* 16:253–259.

71. Uversky, V. N., J. Li, and A. L. Fink. 2001. Evidence for a partially folded intermediate in α-synuclein fibril formation. *J. Biol. Chem.* 276:10737–10744.

72. Sánchez-Puig, N., D. B. Veprintsev, and A. R. Fersht. 2005. Human full-length Securin is a natively unfolded protein. *Protein Sci.* 14:1410–1418.

73. Tantos, A., P. Friedrich, and P. Tompa. 2009. Cold stability of intrinsically disordered proteins. *FEBS Lett.* 583:465–469.

74. Meier, S., M. Strohmeier, …, S. Grzesiek. 2007. Direct observation of dipolar couplings and hydrogen bonds across a β-hairpin in 8 M urea. *J. Am. Chem. Soc.* 129:754–755.

75. Whittington, S. J., B. W. Chellgren, …, T. P. Creamer. 2005. Urea promotes polyproline II helix formation: implications for protein denatured states. *Biochemistry.* 44:6269–6275.

76. Eliezer, D. 2007. Characterizing residual structure in disordered protein states using nuclear magnetic resonance. *Methods Mol. Biol.* 350:49–67.

77. Cai, M., Y. Huang, …, R. Craigie. 1998. An efficient and cost-effective isotope labeling protocol for proteins expressed in Escherichia coli. *J. Biomol. NMR.* 11:97–102.

78. Kazimierczuk, K., A. Zawadzka, and W. Koźmiński. 2008. Optimization of random time domain sampling in multidimensional NMR. *J. Magn. Reson.* 192:123–130.

Supplementary Material for MS 2013BIOPHYSJ302227:

"NMR determines transient structure and dynamics in the disordered C-terminal domain of WASp interacting protein", Haba, Gross et al.

Expression and purification of WIP^C

A coding region for residues 407-503 of WIP was cloned into a pET28 plasmid (Novagen) between its XhoI and BamHI restriction sites. The resulting plasmid was used to transform BL21 competent cells and to overexpress the protein GSSHHHHHH-WIP(407-503)–LEHHHHHH (WIP^C) in M9 minimal medium containing $^{15}NH_4Cl$ (1 g/L), $^{13}C_6$-D-glucose (2.5 g/L) (1). Cultures were grown to OD_{280} ~ 0.6-0.8 and induced with 1 mM isopropyl-thio-D-galactose (IPTG), followed by expression at 27 °C for 14-18 h, yielding uniformly ^{15}N,^{13}C-labeled WIP^C. Cells were lysed and the supernatant loaded on a Ni^{++}-affinity column. WIP^C eluted at 300 mM imidazole at sufficient purity for purposes of NMR data acquisition. For analytical purposes, size exclusion columns were typically run in 20 mM phosphate buffer (pH 7), 150 mM NaCl and 10 mM βME on a Superdex 75 column. The buffer was exchanged and the sample concentrated to 1.0-1.5 mM in a Vivaspin centrifugation tube with a molecular cutoff of 10 kDa. β-mercaptoethanol (βME, 10 mM) or dithiothreitol (DTT, 1-5 mM) were maintained in the sample to ensure residue C^{446} remains reduced throughout the experiments. The protein was assayed on SDS-PAGE prior to NMR measurements, exhibiting a single (>95%) band at 12-13 kDa. Typical yields were 4-6 mg of purified WIP^C per liter of M9 culture.

Acquisition of NMR data

Backbone assignment. Backbone assignment utilized a $^{13}C'$-detected strategy (CON spectrum as readout), based on the 3D-experiments CANCO, CBCACON, CBCANCO, and C-(CC-TOCSY)-CON (2,3), and the 5D-experimnents CACONCACO and NCOCANCO (4,5). 5D-NMR experiments used non-uniform sampling and T_1-relaxation optimized excitation to acquire the high-dimensionality experiment in reasonable time. Experiments for purposes of resonance assignment were performed at 298 K. 2D-CON-based experiments were typically acquired in interleaved in-phase-anti-phase (IPAP) manner with 200-256 complex points, an acquisition time of 35-45 ms in the ^{15}N dimension, and 1024 complex points and an acquisition time of 145 ms in the observed ^{13}C dimension. 3D-experiments were acquired with 32–40 complex points and 7–7.5 ms acquisition time for Cα-based experiments, and 80-120 complex points and 4-6 ms acquisition time for $^{13}C^{\alpha/\beta}$ evolution. ^{15}N evolution was achieved during a 33 ms constant-time period concomitantly with refocusing of the N-C' coupling, and 1024 complex points and an acquisition time of 145 ms were maintained for the observed ^{13}C dimension.

The (H)CACONCACO experiment was measured with the spectral widths set to 6010 (aq) × 4000 ($^{13}C^{\alpha}$) × 2000 (^{15}N) × 1600 ($^{13}C'$) × 4000 ($^{13}C^{\alpha}$) Hz. The (H)NCOCANCO experiment was measured with the spectral widths set to 6010 (aq) × 2000 (^{15}N) × 4000 ($^{13}C^{\alpha}$) × 1600 ($^{13}C'$) × 2000 (^{15}N) Hz. For both experiments, 1024 complex points were collected in the acquisition dimension and 1800 hypercomplex points were randomly distributed using the Poisson disc algorithm (6) over the four indirectly detected dimensions with the maximum evolution periods adjusted to 27, 30 and 50 ms for the $^{13}C\alpha$, $^{13}C'$, and ^{15}N dimensions, respectively. The interscan delay was set to 0.2 s. Figure S1 details the methodology by which the 5D-NMR data were used to perform backbone assignment.

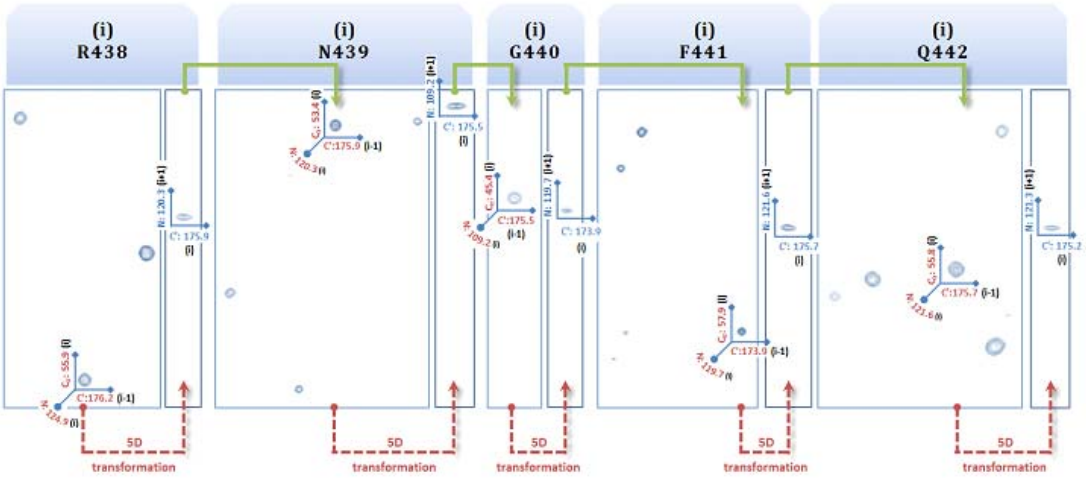

Figure S1. 5D-NMR-based assignment of WIP^{C} resonance frequencies. Assignment of the sequence $R^{438}NGFQ^{442}$ using the 5D (H)NCOCαNCO experiment. Each light blue panel represents a ($^{13}C\alpha^{i}$, $^{13}C'^{i-1}$) plane with given $^{15}N^{i}$ frequency value from the auxiliary 3D experiment, and each dark blue framed spectra represents a ($^{15}N^{i+1}$, $13C'^{i}$) 'hyper-plane' at a given (^{15}N, $^{13}C\alpha$, $^{13}C'$) frequency set. For assignment, the cross-peak correlating the frequencies of $^{13}C`^{437}/^{13}C\alpha^{438}$ with $^{15}N^{438}$ is selected from a 3D spectrum. A 'hyper-plane' corresponding to these coordinates is then extracted from the 5D spectrum, affording the two additional frequencies of $^{13}C`^{438}$ and $^{15}N^{439}$. This process is then repeated for the newly assigned $^{13}C'^{438}/^{13}C\alpha^{439}$ and $^{15}N^{439}$, successively assigning all residues in the N → C direction.

The ubiquitous presence of proline residues required a modification of the original 5D-NMR sequence, due to the unique spin topology of proline ^{15}N nuclei which are coupled to three aliphatic ^{13}C nuclei instead of the usual two. The 50 ms constant-time evolution was originally designed to direct magnetization transfer via the stronger intra-residual interaction and suppress the unwanted inter-residual connectivity, utilizing the difference between the $^{1}J[^{15}N_i, ^{13}C^{\alpha}_i]$ and $^{2}J[^{15}N_i, ^{13}C^{\alpha}_{i-1}]$ couplings (5). However, a side effect of this transfer scheme is a discrimination against the desired intra-residual connectivity in prolines. Therefore, a second version of the experiment was acquired with the ^{15}N evolution period was shortened to 33 ms. The intra- and inter-residual correlations could still be correctly identified on the basis of their intensities.

The $^{1}J(NC^{\alpha})$-selective HCBCANCO experiment (7) was measured with the spectral widths set to 6010 (aq) × 2000 (^{15}N) × 10000 ($C^{\alpha/\beta}$) × 3125 ($H^{\alpha/\beta}$) Hz, and maximum acquisition times of 8, 7, and 40 ms for $^{1}H\alpha/\beta$, $^{13}C\alpha/\beta$, and ^{15}N dimensions, respectively. The experiment was measured with 1024 complex points in the directly detected dimension and 2000 randomly distributed hypercomplex points in the indirectly detected dimensions with 8 scans per increment and the recovery delay set to 0.75 s.

Relaxation experiments. Relaxation delays of 0, 0.2, 0.4, 0.6, 0.8, and 1.0 s were used for measuring ^{15}N longitudinal relaxation (R_1), and spin-lock pulse durations of 4, 24, 44, 64, 94, and 124 ms were applied for measuring the ^{15}N rotating-frame transverse relaxation ($R_{1\rho}$) using a 2.1 kHz spin-lock pulse. In both cases delay durations were randomized and the reference spectrum was repeated after acquisition of all spectra to exclude sample degradation effects. To account for offset effects, R_2 was determined using the relation $R_2 = (R_{1\rho} - R_1\sin^2\theta)/\cos^2\theta$, where $\tan\theta$ is the ratio between the ^{15}N offset and the 2.1 kHz pulse strength. ^{15}N-$\{^{1}H\}$-NOEs were estimated by comparison of two HSQC-like spectra with excitation on steady-state ^{15}N magnetization, with and without saturation of the H^N nuclei. Saturation was effected by a series of 150° pulses for the duration of the recycling delay (4 s).

Supporting References

S1. Cai, M., Y. Huang, K. Sakaguchi, G. M. Clore, A. M. Gronenborn & R. Craigie. (1998). An efficient and cost-effective isotope labeling protocol for proteins expressed in Escherichia coli. *J. Biomol. NMR* **11**, 97-102.

S2. Bermel, W., I. Bertini, I. C. Felli, Y. M. Lee, C. Luchinat & R. Pierattelli. (2006). Protonless NMR experiments for sequence-specific assignment of backbone nuclei in unfolded proteins. *J. Am. Chem. Soc.* **128**, 3918-3919.

S3. Bermel, W., I. Bertini, I. C. Felli, R. Peruzzini & R. Pierattelli. (2010). Exclusively heteronuclear NMR experiments to obtain structural and dynamic information on proteins. *Chemphyschem* **11**, 689-695.

S4. Motackova, V., J. Novacek, A. Zawadzka-Kazimierczuk, K. Kazimierczuk, L. Zidek, H. Sanderova, L. Krasny, W. Kozminski & V. Sklenar. (2010). Strategy for complete NMR assignment of disordered proteins with highly repetitive sequences based on resolution-enhanced 5D experiments. *J Biomol NMR* **48**, 169-177.

S5. Novacek, J., A. Zawadzka-Kazimierczuk, V. Papouskova, L. Zidek, H. Sanderova, L. Krasny, W. Kozminski & V. Sklenar. (2011). 5D ^{13}C-detected experiments for backbone assignment of unstructured proteins with a very low signal dispersion. *J Biomol NMR* **50**, 1-11.

S6. Kazimierczuk, K., A. Zawadzka & W. Kozminski. (2008). Optimization of random time domain sampling in multidimensional NMR. *J Magn Reson* **192**, 123-130.

S7. Novacek, J., N. Y. Haba, J. H. Chill, V. Sklenar & L. Zidek. (2012). 4D non-uniformly sampled HabCabCON/intra-HabCabNCO experiments for the sequential assignment and chemical shift analysis of intrinsically disordered proteins. *J. Biomol. NMR.* **53**, 139-148.