

Reconstruction of Coupled Intra- and Interdomain Protein Motion from Nuclear and Electron Magnetic Resonance

Alexandra Born, Janne Soetbeer, Frauke Breitgoff, Morkos A. Henen, Nikolaos Sgourakis, Yevhen Polyhach, Parker J. Nichols, Dean Strotz, Gunnar Jeschke, and Beat Vögeli*



Cite This: *J. Am. Chem. Soc.* 2021, 143, 16055–16067



Read Online

ACCESS |



Metrics & More

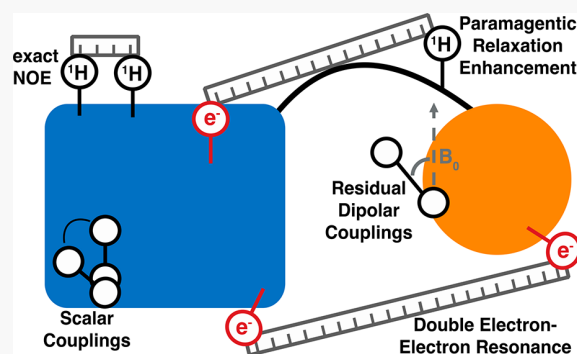


Article Recommendations



Supporting Information

ABSTRACT: Proteins composed of multiple domains allow for structural heterogeneity and interdomain dynamics that may be vital for function. Intradomain structures and dynamics can influence interdomain conformations and *vice versa*. However, no established structure determination method is currently available that can probe the coupling of these motions. The protein Pin1 contains separate regulatory and catalytic domains that sample “extended” and “compact” states, and ligand binding changes this equilibrium. Ligand binding and interdomain distance have been shown to impact the activity of Pin1, suggesting interdomain allostery. In order to characterize the conformational equilibrium of Pin1, we describe a novel method to model the coupling between intra- and interdomain dynamics at atomic resolution using multistate ensembles. The method uses time-averaged nuclear magnetic resonance (NMR) restraints and double electron–electron resonance (DEER) data that resolve distance distributions. While the intradomain calculation is primarily driven by exact nuclear Overhauser enhancements (eNOEs), *J* couplings, and residual dipolar couplings (RDCs), the relative domain distribution is driven by paramagnetic relaxation enhancement (PREs), RDCs, interdomain NOEs, and DEER. Our data support a 70:30 population of the compact and extended states in apo Pin1. A multistate ensemble describes these conformations simultaneously, with distinct conformational differences located in the interdomain interface stabilizing the compact or extended states. We also describe correlated conformations between the catalytic site and interdomain interface that may explain allostery driven by interdomain contact.



INTRODUCTION

Proteins with multiple domains are the norm, not the exception. While over 80% and 67% of eukaryotic and prokaryotic proteins, respectively, include more than one domain, less than 35% of structures deposited in the Protein Data Bank contain multiple domains.¹ Multidomain proteins are often more stable and easier to fold than a single domain while allowing for greater structural and functional plasticity.² Many domains are linked together by an intrinsically disordered region that may act as a hinge to allow for structural heterogeneity between domains while the individual domains internally maintain their 3D structure.^{3–5} Multiple orientations between domains may be vital for the function and activity of a biomacromolecule. Domain orientation and interdomain distance at equilibrium can change with environmental parameters, e.g., ligand concentration.

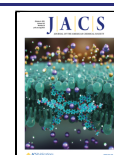
Many properties of a protein, including ligand binding, catalysis, and stability, are influenced by a protein's conformational dynamics. These dynamics range from small movements near an active site to large collective motions over entire domains. In addition, dynamics are believed to be a major factor in allosteric regulation of a protein. Allostery typically

refers to the phenomenon of an effector molecule acting at a distal site that regulates the function of a catalytically active site. For calmodulin and Pin1, both two-domain proteins, it has been shown that ligand binding changes the orientation of the two domains.^{4,6–10} In order to sustain a long-range allosteric mechanism, local motions must be correlated. Although chemical shifts, order parameters, and catalytic activities have been used to characterize allosteric behavior, these signatures only indirectly report on the underlying dynamics, and it has thus been difficult to accurately determine conformational differences.

Studying an isolated domain is useful to learn about its basic structure and function, yet tethering multiple domains together has been shown to significantly alter the structure in 50% and dynamics in 90% of the protein domains studied.¹¹ Therefore,

Received: June 18, 2021

Published: September 27, 2021



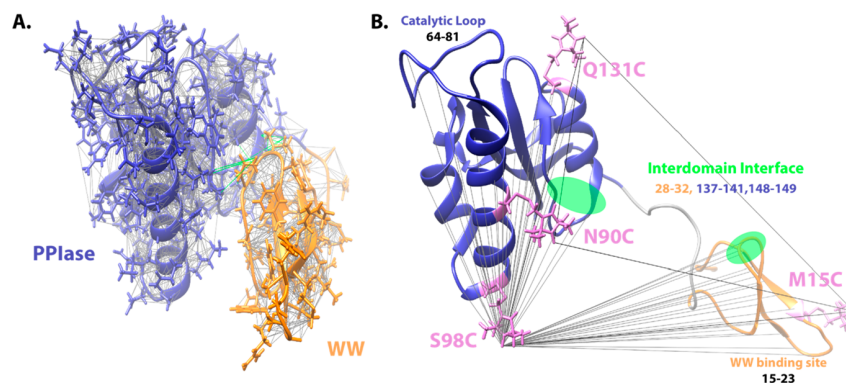


Figure 1. Structure of Pin1. (A) Crystal structure 1pin⁴³ showing the “compact” state with the WW (orange) and PPIase (blue) domains with all eNOEs plotted. Interdomain NOEs are colored in green. (B) Conventional NMR structure 1nmv⁴⁵ showing the extended state. MTSL mutations are colored in purple, and DEER restraints and 98-MTSL PRE distances are overlaid in black and gray, respectively.

investigating individual domains is not sufficient for generating a full understanding of a multidomain protein. Whereas recent computational approaches can predict multidomain protein configuration and interdomain interfaces by *ab initio* folding potential¹² and global structural alignments,¹³ experimental studies of multidomain proteins face a number of challenges in orienting multiple domains. Crystal structures typically capture only one conformation and may feature packing artifacts. Many proteins of interest are too small for cryoelectron microscopy (for now), whereas recent advances in solution NMR have led to structural restraints of macromolecules larger than 100 kDa.¹⁴ Even though solution NMR allows interrogation of structure and dynamics of such multidomain systems, relying solely on interdomain nuclear Overhauser enhancement (NOE) is not sufficient to orient multiple domains.¹⁵ ¹⁵N relaxation data, paramagnetic relaxation enhancement (PRE), and residual dipolar couplings (RDCs) have successfully been utilized to describe interdomain orientation and dynamics,^{9,10,15–20} but these motions have not been linked to intradomain motion.

Here, we introduce a novel method to solve the structure of a two-domain protein that allows for coupling between intra- and interdomain dynamics at atomic resolution. We aim to identify the structural correlations between intradomain structure and interdomain positions. Our approach builds on previous work by the Clore and Vendruscolo labs that proposed to identify correlations within single domains by a combination of conventional NOEs, scalar couplings, RDCs, and relaxation order parameters.^{21–23} The method presented here is primarily based on an innovative combination of both emerging NMR and EPR techniques for precise short- and long-range distance measurements, supplemented by angular restraints. These measurements yield time-averaged short-range distances (within domains) and probability distributions of long-range distances (between domains). The restraints we use in our calculations include exact NOEs, scalar couplings, RDCs, PREs, and double electron–electron resonance (DEER).

Recent advances in the exact quantitative evaluation of the NOE (eNOE)^{24–27} allow us to measure proton–proton distances with less than 0.1 Å error up to 5 Å in favorable cases,^{27,28} and we can detect distances up to ~8 Å. As these eNOEs are motion- and population-averaged observables, structure calculations based on multiple states often achieve a

better agreement with experimental data than an averaged model.²⁵ As such, our method is able to characterize and depict correlated motions from experimental data.

Besides the few eNOEs we observe between the two domains at the interface, we rely on PRE, RDC, and DEER to access longer distances and thereby generate the following restraints. First, PRE allows distance measurements up to 25 Å between the unpaired electron of a paramagnetic spin label and nuclei. Although it does not offer the accuracy of eNOE, it is a powerful tool to characterize the domain orientations and motions. Second, we use RDCs to determine the relative orientation of the two domains. RDCs report on the bond orientation within molecule-fixed frames and thus carry long-range information. Importantly, PREs have also been shown to reduce the degeneracy of RDCs in a multidomain protein and to detect transient, minor states in an exchanging system.^{29,30} Both PRE and RDCs can be used to determine interdomain dynamics, orientation, and motion, yet the averaging of these parameters in NMR is convoluting. Therefore, we also utilize DEER to measure interdomain distances and their populations. DEER measurements are performed using flash-frozen samples so that this electron paramagnetic resonance (EPR) technique provides a distance distribution between spin labels instead of a solution-averaged distance. Importantly, the distances obtained at low temperature can be reliably combined with NMR data at room temperature, due to the slow exchange between the interdomain positions. The distance range depends on the spin environment, and for biological systems typically distances between 15 and 50 Å can be characterized by 4-pulse DEER (4pDEER).³¹ Here, we also use a dynamically decoupled version of the experiment with 5 pulses (5pDEER).³² Together with recent advances in microwave technology that can be exploited to achieve favorable excitation bands, 5pDEER can extend the accessible distance range up to 80 Å.³³ Often used in conjunction with other biophysical techniques such as crystallography, NMR, and small-angle X-ray scattering (SAXS), DEER has determined the orientation of multidomain systems, namely, the HIV-1 RT p66 homodimer,³⁴ fibronectin type III domains of integrin $\alpha 6 \beta 4$,³⁵ tandem POTRA domain pair of Bama,³⁶ calmodulin,³⁷ and the *E. coli* 5'-nucleotidase.³⁸ While the integrin and Bama DEER distance distributions were narrow and indicative of only one conformation, an additional broader ensemble was detected in the case of the p66 reverse

transcriptase and nucleotidase domains. The nucleotidase system even showed that substrate binding could change the equilibrium from completely open to a mixed population of an open and closed populations.³⁸

We applied our method to the two-domain mitotic regulator Pin1. Pin1 (protein interacting with NIMA kinase 1) is a 163 residue peptidyl-prolyl isomerase specific for isomerizing prolines that are immediately preceded by a phosphorylated serine or threonine (pS/TP).^{39–41} Residues 1–39 form the WW interaction domain (named for two conserved tryptophans) that features a three-stranded, antiparallel β -sheet which binds the pS/TP motif *trans*-specifically.⁴² Residues 50–163 form the catalytic PPIase domain responsible for isomerizing the proline in the same motif, and it is composed of a four-stranded core β -sheet with four exterior α -helices.^{43,44} A 10-residue flexible linker separates the two domains.

According to the first crystal structure of Pin1, the two domains assume a “compact” conformation with an interdomain interface composed of residues 28–32 in the WW domain interacting with residues 137–142 and 145–149 in the PPIase domain (Figure 1A).⁴³ The PPIase interdomain interface is located on the opposite side of the catalytic site. Such interaction between the two domains was lacking in an initial NMR ensemble as the domains were found to be in an extended conformation (Figure 1B).⁴⁵ Subsequent NMR relaxation experiments proved that Pin1 tumbles somewhere between two independent domains as a single, rigid unit. Interestingly, ligands or point mutations change the equilibrium of compact and extended states dependent on the ligand sequence. These shifts lead to changes in the catalytic activity of Pin1 via an allosteric mechanism^{42,46–51} and suggest that different conformations of the individual domains exist that stabilize the compact and extended states. Our previously solved eNOE two-state structure of the isolated WW domain of Pin1 supports this notion.⁵² Addition of two different ligands either conserves the correlations between the ligand-binding site and the WW/PPIase interface or induces partial anticorrelation. Importantly, anticorrelation renders both states incompatible with a compact conformation when the PPIase domain is modeled in its typical compact position. This possibly explains why one of the two ligands shifts the equilibrium toward the open conformation. In order to verify this hypothesis, a multistate representation of the full-length Pin1 at atomic resolution is required. Although many structures of Pin1 have been solved, none of them depict Pin1 at equilibrium with compact and extended states.

Yet some work has been done to determine the general orientations and distances in the extended states without considering the spatial sampling within the domains.^{9,53} First of all, while NMR relaxation experiments can report on the degree of compactness in a multidomain protein, this method is unable to provide information about the actual domain positions.^{8,15} Second, RDCs have been utilized on wild-type and I28A mutant Pin1 to generate long-range, orientational bond-vector restraints based on the incomplete averaging of dipole–dipole interactions.⁹ These RDCs were used in conjunction with a Langevin dynamics simulation method optimized for large conformational changes.^{9,20} It is difficult to determine domain positions solely by RDCs due to ambiguities with respect to orientation and insensitivity to translation. Lastly, PRE using a single paramagnetic label has been measured on a Pin1 construct to evaluate the domain distance upon the addition of PEG400.⁵⁴ A recent study also made use

of PREs induced by single label (at position H27).⁵⁵ This study revealed further interdomain contacts located in the first two α helices and connecting loop. Importantly, it also presents PRE evidence that interdomain separation is correlated with compaction of the WW domain.

We demonstrate the power of our method that integrates eNOE, *J* coupling, RDC, PRE, and DEER data to elucidate coupled intra- and interdomain motion by solving multistate ensembles of full-length apo Pin1. Our two-state ensemble satisfies all NMR data and reproduces the compact and extended states. The obtained domain positions are also in close agreement with the DEER distance distributions. We observe distinct intradomain conformations correlated to interdomain distance and propose how the intradomain conformations may stabilize the compact and extended states. Finally, we show structural changes in the catalytic site that are correlated to interdomain contact, supporting a model of interdomain allostery.

RESULTS AND DISCUSSION

Data Collection. Tumbling Time-Specific and Interdomain eNOEs. We applied our eNOE buildup method on Pin1 to extract precise distances between protons with the ultimate goal of determining an accurate multistate structural ensemble. This method is well-established for single-domain proteins of various sizes as well as RNA.^{25,27} Here, we apply this method for the first time to a multidomain protein. In order to convert a cross-relaxation rate into an effective exact distance, it is imperative to determine the tumbling time (τ_c) of the buildup sample. For a single-domain globular protein, the overall tumbling time suffices to determine an accurate distance. In the case of Pin1, its two domains are known to tumble partially independent of one another^{8,56} and we therefore rely on R_1 and $R_{1\rho}$ measurements to quantify the domain-specific tumbling times as 11.3, 14.1, and 3.6 ns for the WW, PPIase, and linker, respectively. By implementing residue-specific τ_c in the eNORA2 program in CYANA, we were able to extract 537 bidirectional and 1731 unidirectional eNOEs as superimposed on the crystal structure in Figure 1A. In addition, 1911 generic-normalized (gn) eNOEs were also determined from spins, where the diagonal decays could not be fitted.⁵⁷ Furthermore, we added 124 $^3J_{\text{HN,H}\alpha}$, 129 $^3J_{\text{H}\alpha,\text{H}\beta(2,3)}$ (also used for the stereospecific assignment), and 12 aromatic $^3J_{\text{N,C}\gamma}$ scalar couplings to increase the intradomain restraint density.

The original NMR structural work of Pin1 (1nmv)⁴⁵ lacked NOEs between the WW and PPIase domain (interdomain NOEs, ID NOEs). However, we were able to identify 20 eNOEs (3 uni-, 10 bi-, and 7 gn-eNOEs) that according to the crystal structure, 1pin,⁴³ lie in the region of the interdomain interface (Figure 1A). Presumably, our higher sample concentration (2 mM) and a 900 MHz cryoprobe spectrometer increased the sensitivity toward ID NOEs compared to room temperature measurements performed on a 800 MHz spectrometer using a sample concentration of 0.6–0.8 mM.⁴⁵ The interface is mostly composed of hydrophobic residues so that half of the eNOE spin pairs were associated with the methyl groups of these residues. As a dynamic system the NOEs are averaged, hence, ID NOEs provide evidence for the compact state, but the simultaneous existence of extended states cannot be discounted because the cross-relaxation rate is dominated by short sampled distances.

Residual Dipolar Couplings To Orient Domains. We measured RDCs on the full-length protein to aid orienting the

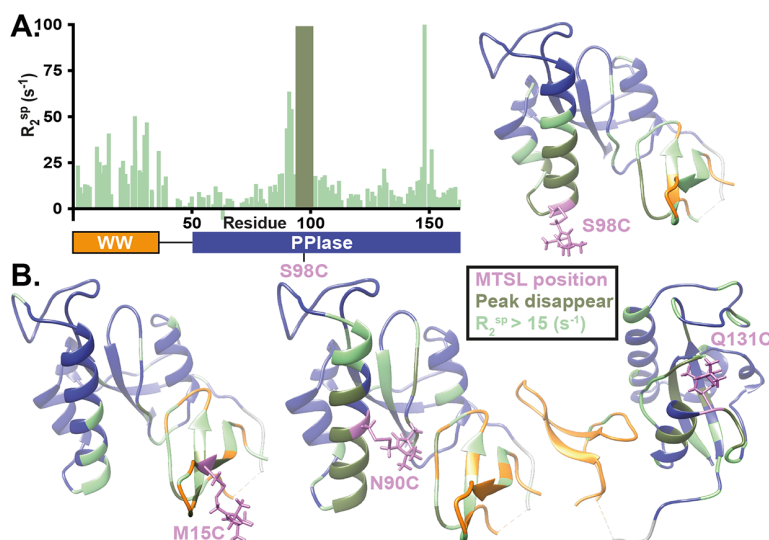


Figure 2. PRE of Pin1. (A, left) Residue vs spin-enhanced relaxation rates of 98-MTSL. (A, right) Major R_2^{sp} from 98-MTSL plotted on the structure 1pin.⁴³ (B) Major R_2^{sp} from 15-, 90-, 131-MTSL plotted on the structure 1pin.⁴³

two domains. RDCs are still averages over dynamic ensembles, and Langevin dynamic simulations have previously been used on Pin1 to generate conformational ensembles that freeze intradomain motion while allowing collective interdomain motion.⁹ Using C12E5 PEG/hexanol, we were able to determine 407 RDCs (140 $^1D_{Ni,HNi}$, 138 $^1D_{C'i,Cai}$, 103 $^1D_{C'i,Ni+1}$, and 97 $^1D_{C'i,HNi+1}$). Given the size difference of the WW and PPIase domains (39 vs 113 residues, respectively) and the fact that the alignment in PEG/hexanol is primarily steric in nature, we expect that the PPIase will cause a greater alignment with the magnetic field. We note that a more favorable way of achieving the alignment of a two-domain protein would be the attachment of a lanthanide tag^{58,59} such that the alignment would be entirely independent of the relative domain positions. In addition to orient one domain relative to another, we also aimed to determine the proper orientation of bond vectors within the domain. While the fit of the WW domain alignment tensor using the X-ray structure (1pin) was excellent (measured vs back-predicted RDCs, $r = 0.84$), our previously solved eNOE-structure of the isolated WW domain (6svc⁵²) achieved an even better agreement ($r = 0.94$), reinforcing the potential of eNOEs. Our ensemble calculation consists of two steps. First, using the WW alignment tensor, we determined bond vector orientations of the WW domain. In the second step, the RDCs of the entire Pin1 were fitted using a PPIase-specific alignment tensor, obtained from an initial fit to 1pin ($r = 0.92$, after removal of outliers 0.96). For this step, the angles of the WW domain were frozen, allowing the domain to move as a rigid body relative to the PPIase domain.

PRE and DEER Mutants for Long-Range Distance Restraints. We relied on PRE and DEER for long-distance restraints that define the relative domain translation and rotation. For these paramagnetic techniques that require adding a nitroxide spin label, we engineered constructs of Pin1 to probe interdomain distances with minimal disruption of the wild-type conformation and enzyme activity. As the MTSL spin label conjugates to free cysteines, we needed to (i) mutate endogenous cysteines (C57 and C113) and (ii) introduce cysteines to various regions of the protein. The

fairly conservative C57S mutation caused severe protein precipitation, while the C57A mutation maintained stability. Originally the endogenous C113 was considered essential for isomerase activity,^{60–62} although recent work has demonstrated that a mutation to aspartate conserves the protein's activity.^{63,64} Therefore, all PRE and DEER mutant constructs also feature C57A and C113D mutations. Using the structure as a guide, we introduced cysteines at M15 (in the WW domain), and N90, S98, and Q131 (in the PPIase domain) as shown in Figure 1B. Chemical shift perturbations were measured on all PRE samples to ensure the absence of long-range perturbations that are indicative of disruption of the overall structure and orientations, as shown in Figure S2. Table S1 reports the isomerase activity of WT Pin1 and of these stable PRE constructs, Figure S3 displays example spectra and fits. All PRE mutants maintained some activity, while mutants M15C and N90C displayed nearly identical activity as WT Pin1. These four PRE constructs were then combined into double mutants for DEER spectroscopy, and Table S1 also reports the associated isomerase activities.

Paramagnetic Relaxation Enhancement across Domains. Due to the high gyromagnetic ratio of an electron, PREs can probe distances up to 25 Å.⁶⁵ We extracted an electron–nucleus distance by measuring the difference of transverse relaxation between a MTSL-labeled and DTT-quenched sample (Figure S4A), termed “spin-enhanced” relaxation rate (R_2^{sp}). While acetyl-MTSL is commonly used for the diamagnetic sample, DTT and ascorbic acid can also be used to remove the MTSL allowing for the use of the exact protein sample for both diamagnetic and paramagnetic measurements.^{54,65–67} Figure 2A displays the measured R_2^{sp} of construct S98C as a function of residue number. Peaks within 13 Å completely disappear from the spectra, while peaks up to 25 Å show quenching due to R_2^{sp} . Residues with an R_2^{sp} greater than 15 s^{-1} are located in a ~ 19 Å sphere surrounding the MTSL spin label and are highlighted in the crystal structure (Figure 2A). Figure 2B shows the analogous representation for M15C, N90C, and Q131C, with associated residue-resolved graphs in Figure S4B–D. In addition to intradomain quenching, constructs M15C, N90C, and S98C feature many

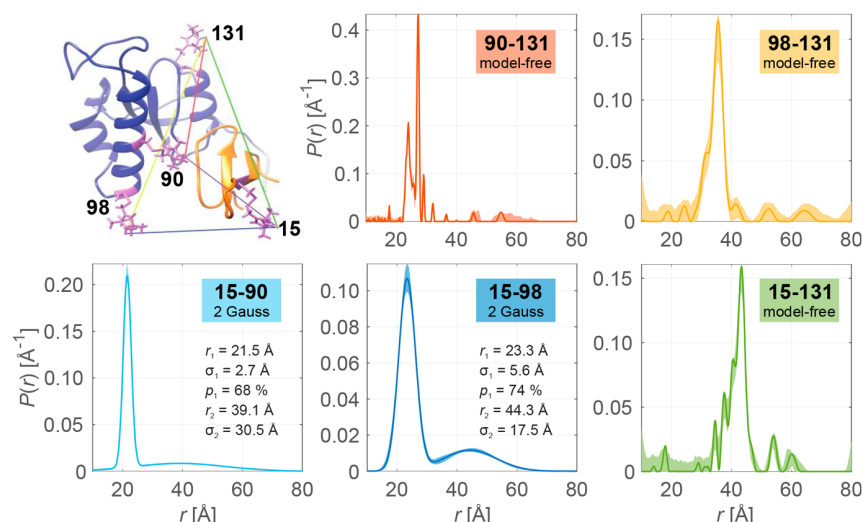


Figure 3. Distance distributions determined by DEER measurements of double-MTSL Pin1 constructs. The experimental population density $P(r)$ and 95% confidence interval are shown for each construct. For 15–90 and 15–98 mutants, the data could be fitted well using a bi-Gaussian $P(r)$, and average distances (r), standard deviations (σ), and populations (p) are reported for the two components.

residues that are quenched across domains, validating this method to probe interdomain distances in Pin1. Similar to the NOE, the R_2^{SP} is solution-averaged, rendering the deconvolution of multiple sampled distances challenging. Nevertheless, on the basis of the increased quenching at shorter distances, PRE has been used to infer the presence of a transient state that is much more compact than the extended state of the major conformer.^{68,69} For Pin1, we are unable to determine populations of these states through PRE alone.

Double Electron–Electron Resonance for Probing Distance Distributions. We combined the labeling sites used for the PRE constructs to form six double mutants for EPR spectroscopy. DEER measurements exploit the dipolar coupling between two paramagnetic labels to detect distances in the range of 15–80 Å. Importantly, DEER can be used to determine distance distributions, i.e., all distances present in the sample at the time of freezing (samples are flash-frozen before measurement). We utilized this information to supplement the NMR-based methods employed here, which read out averaged distances.

Figure 3 shows distance distributions between the nitroxide labels for five of the six double mutants. 90–131 and 98–131 are intradomain distances in the PPIase domain, which could be determined by means of 4pDEER. For the three interdomain (ID) distances between the WW and PPIase domains, it was crucial to employ 5pDEER to reliably measure the longer distances associated with extended states. Whereas the DEER data recorded for the double mutants 98–131, 90–131, and 15–131 could only be fit using a model-free distance distribution (Figure S6), for the constructs 15–90 and 15–98 two distinct dipolar oscillations were observed that could be equally well described by two Gaussian components (Figure S7A,B). The regularization parameters and artifact correction were optimized as the bimodality in these distance distributions influence the width of the peaks. For the double mutant 90–98 the observed dipolar oscillation (Figure S5D) could not be translated to a distance distribution because the close proximity of the spin labels (~ 13 Å expected) violates the point-dipole approximation.

Upon comparing the distance distributions in Figure 3, it becomes evident that distance distributions involving position 90 are typically narrower than for constructs featuring position 98. This suggests that the spin label at position 90 is more restricted in conformational freedom compared to the other labeling positions. The broadened line shape of the room-temperature continuous-wave (CW) EPR spectrum of 15–90 compared to 15–98 (Figure S5A) supports this conclusion as it reports on the restricted motional freedom at the center of the helix (residue 90) relative to the helix end (position 98).

Populations of Extended and Compact States. For 15–90 and 15–98, the DEER-derived distance distributions feature a relatively narrow contribution centered around 21.5 and 23.3 Å, respectively, which matches the compact state seen in the crystal structure 1pin (Figure S8A). In addition, a longer, more dispersed distance is observed for these two constructs which is more akin to the domain distributions seen in the original, extended NMR structure 1nmv (Figure S8B). The accurate description of these distance distributions by two Gaussian components allows us to determine the populations of the two. On the basis of the 95% confidence interval of the population value for the shorter, compact state that extends from 0.67 to 0.72 for 15–90 and from 0.72 to 0.77 for 15–98 (Table S3), we propose that the populations of the compact and extended states are $\sim 70\%$ and $\sim 30\%$, respectively. We note that since each measurement is carried out independently, we cannot determine from these data alone if the major/minor populations for these two mutants are correlated. Whereas involved approaches have been proposed to achieve this,⁷⁰ the correlation in our specific case will become obvious once we calculated structural ensembles (*vide infra*).

Furthermore, the DEER distributions were used for cross-validating the multistate structure calculations based on NOEs, RDCs, and PREs.

Major Conformation: Compact or Extended? We should note that the only other study that quantified the populations of apo Pin1 reported near opposite populations with 71% for the extended state and 29% for the compact state.⁵⁰ This study was based on a chemical shift correlation analysis and on small-angle X-ray scattering (SAXS) data, though neither method

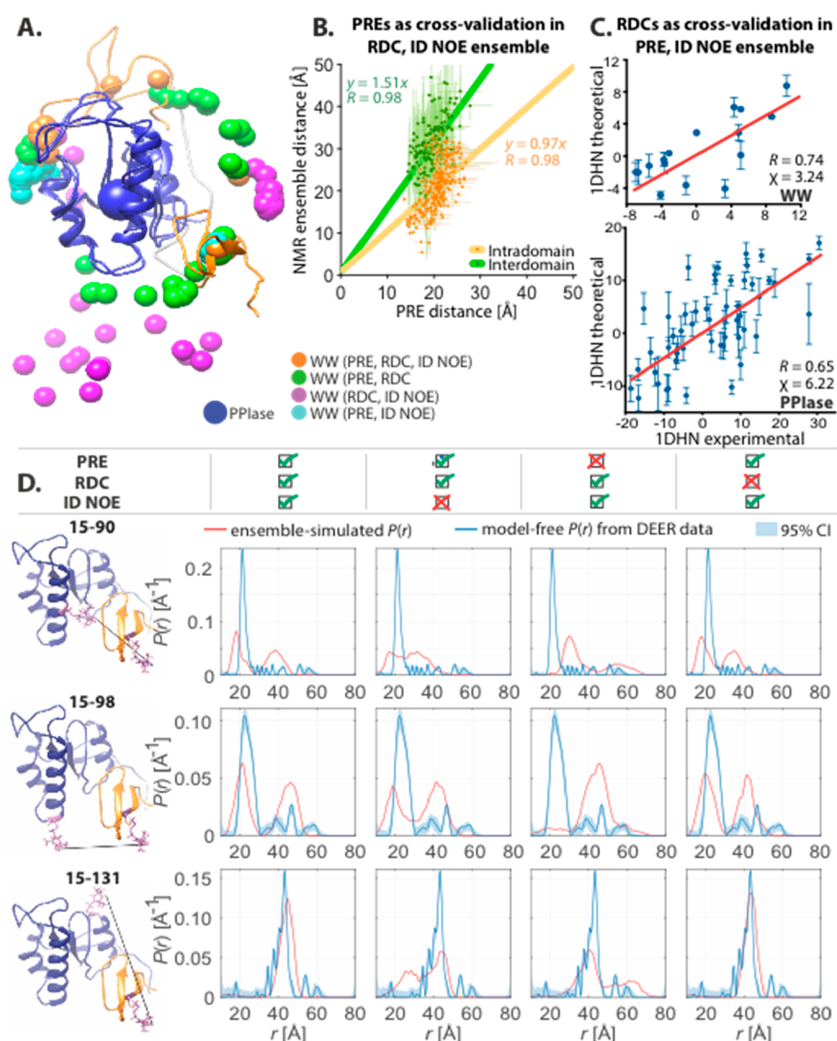


Figure 4. Domain positioning of two-state Pin1 structures. (A) With the PPIase domain overlaid, positions of the center of the WW domain are shown color-coded by the interdomain restraints used to determine the domain orientations. (B) PREs as cross-validation in the structure solved with RDCs and ID NOEs and ensemble error. (C) RDCs as cross-validation in the structure solved with PREs and ID NOEs with ensemble error. (D) Back-calculated DEER distributions from various structure calculations overlaid with experimental DEER distance distributions including the associated 95% confidence interval (CI). Distributions from individual two-state conformers are shown in Figure S8C in the Supporting Information.

directly measures distances between the two domains. It is possible that some distances are beyond the EPR detection limit of >80 Å in the extended state (as we see with the 1nmv back-calculated DEER simulations in Figure S8B), but even when we consider that errors in distance are also expected to increase upon distance increase for a given DEER trace length, it is unlikely that this contribution is large enough to completely interchange the populations. As seen in previous work,^{54,71} many NMR peaks in the interdomain interface have a major and minor species in slow-exchange. Our NOE and relaxation data associated with these slow exchange peaks suggest that the major peak originates from the compact state. For example, Figure S9 shows the major and minor peaks of WW residue T29 in the ^{15}N -HSQC but with NOE cross peaks to PPIase atom 140 H β^* (also atom 137 H α , not shown) only in the major state. This observation implies the major state has interdomain contact and thus corresponds to the compact conformation, while the minor state lacks on average the contact required to produce a NOE. In addition, the slow

exchange peaks in the ^{15}N -HSQC were sufficiently resolved to extract distinct R_1 and $R_{1\rho}$ relaxation rates (Table S5). On the basis of the size and domain behavior of Pin1, more flexible regions are characterized by higher R_1 and lower $R_{1\rho}$ rates, and the interdomain residues in the extended conformation are expected to be less restricted than in the compact conformation. Specifically, for the six interdomain residues with resolved major and minor peaks, the minor peaks are associated with higher R_1 and lower $R_{1\rho}$ rates, implying occupation of the extended state. These data provide evidence that the compact conformation corresponds to the major state of apo Pin1, which further supports the DEER distance distributions.

Multistate Structure Calculation. PREs as Distance Restraints. We converted the PRE relaxation rates into distances (see Supporting Information). Compared to the high precision eNOEs (tenths of Å versus 2–4 Å), PRE restraints correspond to longer distances and are associated with larger uncertainties. Because in CYANA both restraints

are formally used in the same way, the PRE-derived distance restraints would dominate. Therefore, we first excluded the intradomain PREs from the structure calculations (for details on the structure calculations, see the [Supporting Information](#)), though the distances of the obtained structure correlate well with these intradomain PRE distances ($y = 0.97x$, $r = 0.98$, [Figure 4B](#)). Second, we optimized the weight of the interdomain PRE relative to the eNOEs restraints for a two-state structure calculation (as preliminary evidence suggested that two states are sufficient for satisfying the data; see [Figure 6A](#)) that involved all scalar couplings, angle restraints, eNOEs, and RDCs (as described earlier). Note that for this test our multistate ensemble method enforces a 50:50 population, though the DEER data report a major population difference between compact and extended states (this will be addressed later). We calculated five two-state ensembles with various PRE weights (equivalents of NOE weights 1, 0.5, 0.1, 0.01, and 0.001) and evaluated the impact of the different weights on the integrity of the local structure ([Figure S10](#)). As expected, lowering the weight reduces the target function ([Figure S10B](#)); however, the latter nearly levels off at a weight of 0.01, whereas PRE weights greater than 0.01 caused local structural disruptions near spin-labeled sites. This effect is most apparent in the long PPIase helix ([Figure S10C](#)), as a bulge formed near N90 in order to accommodate the PRE restraints to the WW domain. We evaluate the domain orientation by back-calculating DEER distance distributions from these ensembles using the EPR distance simulator within MMM⁷² and overlaying them with our experimental distributions ([Figure S10A](#)). While PRE weights of 1 and 0.5 pull the ID distributions into unrealistic conformations, these DEER distance distributions independently validate the use of a 0.01 PRE weight. All these data confirmed that weighting the PRE data by 0.01 (or reducing the weight by a factor of 100) is sufficient to run calculations that eliminate local artifacts caused by PREs overwriting local restraints. We therefore applied this weight in the following calculations.

Reproduction of EPR-Derived Distance Distributions by Conformation-Averaged NMR Probes. It is not clear *a priori* whether the solution ensemble calculated from averaged NMR restraints will agree with the experimental DEER distance distributions and populations. Therefore, we tested if PRE, RDC, and ID NOEs can be combined to match the range of conformations present in the DEER distance distributions. The NMR restraints for the ensemble calculations are summarized in [Table S6](#).

First, we optimized the calculation to ensure that our ensemble has converged by increasing the number of torsion angle steps and the number of structures calculated ([Figure S10D](#)). Each trial was started from 200 calculations and 50 000 torsion angle steps (as this number is ideal for small proteins) and resulted in similar TF of the ten two-state structures with the lowest values. Increasing the number of steps decreased the RMSD in the PPIase domain, while increasing the total number of calculated structures helped to lower the total RMSD in both domains. For the following calculations, we use 400 calculations each consisting of 100 000 torsion angle steps.

Second, we calculated two-state ensemble structures using our protocol that integrates all NMR probes (eNOE distance restraints, scalar couplings, PREs, RDCs, and ID NOEs). We compare the WW domain positions relative to the PPIase, as visualized in [Figure 4A](#), by overlaying the associated MMM-simulated DEER distance distributions for the calculated

structures (that did not include DEER data in the calculations themselves) with the distance distributions obtained by DEER experiments ([Figure 4D](#)). For an NMR ensemble, MMM will uniformly average the distribution over all conformers, and we show an example of the distributions from individual two-state conformers in [Figure S8C](#). The difference in peak intensity is due to the individual conformers producing a narrow distribution while forcing the same population as the ensemble-averaged distributions. We obtained two distance clusters for 15–90 and 15–98. One narrower cluster is positioned at a short distance of around 18 Å for 15–90 and 21 Å for 15–98, while a broader contribution is centered at a longer distance of 38.5 Å for 15–90 and 46 Å for 15–98. Instead, for 15–131 we only obtain one relatively narrow distribution centered at around 44.5 Å. As mentioned in the previous section, we enforce two states of equal population so that by construction our two-state ensembles yield almost a 50:50 population when two clusters are present. Note that this population could be skewed if one population is much smaller so that both states are more likely to be assigned to the same cluster. Yet our two-state structures place one state into each cluster (rather than both into the compact cluster), which further supports our population estimate of 70:30. Despite the described limitation of our model, the NMR ensemble reproduces the EPR distance distributions remarkably well. First, EPR also detects a bimodal distance distribution for 15–90 and 15–98 though the cluster at shorter distance is more populated. For 15–90, the latter feature is narrower and centered at a larger distance (21.5 Å in EPR, 18 Å in NMR ensemble), while the second cluster is very broad and difficult to quantify in terms of center and width. Many interdomain orientations in our ensembles are also supported by a previous study.⁵⁵ Second, the single cluster obtained for 15–131 is in almost perfect agreement with distance distributions obtained by EPR. We conclude that integration of PRE, RDC, and ID NOEs allows modeling the relative positions of the two domains, as the corresponding interdomain distances agree with DEER experiments, which are able to resolve multimodal distance distributions.

Impact of Different Restraints on Interdomain Orientations. After we optimized our calculations, we cross-validated the different probes used in this study to determine the minimal data needed to restrain the two domains. For the two-state ensemble calculated without ID NOEs, the RDCs and PREs themselves restrain the WW domain along a wide plane that includes the compact position. When the PRE restraints are excluded (but the RDCs and ID NOEs are used), the conformation of the two domains is too extended as the back-calculated interdomain PRE distances are too long compared to the experimental PREs ([Figure 4B](#)). This is also mirrored in [Figure 4A,D](#) that shows that the compact state is not even present in this ensemble. In addition, all ID NOEs were violated in this calculation, suggesting that the ID NOEs alone are not sufficient to produce the compact state, and the PREs are required to generate this conformation. However, we expect that increasing the weight of the ID NOEs may also generate the compact state. On the other hand, removing the RDC restraints does not cause a relevant change in the ensemble because, as long as PREs and ID NOEs are used in the calculations, the relative positions of the two domains are nearly identical ([Figure 4A,D](#)). Back-calculating RDCs from this structure results in a reasonably good agreement with the experimental RDCs for the individual domains ($r_{\text{ww}} = 0.74$ and

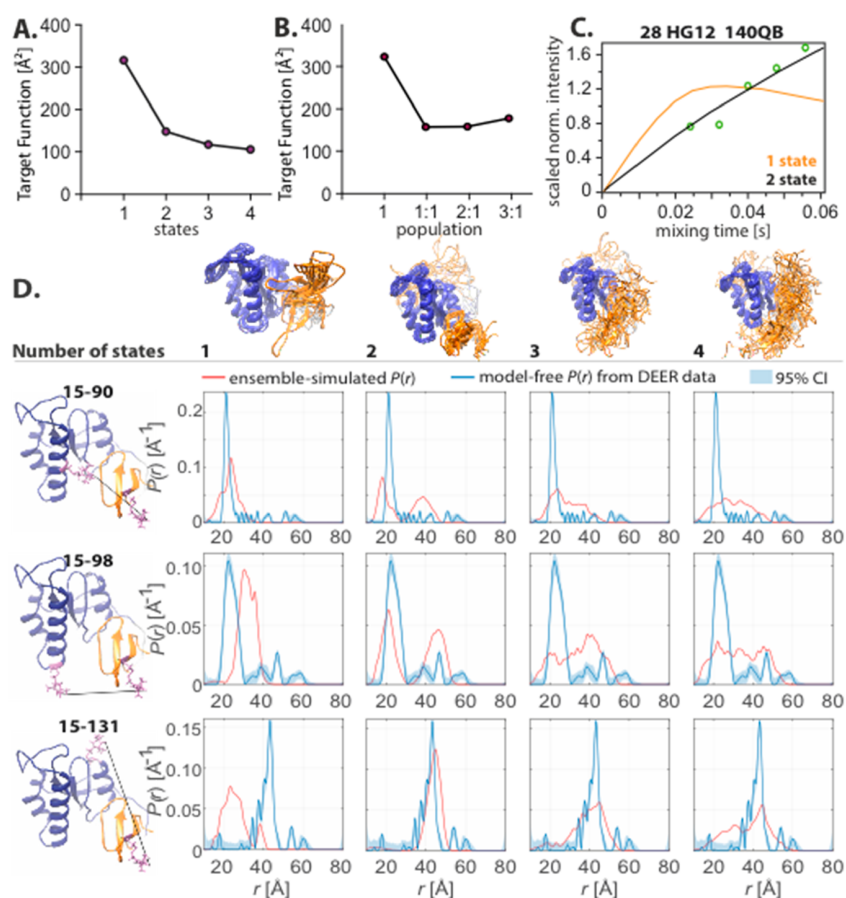


Figure 5. Multistate structures of various numbers of states. (A) Number of states versus CYANA target function. (B) CYANA target function of the two-state structure calculations versus various populations. (C) Experimental unidirectional eNOE buildup intensities (green dots) versus time against back-predicted buildups of representative NOEs for single-state (orange) and two-states (black) ensembles calculated. (D) Relative positions of the two domains in multistate structures with back-calculated DEER distance distributions overlaid with experimental distributions with confidence interval (CI).

$r_{\text{PPIase}} = 0.65$, Figure 4C), though the correlation coefficient of $r_{\text{FL}} = 0.40$ obtained for the full-length Pin1 is significantly smaller. Upon including the RDCs in the calculation, the correlations increase to $r_{\text{ww}} = 0.89$, $r_{\text{PPIase}} = 0.87$, and $r_{\text{FL}} = 0.77$. While the RDCs do not affect the relative global positioning of the two domains, this result suggests that RDCs aid in subtly orienting the individual bond vectors in the internal domain structures. This is perhaps not surprising because RDCs are sensitive to rotation but not to translation. For the compact state, such rotational restraints may be more relevant, whereas the exact orientation of the extended state is less crucial and is possibly partially restricted by the presence of the linker. Overall, the PREs and ID NOEs are sufficient for determining the large-scale conformation of the two domains, while the RDCs support the eNOE restraints in correctly orienting the bond-vectors within the domains as previously demonstrated.⁷³

Two States Are Sufficient for Describing Interdomain Positions. Combining all the PRE, ID NOE, and RDC restraints in addition to our eNOEs and angle restraints, we calculated multistate ensembles of Pin1. We determined the minimal number of states needed to satisfy all the data by checking for convergence of the CYANA target function (TF, proportional to the sum of squared violations) without overfitting (Figure 5A). The TF and NMR restraint violations

of the ten structures with the lowest TF values decrease significantly between the one- and two-state ensembles (TF reduced to less than one-half), as reported in Table S6. Including a third and fourth state further reduces the TF by ~20% and ~10%, respectively. In addition, we also compared the simulated DEER distance distributions of these multistate ensembles to the distributions obtained from experimental DEER distance distributions as shown in Figure 5D. Two-, three-, and four-state ensembles are all in reasonable agreement with the experimental DEER distances. At first glance, the single-state structure may appear similar to the crystal structure 1pin (Figure S8). However, the MMM simulations demonstrate that the 15–98 distance is too long compared to the major DEER peak in the single-state ensemble (~30 Å in ensemble vs 23.3 Å in DEER), while the 15–131 distance is too short (~24 Å in ensemble vs 44.5 Å in DEER). Further, although the contribution to the DEER distribution for 15–90 and 15–98 at longer distances (30–60 Å) is significant, it is not accounted for by the single-state ensemble. As described above, the two-state structure includes both a compact and extended state and matches the experimental data well. In addition, the two-state ensemble models achieve a superior agreement with experimental NOEs buildup intensities accurately compared to the single-state ensemble (Figure 5C, Figure S11). Although a single-state

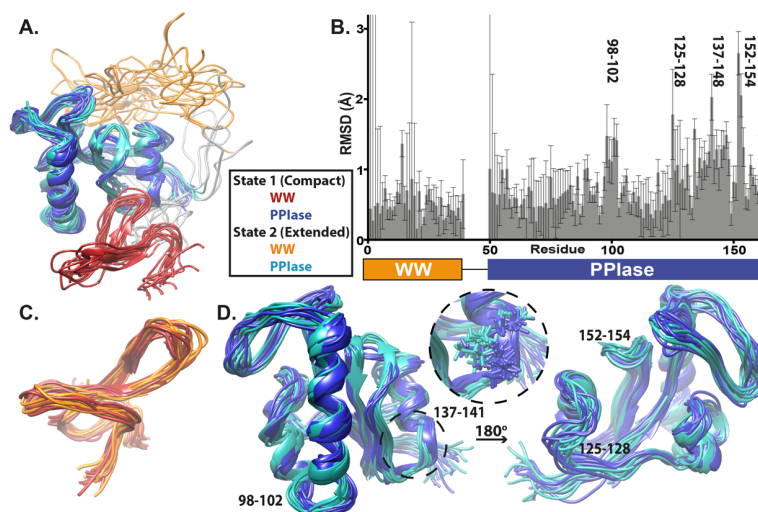


Figure 6. Coupling of inter- and intradomain spatial sampling of Pin1. The two-state ensemble is used for analysis. (A) Extended and compact Pin1 with PPIase overlaid to show the relative position of the WW domain. (B) RMSD between mean compact and extended states of domains versus residue number. Error bars show variance of conformers. (C) Two-state overlay of WW domain. (D) Two-state overlay of PPIase domain. Major RMSD changes from (B) are labeled on the structures. Major differences in interdomain interface are shown in the inset. The 10 conformers with the lowest CYANA TF values are shown.

ensemble describes most NOEs well, some NOEs in regions of interest (i.e., WW binding site, interdomain interface, PPIase catalytic site) fit significantly better to a two-state ensemble, suggesting translational motion in these regions.

When we allow for more than two states, the ensembles feature a WW domain that tends to occupy conformational space comparable to the compact state, while the position near the PPIase catalytic site is not further populated (Figure 5D). This is in agreement with the independently determined compact/extended population ratio of 70:30. To further investigate this proposed population, we performed a pseudo-three-state ensemble calculation, allowing only two distinct conformers, to determine if a 66.6:33.3 population provides a better fit to the experimental data than the 50:50 population of a conventional two-state calculation. We obtain a 3% increase in the TF for 66.6:33.3 compared to the 50:50 calculation, which is within 5 Å² (Figure 5B). Surprisingly, the TF increased significantly (19 Å²; 12% increase) when we ran a pseudo-four-state calculation to test a 75:25 population. This suggests that the minimum is actually between 50:50 and 66.6:33.3. We note that our multistate ensemble calculation is not optimized for population determination in that the first step in the structure calculation only involves the WW domain. The 2:1 weighting should ideally be driven not only by intra-WW domain restraints but also by interdomain restraints. However, this is not possible with the current protocol. Given this limitation, together with the fact that the ideal population distribution is probably more equal than 66.6:33.3, we chose to carry out our structural analysis with the 50:50 distribution. However, we anticipate that the direct use of DEER distance distributions in the calculations would shift the minimum toward 66.6:33.3. Such a calculation would require a way to generate ensembles that reproduce this distribution across the states while reproducing the averaged NMR restraints at the same time.

The interpretation of the sampling of our final two-state ensemble follows Occam's razor approach; that is, we found the simplest representation of the spatial sampling that explains

the data well. This is achieved by the minimal number of states (two in our case) and the minimal difference between these states within the domain (bundling restraints²¹ introduced by Clore and Schwieters and optimized for eNOEs⁷⁴ by us). By construction, the averages of each of the two states over the 10 lowest TF conformers yield the two states sufficient to explain the data, and the standard deviation of a specific state across the 10 conformers is then the uncertainty of the average state.

WW Domain Two-State Ensemble Lacks Structural Correlations. The two-state ensemble produces both compact and extended states (Figure 6A) and we rely on this representation to determine the correlations between intradomain structure and interdomain distance. Figure 6B shows the RMSD between the two states (averaged over the 10 calculations with the lowest TF) on a per-residue basis, with error bars representing the standard deviation within each state. In regions of the standard deviation exceeding the RMSD, structural differences are not likely to be important for changes in orientation. Figure 6B and Figure 6C demonstrate that for the WW domain there are no clear differences between the two. Comparing the WW domain from this two-state structure calculation of the full-length protein to a similar calculation (solved using eNOEs and scalar couplings) of the isolated WW domain (PDB code 1svc),⁵² we observe many changes throughout the domain (Figure 7A). Most notably is the change at the ligand binding site, which may reflect the allosteric influence of the PPIase domain or the rearrangement may be caused by the S18N/W34F mutation introduced in the isolated system. Additionally, there are structural changes at the ID interface, which likely arise due to the presence of the PPIase. We also note that the very high data density for the isolated WW domain allowed for the resolution of two states, revealing structural correlations between the binding and ID interface sites. Because of the lower data density obtained for the full-length Pin1, we are unable to achieve such a resolution.

PPIase Suggests Mode of Interdomain Allostery. Structural changes of interest in the PPIase domain are labeled in Figure 6B,D, including the interdomain interface that involves

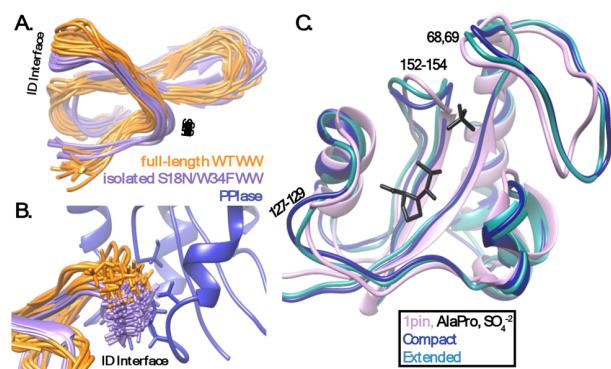


Figure 7. (A) Comparison of ensembles of the WW domain from full-length Pin1 and as an isolated system (PDB code 1svc⁵²). Two-state ensembles are shown with 10 conformers with the lowest CYANA TF for each state, but the individual states are not colored distinctly in order to draw attention to differences in the WW domain dependent on PPIase presence. (B) ID interface residues in compact state. (C) Comparison of catalytic sites of the two-state PPIase mean structures and the PPIase of crystal structure 1pin⁴³ bound to ligand (AlaPro peptide + SO₄²⁻). Regions of particular interest are labeled.

residues 137–148. In the compact state, the side chains of residues A140 and L141 point into the interface, whereas in the extended state they are oriented into the PPIase (Figure 6D inset). The interdomain interface of both domains is mostly composed of hydrophobic residues, suggesting that the hydrophobic effect drives the formation of the compact state. In the extended state, the methyl groups of these residues are oriented toward the core of the protein, away from the interface. Steric clashing also appears to impact the interface, as upon overlaying the WW domains from our full-length structure and from the isolated system (1svc), the hydrophobic residues I28 and T29 of the latter appear to clash with the hydrophobic residues in the PPIase domain (Figure 7B). Hydrophobic-driven interdomain contact is supported by mutagenesis studies as mutations I28A and S138E reduce the hydrophobicity and cause Pin1 to adopt a more extended state.^{49,56,64,75}

By superimposing the secondary structural elements, we also observe major changes in structure at the loop connecting helices 1 and 2 (residues 98–102, Figure 6D), which indicates that this loop may act as a hinge point for these helices. In addition to the interdomain interface, the long helix 1 has previously been identified as an additional pathway for interdomain allostery in addition to the ID interface.⁵¹ Finally, we note a striking difference between the two states in the catalytic site residues 152–154 (Figure 6D). We suspect that this loop repositioning could alter the conformation of the active site, which may lead to activity modulation. Comparing the mean PPIase structures from the two-state ensemble to the crystal structure bound to ligand (1pin) reveals that multiple loops in the catalytic site of the crystal structure are more similar to the extended state than the compact state (Figure 7C). The backbone RMSD between crystal structure and our NMR ensemble for residues 68, 69, 127–129, and 152–154 is higher for the compact state (1.14 Å) than for the extended state (0.97 Å). These regions are critical for catalysis, as residues 68–69 and the loop encompassing residues 127–129 are responsible for phosphate binding and fixing the C-terminal region of the isomerizing peptide, respectively.⁶⁰ Furthermore, residues 152 and 154 are believed to be

important for the hydrogen-bonding network of the catalytic site.^{76,77} The previously mentioned interface mutations that stabilized the extended state, I28A and S138E, also caused Pin1 activity to increase compared to WT.^{49,56,64,75} Conversely, the mutation S138A is predicted to stabilize the compact state and it caused a reduction in activity. While this single-point mutation data support our findings that the extended state has the higher activity in Pin1, another study that involved larger modifications to Pin1 showed no correlation between interdomain contact and activity.⁵⁰ The structural similarity between our extended state and the ligand-bound crystal structure would suggest that the catalytic site is preorganized in the extended state for isomerization even in absence of ligand. Similarly, it was shown that motions necessary for catalysis are an intrinsic property of Pin1, and these motions exist even without ligand present.⁶⁰ Therefore, even in the absence of substrate, Pin1 is primed for ligand binding and catalysis, and the correlated structural changes between the ID interface and the active site may be the driver of interdomain allostery.

CONCLUSION

We have introduced a novel approach that models a multidomain protein in terms of the spatial sampling within each domain, the relative positioning of these domains, and the coupling of the inter- and intradomain sampling by relying on eNOE, scalar couplings, RDC, and PRE data. In addition, we cross-validated the domain positions using distance distributions derived from DEER measurements and we have also utilized these EPR data to improve the accuracy of our model. We note that in general populations obtained from DEER may be altered upon freezing, but this is only the case to a small extent in the current study. In addition, populations determined under *in vitro* conditions may be shifted in physiological cellular environment. To address this issue, the use of molecular crowders in *in vitro* studies has been proposed.⁷⁸

Our multistate ensemble of the two-domain protein Pin1 resolves its compact and extended states, which we independently observed in our experimental DEER distance distributions. In the absence of ligands, our data support that Pin1 slightly favors the compact state with a ~70% population. Furthermore, we were able to determine distinct conformational states within the hydrophobic interdomain interface that stabilize the compact and extended conformers, and we report structural correlations between the interface and the catalytic site. These concerted motions begin to describe how ligand binding in the WW domain alters interdomain contact and thus induces an allosteric change in the PPIase. Having a complete structural ensemble of apo Pin1 in hand, we can further elucidate the allosteric mechanism involved in ligand binding by evaluating the intradomain conformational changes that amend the interdomain dynamics. We expect that our method to generate a multistate ensemble can be optimized and applied to other multidomain proteins of interest.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.1c06289>.

Description of protein expression and purification; description of NMR spectroscopy; description of EPR

spectroscopy; description of data fitting for eNOE extraction, exchange spectroscopy, relaxation and PRE, RDCs, and DEER; description of MMM simulations; description of multistate structure calculation; justification of the use of 5-pulse DEER; table listing isomerization rates of wt and mutant Pin1; three tables listing fit parameters for 4- and 5-pulse DEER; table listing domain-specific R_1 and $R_{1\rho}$ rate constants; table listing statistics of structure calculation; figure showing data used in structure calculation; figure showing CSP of Pin1 mutants; figure showing EXSY spectra of Pin1 mutant isomerizing FFpSPR; figure plotting PRE rates vs residue numbers; figure showing EPR spectra of Pin1 double mutants; two figures showing 4- and 5-pulse DEER analysis; figure showing DEER-, X-ray-, and NMR-derived distance distributions; figure showing NMR spectra with slow-exchange peaks; figure showing optimization of structure calculation; figure showing example NOE-buildup curves (PDF)

AUTHOR INFORMATION

Corresponding Author

Beat Vögeli – Department of Biochemistry and Molecular Genetics, University of Colorado Anschutz Medical Campus, Aurora, Colorado 80045, United States; orcid.org/0000-0003-1176-3137; Email: beat.vogeli@cuanschutz.edu

Authors

Alexandra Born – Department of Biochemistry and Molecular Genetics, University of Colorado Anschutz Medical Campus, Aurora, Colorado 80045, United States

Janne Soetbeer – Laboratory of Physical Chemistry, Department of Chemistry and Applied Biosciences, ETH Zürich, Zürich CH-8093, Switzerland; orcid.org/0000-0003-0008-3494

Frauke Breitgoff – Laboratory of Physical Chemistry, Department of Chemistry and Applied Biosciences, ETH Zürich, Zürich CH-8093, Switzerland

Morkos A. Henen – Department of Biochemistry and Molecular Genetics, University of Colorado Anschutz Medical Campus, Aurora, Colorado 80045, United States; Faculty of Pharmacy, Mansoura University, Mansoura 35516, Egypt

Nikolaos Sgourakis – Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States

Yevhen Polyhach – Laboratory of Physical Chemistry, Department of Chemistry and Applied Biosciences, ETH Zürich, Zürich CH-8093, Switzerland

Parker J. Nichols – Department of Biochemistry and Molecular Genetics, University of Colorado Anschutz Medical Campus, Aurora, Colorado 80045, United States

Dean Strotz – Laboratory of Physical Chemistry, Department of Chemistry and Applied Biosciences, ETH Zürich, Zürich CH-8093, Switzerland

Gunnar Jeschke – Laboratory of Physical Chemistry, Department of Chemistry and Applied Biosciences, ETH Zürich, Zürich CH-8093, Switzerland; orcid.org/0000-0001-6853-8585

Complete contact information is available at:
<https://pubs.acs.org/10.1021/jacs.1c06289>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. David Jones (University of Colorado) for support in NMR spectroscopy and Dr. Ad Bax (NIH, NIDDK) for sharing data measured in his laboratory. This project was supported by NIH Grant R01GM130694-01A1, a start-up package by the University of Colorado to B.V., University of Colorado Cancer Center Support Grant P30 CA046934, and NIH Biomedical Research Support Shared Grant S10 OD025020-01, by a Günthard Foundation Ph.D. scholarship to J.S., and by SNF Grant 200020_169057 to G.J.

REFERENCES

- (1) Zhou, X.; Hu, J.; Zhang, C.; Zhang, G.; Zhang, Y. Assembling Multidomain Protein Structures through Analogous Global Structural Alignments. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (32), 15930–15938.
- (2) Bhaskara, R. M.; Srinivasan, N. Stability of Domain Structures in Multi-Domain Proteins. *Sci. Rep.* **2011**, *1*, 40.
- (3) Wriggers, W.; Chakravarty, S.; Jennings, P. A. Control of Protein Functional Dynamics by Peptide Linkers. *Biopolymers* **2005**, *80* (6), 736–746.
- (4) Shapero, Y. E. NMR Spectroscopy on Domain Dynamics in Biomacromolecules. *Prog. Biophys. Mol. Biol.* **2013**, *112* (3), 58–117.
- (5) Richardson, J. S. The Anatomy and Taxonomy of Protein Structure. *Adv. Protein Chem.* **1981**, *34* (C), 167–339.
- (6) Bertini, I.; Gupta, Y. K.; Luchinat, C.; Parigi, G.; Peana, M.; Sgheri, L.; Yuan, J. Paramagnetism-Based NMR Restraints Provide Maximum Allowed Probabilities for the Different Conformations of Partially Independent Protein Domains. *J. Am. Chem. Soc.* **2007**, *129* (42), 12786–12794.
- (7) 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.
- (8) Jacobs, D. M.; Saxena, K.; Vogtherr, M.; Bernadó, P.; Pons, M.; Fiebig, K. M. Peptide Binding Induces Large Scale Changes in Inter-Domain Mobility in Human Pin1. *J. Biol. Chem.* **2003**, *278* (28), 26174–26182.
- (9) Bouchard, J. J.; Xia, J.; Case, D. A.; Peng, J. W. Enhanced Sampling of Interdomain Motion Using Map-Restrained Langevin Dynamics and NMR: Application to Pin1. *J. Mol. Biol.* **2018**, *430* (14), 2164–2180.
- (10) Baber, J. L.; Szabo, A.; Tjandra, N. Analysis of Slow Interdomain Motion of Macromolecules Using NMR Relaxation Data. *J. Am. Chem. Soc.* **2001**, *123* (17), 3953–3959.
- (11) Vishwanath, S.; de Brevern, A. G.; Srinivasan, N. Same but Not Alike: Structure, Flexibility and Energetics of Domains in Multi-Domain Proteins Are Influenced by the Presence of Other Domains. *PLoS Comput. Biol.* **2018**, *14* (2), e1006008.
- (12) Xu, D.; Jaroszewski, L.; Li, Z.; Godzik, A. AIDA: Ab Initio Domain Assembly for Automated Multi-Domain Protein Structure Prediction and Domain-Domain Interaction Prediction. *Bioinformatics* **2015**, *31* (13), 2098–2105.
- (13) Zhou, X.; Hu, J.; Zhang, C.; Zhang, G.; Zhang, Y. Assembling Multidomain Protein Structures through Analogous Global Structural Alignments. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (32), 15930–15938.
- (14) Takeuchi, K.; Baskaran, K.; Arthanari, H. Structure Determination Using Solution NMR: Is It Worth the Effort? *J. Magn. Reson.* **2019**, *306*, 195–201.
- (15) Fushman, D.; Varadan, R.; Assfalg, M.; Walker, O. Determining Domain Orientation in Macromolecules by Using Spin-Relaxation

and Residual Dipolar Coupling Measurements. *Prog. Nucl. Magn. Reson. Spectrosc.* **2004**, *44* (3–4), 189–214.

(16) Ryabov, Y.; Fushman, D. Interdomain Mobility in Di-Ubiquitin Revealed by NMR. *Proteins: Struct., Funct., Genet.* **2006**, *63* (4), 787–796.

(17) Ryabov, Y. E.; Fushman, D. A Model of Interdomain Mobility in a Multidomain Protein. *J. Am. Chem. Soc.* **2007**, *129* (11), 3315–3327.

(18) Bewley, C. A.; Clore, G. M. Determination of the Relative Orientation of the Two Halves of the Domain-Swapped Dimer of Cyanovirin-N in Solution Using Dipolar Couplings and Rigid Body Minimization. *J. Am. Chem. Soc.* **2000**, *122* (25), 6009–6016.

(19) Qi, Y.; Martin, J. W.; Barb, A. W.; Thélot, F.; Yan, A. K.; Donald, B. R.; Oas, T. G. Continuous Interdomain Orientation Distributions Reveal Components of Binding Thermodynamics. *J. Mol. Biol.* **2018**, *430* (18), 3412–3426.

(20) Wu, X.; Subramaniam, S.; Case, D. A.; Wu, K. W.; Brooks, B. R. Targeted Conformational Search with Map-Restrained Self-Guided Langevin Dynamics: Application to Flexible Fitting into Electron Microscopic Density Maps. *J. Struct. Biol.* **2013**, *183* (3), 429–440.

(21) Clore, G. M.; Schwieters, C. D. Amplitudes of Protein Backbone Dynamics and Correlated Motions in a Small α/β Protein: Correspondence of Dipolar Coupling and Heteronuclear Relaxation Measurements. *Biochemistry* **2004**, *43* (33), 10678–10691.

(22) Clore, G. M.; Schwieters, C. D. How Much Backbone Motion in Ubiquitin Is Required to Account for Dipolar Coupling Data Measured in Multiple Alignment Media as Assessed by Independent Cross-Validation? *J. Am. Chem. Soc.* **2004**, *126* (9), 2923–2938.

(23) Lindorff-Larsen, K.; Best, R. B.; Depristo, M. A.; Dobson, C. M.; Vendruscolo, M. Simultaneous Determination of Protein Structure and Dynamics. *Nature* **2005**, *433* (13), 128–132.

(24) Vögeli, B. The Nuclear Overhauser Effect from a Quantitative Perspective. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *78*, 1–46.

(25) Nichols, P.; Born, A.; Henen, M.; Strotz, D.; Orts, J.; Olsson, S.; Güntert, P.; Chi, C.; Vögeli, B. The Exact Nuclear Overhauser Enhancement: Recent Advances. *Molecules* **2017**, *22* (7), 1176.

(26) Vögeli, B.; Orts, J.; Strotz, D.; Chi, C.; Mingos, M.; Wälti, M. A.; Güntert, P.; Riek, R. Towards a True Protein Movie: A Perspective on the Potential Impact of the Ensemble-Based Structure Determination Using Exact NOEs Dedicated to Prof. Dr. Christian Griesinger. *J. Magn. Reson.* **2014**, *241* (1), 53–59.

(27) Nichols, P. J.; Born, A.; Henen, M. A.; Strotz, D.; Celestine, C. N.; Güntert, P.; Vögeli, B. Extending the Applicability of Exact Nuclear Overhauser Enhancements to Large Proteins and RNA. *ChemBioChem* **2018**, *19* (16), 1695–1701.

(28) Vögeli, B.; Segawa, T. F.; Leitz, D.; Sobol, A.; Choutko, A.; Trzesniak, D.; Van Gunsteren, W.; Riek, R. Exact Distances and Internal Dynamics of Perdeuterated Ubiquitin from NOE Buildups. *J. Am. Chem. Soc.* **2009**, *131* (47), 17215–17225.

(29) Shi, L.; Traaseth, N. J.; Verardi, R.; Gustavsson, M.; Gao, J.; Veglia, G. Paramagnetic-Based NMR Restraints Lift Residual Dipolar Coupling Degeneracy in Multidomain Detergent-Solubilized Membrane Proteins. *J. Am. Chem. Soc.* **2011**, *133* (7), 2232–2241.

(30) Clore, M.; Tang, C.; Iwahara, J. Elucidating Transient Macromolecular Interactions Using Paramagnetic Relaxation Enhancement. *Curr. Opin. Struct. Biol.* **2007**, *17*, 603–616.

(31) Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiess, H. W. Dead-Time Free Measurement of Dipole-Dipole Interactions between Electron Spins. *2000. J. Magn. Reson.* **2011**, *213* (2), 316–325.

(32) Borbat, P. P.; Georgieva, E. R.; Freed, J. H. Improved Sensitivity for Long-Distance Measurements in Biomolecules: Five-Pulse Double Electron-Electron Resonance. *J. Phys. Chem. Lett.* **2013**, *4* (1), 170–175.

(33) Breitgoff, F. D.; Soetbeer, J.; Doll, A.; Jeschke, G.; Polyhach, Y. O. Artefact Suppression in 5-Pulse Double Electron Electron Resonance for Distance Distribution Measurements. *Phys. Chem. Chem. Phys.* **2017**, *19* (24), 15766–15779.

(34) Schmidt, T.; Schwieters, C. D.; Clore, G. M. Spatial Domain Organization in the HIV-1 Reverse Transcriptase P66 Homodimer

Precursor Probed by Double Electron-Electron Resonance EPR. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 17809–17816.

(35) Alonso-García, N.; García-Rubio, I.; Manso, J. A.; Buey, R. M.; Urien, H.; Sonnenberg, A.; Jeschke, G.; De Pereda, J. M. Combination of X-Ray Crystallography, SAXS and DEER to Obtain the Structure of the FnIII-3,4 Domains of Integrin A β 4. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015**, *71*, 969–985.

(36) Ward, R.; Zoltner, M.; Beer, L.; El Mkami, H.; Henderson, I. R.; Palmer, T.; Norman, D. G. The Orientation of a Tandem POTRA Domain Pair, of the Beta-Barrel Assembly Protein BamA, Determined by PELDOR Spectroscopy. *Structure* **2009**, *17* (9), 1187–1194.

(37) Gigli, L.; Andralojć, W.; Dalaloyan, A.; Parigi, G.; Ravera, E.; Goldfarb, D.; Luchinat, C. Assessing Protein Conformational Landscapes: Integration of DEER Data in Maximum Occurrence Analysis. *Phys. Chem. Chem. Phys.* **2018**, *20* (43), 27429–27438.

(38) Krug, U.; Alexander, N. S.; Stein, R. A.; Keim, A.; McHaourab, H. S.; Sträter, N.; Meiler, J. Characterization of the Domain Orientations of E. Coli 5'-Nucleotidase by Fitting an Ensemble of Conformers to DEER Distance Distributions. *Structure* **2016**, *24* (1), 43–56.

(39) Yaffe, M. B.; Schutkowski, M.; Shen, M.; Zhou, X. Z.; Stukenberg, P. T.; Rahfeld, J. U.; Xu, J.; Kuang, J.; Kirschner, M. W.; Fischer, G.; et al. Sequence-Specific and Phosphorylation Dependent Proline Isomerization: A Potential Mitotic Regulatory Mechanism. *Science* **1997**, *278* (5345), 1957–1960.

(40) Shen, M.; Stukenberg, P. T.; Kirschner, M. W.; Lu, K. P. The Essential Mitotic Peptidyl-Prolyl Isomerase Pin1 Binds and Regulates Mitosis-Specific Phosphoproteins. *Genes Dev.* **1998**, *12* (5), 706–720.

(41) Lu, K. P.; Hanes, S. D.; Hunter, T. A Human Peptidyl-Prolyl Isomerase Essential for Regulation of Mitosis. *Nature* **1996**, *380* (6574), 544–547.

(42) Namanja, A. T.; Wang, X. J.; Xu, B.; Mercedes-Camacho, A. Y.; Wilson, K. A.; Etzkorn, F. A.; Peng, J. W. Stereospecific Gating of Functional Motions in Pin1. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (30), 12289–12294.

(43) Ranganathan, R.; Lu, K. P.; Hunter, T.; Noel, J. P. Structural and Functional Analysis of the Mitotic Rotamase Pin1 Suggest Substrate Recognition Is Phosphorylation Dependent. *Cell* **1997**, *89*, 875–886.

(44) Lu, K. P.; Hanes, S. D.; Hunter, T. A Human Peptidyl-Prolyl Isomerase Essential for Regulation of Mitosis. *Nature* **1996**, *380* (6574), 544–547.

(45) Bayer, E.; Goettsch, S.; Mueller, J. W.; Griewel, B.; Guiberman, E.; Mayr, L. M.; Bayer, P. Structural Analysis of the Mitotic Regulator HPin1 in Solution: Insights into Domain Architecture and Substrate Binding. *J. Biol. Chem.* **2003**, *278* (28), 26183–26193.

(46) Bernadó, P.; Fernandes, M. X.; Jacobs, D. M.; Fiebig, K.; García de la Torre, J.; Pons, M. Interpretation of NMR Relaxation Properties of Pin1, a Two-Domain Protein, Based on Brownian Dynamic Simulations. *J. Biomol. NMR* **2004**, *29* (1), 21–35.

(47) Namanja, A. T.; Peng, T.; Zintsmaster, J. S.; Elson, A. C.; Shakour, M. G.; Peng, J. W. Substrate Recognition Reduces Side-Chain Flexibility for Conserved Hydrophobic Residues in Human Pin1. *Structure* **2007**, *15* (3), 313–327.

(48) Behrsin, C. D.; Bailey, M. L.; Bateman, K. S.; Hamilton, K. S.; Wahl, L. M.; Brandl, C. J.; Shilton, B. H.; Litchfield, D. W. Functionally Important Residues in the Peptidyl-Prolyl Isomerase Pin1 Revealed by Unigenic Evolution. *J. Mol. Biol.* **2007**, *365* (4), 1143–1162.

(49) Peng, J. W. Investigating Dynamic Interdomain Allostery in Pin1. *Biophys. Rev.* **2015**, *7* (2), 239–249.

(50) Zhu, W.; Li, Y.; Liu, M.; Zhu, J.; Yang, Y. Uncorrelated Effect of Interdomain Contact on Pin1 Isomerase Activity Reveals Positive Catalytic Cooperativity. *J. Phys. Chem. Lett.* **2019**, *10* (6), 1272–1278.

(51) Guo, J.; Pang, X.; Zhou, H. X. Two Pathways Mediate Interdomain Allosteric Regulation in Pin1. *Structure* **2015**, *23* (1), 237–247.

(52) Strotz, D.; Orts, J.; Kadavath, H.; Friedmann, M.; Ghosh, D.; Olsson, S.; Chi, C. N.; Pokharna, A.; Güntert, P.; Vögeli, B.; Riek, R.

Protein Allostery at Atomic Resolution. *Angew. Chem., Int. Ed.* **2020**, *59*, 22132–22139.

(53) Zhang, M.; Case, D. A.; Peng, J. W. Propagated Perturbations from a Peripheral Mutation Show Interactions Supporting WW Domain Thermostability. *Structure* **2018**, *26* (11), 1474–1485.

(54) Matena, A.; Sinnen, C.; Van Den Boom, J.; Wilms, C.; Dybowski, J. N.; Maltaner, R.; Mueller, J. W.; Link, N. M.; Hoffmann, D.; Bayer, P. Transient Domain Interactions Enhance the Affinity of the Mitotic Regulator Pin1 toward Phosphorylated Peptide Ligands. *Structure* **2013**, *21* (10), 1769–1777.

(55) Zhang, M.; Frederick, T. E.; VanPelt, J.; Case, D. A.; Peng, J. W. Coupled Intra- And Interdomain Dynamics Support Domain Cross-Talk in Pin1. *J. Biol. Chem.* **2020**, *295* (49), 16585–16603.

(56) Wilson, K. A.; Bouchard, J. J.; Peng, J. W. Interdomain Interactions Support Interdomain Communication in Human Pin1. *Biochemistry* **2013**, *52* (40), 6968–6981.

(57) Chi, C. N.; Strotz, D.; Riek, R.; Vögeli, B. Extending the ENOE Data Set of Large Proteins by Evaluation of NOEs with Unresolved Diagonals. *J. Biomol. NMR* **2015**, *62* (1), 63–69.

(58) Pintacuda, G.; Park, A. Y.; Keniry, M. A.; Dixon, N. E.; Otting, G. Lanthanide Labeling Offers Fast NMR Approach to 3D Structure Determinations of Protein-Protein Complexes. *J. Am. Chem. Soc.* **2006**, *128* (11), 3696–3702.

(59) John, M.; Pintacuda, G.; Park, A. Y.; Dixon, N. E.; Otting, G. Structure Determination of Protein-Ligand Complexes by Transferred Paramagnetic Shifts. *J. Am. Chem. Soc.* **2006**, *128* (39), 12910–12916.

(60) Labeikovsky, W.; Eisenmesser, E. Z.; Bosco, D. A.; Kern, D. Structure and Dynamics of Pin1 During Catalysis by NMR. *J. Mol. Biol.* **2007**, *367* (5), 1370–1381.

(61) Barman, A.; Hamelberg, D. Cysteine-Mediated Dynamic Hydrogen-Bonding Network in the Active Site of Pin1. *Biochemistry* **2014**, *53* (23), 3839–3850.

(62) Chen, C.-H.; Li, W.; Sultana, R.; You, M.-H.; Kondo, A.; Shahpasand, K.; Kim, B. M.; Luo, M.-L.; Nechama, M.; Lin, Y.-M.; et al. Pin1 Cysteine-113 Oxidation Inhibits Its Catalytic Activity and Cellular Function in Alzheimer's Disease. *Neurobiol. Dis.* **2015**, *76*, 13–23.

(63) Xu, N.; Tochio, N.; Wang, J.; Tamari, Y.; Uewaki, J. I.; Utsunomiya-Tate, N.; Igarashi, K.; Shiraki, T.; Kobayashi, N.; Tate, S. I. The C113D Mutation in Human Pin1 Causes Allosteric Structural Changes in the Phosphate Binding Pocket of the PPlase Domain through the Tug of War in the Dual-Histidine Motif. *Biochemistry* **2014**, *53* (34), 5568–5578.

(64) Born, A.; Henen, M. A.; Vögeli, B. Activity and Affinity of Pin1 Variants. *Molecules* **2020**, *25* (1), 36.

(65) Battiste, J. L.; Wagner, G. Utilization of Site-Directed Spin Labeling and High-Resolution Heteronuclear Nuclear Magnetic Resonance for Global Fold Determination of Large Proteins with Limited Nuclear Overhauser Effect Data. *Biochemistry* **2000**, *39*, 5355–5365.

(66) Fawzi, N. L.; Fleissner, M. R.; Anthis, N. J.; Kálai, T.; Hideg, K.; Hubbell, W. L.; Clore, G. M. A Rigid Disulfide-Linked Nitroxide Side Chain Simplifies the Quantitative Analysis of PRE Data. *J. Biomol. NMR* **2011**, *51* (1–2), 105–114.

(67) Rumpel, S.; Becker, S.; Zweckstetter, M. High-Resolution Structure Determination of the CylR2 Homodimer Using Paramagnetic Relaxation Enhancement and Structure-Based Prediction of Molecular Alignment. *J. Biomol. NMR* **2008**, *40* (1), 1–13.

(68) Clore, G. M. Practical Aspects of Paramagnetic Relaxation Enhancement in Biological Macromolecules. *Methods Enzymol.* **2015**, *564*, 485–497.

(69) Anthis, N. J.; Doucleff, M.; Clore, G. M. Transient, Sparsely Populated Compact States of Apo and Calcium-Loaded Calmodulin Probed by Paramagnetic Relaxation Enhancement: Interplay of Conformational Selection and Induced Fit. *J. Am. Chem. Soc.* **2011**, *133* (46), 18966–18974.

(70) Hays, J. M.; Boland, E.; Kasson, P. M. Inference of Joint Conformational Distributions from Separately Acquired Experimental Measurements. *J. Phys. Chem. Lett.* **2021**, *12* (6), 1606–1611.

(71) Born, A.; Nichols, P. J.; Henen, M. A.; Chi, C. N.; Strotz, D.; Bayer, P.; Tate, S.-I.; Peng, J. W.; Vögeli, B. Backbone and Side-Chain Chemical Shift Assignments of Full-Length, Apo, Human Pin1, a Phosphoprotein Regulator with Interdomain Allostery. *Biomol. NMR Assignments* **2019**, *13*, 85–89.

(72) Jeschke, G. MMM: A Toolbox for Integrative Structure Modeling. *Protein Sci.* **2018**, *27* (1), 76–85.

(73) Vögeli, B.; Olsson, S.; Riek, R.; Güntert, P. Complementarity and Congruence between Exact NOEs and Traditional NMR Probes for Spatial Decoding of Protein Dynamics. *J. Struct. Biol.* **2015**, *191* (3), 306–317.

(74) Vögeli, B.; Güntert, P.; Riek, R. Multiple-State Ensemble Structure Determination from ENOE Spectroscopy. *Mol. Phys.* **2013**, *111* (3), 437–454.

(75) Wang, J.; Kawasaki, R.; Uewaki, J.-i.; Rashid, A. U. R.; Tochio, N.; Tate, S. Dynamic Allostery Modulates Catalytic Activity by Modifying the Hydrogen Bonding Network in the Catalytic Site of Human Pin1. *Molecules* **2017**, *22* (6), 992.

(76) Matena, A.; Rehic, E.; Hönig, D.; Kamba, B.; Bayer, P. Structure and Function of the Human Parvulins Pin1 and Par14/17. *Biol. Chem.* **2018**, *399* (2), 101–125.

(77) Wang, J.; Tochio, N.; Kawasaki, R.; Tamari, Y.; Xu, N.; Uewaki, J. I.; Utsunomiya-Tate, N.; Tate, S. I. Allosteric Breakage of the Hydrogen Bond within the Dual-Histidine Motif in the Active Site of Human Pin1 PPlase. *Biochemistry* **2015**, *54* (33), 5242–5253.

(78) Collauto, A.; von Bülow, S.; Gophane, D. B.; Saha, S.; Stelzl, L. S.; Hummer, G.; Sigurdsson, S. T.; Prisner, T. F. Compaction of RNA Duplexes in the Cell*. *Angew. Chem., Int. Ed.* **2020**, *59* (51), 23025–23029.

Supporting Information

Reconstruction of coupled intra- and interdomain protein motion from nuclear and electron magnetic resonance

Alexandra Born¹, Janne Soetbeer², Frauke Breitgoff², Morkos A. Henen^{1,3}, Nikolaos Sgourakis⁴, Yevhen Polyhach², Parker Nichols¹, Dean Strotz², Gunnar Jeschke², Beat Vögeli^{1*}

¹ University of Colorado Anschutz Medical Campus, Department of Biochemistry and Molecular Genetics, 12801 East 17th Avenue, Aurora, CO 80045, USA

² Laboratory of Physical Chemistry, Department of Chemistry and Applied Biosciences, ETH Zürich, Vladimir-Prelog-Weg 2, ETH-Hönggerberg, Zürich, CH-8093, Switzerland

³ Faculty of Pharmacy, Mansoura University, Mansoura, 35516, Egypt

⁴ Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA 19104, USA

*Correspondence should be addressed to:

Beat Vögeli, Department of Biochemistry and Molecular Genetics, University of Colorado Denver, Research Center 1 South, Room 9103, 12801 East 17th Avenue, Aurora, CO 80045, USA, Tel: (+1)-303-724-1627, beat.vogeli@cuanschutz.edu, ORCID ID: 0000-0003-1176-3137

Materials and Methods

Protein expression and purification

For measurement of eNOEs, J couplings, and RDCs, wild-type Pin1 was cloned in the pET-28a(+) vector with a 6xHis tag and kanamycin resistance. The plasmid was transformed and expressed into *Escherichia coli* strain BL21 (DE3) as previously reported¹⁻³. In brief, overnight cultures were resuspended in M9 minimal media with kanamycin (50 µg/mL), and ¹⁵N-ammonium chloride (1 g/L) and ¹³C-glucose (2 g/L) for uniform labeling. Cultures were grown at 37°C shaking until optical density A₆₀₀ reached 0.8, then were induced with 1 mM isopropyl-1-thio-d-galactopyranoside. Induced cells were then grown shaking overnight at 25°C. Cells were harvested by centrifugation at 4°C for 20 min at 4000 g.

For protein purification, cells were resuspended in 50 mM potassium phosphate with 1 mM dithiothreitol, 0.3 mM phenylmethylsulfonyl fluoride, and 25 mM imidazole at pH 7.5. Cells were lysed by sonication (4 × 30 s) at 4°C and then centrifuged for 45 min at 52,000 g at 4°C. Cell lysate was filtered, and loaded onto a pre-equilibrated nickel nitrilotriacetic acid column (GE Healthcare) on an AKTA Pure at 4°C. The His-tagged protein was eluted with the same 50 mM phosphate buffer but with 250 mM imidazole. Pin1 was concentrated using a 3000 NMWL cutoff, and then purified using a Superdex 75 10/300 GL size exclusion column in 50 mM sodium phosphate and 150 mM sodium chloride buffer at pH 6.5.

For PRE and DEER measurements, it was necessary to mutate Pin1 to contain either one or two cysteines, respectively, for nitroxide spin labeling. Endogenous cysteines were not ideal to spin label, and therefore were mutated out (C57A and C113D), using replacements that

minimize changes in structure and activity based on previous work^{4,5}. We then conscientiously mutated in cysteines to various sites around Pin1 that should not lead to changes in structure and activity while being able to probe interdomain distances. We obtained four mutants for PRE studies with M15C in the WW domain, and N90C, S98C, and Q131C located in the PPlase domain. For DEER studies, we used all combinations of the PRE mutants to obtain six double-mutant constructs. These mutant constructs were cloned into the same vector for expression as WT Pin1. Expression and purification for the PRE and DEER constructs was done identically as WT Pin1, except the PRE constructs were grown in ¹⁵N-ammonium chloride and natural isotope abundance glucose and the DEER constructs were grown in completely unlabeled media. It should be noted that the DEER constructs had also been grown in ¹⁵N-ammonium chloride and natural isotope abundance glucose to measure HSQC spectra for quality control.

To spin label each ¹⁵N Pin1 mutant sample with MTSL, PD-10 desalting columns were used to change samples into reaction buffer (100 mM NaPO₄, 1 mM EDTA at pH 8) and to remove all DTT (which is a reducing agent and would remove the MTSL label). The protein concentration was measured, and a threefold excess of MTSL (in 50 mM stock dissolved with DMSO) was added to each sample. Samples were incubated at 37°C for 3 hours, and then buffer exchanged back into NMR buffer without DTT (20 mM NaPO₄, 50 mM NaCl, 0.03% NaN₃ at pH 6.5). To remove the MTSL for quenched measurements in PRE, 5 mM DTT was added to the spin-labeled samples (~300 μM in 300 μL) and buffer exchanged using 10 mL of NMR buffer containing 5 mM DTT to remove free MTSL from solution.

NMR Spectroscopy

Unless otherwise noted, experiments were measured at 298 K on a triple-resonance Varian 900 MHz spectrometer equipped with a cryoprobe.

For solving the structure of apo Pin1, the sample contained 2 mM ¹⁵N,¹³C-labeled Pin1 in 20 mM sodium phosphate, 50 mM sodium chloride, 5 mM dithiothreitol (DTT), 0.03% sodium azide, and 3% D₂O (hereby referred to as NMR buffer). As previously published², the backbone was assigned using ¹⁵N-HSQC, HNCACB, HNCA, and CBCA(CO)NH experiments. Side-chains were assigned using ¹³C-resolved aliphatic and aromatic CT-HSQC, HCCH-TOCSY, and HBHA(CO)NH spectra². The chemical shifts have been deposited in the BRMB under accession code 27579.

NOE buildups were measured using a series of 3D simultaneously [¹⁵N,¹³C]-resolved [¹H-¹H-¹³C,¹⁵N] NOESY experiments as previously described^{6,7}. A spectrum with $\tau_{\text{mix}} = 100$ ms was recorded for the resonance assignment, but then diagonal peak decay and cross-peak buildups were analyzed from spectra measured with incremented mixing times $\tau_{\text{mix}} = 24, 32, 40, 48,$ and 56 ms. The spectra were recorded with a linear sampling scheme with $160(^1\text{H}, t_1) \times 50(^{13}\text{C}/^{15}\text{N}, t_2) \times 1024(^1\text{H}, t_3)$ complex points, maximal evolution times of $t_{1\text{max},1\text{H}} = 11.4$ ms, $t_{2\text{max},15\text{N}/13\text{C}} = 16.1$ ms, and $t_{3\text{max},1\text{H}} = 72.9$ ms, spectral widths $\text{SW}_{1,1\text{H}} = 15.6$ ppm, $\text{SW}_{2,15\text{N}} = 34$ ppm, $\text{SW}_{2,13\text{C}} = 30$ ppm, $\text{SW}_{3,1\text{H}} = 15.6$ ppm, an interscan delay of 1.2 s, and 4 scans per increment.

³J_{HN-H α} scalar couplings were measured to restrain backbone angle ϕ with the well documented 3D HNHA COSY-like experiment⁸. To determine the stereospecific assignment of methylene protons based on the χ_1 dihedral angle via the Karplus relationship, we proposed an efficient strategy using a modified 3D HACAHB-COSY pulse sequence³ to eliminate the high power needed for continuous decoupling that may lead to probe damage (original pulse sequence⁹). To increase the maximum evolution time in the indirect dimension without increasing the overall

measurement time, we employed a 50% non-uniform sampling (NUS) scheme. We previously showed that the NUS scheme did not result in spectral artifacts or inaccurate peak intensities in this quantitative experiment³. Aromatic torsion angles χ_1 were restrained by measuring $^3J_{N-C\gamma}$ using the ^{15}N - $^{13}C\gamma$ spin echo difference 1H - ^{15}N HSQC experiment¹⁰.

^{15}N R_1 and $R_{1\rho}$ relaxation rate constants were measured on the same 2 mM ^{15}N , ^{13}C Pin1 sample as used for the NOE-buildup measurement to determine domain-specific tumbling times using standard pulse sequences¹¹. The sampling time points were 0, 40, 80, 130, 190, 280, 390, 590, 680, 790, and 990 ms and 10, 20, 30, 40, 60, 80, 100, and 120 ms for R_1 and $R_{1\rho}$, respectively. A 1,877 Hz strength ^{15}N spin-lock field was applied during the $R_{1\rho}$ relaxation time. R_2 was calculated from R_1 and $R_{1\rho}$, and the global correlation time was obtained under assumption of isotropic tumbling.

RDCs were collected in 5% C12E5 PEG/hexanol¹². The final 2H quadrupolar splitting was 27.4 Hz. For all RDC measurements, ^{15}N , ^{13}C , 2H -labeled Pin1 was used to suppress long-range proton-proton couplings and increase the accuracy of measured RDCs. Amide protons were back-exchanged during protein purification. $^1D_{Ni,HNi}$ RDCs were measured using an ARTSY experiment at an 800 MHz Bruker spectrometer¹³. $^1D_{C'i,C\alpha i}$ couplings were measured at a 500 MHz Bruker spectrometer using a decoupled TROSY-HNCO experiment (without $^{13}C\alpha$ decoupling during $^{13}C'$ evolution). $D_{C'i,Ni+1}$ and $D_{C'i,HNi+1}$ were obtained from a single TROSY-HNCO experiment by removing $^{13}C'$ decoupling during the ^{15}N chemical shift evolution block. The errors in RDC measurements were estimated based on the peak signal-to-noise ratios and linewidths in the spectra to be 1.0 Hz on average for $^1D_{Ni,HNi}$ (individually applied ranging from 0.3 to 3.3 Hz), and 0.5 Hz for $^1D_{C'i,C\alpha i}$, $D_{C'i,Ni+1}$ and $D_{C'i,HNi+1}$ (uniformly applied).

To test catalytic activity of the mutants needed for PRE and DEER studies, the *cis-trans* isomerization rate of ligand FFpSPR was measured using exchange spectroscopy (EXSY)¹⁴. Ligand FFpSPR was synthesized from the Peptide and Protein Chemistry core at University of Colorado Anschutz. EXSY samples contained 2 mM peptide with 50 μM WT or mutant Pin1 in NMR buffer. A series of 2D 1H - 1H NOESYs were recorded on the ligand with mixing times τ_{mix} = 15, 20, 25, 35, 50, 75, 100, 150, and 250 ms. The direct dimension was measured with 64 scans, while the indirect dimension was sampled with a 30% NUS scheme of 256 points. The FFpSPR ligand was assigned using 2D 1H - 1H TOCSY and ROESY experiments on a sample with 4 mM FFpSPR in NMR buffer at pH 6.5.

R_2 PRE rates were measured on the amide protons on the ^{15}N -labeled MTSL-conjugated (paramagnetic) and quenched (diamagnetic) single cysteine mutants using the pulse sequence described here¹⁵. The measured samples contained between 250-300 μM in NMR buffer, but the MTSL-conjugated samples did not contain any DTT. These experiments were recorded on a triple-resonance 600 MHz Bruker equipped with a cryoprobe. An 180° shaped pulse was applied on the transverse 1H during the incremented relaxation delay with values T = 1, 5, 9, 13, 17, and 21 ms. The experiments were carried out with the same parameters as previously published¹⁵, except with a 1.8 ms REBURP1000 shaped pulse with field strength of 3.48 kHz at 8 ppm.

All data were processed with NMRPipe¹⁶ and analyzed using NMRDraw, CcpNmr¹⁷ and SPARKY¹⁸.

EPR Spectroscopy

Continuous-wave EPR spectroscopy

Continuous-wave (CW) EPR spectra were recorded with a Bruker Elexsys E500 spectrometer using the Bruker super-high-Q resonator ER4122SHQ at an X-band frequency of 9.5 GHz. Glass capillaries of 1.5 mm o.d. were filled with 10 μ L of ~ 60 μ M MTSL-labeled Pin1 in buffer. All CW spectra were acquired at room temperature with common settings of 0.6362 mW microwave power (25 dB attenuation), 100 kHz field modulation and 0.1 mT amplitude modulation.

Pulsed EPR spectroscopy: Double Electron-Electron Resonance

Pulsed Double Electron-Electron Resonance (DEER) Q-band EPR measurements were recorded at 50 K with a Bruker Elexsys E580 spectrometer equipped with a home-built resonator¹⁹ and an incoherent arbitrary waveform generator pulse channel. 40 μ L of 30-80 μ M MTSL-labeled Pin1 in buffer:glycerol- d_8 1:1 v:v were filled in 3 mm o.d. quartz capillaries, flash frozen and stored in liquid nitrogen between measurements. Echo-detected field-sweeps were recorded using $\pi/2-\tau-\pi-\tau$ with pulse lengths $t_{\pi/2} = t_{\pi}/2 = 12$ ns and an interpulse delay of $\tau = 400$ ns. DEER data were acquired either with the 4-pulse DEER (4pDEER) sequence $\pi/2_{\text{obs}}-\tau_1-\pi_{\text{obs}}-(\tau_1+t)-\pi_{\text{pump}}-(\tau_2-t)-\pi_{\text{obs}}-\tau_2$ ²⁰ or the 5-pulse DEER (5pDEER) sequence $\pi/2_{\text{obs}}-(\tau/2-t_0)-\pi_{\text{pump}}-t_0-\pi_{\text{obs}}-t-\pi_{\text{pump}}-(\tau-t+\delta)-\pi_{\text{obs}}-(\tau/2+\delta)$ ²¹ featuring a short delay $\delta = 120$ ns to separate the refocused from the stimulated echo²². 4pDEER measurements were performed with monochromatic, rectangular pulses of length $t_{\pi,\text{pump}} = 12$ ns, applied at the maximum of the nitroxide Q-band spectrum and observer pulses with $t_{\pi/2,\text{obs}} = t_{\pi,\text{obs}} = 16$ ns, offset 100 MHz from the pump frequency. 5pDEER measurements were recorded with HS{1,6} pump pulses of 150 MHz width and monochromatic, rectangular observer pulses with $t_{\pi/2,\text{obs}} = t_{\pi,\text{obs}} = 32$ ns, placed 70 MHz away from the pump position, corresponding to an optimized pulse setup that suppresses overlap artefacts²². In each case, nuclear modulations were averaged using an eight-step phase cycle with 16 ns steps.

Due to non-ideal excitation, an artefact is present in 5pDEER data²¹. To avoid interference of this artefact with the extraction of distance distributions, 5pDEER data were acquired with two different t_0 values and artefact correction was carried out as described in²³. In brief: the difference in t_0 leads to a different relative timing of the artefact with respect to the main signal. This difference can be used for artefact identification and correction. The time shift Δt_0 was 160 ns, 112 ns and 80 ns for 15-131, 15-98 and 15-90, respectively.

Data Fitting

eNOE distance extraction

As described previously⁷, the NOESY spectrum from the 100 ms experiment was assigned in CcpNmr¹⁷, and then the peak lists were exported into NMRPipe format. Using the NlinLS autofit script¹⁶, diagonal- and cross-peak intensities were extracted for all mixing times. The auto-relaxation rate constants (ρ) and initial magnetization (M_0) values were fitted to the diagonal peak intensities, and used with the cross-peak intensities to calculate cross-relaxation rate constants (σ) using the full-relaxation matrix (FRM) approach in the eNORA2 package that has been incorporated into CYANA²⁴⁻²⁷. A conventional, preliminary NMR structure was calculated from assigned peak intensities from the 100 ms NOESY in CYANA to produce a model of Pin1 to

be used for spin diffusion corrections for the FRM approach. Cross-relaxation rates were converted into effective distances (where motional effects are absorbed) with

$$r_{ij}^{eff} = \sqrt[6]{\frac{56.94 \cdot \tau_c / \text{ns}}{\sigma_{ij} / \text{s}^{-1}}}$$

with tumbling assumed to be isotropic²⁸. Due to the overall correlation time τ_c impacting σ and the spin diffusion corrections²⁸, the buildup fits had to be split into three separate calculations due to each domain (and linker) having a distinct τ_c . The τ_c values used for the buildup rate calculation were 11.3, 3.6, and 14.1 ns for the WW domain, linker, and PPIase domain, respectively. All uni-directional buildups were normalized to the spin of destination ($j \rightarrow i$), except when the diagonal peak was missing or overlapped and then the eNOE was normalized to spin of origin ($i \rightarrow j$) if available²⁹. As described previously⁷, the quality of fits was inspected visually to determine an upper limit χ_N at 29,0000 as a cutoff to discard eNOEs above. Error tolerances for bi- and uni-directional eNOEs were applied in CYANA. Generic normalized NOE upper limit distances were created by supplying upper limit M_0 and average ρ calculated from good fits of relevant atom types, and given an additional error tolerance of 20%³⁰.

Exchange spectroscopy

Exchange rates of the isomerizing ligand FFpSPR were determined by fitting the ratios of the intensities of the cross peaks to the diagonal peaks of the *cis* and *trans* peaks from atom H^N in the ligand residue R5. Cross peaks between the *cis* and *trans* peaks do not appear without enzyme present, due to the timescale of uncatalyzed isomerization being too slow for this NOESY series. The mixing time-modulated diagonal and cross intensities were fitted globally to equations having ρ_c , ρ_t , k_{ct} , and k_{tc} as free parameters as previously described³¹. The uncertainties in the rate constants were estimated using 200 Monte Carlo simulations. An example spectrum and fits are shown in **Figure S3**, and fitted k_{ct} and k_{tc} values in **Table S1**, respectively. The exchange rate $k_{EXSY} = k_{CT} + k_{TC}$.

Relaxation and paramagnetic relaxation enhancements

Peak intensities for R_1 , $R_{1\rho}$, and R_2 (PRE) measurements were fitted as a function of the relaxation delay to single exponential decay using the non-linear least-squares (NlinLS) algorithm in NMRPipe¹⁶. Uncertainties were estimated using Monte Carlo simulations. ¹⁵N amide relaxation rates were used to define the domain-specific tumbling times, while ¹H amide R_2 were measured for PRE. To measure the PRE, we measured both the transverse relaxation of the oxidized, spin-labeled sample (R_2^*), and the reduced, unlabeled sample (R_2). The difference between the two rates results in the R_2^{sp} ($R_2^* - R_2 = R_2^{sp}$), which can then be converted into a distance between the free electron and the nuclei using a simplified the Solomon-Bloembergen equation³²:

$$r = \sqrt[6]{\left(\frac{\mu_0}{4\pi}\right)^2 \frac{K}{R_2^{sp}} \left\{ 4\tau_c + 3 \frac{\tau_c}{1 + (\omega_I \tau_c)^2} \right\}}$$

with constant $K = 1.23 \times 10^{-32} \text{ cm}^6 \text{ s}^{-2}$, μ_0 is the magnetic permeability of a vacuum, and ω_I is the Larmor frequency of the proton. Domain-specific τ_c values were applied. The error in R_2^{sp} (ΔR_2^{sp}) is obtained from the errors of the measured relaxation rates as:

$$\Delta R_2^{sp} = \sqrt{\Delta R_2^{*2} + \Delta R_2^2}$$

In addition to this symmetrically propagated error, 2 Å was linearly added for upper and lower limit restraints for structure calculation.

Residual dipolar couplings

To find the optimal domain-specific alignment tensors, the experimental RDCs were compared to back-calculated RDCs for various structures. The structures tested include 1pin³³ (crystal structure), 1nmv¹ (conventional NMR structure), preliminary 1- and 2-state eNOE structures, and 6svc³⁴ (1-state eNOE structure of the isolated WW domain). Final tensors were determined using the “FindTensor” script in CYANA using 6svc (Pearson’s correlation coefficient $r = 0.944$) and 1pin ($r = 0.917$) to fit the WW and PPlase domain, respectively. The WW tensor had a magnitude D_a of -6.535 with rhombicity R of 0.449 (as defined in CYANA), while the PPlase tensor had D_a of 16.46 with R of 0.242. During structure calculation, CYANA keeps D_a and R fixed, but allows the tensor to reorient. We have previously shown that for a rigid molecule (GB3) D_a is ~5% larger when fitted to a two-state ensemble than a single-state structure because averaging of the RDCs is partially absorbed into D_a in case of a single state. Multiple states effectively model this averaging such that the tensor is the tensor expected for a (hypothetically) rigid structure. Given the experimental error of the RDCs, addition of 5% to D_a had no noticeable impact on our calculations and we proceeded with the values obtained from the single-state fits.

Double Electron-Electron Resonance

All DEER data were analyzed using an earlier version of DeerLab (DL) based on Matlab³⁵ (release 0.9.2, available under <https://github.com/JeschkeLab/DeerLab-Matlab>) by modelling the background decay by a stretched exponential function (bg_strexp, $B(t) = \exp(-\kappa |t|^d)$) with the decay rate κ and the stretch factor d restraint to 0.02-1 μs^{-1} and 0.9-1.2, respectively. 4pDEER and artefact corrected 5pDEER data²³ were analyzed using the single-pathway 4pDEER experiment model (ex_4pdeer) including the modulation depth λ as a fit parameter. Compared to the latter approach, highly similar results were obtained by analyzing the primary 5pDEER data with the multi-pathway 5pDEER model (ex_5pdeer), consisting of an unmodulated pathway, the 5pDEER pathway and 4pDEER pathway (“artefact” that refocused at time $T_0^{(2)}$) with amplitudes Λ_0 , λ_1 and λ_2 , respectively. Parameter-free and Gaussian distance distributions $P(r)$ were computed via Tikhonov regularization, using generalized cross-validation (GCV) to select the optimal regularization parameter. Bootstrapping with 200 samples produced converged 95% confidence intervals, listed in **Table S2-S4** for all fitting parameters.

MMM simulations

DEER simulations of structural ensembles were produced using the Multiscale Modeling of Macromolecules (MMM) MATLAB toolbox^{36,37}. MTSL-rotamers were simulated in positions 15, 90, 98, and 131 to match our experimental MTSL sites. At least 20 rotamers at each site were simulated with the ambient-temperature rotamer library that has been shown to provide the best fits for samples obtained by flash-freezing. Distance distributions of an ensemble were then calculated for the different combinations of spin label positions.

Multi-state structure calculation

As described previously⁴⁰, bi-directional eNOEs (NOEs evaluated with both cross peaks and diagonals) had no error tolerance applied and the upper and lower limit (CYANA “upl” and “lol”, respectively) restraints were equal, as these are exact distances. For peaks with only one cross peak able to be evaluated, a tolerance of $\pm 20\%$ was applied to the conversion of lols and upls²⁹. Generic-normalized (gn) eNOEs were converted into upl with 20% error added³⁰. Additional methylene and methyl group errors were 7% and 10%, respectively. Multi-state structures used the same input restraints as a single state structure, but a symmetry restraint weight of 0.1 (0.01 for side chains) was used to keep identical heavy atoms together with a width of 1.2 Å. Scalar couplings (“cco”), RDCs (“rdc”) and loose helix angle restraints (“aco”) were used with annealing weights of [0, 0.5, 1, 1], [0, 0.5, 1, 1] and [1, 1, 0, 0], respectively.

To use PRE restraints in structure calculations, upper (upl) and lower (lol) limit distance restraints were given according to different PRE parameters in a similar fashion as described previously³². Residues too close to the MTSL label disappear completely from the spectra, preventing the measurement of relaxation rates. As the largest R_2^{sp} measured results in a distance ~ 13 Å, residues that have completely disappeared were given an upl of 15 Å with addition of 2 Å to account for error. Conversely, residues with a relaxation rate difference less than 5 s^{-1} have a large relative uncertainty. These residues were given only a lol restraint of 21 Å, as a $R_2^{sp} = 5\text{ s}^{-1}$ gives a distance of 23 Å with a 2 Å error. The ideal distance restraints originated from residues between 13-23 Å away from the spin-label as there was a clear difference in the plotted fits. These residues have a $R_2^{sp} > 10\text{ s}^{-1}$, and were given both a lol and upl according to the extracted distance with the experimental error and additional 2 Å error. Lastly, residues that had R_2^{sp} between 5 and 10 s^{-1} as the plotted rates had only a very minor difference in relaxation rate. Therefore, the rate was converted into a distance, but only the lol distance was used as a structural restraint. The distance restraints originating from the spin-labeled residue were attached to that residue’s C β atom, in order to restrain near the backbone rather than just moving the side chain. Therefore, the PRE restraints were implemented in CYANA similar to the NOE restraints but between C β and H^N atoms. While the weight of the PRE restraints was optimized, the final weight used for structure calculations was 0.01 to compensate for the larger absolute distances and errors. Only the interdomain (ID) PRE restraints were used in the structure calculations.

Due to the symmetry restraints needed to solve a multi-state structure in CYANA, a two-step calculation was performed to allow the domains to sample various positions. For example in the two-state calculation, the interdomain NOEs were almost entirely fulfilled in one state (compact), and therefore the domains can occupy an extended state for the other conformation. But, due to the symmetry restraint, the second state would be unable to form a more extended configuration due to the 1.2 Å distance limit. Therefore, we use a two-step calculation with only the WW domain symmetry restraints used in the first step, and then only the PPlase domain symmetry restraints in the second step. Restraints were used according to **Figure S1**. For Step 1, the WW domain is calculated using NOEs, scalar couplings, WW symmetry restraints (if multi-state), and WW RDCs fitted using the WW alignment tensor ($D_a = -6.535$, $R = 0.449$). The 10 calculations with the lowest WW target function (TF; proportional to the sum of squared violations) were then used as input for the second calculation by fixing the WW angles (residues

1-39). With the WW angles fixed, the PPlase domain is then calculated in Step 2 around the WW domain, with the orientation determined by the interdomain PREs, interdomain NOEs, and the RDCs. The RDC alignment tensor for Step 2 is based on the PPlase domain ($D_a = 16.46$, $R = 0.242$) but RDCs from both domains were utilized. In this way, the different overall alignment of the WW domain can be deconvoluted from the individual bond orientations giving rise to the individual WW domain RDCs, and will result in different positioning of the WW domain within the coordinate frame of the PPlase tensor. Note that for such calculations, the tensor of the stronger aligned domain (in our case the PPlase) must be used for calculations using all RDCs, such that the stronger RDC averaging of the less aligned domain will result in more diverse positions relative to the stronger aligned domain. CYANA is only able to read one conformer as angle input at a time, so therefore the 10 lowest TF WW domains were used as input for Step 2, with Step 2 needing to be calculated for each WW input.

To estimate the population of the two states, pseudo three- and four-state calculations were set up allowing only two distinct states (with populations of 66.6:33.3 and 75:25, respectively) through strict symmetry restraints (0.01 Å distance limit between heavy atoms). Similar to the conventional multi-state calculation, the WW domain was calculated first and then the five conformers with the lowest TF were used for subsequent PPlase calculation. We compare the average TF of these calculations to the TF of the conventional 2-state ensemble in **Figure 5B**. We do not externally specify which state (compact or extended) is higher populated, so both states could theoretically occupy the higher population.

All final single- and multi-state structure calculations were performed in CYANA-3.98 with 400 structures with random torsion angle values (except in Step 2 with the pre-calculated WW angles as fixed input) using the standard simulated annealing protocol with 100,000 torsion angle steps. The lowest TF conformer from each of the 10 calculations was selected for the final ensemble.

Supporting Information for Results and Discussion section

Five-pulse DEER is necessary to measure longer interdomain distances

The broadened echo-detected field-sweep of mutant 90-98 compared to other constructs, *e.g.* 15-90 shown in **Figure S5B**, points to a close proximity of the MTSL labels (~13 Å expected). At such a short distance, exchange interaction between the electron spins adds to their dipolar coupling. For this reason, we cannot reliably compute the distance distribution from the 90-98 4pDEER signal (**Figure S5D**). **Figure S6A-B** displays the 4pDEER data analysis for the other two intradomain (PPlase) mutants 98-131 and 90-131 and **Table S2** specifies the underlying fit parameters. Contrary to the model-free distance distribution, the Gaussian model does not accurately describe the time-domain 4pDEER data.

In order to access the longer interdomain distances between the WW and PPlase domain, we recorded 5pDEER traces for mutants 15-90, 15-98 and 15-131 (**Figure S7**). In the latter case, the 4pDEER contribution (“artefact”) to the 5pDEER signal coincides with the period of the dipolar oscillation (**Figure S5C**). Even in this challenging case, both the artefact correction (**Figures S5C**) combined with a 4pDEER analysis (**Figures S6C**) and a direct 5pDEER analysis (**Figure S7C**) in terms of dipolar pathways (**Table S4**) result in highly comparable distance distribution. Whereas the time-domain DEER data of 15-90 and 15-98 can be well described by a sum of two Gaussian

components (2Gauss) (**Figures S7A-B**), in case of 15-131 the Gaussian model somewhat oversimplifies the model-free distance distribution (**Figures S6C, S7C**).

Table S1: Isomerization rates measured using exchange spectroscopy (EXSY). Error is reported in parentheses. The peaks from 90-98 mutant were unable to be fit simultaneously which resulted in a large error to the exchange rate.

Pin1 Variant	$k_{ct} (s^{-1})$	$k_{tc} (s^{-1})$	$k_{EXSY} (s^{-1})$
WT	43.77 (1.17)	4.54 (0.38)	48.31(1.55)
AD+M15C	44.11 (4.22)	4.00 (0.42)	48.11 (4.64)
AD+N90C	42.09 (6.01)	3.95 (0.36)	46.04 (6.37)
AD+S98C	23.27 (1.34)	2.94 (0.45)	26.21 (2.79)
AD+Q131C	5.25 (0.45)	1.34 (0.08)	6.59 (0.53)
AD+M15C+N90C	35.03 (2.00)	4.59 (0.20)	39.63 (2.20)
AD+M15C+S98C	42.61 (6.67)	4.18 (0.53)	46.79 (7.20)
AD+M15C+Q131C	21.04 (2.09)	1.74 (0.13)	22.78 (2.22)
AD+N90C+Q131C	10.27 (0.16)	1.44 (0.08)	11.71 (0.24)
AD+S98C+Q131C	8.71 (0.91)	0.51 (0.05)	9.21 (0.96)
AD+N90C+S98C *	25.49 (13.33)	1.15 (0.58)	26.64 (13.91)

Table S2: Fit parameters for 4pDEER (90-131, 98-131) and corrected 5pDEER (15-131) analyzed in terms of a model-free and Gaussian distance distribution (see Figure S5). The values in brackets defines the 95% confidence interval.

mutant	ex_4pdeer	background		1 Gauss	
	λ	κ [μs^{-1}]	d	$\langle r \rangle$ [$\text{\AA}$]	σ [$\text{\AA}$]
90-131 Fig. S6A	0.22985 [0.22271 0.2362]	0.057594 [0.045669 0.096954]	1.1608 [0.93048 1.1811]	-	-
90-131 Fig. S6A	0.21491 [0.21342 0.21749]	0.12499 [0.10947 0.12839]	0.9 [0.9 1.0006]	26.127 [26.064 26.194]	4.4063 [4.2316 4.616]
98-131 Fig. S6B	0.21092 [0.20167 0.23071]	0.069802 [0.022226 0.093431]	0.93295 [0.91211 1.1627]	-	-
98-131 Fig. S6B	0.19047 [0.18811 0.19562]	0.12695 [0.094362 0.1325]	0.9 [0.9 1.085]	35.012 [34.823 35.213]	5.2584 [4.8203 5.7794]
15-131 Fig. S6C	0.14836 [0.14499 0.15475]	0.10664 [0.08486 0.1149]	0.94524 [0.91021 1.0038]	-	-
15-131 Fig. S6C	0.13807 [0.13644 .14034]	0.13186 [0.11962 0.13406]	0.9 [0.9 0.93818]	42.238 [42.077 42.381]	6.4904 [6.163 6.9201]

Table S3: Fit parameters for 5pDEER (15-90,15-98) analyzed in terms of a model-free and a sum of two Gaussian distance distribution (see Figure S7A-B). The values in brackets defines the 95% confidence interval.

Mutant	ex_5pdeer				background		2 Gauss				
	Λ_0	λ_1	λ_2	$T_0^{(2)}$ [μ s]	κ [μ s ⁻¹]	d	$\langle r_1 \rangle$ [$\text{\AA}$]	σ_1 [$\text{\AA}$]	ρ_1	$\langle r_2 \rangle$ [$\text{\AA}$]	σ_2 [$\text{\AA}$]
15-90 Fig. S7A	0.57727 [0.57274 0.58008]	0.3336 [0.33274 0.33985]	0.09287 [0.091697 0.096806]	3.0674 [3.0663 3.0684]	0.034349 [0.026318 0.038254]	0.96081 [0.91199 1.0642]	-	-	-	-	-
15-90 Fig. S7A	0.59131 [0.60077 0.61683]	0.35612 [0.35612 0.36652]	0.11018 [0.11056 0.11466]	3.0672 [3.0662 3.0683]	0.020039 [0.020274 0.035608]	1.1458 [0.9052 1.14]	21.546 [21.503 21.604]	2.6808 [2.5458 2.8532]	0.68184 [0.67339 0.71668]	39.134 [37.955 40.909]	30.458 [26.307 31.442]
15-98 Fig. S7B	0.5013 [0.46579 0.50312]	0.37413 [0.35067 0.37414]	0.12804 [0.12309 0.13389]	3.7576 [3.7549 3.7603]	0.096734 [0.09747 0.12078]	1.1957 [1.1014 1.1885]	-	-	-	-	-
15-98 Fig. S7B	0.48264 [0.47938 0.49166]	0.36418 [0.35896 0.36907]	0.12931 [0.12665 0.13365]	3.759 [3.7562 3.7618]	0.099532 [0.09554 0.12184]	1.1818 [1.0969 1.1979]	23.342 [23.195 23.505]	5.599 [5.1074 6.1964]	0.74406 [0.71934 0.76761]	44.348 [42.939 45.424]	17.479 [14.655 20.189]

Table S4: Fit parameters for 5pDEER (15-131) analyzed in terms of a model-free and Gaussian distance distribution (see Figure S7C). The values in brackets defines the 95% confidence interval.

Mutant	ex_5pdeer				background		1 Gauss	
	Λ_0	λ_1	λ_2	$T_0^{(2)}$ [μ s]	κ [μ s ⁻¹]	d	$\langle r \rangle$ [$\text{\AA}$]	σ [$\text{\AA}$]
15-131 Fig. S7C	0.85185 [0.837 0.89986]	0.11328 [0.10934 0.12593]	0.0489 [0.047385 0.056912]	3.7021 [3.6971 3.7357]	0.075026 [0.057658 0.10081]	1.0568 [0.94156 1.1319]	-	-
15-131 Fig. S7C	0.86045 [0.85034 0.8574]	0.10223 [0.10093 0.10429]	0.046096 [0.045023 0.04679]	3.6978 [3.6912 3.7047]	0.12069 [0.087381 0.1223]	0.9 [0.9 1.0315]	42.948 [42.787 43.171]	8.6381 [8.4334 9.179]

Table S5: Top) Domain-specific relaxation parameters. Bottom) Residue-specific relaxation parameters of major and minor peaks in the interdomain interface.

Domain	R_1 (s ⁻¹), averaged over residues	$R_{1\rho}$ (s ⁻¹) averaged over residues
WW	1.22	17.9
Linker	1.71	4.97
PPlase	0.87	23.4

Residue	State	R_1 (s ⁻¹)	$R_{1\rho}$ (s ⁻¹)	Residue	State	R_1 (s ⁻¹)	$R_{1\rho}$ (s ⁻¹)
29	Major	1.01	20.41	137	Major	0.72	28.57
	Minor	1.19	18.52		Minor	0.89	20.00
31	Major	1.07	21.74	140	Major	0.71	30.30
	Minor	1.42	11.90		Minor	0.83	19.23
32	Major	0.94	18.87	141	Major	0.77	27.78
	Minor	1.04	11.24		Minor	1.07	27.63

Table S6: Structural statistics and CYANA input data for apo Pin1.

NMR distance and dihedral constraints		
NOE distance constraints		
Total eNOEs	2268	
Bi-directional eNOEs	537	
Uni-directional eNOEs	1731	
gnNOEs	1937	
Interdomain NOEs	20	
Intra-residue, $ i-j =0$	569	
Sequential, $ i-j =1$	850	
Medium-range, $1< i-j <5$	727	
Long-range, $ i-j \geq 5$	947	
PRE interdomain restraints		
Total ID PRE restraints	250	
upl and lol	104	
upl only	11	
lol only	135	
Dihedral angle restraints		
$^3J_{\text{HNH}\alpha}$	124	
$^3J_{\text{H}\alpha\text{H}\beta}$	129	
$^3J_{\text{NC}\gamma}$	12	
Helix angle restraints ($\phi+\psi$)	66	
Residual dipolar couplings		
Total RDCs	407	
$^1D_{\text{NH}}$	114	
$^1D_{\text{C}'\text{C}\alpha}$	121	
$^1D_{\text{NC}'}$	86	
$^1D_{\text{C}'\text{H}}$	86	
	1-state ensemble	2-state ensemble
Structure Statistics		
CYANA target function ($\text{\AA}^2$)	316.25±7.62	148.30±3.89
NOE violations (>0.8 $\text{\AA}$)	43	15
Scalar coupling violations (>2 Hz)	25	5
RDC violations (>8 Hz)	14	5
PRE violations (>6 $\text{\AA}$)	30	7
Deviations from mean		
Backbone full WW ($\text{\AA}$)	0.51±0.16	0.90±0.11
Backbone 2° elements WW ($\text{\AA}$)	0.42±0.20	0.75±0.11
Backbone full PPlase ($\text{\AA}$)	1.08±0.35	0.94±0.08
Backbone 2° elements PPlase ($\text{\AA}$)	0.78±0.28	0.73±0.09

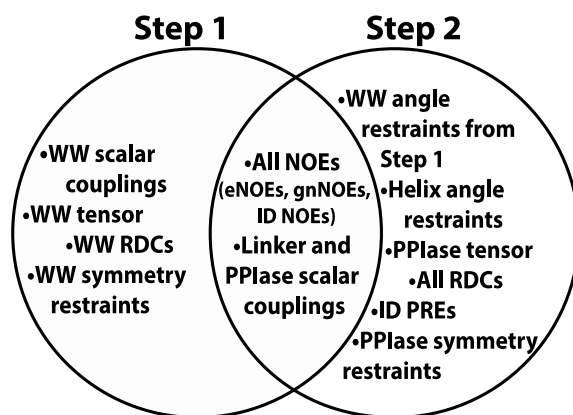


Figure S1: Diagram showing the data used in the two-step structure calculation in CYANA required for two-domain proteins.

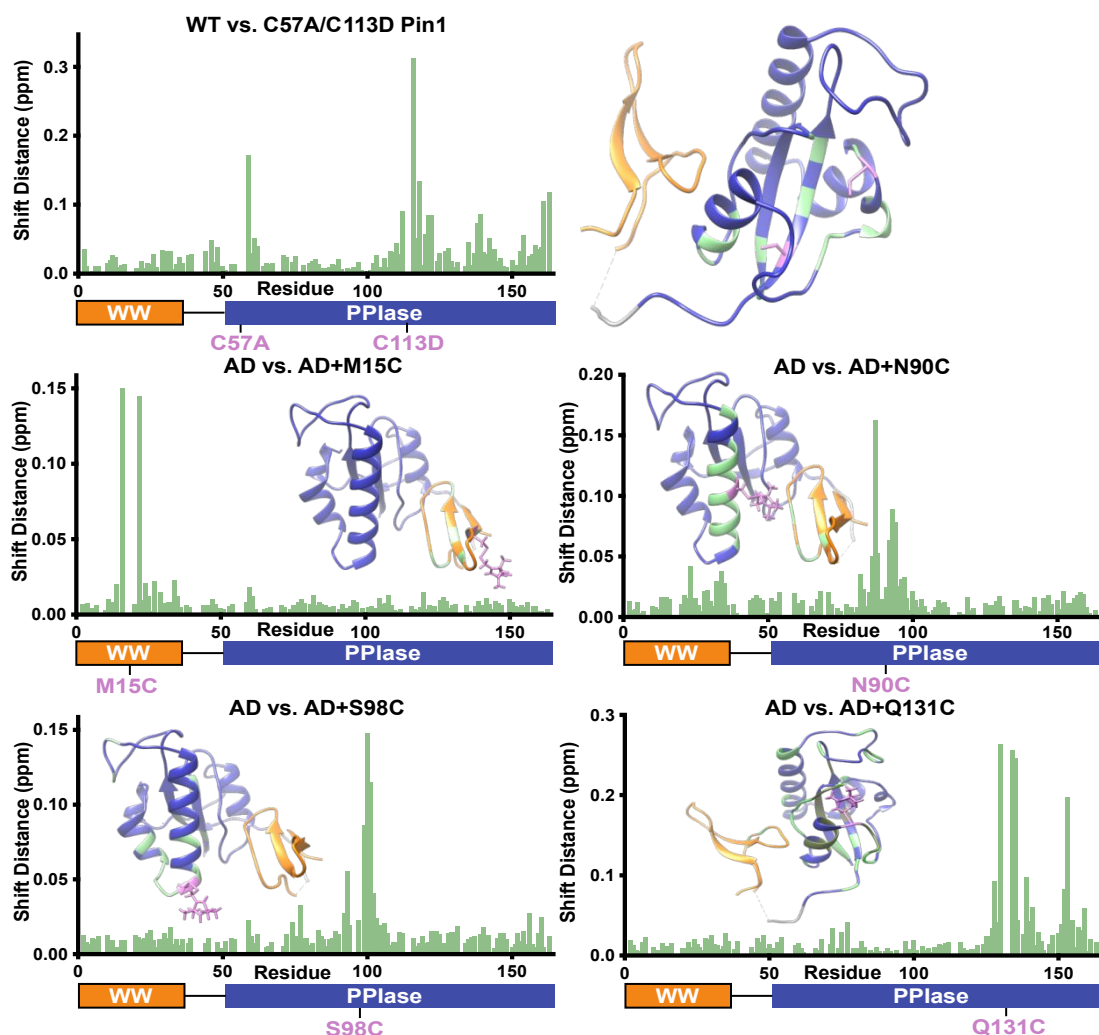


Figure S2: Chemical shift perturbations due to mutations of Pin1. Perturbations are calculated using the equation $\Delta\delta(ppm) = \sqrt{\Delta\delta_H^2 + 0.15(\Delta\delta_N^2)}$. Major perturbations are drawn in green onto the structure using cutoffs 0.05, 0.02, 0.02, 0.02, and 0.04 ppm for C57A/C113D, M15C, N90C, S98C, and Q131C, respectively. The mutated residues are shown in pink stick representation.

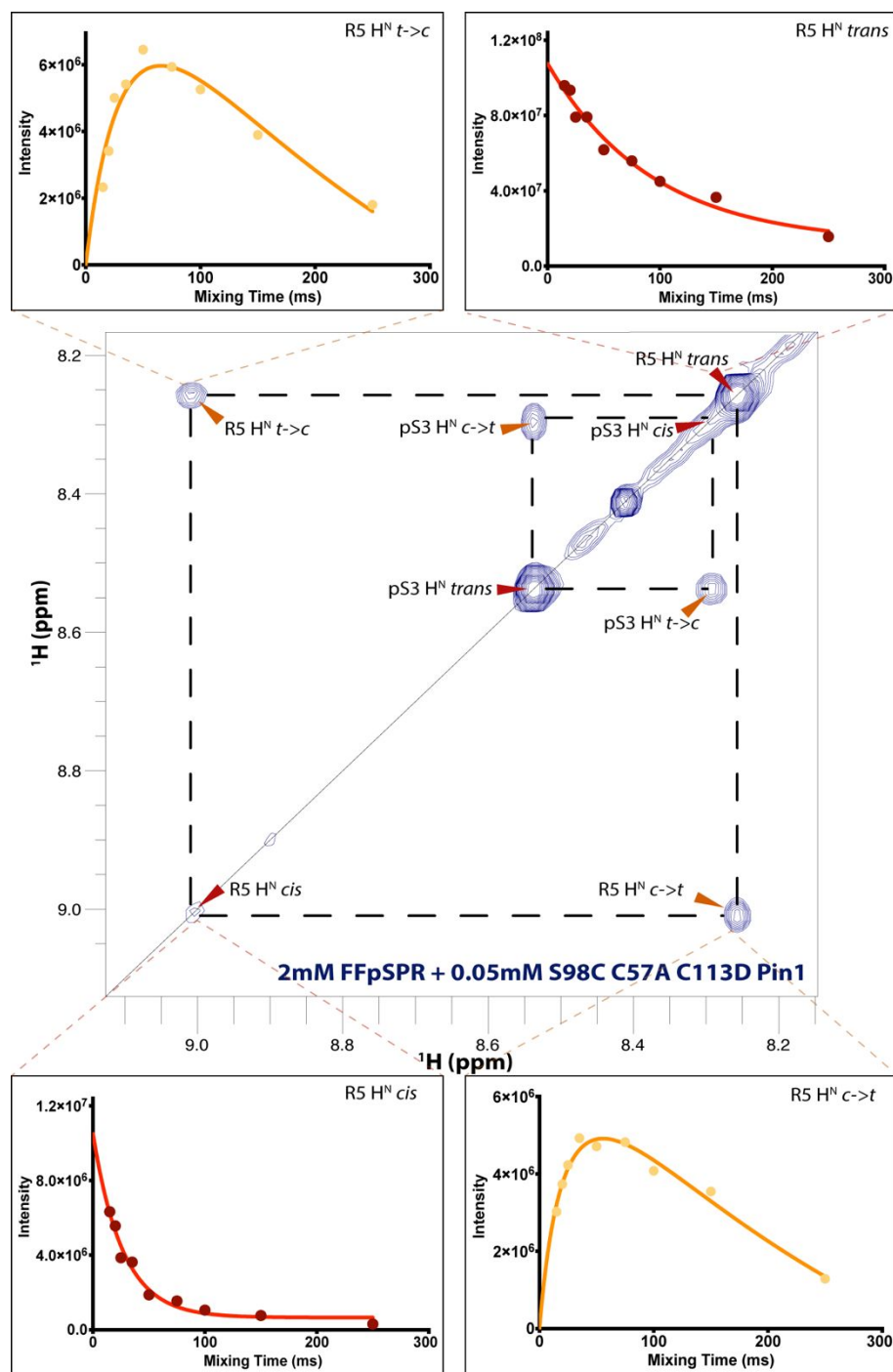


Figure S3: ^1H - ^1H exchange spectroscopy (EXSY) of ligand FFpSPR undergoing isomerization by C57A/C113D+S98C Pin1. Cross peaks from residue 3 and residue 5 are clearly present. Peak intensities versus mixing time of residue 5 and fits using the equations previously described³¹ are shown in the insets. Fitted k_{ct} and k_{tc} values are shown in Supplemental Table S1. Cross peaks do not appear without enzyme present, due to the timescale of uncatalyzed isomerization occurring too slowly for this EXSY series' timescale.

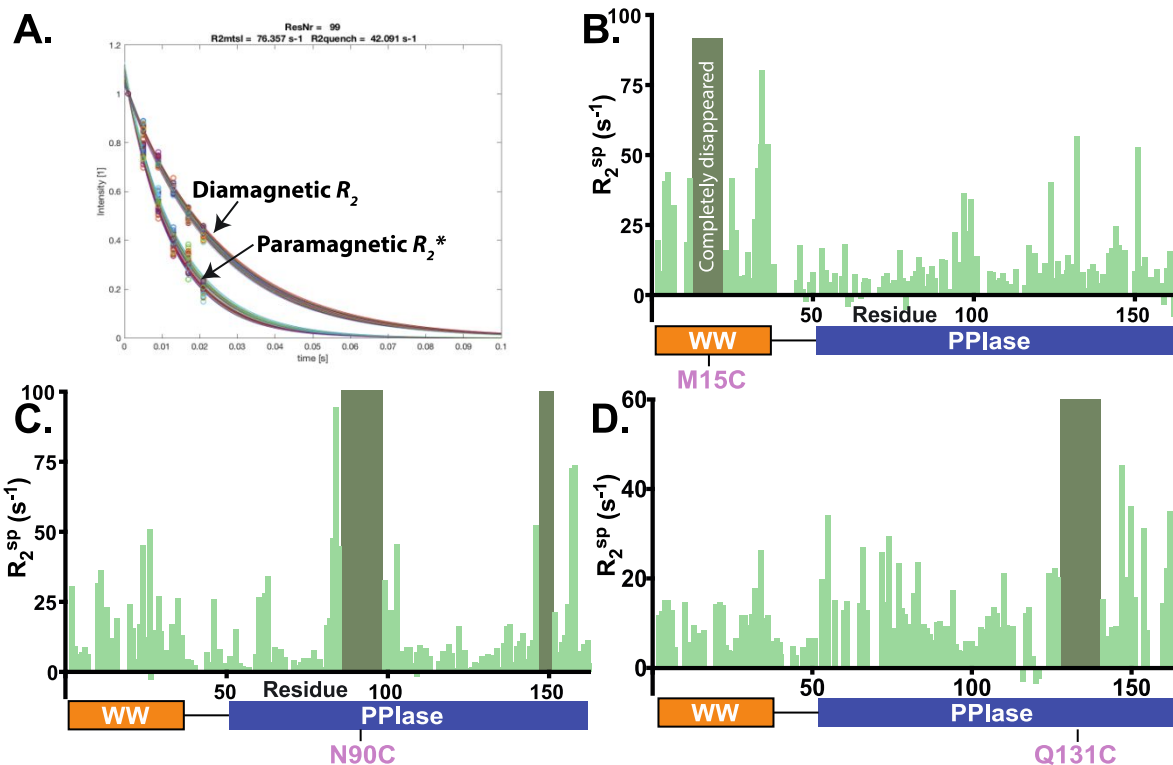


Figure S4: PRE on Pin1 mutants M15C, N90C and Q131C. **A)** Example of PRE intensity plot. 100 fits are shown for residues 99 from the diamagnetic and paramagnetic M15C construct, each obtained after adding normally distributed noise mirroring the experimental signal-to-noise ratio to each measurement point. The thus obtained errors are propagated into R_2^{sp} , which is the difference between the two rates. R_2^{sp} versus residue number are shown for constructs with MTSL at position **B)** M15C, **C)** N90C, and **D)** Q131C. Fits and plot were done in an inhouse-written MATLAB script.

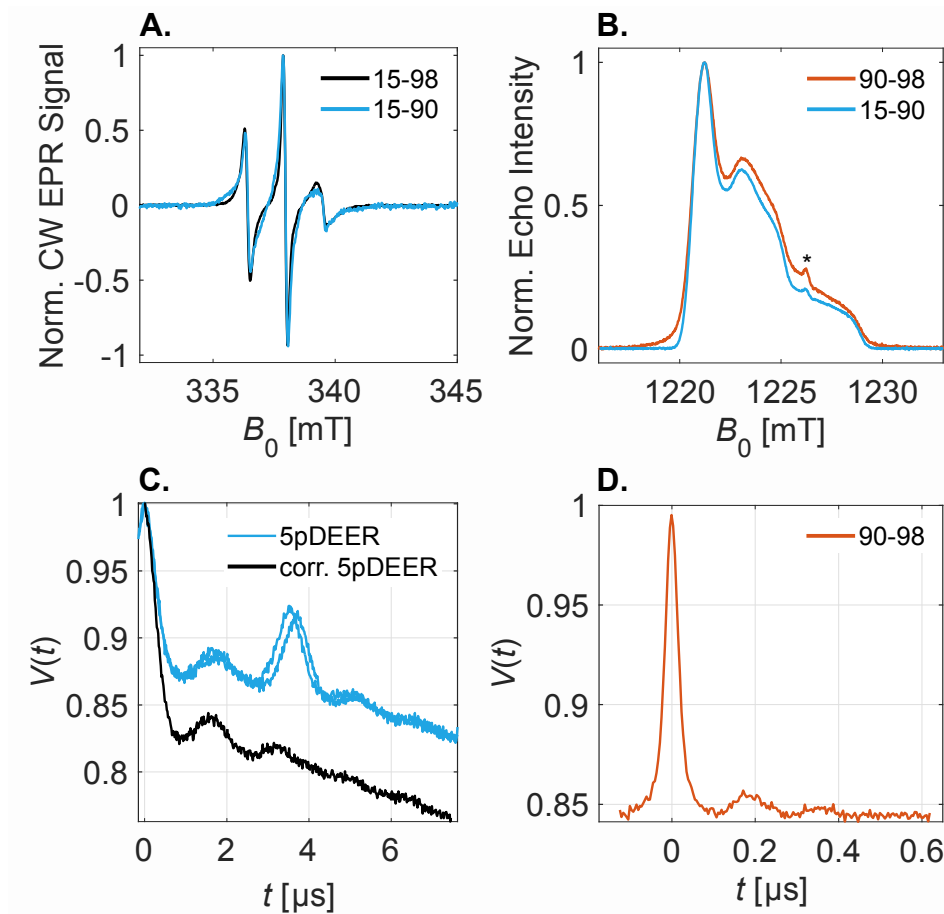


Figure S5: EPR on Pin1 double mutants. **A.** X-band CW spectra recorded at room temperature and 9.5 GHz. **B.** Q-band echo-detected field-sweep spectrum recorded at 50 K and 34.4 GHz. The asterisk marks a small peak that originates from E' centers in the quartz capillaries. **C.** Time-shifted 5pDEER traces ($\Delta t_0 = 160$ ns) recorded for 15-131 (blue) and artefact corrected 5pDEER trace (black). **D.** 4pDEER data of 90-98.

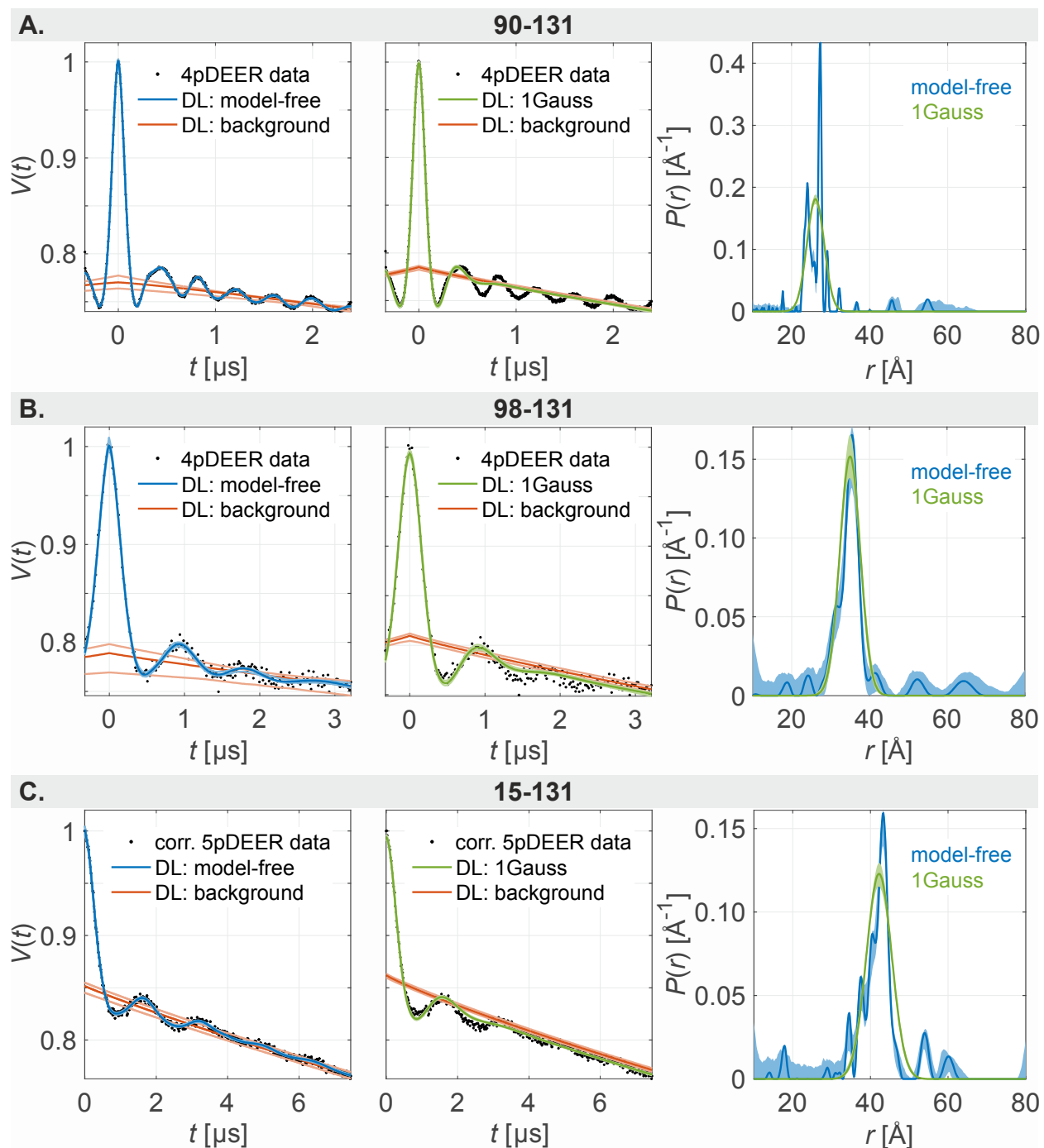


Figure S6: 4pDEER analysis of Pin1 double mutants 90-131, 98-131 and 15-131. DeerLab (DL) analysis of DEER data $V(t)$ using the 4pDEER model (ex_4pdeer), a stretched exponential background (orange) (see **Table S2** for fit parameters κ , d , λ) and different shapes of the distance distribution $P(r)$ to calculate the fit: model-free (left, blue) or a Gaussian distance distribution $P(r)$ (middle, green). The corresponding $P(r)$ are compared on the right for **A.** 90-131 **B.** 98-131 **C.** 15-131 with shaded areas representing the 95% confidence interval. See **Table S2** for the fit parameters of the Gaussian $P(r)$.

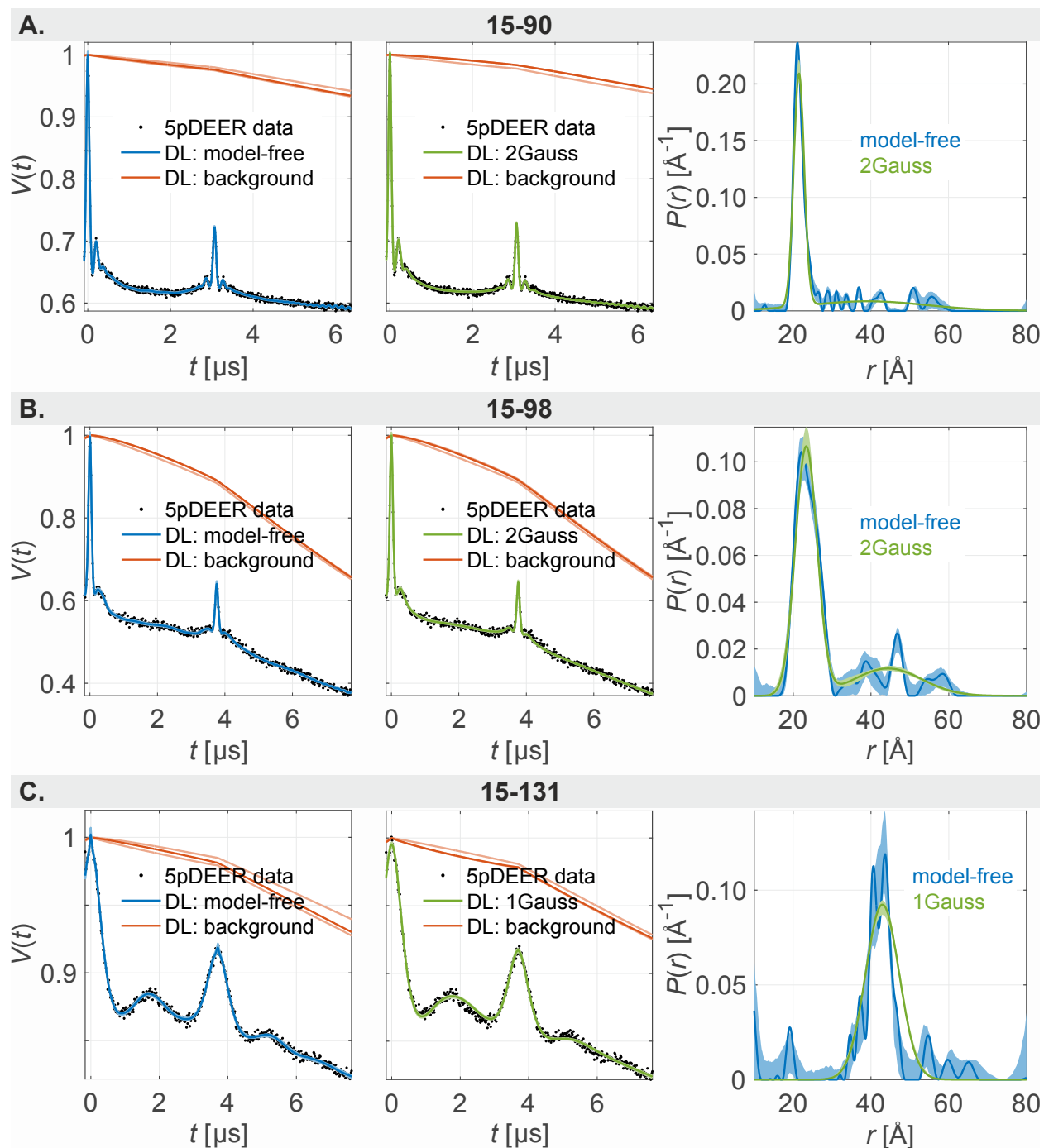


Figure S7: 5pDEER analysis of Pin1 double mutants 15-90, 15-98 and 15-131. DeerLab (DL) analysis of the DEER data $V(t)$ using the 5pDEER model (ex_5pdeer), a stretched exponential background (orange) (see **Table S3-4** for fit parameters κ , d , Λ_0 , λ_1 , λ_2 , $T_0^{(2)}$) and different shapes of the distance distribution $P(r)$ to calculate the fit: model-free (left, blue) or a Gaussian (1Gauss) or sum of two Gaussian (2Gauss) $P(r)$ (middle, green). The corresponding $P(r)$ are compared on the right for **A.** 15-90 **B.** 15-98 **C.** 15-131 with shaded areas representing the 95% confidence interval. See **Table S3-4** for fit parameters of the Gaussian $P(r)$.

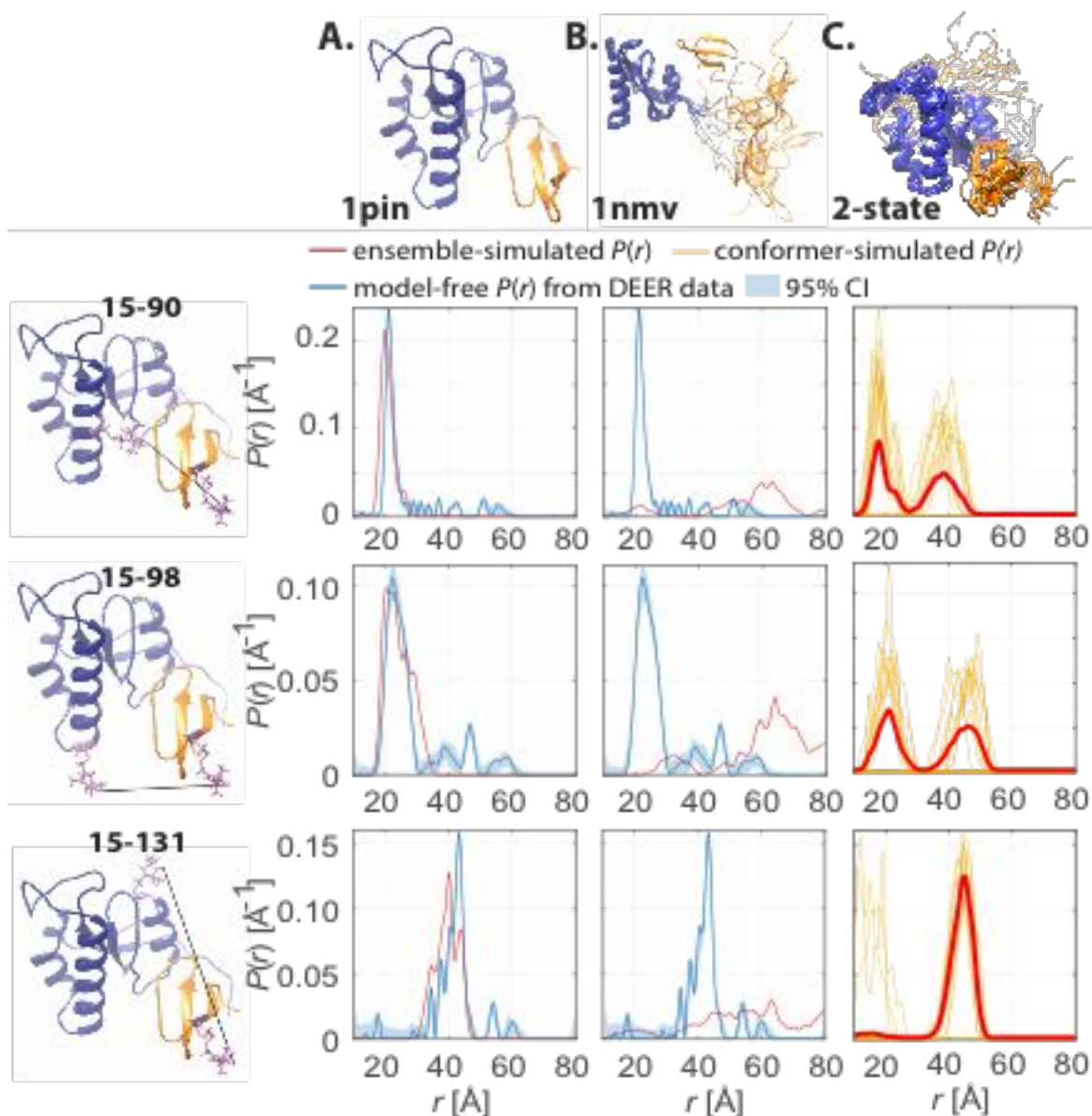


Figure S8. DEER-derived distance distributions with X-ray and NMR structures. Back-calculated, expected DEER-derived distance distributions from the published A) crystal structure 1pin³³ (red) and B) NMR structure 1nmv¹ overlaid with the experimental DEER distributions (blue). Note that the single-model structure of 1pin still results in a distribution of distances because all possible conformations of the electron-spin labeled side chain are modeled in MMM. C) MMM-simulated DEER distance distributions of the individual 2-state conformers (orange) in our ensemble overlaid with the ensemble-average over all conformers.

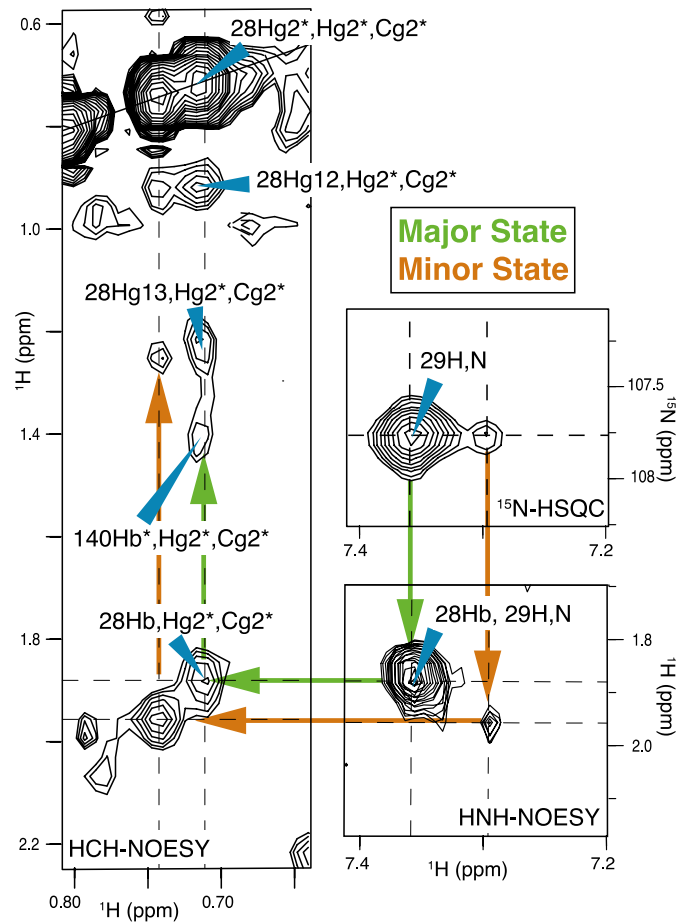


Figure S9: Interdomain interface peaks in ^{15}N -HSQC and 3D NOESY spectra show slow exchange between compact and extended states.

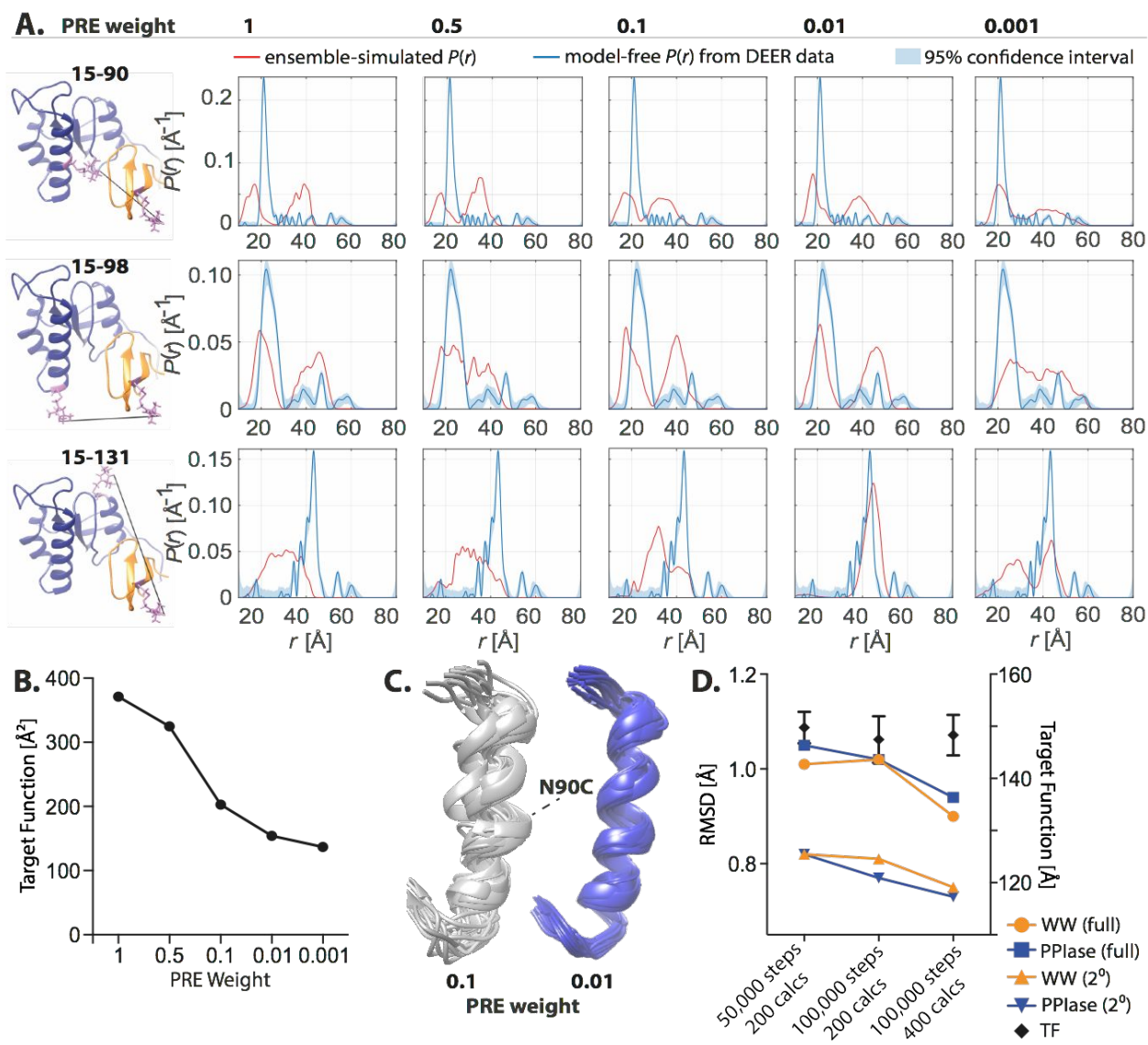
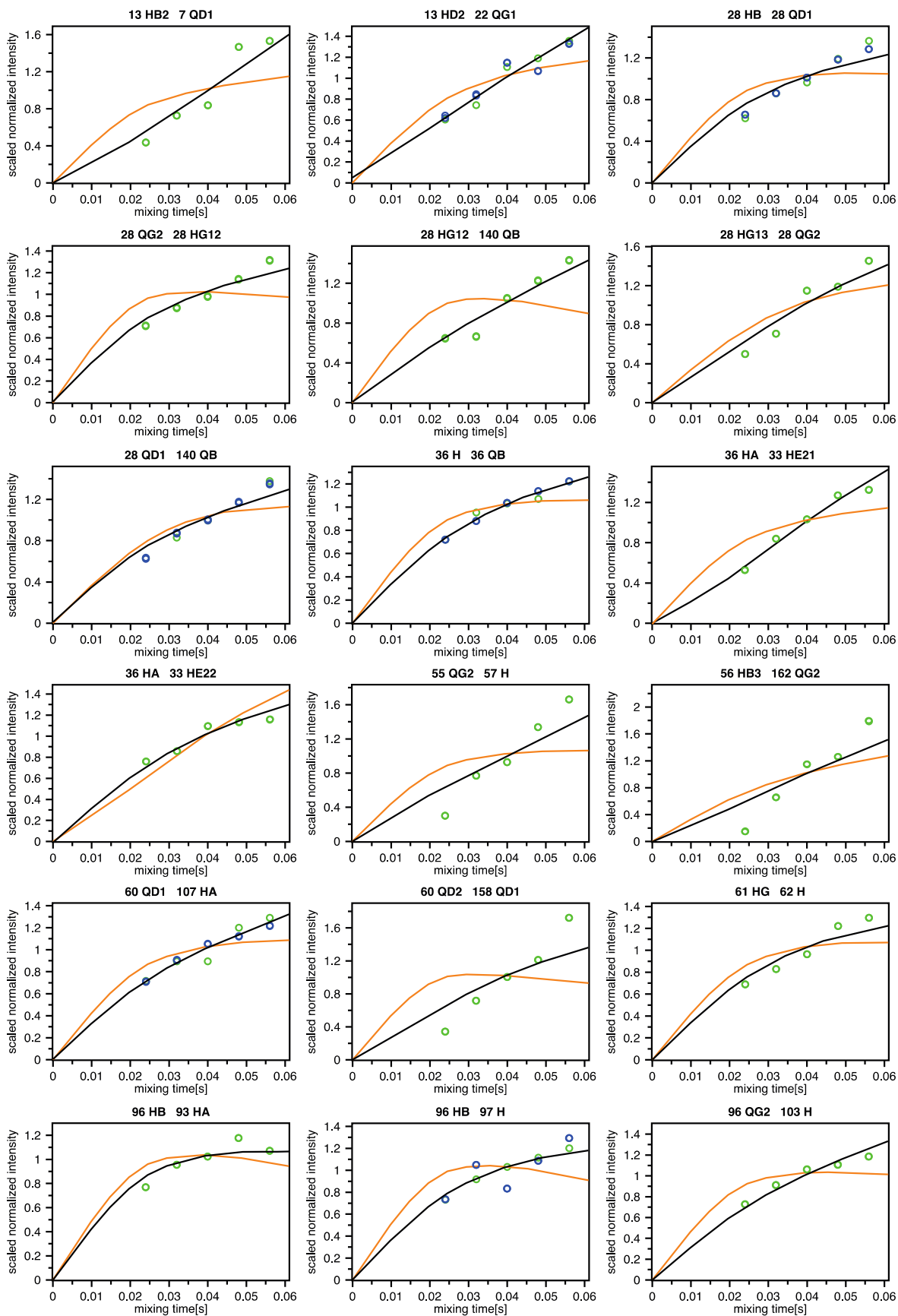
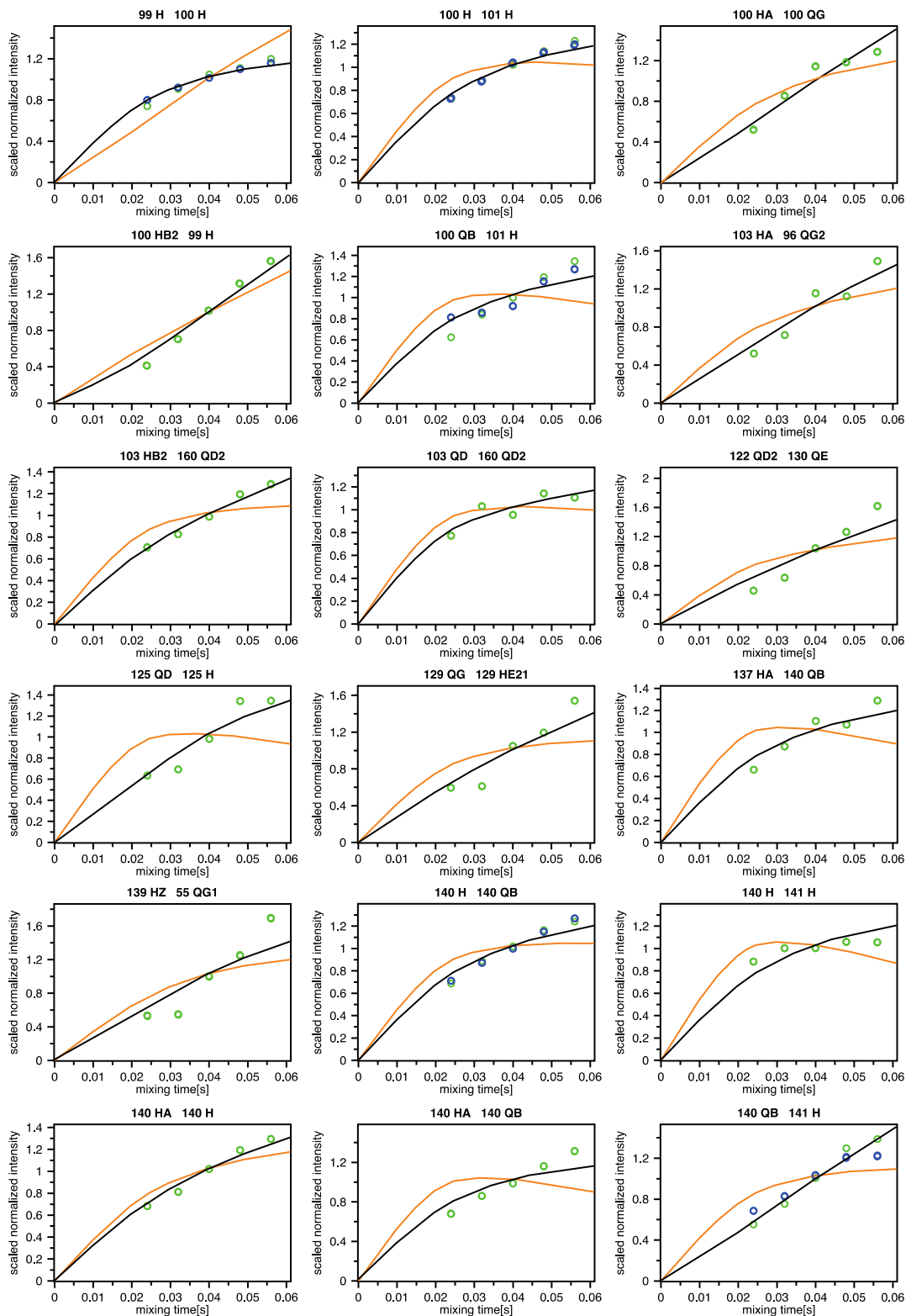


Figure S10: Optimization of structure calculation parameters. A) Back-calculated DEER distributions of two-state structure calculations with PRE distance restraints at different weights (in addition to RDC and ID NOE restraints) overlaid with experimental DEER distributions. B) PRE weight versus CYANA target function in two-state calculations. C) PRE weight versus $\alpha 1$ -helix structure in the two-state calculations. D) Number of torsion angle steps and number of initial structures were optimized by following RMSD (full domain and secondary structural elements) and target function in two-state calculations.





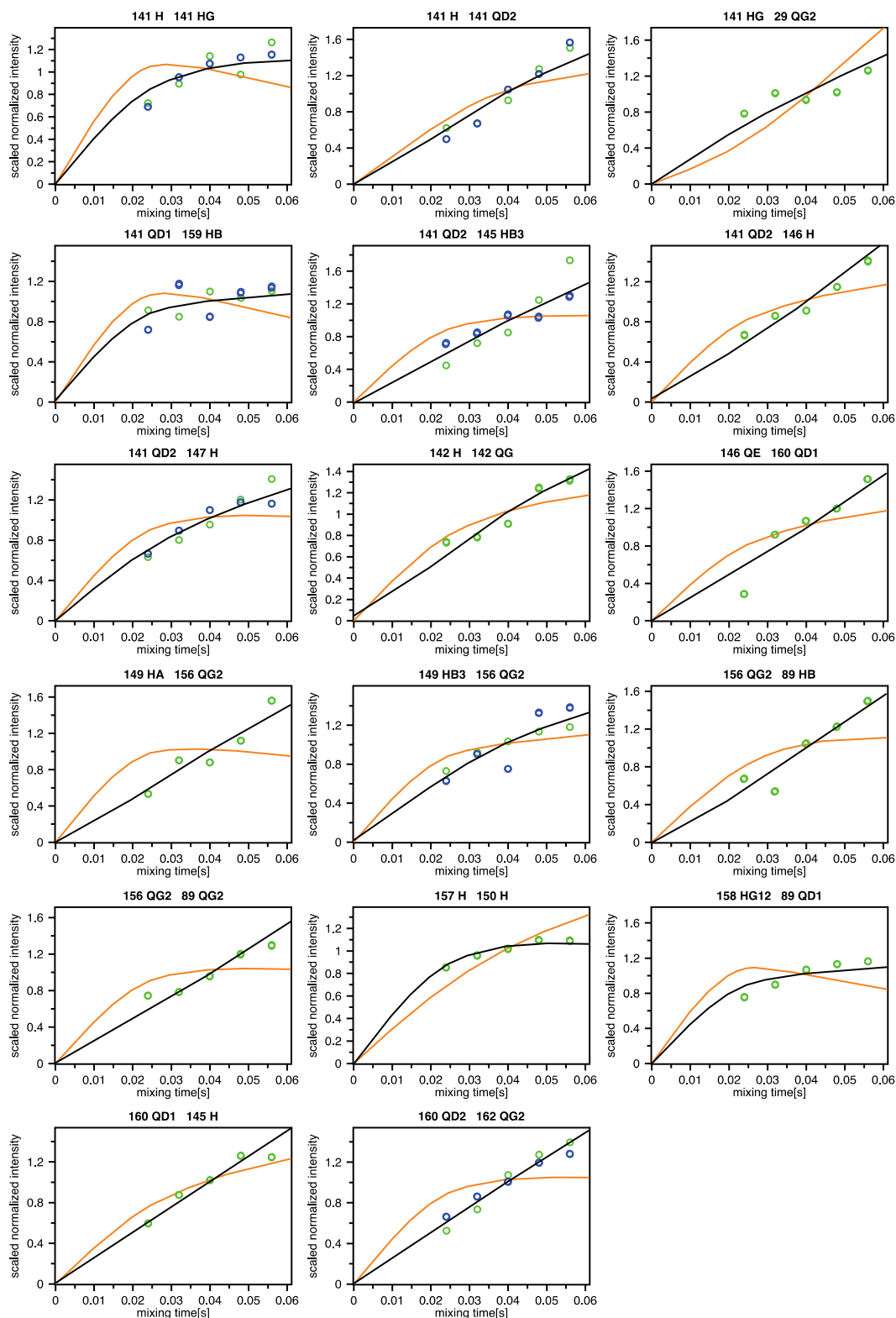


Figure S11. Back-predicted buildups of representative NOEs for single-state (orange) and two-states (black) ensembles calculated against experimental bidirectional eNOE buildups (green and blue dots). In regions of interest (i.e. WW binding site, ID interface, PPIase catalytic site, PPIase

hinge), the two-state ensemble fulfills the experimental data better than the single-state ensemble. The back-predicted buildups were calculated using eNORA2 in CYANA.

References

- (1) Bayer, E.; Goettsch, S.; Mueller, J. W.; Griewel, B.; Guiberman, E.; Mayr, L. M.; Bayer, P. Structural Analysis of the Mitotic Regulator HPin1 in Solution: Insights into Domain Architecture and Substrate Binding. *J. Biol. Chem.* **2003**, *278* (28), 26183–26193.
- (2) Born, A.; Parker, -; Nichols, J.; Henen, M. A.; Chi, C. N.; Dean Strotz, -; Bayer, - Peter; Tate, - Shin-Ichi; Peng, J. W.; Beat Vögeli, -. Backbone and Side-Chain Chemical Shift Assignments of Full-Length, Apo, Human Pin1, a Phosphoprotein Regulator with Interdomain Allostery. *Biomol. NMR Assign.* **2018**.
- (3) Born, A.; Henen, M.; Nichols, P.; Wang, J.; Jones, D.; Vögeli, B.; Born, A.; Henen, M. A.; Nichols, P.; Wang, J.; et al. Efficient Stereospecific H β 2/3 NMR Assignment Strategy for Mid-Size Proteins. *Magnetochemistry* **2018**, *4* (2), 25.
- (4) Matena, A.; Sinnen, C.; Van Den Boom, J.; Wilms, C.; Dybowski, J. N.; Maltaner, R.; Mueller, J. W.; Link, N. M.; Hoffmann, D.; Bayer, P. Transient Domain Interactions Enhance the Affinity of the Mitotic Regulator Pin1 toward Phosphorylated Peptide Ligands. *Structure* **2013**, *21* (10), 1769–1777.
- (5) Xu, N.; Tochio, N.; Wang, J.; Tamari, Y.; Uewaki, J. I.; Utsunomiya-Tate, N.; Igarashi, K.; Shiraki, T.; Kobayashi, N.; Tate, S. I. The C113D Mutation in Human Pin1 Causes Allosteric Structural Changes in the Phosphate Binding Pocket of the Ppiase Domain through the Tug of War in the Dual-Histidine Motif. *Biochemistry* **2014**, *53* (34), 5568–5578.
- (6) Vögeli, B.; Güntert, P.; Riek, R. Multiple-State Ensemble Structure Determination from ENOE Spectroscopy. *Mol. Phys.* **2013**, *111* (3), 437–454.
- (7) Nichols, P. J.; Born, A.; Henen, M. A.; Strotz, D.; Jones, D. N.; Delaglio, F.; Vögeli, B. Reducing the Measurement Time of Exact NOEs by Non-Uniform Sampling. *J. Biomol. NMR* **2020**, *74* (12), 717–739.
- (8) Kuboniwa, H.; Grzesiek, S.; Delaglio, F.; Bax, A. Measurements of HN-Ha J Couplings in Calcium-Free Calmodulin Using 2D and 3D Water Flip-Back Methods. *J. Biomol. NMR* **1994**, *4* (6), 871–878.
- (9) Grzesiek, S.; Kuboniwa, H.; Hinck, A. P.; Bax, A. Multiple-Quantum Line Narrowing for Measurement of H-Alpha-H-Beta J-Couplings in Isotopically Enriched Proteins. *J. Am. Chem. Soc.* **1995**, *117* (19), 5312–5315.
- (10) Hu, J.-S.; Grzesiek, S.; Bax, A. Two-Dimensional NMR Methods for Determining Angles of Aromatic Residues in Proteins From. *J. Am. Chem. Soc.* **1997**, *119* (96), 1803–1804.
- (11) Kay, L. E.; Torchia, D. A.; Bax, A. *Backbone Dynamics of Proteins As Studied by 15N Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease?*; London, 1989; Vol. 28.
- (12) Rückert, M.; Otting, G. Alignment of Biological Macromolecules in Novel Nonionic Liquid Crystalline Media for NMR Experiments. *J. Am. Chem. Soc.* **2000**, *122* (32), 7793–7797.
- (13) Fitzkee, N. C.; Bax, A. Facile Measurement of 1H-15N Residual Dipolar Couplings in Larger Perdeuterated Proteins. *J. Biomol. NMR* **2010**, *48* (2), 65–70.
- (14) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. Investigation of Exchange Processes by Two-Dimensional NMR Spectroscopy. *J. Chem. Phys.* **1979**, *71* (11), 4546–4553.
- (15) Donaldson, L. W.; Skrynnikov, N. R.; Choy, W. Y.; Muhandiram, D. R.; Sarkar, B.; Forman-Kay, J. D.; Kay, L. E. Structural Characterization of Proteins with an Attached ATCUN Motif

- by Paramagnetic Relaxation Enhancement NMR Spectroscopy. *J. Am. Chem. Soc.* **2001**, *123* (40), 9843–9847.
- (16) Delaglio, F.; Grzesiek, S.; Vuister, G.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* **1995**, *6* (3).
 - (17) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins Struct. Funct. Bioinforma.* **2005**, *59* (4), 687–696.
 - (18) Lee, W.; Tonelli, M.; Markley, J. L. NMRFAM-SPARKY: Enhanced Software for Biomolecular NMR Spectroscopy. *Bioinformatics* **2015**, *31* (8), 1325–1327.
 - (19) Polyhach, Y.; Bordignon, E.; Tschaggelar, R.; Gandra, S.; Godt, A.; Jeschke, G. High Sensitivity and Versatility of the DEER Experiment on Nitroxide Radical Pairs at Q-Band Frequencies. *Phys. Chem. Chem. Phys.* **2012**, *14* (30), 10762–10773.
 - (20) Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiess, H. . Dead-Time Free Measurement of Dipole–Dipole Interactions between Electron Spins. *J. Magn. Reson.* **2000**, *142* (2), 331–340.
 - (21) Borbat, P. P.; Georgieva, E. R.; Freed, J. H. Improved Sensitivity for Long-Distance Measurements in Biomolecules: Five-Pulse Double Electron-Electron Resonance. *J. Phys. Chem. Lett.* **2013**, *4* (1), 170–175.
 - (22) Breitgoff, F. D.; Soetbeer, J.; Doll, A.; Jeschke, G.; Polyhach, Y. O. Artefact Suppression in 5-Pulse Double Electron Electron Resonance for Distance Distribution Measurements. *Phys. Chem. Chem. Phys.* **2017**, *19* (24), 15766–15779.
 - (23) Breitgoff, F. D.; Polyhach, Y. O.; Jeschke, G. Reliable Nanometre-Range Distance Distributions from 5-Pulse Double Electron Electron Resonance. *Phys. Chem. Chem. Phys.* **2017**, *19* (24), 15754–15765.
 - (24) Strotz, D.; Orts, J.; Chi, C. N.; Riek, R.; Vögeli, B. ENORA2 Exact NOE Analysis Program. *J. Chem. Theory Comput.* **2017**, *13* (9), 4336–4346.
 - (25) Downing, A. K.; Güntert, P. Automated NMR Structure Calculation With CYANA. In *Protein NMR Techniques*; Humana Press: New Jersey, 2004; pp 353–378.
 - (26) Orts, J.; Vögeli, B.; Riek, R. Relaxation Matrix Analysis of Spin Diffusion for the NMR Structure Calculation with ENOEs. *J. Chem. Theory Comput.* **2012**, *8* (10), 3483–3492.
 - (27) Nichols, P. J.; Henen, M. A.; Born, A.; Strotz, D.; Güntert, P.; Vögeli, B. High-Resolution Small RNA Structures from Exact Nuclear Overhauser Enhancement Measurements without Additional Restraints. *Commun. Biol.* **2018**, *1* (1), 61.
 - (28) Vögeli, B. The Nuclear Overhauser Effect from a Quantitative Perspective. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *78*, 1–46.
 - (29) Strotz, D.; Orts, J.; Mingos, M.; Vögeli, B. The Experimental Accuracy of the Uni-Directional Exact NOE. *J. Magn. Reson.* **2015**, *259*, 32–46.
 - (30) Chi, C. N.; Strotz, D.; Riek, R.; Vögeli, B. Extending the ENOE Data Set of Large Proteins by Evaluation of NOEs with Unresolved Diagonals. *J. Biomol. NMR* **2015**, *62* (1), 63–69.
 - (31) Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. A Heteronuclear Correlation Experiment for Simultaneous Determination of ¹⁵N Longitudinal Decay and Chemical Exchange Rates of Systems in Slow Equilibrium. *J. Biomol. NMR* **1994**, *4* (5), 727–734.

- (32) Battiste, J. L.; Wagner, G. Utilization of Site-Directed Spin Labeling and High-Resolution Heteronuclear Nuclear Magnetic Resonance for Global Fold Determination of Large Proteins with Limited Nuclear Overhauser Effect Data. *Biochemistry* **2000**, *39*, 5879–5896.
- (33) Ranganathan, R.; Lu, K. P.; Hunter, T.; Noel, J. P. Structural and Functional Analysis of the Mitotic Rotamase Pin1 Suggest Substrate Recognition Is Phosphorylation Dependent. *Cell* **1997**, *89*, 875–886.
- (34) Strotz, D.; Orts, J.; Kadavath, H.; Friedmann, M.; Ghosh, D.; Olsson, S.; Chi, C. N.; Pokharna, A.; Güntert, P.; Vögeli, B.; et al. Protein Allostery at Atomic Resolution. *Angew. Chemie - Int. Ed.* **2020**, *59*, 2–10.
- (35) Fábregas Ibáñez, L.; Jeschke, G.; Stoll, S. DeerLab: A Comprehensive Software Package for Analyzing Dipolar Electron Paramagnetic Resonance Spectroscopy Data. *Magn. Reson.* **2020**, *1* (2), 209–224.
- (36) Jeschke, G. MMM: A Toolbox for Integrative Structure Modeling. *Protein Sci.* **2018**, *27* (1), 76–85.
- (37) Polyhach, Y.; Bordignon, E.; Jeschke, G. Rotamer Libraries of Spin Labelled Cysteines for Protein Studies. *Phys. Chem. Chem. Phys.* **2011**, *13* (6), 2356–2366.