VIP Structural Biology Very Important Paper

Zitierweise:
Internationale Ausgabe: doi.org/10.1002/anie.202008734
Deutsche Ausgabe: doi.org/10.1002/ange.202008734

Protein Allostery at Atomic Resolution

Dean Strotz, Julien Orts, Harindranath Kadavath, Michael Friedmann, Dhiman Ghosh, Simon Olsson, Celestine N. Chi, Aditya Pokharna, Peter Gntert, Beat Vçgeli* und Roland Riek*

Abstract: Protein allostery is a phenomenon involving the long range coupling between two distal sites in a protein. In order to elucidate allostery at atomic resoluion on the ligand-binding WW domain of the enzyme Pin1, multistate structures were calculated from exact nuclear Overhauser effect (eNOE). In its free form, the protein undergoes a microsecond exchange between two states, one of which is predisposed to interact with its parent catalytic domain. In presence of the positive allosteric ligand, the equilibrium between the two states is shifted towards domain–domain interaction, suggesting a population shift model. In contrast, the allostery-suppressing ligand decouples the side-chain arrangement at the inter-domain interface thereby reducing the inter-domain interaction. As such, this mechanism is an example of dynamic allostery. The presented distinct modes of action highlight the power of the interplay between dynamics and function in the biological activity of proteins.

Introduction

Allostery in proteins describes the process by which a signal such as ligand binding on one site of a protein or protein complex is transmitted to another distal functional site thereby regulating biological activities.[1] Several models on the mechanism of allostery have been postulated including the sequential mechanism,[2] the population shift model (including the conformational selection mechanism originally termed the symmetric model; Monod[3]), and the dynamic allostery model.[4] While the sequential mechanism assumes adaptability of the structure upon ligand binding, the model by Monod is based on the existence of two pre-existing exchanging states whose population equilibrium shifts upon ligand binding since the ligand selects one of the two states. The dynamic allostery model assumes that ligand binding changes the frequency and amplitude of thermal fluctuations within a protein without perturbing the average structure.

Experimental elucidation of allostery as „an action at a distance“ phenomenon is challenging.[1] The challenge is due to the availability of mostly low resolution, local data in NMR—including relaxation studies [5, 6]—or individual determined structures of trapped states (such as free and ligand-bound states, or intermediate states stabilized by for example mutagenesis). Integrating analysis of experimental data with molecular dynamic simulations using for example recently developed statistical methods is however emerging as an interesting and powerful approach.[7–11]

Recent progress in NMR-based methods opened an avenue towards a more holistic description of motion and ensembles of structures. These include residual dipolar coupling (RDC) measurements, relaxation dispersion NMR experiments, cross-correlated relaxation (CCR), paramagnetic relaxation enhancement (PRE), and exact Nuclear Overhauser Enhancement or Effect (eNOE) data in combination with molecular dynamics simulation, structure prediction software, or ensemble-based structure calculations.[12–26]

Here, we made use of the eNOE approach, which allows for the multi-state structure determination of well behaving proteins because of the high accuracy (i.e. < 0.1) of the ensemble-averaged restraints obtained[23, 25, 27, 28] and applied it to a variant of the prototypical allostery-comprising WW domain of Pin1 (see the Material and Methods section in the Supporting Information). Pin1 is a peptidyl-prolyl cis-trans isomerase (PPIase). Its biological significance includes amongst others an involvement in the regulation of mitosis,[29] a protective function against Alzheimers disease,[30] increase of hepatitis C infection[31] and it is overexpressed in many human cancer cells.[32] Pin1 contains an N-terminal WW domain (the name has its origin in the presence of two Trp

[*] Dr. D. Strotz, Dr. J. Orts, Dr. H. Kadavath, Dr. M. Friedmann, D. Ghosh, A. Pokharna, Prof. Dr. P. Gntert, Prof. Dr. R. Riek
Laboratory of Physical Chemistry,
Swiss Federal Institute of Technology, ETH-Hçnggerberg
8093 Zrich (Switzerland)
E-Mail: roland.riek@phys.chem.ethz.ch

Dr. S. Olsson
Department of Mathematics and Computer Science,
Freie Universitt Berlin, Arnimallee 6, 14195 Berlin (Germany)

Dr. C. N. Chi
Department of Medical Biochemistry and Microbiology,
Uppsala Biomedical Center, Uppsala University
751 23 Uppsala (Sweden)

Prof. Dr. P. Gntert
Institute of Biophysical Chemistry,
Center for Biomolecular Magnetic Resonance, and
Frankfurt Institute for Advanced Studies,
J.W. Goethe-Universitt
Max-von-Laue-Str. 9, 60438 Frankfurt am Main (Germany),
and
Graduate School of Science, Tokyo Metropolitan University
Hachioji, Tokyo 192-0397 (Japan)

Dr. B. Vçgeli
Department of Biochemistry and Molecular Genetics,
University of Colorado at Denver
12801 East 17th Avenue, Aurora, CO 80045 (USA)
E-Mail: beat.vogeli@cuanschutz.edu

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.202008734.

residues) separated by a flexible linker from the C-terminal catalytic PPIase domain (Figure 1 A).[33]

The 34-residues-long N-terminal WW domain is thought to be responsible for ligand recognition and binding as evidenced by NMR titration experiments (Figure 1), the C-terminal domain contains the catalytically active site. The two domains, which interact loosely via Loop 2 of the WW domain (i.e. residues I28-Q33), form the ligand binding site and the extent of interaction depends on the ligand that binds at a distal Loop 1 comprising residues M15-R21.[34–36] One family of substrate (such as the peptide pCdc25C of interest here) reduces the inter-domain contact, while other peptide families

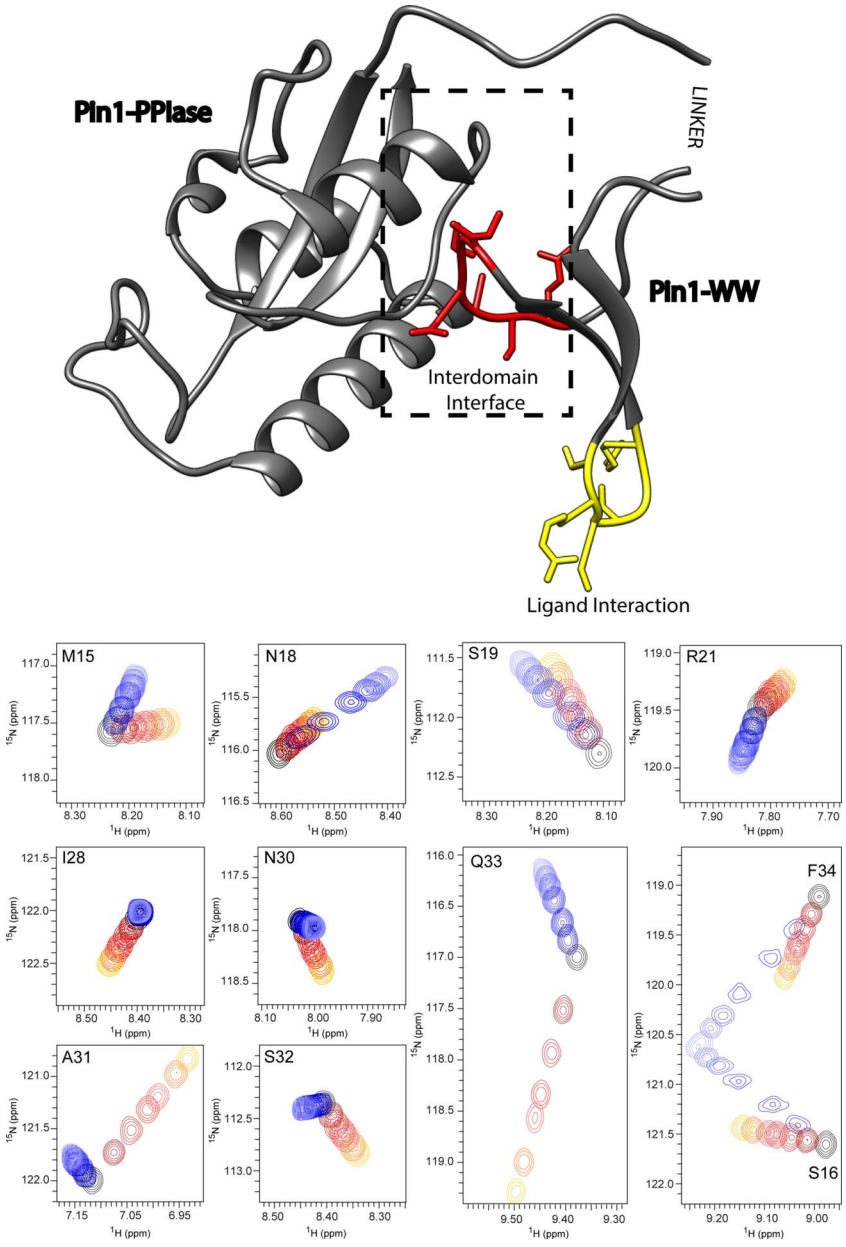


Figure 1.The 3D structure of Pin1 with its postulated allosteric interaction within the variant WW domain. The allostery within the WW domain occurs between the ligand-binding site (indicated in yellow) and the interaction site with its catalytic domain is highlighted in red and labeled as the inter-domain interface.[38–41, 43–45] This interaction is shown on top of the 3D crystal structure (PDB code 1PIN) represented by a ribbon with the residues of interest also highlighted by side chains. The NMR chemical shift titrations of the ^{15}N-labeled variant WW with the positive allosteric peptide FFpSPR and the negative allosteric peptide pCdc25C are shown for the relevant residues (i.e. M15-R21 form the ligand binding site and I28-Q33 are residues in the inter-domain interface) measured by ^{15}N[^{1}H]-HSQC experiments. The black cross peak corresponds to the ^{15}N-^{1}H moiety of the apo form. Upon titration with FFpSPR highlighted by the color code ranging from black over red to yellow, the cross peaks move with increasing concentration away from the apo form. Similarly, the color changes from dark to light blue indicate the chemical shift changes due to the interaction with the ligand pCdc25C. The shift changes indicate a fast exchange regime (i.e. ms time regime). In the binding site, the cross peaks move in the same direction for both ligands, while at the interface they shift in opposite directions. While the interface peaks undergo smaller shifts for the negative-allosteric ligand pCdc25C in comparison with FFpSPR, the binding-site peaks show similar shift magnitudes.

(such as the peptide FFpSPR of interest here) enhance the inter-domain contact.[35, 37, 38] Extensive studies at low resolution by others show that these properties require a substantial allosteric cross-talk between Loop 2 and the ligand binding site Loop 1 of the WW domain. [38–45] Thus, the WW domain possesses a ligand-mediated allosteric coupling. Experimental evidence includes the ligand titration NMR experiments with the isolated WW domain, showing chemical shift changes at the distal site Loop 2 upon ligand binding at Loop 1 (Figure 1).[38–41, 43–45] Furthermore, the two ligands induce distinct chemical shift changes in direction and magnitude on Loop 2 in line with their opposing property in the inter-domain interaction (Figure 1).[38] In order to explore the nature of this allosteric coupling at atomic resolution eNOE ensemble structures were determined of the apo-state of the WW domain as well as the WW domain in presence of either of the two peptide ligands FFpSPR and pCdc25C.

Results

Multi-State Structure Determination of the WW domain in Absence and Presence of Ligands

Following an established protocol[23, 25] with the eNORA2 program[46] ([47] CYANA version), ensemble structure calculations were performed for all three systems with eNOE-based distance restraints (Figure 2 A) and scalar couplings (Supplementary Tables S1–S3 and Material for more details). In Figure 2 B, the large number of restraints is demonstrated for Trp11 in the apo form of the WW domain, for which ca. 60 distance restraints have been collected, while on average there are roughly 20 eNOE-derived distance restraints per residue. As a measure of the quality of the calculated structures, the CYANA target function (TF), which is a weighted sum of all squared violations of the experimental restraints, is used. It drops significantly from one state to two states and levels off after three states (Figure 2 C). In a ten-state structure calculation, the two states are still observed (as exemplified in the Ramachandran plot for Thr29 in Figure S1) further supporting the two-state nature of the system. While the TF is an insensitive measure for determining the populations of the two states (Figure S2A), the details of the two

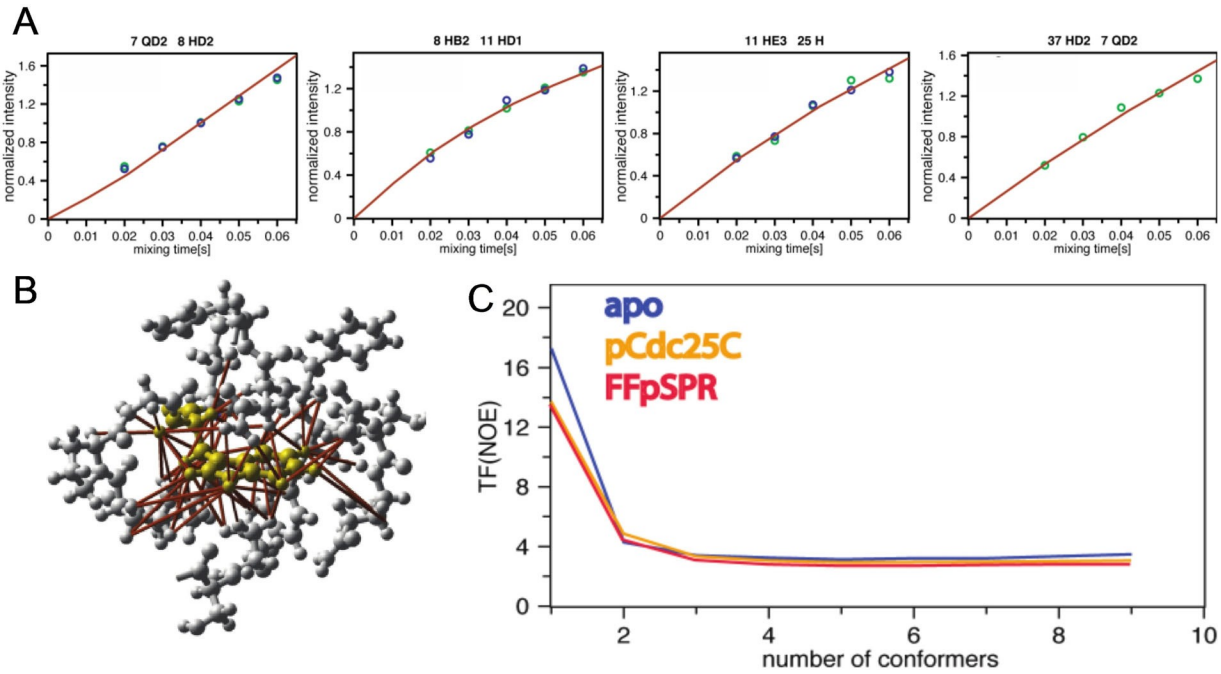


Figure 2. Distance restraint collection and structure calculation. A) Five experimental eNOE measured at 20, 30, 40, 50, and 60 ms NOESY mixing time of the apo WW domain, color coded by green and blue dots corresponding to the two respective NOEs on both sides of the diagonal versus time are shown. In addition, these data are superimposed with back-predicted buildups derived from the calculated two-state ensembles. The NOEs are between residues indicated at the top of each graph. The intensities are normalized to the average value of each buildup. The back-predicted theoretical model fulfils the experimental data very well (see also Figure S2). The back-predicted buildups were calculated using eNORA2 implemented in CYANA[46] ([47] CYANA version). B) eNOE-derived distance restraints around Trp11 of apo WW are mapped onto the 3D structure, indicating the large size of the data set. Trp11 and residues around Trp11 are shown in yellow and grey, respectively. Over 60 eNOE distance restraints (highlighted in red) were collected contrasting the four degrees of freedom of a Trp. This highlights the high density of information obtained by eNOE-based structure determination. C) CYANA target function (TF) values of various ensemble-based structure calculations, demonstrating the importance of the ensemble-based structure. The CYANA TF, which is the (weighted) sum of the squared violations of the conformational restraints versus number of simultaneously calculated states, is shown for all three calculations. The decrease of the TF with an increasing number of states indicates that at least two states are required to describe the experimental data well.

states are preserved (including the correlated/non-correlated configurations of Thr29 and Ala31 discussed below) for a population range 1:1— 1:3 for apo WW and WW domain in complex with pCdc25C, while for WW—FFpSPR it is conserved in the range of 1:9–4:6 (Figure S2A).

In order to get further insights into the relative population of the two states, we conducted titration experiments with FFpSPR using a T2-filtered [^{15}N,^{1}H]-HSQC experiment to enhance the signal broadening. The broadest signal can be attributed to a 1:1 population allowing a determination of the relative populations of the two states of 1:3 for apo WW (Figure S2B). To support this finding a ^{15}N-resolved CEST—[^{15}N,^{1}H]-TROSY experiment was conducted for apo WW (Figure S2C). The CEST data of the allosteric sites Ala31 and Gln33 albeit at the noise level show the presence of two states, one corresponding to the fully FFpSPR-bound state (Figure S2C, blue arrow), while the other allosteric state can also be identified (Figure S2C, cyan arrow) and agrees well with the elucidation of the shifts of the two states by the titration (Figures S2B and S2C). With the knowledge of the chemical shifts of the samples used for the structure determination, it can be estimated that the two states of the apo WW domain are present in a ratio of 3:1, while in the WW-FFpSPR complex the populations are 1:4. ^{15}N relaxation measurements at two magnetic fields in concert with the titration experiments revealing the chemical shift difference between the two states yielded individual exchange rates (Figure S2D). A similar exchange rate of ca 100 kHz is observed for both the ligand binding site (i.e. Ser19) as well as the allosteric site (i.e. Ile28, Asn30, Ala31, Ser32, and Gln33) supporting again the presence of two states that exchange in a concerted fashion.

Overall, these findings indicate that, in contrast to the single-state structure,multi-state ensembles describe the experimental data well (Tables S1–S3).

Validating the Multi-State Structures

The agreement of the model with the experimental data is also illustrated by the superposition of experimental NOE data and back-calculated NOE buildups (Figure 2 A). The improvement of the model with respect to the experimental data is shown even more explicitly by comparing back-predictions derived from single-state structures in comparison to the model derived from two-state structures (Figure S3). Furthermore, a cross-validation test with cross-correlated relaxation data (not used in the structure calculations) fit better with the two-state structural ensembles than the single-state structures (Figure S4). Finally, a cross-validation test was performed with a jackknife procedure that repeats the structure calculation twenty times with 5 % of the experimental input data randomly deleted such that each distance restraint is omitted exactly once. These obtained structures are similar to the original structures including the correlated states between Thr29 and Ala31 of interest below (Figure S5). As a representative for the following discussion, the two state ensembles described by a structural bundle of 2 20 conformers (Figure 2) are used.

Discussion

The Apo State of the WW Domain Comprises Two Distinct Conformational States

Inspection of the apo WW domain two-states ensemble reveals two spatially well-separated states from the ligand-binding site including Loop 1 via the backbone of the b-strand b2 and Asn26 to the inter-domain site Loop 2 (Figure 3 A). The two states are well separated both by the side chains shown (i.e. Arg14, Met15, Ser16, Ser19, Tyr23, Phe25, Asn26, Ile28, Thr29, Asn30, Ala31, Gln33) as well as the backbone angles for Thr29 and Ala31 highlighted in Ramachandran plots (Figure S6). Since only one set of chemical shifts is observed, it is suggested that the two states interchange in the micro-millisecond time range in a concerted fashion between the side chain of the ligand binding site (such as Arg14, Met15, Ser16, Ser19, Tyr23 and possibly Phe25) to the side chains of the inter-domain interaction site (i.e. Ile28, Thr29, Asn30, Ala31 and Gln33) via the backbone of the b-strand b2 and the side chain Asn26 (Figure 3 A). Of particular interest for the following comparison is thereby the relative side-chain arrangements of Thr29 and Ala31 side chains (Figures 3 A and D), which can be illustrated by a seesaw sketch representing the two states (Figures 3 D and F): if the side chain of Thr29 is „below“ the backbone in Figure 3 D, the side chain of Ala31 is „above“ it and vice versa if the side chain of Thr29 is „above“ the backbone, the side chain of Ala31 is „below“ it, respectively.

The Mechanism of Action of the Positive Allosteric Ligand Shifts the Population of States

Inspection of the two-state structure of the WW domain in complex with the positive allosteric ligand FFpSPR shows that the two-state ensemble of the WW domain only shifts the population of the two states towards the dark blue state identified in the apo form (Figures 3 A and D). This includes the seesaw arrangement (Figure 3 E) between the two side chains of Thr29 and Ala31 as illustrated in Figure 3 D for which the yellow state of the complex superimposes well with the cyan state of the apo WW, while the red state of the complex superimposes well with the blue state of the apo WW domain. These arrangements are also found in the Ramachandran plots of Thr29 and Ala31 (Figure S6). Thus, the FFpSPR peptide appears to select the dark blue state such that the mode of allosteric action is proposed to be conformational selection.[3]

The Negative Allosteric Ligand Follows the Dynamic Allostery Model

In striking contrast to the peptide FFpSPR, peptide pCdc25C influences the inter-domain interaction between the WW domain and its catalytic domain negatively. Based on the finding that the peptide FFpSPR acts by the conformational selection model discussed above,it would seem logical to

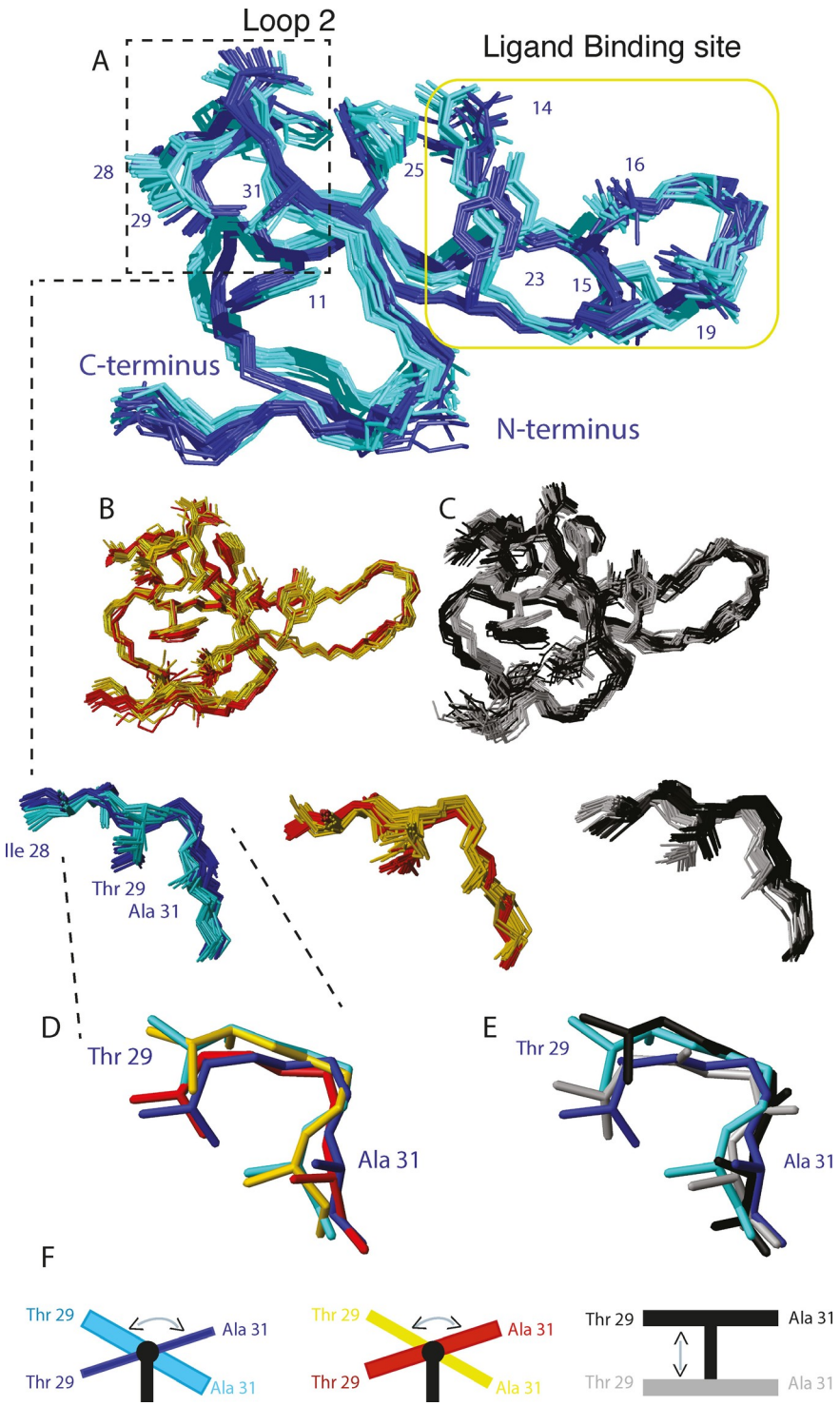


Figure 3. Two-state structural ensemble of apo WW (A), WW in complex with the positive (B) and negative (C) allosteric ligand, highlighting the presence of two distinct states. Backbone traces of 20 structural ensembles of the WW domain each representing two different states are shown. In addition, several side chains are shown and labeled. The WW states are color coded with cyan and blue for the apo WW, yellow and red for the positive allosteric ligand FFpSPR complex, and grey and black for the negative allosteric ligand pCdc25C complex. The two states of the catalytic-domain-interacting Loop 2 are enlarged as indicated. In addition, the two states for residues 29–31 are shown for one conformer only for clarity in (D) and (E) by superimposing the two apo states with either the two states of the positive (D) or the two states of the negative (E) allosteric-ligand–WW complex. These superpositions illustrate that the positive allosteric ligand does not alter the two-state structures as the yellow state superimposes with the cyan state and the red state with the blue state for both residues Thr29 and Ala31. In the case of the negative allosteric ligand, the grey state superimposes with the blue state for residue Thr29, but for Ala31, the blue state superimposes with the black state, and the black state superimposes with the cyan state for Thr29, while the black state superimposes with blue state of Ala31. In (F) this observed change is illustrated by a seesaw model. In the case of apo WW, Thr29 and Ala31 alter their states like a seesaw with the cyan state more populated (drawn thicker). When bound to the positive allosteric ligand, the seesaw states are preserved, however with different populations of the two states. In contrast, no seesaw-like two states are observed in the case of the negative allosteric ligand, shown in grey/black.

assume that the pCdc25C peptide selects the binding-incompetent state and thereby interferes with the inter-domain interaction. However, this is not the case as revealed by the two-state structure calculation of the WW domain in presence of the peptide ligand pCdc25C. Still, two states are observed (Figure 2 C). Furthermore, both states are distinct in the backbone (represented by Ramachandran plots in Figure S6) as well as the side chains of Loop 2 (Figure 3 C and D). However, the backbone and side-chain states in Loop 2 between residues Ile28/Thr29 versus Ala31 are anti-correlated with each other when compared with the apo structure and the FFpSPR structure (Figure 3 D). Hence, the two states of the seesaw of the Thr29/Ala31 side chains do not exist anymore but either both side chains are up or down simultaneously as illustrated in Figure 3 D and E. In detail, when the side chain of Thr29 is close to the dark blue state of the apo structure, the side chain of Ala31 is superimposable with the cyan state, while if the side chain of Ala31 is close to the cyan apo structure, the side chain of Thr29 aligns with the dark blue state of the apo structure. Thus, when Loop 2 is locally averaged over the two states the apo structure is locally not distinguishable from the WW domain structure in complex with pCdc25C as supported by minor observed chemical shift changes contrasting the situation for the positive allosteric case (Figure 1). However, when viewed in a time-resolved manner the free WW domain and in complex behave differently because the dynamics altered. The pCdc25C peptide—WW domain complex can thus be regarded as an example of a dynamic allostery model, where the local structure is not perturbed on average, but on the global level there is a change from a correlated motion in the apo-state to an anticorrelated motion in the complex (Figure 3 D and E).

The Ligand-Induced Allostery of the WW Domain in the Context of Full-Length Pin1

It has been demonstrated that by binding to Loop 1 of the WW domain the peptide pCdc25C reduces the loose inter-domain contact with the C-terminal catalytic PPIase domain via an allosteric mechanism,[33] while the peptide FFpSPR enhances this inter-domain contact.[35, 37, 38] Since Loop 2 is the side of the WW domain which interacts with the PPIase domain, the above findings on the two states on Loop 2 (Figure 3) give insights into the mechanism of positive and negative influence on the interaction between the two domains. For this the WW domain structures of all three systems studied here were superimposed with the WW domain of the crystal structure of full-length Pin 1 (Figure 4; 1PIN.pdb). The superpositions show that the cyan and yellow states of the apo WW and WW domain in complex with FFpSPR clash in part with the catalytic PPIase domain in contrast to the corresponding blue and red states, respectively. Since the predicted clash is not possible because of very large van der Waals energies, it is assumed that the cyan and yellow states are not fit in binding with the catalytic PPIase domain. Indeed, upon FFpSPR peptide ligand binding the population changes in favor of the red non-clashing state and thus the inter-domain interaction is enhanced as demonstrated.[35, 37, 38] In the case of the WW domain bound with pCdc25C both states clash into the catalytic domain, which is interpreted that both states are not able to interact well with the catalytic domain as demonstrated.[35, 37, 38]

The mechanism of the allosteric coupling between ligand-binding Loop 1 and inter-domain interacting Loop 2 is further illustrated in Figure 5 (clay colored structure). The WW domain possesses two distinct states likely of similar energy that cover both the ligand binding site (i.e. Loop 1) as well as the inter-domain interacting site (i.e. Loop 2). The interchange between the catalytic domain-binding competent and non-competent states of Loop 2 in the micro-second time range of

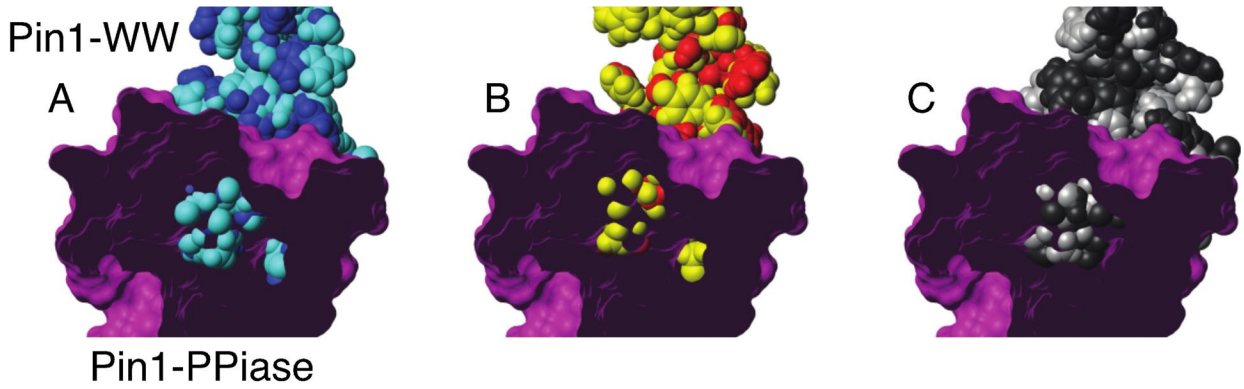


Figure 4. The multi-state structures of the WW domain in the context of its predicted interaction with the catalytic domain highlight state-specific clashes. The two-state structures are shown by space filling calotte (CPK) models of A) the apo WW in cyan and blue, B) the WW domain in complex with the positive allosteric ligand FFpSPR (shown in yellow and red), and C) the WW domain in complex with the negative allosteric peptide pCdc25C (shown in grey and black), which have been superimposed onto the WW domain structure of full-length Pin1 (PDB code 1PIN). The contact surface of the catalytic domain Pin1-PPIase is shown in pink, cut from the front in order to illustrate the eventual clashing of the WW domain with the Pin1-PPIase domain. Note, there is only a clash if the center of the spheres of the WW domain are visible inside the PPIase. The inspection of the figures shows that in (A) the cyan state clashes with the pink PPIase domain, while the blue state does not; in (B) the yellow state clashes with the PPIase domain, while the red state does not, and in (C) both the black and grey states clash with the PPIase domain.

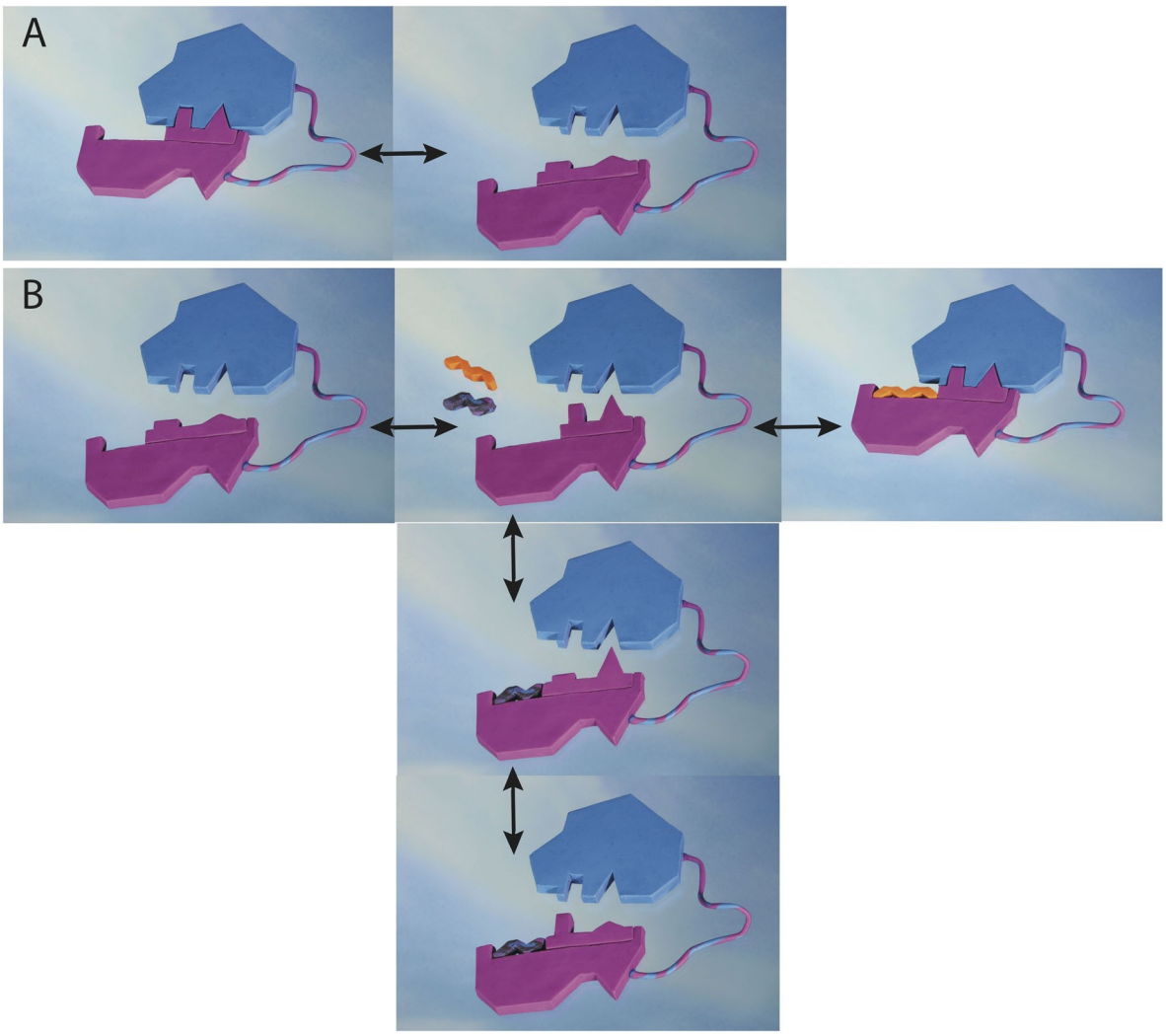


Figure 5.Allosteric mechanisms of action of the WW domain. A) The apo form of the WW domain (represented by a pink clay form) is undergoing exchange between two states, one of which (on the left) is able to bind the catalytic domain colored in blue. B) In presence of either the positive allosteric ligand FFpSPR (shown in orange) or the negative allosteric peptide pCdc25C (shown in black) two distinct allosteric mechanisms are active. The positive allosteric ligand FFpSPR selects the state that interacts with the catalytic domain, enhancing interaction with the catalytic domain (arrow to the right). This proposed mechanism is thus based on the population shift model. The negative allosteric peptide pCdc25C acts via the dynamic allostery model, where the average local structure is not perturbed, but at any given time it is incompatible with interacting with the catalytic site.

the apo WW domain thereby perturbs the inter-domain contact. The ligand FFpSPR selects the binding-competent state and enhances the inter-domain interaction. In contrast, in complex with pCdc25C both Loop 2 conformations show clashes interpreted here as interference of a domain–domain interaction (Figure 5). While the average structure of Loop 2 did not get perturbed upon pCdc25C binding (in Figure 5: the two triangles and two rectangles are on average the same as in the apo state) it is the dynamics that changed from being concerted in the apo state to anti-correlated in the complex with pCdc25C yielding binding interference to the catalytic domain. Hence, depending on the peptide not only the outcome of allostery but also the mechanism of allostery is altered. This is possible since all the structural states involved have similar energies with low activation barriers between them enabling different processesand pathways by small perturbations.

Conclusion

There are several mechanisms of action of allostery including the population shift model and the dynamic allostery model. Because of recent advances in NMR methodology a large collection of highly accurate experimental data was obtained and allowed to elucidate the mechanism of allostery for the WW domain at atomic resolution. A ligand-dependent mechanism of action of allostery was thereby revealed, inferring for one ligand the population shift model and for the other ligand the dynamic allostery model. These mechanisms of action highlight also the possible multi-dimensional interplay between dynamics and structure that amount to evolutionary selection for fittest performance. It further indicates the astonishing multifaceted possibilities this multi-dimensional dynamic structure landscape possesses.

Conflict of interest

The authors declare no conflict of interest.

Stichwçrter: allostery ·ligand binding · NMR spectroscopy · protein dynamics ·protein structure

[1] H. N. Motlagh, J. O. Wrabl, J. Li, V. J. Hilser, Nature 2014, 508, 331 – 339.
[2] D. E. Koshland, G. Nemethy, D. Filmer, Biochemistry 1966, 5, 365 – 385.
[3] J. Monod, J. P. Changeux, F. Jacob, J. Mol. Biol. 1963, 6, 306 – 329.
[4] A. Cooper, D. T. F. Dryden, Eur. Biophys. J. Biophys. Lett. 1984, 11, 103 – 109.
[5] L. E. Kay, J. Magn. Reson. 2005, 173, 193 – 207.
[6] L. E. Kay, D. A. Torchia, A. Bax, Biochemistry 1989, 28, 8972 – 8979.
[7] W. C. Swope, J. W. Pitera, F. Suits, M. Pitman, M. Eleftheriou, B. G. Fitch, R. S. Germain, A. Rayshubski, T. J. C. Ward, Y. Zhestkov, R. Zhou, J. Phys. Chem. B 2004, 108, 6582 – 6594.
[8] F. No, D. Krachtus, J. C. Smith, S. Fischer, J. Chem. Theory Comput. 2006, 2, 840 – 857.
[9] D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. B. Shan, W. Wriggers, Science 2010, 330, 341 – 346.
[10] S. Olsson, F. Noe, J. Am. Chem. Soc. 2017, 139, 200 – 210.
[11] S. Olsson, H. Wu, F. Paul, C. Clementi, F. Noe, Proc. Natl. Acad. Sci. USA 2017, 114, 8265 – 8270.
[12] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, Nat. Struct. Biol. 1997, 4, 292 – 297.
[13] L. Yao, B. Vçgeli, D. A. Torchia, A. Bax, J. Phys. Chem. B 2008, 112, 6045 – 6056.
[14] P. Vallurupalli, D. F. Hansen, L. E. Kay, Proc. Natl. Acad. Sci. USA 2008, 105, 11766 – 11771.
[15] P. Neudecker, A. Zarrine-Afsar, A. R. Davidson, L. E. Kay, Proc. Natl. Acad. Sci. USA 2007, 104, 15717 – 15722.
[16] E. J. Fuentes, S. A. Gilmore, R. V. Mauldin, A. L. Lee, J. Mol. Biol. 2008, 375, 1489 – 1490.
[17] E. J. Fuentes, C. J. Der, A. L. Lee, J. Mol. Biol. 2004, 335, 1105 – 1115.
[18] A. Dhulesia, J. Gsponer, M. Vendruscolo, J. Am. Chem. Soc. 2008, 130, 8931 – 8939.
[19] G. M. Clore, C. D. Schwieters, J. Am. Chem. Soc. 2004, 126, 2923 – 2938.
[20] G. M. Clore, C. D. Schwieters, Biochemistry 2004, 43, 10678 – 10691.
[21] K. Lindorff-Larsen, R. B. Best, M. A. DePristo, C. M. Dobson, M. Vendruscolo, Nature 2005, 433, 128 – 132.
[22] G. Bouvignies, P. Bernado, S. Meier, K. Cho, S. Grzesiek, R. Bruschweiler, M. Blackledge, Proc. Natl. Acad. Sci. USA 2005, 102, 13885 – 13890.
[23] B. Vçgeli, S. Kazemi, P. Gntert, R. Riek, Nat. Struct. Mol. Biol. 2012, 19, 1053 – 1057.
[24] N. A. Lakomek, K. F. A. Walter, C. Fares, O. F. Lange, B. L. de Groot, H. Grubmuller, R. Bruschweiler, A. Munk, S. Becker, J. Meiler, C. Griesinger, J. Biomol. NMR 2008, 41, 139 – 155.
[25] C. N. Chi, D. Strotz, R. Riek, B. Vçgeli, J. Biomol. NMR 2015, 62, 63 – 69.
[26] S. Olsson, D. Strotz, B. Vogeli, R. Riek, A. Cavalli, Structure 2016, 24, 1464 – 1475.
[27] P. J. Nichols, A. Born, M. A. Henen, D. Strotz, J. Orts, S. Olsson, P. Gntert, C. N. Chi, B. Vçgeli, Molecules 2017, 22, 1176.
[28] P. J. Nichols, A. Born, M. A. Henen, D. Strotz, C. N. Celestine, P. Gntert, B. Vçgeli, ChemBioChem 2018, 19, 1695 – 1701.
[29] K. P. Lu, S. D. Hanes, T. Hunter, Nature 1996, 380, 544 – 547.
[30] S. L. Ma, L. Pastorino, X. Z. Zhou, K. P. Lu, J. Biol. Chem. 2012, 287, 6969 – 6973.
[31] Y. S. Lim, H. T. L. Tran, S. J.Park, S. A. Yim, S. B. Hwang, J. Virol. 2011, 85, 8777 – 8788.
[32] K. Lu, Cancer Cell 2003, 4, 175 – 180.
[33] R. Ranganathan, K. P. Lu, T. Hunter, J. P.Noel, Cell 1997, 89, 875 – 886.
[34] E. Bayer, S. Goettsch, J. W. Mueller, B. Griewel, E. Guiberman, L. M. Mayr, P. Bayer, J. Biol. Chem. 2003, 278, 26183 – 26193.
[35] K. A. Wilson, J. J. Bouchard, J. W. Peng, Biochemistry 2013, 52, 6968 – 6981.
[36] J. J. Guo, X. D. Pang, H. X. Zhou, Structure 2015, 23, 237 – 247.
[37] D. M. Jacobs, K. Saxena, M. Vogtherr, P. Bernado, M. Pons, K. M. Fiebig, J. Biol. Chem. 2003, 278, 26174 – 26182.
[38] J. W. Peng, Biophys. Rev. 2015, 7, 239 – 249.
[39] T. Peng, J. S. Zintsmaster, A. T. Namanja, J. W. Peng, Nat. Struct. Mol. Biol. 2007, 14, 325 – 331.
[40] J. W. Peng, B. D. Wilson, A. T. Namanja, J. Biomol. NMR 2009, 45, 171 – 183.
[41] F. Morcos, S. Chatterjee, C. L. McClendon, P. R. Brenner, R. Lpez-Rendn, J. Zintsmaster, M. Ercsey-Ravasz, C. R. Sweet, M. P. Jacobson, J. W. Peng, J. A. Izaguirre, PLoS Comput. Biol. 2010, 6, e1001015.
[42] R. B. Fenwick, L. Orellana, S. Esteban-Martin, M. Orozco, X. Salvatella, Nat. Commun. 2014, 5, 4070.
[43] J. C. Crane, E. K. Koepf, J. W. Kelly, M. Gruebele, J. Mol. Biol. 2000, 298, 283 – 292.
[44] M. Socolich, S. W. Lockless, W. P. Russ, H. Lee, K. H. Gardner, R. Ranganathan, Nature 2005, 437, 512 – 518.
[45] W. P.Russ, D. M. Lowery, P. Mishra, M. B. Yaffe, R. Ranganathan, Nature 2005, 437, 579 – 583.
[46] J. Orts, B. Vçgeli, R. Riek, J. Chem. Theory Comput. 2012, 8, 3483 – 3492.
[47] D. Strotz, J. Orts, C. N. Chi, R. Riek, B. Vçgeli, J. Chem. Theory Comput. 2017, 13, 4336 – 4346.

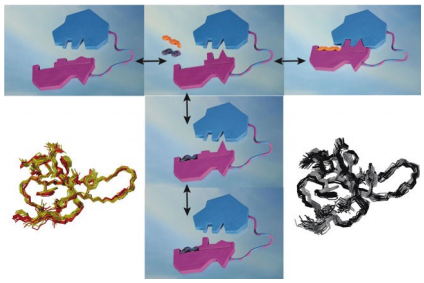
Manuskript erhalten: 22. Juni 2020
Vernderte Fassung erhalten: 23. Juli 2020
Akzeptierte Fassung online: 14. August 2020
Endgltige Fassung online: && . && &&&&

Forschungsartikel

Structural Biology

D. Strotz, J. Orts, H. Kadavath, M. Friedmann, D. Ghosh, S. Olsson, C. N. Chi, A. Pokharna, P. Gntert, B. Vçgeli,* R. Riek* _____ &&&&–&&&&

Protein Allostery at Atomic Resolution



Multi-state structure calculations from exact nuclear Overhauser effect (eNOE) rates are used to elucidate long range, coupled structural rearrangements within the Pin1 enzyme in the absence and presence of two antagonizing allosteric ligands. This enables the elucidation of the allosteric mode of action of these ligands, with one ligand following a population shift mechanism and the other a dynamic allosteric mechanism.

Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Protein Allostery at Atomic Resolution

Dean Strotz, Julien Orts, Harindranath Kadavath, Michael Friedmann, Dhiman Ghosh, Simon Olsson, Celestine N. Chi, Aditya Pokharna, Peter Gntert, Beat Vçgeli,* and Roland Riek*

ange_202008734_sm_miscellaneous_information.pdf

Material and Methods

Preparation of samples

The Pin1-WW construct S18N/W34F was used because it behaves experimentally more favourable and it is more resistant to aggregation than the wild-type. [5-6] Preparation of the $^{15}N/^{13}C$-labeled WW domain variant comprising S18N/W34F was done as follows: The S18N genetic sequence was delivered by GL Biochem Ltd. preassembled in peT32 with a HIS6-tag cleavable by TEV. Transformed E. coli BL21/DE3 cells were grown in pre-cultures started from fresh glycerol stock. In 2L M9 cultures (with either ^{15}N NH_4Cl or ^{15}N NH_4Cl/ ^{13}C glucose) the cells were grown form OD_{600} of 0.1 to 0.7 at 37˚C and shaken at 120 RPM, then induced with IPTG. Following induction, the temperature was reduced to 25˚C and left for expression for 4 hrs before harvesting. Following two-step Ni-column purification the sample was desalted and TEV protease added in 1:50 (m/m) ratio and left over-night at room temperature. A further Ni-column purification step then provided the clean NMR sample. The sample's buffer was exchanged to the NMR buffer (10 mM K_2PO_4, 100 mM NaCl, 0.02 % NaN_3, pH 6.0) using dialysis, then concentrated to 1.2 mM using 2 kDa cutoff concentrator tubes (Sartorius Vivaspin 15R). The gene of full-length Pin1 S18N/W34F variant was bought from genescript, sub-cloned into a pET28a vector containing an N-terminal His-tag with a thrombin cleavage site (MHHHHHHLVPRGS). For expression the cDNA was transformed into E. coli BL21 cells and plated on a kanamycin-containing plate (50 ug/ml). The cells were grown over-night at 37 ˚C and then used to inoculate a 10 ml pre-culture. The culture was grown at 37 ˚C for three hours and thereafter used to inoculate a 1-liter culture (kanamycin 50 ug/ml) of M9 medium for $^{15}N/^{13}C$- or ^{15}N-labeling. Cells were grown to an OD_{600} of 0.9. Protein expression was initiated by adding 1 mM IPTG (isopropyl β-D-thiogalactopyranoside). The cells were then allowed to express over night at 18 ˚C and harvested by spinning at 5,000 g for 15 minutes and re-suspended in purification buffer (10 mM Tris/HCl, 200 mM NaCl). The cells were lyophilized and spun at 40,000 g for 20 minutes. The supernatant was filtered (0.4 μm and 0.2 μm filters) and loaded onto a nickel (II)-charged chelating sepharose FF column (Amersham Biosciences), equilibrated with purification buffer as above and washed with 400 ml of the same buffer. The bound sample was eluted with 250

mM imidazole at pH 7.9, in aliquots of 10 ml. Fractions containing partially pure proteins were pooled, desalted and passed through a DEAE column equilibrated with purification buffer. The sample was collected as flow-through. The purity was checked on SDS PAGE stained with coomassie brilliant blue. The pure protein preparation was concentrated to experimental concentration of 0.4 mM. The concentration was determined by absorption measurements using the molar absorption coefficient calculated.

The phosphorylated ligands pCdc25C (i.e. EQPLpTPVTDL) and FFpSPR were ordered from Bachem AG, Switzerland.

The NMR buffer was 10 mM K_2PO_4, 100 mM NaCl, 0.02 % NaN_3, in 3 % D_2O and pH 6.0 with sample concentrations of 1.2 mM WW domain or 0.4 mM full-length Pin1. For the sample in complex with pCdc25C a 4-fold excess of ligand was used, while for the complex with FFpSPR a 10-fold excess of the ligand was used, respectively. Using the isotherms of the chemical shift perturbations for the ligand-binding site residues (excluding allosteric site residues) and two-state exchange models (chemical shift perturbation versus concentration) the pCdc25C affinity (K_d) to the WW domain was determined to be 526 +/- 146 μM in line with isothermal titration calorimetry (ITC) measurements (Supplementary Figure S2) that yielded a similar affinity (K_d) of 158 +/- 70 μM (overall K_d as opposed to binding site). Correspondingly, the K_d for FFpSPR was determined to be 708 +/- 38 μM. Using the above values the occupancy of the NMR samples were calculated as follows:

fraction $= (Pt + Lt + K_d - ((Pt + Lt + K_d)^2 - 4 * Pt * Lt)^{1/2}) / (2 * Pt)$

Where Lt and Pt are the ligand and protein concentration, respectively, in ∝M and the K_d was the calculated K_d from NMR titrations (∝M).

The binding affinities determined are significantly lower than the reported one for wild-type WW domain (i.e. 43 μM for FFpSPR and 6 μM for pCdC25C) attributed to the mutation.

NMR experiments

All experiments were recorded on a Bruker 700 MHz spectrometer and at 5 ˚C, except where described otherwise. All spectra were processed and analyzed using the software package NMRPipe [7], assignment was done in CcpNMR [8]. The measurement and analysis

of eNOEs using eNORA2 [1] ([2] CYANA version) was described previously in detail [9], in short: series of 3D [^{15}N, ^{13}C]-resolved [^{1}H, ^{1}H]-NOESY-HSQC experiments were recorded to measure NOE buildups.[10] The inter-scan delay was 0.8 s. Simultaneous [^{15}N, ^{1}H]-HSQC and [^{13}C, ^{1}H]-HSQC elements were employed, following indirect proton chemical shift evolution and [^{1}H, ^{1}H]-NOE mixing (τ_m). Diagonal-peak decays and cross-peak buildups were measured with τ_m of 20, 30, 40, 50, and 60 ms for all the three samples.

Titration experiments with the ^{13}C,^{15}N-labeled WW domain were done by two different approaches: (i) For a fixed amount of labeled ^{15}N/^{13}C WW domain, increasing amounts of peptides were titrated into the protein solution and the extent of binding was determined by measuring the chemical shift changes in the [^{15}N, ^{1}H]-HSQC spectrum for several resonances and averaged thereafter. In the second approach T2-filtered [^{15}N, ^{1}H]-HSQC experiments on ^{1}H and ^{15}N with a T2 times of 10 ms and 75 ms were measured that strengthen the line broadening as can be seen by a comparison between Figure 1 and Figure 2SB.

The rotational correlation times τ_c of the individual samples were determined using ^{15}N-relaxation measurements as described previously in detail [9] yielding a τ_c for the apo WW domain to be 4.25 ns at 5 ˚C and 1.2 mM concentration, a τ_c of 5.67 ns for the WW domain in complex with pCdc25C at 5 ˚C and 1.2 mM concentration and a τ_c of 5.13 ns at 5 ˚C and 1.2 mM concentration for the WW domain in complex with FFpSPR, respectively.

$^{3}J_{HN,H\alpha}$ scalar coupling were measured as described previously in detail. [9] $^{3}J_{H\alpha,H\beta}$ scalar couplings were obtained from 3D ^{13}Cα-separated H^{α}-H^{β} in-phase COSY (HACAHB-COSY) experiments [11] in D_2O. The experiment was recorded with 50(MQ[Cα], t_1) × 54(Hβ, t_2) × 2048(Hα, t_3) complex points, giving $t_{1max,\,13C}$ = 22.5 ms, $t_{2max,1H}$ = 10.8 ms, $t_{3max,1H}$ = 204.8 ms. The time domain data were multiplied with a square cosine function in the direct dimension and cosine functions in the indirect dimensions and zero-filled to 256 × 512 × 2048 complex points. The Karplus parameters used in structure calculations were from Hu et al.[12] $^{3}J_{C',C\gamma}$ and $^{3}J_{N,C\gamma}$ scalar couplings for aromatic side chains were obtained from ^{13}C'-{$^{13}C^{\gamma}$} and ^{15}N-{$^{13}C^{\gamma}$} spin-echo difference [^{15}N,^{1}H]-HSQC experiments [13] performed on a Bruker 600 MHz spectrometer. The experiments were recorded with 100(^{15}N, t_1) or 200(^{15}N, t_1) × 512(^{1}HN, t_2) complex points, giving $t_{1max,15N}$ = 50 ms or $t_{1max,15N}$ = 100 ms and $t_{2max,1H}$ = 51.2 ms, respectively. The time domain data were multiplied with a square cosine function in the

direct dimension and cosine functions in the indirect dimensions and zero-filled to 512 × 2048 complex points. The Karplus parameters used in structure calculations were from Hu et al.[12]

Cross-correlated relaxation rates $\Gamma_{HNiNi/H\alpha iC\alpha i} + \Gamma_{H\alpha iNi/HNiC\alpha i}$ were obtained from two experiments performed on a Bruker 600 MHz spectrometer equipped with a z-axis gradient cryogenic probe. A DIAI (double in-phase/anti-phase inter-conversion) method was realized with a pair of 3D HNCA pulse sequences ("reference" and "trans")[14] for the first experiment. A 3D ct-HNCA MMQ (mixed multi-quantum, with zero- and double-quantum coherence evolution averaged) experiment was used for the second experiment. The ZQ (zero quantum) and DQ (double quantum) coherences were superimposed, resulting in four components to be evaluated [15-16]. The experiments were recorded with τ_{MQ} = 31.0 ms or τ_{MQ} = 33.5 ms, 50(MQ[N,Cα], t_1) or 55(MQ[N, Cα], t_1) × 36(N, t_2) × 512(HN, t_3) complex points, t_{1max} = 25.0 or 27.5 ms, $t_{2max,15N}$ = 18.0 ms, $t_{3max,1H}$ = 51.2 ms. The time domain data were multiplied with a square cosine function in the direct dimension and cosine functions in the indirect dimensions and zero-filled to 256 × 128 × 2048 complex points. The back-calculation of the cross-correlated relaxation rates followed the procedure described previously in detail. [17]

The 2D CEST-[^{15}N,^{1}H]-TROSY spectra [18] were recorded on a Bruker 700 MHz spectrometer for both ^{15}N-labeled apo WW. The continuous wave 5% truncated Gaussian pulse used with a length of 75 ms showed an excitation profile of 8.5 Hz. The saturation was obtained by 6 such pulses. The experiments were recorded with 32(^{15}N, t_1) × 512(^{1}HN, t_2) complex points, giving $t_{1max,15N}$ = 16 ms and $t_{2max,1H}$ = 51.2 ms. Along the ^{15}N-CEST dimension 90 saturations were measured around the Ala31 and Gln33 resonances in steps of 8 Hz.

For the evaluation of the exchange rate the following ^{15}N relaxation measurements were undertaken: Longitudinal relaxation rate, R1, rotating-frame relaxation rate, R1ρ and {^{1}H}^{15}N NOE rate were determined using pseudo 3D TROSY-based experiments at 600 MHz as well as 900 MHz spectrometer. The recorded data was analysed in NMRPipe and was fed to the RELAX software suite [19], to compute parameters like S2, R_{ex}, τ_c and τe. The equilibrium rate constant k_{ex}, was extracted using acquired Rex, population of the two states pA and pB and their chemical shift difference $\Delta\omega$ using the equation $R_{ex} = p_A p_B \Delta\omega^2 / k_{ex}$ valid in the fast k_{ex} exchange regime. Furthermore, multiple consistency checks in-built in RELAX were performed using the schemes proposed. [20]

Structure calculation

The structure calculation followed the established ensemble-based protocol [21-22] using the software packages eNORA2 [1] ([2] CYANA version) and CYANA. [23-24] As input for the structure calculation we used upper and lower distance restraints from eNOEs together with backbone, Hβ and aromatic side-chain scalar couplings and conservative Φ and Ψ dihedral angle restraints derived from $^{13}C^{\alpha}$ chemical shifts (Tables S1-S3). [25] The weight of the dihedral angle restraints was reduced to zero in the final steps of the structure calculation. Calculations were done with 50'000 torsion angle dynamics steps for 100 conformers with random torsion angles by simulated annealing. The multi-states structural ensembles were each calculated simultaneously and averaged. A weak harmonic well potential with bottom width of 1.2 Å was used to keep identical heavy atoms from the different states together. [10, 21] The 20 conformers with the lowest final target function values were selected and analyzed. The calculated coordinates and the complete data set consisting of the eNOEs together with the upper/lower distance limit tables are deposited in the PDB under 6SVC, 6SVE and 6SVH.

Table S1: Structural statistics and CYANA input data for the apo WW domain

Structural input and statistics

Table S1: Structural statistics and cyana input data for apo Pin1-WW.

NMR distance and dihedral constraints			
Distance constraints			
Total eNOEs	686		
eNOEs from one pathway	415 ($NORM_{orig}$ and $NORM_{dest}$)		
eNOEs from two pathways	271		
Intra-residue, $\mid i-j \mid = 0$	255		
Sequential, $\mid i-j \mid = 1$	168		
Short-range, $\mid i-j \mid <= 1$	423		
Medium-range, $1 < \mid i-j \mid < 5$	74		
Long-range, $\mid i-j \mid >= 5$	189		
Dihedral angle restraints			
3JHNα scalar couplings	26		
3JHαHβ scalar couplings	24		
3JHNCG scalar couplings (aromatic)	6		
3JHNCOCG scalar couplings (aromatic)	6		
13Cα chemical shifts	8		
	One-state ensemble	Two-states ensemble	
Structure statistics			
Average CYANA target function value ($Å_2$)	20.38 ± 0.07	7.22 ± 0.16	
Violations			
Distance constraints (> 0.5Å)	14	0	
Dihedral angle constraints (> 5°)	0	0	
Deviations from idealized geometry			
RMSD (Å)			
Backbone to mean	0.02 ± 0.02	0.43 ± 0.05	
Heavy atoms to mean	0.46 ± 0.04	0.88 ± 0.06	
		1st state	2nd state
Backbone to mean		0.30 ± 0.13	0.30 ± 0.14
Heavy atoms to mean		0.76 ± 0.14	0.80 ± 0.13
RMSD to X-ray structure (Å)			
Backbone			
1Pin.pdb	0.69	0.98	0.88
2ZQT.pdb (M130A)	0.63	0.90	0.81
Heavy atoms			
1Pin.pdb	1.50	1.49	1.43
2ZQT.pdb (M130A)	1.36	1.37	1.32

Table S2: Structural statistics and CYANA input data for the WW domain in complex with pCdc25C

Table S2: Structural statistics and Cyana input data for the Pin1-WW-pCdc25C complex.

NMR distance and dihedral constraints			
Distance constraints			
Total eNOEs	711		
eNOEs from one pathway	450 ($NORM_{orig}$ and $NORM_{dest}$)		
eNOEs from two pathways	261		
Intra-residue, $\mid i-j \mid = 0$	258		
Sequential, $\mid i-j \mid = 1$	169		
Short-range, $\mid i-j \mid <= 1$	427		
Medium-range, $1 < \mid i-j \mid < 5$	91		
Long-range, $\mid i-j \mid >= 5$	193		
Dihedral angle restraints			
^3JHNα scalar couplings	24		
^3JHαHβ scalar couplings	23		
^3JHNCG scalar couplings (aromatic)	6		
^3JHNCOCG scalar couplings (aromatic)	6		
13Cα chemical shifts	6		
	One-state ensemble	Two-states ensemble	
Structure statistics			
Average CYANA target function value (Å^2)	18.05 ± 0.04	8.35 ± 0.25	
Violations			
Distance constraints (> 0.5Å)	8	0	
Dihedral angle constraints (> 5°)	0	0	
Deviations from idealized geometry			
RMSD (Å)			
Backbone to mean	0.09 ± 0.04	0.51 ± 0.10	
Heavy atoms to mean	0.59 ± 0.09	0.95 ± 0.09	
		1st state	2nd state
Backbone to mean		0.35 ± 0.06	0.40 ± 0.09
Heavy atoms to mean		0.84 ± 0.08	0.86 ± 0.09
RMSD to X-ray structure (Å)			
Backbone			
1Pin.pdb	0.55	0.76	0.73
2ZQT.pdb (M130A)	0.51	0.72	0.69
Heavy atoms			
1Pin.pdb	1.59	1.29	1.29
2ZQT.pdb (M130A)	1.54	1.30	1.29

Table S3: Structural statistics and CYANA input data for the WW domain in complex with FFpSPR

Table S3: Structural statistics and Cyana input data for the Pin1-WW-FFpSPR complex.

NMR distance and dihedral constraints			
Distance constraints			
Total eNOEs	760		
eNOEs from one pathway	473 ($NORM_{orig}$ and $NORM_{dest}$)		
eNOEs from two pathways	287		
Intra-residue, $\mid i-j \mid = 0$	257		
Sequential, $\mid i-j \mid = 1$	192		
Short-range, $\mid i-j \mid <= 1$	449		
Medium-range, $1 < \mid i-j \mid < 5$	93		
Long-range, $\mid i-j \mid >= 5$	218		
Dihedral angle restraints			
^3JHNα scalar couplings	26		
^3JHαHβ scalar couplings	19		
^3JHNCG scalar couplings (aromatic)	6		
^3JHNCOCG scalar couplings (aromatic)	6		
13Cα chemical shifts	8		
	One-state ensemble	Two-states ensemble (pop 1:4)	
Structure statistics			
Average CYANA target function value (Å^2)	16.78 ± 0.04	9.25 ± 0.74	
Violations			
Distance constraints (> 0.5Å)	7	0	
Dihedral angle constraints (> 5°)	0	0	
Deviations from idealized geometry			
RMSD (Å)			
Backbone to mean	0.09 ± 0.07	0.33 ± 0.08	
Heavy atoms to mean	0.49 ± 0.11	0.77 ± 0.08	
		state (pop 1)	state (pop 4)
Backbone to mean		0.31 ± 0.09	0.13 ± 0.04
Heavy atoms to mean		0.72 ± 0.09	0.58 ± 0.10
RMSD to X-ray structure (Å)			
Backbone			
1Pin.pdb	0.55	0.73	0.56
2ZQT.pdb (M130A)	0.51	0.68	0.51
Heavy atoms			
1Pin.pdb	1.25	1.35	1.25
2ZQT.pdb (M130A)	1.23	1.33	1.23

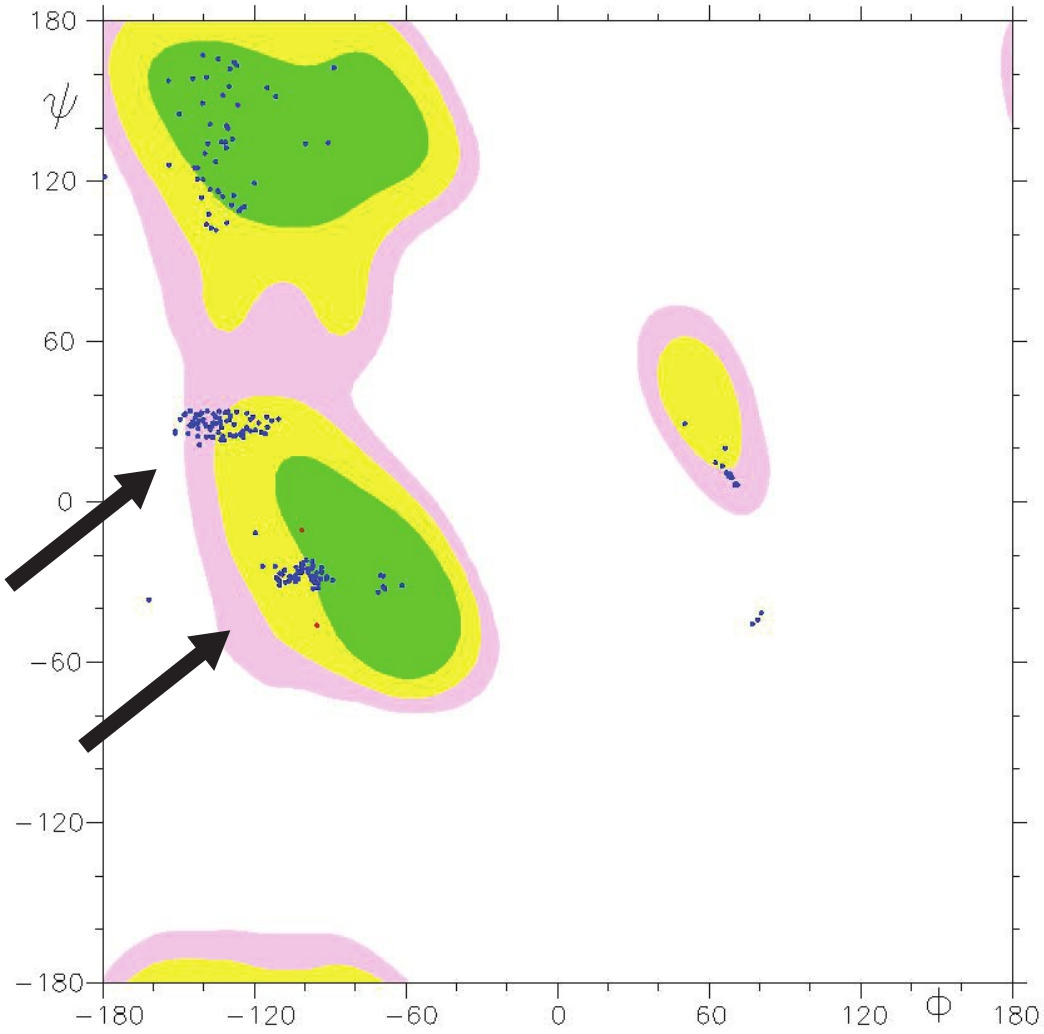


Figure S1: The two states are preserved also for a ten-state structure calculation of apo WW as exemplified by the Ramachandran plot of Thr29. The Ramachandran plot of Thr29 of all the 20 conformers of the ten-state structure calculation (i.e. 200 conformers in total) is shown. While there are outliers, the two states of interest (highlighted by arrows; compare also with Figure 3) are still present.

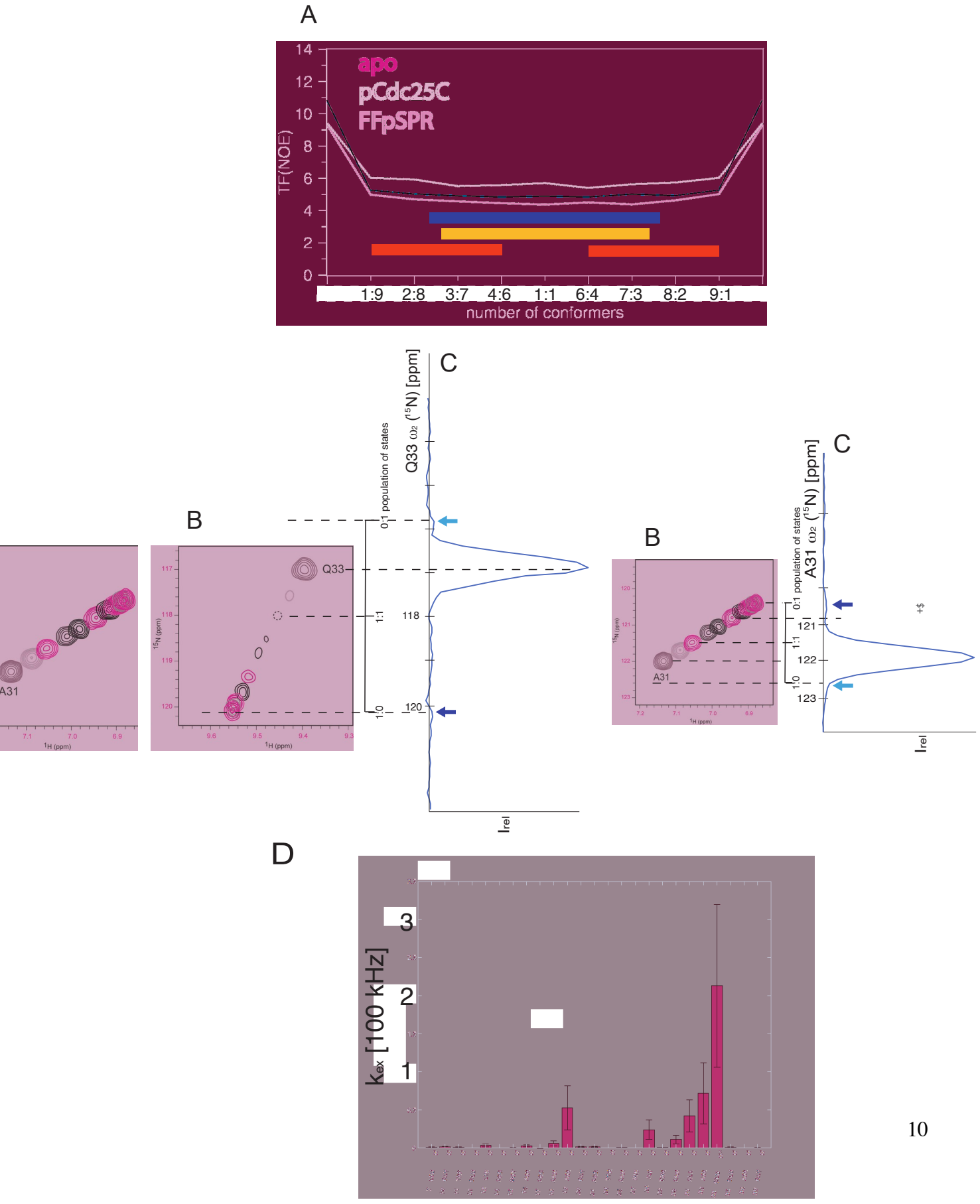
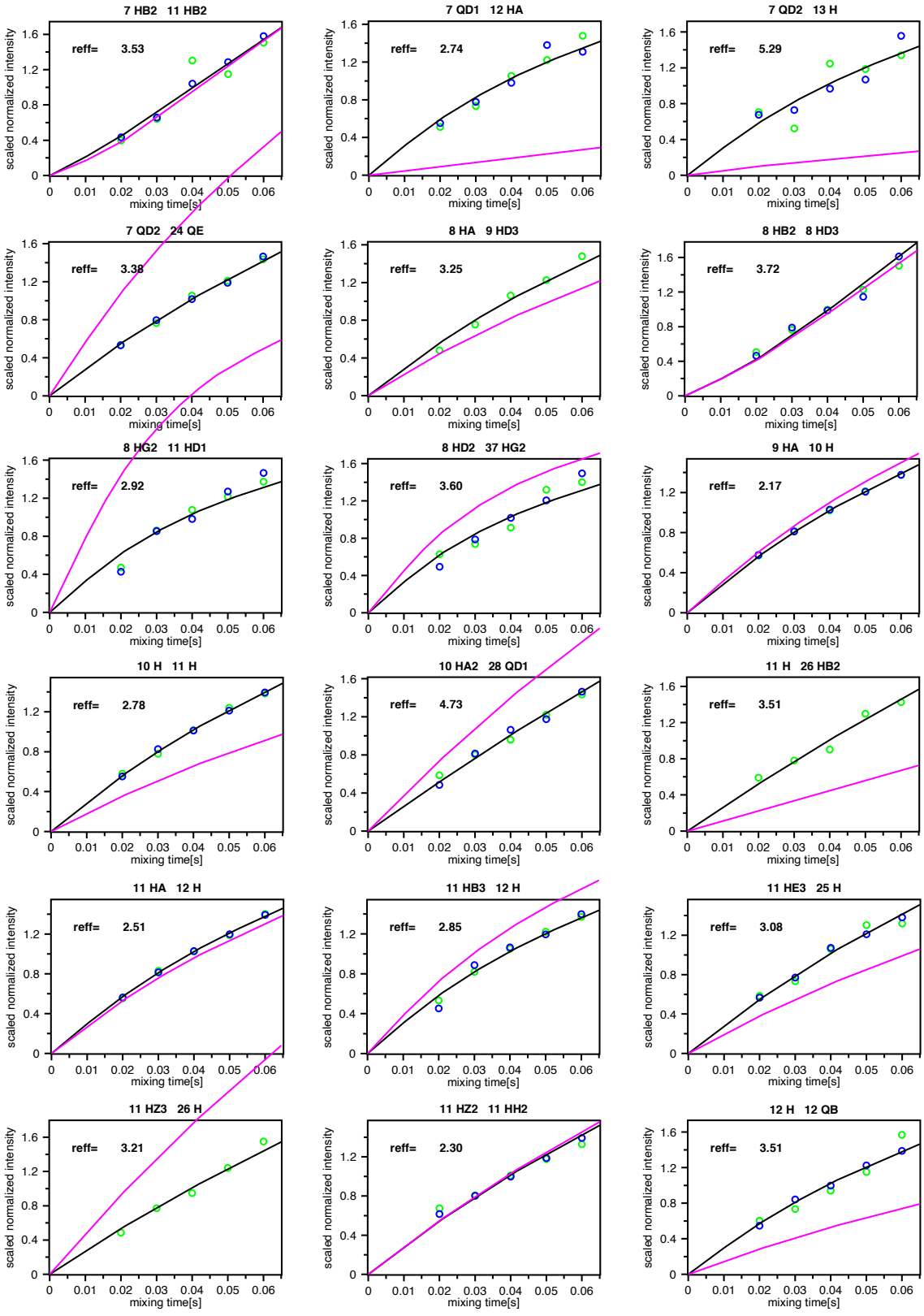
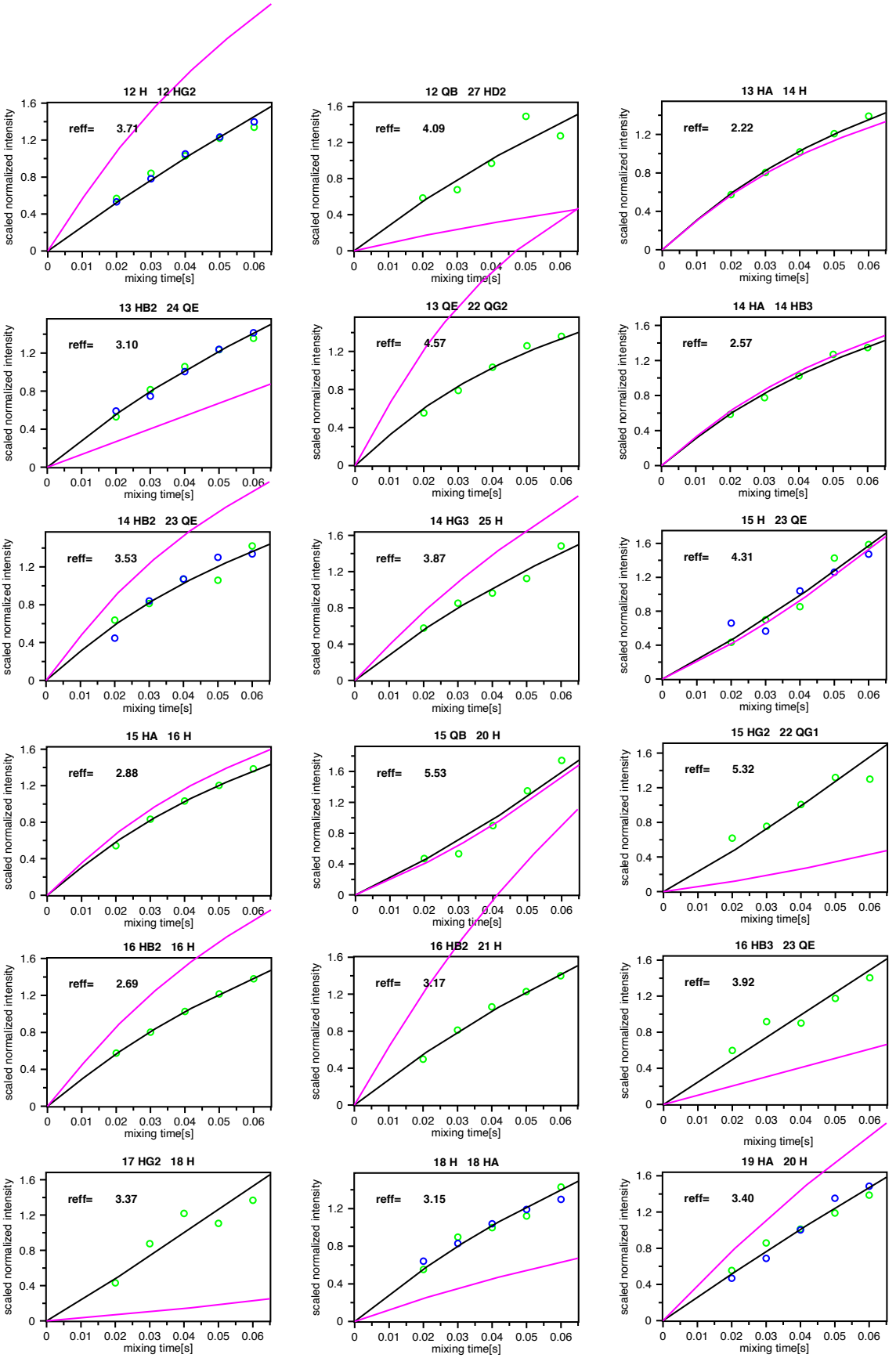
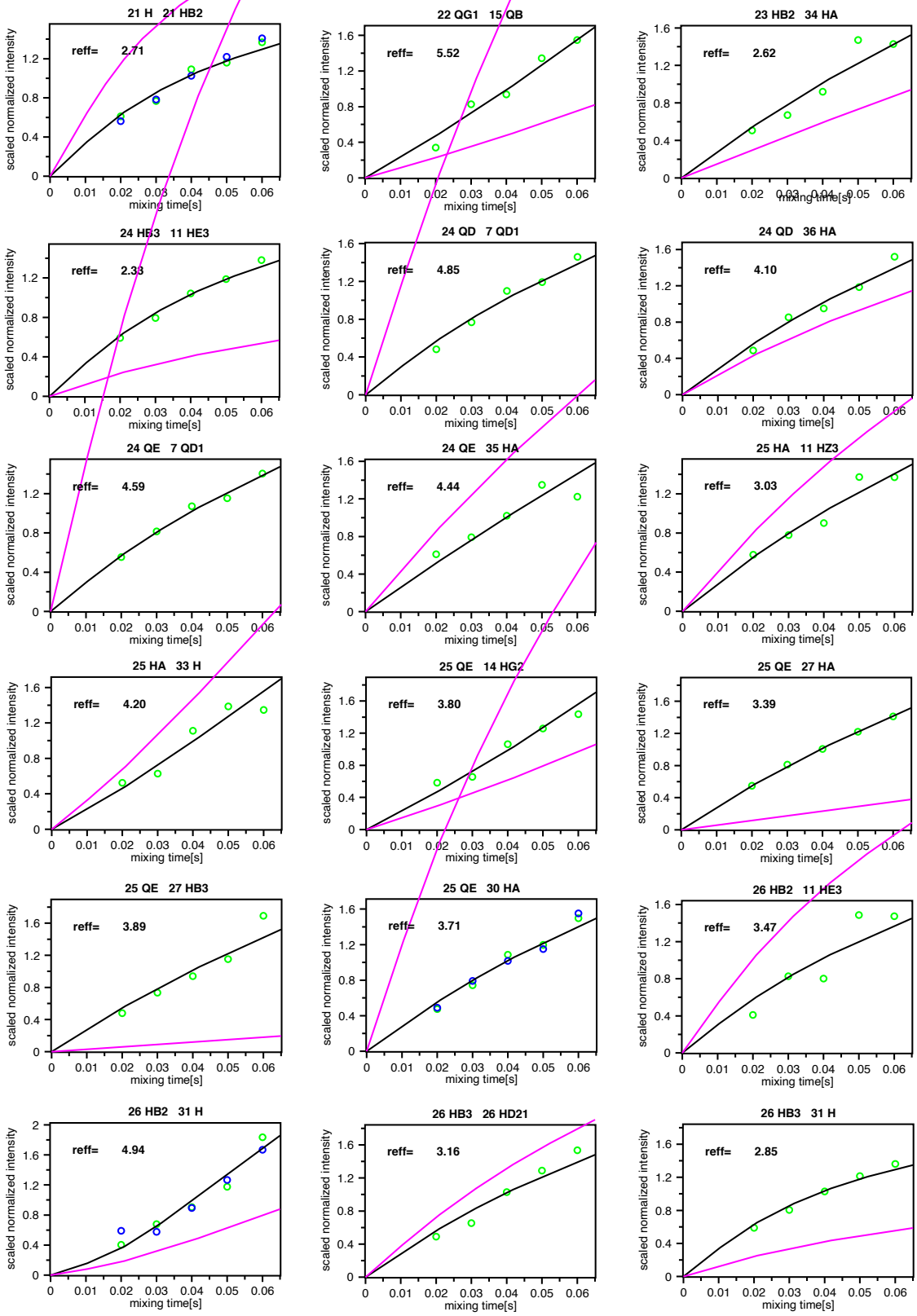

A
apo
pCdc25C
FFpSPR
TF(NOE)
14
12
10
8
6
4
2
0
1:9
2:8
3:7
4:6
1:1
6:4
7:3
8:2
9:1
number of conformers
B
C
Q33
A31
Q33 ω2 (15N) [ppm]
A31 ω2 (15N) [ppm]
0:1 population of states
1:1
1:0
118
120
121
122
123
Irel
1H (ppm)
15N (ppm)
D
kex [100 kHz]
3
2
1

Figure S2: Population determination of the states of apo WW domain. (A) shows the CYANA target function (TF) of the two-state structure calculations versus various populations. For this a pseudo ten-state structure calculation was set up allowing only two distinct states with various populations between 1:9 to 9:1 through symmetry restraints. These calculations differ thus from the multi-state structure calculations performed in Fig. 2 explaining the different TF values. From the Figure it is evident that the TF cannot determine the populations between 1:9 and 9:1. The bars below show the area of populations for which the two-state structures including the correlation between Thr29 and Ala31 discussed in detail in the main text are conserved. In the case of the apo WW domain the same two states are present between 1:3 - 1:1, while in the case of the WW domain in complex with FFpSPR the same two states are obtained in the population range between 1:9 and 4:6. The color code used for the data is blue for apo WW, yellow for WW in complex with pCdc25C and red for WW in complex with FFpSPR. (B) Population determination via a WW titration experiment using the FFpSPR peptide using a T2-filtered [^{15}N,^{1}H]-HSQC experiment, which enhances the line broadening as can be seen by comparing the spectra in figure b with Figure 1. The decrease and increase of signal intensity during the titration are determined by the relative populations between the two states that interchange in the fast/intermediate time regime. The weakest signal is observed when the two states are equally populated (i.e. 1:1; yellow cross peak for Ala31, while for Gln33 the peak is very weak and its position is indicated by a dashed circle). In concert with the knowledge of the chemical shift population of the titration end point, the cross peak at population of 1:1 allows the determination of the chemical shift of the other state. With the knowledge of the chemical shift of apo WW in absence of ligand, the population of the two states can be determined and is in the order of 1:3 as indicated. (C) ^{15}N CEST NMR of Ala31 and Gln33 for the apo WW domain. As indicated by a blue arrow, the apo WW domain shows a resonance that corresponds to the ^{15}N frequency when fully occupied with the FFpSPR ligand attributed to the chemical shift resonance of one of the allosteric states, while the cyan arrow indicates a resonance that is in line with the other state identified by the titration experiment described in (B). These data again indicate a population of about 1:3 between the two states. While the signal to noise ratio of the ^{15}N CEST NMR is rather good as can be seen by the flat baseline, the saturation-derived signals are rather weak (i.e. the signals indicated by arrows), which may be attributable to the possibly rather fast exchange rate for CEST (see below). In addition, in the case of Gln33 between the main resonance and the blue-indicated state (i.e. between 118 and 119.5 ppm) there appear to be other states that lie between the two extreme states identified indicating a continuous sampling between the two states. (D) The exchange rate for individual amide moieties was determined using ^{15}N relaxation measurements (i.e. T1, T2, and {^{1}H}^{15}N -NOEs) at two fields (i.e. 600 MHz and 900 MHz) and the chemical shift differences derived from the titration experiments (B) yielding an exchange of the two states in the 100 kHz regime. The exchange rates in addition with the analysed output data are also listed in Table S4.


7 HB2 11 HB2
reff= 3.53
7 QD1 12 HA
reff= 2.74
7 QD2 13 H
reff= 5.29
7 QD2 24 QE
reff= 3.38
8 HA 9 HD3
reff= 3.25
8 HB2 8 HD3
reff= 3.72
8 HG2 11 HD1
reff= 2.92
8 HD2 37 HG2
reff= 3.60
9 HA 10 H
reff= 2.17
10 H 11 H
reff= 2.78
10 HA2 28 QD1
reff= 4.73
11 H 26 HB2
reff= 3.51
11 HA 12 H
reff= 2.51
11 HB3 12 H
reff= 2.85
11 HE3 25 H
reff= 3.08
11 HZ3 26 H
reff= 3.21
11 HZ2 11 HH2
reff= 2.30
12 H 12 QB
reff= 3.51
scaled normalized intensity
mixing time[s]


12 H 12 HG2
reff= 3.71
12 QB 27 HD2
reff= 4.09
13 HA 14 H
reff= 2.22
13 HB2 24 QE
reff= 3.10
13 QE 22 QG2
reff= 4.57
14 HA 14 HB3
reff= 2.57
14 HB2 23 QE
reff= 3.53
14 HG3 25 H
reff= 3.87
15 H 23 QE
reff= 4.31
15 HA 16 H
reff= 2.88
15 QB 20 H
reff= 5.53
15 HG2 22 QG1
reff= 5.32
16 HB2 16 H
reff= 2.69
16 HB2 21 H
reff= 3.17
16 HB3 23 QE
reff= 3.92
17 HG2 18 H
reff= 3.37
18 H 18 HA
reff= 3.15
19 HA 20 H
reff= 3.40
scaled normalized intensity
mixing time[s]


21 H 21 HB2
reff= 2.71
22 QG1 15 QB
reff= 5.52
23 HB2 34 HA
reff= 2.62
24 HB3 11 HE3
reff= 2.33
24 QD 7 QD1
reff= 4.85
24 QD 36 HA
reff= 4.10
24 QE 7 QD1
reff= 4.59
24 QE 35 HA
reff= 4.44
25 HA 11 HZ3
reff= 3.03
25 HA 33 H
reff= 4.20
25 QE 14 HG2
reff= 3.80
25 QE 27 HA
reff= 3.39
25 QE 27 HB3
reff= 3.89
25 QE 30 HA
reff= 3.71
26 HB2 11 HE3
reff= 3.47
26 HB2 31 H
reff= 4.94
26 HB3 26 HD21
reff= 3.16
26 HB3 31 H
reff= 2.85
scaled normalized intensity
mixing time[s]

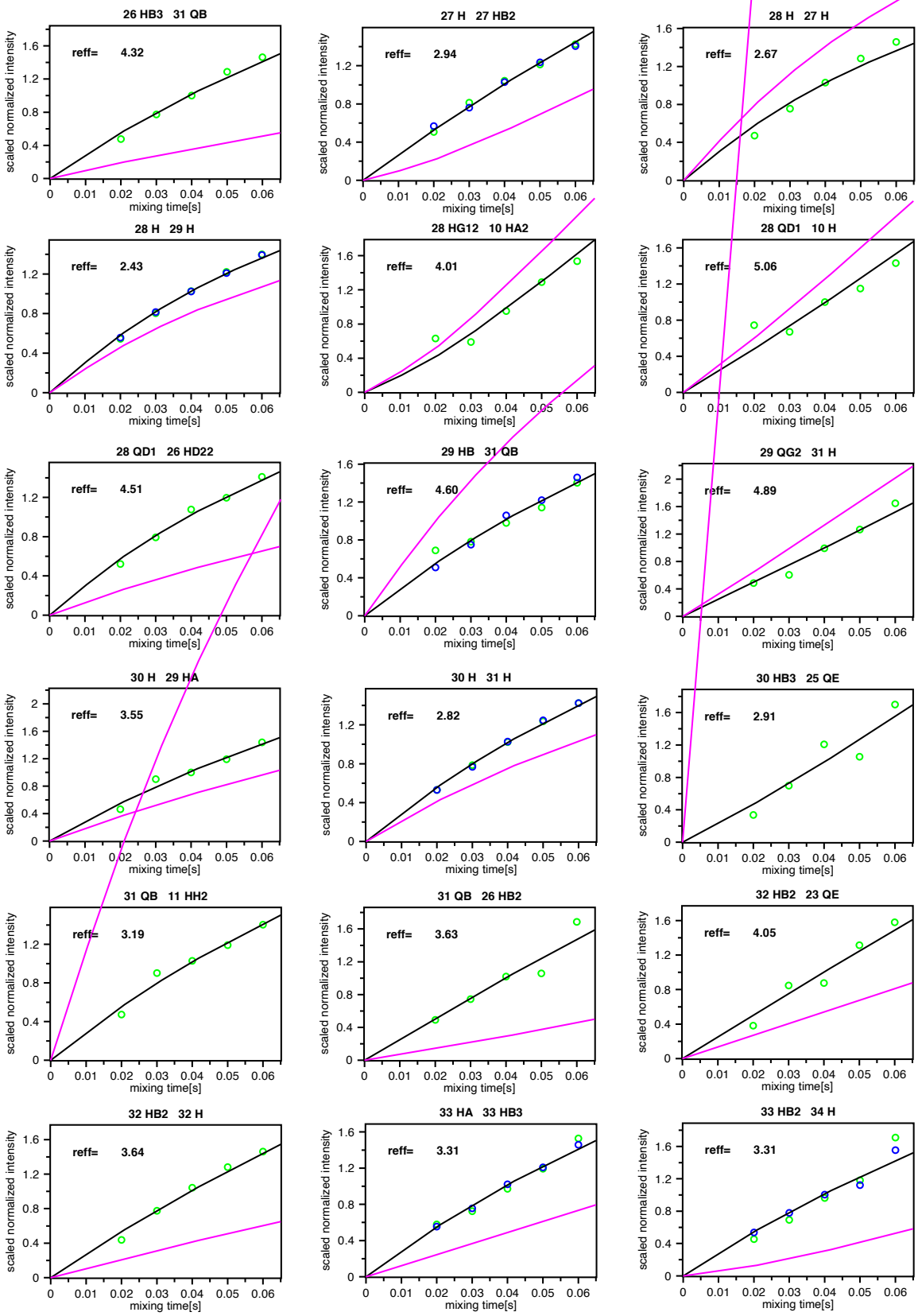

26 HB3 31 QB
reff= 4.32
27 H 27 HB2
reff= 2.94
28 H 27 H
reff= 2.67
28 H 29 H
reff= 2.43
28 HG12 10 HA2
reff= 4.01
28 QD1 10 H
reff= 5.06
28 QD1 26 HD22
reff= 4.51
29 HB 31 QB
reff= 4.60
29 QG2 31 H
reff= 4.89
30 H 29 HA
reff= 3.55
30 H 31 H
reff= 2.82
30 HB3 25 QE
reff= 2.91
31 QB 11 HH2
reff= 3.19
31 QB 26 HB2
reff= 3.63
32 HB2 23 QE
reff= 4.05
32 HB2 32 H
reff= 3.64
33 HA 33 HB3
reff= 3.31
33 HB2 34 H
reff= 3.31
scaled normalized intensity
mixing time[s]

Figure S3: Experimental bidirectional eNOE buildups (green and blue dots) versus time against back-predicted buildups of representative NOEs for single-state (pink) and two-states (black) ensembles calculated. In general, the two-states ensemble fulfils the data better than the single-state structure. For example, the last 33 HB2 – 34 H buildup is well fit by the two-state structure (black line), while not well fit in a one state structure calculation (in pink). The rather significant differences are due to the power to the six dependence of the NOE versus 1/distance. The back-predicted buildups were calculated using eNORA2 ([1]([2] CYANA version). The effective distance in the 2 state structure calculation is also listed.

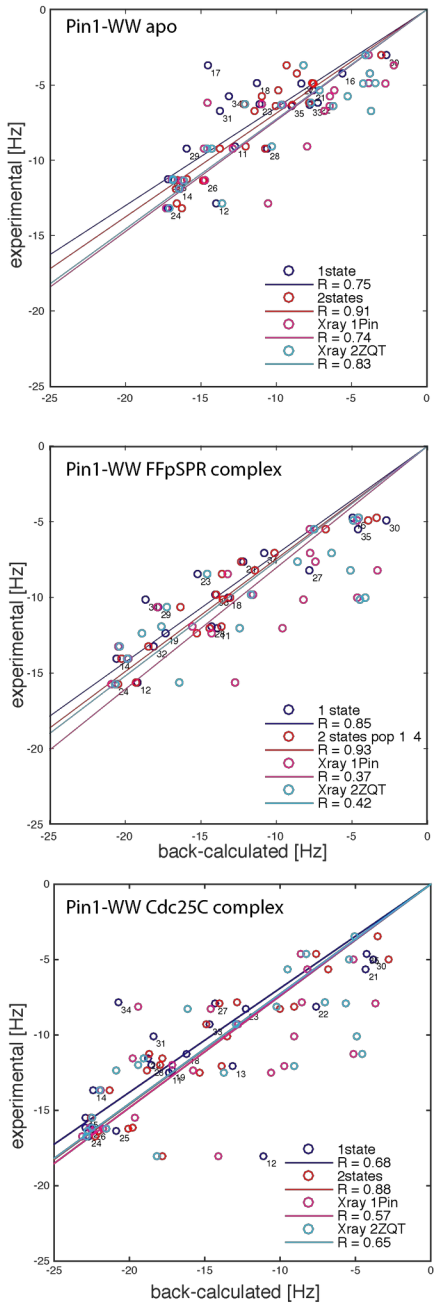


Figure S4: Cross validation of the two-state ensembles using cross-correlated relaxation rates not used in the structure calculation. Cross-correlated relaxation rates $\Gamma_{HNiNi/H\alpha iC\alpha i} + \Gamma_{H\alpha iNi/HNiC\alpha i}$ were obtained as described in Material and Methods. The procedure for the back-calculation of the cross-correlated relaxation rates has previously been described in detail [3]. The increase of Pearson's correlation coefficient *R* from the one-state structure to the two-state structure shows that the two-state structure calculation fulfils the cross-correlated relaxation data better than the one-state structure. In addition, the experimental cross-correlated relaxation rates were compared with back-calculated values using the x-ray structure (pdb 2ZQT). As these *R* values are rather low the x-ray structure does not fulfil the experimental data well.

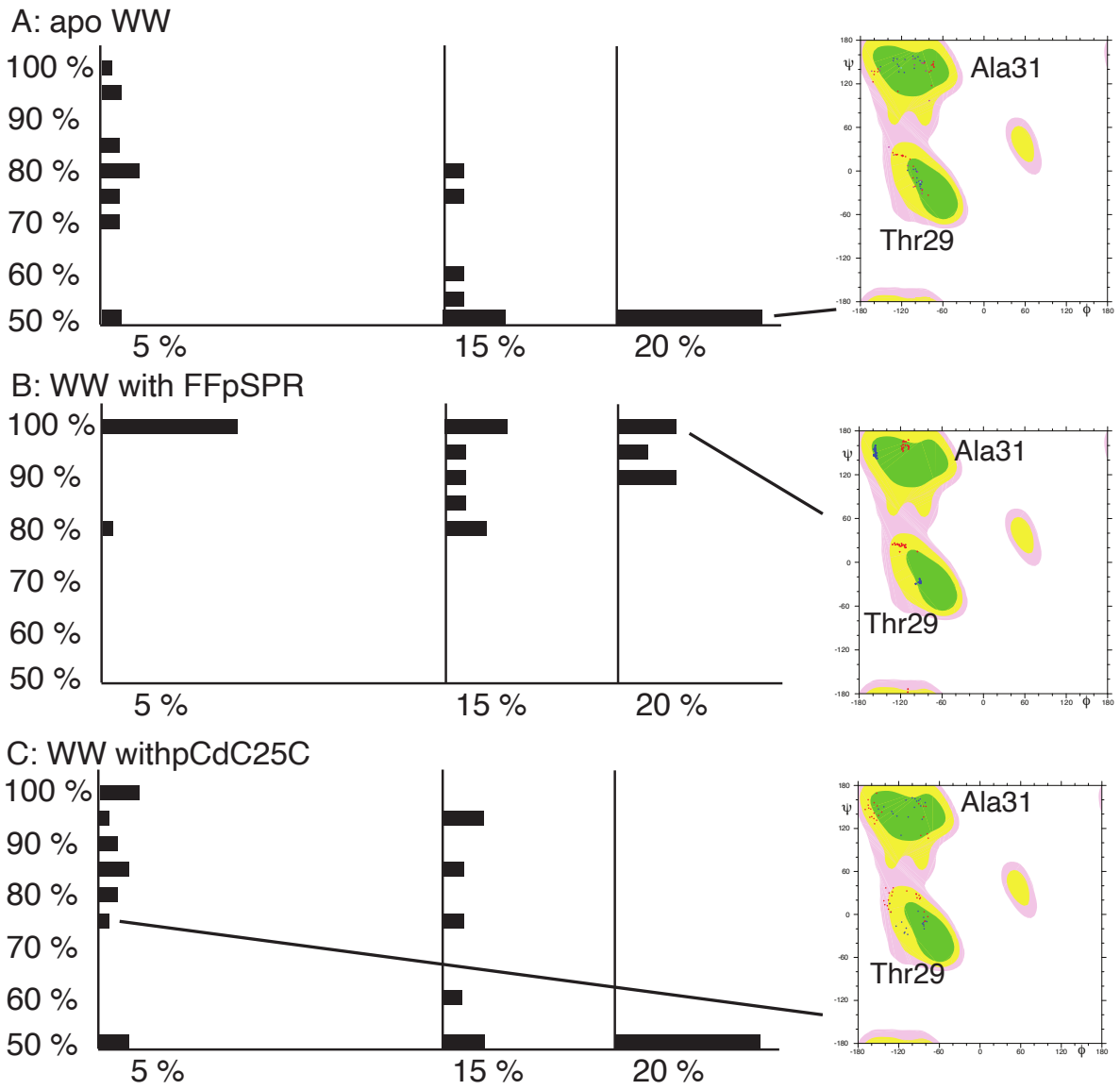


Figure S5: Cross-validation test performed with a jackknife procedure shows that the experimental data for the two-state structure calculations are slightly overdetermined. In the jackknife procedure, the structure calculations were repeated twenty times with 5% (7 times with 15%, and 5 times with 20%, respectively) of the experimental input data randomly deleted such that each distance restraint is omitted exactly once. The presence of the two states including the angular correlation between Ala31 and Thr29 discussed in the text was checked as exemplified for three Ramachandran examples on the right. If in the entire calculation the two states including the angular correlation between Ala31 and Thr29 was observed the outcome was included in the bar at 100% (see middle Ramachandran plot). In the absence of a correlation between the two states (as exemplified with the top Ramachandran plot), the bar at 50% (which means entirely random) was added a value (bottom plot). Otherwise, the value between the extremes was accordingly added. The vertical bar diagram summarizes the jackknife procedure and shows the robustness of the two-state structure calculations in the case of 5% data deletion for all three systems, while in the case of 15% data deletion only the WW in complex with FFpSPR was still robust.

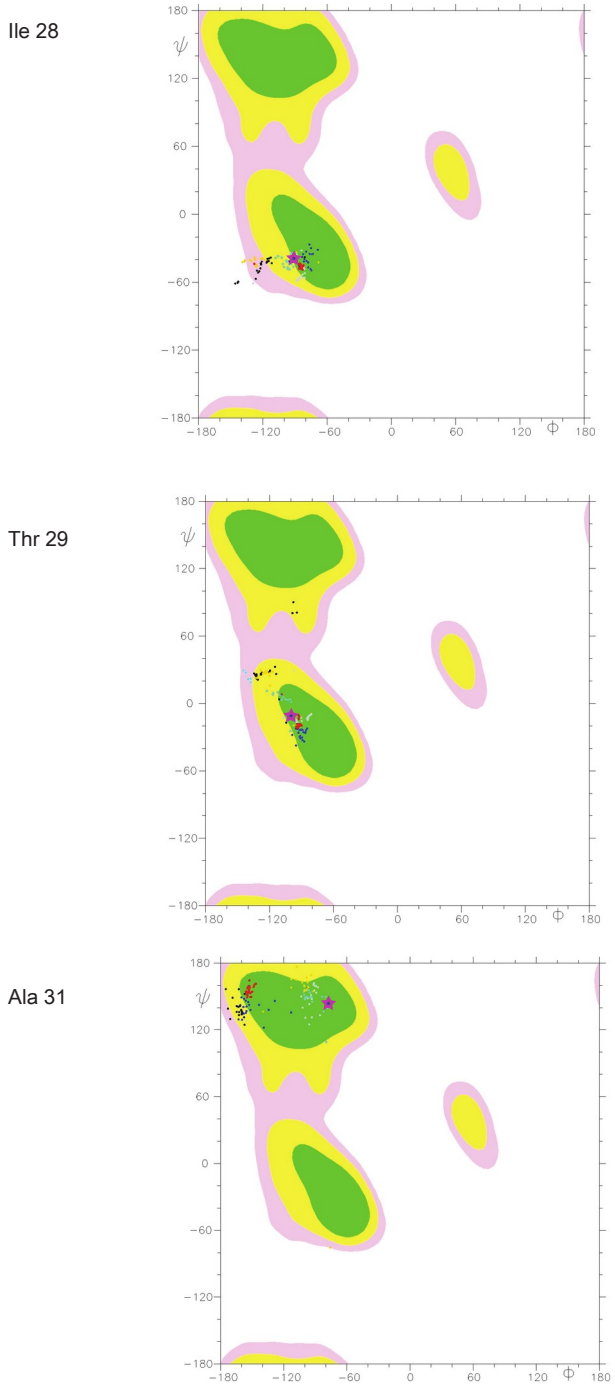


Figure S6: Ramachandran plots for Ile28, Thr29 and Ala31 show the two states of Thr29 and Ala31 in each of the structures determined. The same color code is used as for the 3D structures shown in Figure 3**.** In addition, the Ramachandran angle of the x-ray structure (pdb code 1PIN) is shown with a pink star.

Table S4: ^{15}N relaxation data analysis by RELAX [4]

Assignment	ω_1 [ppm]	ω_2 [ppm]	$\Delta\omega$ [ppm]	R_{ex}(field indep.) [1/s]	R_{ex} (700MHz) [1/s]	R_{ex} error (700 MHz) [1/s]	R_{ex} (900MHz) [1/s]	R_{ex} error (900 MHz) [1/s]	k_{ex} [1/s]	k_{ex} error [1/s]
L7N-H	124.735	124.805	0.07	1.02E-20	0.19741492	0.090849599	0.32633895	0.15017995	787	535
G10N-H	111.752	111.658	0.094	1.05E-20	0.203275303	0.084618	0.336026521	0.139878735	1378	897
W11N-H	117.612	117.326	0.286	1.04E-19	2.018932628	0.053687362	3.337419242	0.088748497	1284	643
E12N-H	120.649	120.662	0.013	2.07E-20	0.399519409	0.064430088	0.660430044	0.106506881	13	7.
K13N-H	125.843	125.636	0.207	2.00E-20	0.38633454	0.062107379	0.638634648	0.1026673	3517	1847
R14N-H	128.035	128.044	0.009	1.36E-20	0.262230633	0.074964379	0.433483292	0.123920709	10	6
M15N-H	117.571	117.46	0.111	4.02E-20	0.777101244	0.078777551	1.284595934	0.130224116	503	257
S16N-H	121.559	121.382	0.177	1.73E-20	0.334463448	0.052539789	0.552888557	0.086851488	2971	1557
R17N-H	128.047	128.047	0	9.85E-20	1.905064964	0.149457648	3.149189023	0.247062643	0	0
N18N-H	116.019	115.662	0.357	3.29E-20	0.635619717	0.11325656	1.050718308	0.187220028	6359	3375
S19N-H	112.29	111.334	0.956	2.84E-20	0.549764276	0.11993689	0.908794007	0.198263022	5273	28761
G20N-H	111.565	111.804	0.239	6.22E-20	1.201806585	0.049405319	1.986659865	0.081670017	1507	756
R21N-H	119.485	119.179	0.306	1.00E-19	1.937076253	0.088374653	3.202105642	0.146088712	1533	769
Y24N-H	117.05	117.107	0.057	1.96E-20	0.378104833	0.073297058	0.625030438	0.121164525	273	146
N26N-H	129.363	129.37	0.007	5.62E-20	1.085848762	0.072528225	1.794974484	0.119893596	1	1
H27N-H	119.9	119.824	0.076	5.47E-20	1.057105518	0.101299514	1.747460141	0.167454298	173	88
I28N-H	122.101	122.672	0.571	2.21E-20	0.426961024	0.065385374	0.705792713	0.108086026	24218	12664
T29N-H	107.842	107.986	0.144	4.94E-20	0.954940541	0.068279449	1.57857518	0.11287011	688	347
N30N-H	117.937	118.546	0.609	5.42E-20	1.048357953	0.050541227	1.732999882	0.083547743	11219	5635
A31N-H	122.001	120.433	1.568	9.57E-20	1.850625758	0.048503333	3.059197681	0.080178978	42134	21095
S32N-H	112.371	113.017	0.646	9.49E-21	0.183469426	0.043847851	0.303286194	0.072483181	72137	39977
Q33N-H**	117.012	120.028	3.016	7.01E-20	1.35550984	0.042258674	2.240740756	0.069856175	212823	106618
E35N-H	118.854	118.702	0.152	2.98E-20	0.576998859	0.037721642	0.95381444	0.062356184	1270	640
S38N-H	115.762	115.734	0.028	4.82E-20	0.931806172	0.084630338	1.540332652	0.139899131	27	13
G39N-H	116.793	116.668	0.125	3.97E-20	0.767988	0.058603899	1.269531184	0.096875833	645	326

R_{ex} is the extracted motion slower than the rotational correlation time of the molecule at the two fields indicated. The k_{ex} are shown also shown in a graph representation in FigureS2D. The k_{ex} and its error are rounded to the digit. **The data for Q33 was analysed at 288 K.

References

[1] J. Orts, B. Vögeli, R. Riek, J. Chem. Theory Comput. **2012**, 8,

[2] D. Strotz, J. Orts, C. N. Chi, R. Riek, B. Vögeli, J. Chem. Th
4336–4346.

[3] B. Vogeli, L. Vugmeyster, Chemphyschem **2019**, 20, 178-196.

[4] M. Bieri, E. J. d'Auvergne, P. R. Gooley, J. Biomol. NM

[5] J. C. Crane, E. K. Koepf, J. W. Kelly, M. Gruebele, J. Mo

[6] J. L. Price, D. Shental-Bechor, A. Dhar, M. J. Turner, E.
Levy, J. W. Kelly, J. Am. Chem. Soc. **2010**, 132, 15359-15367.

[7] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfe
1995, 6, 277–293.

[8] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fo
Ulrich, J. L. Markley, J. Ionides, E. D. Laue, Proteins **20**

[9] D. Strotz, J. Orts, M. Minges, B. Vögeli, J. Magn. Reson. **2015**, 25

[10] B. Vögeli, P. Güntert, R. Riek, Mol. Phys. **2013**, 111, 437–454.

[11] S. Grzesiek, H. Kuboniwa, A. P. Hinck, A. Bax, J. Am. Ch
5315.

[12] C. Perez, F. Löhr, H. Rüterjans, J. M. Schmidt, J. Am. Ch
7093.

[13] J. S. Hu, S. Grzesiek, A. Bax, J. Am. Chem. Soc. **1997**, 11
[14] P. Pelupessy, E. Chiarparin, R. Ghose, G. Bodenhausen, J. Bio
277-280.
[15] R. B. Fenwick, C. D. Schwieters, B. Vogeli, J. Am. Chem. So
8421.
[16] B. Vogeli, J. Biomol. NMR **2017**, 67, 211-232.
[17] B. Vogeli, J. Chem. Phys. **2010**, 133.
[18] P. Vallurupalli, G. Bouvignies, L. E. Kay, J. Am. Chem. Soc. **20**
[19] E. J. d'Auvergne, P. R. Gooley, J. Biomol. NMR **2008**, 40
[20] D. Fushman, N. Tjandra, D. Cowburn, J. Am. Chem. Soc. **19**
[21] B. Vögeli, S. Kazemi, P. Güntert, R. Riek, Nat. Struct. Mol. Bio
1057.
[22] B. Vögeli, Prog. Nucl. Magn. Reson. Spectrosc. **2014**, 78, 1–46.
[23] P. Güntert, C. Mumenthaler, K. Wüthrich, J. Mol. Biol. **19**
[24] P. Güntert, Eur. Biophys. J. **2009**, 38, 129–143.
[25] P. Luginbühl, T. Szyperski, K. Wüthrich, J. Magn. Reson. B **19**