Supplementary Information: Protein allostery of the WW domain at atomic resolution

Dean Strotz[1], *Julien Orts*[1], *Harindranath Kadavath*[1], *Michael Friedmann*[1], *Dhiman Ghosh*[1], *Simon Olsson*[2], *Celestine N. Chi*[3], *Peter Güntert*[1,4,5], *Beat Vögeli*[6 *] and *Roland Riek*[1*]

[1] Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH-Hönggerberg, CH-8093 Zürich, Switzerland.

[2] Department of Mathematics and Computer Science, Freie Universität Berlin, Arnimallee 6, 14195 Berlin, Germany

[3] Department of Medical Biochemistry and Microbiology, Uppsala Biomedical Center, Uppsala University, 751 23 Uppsala, Sweden

[4] Institute of Biophysical Chemistry, Center for Biomolecular Magnetic Resonance, and Frankfurt Institute for Advanced Studies, J.W. Goethe-Universität, Max-von-Laue-Str. 9, 60438 Frankfurt am Main, Germany.

[5] Graduate School of Science, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan.

[6] 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 and roland.riek@phys.chem.ethz.ch

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

NMR distance and dihedral constraints			
Distance constraints			
Total eNOEs	686		
eNOEs from one pathway	415 ($NORM_{orig}$ and $NORM_{dest}$ [8])		
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			
3*J*HNα scalar couplings	26		
3*J*HαHβ scalar couplings	24		
3*J*HNCG scalar couplings (aromatic)	6		
3*J*HNCOCG 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 2: Structural statistics and CYANA input data for the WW domain in complex with pCdc25C

NMR distance and dihedral constraints			
Distance constraints			
Total eNOEs	711		
eNOEs from one pathway	450 ($NORM_{orig}$ and $NORM_{dest}$ [8])		
eNOEs from two pathways	261		
Intra-residue, $\| i-j \| = 0$	258		
Sequential, $\|i-j\| = 1$	169		
Short-range, $\| i-j \| <= 1$	427		
Medium-range, $1 < \|i-j\| < 5$	91		
Long-range, $\|i-j \| >= 5$	193		
Dihedral angle restraints			
3*J*HNα scalar couplings	24		
3*J*HαHβ scalar couplings	23		
3*J*HNCG scalar couplings (aromatic)	6		
3*J*HNCOCG 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

NMR distance and dihedral constraints			
Distance constraints			
Total eNOEs	760		
eNOEs from one pathway	473 ($NORM_{orig}$ and $NORM_{dest}$ [8])		
eNOEs from two pathways	287		
Intra-residue, $\|i-j\| = 0$	257		
Sequential, $\|i-j\| = 1$	192		
Short-range, $\|i-j\| <= 1$	449		
Medium-range, $1 < \|i-j\| < 5$	93		
Long-range, $\|i-j\| >= 5$	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

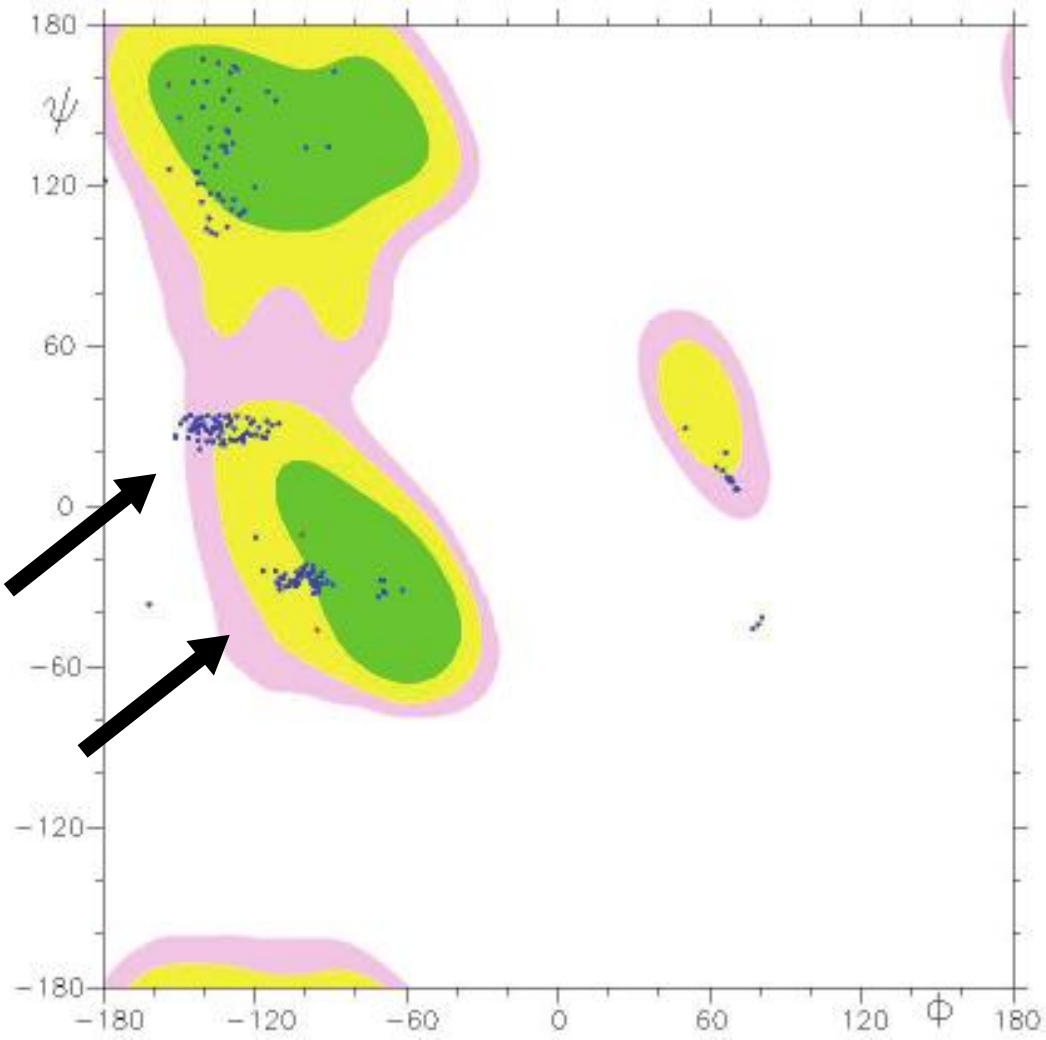
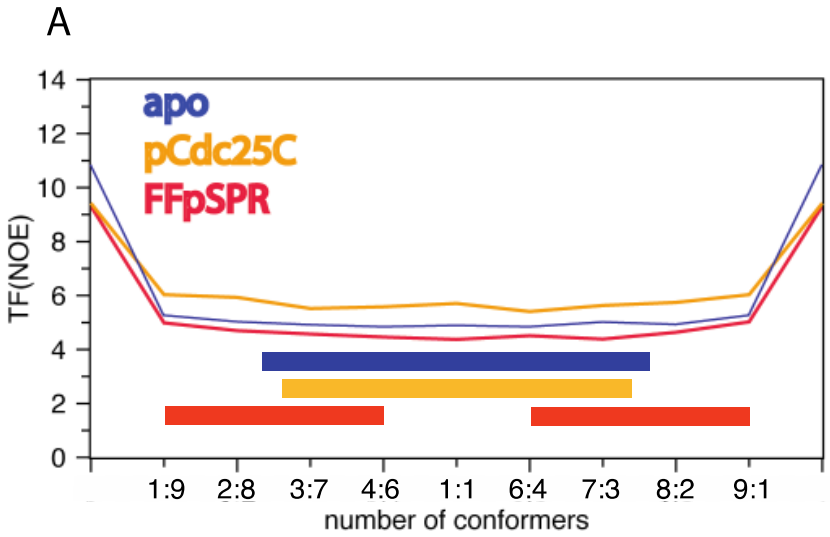


Figure S1: The two states are preserved also for a ten-state structure calculation of apo WW as exemplified by the Ramachandran plot of Thr29. The Ramachandran plot of Thr29 of all the 20 conformers of the ten-state structure calculation (i.e. 200 conformers in total) is shown. While there are outliers,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


A
apo
pCdc25C
FFpSPR
TF(NOE)
number of conformers
1:9
2:8
3:7
4:6
1:1
6:4
7:3
8:2
9:1

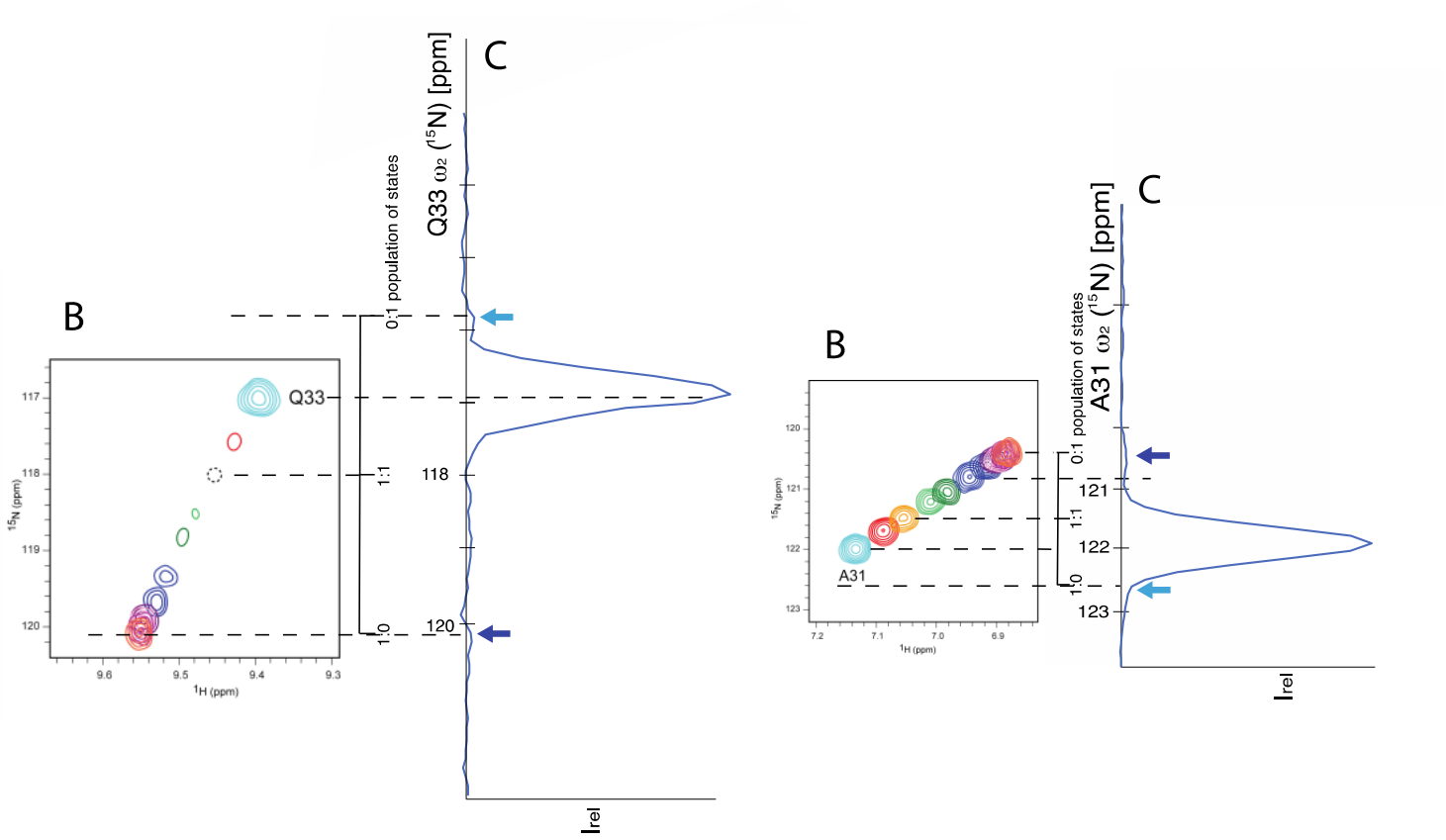

B
C
Q33
Q33 ω2 (15N) [ppm]
0:1 population of states
1:1
1:0
118
120
Irel
A31
A31 ω2 (15N) [ppm]
121
122
123

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. 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). 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.

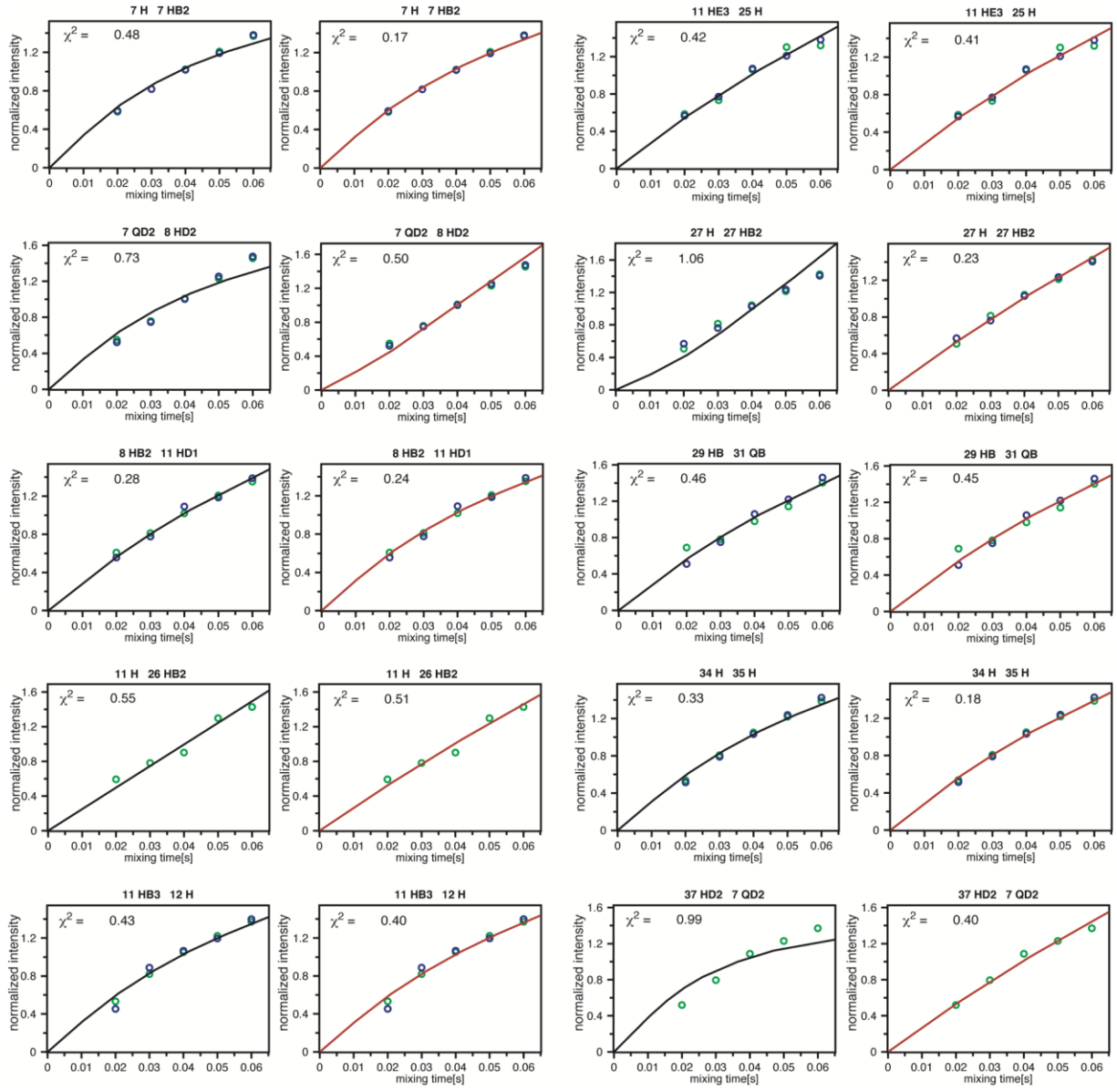


Figure S3: Experimental bidirectional eNOE buildups (green and blue dots) versus time against back-predicted buildups of representative NOEs for single-state (black) and two-states (red) ensembles calculated. The connecting lines are drawn to guide the eye. The two-states ensemble fulfils the data better than the single-state structure. For example, the last 37 HD2 – 7 QD2 buildup is well fit by the two-state structure (red line), while not well fit in a one state structure calculation (in black). The back-predicted buildups were calculated using eNORA2 (Orts et al., 2012, Strotz et al., 2017, CYANA version). χ^2 is the sum of the squared violations between the measured and modeled intensities (Vögeli et al., 2013).

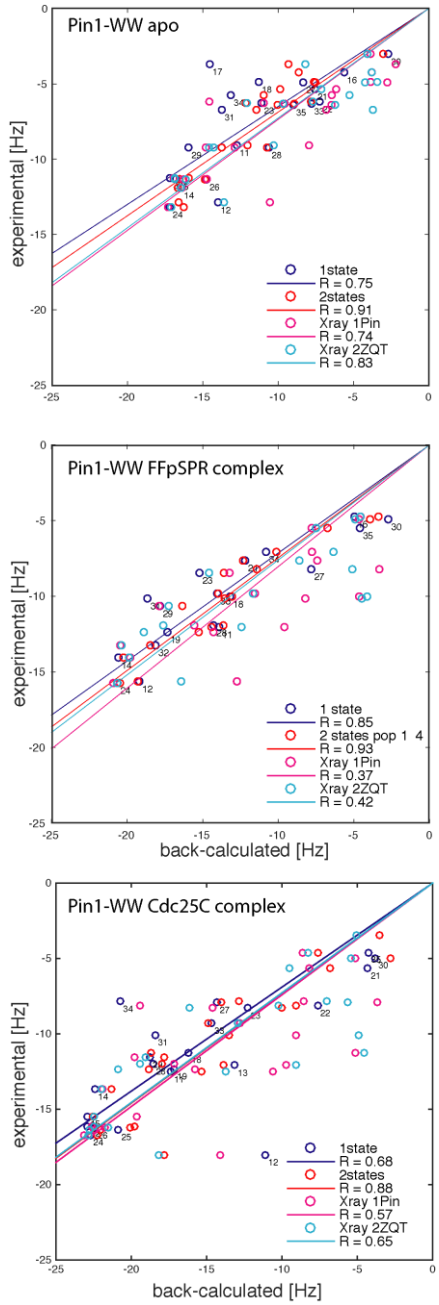


Figure S4: Cross validation of the two-state ensembles using cross-correlated relaxation rates not used in the structure calculation. Cross-correlated relaxation rates $\Gamma_{HNiNi/H\alpha iC\alpha i} + \Gamma_{H\alpha iNi/HNiC\alpha i}$ were obtained as described in Material and Methods. The procedure for the back-calculation of the cross-correlated relaxation rates has previously been described in detail (Vögeli B et al. 2019). 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 fulfills 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 fulfill the experimental data well.

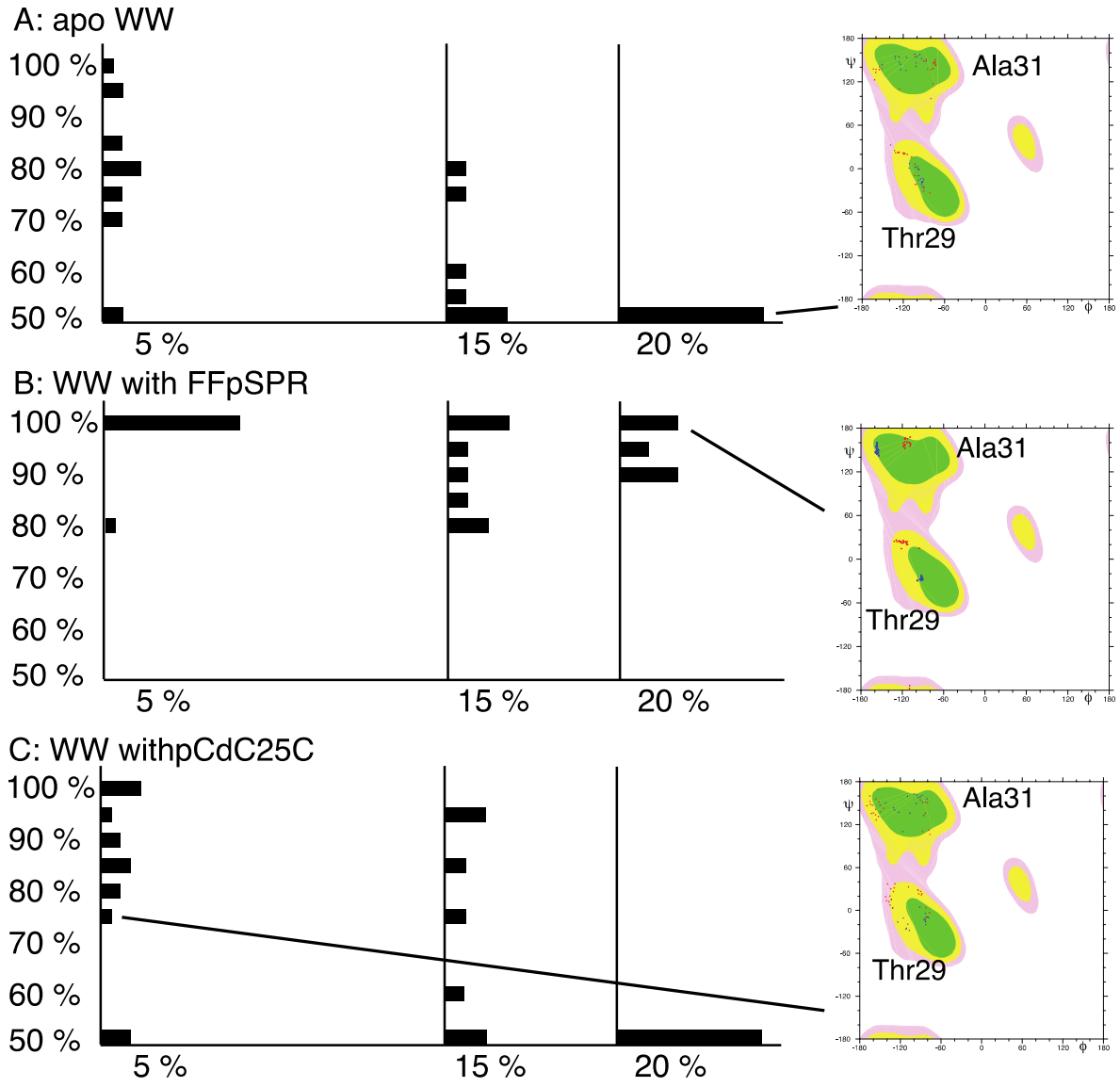


Figure S5: Cross-validation test performed with a jackknife procedure shows that the experimental data for the two-state structure calculations are slightly overdetermined. In the jackknife procedure, the structure calculations were repeated twenty times with 5% (7 times with 15%, and 5 times with 20%, 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.

References

Orts J, Vögeli B, Riek R (2012) Relaxation Matrix Analysis of Spin Diffusion for the NMR Structure Calculation with eNOEs. J ChemTheory Comput 8:3483-3492

Strotz D, Orts J, Chi CN, Riek R, Vögeli B (2017) eNORA2 Exact NOE Analysis Program. J Chem Theory Comput 13(9):4336-4346

Vögeli B, Güntert P, Riek R (2013) Multiple-state ensemble structure determination from eNOE spectroscopy. Mol Phys 111(3):437-454

Vögeli B, Vugmeyster L (2019) Distance-independent cross-correlated relaxation and isotropic chemical shift modulation in protein dynamics studies. ChemPhysChem 20:178-196