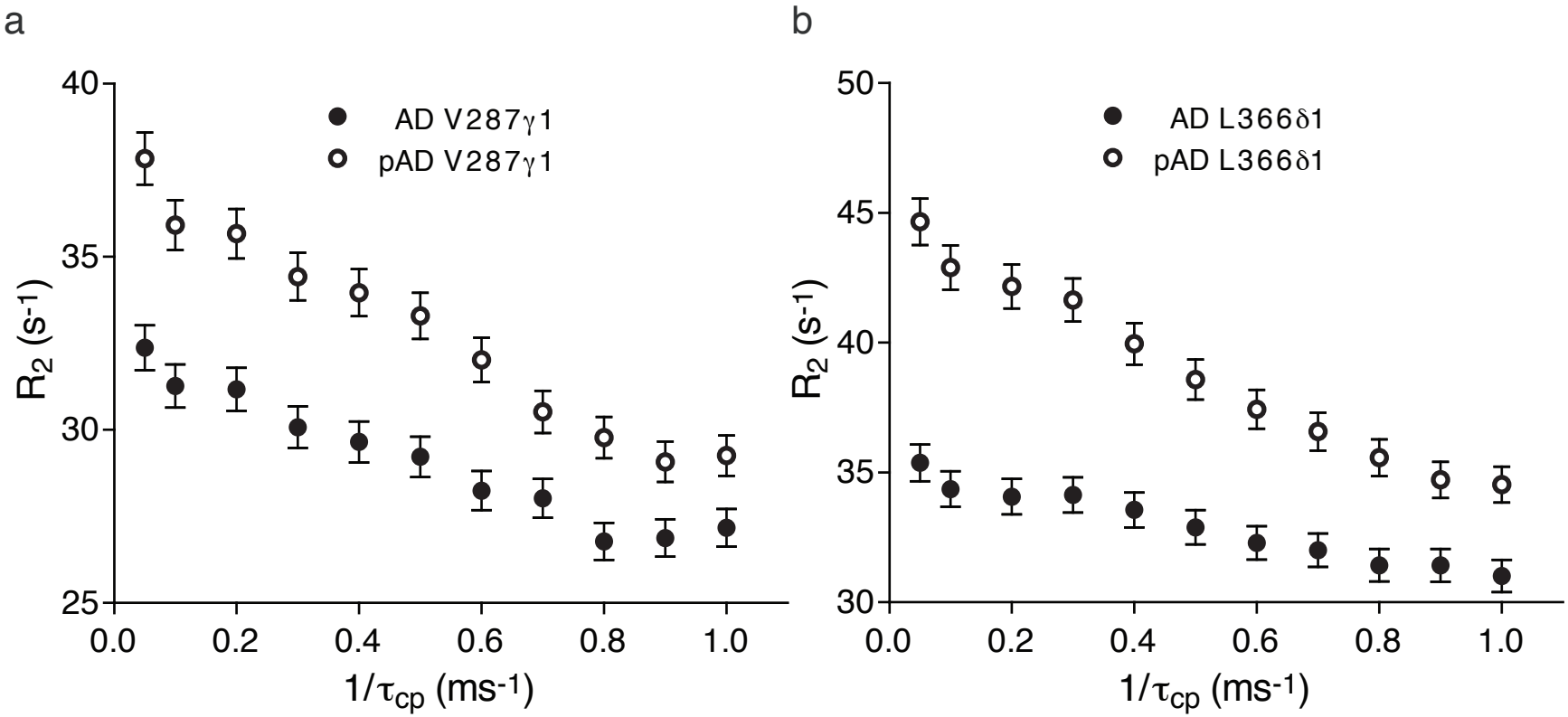
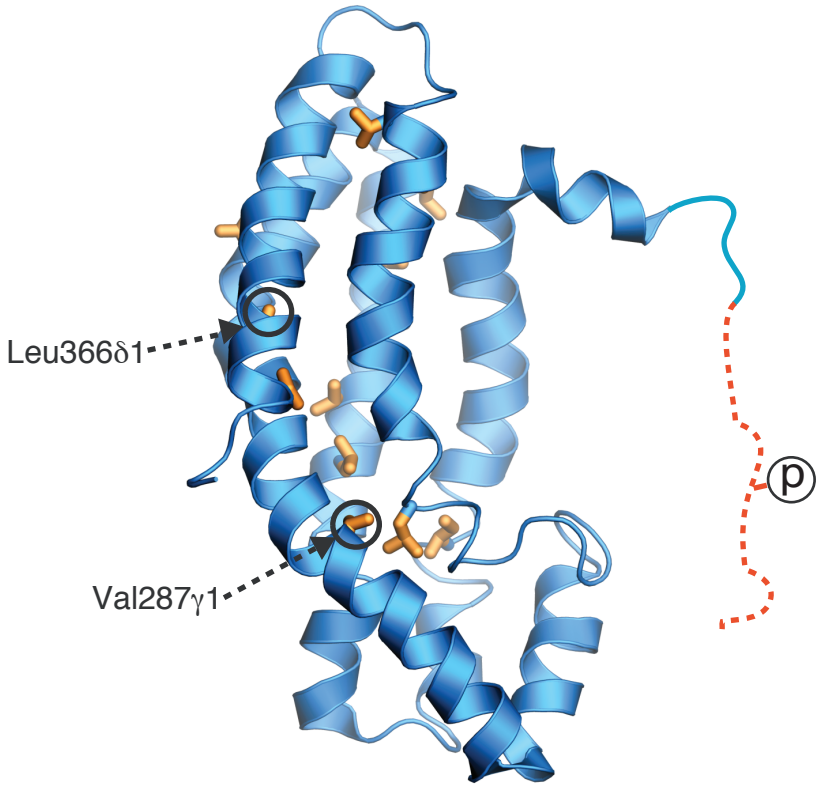
Internal Dynamics Control Activation and Activity of the Autoinhibited Vav DH Domain

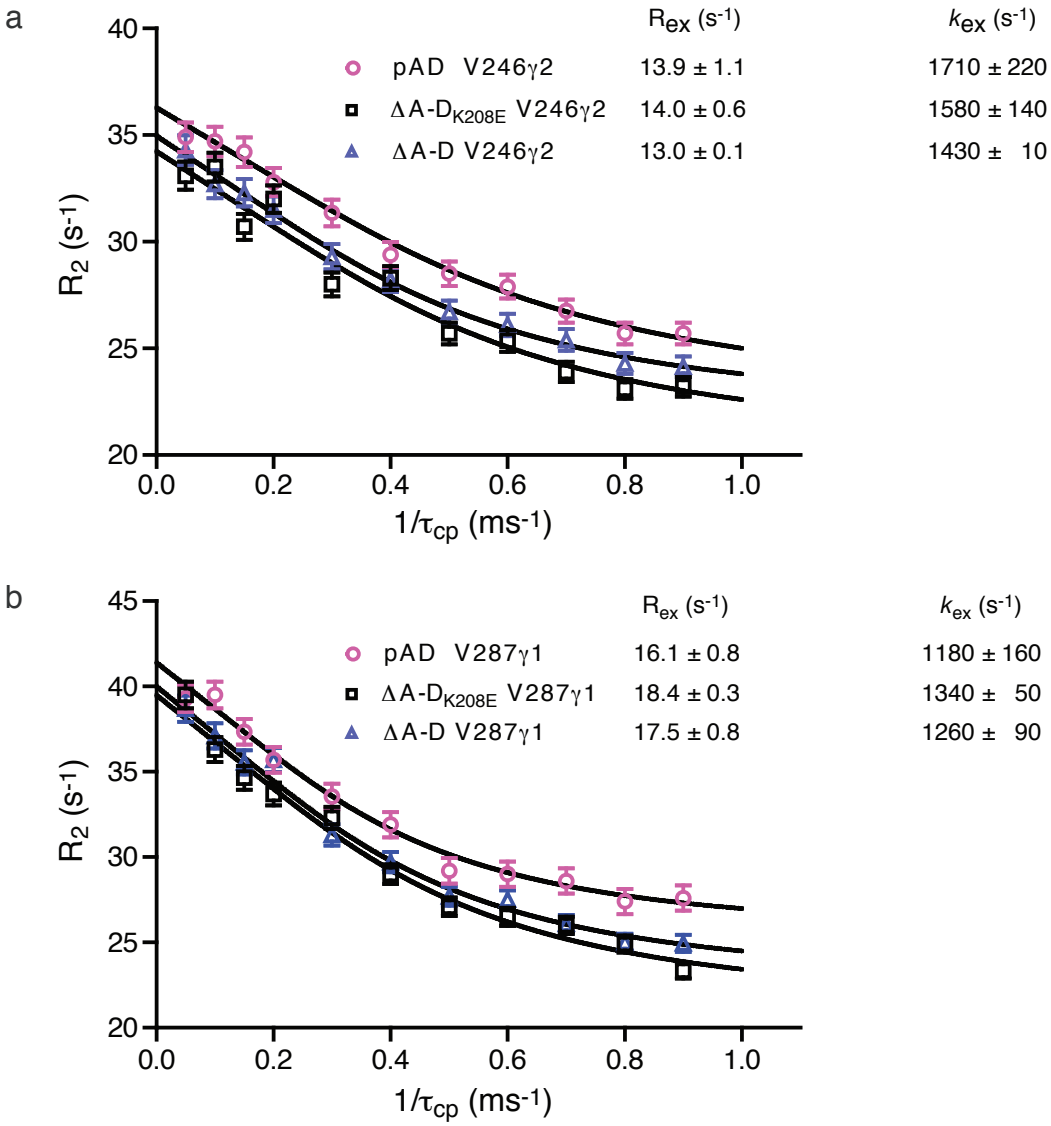
Pilong Li, Ilídio R. S. Martins, Gaya K. Amarasinghe, Michael K. Rosen



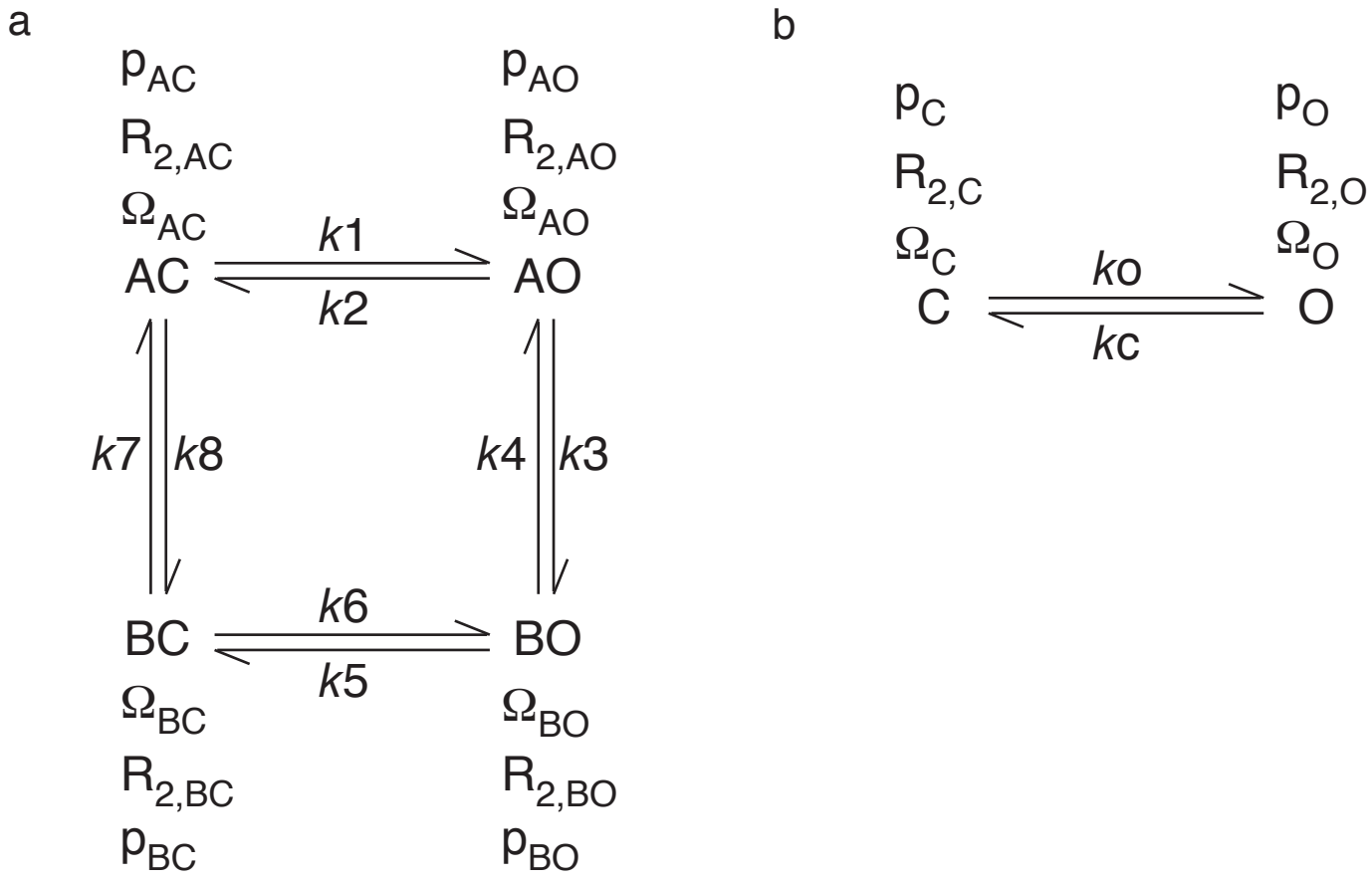
Supplementary Figure 1. Single quantum ^{13}C methyl relaxation dispersion curves for representative resonances that show increased relaxation dispersion amplitude upon phosphorylation. Data are shown for Val287γ1 (panel **a**) and Leu366δ1 (panel **b**) in AD (solid circle) and pAD (open circle). Error bars are obtained from the average of the standard deviations of at least two duplicate points if this average is above 2.0 % of the R_2 value, and are set to 2.0 % of the corresponding R_2 if below.



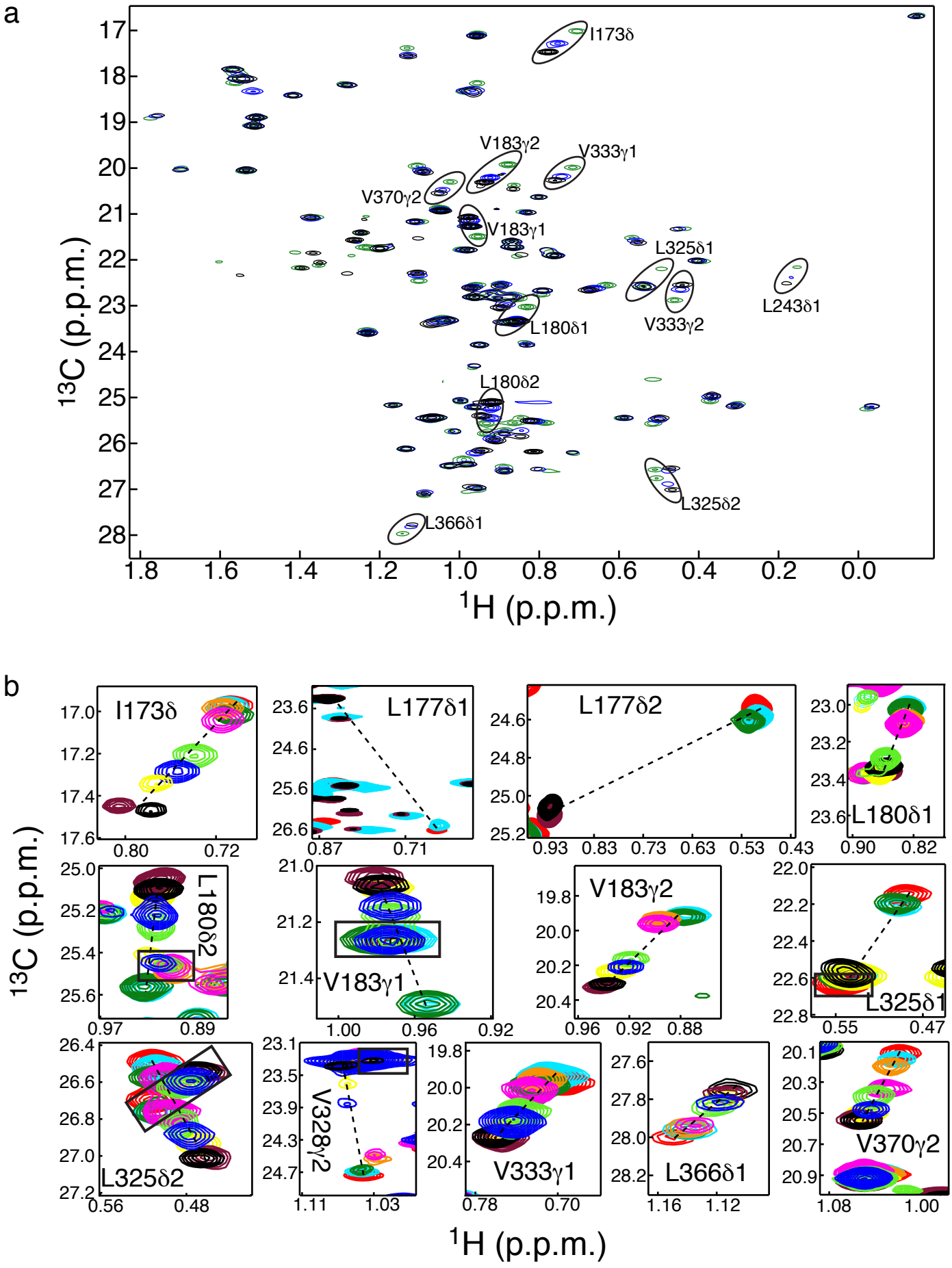
Supplementary Figure 2. Ribbon diagram of the pAD structure model based on AD solution structure (pdbID, 1F5X1) with methyl-containing sidechains showing $\Delta R_2 > 2.0\ s^{-1}$ drawn as sticks. The sidechains of Val287γ1 and Leu366δ1, whose resonances are shown in Supplementary Figure 1 are circled. Structure rendered using Pymol[2].



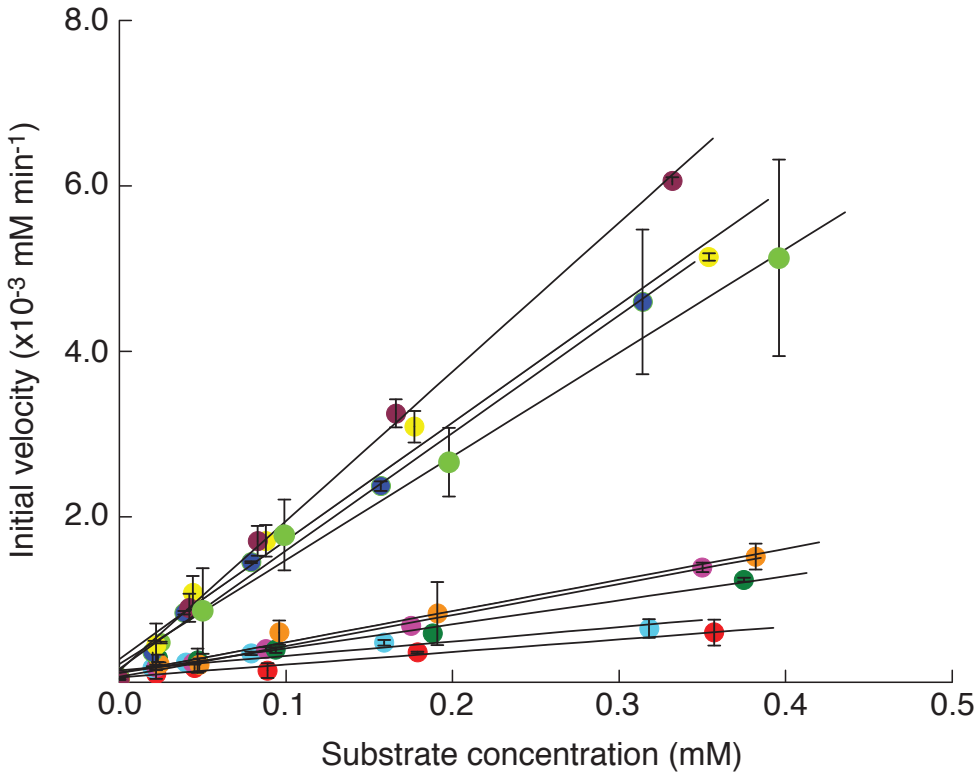
Supplementary Figure 3. Single quantum ^{13}C methyl relaxation dispersion curves for pAD and the ΔA-D (residues 181-375, containing only the DH domain) and ΔA-D_{K208E}. Data are shown for Val246γ2 (panel **a**) and Val287γ1 (panel **b**) in pAD (pink open circle), ΔA-D (blue open triangle) and ΔA-D_{K208E} (black open square). Both resonances have large exchange contributions from the dynamic process that is intrinsic to the DH domain (see main text and Supplementary Methods text for discussion of four-state dynamics in AD), and thus provide a sensitive comparison of the three proteins. Data were acquired at 10 °C on a 600 MHz instrument. Error bars are obtained from the average of the standard deviations of the four duplicate points if this average is above 2 % of the R_2 value, and are set to 2 % of the corresponding R_2 if below. Data were fit to a two-state model to yield values for R_{ex}, the exchange contribution to R_2, and k_{ex}, the sum of the forward and reverse rate constants. Correspondence of the raw data and fitted values between pAD and ΔA-D indicates that the former is a good mimic of a protein lacking the inhibitory helix entirely, and thus of a fully open state. This conclusion is also supported by the fact that methyl ^{1}H and ^{13}C chemical shifts of ΔA-D are identical within experimental error to those from the DH domain of pAD, with the exception of peaks from Val183γ1, γ2 (which probably are affected by the nearby charged N-terminus in the shorter construct). Correspondence of the raw data and fitted values between ΔA-D and ΔA-D_{K208E} indicates that the mutation does not affect the equilibrium intrinsic to the DH domain. Again, this conclusion is also supported by the similarity between ΔA-D and ΔA-D_{K208E} chemical shifts (not shown).



Supplementary Figure 4. Models for the four-state equilibrium and the simplified two-state equilibrium involving the Ac helix in the AD protein. **(a)** The system is described by two coupled processes, one involving dissociation of the inhibitory helix, and a second within the DH domain itself. AC and BC represent the helix-bound (closed) states of the system; AO and BO represent the helix-dissociated (open) states. The dynamics within the DH domain are modeled as a second two-state equilibrium between states A and B. Fluctuations across this second process thus convert AO/AC to BO/BC. **(b)** In the simplified model, only the process involving dissociation of the inhibitory helix is depicted. In both **(a)** and **(b)**, $R_{2,i}$ is the transverse relaxation rate of species i in the absence of exchange. pi is the fractional population of species i. Ω_i is the chemical shift of species i. The ks are first-order rate constants for the indicated transitions.


a
I173δ
V183γ2
V333γ1
V370γ2
V183γ1
L325δ1
L243δ1
L180δ1
V333γ2
L180δ2
L325δ2
L366δ1
13C (p.p.m.)
1H (p.p.m.)
b
I173δ
L177δ1
L177δ2
L180δ1
L180δ2
V183γ1
V183γ2
L325δ1
L325δ2
V328γ2
V333γ1
L366δ1
V370γ2
13C (p.p.m.)
1H (p.p.m.)

Supplementary Figure 5. Overlay of $^{1}H/^{13}C$ HSQC spectra. (a) Methyl region of $^{1}H/^{13}C$ HSQC spectra (Alaβ, Valγ1/2, Leuδ1/2 and Ileγ2) of AD (dark green), AD_{K208A} (blue), and pAD (black) are shown. Note that the fractional displacement of resonances in AD_{K208A} relative to AD and pAD is very consistent among the various peaks. (b) Expansions of overlays for all peaks used in the analysis of Figure 3, except for those shown in Figure 2a. Dotted lines along peak shift trajectory help identify the relevant peaks in the selected regions. Peaks are colored according to Figure 2a. Boxes indicate overlapping resonances that are not part of the mutant trajectory.



Supplementary Figure 6. Vav mutants show different phosphorylation rates. Initial phosphorylation rate versus substrate concentration is shown for a series of AD constructs. Phosphorylation data were collected in the linear region of the Michaelis-Menten curve at substrate concentrations below K_M for the most open construct (AD_{K208E}). Data for the following DH domain mutants are shown: K208E (burgundy), K208S (yellow), K208A (blue), E169K K208A (light green), K208V (pink), E169K K208V (orange), wild type (dark green), E169K (cyan), and E169K E207K (red).

Supplementary Figure References

1. Aghazadeh, B., Lowry, W.E., Huang, X.Y. & Rosen, M.K. Structural basis for relief of autoinhibition of the Dbl homology domain of proto-oncogene Vav by tyrosine phosphorylation. *Cell* **102**, 625-33 (2000).
2. Delano, W.L. The PyMOL Molecular Graphics System *DeLano Scientific, San Carlos, CA, USA* (2002).

Supplementary Methods

Four-state dynamics in the AD protein

As shown by Kay and colleagues for three-state[1,2], and below for four-state, in systems of this complexity two-state fitting in the presence of realistic noise in the data can yield highly inaccurate values for populations and rates across the various equilibria. We have not yet assessed whether the six NMR experiments used successfully by Kay and colleagues to solve the three-state case[2] would be successful in the four-state system as well. However, even in the three-state case, only the linear system is currently tractable, and even there, the fitted parameters have significant uncertainties when real data are analyzed[1,2]. In the four-state case, the number of fitted parameters increases substantially over the three-state. This not only increases the number of adjustable parameters in the fitting, but also significantly increases the interdependence of the parameters, making it very difficult to find unique solutions. Thus, it is highly unlikely that the four-state case is tractable without new NMR methods and/or biochemical simplifications.

In order to understand the limitations of two-state fitting of complex equilibria, we first modeled the four-state system shown in Supplementary Figure 4a (see methods below). This system is a likely representation of the Vav AD protein, where an open-closed equilibrium of the inhibitory helix (AC⇔AO and BC⇔BO) is thermodynamically coupled to a second dynamic process in the DH domain that is assumed to be two-state (AC⇔BC and AO⇔BO). In order to examine more favorable cases (i.e. where a two-state approximation would appear to be a reasonable approximation) the chemical shifts of AO and BO were set identically to 0 ppm and those of AC and BC were set to 0.5

ppm. That is, the A/B equilibrium was assumed to have no effect on chemical shift, such that only the open/closed equilibrium should give rise to an exchange contribution to R_2 directly. This could roughly mimic the situation of the inhibitory arm of the AD protein. Intrinsic R_2 values of AO and BO were assumed to be 5 s^{-1} and those of AC and BC were assumed to be 30 s^{-1}, again to mimic the inhibitory arm in the AD protein. In all cases, rate constants k1-k7 were chosen either randomly (Cases 3-6) or by design (Cases 1-2) and k8 was calculated based on thermodynamic constraints. Fractional state populations were calculated from the rate values to satisfy the equilibrium around the thermodynamic box. Based on these conditions, we generated theoretical noise-free ^{13}C relaxation dispersion data using modified Bloch equations, as described below. Data were simulated at 600 MHz and 800 MHz static field strengths, with applied transverse fields ranging from 50 Hz to 1000 Hz in 50 Hz steps. The resulting simulated data were then fit to the Carver-Richards equation describing the two-state equilibrium shown in Supplementary Figure 4b using a program kindly provided by Drs. N. Korzhnev and L. Kay. The initial four-state inputs and fitted two-state parameters are shown in Cases 1-6 below.

Case 1:

$p_{AO} + p_{BO} = 0.5$; $p_{AC} + p_{BC} = 0.5$

Input parameters

k1	k2	k3	k4	k5	k6	k7	k8	p_{AO}	p_{BO}	p_{AC}	p_{BC}	$\Delta\omega$

1500	1500	100	900	300	300	900	100	0.45	0.05	0.45	0.05	0.5

Fitted parameters assuming two-state equilibrium*

k_o	k_c	p_O	p_C	$\Delta\omega$
1187	2755	0.301	0.699	0.647

-

Case 2:

$p_{AO} + p_{BO} = 0.5$; $p_{AC} + p_{BC} = 0.5$

Input parameters

k1	k2	k3	k4	k5	k6	k7	k8	p_{AO}	p_{BO}	p_{AC}	p_{BC}	$\Delta\omega$
1500	1500	900	100	300	300	100	900	0.05	0.45	0.05	0.45	0.5

Fitted parameters assuming two-state equilibrium*

k_o	k_c	p_O	p_C	$\Delta\omega$
627	645	0.493	0.507	0.679

-

Case 3:

$p_{AO} + p_{BO} = 0.768$; $p_{AC} + p_{BC} = 0.232$

Input parameters

k1	k2	k3	k4	k5	k6	k7	k8	p_{AO}	p_{BO}	p_{AC}	p_{BC}	$\Delta\omega$
1351	1375	1340	864	1660	347	641	4673	0.093	0.675	0.091	0.141	0.5

Fitted parameters assuming two-state equilibrium*

k_o	k_c	p_O	p_C	$\Delta\omega$
2925	878	0. 769	0. 231	0. 68

-

Case 4:

$p_{AO} + p_{BO} = 0.586$; $p_{AC} + p_{BC} = 0.414$

Input parameters

k1	k2	k3	k4	k5	k6	k7	k8	p_{AO}	p_{BO}	p_{AC}	p_{BC}	$\Delta\omega$
1822	1842	1270	1628	1800	932	1194	1780	0.235	0.351	0.233	0.181	0.5

Fitted parameters assuming two-state equilibrium*

k_o	k_c	p_O	p_C	$\Delta\omega$
2712	2846	0.512	0.488	0. 65

-

Case 5:

$p_{AO} + p_{BO} = 0.641$; $p_{AC} + p_{BC} = 0.359$

Input parameters

k1	k2	k3	k4	k5	k6	k7	k8	p_{AO}	p_{BO}	p_{AC}	p_{BC}	$\Delta\omega$
185	1221	1002	356	118	1569	239	7.6	0.621	0.02	0.094	0.265	0.5

Fitted parameters assuming two-state equilibrium*

k_o	k_c	p_O	p_C	$\Delta\omega$
473	125	0.791	0.209	0.65

-

Case 6:

$p_{AO} + p_{BO} = 0.762$; $p_{AC} + p_{BC} = 0.238$

Input parameters

k1	k2	k3	k4	k5	k6	k7	k8	p_{AO}	p_{BO}	p_{AC}	p_{BC}	$\Delta\omega$
131	1049	640	416	140	1433	1012	19	0.748	0.014	0.094	0.144	0.5

Fitted parameters assuming two-state equilibrium*

k_o	k_c	p_O	p_C	$\Delta\omega$
760	97	0.887	0.113	0. 82

*Fitted parameters, k_o, k_c, p_o and p_c, have two-fold degeneracy and the more favorable situation for populations is listed.

In cases 2, 3 and 4, the two-state model yields reasonably accurate population distributions across the open/closed equilibrium (i.e. $p_{AO} + p_{BO}$ used to generate the data is close to p_o obtained in the fit). But for cases 1, 5 and 6, the fitted values show substantial errors. It is not entirely clear why some parameter combinations yield poorer results than others. However, we note that in favorable cases the populations of the A and B components (i.e. AO+AC and BO+BC) were such that the highest populations occurred for the open/closed transitions that intrinsically make a greater contribution to exchange-mediated relaxation (because of exchange rate closer to intermediate exchange). In contrast, when A and B are distributed oppositely, the two-state fits are incorrect. For

example, in Case 2, the B states are 9 times more populated than the A states ($p_{AO}+p_{AC}$ = 0.1, $p_{BO}+p_{BC}$ = 0.9). Also the AO/AC transition has k_{ex} = k1+k2 = 3000 s^{-1}, compared to the BO/BC transition which has k_{ex} = k5+k6 = 600 s^{-1}. On a 600 MHz instrument, $\Delta\omega$ = 470 rad s^{-1}. Thus, the BO/BC transition is closer to intermediate exchange than AO/AC. Populations are essentially matched to exchange, and the two-state model is a reasonable fit. In contrast, in Case 1, the B states are 9 times less populated than the A states, while the k_{ex} rates are identical to Case 2. Here, the less populated arm of the four-state box makes the inherently greater contribution to exchange relaxation, and the two-state fit is much poorer. In the absence of a priori knowledge of which states and transitions are dominant in a system, in cases of four-state equilibria a two-state fit is generally not applicable.

Simulation of a Four-State Equilibrium

The derivations below are based on modified Bloch equations, following the approach of Kay and coworkers in analysis of a 3-state equilibrium[2]. In contrast to that work, in the current situation it is not justified to assume that there is a single dominant species. Therefore R_2 is calculated based on all four states (A1) given the assumption that both processes are in the fast NMR time regime:

$$R_2 = -\frac{1}{T}\ln\frac{\sum M_i(T)}{\sum M_i(0)} \qquad \text{(A1)}$$

where T is the constant relaxation period in seconds; i = AC, AO, BC, BO are defined as in Supplementary Figure 4a; $M_i(0)$ is the initial magnetization of the i^{th} state, which is proportional to its corresponding fractional population in thermal equilibrium; and $M_i(T)$ is the magnetization of the i^{th} state at the end of the constant relaxation time period.

Magnetization $M_\pm$ ($M = M_+ + M_-$) that evolves during the period T is of the form $X_\pm$ (X = ^{13}C and ^{15}N) for methyl ^{13}C single quantum CPMG and amide ^{15}N single quantum CPMG measurements. The application of a $(\delta\text{-}180^o\text{-}\delta)_{2n}$ spin-echo train of length $T = 4n\delta$ gives rise to M(T) as:

$$\mathbf{M}_\pm(\mathbf{T}) = (\exp(\mathbf{A}_\pm\delta)\exp(\mathbf{A}_\pm{*}\delta)\exp(\mathbf{A}_\pm{*}\delta)\exp(\mathbf{A}_\pm\delta))^n\mathbf{M}_\pm(\mathbf{0}) \qquad \text{(A2)}$$

where * denotes complex conjugate; M(0) is proportional to $(p_{AC}, p_{AO}, p_{BC}, p_{BO})^T$ and p_{AC}, p_{AO}, p_{BC}, p_{BO} are the fractional populations of AC, AO, BC, and BO states, respectively. The evolution matrix A is given by:

$$A\pm = \begin{bmatrix} -k1 - k8 - R_{2,AC} \pm i * \Omega_{AC} & k2 & 0 & k7 \\ k1 & -k2 - k3 - R_{2,AO} \pm i * \Omega_{AO} & k4 & 0 \\ 0 & k3 & -k4 - k5 - R_{2,BO} \pm i * \Omega_{BO} & k6 \\ k8 & 0 & k5 & -k6 - k7 - R_{2,BC} \pm i * \Omega_{BC} \end{bmatrix} \quad \text{(A3)}$$

where all parameters are defined in Supplementary Figure 4a.

Derivation 1: Linear behavior of chemical shift in perturbations of a four-state equilibrium

Statement: the resonances derived from fast four-state equilibrium will be collinear given the assumptions that mutations only directly perturb kinetics and thermodynamics of the two edges involving opening-closing (AC-AO and BC-BO) and not the other two edges involving the second dynamic process (AC-BC and AO-BO) and mutations do not change chemical shift of any of the states.

Evidence supporting these assumptions:

First, the sites of our mutations are directly under the helix at its interface with the DH domain. They are far from the resonances that are strongly affected by the DH dynamics (i.e. that show large ΔR_2 in the phosphorylated protein). Thus, from a structural standpoint the mutations should affect the helix dynamics much more than the DH dynamics. Second, ΔA-D and $\Delta A\text{-}D_{K208E}$ have identical chemical shifts (data not shown) and relaxation behaviors (compare black open squares ($\Delta A\text{-}D_{K208E}$) and blue open triangles (ΔA-D) in Supplementary Fig. 3). If the K208E mutation were directly affecting the DH dynamics and structures of the different states, these NMR properties would be different between mutant and wild type proteins. Thus, even in the mutant that has the strongest effect on the dynamics of the AD protein, we can ascribe this effect with great confidence to selective perturbation of the helix equilibrium.

Refer to Supplementary Figure 4 for definitions of all parameters used below.

The observed ^{13}C resonance derived from a methyl group in a particular construct i is given by:

$$\Omega_{observed}^{i} = p_O^{i} * \Omega_O^{i} + p_C^{i} * \Omega_C^{i} \quad \text{(A4)}$$

where p_O^{i}, p_C^{i} are the open and closed fractional population of the construct i and the sum of p_O^{i} and p_C^{i} is unity; Ω_O^{i} and Ω_C^{i} are the effective open and closed state chemical shifts given by:

$$\Omega_O^{i} = [p_{AO}^{i}/(p_{AO}^{i} + p_{BO}^{i})]*\Omega_{AO}^{i} + [p_{BO}^{i}/(p_{AO}^{i} + p_{BO}^{i})]*\Omega_{BO}^{i} \quad \text{(A5)}$$

$$\Omega_C^{i} = [p_{AC}^{i}/(p_{AC}^{i} + p_{BC}^{i})]*\Omega_{AC}^{i} + [p_{BC}^{i}/(p_{AC}^{i} + p_{BC}^{i})]*\Omega_{BC}^{i} \quad \text{(A6)}$$

where index i represents that the parameters are specific for the construct i.

According to the assumptions in the statement, none of the parameters on the right hand side of (A5) and (A6) are affected by mutations and therefore the resulting left hand sides of (A5) and (A6) are mutation independent and the index i of Ω_O^{i} and Ω_C^{i} in (A4), (A5), (A6) can be removed. Therefore (A4) can be rewritten as:

$$\Omega_{observed}^{i} = p_O^{i} * \Omega_O + p_C^{i} * \Omega_C \quad \text{(A7)}$$

which places $\Omega_{observed}^{i}$ on the straight line defined by the two extremes, Ω_O and Ω_C.

Derivation 2: In highly biased equilibria relaxation dispersion is much more sensitive than chemical shift to changes in equilibrium populations.

For simplicity, we take two-state equilibrium (A⇔B) as an example to demonstrate this statement. In fast exchange:

$$\Omega_{observed} = p_A * \Omega_A + (1-p_A) * \Omega_B = p_A * (\Omega_A - \Omega_B) + \Omega_B$$

So,

$$d(\Omega_{observed})/d(p_A) = (\Omega_A - \Omega_B)$$

which is constant for all values of p_A.

The exchange contribution to R_2, R_{ex}, is given by:

$R_{ex} = (p_A * p_B)(\Omega_A - \Omega_B)^2/k_{ex} = (p_A)(1-p_A)(\Omega_A - \Omega_B)^2/k_{ex}$

So,

$d(R_{ex})/d(p_A) = (1-2p_A)(\Omega_A - \Omega_B)^2/k_{ex}$

which is dependent on p_A, and has largest absolute value as p_A approaches 0 or 1.

Thus, R_{ex} changes most with p_A for small or large values of p_A, while changes is chemical shift with p_A are constant across the entire $\Delta\omega$ range. For example, when $p_A = 0.96$, a population change of 0.02, to $p_A = 0.98$, will introduce a 2% change in chemical shift but a 50% change in R_{ex}.

References

1. Korzhnev, D.M. et al. Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* **430**, 586-90 (2004).
2. Korzhnev, D.M., Neudecker, P., Mittermaier, A., Orekhov, V.Y. & Kay, L.E. Multiple-site exchange in proteins studied with a suite of six NMR relaxation dispersion experiments: an application to the folding of a Fyn SH3 domain mutant. *J Am Chem Soc* **127**, 15602-11 (2005).