

ARTICLE

DOI: 10.1038/s41467-018-06133-0

OPEN

Identifying weak interdomain interactions that stabilize the supertertiary structure of the N-terminal tandem PDZ domains of PSD-95

Inna S. Yanez Orozco¹, Frank A. Mindlin², Junyan Ma^{3,4}, Bo Wang¹, Brie Levesque², Matheu Spencer¹, Soheila Rezaei Adariani¹, George Hamilton¹, Feng Ding¹, Mark E. Bowen² & Hugo Sanabria^{1,4}

Previous studies of the N-terminal PDZ tandem from PSD-95 produced divergent models and failed to identify interdomain contacts stabilizing the structure. We used ensemble and single-molecule FRET along with replica-exchange molecular dynamics to fully characterize the energy landscape. Simulations and experiments identified two conformations: an open-like conformation with a small contact interface stabilized by salt bridges, and a closed-like conformation with a larger contact interface stabilized by surface-exposed hydrophobic residues. Both interfaces were confirmed experimentally. Proximity of interdomain contacts to the binding pockets may explain the observed coupling between conformation and binding. The low-energy barrier between conformations allows submillisecond dynamics, which were time-averaged in previous NMR and FRET studies. Moreover, the small contact interfaces were likely overridden by lattice contacts as crystal structures were rarely sampled in simulations. Our hybrid approach can identify transient interdomain interactions, which are abundant in multidomain proteins yet often obscured by dynamic averaging.

¹Department of Physics and Astronomy, Clemson University, Clemson, SC, USA. ²Department of Physiology and Biophysics, Stony Brook University, Stony Brook, NY, USA. ³Department of Chemistry, Clemson University, Clemson, SC, USA. ⁴Center for Optical Materials Science and Engineering Technology, Clemson, SC, USA. These authors contributed equally: Inna S. Yanez Orozco, Frank A. Mindlin. Correspondence and requests for materials should be addressed to F.D. (email: fding@clemson.edu) or to M.E.B. (email: mark.bowen@stonybrook.edu) or to H.S. (email: hsanabr@clemson.edu)

Intramolecular interactions within the primary amino acid sequence drives polypeptides to fold. High-affinity interactions, such as those forming the hydrophobic core, produce a relatively static conformation. Weak intramolecular interactions permit a dynamic ensemble of alternate conformations, which is difficult to predict and challenging to identify experimentally. Prediction of tertiary protein folding is now quite accurate for small protein domains. However, many proteins contain independently folded subdomains that subsequently assemble into a supertertiary structure¹. For the folding of multidomain proteins, the subdomains act as the primary sequence and their intramolecular interactions drive supertertiary folding. The same forces govern tertiary and supertertiary folding. However, the surface of folded subdomains is generally polar resulting in low affinity intramolecular interactions. Hence, the resulting supertertiary structure may sample different conformations with similar free energy over a broad range of timescales.

Structure determination of multidomain proteins remains a major challenge because of their dynamic and heterogeneous nature. Thus, many structural biology methods cannot describe supertertiary proteins. Moreover, the size of multidomain proteins, and the long dynamic timescales limit traditional all-atom molecular dynamics (MD) simulations. As such, few computational methods can predict supertertiary structure, even when all subdomains structures are known^{2,3}. Förster Resonance Energy Transfer (FRET) is not bound by these limitations so it can probe the structures of dynamic biomolecules *in vitro* and *in vivo*^{4–9}.

Here, we present an integrative approach to supertertiary structure determination combining simulations and FRET experiments applied to the postsynaptic density protein of 95 kDa (PSD-95, Fig. 1a), which is a prototypical dynamic, multidomain protein^{10–14}. PSD-95 contains five independently folded subdomains: three tandem PDZ domains, an SH3 domain and a Guanylate kinase-like domain, which are all involved in protein interactions at excitatory synapses^{15,16}. The structures of *all* five subdomains are known^{14,17,18}. Moreover, the PDZ domains of PSD-95 bind critical synaptic proteins such as ionotropic glutamate receptors, neuronal nitric oxide synthase¹⁹, the synaptic adhesion-protein neuroligin and synGAP, a GTPase linked to synaptic plasticity¹². Previously, we probed the supertertiary structure of PSD-95 with single-molecule FRET²⁰ revealing that the first two PDZ domains form a structurally independent supramodule²¹. There is particular interest in this PDZ1-PDZ2 tandem, which is the target of pharmaceutical compounds in clinical trials to treat ischemic stroke^{22,23}.

PDZ domains independently fold into a conserved tertiary structure¹⁶. They appear within multidomain proteins, often in tandem, with up to 13 PDZ domains in a single protein^{11,24}. The Protein Data Bank contains nearly 500 structures of PDZ domain but less than 30 are of tandem PDZ domains and none contain more than two domains. Thus, we lack information about the supertertiary PDZ interactions, which influence their function. PDZ domains typically interact with C-terminal peptides. As such, tandem PDZ domains are involved in multiprotein complex formation and often linked to signal transduction.

Studies of tandem PDZ domains revealed a range of interdomain affinities, which affect their supertertiary dynamics. For example, the Mint2 PDZ tandem was sufficiently dynamic that only the first domain of the tandem was seen as ordered within a protein crystal²⁵. In contrast, intramolecular interactions within Mint1 lock the PDZ tandem into a single conformation that results in autoinhibition of peptide binding²⁶. Thus, supertertiary interactions are an important regulatory mechanism for tandem PDZ domains.

NMR observed weak interdomain affinity within the PSD-95 PDZ tandem²⁷ while FRET detected no interactions between the

domains when unlinked²¹. Despite the weak interdomain interactions, NMR data suggested a “restrained conformation” in the unbound state, but identified no interdomain contacts to account for this restraint²⁷ (Fig. 1b). In contrast, peptide binding unlocked the PDZ tandem resulting in a “dramatic change of protein dynamics”²⁸. Thus, interdomain interactions appear incompatible with peptide binding suggesting the potential for autoinhibition of PSD-95. The PDZ tandem was also crystalized, which inherently suggests limited dynamics²⁹. The crystal contained two different conformations (Fig. 1b). However, both lacked interdomain contacts that could explain the restricted dynamics. Finally, single molecule Total Internal Reflection FRET Microscopy (smTIRFM) experiments also observed limited dynamics but pointed to a model for the mean-occupied conformation that similarly lacked interdomain contacts (Fig. 1b)²¹. However, the time resolution (10 Hz) was insufficient to identify limiting conformational states.

Crystallography²⁹, NMR²⁷, and smTIRFM²¹ all suggest limited dynamics in the PDZ tandem, but the experimentally observed conformations show large divergence (Fig. 1b). Importantly, none of the models account for the restrictions on mobility suggesting that a limiting state with the PDZ domains in contact remains unknown. Each of these studies captured a different snapshot, but none provided a complete picture of the energy landscape.

To resolve the elusive limiting conformational states and detect supertertiary dynamics, we revisited FRET experiments on the PSD-95 PDZ tandem with orders of magnitude faster in time resolution. Additionally, we used Discrete Molecular Dynamics (DMD) simulations to map the energy landscape and identify interdomain interactions that stabilize the PDZ tandem. Experiments and simulations identified two limiting conformations for the PDZ tandem that involve different interdomain interactions: an open-like (OL) conformation characterized by a small contact interface area of 440 Å² stabilized by interdomain salt bridges, and a closed-like (CL) conformation characterized by a larger contact interface with 701 Å² of buried surface area, which is stabilized by surface-exposed hydrophobic residues. DMD suggests a relatively low-energy barrier that permits interconversion between these states, which fall within a relatively limited conformational space. Multiparameter fluorescence detection (MFD) confirmed the presence of submillisecond dynamics. Using disulfide mapping, we confirmed the location of the CL interdomain interface while salt-bridge mutations confirmed the OL interdomain interface. Thus, our hybrid approach, combining simulations and experiments, was able to resolve the two closely related conformations that confounded previous structural studies.

Results

A network of FRET restraints to probe supertertiary dynamics.

We used a set of ten PDZ tandem variants with one unique cysteine in each PDZ domain (Fig. 1c), which were fluorescently labeled for FRET measurements. The labeling positions visually maximized the spatial distribution across the PDZ tandem within the constraints of surface accessibility and maintaining tertiary structural stability (Fig. 1d). The uncertainty in FRET-based structural models strongly depends on the number of restraints³⁰. Assuming the domains to be rigid bodies, ten variants should provide sufficient restraints to determine the supertertiary structure of the PDZ tandem.

To minimize double labeling with the donor, which complicates data interpretation, PDZ tandem variants were substoichiometrically labeled with donor. For a donor-only (DOnly) control sample, an aliquot of this reaction was quenched. The remaining sample was labeled with an excess of the acceptor

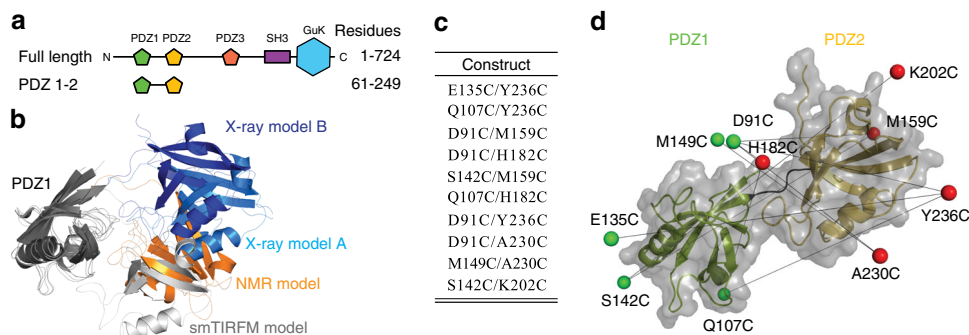


Fig. 1 The PDZ1-2 tandem from PSD-95. **a** PSD-95 contains five protein-interaction domains connected by linkers of varying length. The N-terminal PDZ tandem contains the first two PDZ domains (residues 61-249). **b** Structural models of the PDZ tandem determined with different experimental methods. PDZ1 (gray) from each model was aligned to highlight conformational differences. PDZ2 is colored according the method used to resolve the structure. Models from the crystal structure (blue and cyan) [PDB ID: 3GSL]. Model from NMR based on residual dipolar coupling, which was kindly provided by M. Zhang (orange). Model from smFRET (white). **c** Cysteine mutations used for fluorescent labeling in the PDZ tandem of PSD-95 (Residues 60-249). The first residue is in PDZ1 while the second is in PDZ2. Each pair forms a single distance restraint that is measured independently. **d** Network of FRET restraints in the PDZ Tandem. Mean position of each dye (spheres) at the associated cysteine residue as determined by modeling. PDZ1 sites colored green. PDZ2 sites colored red. Black lines are shown between mean positions of dyes

(Alexa Fluor 647). Thus, the resulting samples are randomly labeled with a ~1:4 ratio of donor and acceptor fluorophores. However, the doubly labeled samples still contained a fraction of apparent DOnly protein, presumably due to the presence of an inactive acceptor³¹.

Identification of limiting conformational states. By exciting the dyes with picosecond laser pulses, ensemble time correlated single photon counting (eTCSPC) can detect structural states that persist on the timescale of fluorescence emission (ns). Interdomain motions are much slower (μ s-ms) so the PDZ tandem is essentially static during a round of fluorescence emission. The flexibly tethered dyes move at timescales slightly faster than the fluorescence emission, which uncouples motion of the dyes from the tandem³². We measured fluorescence lifetime decays for the 10 FRET variants and the 10 corresponding DOnly controls (Fig. 2a and Supplementary Fig. 1). To describe the donor de-excitation, we compared fitting the DOnly controls with a model containing one- or two-exponential decays (Supplementary Eq. 1). The weighted residuals and the autocorrelation of the residuals improved with two-exponential decays (Fig. 2b). We used the DOnly models to fit the corresponding DA sample under the assumption that all donor states are quenched equally by FRET. Each conformation has an associated rate of energy transfer (k_{RET}) so the full decay is a superposition of the conformational states that are present (Supplementary Eqs. 3 & 4). Thus, eTCSPC results provide the number of conformations along with a model of the interdye distances for each conformation^{5,33}.

To arrive at the correct number of conformations, we fit eTCSPC decays individually with an increasing number of states, and obtained the best results with three conformational states. The model also included a no FRET state to account for inactive acceptor molecules³¹. Thus, each DA sample yielded three interdye distances that arose from each of the conformational states. However, to use these FRET distances for modeling, it is necessary to assign each distance to its conformational state, which is typically done by population assignment. In this case, the states were distributed somewhat uniformly, which complicated the assignment. To ensure a consistent population distribution across all FRET samples and a proper assignment of states, we implemented a global fit for all ten FRET variants with a shared population distribution across the three states. The DA

fluorescence decays were best described by three states based on improvements in the residuals (Fig. 2c and Supplementary Fig. 1). Moreover, the figure of merit (χ^2) improved from an average of 1.26 to 1.15 (Supplementary Fig. 1K). For four variants, the improvement in χ^2 was above 10% (Fig. 2d). The improvement in χ^2 with three states was statistically significant based on the F-test (Supplementary Eq. 15)^{34,35} with a confidence interval greater than 99.9%.

With the global fit, we could simultaneously describe all 10 FRET decays with a unified, unbiased model. The most-populated state accounts for 43.5% of the population with 32.4% in the second-most-populated state and 24.1% in the least-populated state (Supplementary Table 1). Thus, the PDZ tandem samples three conformational states that give rise to different interdye distances.

The two main sources of uncertainty in the reported distances are: (i) the statistical uncertainty of fitting, and (ii) the uncertainty regarding the orientation of the dyes (κ^2), which is an important parameter for converting FRET to distance. We estimated κ^2 uncertainty from the anisotropy decays of the DOnly and DA samples including the directly excited acceptor and acceptor-sensitized emission (Supplementary Fig. 2). The κ^2 distribution was calculated with a wobbling-in-cone model³⁶ and compared to the standard assumption of $\langle \kappa^2 \rangle = 2/3$ (Supplementary Fig. 3). We used these sources of error to determine the accuracy of our reported distances (Supplementary Eq. 16; Supplementary Table 1).

Experimental confirmation of dynamic averaging. To detect dynamic transitions between the states identified by eTCSPC, we used MFD, which simultaneously records the average fluorescence lifetime, intensity and anisotropy of the photons emitted by each molecule. The conformational changes that occur while the molecule diffuses through the confocal excitation volume affect the fluorescence lifetime and intensity measurements differently. If a molecule samples multiple states, the FRET intensity indicator (F_D/F_A) and the average fluorescence lifetime ($\langle \tau_{D(A)} \rangle$) would differentially reflect changes to the mean and variance of the photon counts, respectively. This shifts the peak in the 2D histogram from the expected value for a static molecule in a single state^{37,38}. This analysis is similar to mean-variance histograms in patch clamp electrophysiology³⁹. Thus, plotting F_D/F_A against

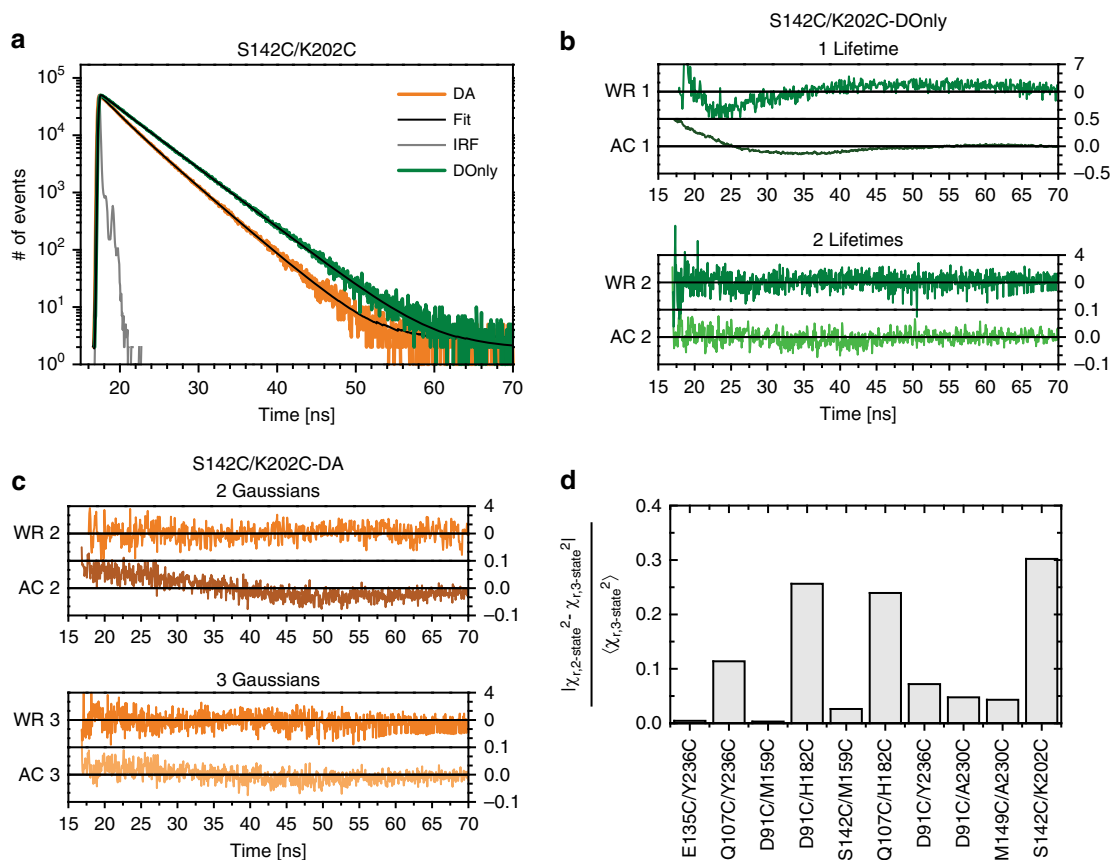


Fig. 2 Ensemble time correlated single photon counting (eTCSPC) to identify limiting states. Representative data are shown for the variant S142C/K202C. The remaining variants are shown in Supplementary Fig. 1. **a** Time-resolved fluorescence decays for the Donor-only control (DOnly, green) and Donor-Acceptor FRET sample (DA, orange). Instrument response function is shown in gray. DA decays were fit using a global analysis of all ten variants (black lines). **b** Weighted residuals (WR, green) and autocorrelation (AC, lighter green) for fitting the DOnly decays with one (top panel) and two (bottom panel) fluorescence lifetimes. **c** Weighted residuals (WR, orange) and autocorrelation (AC, brown and cream) for the global fit using either two (top panel) or three (bottom panel) Gaussian-distributed states (Supplementary Table 1). **d** Improvement in χ^2 when fitting with two or three Gaussian-distributed states. The relative change of χ^2 is shown for each FRET variant (labels beneath panel)

$\langle \tau_{D(A)} \rangle_f$ per single-molecule event is a useful tool to identify dynamics³⁷.

For all 10 variants, we plotted F_D/F_A against $\langle \tau_{D(A)} \rangle_f$ as a 2D frequency histogram. The separate 1D frequency histograms for F_D/F_A and $\langle \tau_{D(A)} \rangle_f$ are aside and atop the main panel, respectively. Figure 3a–c shows three representative examples (A Q107C/H182C-DA, B D91C/M159C-DA, and C S142C/K202C-DA). For most variants, the frequency histograms appear as a single asymmetric population distribution (Supplementary Fig. 4).

We include guidelines to help interpret these frequency histograms. First, we calculated the expected F_D/F_A for the three states identified by eTCSPC (purple, blue, and cyan lines in Fig. 3a–c). Next, we calculated the expected relationship between F_D/F_A and $\langle \tau_{D(A)} \rangle_f$ for molecules with no conformational dynamics (Fig. 3a–c red “static FRET” lines). Finally, we calculated the expected relationship between F_D/F_A and $\langle \tau_{D(A)} \rangle_f$ for molecules undergoing transitions between states (Fig. 3a–c green “dynamic FRET” lines). Consequently, PDZ tandem molecules in a stable conformation will fall along the red static-FRET line while molecules undergoing conformational dynamics as they pass through the confocal excitation volume will fall off this line.

Given these guidelines, we observe that the Q107C/H182C variant is close to mid FRET (Fig. 3a) while D91C/M159C is

centered at high FRET (low F_D/F_A , Fig. 3b), and S142C/K202C-DA shows low FRET (high F_D/F_A , Fig. 3c) consistent with previous measurements²¹. In these cases, the equilibrium lays closer to the most-populated state identified from eTCSPC (Fig. 3a–c, purple lines). Moreover, the population maxima for most variants fell within the area enclosed by the dynamic FRET lines indicating submillisecond dynamic averaging (Supplementary Fig. 4). It is important to note that at the time resolution of MFD, the limiting states from eTCSPC do not appear as discrete peaks due to averaging.

To confirm the presence of dynamic averaging, we varied the time binning of the data and used Photon Distribution Analysis (PDA)³⁸ to model the F_D/F_A distributions. For static molecules, the model function and χ^2 would be unaffected by changing the time binning. In contrast, we observed that χ^2 changed significantly when the time-binning window changed, which confirms dynamic averaging (Supplementary Fig. 5). Even our shortest time binning was unable to resolve the distribution into its underlying limiting states. This suggests that dynamics occur on timescales faster than the diffusion time, where time-binning analysis is optimal.

DMD simulations of the PDZ tandem. To produce an unbiased representation of the conformational landscape, we performed MD simulations, which we compared with the experimental

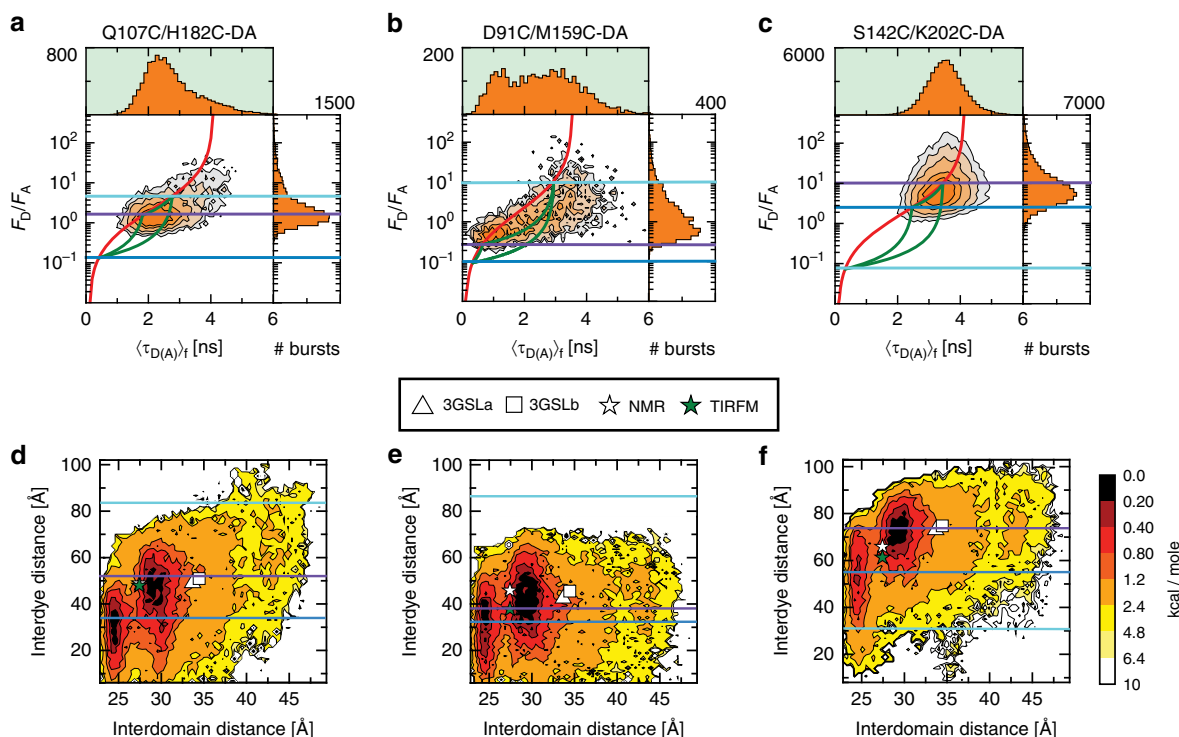


Fig. 3 Multiparameter Fluorescence Detection of FRET and energy landscape of the PDZ1-2 tandem variants. Each panel plots two FRET indicators for each single-molecule event. The donor over acceptor fluorescence ratio (F_D/F_A) on the y-axis and the average fluorescence lifetime ($\langle\tau_{D(A)}\rangle_t$) on the x-axis. Shown are the 2D histograms (orange contours) for (a) Q107C/H182C-DA, (b) D91C/M159C-DA, and (c) S142C/K202C-DA. The 1D frequency histograms are shown along the axes for each FRET indicator. The static FRET lines (Supplementary Eq. 10, Supplementary Table 3) are shown in red. The dynamic FRET lines (Supplementary Eq. 11, Supplementary Table 4) are shown in green and connect the states identified with eTCSPC. The mean interdyer distances for each of the eTCSPC states was converted to F_D/F_A and shown as lines colored purple, blue, and cyan in decreasing order of population occupancy. **d–f** Histograms of the potential mean force (PMF) calculated from DMD simulations for the variants (d) Q107C/H182C-DA, (e) D91C/M159C-DA, and (f) S142C/K202C-DA. The remaining histograms are shown in Supplementary Fig. 6. Histograms show the interdomain distance between the center of mass of the two PDZ domains plotted against the interdyer distance distribution, which was calculated as $p(R_{DA})_{PMF} = \exp(-E_{PMF}/k_B T)$ where E_{PMF} is the integrated free energy over interdomain distance as the function of the interdyer distance, R_{DA} ; and k_B is Boltzmann constant. Contours are scaled in kcal/mole. Lines represent the mean interdyer distances for each of the states identified by eTCSPC and are colored purple, blue, and cyan in decreasing order of population occupancy. Interdomain distances for previously reported models of the PDZ tandem are marked by symbols and denoted beneath the panels. Two basins are clearly identified, with residual states that are heterogeneously distributed with respect to the interdomain distance

results. Given the size of the PDZ tandem and the relatively long timescale associated with domain motions, we used the atomistic DMD⁴⁰. DMD is a rapid MD algorithm that has been benchmarked with ab initio protein folding⁴¹ and used to engineer de novo allosterically regulated proteins⁴² by capturing their coupled conformational dynamics.

As the starting conformation, we used the smFRET model²¹ and included the fluorescent dyes at all labeling sites to relate the simulations back to experiments. Importantly, the dyes were rendered non-interactive so as not to affect the simulation. To sufficiently sample conformational dynamics in the PDZ tandem, we performed replica-exchange DMD simulations with a cumulative total simulation time of $\sim 1.6 \mu\text{s}$. A previous benchmark study indicated that protein dynamics in DMD simulations with implicit solvent is two or three orders of magnitude faster than observed in experiments⁴³. Hence, with efficient sampling enabled by DMD and the replica-exchange, we expected to observe dynamics beyond μs in experimental time. We applied the weighted histogram analysis method (WHAM)⁴⁴ to analyze the conformational distribution from the replica-exchange trajectories, which give estimates of the potential mean force (PMF) at room temperature (300 K). We computed the PMF as a function of the interdomain distance between centers of mass of

the two PDZ domains (x-axis, Fig. 3d–f and Supplementary Fig. 6). To facilitate a comparison between DMD simulations and experimental distances, we also computed the PMF as a function of the interdyer distance for each FRET variant (y-axis, Fig. 3d–f and Supplementary Fig. 6). These two representations of the PMF are plotted against each other as a contour histogram in units of kcal/mole.

In this representation, we observed two distinct minima in the PMF energy landscape corresponding to two distinct conformational states along with a broad sampling of alternate conformations with larger interdomain distances. By projecting the free energy surface onto various reaction coordinates, we estimated the apparent energy barrier to be only a few kcal/mole (Fig. 3d–f and Supplementary Fig. 7 & 8), which is consistent with the dynamic averaging observed in MFD (Supplementary Fig. 4). Although the two conformations are reasonably separated in terms of interdomain distance, the associated interdyer distributions are substantially overlapped for many variants (Supplementary Fig. 6). Both the simulated interdyer distances ($p(R_{DA})_{PMF}$) and, for most variants, the experimentally measured distances displayed a relatively unimodal distance distribution that did not resolve the two conformations into separate peaks (Supplementary Fig. 9). Thus, several variants were not sensitive to the conformational differences.

Interdomain contacts that stabilize the PDZ tandem. The published structural models were unable to identify the interdomain contacts that stabilize the PDZ tandem. For the two identified minima, we selected a representative structure with the smallest average root mean squared displacement (RMSD) to the entire ensemble (Fig. 4a, b, respectively). The RMSD between the two representative models is 10.4 Å, which supports these being distinct conformations. The CL state model showed a buried surface area of 701 Å². In contrast, the OL state model showed a smaller contact interface of 440 Å². It should be noted that these representative structures fall within a rather shallow basin in the energy landscape. As such, these minimal surfaces do not represent the entire contact area for each conformational ensemble.

For each representative structure from the two basins identified in the energy landscape (Fig. 3d–f), we computed the frequency map of pair-wise residue contacts between the PDZ domains and plotted the location of persistent contacts (Fig. 4c–d). We found that the closed-like (CL) state is primarily stabilized by interactions between surface-exposed hydrophobic residues (highlighted next to the axes, Figure 4c). In contrast, the open-like (OL) state involved primarily electrostatic interactions between oppositely charged residues in PDZ1 and PDZ2 (Fig. 4d). An overlay of the two models highlights the rotation of PDZ2 with respect to PDZ1 that accompanies the transition between states (Supplementary Fig. 10).

Having identified the interdomain contacts stabilizing each state, we computed the fraction of native contacts for each state (Q_{CL} and Q_{OL}) and used these Q -values^{45,46} to represent the energy landscape at the room temperature. By tracking the formation of these interfaces, we observed multiple interconversions between states during our simulations (Supplementary Fig. 7). Moreover, Q -values showed that the two states occupy distinct basins in the energy landscape (Supplementary Fig. 7). Finally, the states are mutually exclusive as interdomain contacts in either state must break to transition to the other state (Supplementary Fig. 8).

Comparison of simulations to the experimental distances. To compare DMD and eTCSPC, we added guidelines for distances of the three states from eTCSPC to the PMF plots from DMD (Fig. 3d–f and Supplementary Fig. 9). This shows that the most-populated and second-most populated states from eTCSPC generally correspond to the OL and CL states from DMD, respectively. The least-populated eTCSPC state, which has no direct corollary in DMD simulations, sampled a heterogeneous ensemble of states outside the predominant minima (Supplementary Fig. 8). Although DMD does not identify a discrete third state, we note that fitting the eTCSPC to a two-state model failed to provide positive correlations with the interdy distances from DMD simulations (Supplementary Fig. 10A). We conclude that the least-populated state (~24% of molecules) could encompass

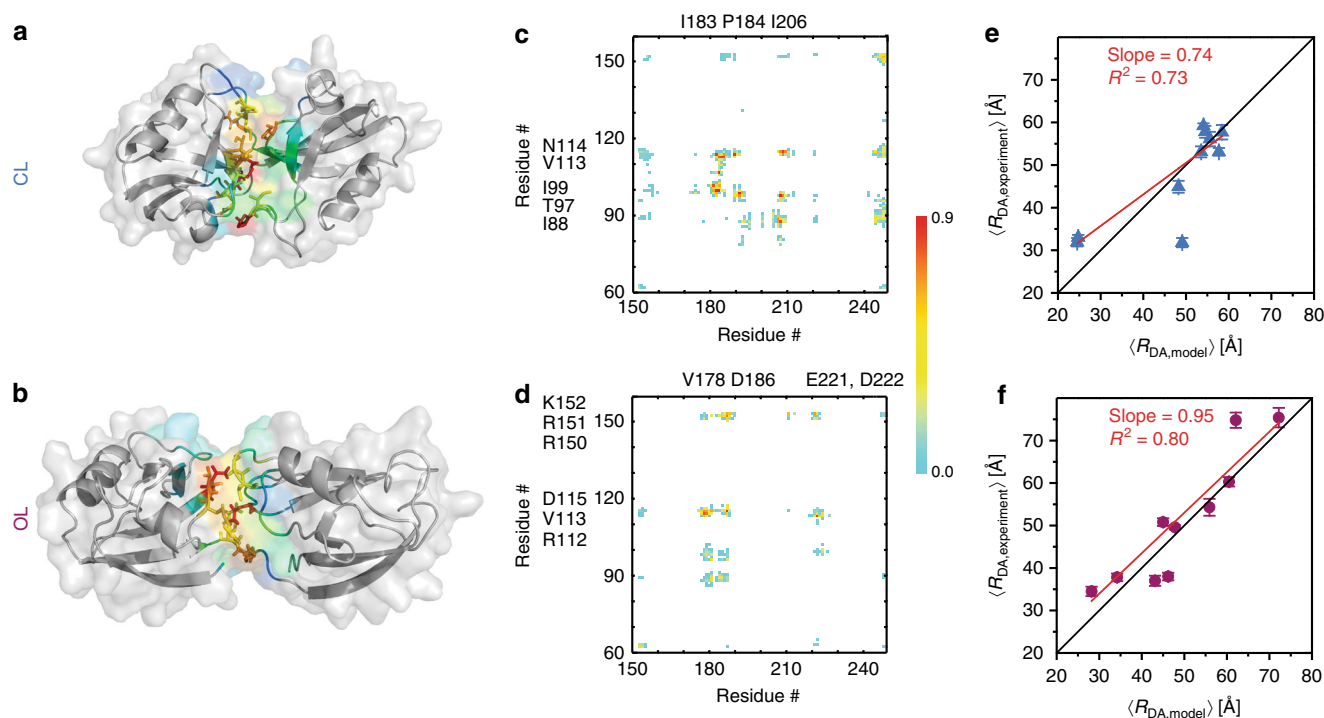


Fig. 4 Representative structures, contact maps and correlation between experimental and DMD simulated interdy distances. **a** Cartoon representation for the closed-like (CL) state, which has the shorter interdomain distance and a larger contact interface. PDZ domains are shown in cartoon representation with a transparent surface. Interfacial residues are colored according to their contact frequency (cyan to red denoting low to high frequency). **b** Representative structure for the open-like (OL) state with increased interdomain separation and smaller contact interface. Rendered and colored as in panel **a**. **c** The interdomain contact frequency map per residue for the conformational ensembles of the CL state. Contact frequency uses same color palette as in panels **A** and **B**; no contact white, low contact probability cyan and high contact red. **d** The interdomain contact frequency map for the OL state. The color palette is the same as in **(c)**. Residues with the highest interdomain contact frequencies are highlighted along the axis for each PDZ domain. **e, f** Measured interdy distances from eTCSPC (y-axis) plotted against the simulated interdy distances for the representative models shown in **(a)** and **(b)** (x-axis). **e** Distances for the CL state plotted against those for the second-most populated state in eTCSPC. **f** Distances for the OL state plotted against those for the most populated state in eTCSPC. Data was fit to a line to extract the slope and correlation. Error bars in **(e)** and **(f)** correspond to the uncertainty on the measured distances shown in Supplementary Table 1 using an error propagation rule (Supplementary Eq. 16)

the wide range of outlying DMD conformations (~20% of the simulation time). The presence of outlying DMD states is readily apparent from the *Q*-value distributions, which reveals a population with *Q*-values corresponding to neither state (Supplementary Fig. 8B). This DMD “population” displayed a highly variable interdomain distance with no distinct minimum on the PMF histograms; hence, we are unable to reliably select a representative example for comparison with experiments.

To compare the experiments and simulations, we correlated the interdye distances from the DMD models to those from eTCSPC. The CL state showed a positive correlation indicating good agreement with an R^2 of 0.73 and a slope of 0.74 (Fig. 4e). The OL state showed an R^2 of 0.80 and slope of 0.95 representing even better agreement between experiment and simulation (Fig. 4f). The distribution of distances from DMD and MFD did show differences (Supplementary Fig. 9). This is not unexpected as variance in DMD arises from conformational sampling, while variance in MFD depends on signal to noise and the rate of conformational exchange. However, the mean interdye distances from MFD also correlated well with mean interdye distances from DMD with $R^2 = 0.80$ (Supplementary Fig. 9K).

Disulfide mapping to probe the predicted contact interface. To validate our model for the CL state, we monitored the rate of disulfide formation between engineered cysteine residues at different positions within PDZ1 and PDZ2. Disulfide bond formation is distance dependent so the rate is determined by structural proximity⁴⁷. Therefore, we introduced a disulfide pair at the predicted contact interface, A106C/P184C (Fig. 5a, orange). These residues are ~5 Å apart and should have the highest rate of

disulfide formation. The proline is in a loop region so the mutation was well tolerated. As a positive control, we selected a FRET variant with the high FRET efficiency, Q107C/H182C (Fig. 5a, blue). These residues are ~15 Å apart so these residues should have reduced disulfide formation. As a negative control, we chose a FRET variant with low FRET efficiency, S142C/K202C (Fig. 5a, red). These residues are over 40 Å away and should form disulfides poorly.

We initiated disulfide formation with copper (II) phenanthroline followed by non-reducing SDS-PAGE, which revealed increased mobility upon disulfide formation (Fig. 5b). Additionally, a small amount of higher order disulfide formation occurred (Supplementary Fig. 11). For each variant, we measured the disulfide formation in triplicate (Fig. 5c). Each reaction was well fit to a single exponential function to obtain the initial and final extent of disulfide formation along with the reaction rate for each variant. Analysis of the extent of disulfide formation revealed that less than 5% of the control samples had formed disulfides before initiating the reaction, while the A106C/P184C mutant showed $11.6 \pm 0.08\%$ disulfide formation before the reaction had started (Fig. 5d). The negative control showed no significant disulfide formation during the reaction (initial $3.37 \pm 0.32\%$ and final $3.69 \pm 5.06\%$) (Fig. 5d). The positive control showed $46.94 \pm 2.83\%$ disulfide formation. However, the mutant designed to be at the CL interface (A106C/P184C), showed the highest extent of disulfide formation ($84.58 \pm 0.01\%$). The differences in the extent of disulfide formation were mirrored by differences in reaction rates (Fig. 5e). Although the negative control showed minimal disulfide bond formation, we estimated the reaction rate to be $5.27 \pm 2.31\%/min$. The positive control formed disulfides 2.5 fold

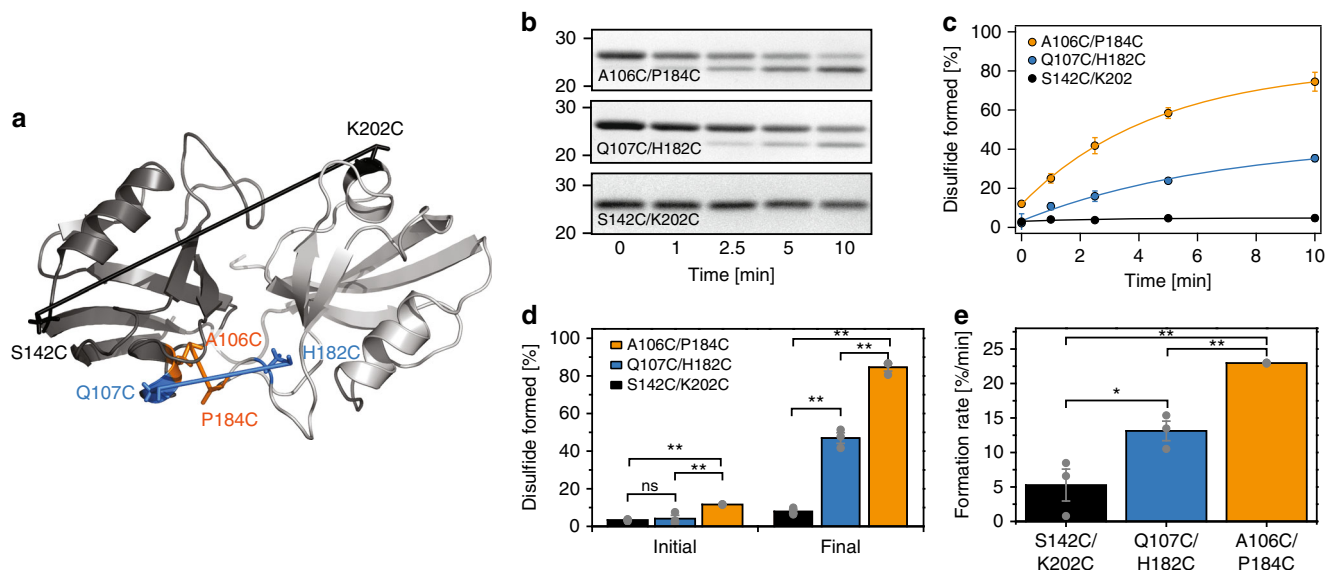


Fig. 5 Disulfide mapping of the interdomain contact interface in the PDZ tandem. **a** Location of the engineered cysteine pairs used to test the contact interface. Cartoon representation of the CL state of the PDZ tandem. PDZ1, gray. PDZ2, white. Cysteine side chains are shown as lines with the pairs connected by their sulfur atoms. A106C/P184C, orange. Q107C/H182C, blue. S142C/K202C, black. **b** Representative SDS-PAGE showing individual time points in the disulfide bond formation reaction used for the analysis. Disulfide formation was initiated by shifting the samples to oxidizing conditions. The location of the introduced cysteine residues for each mutant is shown within each panel. Duration of the reaction for each time point is indicated below. The full gels are shown in Supplementary Fig. 11. **c** Kinetic analysis of the disulfide formation reactions. A106C/P184C, orange. Q107C/H182C, blue. S142C/K202C, black. Gel band intensities were measured in triplicate to obtain the percentage of disulfide for each time point. Each time course was fit with an exponential function to obtain the initial and final percentage of disulfide bonding along with the rate of disulfide formation. **d** Extent of disulfide formation for each mutant taken from the fits to panel B. The initial percentage (left) showed the presence of disulfide before the oxidation reaction. The final percentage showed the amount of disulfide formed after the oxidation reaction has gone to completion. ** indicates $p > 0.01$; ns stands for no significance in an unpaired student *t*-test. **e** Disulfide formation rate for each mutant as taken from the fits to panel b. ** indicates $p > 0.002$. * indicates $p > 0.05$ in an unpaired student *t*-test. For all graphs, error bars indicate SEM for three replicate measurements. Individual values are shown as dots

faster ($13.13 \pm 1.4\%/min$). Thus, the disulfide formation rates correlate with differences in FRET efficiency. In contrast, the A106C-P184C mutant formed disulfides 4.4 fold faster ($22.97 \pm 0.04\%/min$) and 1.7 fold faster than the positive control. Thus, the extent and rate of disulfide formation agree with predictions of proximity based on our CL state model. This confirms the predicted interdomain contact interface.

To probe the OL state, we engineered a salt bridge by introducing the mutations: V113K, N114K, V223D, and M224E into the variant Q107C/H182C (Supplementary Fig. 12). V113 and N114 are within the OL interface. Replacing these uncharged residues with lysines creates an electropositive patch on PDZ1. In contrast, replacing V223 and M224 with aspartate and glutamate creates an electronegative patch on PDZ2. V223D and M224E are not within the OL interface, but are close enough that the electrostatic interactions should disfavor the CL state. By measuring smFRET in MFD (Supplementary Fig. 12B-C), we observed a shift towards higher F_D/F_A (i.e., lower FRET) indicating a larger interdyer distance in the “salt-bridge” variant compared to the control. The domain rotation needed to bring complementary charges together would move these labeling sites farther apart. Thus, the observed FRET shift is consistent with the expected motion. Thus, by adding additional salt bridges to the ionic interactions in the OL state, we observe a clear redistribution of FRET.

In summary, by using a hybrid approach combining eTCSPC, DMD, and MFD, we provide a self-consistent view of the conformation and dynamics of the PDZ tandem, which we independently verified by experimentally probing the predicted interdomain contacts.

Discussion

The PDZ1-2 tandem from PSD-95 has been extensively studied but none of the reported structures identified any interdomain contacts that explain the limited dynamics observed in these reports (Fig. 1). Here, we used replica-exchange DMD simulations, totaling 1.6 μ s, to identify two predominant low-energy basins in the conformational landscape, which correspond to two limiting state structures that involve different interdomain contacts (Fig. 4). Our eTCSPC analysis identified three states (Fig. 2) with the two predominant states corresponding to those from DMD with a correlation coefficient above 73% (Fig. 4). MFD analysis confirmed the presence of dynamic averaging at timescales faster than milliseconds (Fig. 3a-c). Moreover, time window analysis showed significant changes in the figure of merit at different time binning (Supplementary Fig. 5). This is a clear signature of dynamic averaging at timescales faster than milliseconds. Therefore, MD simulations and experiments agree that the PDZ tandem is dynamic and samples two low-energy conformations that are similar but not identical.

The coexistence of two similar conformational states revealed the difficulty in designing a network of FRET distance restraints. Including the dyes in DMD simulations revealed that several restraints showed similar interdyer distances for the two conformations. This lack of conformational sensitivity for some variants explains why treating measurements separately was less successful than a global fit at consistent assignment of state distributions. The two most populated states, which agree with DMD, account for 75.9% of the FRET observations. The remaining 24.1% show significant heterogeneity among variants and, in some cases, correspond to distances beyond those sampled by DMD. Despite these limitations, global treatment of the ten fluorescence decays provided a unified description of the two predominant conformational states.

Moreover, we validated both limiting states by (i) engineering a disulfide bridge in the interface of the CL state (Fig. 5) and (ii) enhancing the salt-bridge interactions that stabilize the OL state. Our measurements of these variants were in agreement with the expectations based on DMD (Fig. 5 and Supplementary 11). Thus, we independently corroborated the proposed structural models, which sets a new standard for accuracy in FRET-based protein modeling.

To compare previous structural models of the PDZ1-2 tandem with DMD, we calculated the interdomain distances for the two crystallographic models²⁹, the representative NMR model²⁷ and the smTIRFM model²¹. The interdyer distances were estimated from the accessible volume for dyes in each structure. We mapped these interdomain and interdyer distances on to the PMF histograms from DMD (Fig. 3d-f). For the crystallographic models, the interdomain distances (33.7 Å and 34.5 Å for 3GSLa and 3GSLb, respectively) are larger than the interdomain separation observed in DMD simulations (Fig. 3a-c). Thus, the crystal structures were infrequently sampled in simulations. The extensive crystal lattice interactions may have overridden the stabilization provided by the weaker intramolecular interface. For the NMR model, the interdomain distance falls between the two states and closer to the OL state, which was the most populated conformation in eTCSPC. The same is also true for the smTIRFM model, which was between the two states from DMD and similarly closer to the OL state. This is understandable given that the measurement timescale was slower than the submillisecond dynamic averaging observed in MFD. Thus, the dynamic exchange between two closely related conformations complicated the derivation of structural models by NMR and smTIRFM.

NMR⁴⁸ and SAXS²³ experiments showed that peptide binding to the PDZ tandem increased both interdomain flexibility and interdomain separation such that the defined orientation between PDZ domains was lost. Thus, peptide binding may be incompatible with the CL conformations we observed. Neither of the peptide binding sites in the PDZ tandem are sterically occluded in either state. However, the binding affinity of PDZ domains is known to be regulated through allosteric interactions⁴⁹. Such allosteric networks have been identified operating in PDZ2 from PSD-95⁵⁰. Moreover, ligand binding to PDZ domains is coupled to widespread changes in sidechain dynamics such that interactions distal from the binding pocket can affect affinity⁵¹. Closer examination of the two limiting states reveals that the contact interfaces involve residues that surround the ligand binding sites. PDZ domains bind ligands in the β 2- α 2 groove near the long α 2-helix. The CL state interface involves residues in the β 2- β 3 loops of both domains, which are near the exit of the ligand-binding pocket. In the OL state, the β 2- β 3 loop and the β 5 sheet, which abuts α 2, interact with the interdomain linker and with β 4 on the backside of PDZ1. Such interactions between elements around the binding sites could easily modulate affinity resulting in autoinhibition. Validation of autoinhibition in either state remains for future studies.

We combined ensemble and single-molecule FRET experiments with high temporal resolution and DMD simulations to describe the supertertiary structure and dynamics of the PDZ1-2 tandem from PSD-95. DMD simulations observed a heterogeneous free energy landscape with two clear minima. These two conformations were confirmed by eTCSPC while the temporal resolution of MFD showed that the PDZ1-2 tandem is dynamically exchanging. The upper limit of the dynamic timescale is set by time binning in MFD experiments, with a mean diffusion time of \sim 1 ms. Such fast interconversion is consistent with a low free energy barrier separating the two states as suggested by simulations. This suggests weak interdomain interactions, which is consistent with the relatively small contact areas observed for

each state. While neither state was very stable, it is likely that the presence of two nearby contact states minimizes the time spent in unrestrained conformations associated with peptide binding. Our results explain the highly divergent models from previous structural studies.

Finally, our approach is uniquely positioned to study weak and transient interdomain interactions. Although often hidden from biophysical characterization, they are highly abundant in multidomain proteins such as poly-PDZ proteins¹⁴, ion channels⁵², ATPases⁵³, enzymes⁵⁴, and intrinsically disordered proteins with dynamic supertertiary structures¹. When identified, these transient and short-lived conformational states could serve as potential targets for drug development and may reveal novel mechanisms of action^{55–57}.

Methods

Protein expression and purification. The PDZ1-2 tandem from *Rattus norvegicus* PSD-95 (residues 61–249) was expressed from pPROEX HTB in the BL21 *E. coli* strain induced with 0.5 mM IPTG for 2 h at 30 °C. Proteins were purified by a Ni-affinity and eluted with 250 mM imidazole. The 6 His tag was removed by TEV protease cleavage during dialysis into 20 mM Tris pH 7.4 100 mM NaCl 1 mM DTT 1 mM EDTA followed by anion exchange using HiTrap Q columns (GE Healthcare). Finally, proteins were purified by size exclusion chromatography on Superdex 75 (GE Healthcare)²¹. Proteins were first labeled with a 1:2 ratio of Alexa 488 C₅ maleimide for 1 h at 4 °C followed by addition of a 5:1 molar ratio of Alexa Fluor 647 C₂ maleimide, which was reacted overnight at 4 °C. Unconjugated dye was removed by desalting with Sephadex G50 (GE Healthcare) followed by dialysis.

Ensemble time correlated single photon counting. Ensemble time-correlated single-photon-counting (eTCSPC) measurements were performed using a Fluorolog3 spectrofluorometer in T format with a PDX detector and Horiba Yvon photon system (Horiba, USA). The excitation sources were a pulsed 485 nm and a 635 nm NanoLED diode lasers (Horiba Yvon, USA) operating at 1 MHz, for donor and acceptor, respectively. The emission monochromator was set to the maximum emission wavelength of the fluorophore being measured, while the slit was set at a maximum of 10 nm in order to get an intensity count rate lower than 2% of the repetition rate. Fluorescence decay analysis is done using ChiSurf written in Python following the algorithm described in Supplementary Methods.

Multiparameter fluorescence detection. Freely diffusing molecules in pM concentration were excited as they passed through the focal volume of a 60 × 1.2 NA collar (0.17) corrected Olympus objective. Pulsed interleaved excitation (PIE)⁵⁸ with diode lasers at 485 nm and 640 nm (PicoQuant, Germany) was operated at 40 MHz with 25 ns interleaved time. The power at the objective was 120 μW at 485 nm and 39 μW at 640 nm. Emitted photons were collected through the same objective and spatially filtered through a 70 μm pinhole to limit the effective confocal detection volume.

Emission was separated into parallel and perpendicular polarization components at two different spectral windows using band pass filters ET525/50 and ET720/150 (Chroma Technology Co.) for donor and acceptor, respectively. Four photon-detectors are used—two for donor (PMA Hybrid model 40 PicoQuant, Germany) and two for acceptor (PMA Hybrid model 50, PicoQuant, Germany). To ensure temporal data registration of the four synchronized input channels, we used a HydraHarp 400 TCSPC module (PicoQuant, Germany) in Time-Tagged Time-Resolved mode.

Labeled samples were diluted to pM concentration in PBS buffer (50 mM sodium phosphate, pH 7.5, 150 mM NaCl, 40 μM TROLOX), which had been charcoal filtered to remove residual impurities. At pM concentrations, we observe ~1 molecule per second in the focal volume. Samples were imaged in NUNC chambers (Lab-Tek, Thermo Scientific, Germany) that were pre-coated with a solution of 0.01% Tween 20 (Thermo Scientific) in water for 30 min to minimize surface adsorption. We obtained the instrument response function (IRF) by measuring water while protein-free buffer was used for background subtraction. Calibration experiments and details on data collection were recently reported^{59,60}.

DMD simulations. A detailed description of the DMD algorithm can be found elsewhere^{40,43}. Briefly, proteins were represented by the united-atom model containing all heavy atoms and polar hydrogens with an implicit solvent model. The step-wise potential functions were obtained by mimicking the continuous interatomic interaction potentials in the molecular mechanics-based Medusa force field⁶¹. The Medusa force field has been shown to be effective in accurate prediction of protein stability changes upon mutations for a large set of experimental data⁶². The bonded interactions include covalent bonds, bond angles, and dihedrals. The interatomic interactions include van der Waals, solvation, hydrogen bond, and electrostatics. The solvation energy is modeled by the Lazaridis-Karplus implicit solvent model⁶³. An implicit solvent approach is appropriate for studying

long timescale dynamics of biomolecules due to the separation from the fast water dynamics. Screened electrostatic interactions are computed by the Debye-Hückel approximation. A Debye length of 1 nm is used by assuming a water dielectric constant of 80 and a monovalent electrolyte concentration of 0.1 M. The Anderson's thermostat is used to maintain constant temperature and a periodic boundary condition is applied. Because they follow the same physical laws, the dynamics observed in DMD are equivalent to continuous potential MD at timescales longer than picoseconds with differences mainly at short timescales within the sub-picosecond range (i.e., the average time step between two consecutive interatomic collisions where a potential energy step is encountered).

DMD simulations included the fluorescent dyes, which were coupled to the corresponding cysteine residues as in experiments. The atoms in the linker were rendered non-interacting by setting a very small hardcore distance with all other atoms (0.001 Å). The atoms within the dye moiety have a hard-sphere interaction with other atoms of the dye through a hardcore distance of 3 Å, while the dyes themselves are effectively non-interactive through a small hardcore distance of 0.001 Å.

We performed replica-exchange DMD simulations with eight replicas at different temperatures: 275, 287, 300, 315, 330, 345, 360 and 375 K. The exchange between replicas with neighboring temperatures was carried out every 50 ps. Each replica lasted 200 ns with an accumulative simulation time of 1.6 μs. The PMF was calculated by the WHAM using the last 175 ns from each of the eight simulations. The DMD program is available online (www.moleculesinaction.com).

Disulfide mapping in the PDZ1-2 tandem of PSD-95. Proteins were purified under reducing conditions. Immediately before the reaction, proteins were fully reduced by incubation in fresh 5 mM DTT for 1 h at 4 °C followed by desalting into non-reducing conditions (20 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA). Disulfide oxidation reactions were performed using a protein concentration of 2 μM at 25 °C. Disulfide formation was initiated by the addition of 0.5 mM CuSO₄ and 1.75 mM 1, 10-phenanthroline. Time points were quenched by adding 40 mM N-ethylmaleimide to alkylate unbound cysteines and 10 mM EDTA followed immediately by boiling at 95 °C for 5 min in non-reducing Laemmli sample buffer⁴⁷. Samples were run on 15% SDS-PAGE. All experiments were carried out in triplicate. Intensities for both the native and shifted bands were measured in ImageJ. Percentages of disulfide formation were calculated for each time point and corrected for the presence of higher order oligomers. Each reaction was well fit to a single exponential function to obtain the initial and final extent of disulfide formation along with the reaction rate for each mutant. Replicates were analyzed separately to obtain the average and standard error of the mean (SEM) as well as to estimate the error in the fitted parameters.

Code availability. MFD is made available at <http://www.mpc.hhu.de/en/software>. ChiSurf used for fluorescence decay analysis is available at <http://www.fret.at/tutorial/chisurf/>. DMD simulation engine is available at <http://www.moleculesinaction.com>.

Data availability

The authors declare that all data supporting the findings of this study are available within the paper and its supplementary information file, and available from the corresponding authors upon reasonable request.

Received: 20 October 2017 Accepted: 16 August 2018

Published online: 13 September 2018

References

1. Tompa, P. On the supertertiary structure of proteins. *Nat. Chem. Biol.* **8**, 597–600 (2012).
2. Xu, D., Jaroszewski, L., Li, Z. & Godzik, A. AIDA: ab initio domain assembly for automated multi-domain protein structure prediction and domain-domain interaction prediction. *Bioinformatics* **31**, 2098–2105 (2015).
3. Hertig, S., Goddard, T. D., Johnson, G. T. & Ferrin, T. E. Multidomain assembler (MDA) generates models of large multidomain proteins. *Biophys. J.* **108**, 2097–2102 (2015).
4. Schuler, B., Lipman, E. A., Steinbach, P. J., Kumke, M. & Eaton, W. A. Polyproline and the “spectroscopic ruler” revisited with single-molecule fluorescence. *Proc. Natl Acad. Sci. U. S. A.* **102**, 2754–2759 (2005).
5. Kalinin, S., et al. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat. Methods* **9**, 1218–1225 (2012).
6. Tomescu, A. I., Robb, N. C., Hengrungen, N., Fodor, E. & Kapanidis, A. N. Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. *Proc. Natl Acad. Sci. USA* **111**, E3335–E3342 (2014).

7. Hellenkamp, B., Wortmann, P., Kandzia, F., Zacharias, M. & Hugel, T. Multidomain structure and correlated dynamics determined by self-consistent FRET networks. *Nat. Methods* **14**, 174–180 (2017).
8. Eilert, T., Beckers, M., Drechsler, F. & Michaelis, J. Fast-NPS—A Markov Chain Monte Carlo-based analysis tool to obtain structural information from single-molecule FRET measurements. *Comput. Phys. Commun.* **219**, 377–389 (2017).
9. Reinartz, I. et al. Simulation of FRET dyes allows quantitative comparison against experimental data. *J. Chem. Phys.* **148**, 123321 (2018).
10. Sheng, M. PDZs and receptor/channel clustering: rounding up the latest suspects. *Neuron* **17**, 575–578 (1996).
11. Hung, A. Y. & Sheng, M. PDZ domains: structural modules for protein complex assembly. *J. Biol. Chem.* **277**, 5699–5702 (2002).
12. Kim, E. & Sheng, M. PDZ domain proteins of synapses. *Nat. Rev. Neurosci.* **5**, 771–781 (2004).
13. Feng, W. & Zhang, M. Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nat. Rev. Neurosci.* **10**, 87–99 (2009).
14. Lee, H. J. & Zheng, J. J. PDZ domains and their binding partners: structure, specificity, and modification. *Cell. Commun. Signal.* **8**, 8 (2010).
15. Cho, K. O., Hunt, C. A. & Kennedy, M. B. The rat brain postsynaptic density fraction contains a homolog of the *Drosophila* discs-large tumor suppressor protein. *Neuron* **9**, 929–942 (1992).
16. Fanning, A. S. & Anderson, J. M. Protein-protein interactions: PDZ domain networks. *Curr. Biol.* **6**, 1385–1388 (1996).
17. Doyle, D. A. et al. Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. *Cell* **85**, 1067–1076 (1996).
18. Erlendsson, S., & Madsen, K. L. Membrane binding and modulation of the PDZ domain of PICK1. *Membranes (Basel)* **5**, 597–615 (2015).
19. Toto, A. et al. Ligand binding to the PDZ domains of postsynaptic density protein 95. *Protein Eng. Des. Sel.* **29**, 169–175 (2016).
20. McCann, J. J. et al. Supertertiary structure of the synaptic MAGuK scaffold proteins is conserved. *Proc. Natl Acad. Sci. USA* **109**, 15775–15780 (2012).
21. McCann, J. J., Zheng, L., Chiantia, S. & Bowen, M. E. Domain orientation in the N-Terminal PDZ tandem from PSD-95 is maintained in the full-length protein. *Structure* **19**, 810–820 (2011).
22. Cook, D. J., Teves, L. & Tymianski, M. Treatment of stroke with a PSD-95 inhibitor in the gyrencephalic primate brain. *Nature* **483**, 213 (2012).
23. Bach, A. et al. A high-affinity, dimeric inhibitor of PSD-95 bivalently interacts with PDZ1-2 and protects against ischemic brain damage. *Proc. Natl Acad. Sci. USA* **109**, 3317–3322 (2012).
24. Ye, F. & Zhang, M. Structures and target recognition modes of PDZ domains: recurring themes and emerging pictures. *Biochem. J.* **455**, 1–14 (2013).
25. Xie, X. et al. Open-closed motion of Mint2 regulates APP metabolism. *J. Mol. Cell Biol.* **5**, 48–56 (2013).
26. Long, J.-F. et al. Autoinhibition of X11/Mint scaffold proteins revealed by the closed conformation of the PDZ tandem. *Nat. Struct. Mol. Biol.* **12**, 722–728 (2005).
27. Long, J.-F. et al. Supramodular structure and synergistic target binding of the N-terminal tandem PDZ domains of PSD-95. *J. Mol. Biol.* **327**, 203–214 (2003).
28. Wang, W., Weng, J., Zhang, X., Liu, M. & Zhang, M. Creating conformational entropy by increasing interdomain mobility in ligand binding regulation: a revisit to N-terminal tandem PDZ domains of PSD-95. *J. Am. Chem. Soc.* **131**, 787–796 (2008).
29. Sainlos, M. et al. Biomimetic divalent ligands for the acute disruption of synaptic AMPAR stabilization. *Nat. Chem. Biol.* **7**, 81–91 (2011).
30. Knight, J. L., Mekler, V., Mukhopadhyay, J., Ebright, R. H. & Levy, R. M. Distance-restrained docking of rifampicin and rifamycin SV to RNA polymerase using systematic FRET measurements: developing benchmarks of model quality and reliability. *Biophys. J.* **88**, 925–938 (2005).
31. Ha, T. & Tinnefeld, P. Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. *Annu. Rev. Phys. Chem.* **63**, 595–617 (2012).
32. Dimura, M. et al. Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Curr. Opin. Struct. Biol.* **40**, 163–185 (2016).
33. Sindbert, S. et al. Accurate distance determination of nucleic acids via forster resonance energy transfer: implications of dye linker length and rigidity. *J. Am. Chem. Soc.* **133**, 2463–2480 (2011).
34. Box George, E. P. Fitting empirical data. *Ann. N. Y. Acad. Sci.* **86**, 792–816 (1960).
35. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. 3rd edn (Springer, New York, 2006).
36. Dale, R. E., Eisinger, J. & Blumberg, W. E. The orientational freedom of molecular probes. *Oriental. Factor Intramol. Energy Transf. Biophys. J.* **26**, 161–193 (1979).
37. Sisamakos, E., Valeri, A., Kalinin, S., Rothwell, P. J. & Seidel, C. A. Accurate single-molecule FRET studies using multiparameter fluorescence detection. *Methods Enzymol.* **475**, 455–514 (2010).
38. Kalinin, S., Valeri, A., Antonik, M., Felekyan, S. & Seidel, C. A. M. Detection of structural dynamics by FRET: a photon distribution and fluorescence lifetime analysis of systems with multiple states. *J. Phys. Chem. B* **114**, 7983–7995 (2010).
39. Patlak, J. B. Measuring kinetics of complex single ion channel data using mean-variance histograms. *Biophys. J.* **65**, 29–42 (1993).
40. Ding, F., Dokholyan, N. V. in *Computational Modeling of Biological Systems* (Springer, Berlin, 2012).
41. Ding, F., Tsao, D., Nie, H. & Dokholyan, N. V. Ab initio folding of proteins with all-atom discrete molecular dynamics. *Structure* **16**, 1010–1018 (2008).
42. Karginov, A. V., Ding, F., Kota, P., Dokholyan, N. V. & Hahn, K. M. Engineered allosteric activation of kinases in living cells. *Nat. Biotechnol.* **28**, 743–747 (2010).
43. Shirvanyants, D., Ding, F., Tsao, D., Ramachandran, S. & Dokholyan, N. V. Discrete molecular dynamics: an efficient and versatile simulation method for fine protein characterization. *J. Phys. Chem. B* **116**, 8375–8382 (2012).
44. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **13**, 1011–1021 (1992).
45. Best, R. B., Hummer, G. & Eaton, W. A. Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci. USA* **110**, 17874–17879 (2013).
46. Noel, J. K., Chahine, J., Leite, V. B. & Whitford, P. C. Capturing transition paths and transition states for conformational rearrangements in the ribosome. *Biophys. J.* **107**, 2881–2890 (2014).
47. Bass, R. B., Butler, S. L., Chervitz, S. A., Gloor, S. L., Falke, J. J. in *Methods in Enzymology* (Elsevier, Amsterdam, 2007).
48. Wang, W., Weng, J., Zhang, X., Liu, M. & Zhang, M. Creating conformational entropy by increasing interdomain mobility in ligand binding regulation: a revisit to N-terminal tandem PDZ domains of PSD-95. *J. Am. Chem. Soc.* **131**, 787–796 (2009).
49. Kumawat, A. & Chakrabarty, S. Hidden electrostatic basis of dynamic allostery in a PDZ domain. *Proc. Natl. Acad. Sci. USA* **114**, E5825–E5834 (2017).
50. Gianni, S. et al. Sequence-specific long range networks in PSD-95/discs large/ZO-1 (PDZ) domains tune their binding selectivity. *J. Biol. Chem.* **286**, 27167–27175 (2011).
51. Petit, C. M., Zhang, J., Sapienza, P. J., Fuentes, E. J. & Lee, A. L. Hidden dynamic allostery in a PDZ domain. *Proc. Natl. Acad. Sci. USA* **106**, 18249–18254 (2009).
52. Varnum, M. D. & Zagotta, W. N. Interdomain interactions underlying activation of cyclic nucleotide-gated channels. *Science* **278**, 110–113 (1997).
53. Hunt, J. F. et al. Nucleotide control of interdomain interactions in the conformational reaction cycle of SecA. *Science* **297**, 2018–2026 (2002).
54. Bouvignies, G. et al. Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature* **477**, 111–114 (2011).
55. Uversky, V. N. Intrinsically disordered proteins and their (disordered) proteomes in neurodegenerative disorders. *Front. Aging Neurosci.* **7**, 18 (2015).
56. Callebaut, I., Hoffmann, B. & Mornon, J. P. The implications of CFTR structural studies for cystic fibrosis drug development. *Curr. Opin. Pharmacol.* **34**, 112–118 (2017).
57. Yang, M. et al. Single-molecule probing of the conformational homogeneity of the ABC transporter BtuCD. *Nat. Chem. Biol.* **14**, 715–722 (2018).
58. Kudryavtsev, V. et al. Combining MFD and PIE for Accurate Single-Pair Förster Resonance Energy Transfer Measurements. *Chemphyschem* **13**, 1060–1078 (2012).
59. Ma, J., et al. High precision FRET at single-molecule level for biomolecule structure determination. *J. Vis. Exp.* (123), e55623 (2017).
60. Hellenkamp, B. et al. Precision and accuracy of single molecule FRET measurements - a multi-laboratory benchmark study. *Nat. Methods* **15**, 669–676 (2018).
61. Ding, F. & Dokholyan, N. V. Emergence of protein fold families through rational design. *PLoS. Comput. Biol.* **2**, e85 (2006).
62. Yin, S., Ding, F. & Dokholyan, N. V. Modeling backbone flexibility improves protein stability estimation. *Structure* **15**, 1567–1576 (2007).
63. Lazaridis, T. & Karplus, M. Effective energy function for proteins in solution. *Proteins* **35**, 133–152 (1999).

Acknowledgements

The authors would like to thank Mingjie Zhang for providing coordinates for the NMR model. We thank Thomas O. Peulen for developing ChiSurf. This work was supported by NIH (GM094246 to H.S., MH081923 to M.E.B. and GM119691 to F.D.), NSF (CAREER MCB- 1749778 to H.S. and CAREER CBET-1553945 to F.D.) start-up funds from

Clemson University (H.S.), Center for Optical Materials Science and Engineering Technologies (COMSET), and the Creative Inquiry program (H.S.).

Author contributions

F.D., M.E.B. and H.S. conceived and designed the project. B.L. prepared and labeled FRET variants. I.S.Y.O., F.A.M., J.M., S.R.A. and G.H. measured and analyzed samples. M.S. carried out AV simulations. B.W. and F.D. performed DMD simulations. F.D., M.E. B. and H.S. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-018-06133-0>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018

Nature Communications: Supplement

Identifying Weak Interdomain Interactions that Stabilize the Supertertiary Structure of the N-Terminal Tandem PDZ Domains of PSD-95

Authors: Inna S. Yanez Orozco^{1§}, Frank Mindlin^{2§}, Junyan Ma^{3,4}, Bo Wang¹, Brie Levesque², Matheu Spencer¹, Soheila Rezaei Adariani¹, George Hamilton¹, Feng Ding¹, Mark E. Bowen², Hugo Sanabria^{1,4*}

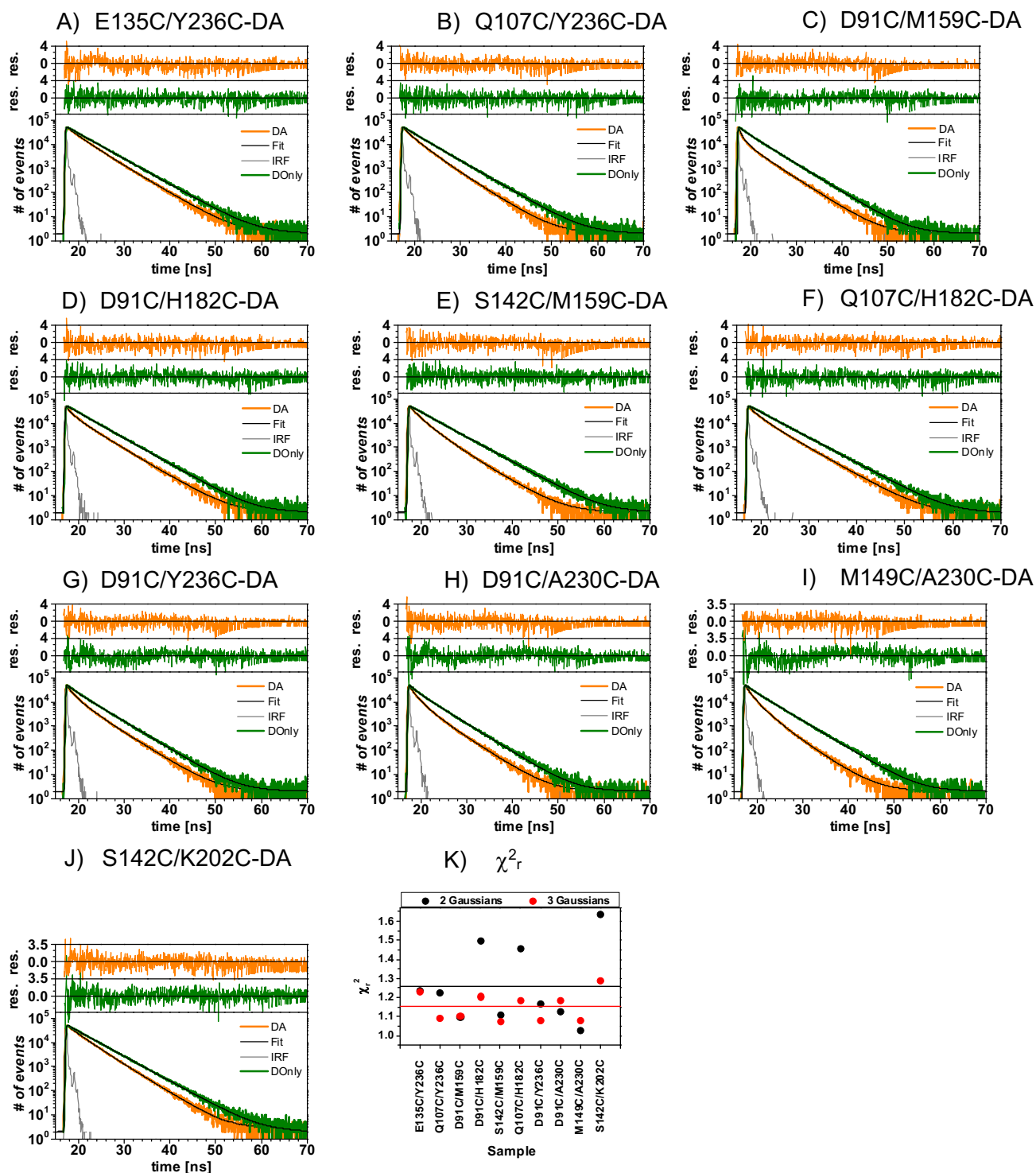
§: contributed equally, *: corresponding authors

| | Title | Page |
|--------------------------------|--|-----------|
| Supplementary Figure 1 | Time Resolved Fluorescence of donor only (DOnly) and donor-acceptor (DA) labeled PDZ1-2 tandems. | 3 |
| Supplementary Figure 2 | Time Resolved Fluorescence Anisotropy of donor only (DOnly) and donor-acceptor (DA) labeled PDZ1-2 tandems. | 4 |
| Supplementary Figure 3 | κ^2 distribution using wobble in a cone model. | 5 |
| Supplementary Figure 4 | Multiparameter Fluorescence Detection histograms. | 6 |
| Supplementary Figure 5 | Dynamic averaging observed by time window analysis. | 8 |
| Supplementary Figure 6 | Energy landscape of PDZ1-2 tandem from DMD simulations. | 10 |
| Supplementary Figure 7 | Replica Exchange DMD simulation trajectories. | 12 |
| Supplementary Figure 8 | Two dimensional PMF. | 13 |
| Supplementary Figure 9 | Comparison of the interdye distributions as derived from DMD simulation ($p(R_{DA})_{PMF} = \exp\left(-\frac{E_{PMF}}{0.6}\right)$) and the distances derived from MFD. | 14 |
| Supplementary Figure 10 | Distance comparison between experiment and simulations of ensemble representatives for the open-like (OL) and closed-like (CL) states. | 16 |
| Supplementary Figure 11 | Uncropped Scans of SDS-PAGE gels used for Disulfide mapping of the interdomain contact interface in the PDZ tandem. | 17 |
| Supplementary Figure 12 | Engineered electro-negative and electropositive regions on PDZ domains. | 18 |
| Supplementary Table 1 | Time Resolved fit results. A) Donor Acceptor Distances and fractions. B) Donor only Lifetime Decay. C) Acceptor Only Lifetime Decay. | 19 |
| Supplementary Table 2 | Time Resolved Anisotropy fit results A) Fit results of the time resolved DOnly anisotropy. B) Fit results of the time resolved acceptor anisotropy from direct excitation of the acceptor. C) Fit results of the time resolved acceptor anisotropy sensitized by FRET. | 21 |
| Supplementary Table 3 | Static FRET lines | 22 |
| Supplementary Table 4 | Dynamic FRET Lines | 23 |
| Supplementary Methods | <ul style="list-style-type: none"> • Ensemble Time Correlated Single Photon Counting (eTCSPC) analysis • Accessible Volume (AV) simulations to estimate measured distance • Analysis of Multiparameter Fluorescence Detection (MFD) for single-molecule FRET (smFRET) experiments • Calculation of the Static and Dynamic FRET Lines • Determination of Quantum Yields • Estimation of $\langle \kappa^2 \rangle$ and κ^2-distributions along with the associated uncertainty | 25 |

| | | |
|--|--|--|
| | <ul style="list-style-type: none">• <i>Error propagation</i>• <i>Global Analysis and goodness of fit.</i> | |
|--|--|--|

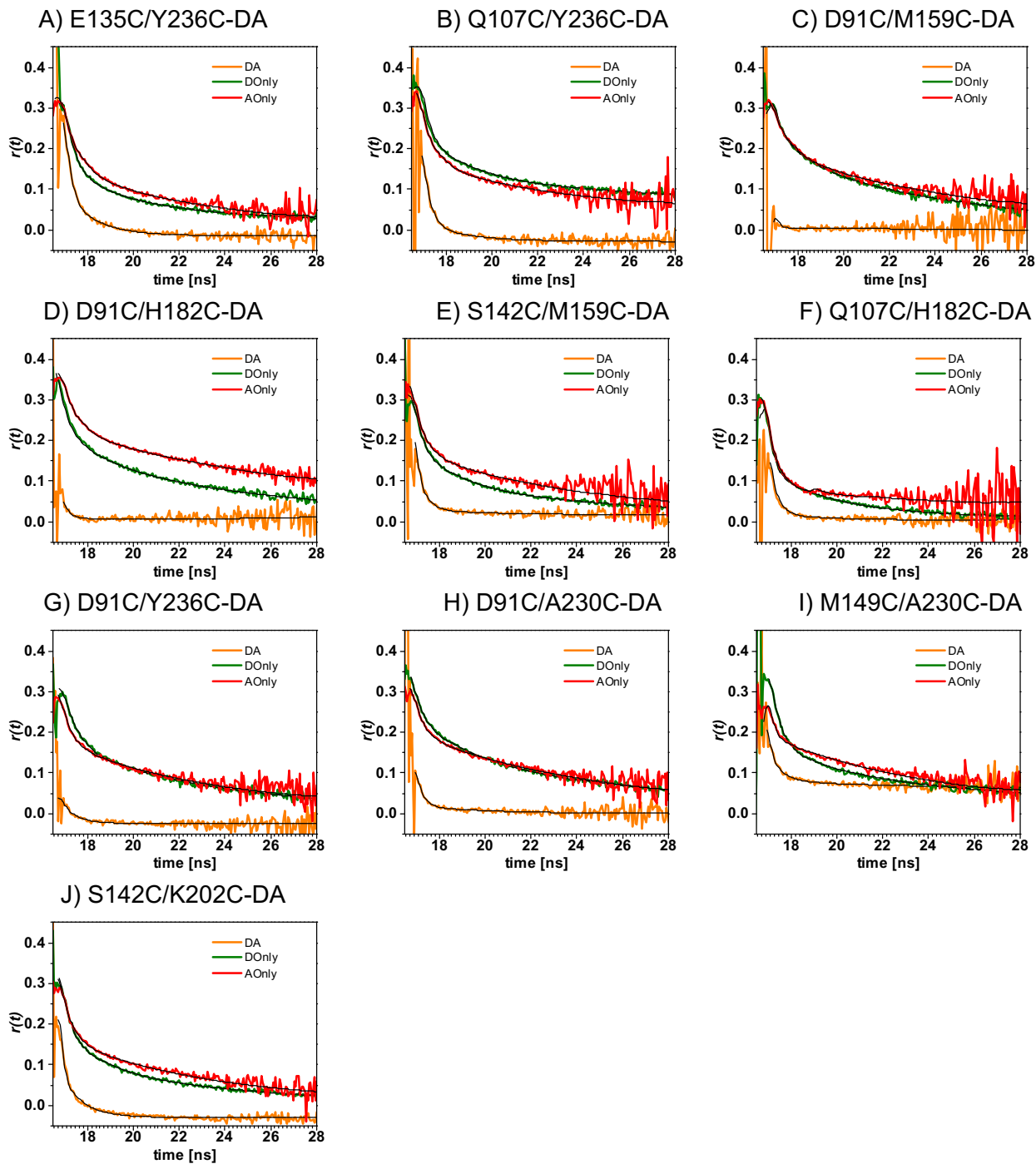
Supplementary Figures

Supplementary Figure 1



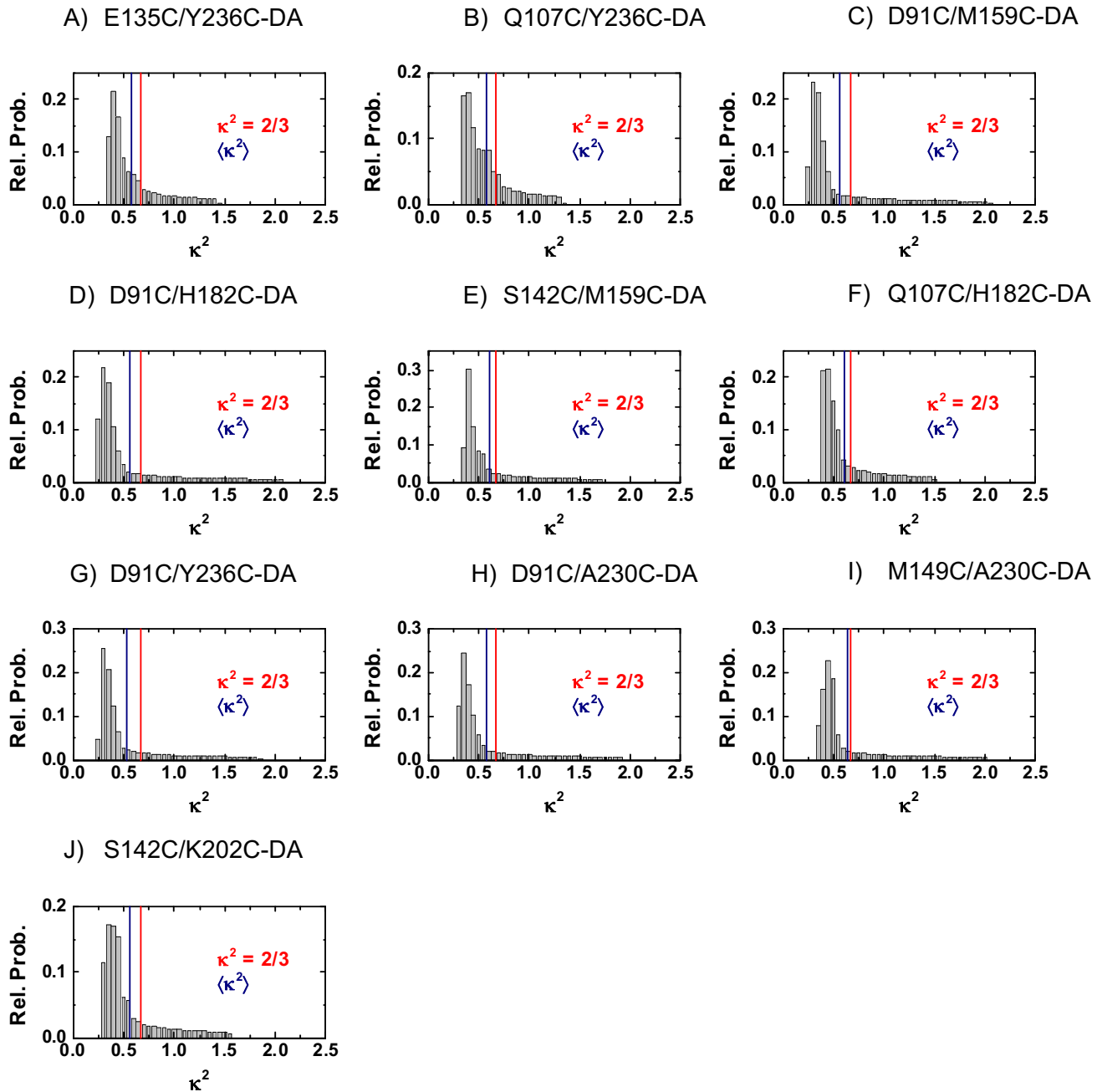
Supplementary Figure 1 Time Resolved Fluorescence of donor only (DOnly) and donor-acceptor (DA) labeled PDZ1-2 tandems. DOnly decays shown in green, and corresponding black overlay uses the Eq. 1. DA decays are shown in orange and model fit function uses Eqs. 3 and 4 to describe the Gaussian distribution of states. Instrument response function (IRF) is shown in gray. A) E135C/Y236C-DA B) Q107C/Y236C-DA C) D91C/M159C-DA D) D91C/H182C-DA, E) S142C/M159C-DA, F) Q107C/H182C-DA, G) D91C/Y236C-DA, H) D91C/A230C-DA, I) M149C/A230C-DA, J) S142C/K202C-DA. K) χ^2 comparison of two-Gaussian distributed states and three-Gaussian distributed states. Supplementary Table 1 summarizes the fit results.

Supplementary Figure 2



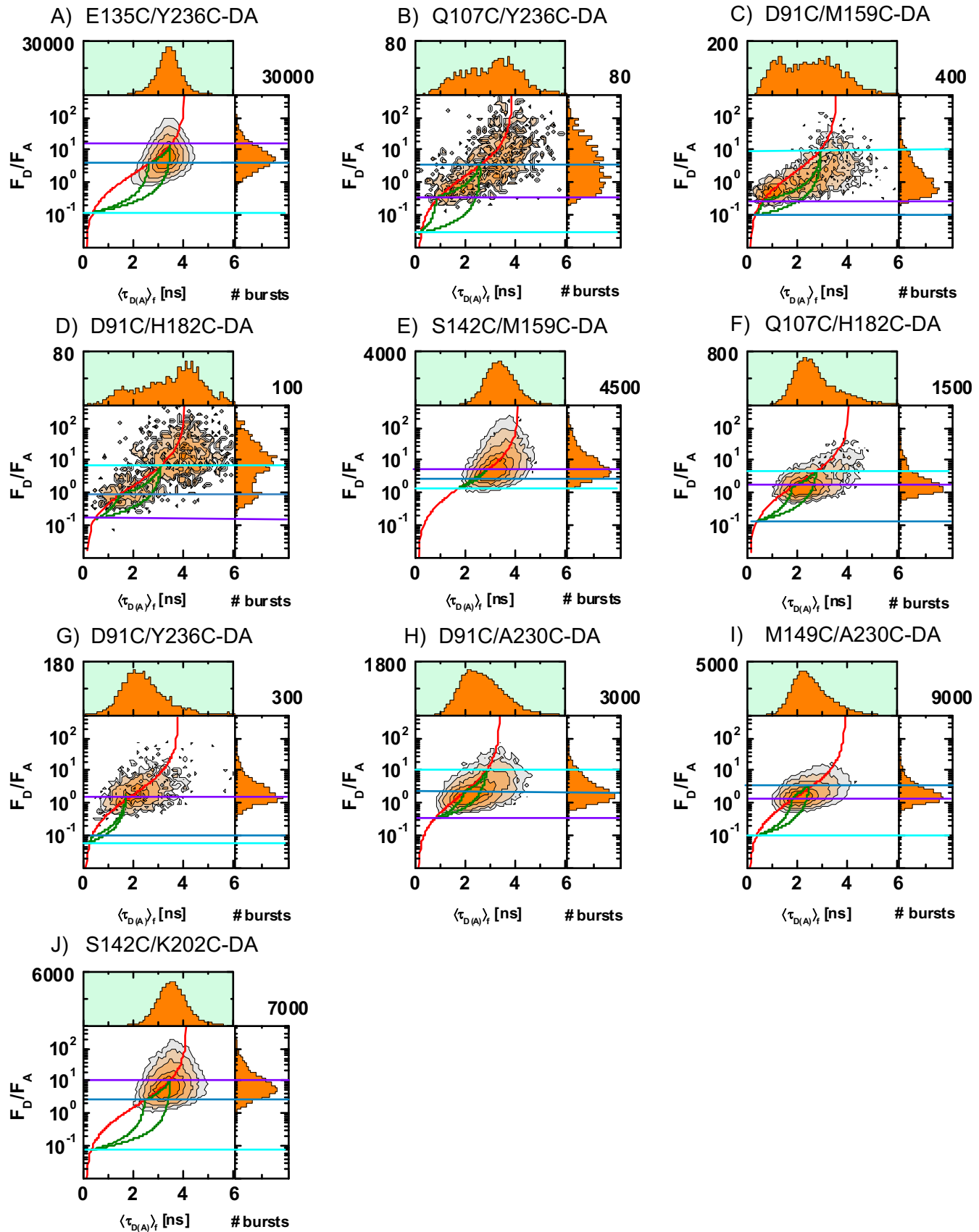
Supplementary Figure 2 Time Resolved Fluorescence Anisotropy of donor only (DOnly) and donor-acceptor (DA) labeled PDZ1-2 tandems. DOnly decays shown in green and corresponding black overlay use the Eqs. 1 and 6. DA decays are shown in orange and model fit function use Eqs. 4 and 6. Direct excitation of acceptor Instrument response function (IRF) is shown in gray. A) E135C/Y236C-DA B) Q107C/Y236C-DA C) D91C/M159C-DA D) D91C/H182C-DA, E) S142C/M159C-DA, F) Q107C/H182C-DA, G) D91C/Y236C-DA, H) D91C/A230C-DA, I) M149C/A230C-DA, J) S142C/K202C-DA. Supplementary Table 1 summarizes the fit results.

Supplementary Figure 3



Supplementary Figure 3 κ^2 distribution using wobble in a cone model. A) E135C/Y236C-DA B) Q107C/Y236C-DA C) D91C/M159C-DA D) D91C/H182C-DA, E) S142C/M159C-DA, F) Q107C/H182C-DA, G) D91C/Y236C-DA, H) D91C/A230C-DA, I) M149C/A230C-DA, J) S142C/K202C-DA. Supplementary Table 1 summarizes the fit results. Red and blue vertical lines are the assumed value of $2/3$ and the average value of the distribution, respectively. Distributions were obtained using time resolved anisotropy decays following Eqs. 12-14

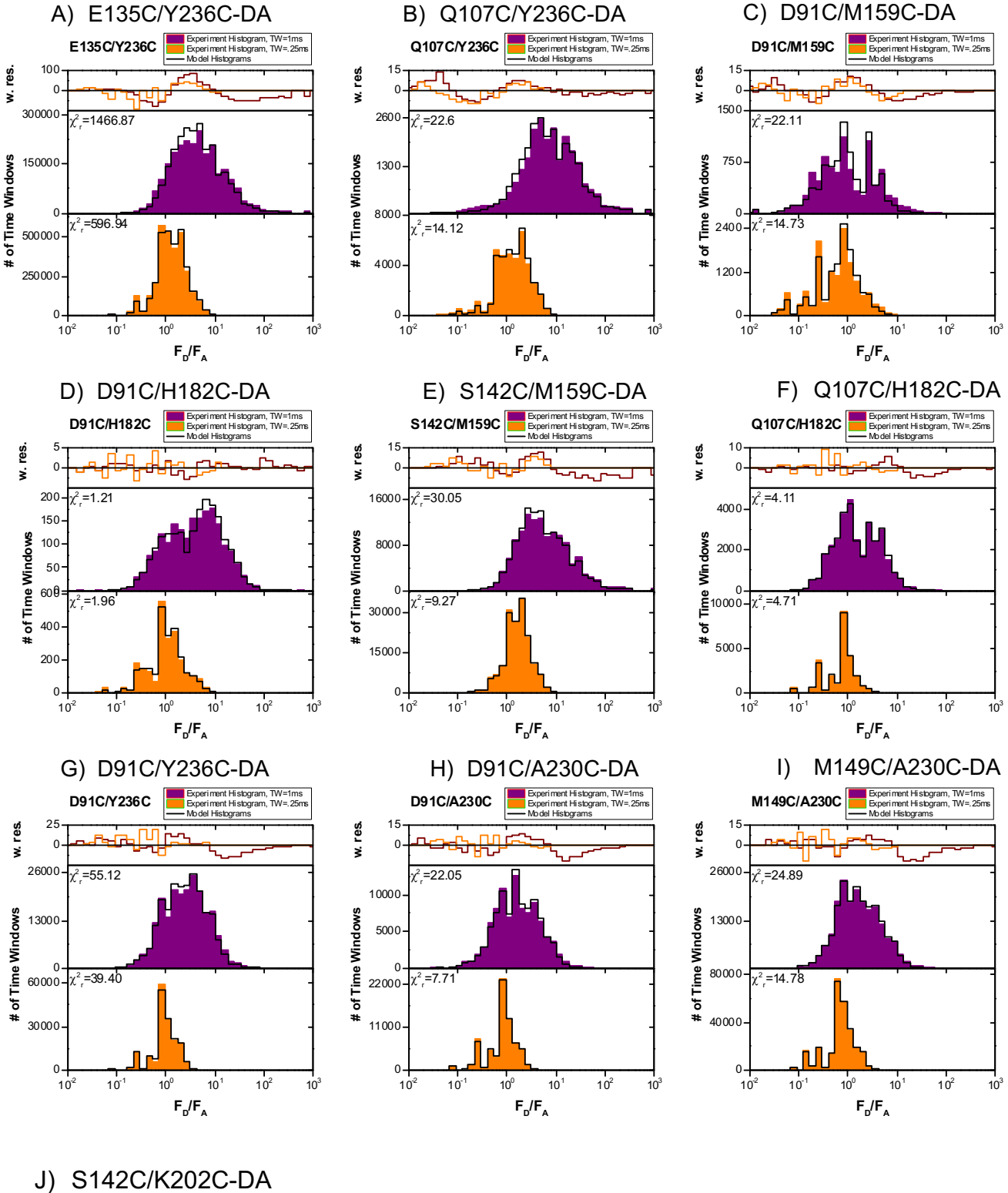
Supplementary Figure 4

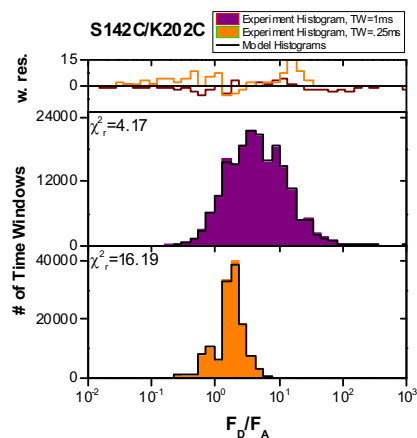


Supplementary Figure 4 Multiparameter Fluorescence Detection histograms. Each panel plots two FRET indicators for each single molecule event. The donor over acceptor fluorescence ratio (F_D/F_A) on the y-axis and

the average fluorescence lifetime ($\langle\tau_{D(A)}\rangle_f$) on the x-axis. Shown are the 2D histograms (orange contours) for A) E135C/Y236C-DA B) Q107C/Y236C-DA C) D91C/M159C-DA D) D91C/H182C-DA, E) S142C/M159C-DA, F) Q107C/H182C-DA, G) D91C/Y236C-DA, H) D91C/A230C-DA, I) M149C/A230C-DA, J) S142C/K202C-DA. Supplementary Table 1 summarizes the fit results. Horizontal guidelines for the major and minor states 1 and 2 are shown in purple and blue, respectively. Red and green lines correspond to the static and dynamic FRET lines (Supplementary Tables 3, 4), respectively.

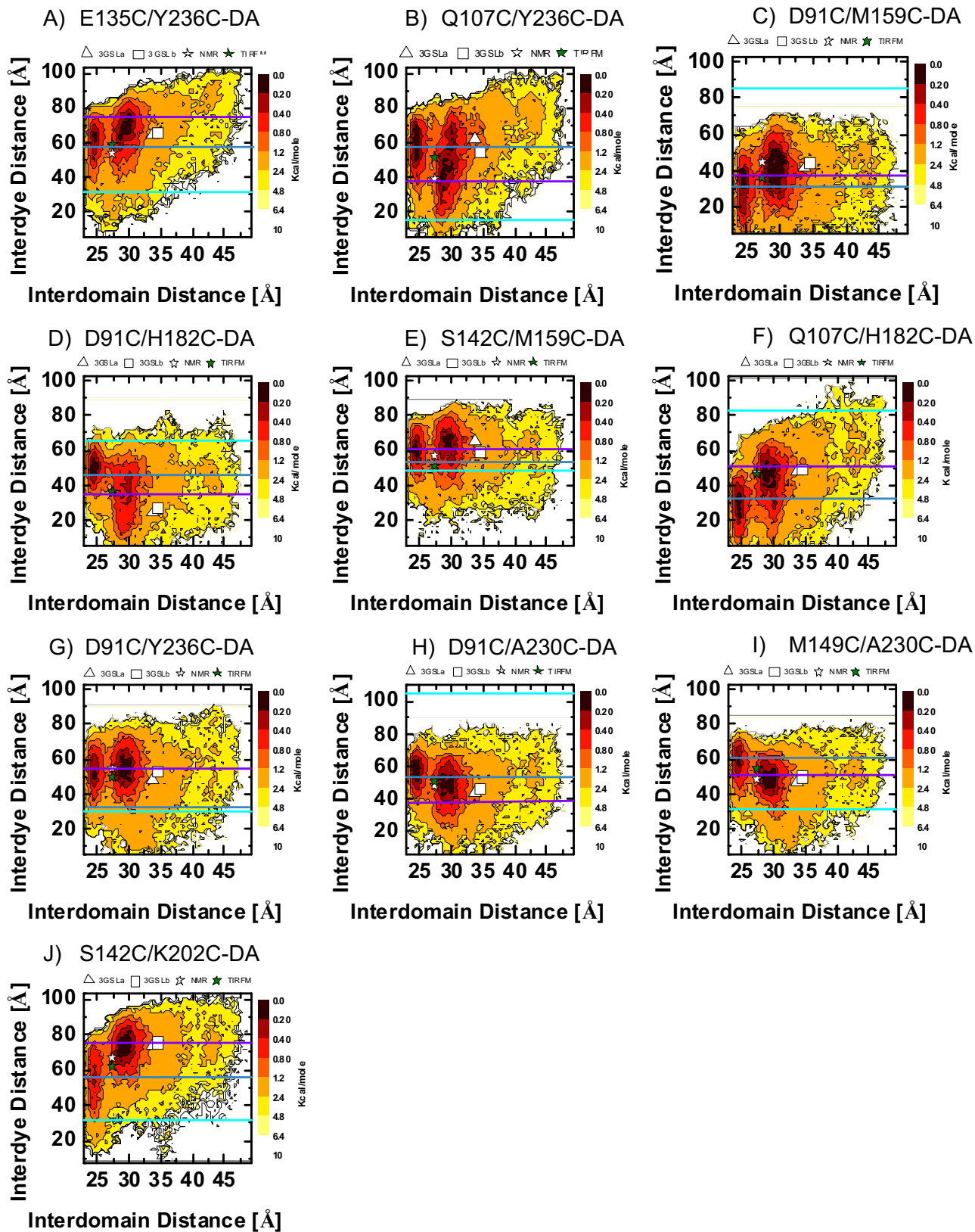
Supplementary Figure 5





Supplementary Figure 5 Dynamic averaging observed by time window analysis. Two time windows (0.25 ms and 1 ms) were globally fit for each FRET variant. A) E135C/Y236C-DA B) Q107C/Y236C-DA C) D91C/M159C-DA D) D91C/H182C-DA, E) S142C/M159C-DA, F) Q107C/H182C-DA, G) D91C/Y236C-DA, H) D91C/A230C-DA, I) M149C/A230C-DA, J) S142C/K202C-DA. Each time window has its corresponding χ^2_r , which would not vary significantly between selected time windows for static molecules. The model used considers a single Gaussian distributed interdye distance with free standard deviation. The fit quality is poor indicating that a single state does not describe the system. Moreover, the figure of merit χ^2_r changes between time windows as indicative of dynamic effects.

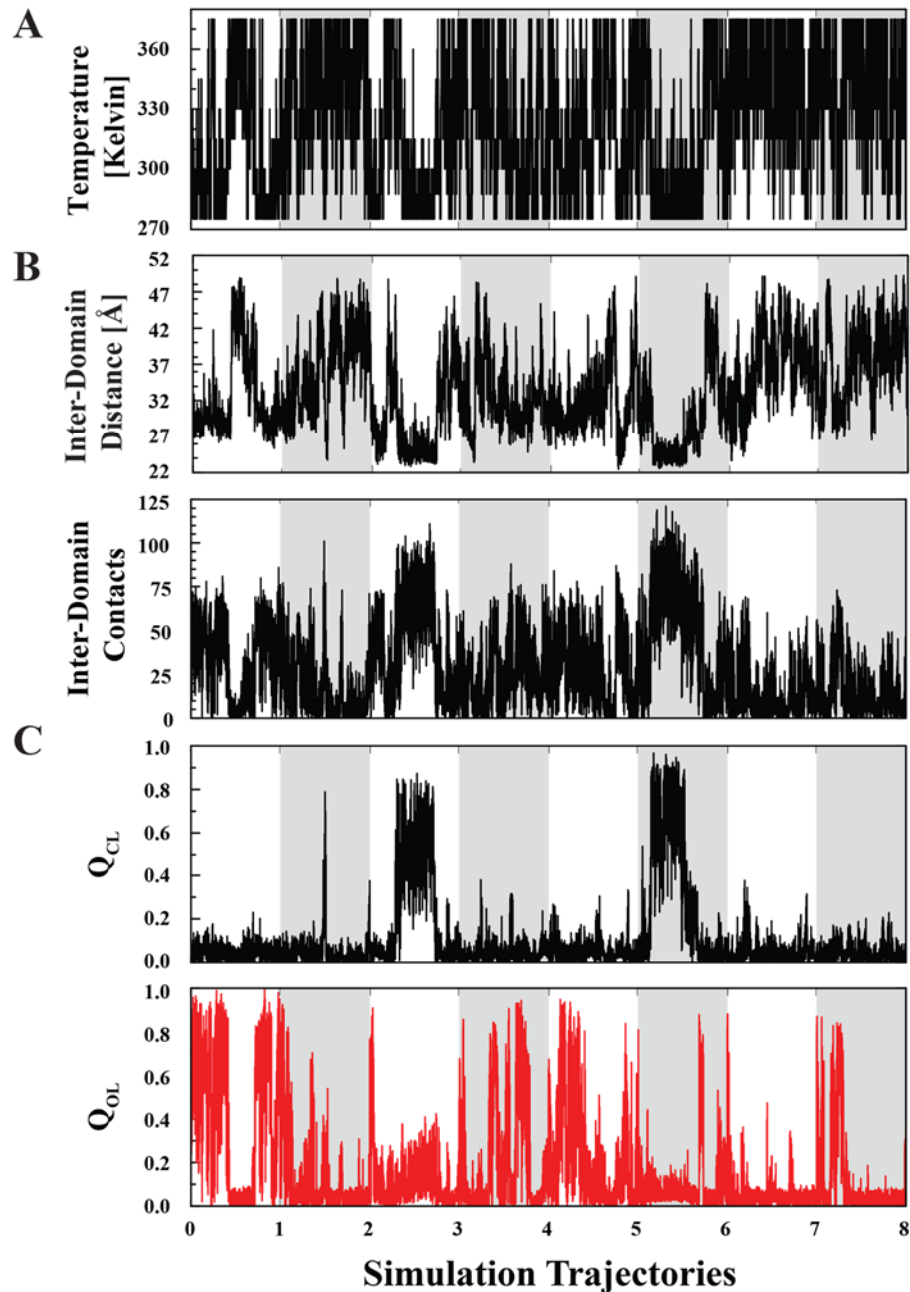
Supplementary Figure 6



Supplementary Figure 6 Energy landscape of PDZ1-2 tandem from DMD simulations. A) E135C/Y236C-DA B) Q107C/Y236C-DA C) D91C/M159C-DA D) D91C/H182C-DA, E) S142C/M159C-DA, F) Q107C/H182C-DA, G) D91C/Y236C-DA, H) D91C/A230C-DA, I) M149C/A230C-DA, J) S142C/K202C-DA. Supplementary

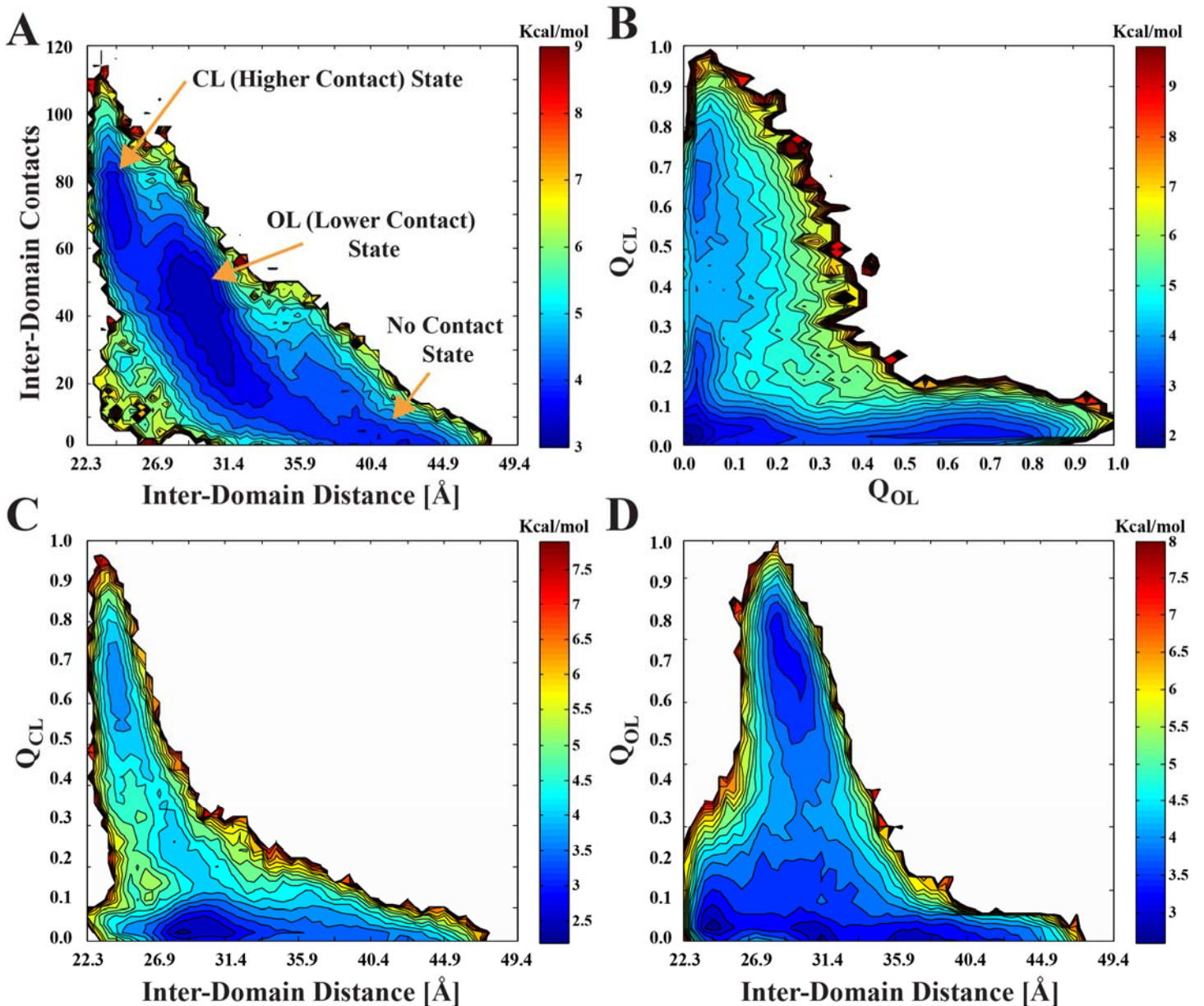
Table 1 summarizes the fit results. Horizontal guidelines for the mean interdye states determined by eTCSPC are shown in purple and blue, respectively. Symbols represent the X-ray, NMR and TIRFM structural models.

Supplementary Figure 7



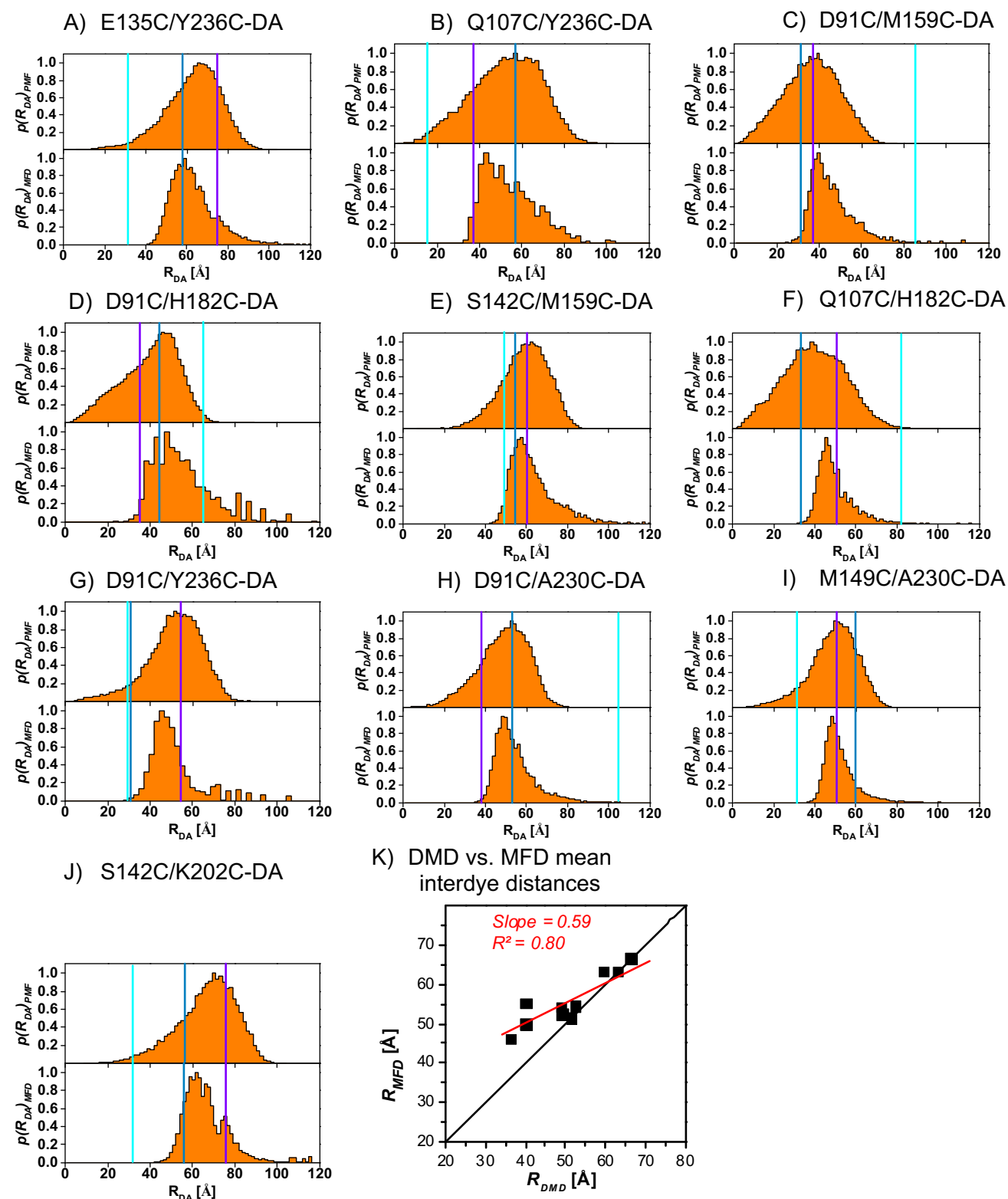
Supplementary Figure 7 Replica Exchange DMD simulation trajectories. A) Temperature as a function of the simulation time represented in terms of the replica or trajectory, B) Time evolution of the center of mass distance between two domains (Inter-Domain distance) and the number of Inter-Domain contacts between residues. Two residues are in contact if they have a least one atomic contact with a cutoff of 6.5 Å. C) The simulation trajectories of eight replicas in terms of the fraction of native contacts for the CL (closed-like) state (Q_{CL}) and the OL (open-like) state (Q_{OL}). Q-values for each state were computed according to native contact determined from corresponding structural ensembles (Fig. 4, and Supplementary Fig. 8). The total time for each replica is 200 ns, and there are 8 trajectories (0-7). The shading is used to separate different replicas.

Supplementary Figure 8



Supplementary Figure 8 Two dimensional PMF. The conformational dynamics of PDZ1-2 depicted by two-dimensional PMF (2D-PMF) plots projected onto different reaction coordinates, including the number of residue-wise contacts between two PDZ domains (Inter-Domain Contacts), the center of mass distance between two domains (Inter-Domain distance), and fractions of native contacts for the closed-like (CL) (Q_{CL}) and the open-like state (Q_{OL}). (A) From the PMF plot of Inter-Domain Contacts versus Inter-Domain Distance, two energy basins corresponding to the CL state and the OL state can be identified. (B) The 2D-PMF as the function of Q_{CL} and Q_{OL} suggests that these two states are exclusive to each other with different interface contacts (see also Fig. 4C,D) and the tandem domain in other state has to break all contacts in order to fold to the other state. The 2D-PMF as functions of (C) Q_{CL} versus Inter-domain and (D) Q_{OL} versus Inter-domain also highlight the separation of two states. The color bar denotes free energy scale in a unit of Kcal/mol.

Supplementary Figure 9

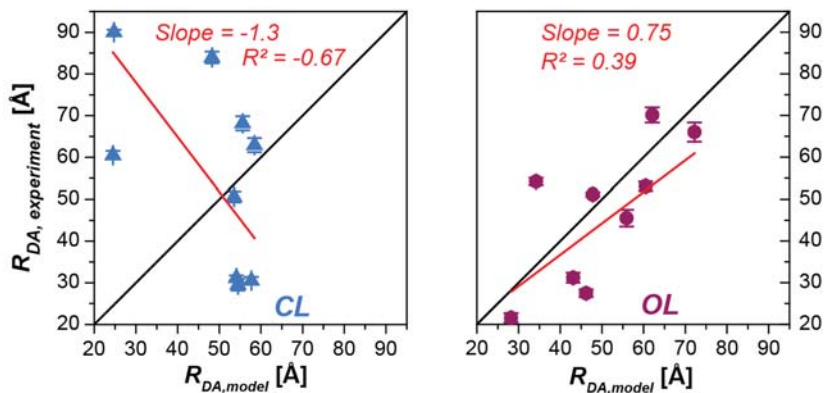


Supplementary Figure 9 Comparison of the interdyne distributions as derived from DMD simulation ($p(R_{DA})_{PMF} = \exp\left(-\frac{E_{PMF}}{0.6}\right)$) and the distances derived from MFD. A) E135C/Y236C-DA B) Q107C/Y236C-DA C) D91C/M159C-DA D) D91C/H182C-DA, E) S142C/M159C-DA, F) Q107C/H182C-DA, G) D91C/Y236C-DA, H) D91C/A230C-DA, I) M149C/A230C-DA, J) S142C/K202C-DA. F_D/F_A distributions are converted into

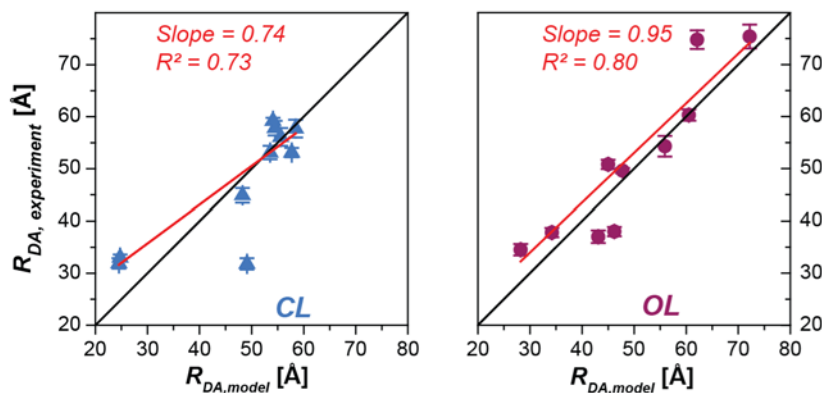
interdye distances following $R_{DA} = R_0 \left(\frac{\Phi_{FA}}{\Phi_{FD(0)}} \cdot \frac{F_D}{F_A} \right)^{1/6}$ 12. K) Correlation between the mean interdye distances as derived for each variant using DMD interdye distributions and mean interdye distances derived from MFD distance distributions.

Supplementary Figure 10

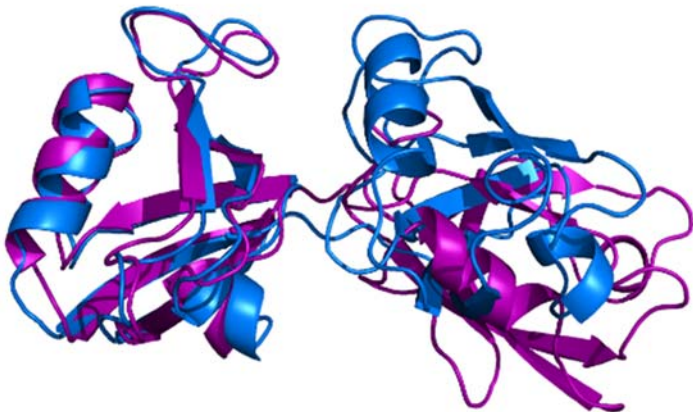
A) Two-Gaussian distributed states



B) Three-Gaussian distributed states

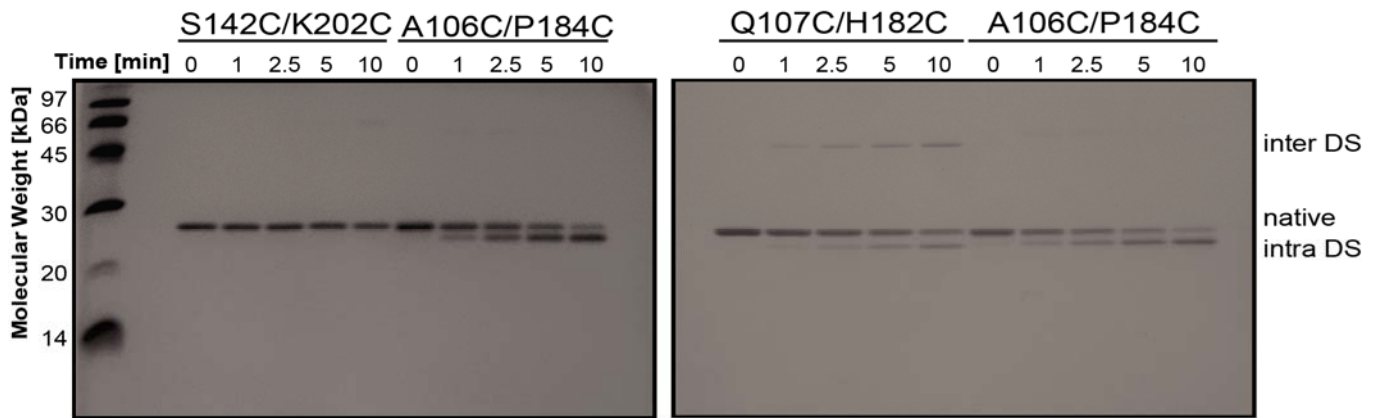


C) Overlay of the two representative structures



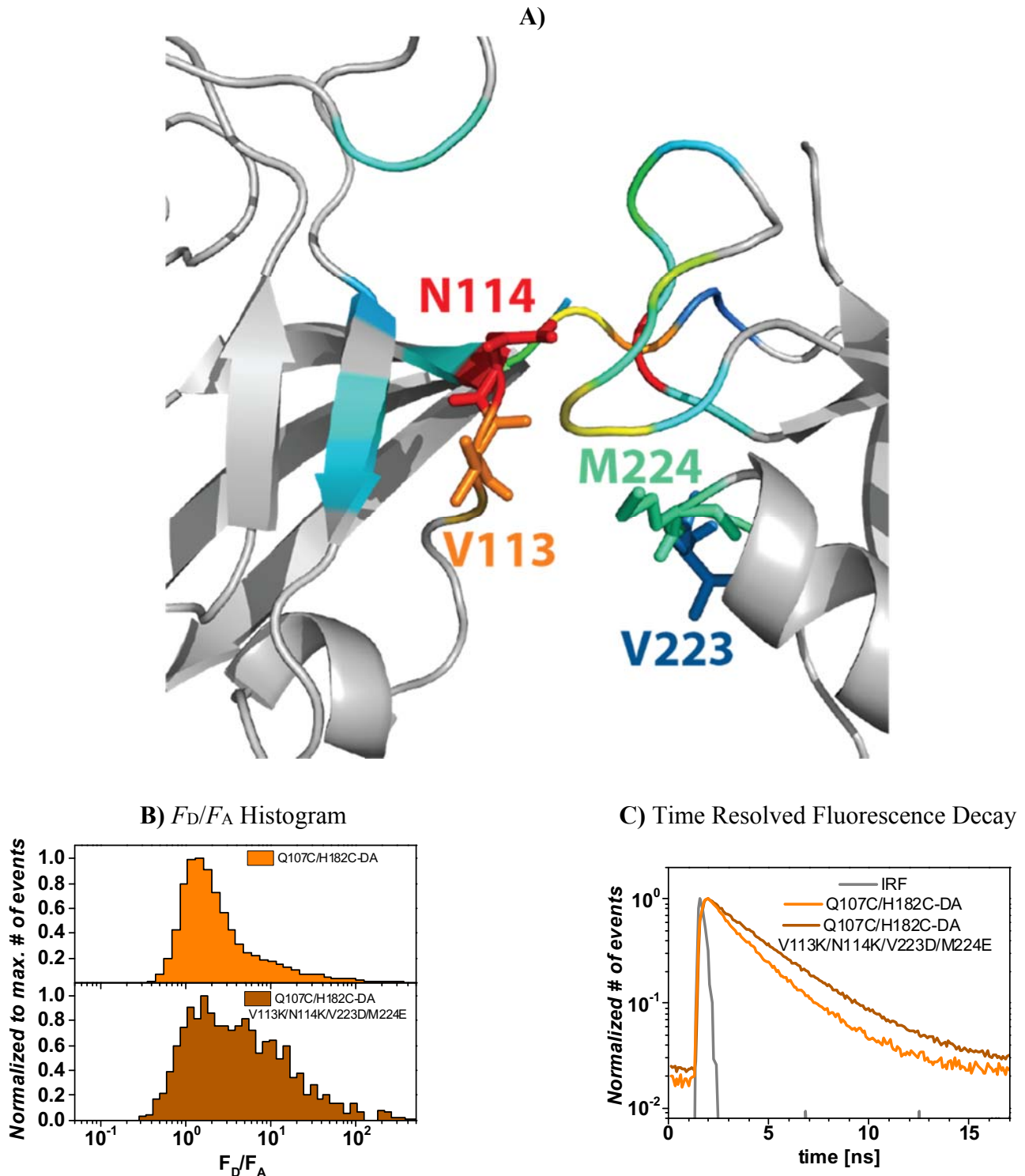
Supplementary Figure 10 Distance comparison between experiment and simulations of ensemble representatives for the open-like (OL) and closed-like (CL) states. A) Using a two-Gaussian distributed state model and B) with a three-Gaussian distributed state. Error bars on panel A and B correspond to uncertainties on the measured interdyne distance. C) Cartoon representation of the two representative structures (Fig. 4A-B) overlaying PDZ1. The significant rotation of PDZ2 with respect of PDZ1 is observed. Blue is the CL state and magenta corresponds to the OL state.

Supplementary Figure 11



Supplementary Figure 11 Uncropped Scans of SDS-PAGE gels used for Disulfide mapping of the interdomain contact interface in the PDZ tandem. A cartoon representation of the location of engineered disulfides is shown in Fig. 5 in the main text. Representative SDS-PAGE showing individual time points used for the analysis of the extent and kinetics of the disulfide bond formation reactions. Disulfide formation was initiated by the addition of 0.5 mM CuSO_4 and 1.75 mM 1, 10-phenanthroline. Time points were quenched by adding 40 mM N-ethylmaleimide to alkylate unbonded cysteines and 10 mM EDTA followed immediately by boiling at 95 °C for 5 minutes in non-reducing Laemmli sample buffer. The location of the introduced cysteine residues for each mutant is labeled above the panels. Duration of the reaction for each time point is indicated above each lane. The molecular weights of the standards (GE Healthcare) are indicated to the left of the panels. The position of the native PDZ Tandem along with the intramolecular and intermolecular disulfide products are indicated to the right of the panels. Reactions were run in triplicate to obtain the percentage of disulfide for each time point. Each time course was fit with an exponential function to obtain the initial and final percentage of disulfide bonding along with the rate of disulfide formation as shown in Fig. 5.

Supplementary Figure 12



Supplementary Figure 12 Engineered electro-negative and electropositive regions on PDZ domains. A) Cartoon representation of the open-like state showing the residues that were selected for mutations (V113K/N114K/V223D/M224E) B) F_D/F_A Histogram of the Q107C/H182C (**Orange**) and the engineered salt bridge V113K/N114K/V223D/M224E (**brown**). This last one shifts the equilibrium towards lower FRET efficiencies or higher F_D/F_A ratio. C) Time Resolved Fluorescence of the Donor emission in the presence of acceptor for the Q107C/H182C and the engineered salt bridge V113K/N114K/V223D/M224E. The mutated interface shows a slower decay, consistent with the F_D/F_A ratio. The shift towards lower FRET is consistent

with the anticipated clockwise motion of PDZ2 (right domain in panel A) with respect to PDZ1 (left domain in panel A), which in turn it will push farther the labels located at Q107C and H182C to center the engineered interface.

Supplementary Tables

Supplementary Table 1 Time Resolved fit results.

A) Donor Acceptor Distances and fractions as model with a three Gaussian distributed and the addition of a No FRET state (Eq. 4). There is a major population representing the 43.5%. The second most populated accounts for 32.4%, and the least populated corresponds to 24.1%.

| Sample | σ_{1DA} [Å] | $\langle R_{DA} \rangle_1 \pm \epsilon$ [Å] | σ_{2DA} [Å] | $\langle R_{DA} \rangle_2 \pm \epsilon$ [Å] | σ_{3DA} [Å] | $\langle R_{DA} \rangle_3 \pm \epsilon$ [Å] | No FRET | χ_r^2 |
|----------------|-----------------------|--|--------------------|--|-----------------------|--|---------|------------|
| E135C/Y236C-DA | 2.6 | 74.8 ± 1.8 | 8.7 | 57.7 ± 1.7 | 5.8 | 32.6 ± 0.8 | 0.28 | 1.233 |
| Q107C/Y236C-DA | 1.3 | 37.8 ± 0.9 | 8.4 | 57.8 ± 1.4 | 0.7 | 15.7 ± 1.1 | 0.15 | 1.095 |
| D91C/M159C-DA | 9.1 | 37.0 ± 1.2 | 5.0 | 31.9 ± 1.0 | 3.0 | 85.2 ± 2.6 | 0.09 | 1.104 |
| D91C/H182C-DA | 8.0 | 34.5 ± 1.1 | 6.1 | 44.9 ± 1.4 | 8.2 | 65.7 ± 2.1 | 0.21 | 1.206 |
| S142C/M159C-DA | 0.5 | 60.3 ± 1.1 | 3.4 | 53.1 ± 0.9 | 13.9 | 47.8 ± 0.8 | 0.09 | 1.078 |
| Q107C/H182C-DA | 4.4 | 50.8 ± 0.9 | 4.5 | 33.0 ± 0.6 | 4.8 | 81.7 ± 1.4 | 0.12 | 1.183 |
| D91C/Y236C-DA | 3.9 | 54.3 ± 2.0 | 15.5 | 31.8 ± 1.1 | 2.8 | 29.7 ± 1.1 | 0.22 | 1.083 |
| D91C/A230C-DA | 10.6 | 38.0 ± 0.9 | 3.7 | 53.1 ± 1.3 | 2.7 | 105.7 ± 2.7 | 0.11 | 1.183 |
| M149C/A230C-DA | 3.3 | 49.6 ± 0.5 | 2.3 | 59.2 ± 0.6 | 11.1 | 31.8 ± 0.3 | 0.07 | 1.078 |
| S142C/K202C-DA | 5.7 | 75.4 ± 2.3 | 4.5 | 56.1 ± 1.7 | 1.9 | 31.2 ± 1.0 | 0.00 | 1.290 |

B) Donor only Lifetime Decay. $\Phi_{FD(0)}$ determination is described in methods section. Goodness of fit evaluated for 1 (χ_{r1}^2) and 2 (χ_{r2}^2) exponential decays.

| Sample | x_1 | τ_1 [ns] | x_2 | τ_2 [ns] | $\langle \tau \rangle_x$ [ns] | $\langle \tau \rangle_f$ [ns] | $\Phi_{D(0)}$ | χ_{r2}^2 | χ_{r1}^2 |
|----------------|-------|------------------|-------|------------------|----------------------------------|----------------------------------|---------------|---------------|---------------|
| E135C/Y236C-DA | 0.92 | 4.13 | 0.08 | 1.24 | 3.90 | 4.06 | 0.78 | 1.172 | 2.613 |
| Q107C/Y236C-DA | 0.87 | 3.99 | 0.13 | 1.46 | 3.65 | 3.85 | 0.73 | 1.288 | 4.091 |
| D91C/M159C-DA | 0.77 | 3.76 | 0.23 | 1.17 | 3.16 | 3.54 | 0.63 | 1.386 | 12.53 |
| D91C/H182C-DA | 0.92 | 4.10 | 0.08 | 1.10 | 3.86 | 4.03 | 0.77 | 1.361 | 2.448 |
| S142C/M159C-DA | 0.90 | 4.24 | 0.10 | 1.54 | 3.98 | 4.14 | 0.80 | 1.343 | 2.791 |
| Q107C/H182C-DA | 0.94 | 4.14 | 0.06 | 1.15 | 2.85 | 4.09 | 0.79 | 1.295 | 2.020 |
| D91C/Y236C-DA | 0.74 | 3.29 | 0.26 | 1.03 | 2.71 | 3.07 | 0.54 | 1.556 | 14.98 |
| D91C/A230C-DA | 0.66 | 3.72 | 0.34 | 1.15 | 2.85 | 3.37 | 0.57 | 1.473 | 26.35 |
| M149C/A230C-DA | 0.76 | 3.85 | 0.24 | 1.13 | 3.20 | 3.62 | 0.64 | 1.711 | 13.89 |
| S142C/K202C-DA | 0.89 | 4.22 | 0.11 | 1.05 | 3.86 | 4.12 | 0.77 | 1.674 | 4.384 |

C) Acceptor Only Lifetime Decay. Φ_{FA} determination is described in methods section

| Sample | x_1 | τ_1 | x_2 | τ_2 | $\langle \tau \rangle_x$ [ns] | $\langle \tau \rangle_f$ [ns] | Φ_{FA} | χ_r^2 |
|----------------|-------|----------|-------|----------|----------------------------------|----------------------------------|-------------|------------|
| E135C/Y236C-DA | 0.68 | 1.54 | 0.32 | 0.77 | 1.30 | 1.40 | 0.35 | 1.174 |
| Q107C/Y236C-DA | 0.48 | 1.71 | 0.52 | 1.19 | 1.44 | 1.49 | 0.39 | 1.748 |
| D91C/M159C-DA | 0.60 | 1.65 | 0.40 | 0.72 | 1.28 | 1.44 | 0.35 | 1.849 |
| D91C/H182C-DA | 0.66 | 1.73 | 0.34 | 0.72 | 1.38 | 1.55 | 0.38 | 1.815 |
| S142C/M159C-DA | 0.58 | 1.53 | 0.42 | 0.86 | 1.25 | 1.34 | 0.34 | 1.447 |

| | | | | | | | | |
|-----------------------|------|------|------|------|------|------|------|-------|
| Q107C/H182C-DA | 0.59 | 1.57 | 0.41 | 0.67 | 1.20 | 1.36 | 0.33 | 0.885 |
| D91C/Y236C-DA | 0.65 | 1.69 | 0.35 | 0.86 | 1.40 | 1.51 | 0.38 | 1.643 |
| D91C/A230C-DA | 0.62 | 1.83 | 0.38 | 0.93 | 1.49 | 1.62 | 0.41 | 1.423 |
| M149C/A230C-DA | 0.66 | 1.79 | 0.34 | 0.90 | 1.49 | 1.61 | 0.41 | 1.413 |
| S142C/K202C-DA | 0.60 | 1.91 | 0.40 | 1.00 | 1.54 | 1.67 | 0.42 | 1.423 |

Supplementary Table 2 Time Resolved Anisotropy.

A) Fit results of the time resolved DOnly anisotropy.

| Sample | b_1 | ρ_1 [ns] | b_2 | ρ_2 [ns] | b_3 | ρ_3 [ns] | $\chi_{r,sum}^2$ | $\chi_{r,diff}^2$ | $r_{ss} \pm \text{std}$ |
|----------------|-------|---------------|-------|---------------|-------|---------------|------------------|-------------------|-------------------------|
| E135C/Y236C-DA | 0.12 | 0.33 | 0.09 | 1.99 | 0.06 | 12.24 | 2.86 | 1.03 | 0.1 ±0.01 |
| Q107C/Y236C-DA | 0.16 | 0.31 | 0.11 | 2.32 | 0.12 | 37.00 | 2.90 | 1.14 | 0.2 ±0.00 |
| D91C/M159C-DA | 0.10 | 0.22 | 0.09 | 1.27 | 0.17 | 8.81 | 1.69 | 1.10 | 0.2±0.00 |
| D91C/H182C-DA | 0.12 | 0.35 | 0.08 | 1.91 | 0.15 | 10.97 | 2.59 | 0.95 | 0.2 ±0.00 |
| S142C/M159C-DA | 0.15 | 0.29 | 0.10 | 2.06 | 0.08 | 14.00 | 4.40 | 1.51 | 0.1 ±0.02 |
| Q107C/H182C-DA | 0.25 | 0.22 | 0.10 | 2.63 | 0.03 | 13.49 | 4.17 | 1.51 | 0.1 ±0.00 |
| D91C/Y236C-DA | 0.14 | 0.62 | | | 0.16 | 7.39 | 2.23 | 1.16 | 0.1 ±0.01 |
| D91C/A230C-DA | 0.11 | 0.35 | 0.12 | 2.89 | 0.12 | 15.47 | 3.42 | 1.28 | 0.2 ±0.01 |
| M149C/A230C-DA | 0.17 | 0.36 | 0.07 | 2.59 | 0.10 | 14.94 | 3.51 | 1.32 | 0.1 ±0.02 |
| S142C/K202C-DA | 0.14 | 0.27 | 0.09 | 1.45 | 0.10 | 7.91 | 3.68 | 1.33 | 0.1 ±0.01 |

B) Fit results of the time resolved acceptor anisotropy from direct excitation of the acceptor.

| Sample | b_1 | ρ_1 [ns] | b_2 | ρ_2 [ns] | b_3 | ρ_3 [ns] | $\chi_{r,sum}^2$ | $\chi_{r,diff}^2$ | $r_{ss} \pm \text{std}$ |
|----------------|-------|---------------|-------|---------------|-------|---------------|------------------|-------------------|-------------------------|
| E135C/Y236C-DA | 0.19 | 0.58 | 0.15 | 6.94 | | | 9.28 | 1.74 | 0.2 ±0.01 |
| Q107C/Y236C-DA | 0.17 | 0.48 | 0.12 | 6.04 | 0.05 | 5691 | 6.55 | 1.36 | 0.2 ±0.06 |
| D91C/M159C-DA | 0.14 | 0.64 | 0.19 | 9.83 | | | 6.14 | 1.33 | 0.2 ±0.02 |
| D91C/H182C-DA | 0.15 | 0.56 | 0.22 | 13.79 | | | 31.58 | 4.52 | 0.2 ±0.02 |
| S142C/M159C-DA | 0.15 | 0.44 | 0.17 | 8.81 | | | 2.78 | 1.30 | 0.2 ±0.03 |
| Q107C/H182C-DA | 0.26 | 0.43 | 0.08 | 18.24 | | | 3.17 | 1.18 | 0.1 ±0.01 |
| D91C/Y236C-DA | 0.12 | 0.41 | 0.17 | 7.46 | | | 3.16 | 1.29 | 0.2 ±0.03 |
| D91C/A230C-DA | 0.11 | 0.39 | 0.19 | 8.94 | | | 3.52 | 1.11 | 0.2 ±0.02 |
| M149C/A230C-DA | 0.17 | 0.19 | 0.19 | 8.28 | | | 4.35 | 1.37 | 0.2 ±0.05 |
| S142C/K202C-DA | 0.14 | 0.43 | 0.16 | 6.97 | | | 5.03 | 1.82 | 0.2 ±0.04 |

C) Fit results of the time resolved acceptor anisotropy sensitized by FRET.

| Sample | b_1 | ρ_1 [ns] | b_2 | ρ_2 [ns] | b_∞ | ρ_∞ [ns] | $\chi_{r,sum}^2$ | $\chi_{r,diff}^2$ | $r_{ss} \pm \text{std}$ |
|----------------|-------|---------------|-------|---------------|------------|--------------------|------------------|-------------------|-------------------------|
| E135C/Y236C-DA | 0.22 | 0.16 | 0.10 | 1.28 | -0.01 | 285 | 3.31 | 1.10 | 0.0 ±0.03 |
| Q107C/Y236C-DA | 0.10 | 0.32 | 0.04 | 1.78 | -0.03 | 61887 | 2.44 | 1.11 | 0.0 ±0.08 |
| D91C/M159C-DA | 0.003 | 0.31 | | | 0.005 | 15 | 1.79 | 1.08 | 0.0 ±0.03 |
| D91C/H182C-DA | 0.02 | 0.24 | | | 0.007 | 154633 | 1.84 | 1.07 | 0.0 ±0.02 |
| S142C/M159C-DA | 0.07 | 0.40 | | | 0.03 | 18 | 2.78 | 0.98 | 0.0 ±0.11 |
| Q107C/H182C-DA | 0.09 | 0.34 | | | 0.01 | 5 | 2.50 | 0.98 | 0.0 ±0.03 |
| D91C/Y236C-DA | 0.04 | 0.59 | | | -0.03 | 59417 | 2.43 | 1.05 | -0.0±0.07 |
| D91C/A230C-DA | 0.04 | 0.28 | | | 0.02 | 3 | 2.28 | 0.96 | 0.0 ±0.07 |
| M149C/A230C-DA | 0.06 | 0.51 | | | 0.08 | 45 | 1.98 | 1.01 | 0.1 ±0.14 |
| S142C/K202C-DA | 0.10 | 0.82 | | | -0.03 | 164115 | 3.34 | 1.19 | -0.0±0.08 |

Supplementary Table 3 Static FRET lines

| Sample | FRET Lines |
|-----------------------|--|
| E135C/Y236C-DA | $(0.7798/0.35)/((4.0565/((-0.0381*\langle\tau_{D(A)}\rangle_f^3)+(0.2694*\langle\tau_{D(A)}\rangle_f^2)+0.5465*\langle\tau_{D(A)}\rangle_f + -0.0505))-1)$ |
| Q107C/Y236C-DA | $(0.7718/0.39)/((3.8588/((-0.0497*\langle\tau_{D(A)}\rangle_f^3)+(0.3622*\langle\tau_{D(A)}\rangle_f^2)+0.3548*\langle\tau_{D(A)}\rangle_f + -0.0509))-1)$ |
| D91C/M159C-DA | $(0.6329/0.35)/((3.5397/((-0.0546*\langle\tau_{D(A)}\rangle_f^3)+(0.2997*\langle\tau_{D(A)}\rangle_f^2)+0.6392*\langle\tau_{D(A)}\rangle_f + -0.0600))-1)$ |
| D91C/H182C-DA | $(0.7720/0.38)/((4.0316/((-0.0386*\langle\tau_{D(A)}\rangle_f^3)+(0.2699*\langle\tau_{D(A)}\rangle_f^2)+0.5508*\langle\tau_{D(A)}\rangle_f + -0.0508))-1)$ |
| S142C/M159C-DA | $(0.7940/0.34)/((4.1353/((-0.0370*\langle\tau_{D(A)}\rangle_f^3)+(0.2645*\langle\tau_{D(A)}\rangle_f^2)+0.5502*\langle\tau_{D(A)}\rangle_f + -0.0515))-1)$ |
| Q107C/H182C-DA | $(0.7912/0.33)/((4.0832/((-0.0372*\langle\tau_{D(A)}\rangle_f^3)+(0.2681*\langle\tau_{D(A)}\rangle_f^2)+0.5364*\langle\tau_{D(A)}\rangle_f + -0.0492))-1)$ |
| D91C/Y236C-DA | $(0.7578/0.38)/((3.7890/((-0.0424*\langle\tau_{D(A)}\rangle_f^3)+(0.2987*\langle\tau_{D(A)}\rangle_f^2)+0.4866*\langle\tau_{D(A)}\rangle_f + -0.0416))-1)$ |
| D91C/A230C-DA | $(0.5692/0.41)/((3.3669/((-0.0621*\langle\tau_{D(A)}\rangle_f^3)+(0.2944*\langle\tau_{D(A)}\rangle_f^2)+0.7327*\langle\tau_{D(A)}\rangle_f + -0.0708))-1)$ |
| M149C/A230C-DA | $(0.7802/0.41)/((3.9010/((-0.0397*\langle\tau_{D(A)}\rangle_f^3)+(0.2882*\langle\tau_{D(A)}\rangle_f^2)+0.4903*\langle\tau_{D(A)}\rangle_f + -0.0425))-1)$ |
| S142C/K202C-DA | $(0.7743/0.42)/((4.1254/((-0.0371*\langle\tau_{D(A)}\rangle_f^3)+(0.2576*\langle\tau_{D(A)}\rangle_f^2)+0.5814*\langle\tau_{D(A)}\rangle_f + -0.0551))-1)$ |

Supplementary Table 4 Dynamic FRET Lines

| Sample | Transition States | Dynamic FRET Lines |
|----------------|-------------------|---|
| E135C/Y236C-DA | 1-2 | $0.1922/(0.35*((1/2.6182+1/3.4779)-(1.1192*\langle\tau_{D(A)}\rangle_f +- 0.4056))/(3.4779*2.6182))-1/4.0565))$ |
| | 2-3 | $0.1922/(0.35*((1/0.4107+1/2.6180)-(2.1957*\langle\tau_{D(A)}\rangle_f +- 3.1428))/(2.6180*0.4107))-1/4.0565))$ |
| | 3-1 | $0.1922/(0.35*((1/3.4779+1/0.4110)-(2.1484*\langle\tau_{D(A)}\rangle_f +- 3.9927))/(0.4110*3.4779))-1/4.0565))$ |
| Q107C/Y236C-DA | 1-2 | $0.1898/(0.39*((1/2.5947+1/0.8567)-(1.6334*\langle\tau_{D(A)}\rangle_f +- 1.6540))/(0.8567*2.5947))-1/3.8588))$ |
| | 2-3 | $0.1898/(0.39*((1/0.1616+1/2.5950)-(3.1465*\langle\tau_{D(A)}\rangle_f +- 5.5739))/(2.5950*0.1616))-1/3.8588))$ |
| | 3-1 | $0.1898/(0.39*((1/0.8567+1/0.1620)-(3.3811*\langle\tau_{D(A)}\rangle_f +- 2.1081))/(0.1620*0.8567))-1/3.8588))$ |
| D91C/M159C-DA | 1-2 | $0.1788/(0.35*((1/0.3363+1/0.6581)-(2.4540*\langle\tau_{D(A)}\rangle_f +- 1.1029))/(0.6581*0.3363))-1/3.5397))$ |
| | 2-3 | $0.1788/(0.35*((1/2.9473+1/0.3360)-(2.0567*\langle\tau_{D(A)}\rangle_f +- 3.1084))/(0.3360*2.9473))-1/3.5397))$ |
| | 3-1 | $0.1788/(0.35*((1/0.6581+1/2.9470)-(1.5663*\langle\tau_{D(A)}\rangle_f +- 1.6555))/(2.9470*0.6581))-1/3.5397))$ |
| D91C/H182C-DA | 1-2 | $0.1915/(0.38*((1/1.4315+1/0.5166)-(2.1498*\langle\tau_{D(A)}\rangle_f +- 1.7526))/(0.5166*1.4315))-1/4.0316))$ |
| | 2-3 | $0.1915/(0.38*((1/3.0955+1/1.4310)-(1.3838*\langle\tau_{D(A)}\rangle_f +- 1.1915))/(1.4310*3.0955))-1/4.0316))$ |
| | 3-1 | $0.1915/(0.38*((1/0.5166+1/3.0950)-(1.9710*\langle\tau_{D(A)}\rangle_f +- 3.0072))/(3.0950*0.5166))-1/4.0316))$ |
| S142C/M159C-DA | 1-2 | $0.1920/(0.34*((1/2.2646+1/2.8450)-(1.2376*\langle\tau_{D(A)}\rangle_f +- 0.7078))/(2.8450*2.2646))-1/4.1353))$ |
| | 2-3 | $0.1920/(0.34*((1/1.7475+1/2.2650)-(1.4044*\langle\tau_{D(A)}\rangle_f +- 1.0204))/(2.2650*1.7475))-1/4.1353))$ |
| | 3-1 | $0.1920/(0.34*((1/2.8450+1/1.7470)-(1.3365*\langle\tau_{D(A)}\rangle_f +- 0.9823))/(1.7470*2.8450))-1/4.1353))$ |
| Q107C/H182C-DA | 1-2 | $0.1938/(0.33*((1/0.4287+1/1.8140)-(2.2672*\langle\tau_{D(A)}\rangle_f +- 2.3565))/(1.8140*0.4287))-1/4.0832))$ |
| | 2-3 | $0.1938/(0.33*((1/2.7692+1/0.4290)-(2.1777*\langle\tau_{D(A)}\rangle_f +- 3.2727))/(0.4290*2.7692))-1/4.0879))$ |
| | 3-1 | $0.1938/(0.33*((1/1.8140+1/2.7690)-(1.3466*\langle\tau_{D(A)}\rangle_f +- 1.0030))/(2.7690*1.8140))-1/4.0879))$ |
| D91C/Y236C-DA | 1-2 | $0.1763/(0.38*((1/0.3277+1/1.7191)-(2.0601*\langle\tau_{D(A)}\rangle_f +- 1.8213))/(1.7191*0.3277))-1/3.0660))$ |
| | 2-3 | $0.1763/(0.38*((1/0.2584+1/0.3280)-(2.9289*\langle\tau_{D(A)}\rangle_f +- 0.8539))/(0.3280*0.2584))-1/3.0660))$ |
| | 3-1 | $0.1763/(0.38*((1/1.7191+1/0.2580)-(2.2776*\langle\tau_{D(A)}\rangle_f +- 2.1956))/(0.2580*1.7191))-1/3.0660))$ |
| D91C/A230C-DA | 1-2 | $0.1691/(0.41*((1/1.9750+1/0.7697)-(1.4054*\langle\tau_{D(A)}\rangle_f +- 0.7674))/(0.7697*1.9750))-1/3.3669))$ |
| | 2-3 | $0.1691/(0.41*((1/2.8398+1/1.9750)-(0.9357*\langle\tau_{D(A)}\rangle_f +- 0.7674))/(1.9750*2.8398))-1/3.3669))$ |

| | | |
|-----------------------|-----|---|
| | | $+0.2506)/(1.9750*2.8398))-1/3.3669))$ |
| | 3-1 | $0.1691/(0.41*((1/0.7697+1/2.8400-(1.3289*\langle\tau_{D(A)}\rangle_f +- 0.9082)/(2.8400*0.7697))-1/3.3669))$ |
| M149C/A230C-DA | 1-2 | $0.1767/(0.41*((1/2.4741+1/1.8263-(1.1220*\langle\tau_{D(A)}\rangle_f +- 0.2544)/(1.8263*2.4741))-1/3.6193))$ |
| | 2-3 | $0.1767/(0.41*((1/0.3820+1/2.4740-(1.9521*\langle\tau_{D(A)}\rangle_f +- 2.3469)/(2.4740*0.3820))-1/3.6193))$ |
| | 3-1 | $0.1767/(0.41*((1/1.8263+1/0.3820-(2.0101*\langle\tau_{D(A)}\rangle_f +- 1.8497)/(0.3820*1.8263))-1/3.6193))$ |
| S142C/K202C-DA | 1-2 | $0.1877/(0.42*((1/2.4587+1/3.4466-(1.1154*\langle\tau_{D(A)}\rangle_f +- 0.3782)/(3.4466*2.4587))-1/4.1254))$ |
| | 2-3 | $0.1877/(0.42*((1/0.3370+1/2.4590-(2.3042*\langle\tau_{D(A)}\rangle_f +- 3.2171)/(2.4590*0.3370))-1/4.1254))$ |
| | 3-1 | $0.1877/(0.42*((1/3.4466+1/0.3370-(2.2527*\langle\tau_{D(A)}\rangle_f +- 4.3154)/(0.3370*3.4466))-1/4.1254))$ |

Supplementary Methods

Ensemble Time Correlated Single Photon Counting (eTCSPC) analysis

Time resolved fluorescence decays ($F(t)$) were described using a multi-exponential model as follows

$$F(t) = \sum_i x^{(i)} \exp\left(-\frac{t}{\tau^{(i)}}\right), \quad (1)$$

where $x^{(i)}$ is the i -th population fraction and $\tau^{(i)}$ is the fluorescence lifetime of that population. The average lifetimes for each species and the overall average fluorescence lifetimes are

$$\langle \tau \rangle_x = \sum_i x^{(i)} \cdot \tau^{(i)}, \text{ and } \langle \tau \rangle_f = \frac{\sum_i x^{(i)} \cdot (\tau^{(i)})^2}{\langle \tau \rangle_x}, \quad (2)$$

respectively.

To model a superposition of multiple Gaussian-distributed FRET states, with population fractions “ j ” corresponding to $x_{DA}^{(j)}$, one needs to generate the interdye distance distribution $p(R_{DA})$ as follows

$$p(R_{DA}) = \sum_{j=1,2,3} x_{DA}^{(j)} \frac{1}{\sqrt{2\pi} \cdot \sigma_{DA}} \exp\left(-\frac{(R_{DA} - \langle R_{DA}^{(j)} \rangle)^2}{2 \sigma_{DA}^2}\right), \quad (3)$$

where σ_{DA} is the width of the distribution, which comes from the Accessible Volume of the fluorescent dyes. $\langle R_{DA} \rangle$ is the mean interdye distance. The final model included a no FRET state, whose fluorescence decay is the same as that for the donor-only ($D_{(0)}$) control.

$$F(t) = \sum_{i=1,2} x_{D(0)}^i \exp\left(-\frac{t}{\tau_{D(0)}^{(i)}}\right) \cdot \sum_{j=1,2,3} x_{DA}^{(j)} \int p(R_{DA}^{(j)}) \exp(-t \cdot k_{RET}^{(j)}) dR_{DA} \quad (4)$$

where $k_{RET}^{(j)}$ is the rate of energy transfer of the j -th FRET state and is a function of the interdye distance

$$k_{RET} = k_F \langle \kappa^2 \rangle \left(\frac{R_0}{R_{DA}}\right)^6, \quad (5)$$

where k_F is the radiative rate constant, R_0 is the Förster radius, $\langle \kappa^2 \rangle$ is the interdye orientation factor. The isotropic assumption ($\langle \kappa^2 \rangle = \frac{2}{3}$) was verified by time resolved anisotropy.

To obtain the time-resolved anisotropy ($r(t)$), fluorescence intensity decays in both polarizations (F_{\parallel} and F_{\perp} with respect to the excitation source) were collected to determine the proper detection-efficiency factor as

$$F_{\parallel}(t) = \frac{1}{3} F(t) [1 + 2r(t)] \quad (6)$$

$$F_{\perp}(t) = \frac{1}{3} F(t) [1 - r(t)]$$

Time-resolved anisotropy and fluorescence intensity decays for both polarizations were jointly fit using a multi-exponential decay function

$$r(t) = \frac{F_{\parallel}(t) - F_{\perp}(t)}{F_{\parallel}(t) + 2 \cdot G \cdot F_{\perp}(t)} \quad (7)$$

$$r(t) = \sum_i b_i \exp\left(-\frac{t}{\rho_i}\right)$$

where b_i is the fractional anisotropy, ρ is the rotational correlation time, and G is a detection correction parameter (i.e. G-factor).

Accessible Volume (AV) simulations to estimate measured distance

To account for dye linker mobility, we calculated AVs for donor and acceptor dyes attached to each labeling site. The AV approximates the dye as a hard sphere connected to the protein via a linker modeled as a flexible cylinder with the dimensions based on the chemical structure¹⁻⁴. For Alexa 488, the five-carbon linker length was set to 20 Å, the width of the linker is 4.5 Å and three dye radii 5.0 Å, 4.0 Å, and 1.5 Å. Similarly, for

Alexa 647 the dimensions used were: length = 22 Å, width = 4.5 Å and the three dye radii 11.0 Å, 3.0 Å, and 1.5 Å.

For each FRET pair, we calculated the distance between dye mean positions ($\langle R_{DA} \rangle$) within the AV.

$$\langle R_{DA} \rangle = |\langle R_D^{(i)} - R_A^{(j)} \rangle| = \frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m |R_D^{(i)} - R_A^{(j)}|, \quad (8)$$

where $\overline{R_D^{(i)}}$ and $\overline{R_A^{(j)}}$ are all the possible positions (n, m) that the donor and acceptor fluorophores adopt.

Analysis of Multiparameter Fluorescence Detection (MFD) for single-molecule FRET (smFRET) experiments

Single molecule events were identified as events with intensity 2σ above the mean background signal. Cut-off times varied from sample to sample with a minimum of 60 photons collected for each burst^{5,6}. Each burst was then processed and fit using a maximum-likelihood algorithm as previously described⁷. The single molecule selection rules are: the difference in single molecule duration on green channels given donor excitation (T_{GX}) and duration on red channels given direct acceptor excitation (T_{RR}) was $-1.5 \text{ ms} < T_{GX} - T_{RR} < 1.5 \text{ ms}$; and satisfy the FRET Stoichiometry (S_{PIE}) parameter of $0.13 < S_{PIE} < 0.6$, which selects for bursts which both fluorophores present. The PIE stoichiometry is given by:

$$S_{PIE} = \frac{F_{R|D} - \beta F_{R|A} - \alpha F_{G|D} + \gamma F_{G|D}}{F_{R|D} - \beta F_{R|A} - \alpha F_{G|D} + \gamma F_{G|D} + F_{R|A}}, \quad (9)$$

where $F_{G|D}$, $F_{R|D}$, and $F_{R|A}$ are background-corrected fluorescence intensities measured in the donor channel after donor excitation ($G|D$), in the acceptor channel after donor excitation ($R|D$) and direct acceptor excitation ($R|A$), respectively, α is the correction factor for donor crosstalk into the acceptor channel, β is the correction factor for direct excitation of the acceptor by the donor excitation source, and $\gamma = \frac{\Phi_{FA}}{\Phi_{FD(0)}} \cdot \frac{g_R}{g_G}$ is a function of the quantum yield of the donor $\Phi_{FD(0)}$, the quantum yield of the acceptor Φ_{FA} and detection efficiencies g . Fluorescent bursts were plotted in 2D histograms (Origin 8.6, OriginLab Co).

Calculation of the Static and Dynamic FRET Lines

The relationship between FRET intensity (F_D/F_A) and the FRET lifetime ($\langle \tau_{D(A)} \rangle$) is different for static molecules and those undergoing dynamic averaging. To facilitate interpretation of the MFD experiments, we include guidelines that relate FRET intensity (F_D/F_A) to the FRET lifetime ($\langle \tau_{D(A)} \rangle$) for molecules that are static and those that are dynamic. The parametric line that relates the FRET indicators, $E_{static,L}$ (i.e. FRET Efficiency) or $(F_D/F_A)_{static,L}$ and $\langle \tau_{D(A)} \rangle_f$, in the absence of dynamics is defined as

$$E_{static,L} = 1 - \frac{\sum_{i=0} A^{(i)} \cdot \langle \tau_{D(A)} \rangle_f^i}{\tau_{D(0)}} \quad (10)$$

$$\left(\frac{F_D}{F_A} \right)_{static,L} = \left(\gamma' \cdot \left(\frac{\tau_{D(0)}}{\sum_{i=0} A^{(i)} \cdot \langle \tau_{D(A)} \rangle_f^i} - 1 \right) \right)^{-1},$$

where A_i are the coefficients of an empirical polynomial function that takes into account the intrinsic linker dynamics (L) of the dyes, and $\gamma' = \Phi_{FA}/\Phi_{FD(0)}$, which is the ratio of the quantum yields of acceptor to donor. Hereafter, we used the simplified notation of $F_{G|D} = F_D$ and $F_{R|D} = F_A$.

The parametric line that relates the FRET indicators in the presence of dynamics is defined as

$$E_{dyn,L} = 1 - \frac{\langle \tau_{D(A)}^{(1)} \rangle_f \cdot \langle \tau_{D(A)}^{(2)} \rangle_f}{\tau_{D(0)} \left(\langle \tau_{D(A)}^{(1)} \rangle_f + \langle \tau_{D(A)}^{(2)} \rangle_f - \sum_{i=0}^3 C_{i,L} (\langle \tau_{D(A)} \rangle_f)^i \right)} \quad (11)$$

$$\left(\frac{F_D}{F_A}\right)_{dyn,L} = \frac{1}{\gamma' \cdot \tau_{D(0)}} \cdot \frac{\langle \tau_{D(A)}^{(1)} \rangle_f \langle \tau_{D(A)}^{(2)} \rangle_f}{\left(\langle \tau_{D(A)}^{(1)} \rangle_f + \langle \tau_{D(A)}^{(2)} \rangle_f - \sum_{i=0}^3 C_{i,L} (\langle \tau_{D(A)} \rangle_f)^i \right) - \frac{\langle \tau_{D(A)}^{(1)} \rangle_f \langle \tau_{D(A)}^{(2)} \rangle_f}{\tau_{D(0)}}$$

where C_i is the coefficient of an empirical polynomial function that takes into account the intrinsic linker dynamics (L) of the dyes. The superscripts (1) and (2) indicate the two limiting states that are in exchange, which will be connected by the dynamic FRET line.

Determination of Quantum Yields

We assume that only dynamic quenching takes place such that $\Phi_{FD(0)}$ and Φ_{FA} are proportional to the species-averaged fluorescence lifetime $\langle \tau_{D(A)} \rangle_x$ of donor and acceptor, respectively. As reference samples we used Alexa488-labeled DNA $\langle \tau_{D(0)} \rangle_x = 4.0$ ns, $\Phi_{FD(0)} = 0.8$ and for the acceptor we used Cy5-labeled DNA with $\langle \tau_A \rangle_x = 1.17$ ns and $\Phi_{FA} = 0.32$ ⁸. The obtained donor and acceptor quantum yields are presented in Supplementary Table 1B, 1C. For Alexa488 and Alexa647, we considered a reduced Förster distance of 52 Å where we assumed isotropic reorientation of the dyes using $\kappa^2 = 2/3$ due to the long linkers.

Estimation of $\langle \kappa^2 \rangle$ and κ^2 -distributions along with the associated uncertainty

Experimentally, we probe if the assumption of $\kappa^2 = 2/3$ is justifiable or not. Considering that fluorophores follow the “wobble-in-cone” model⁹, it is possible to calculate a distribution of all possible values of κ^2 . For that, we need the residual anisotropies obtained by fitting the time resolved anisotropy ($r_\infty = b_3$, Eq. 7) for the various sources (Donly - donor, A - acceptor and A(D) – FRET-sensitized emission of acceptor). Then, all κ^2 values can be described as follows.

$$\begin{aligned} \kappa^2 = & \frac{2}{3} + \frac{2}{3} S_D^{(2)} S^{(2)}(\beta_1) + \frac{2}{3} S_A^{(2)} S^{(2)}(\beta_2) \\ & + \frac{2}{3} S_D^{(2)} S_A^{(2)} (S^{(2)}(\delta) + 6S^{(2)}(\beta_1)S^{(2)}(\beta_2) + 1 + 2S^{(2)}(\beta_1) + 2S^{(2)}(\beta_1) + 2S^{(2)}(\beta_2) \\ & - 9\cos\beta_1\cos\beta_2\cos\delta) \end{aligned} \quad (12)$$

where β_1 and β_2 are the angles between the symmetry axes of each dye rotation, and δ is the angle between them. The necessary second-rank order parameters $S^{(2)}$ are defined by:

$$\begin{aligned} S^{(2)}(\delta) &= \frac{1}{2} (3 \cos^2 \delta - 1) = \frac{r_{\infty, A(D)}}{r_0 S_D^{(2)} S_A^{(2)}} \\ S^{(2)}(\beta_1) &= \frac{1}{2} (3 \cos^2 \beta_1 - 1) \quad \text{and} \\ S^{(2)}(\beta_2) &= \frac{1}{2} (3 \cos^2 \beta_2 - 1) \end{aligned} \quad (13)$$

where r_0 is the fundamental anisotropy of the dyes, which were 0.38 and 0.39 for the donor and acceptor fluorophores, respectively⁴. The dye motions are characterized by the second-rank order parameters $S_D^{(2)}$ and $S_A^{(2)}$ as

$$\begin{aligned} \text{Donor:} \quad & \frac{1}{2} \cos^2 \theta_{\text{disk}} = \sqrt{\frac{r_{\infty, \text{Donly}}}{r_0}} = -S_D^{(2)} \\ \text{Acceptor:} \quad & \frac{1}{2} \cos^2 \theta_{\text{cone}} (1 + \cos \theta_{\text{cone}}) = \sqrt{\frac{r_{\infty, A}}{r_0}} = S_A^{(2)}, \end{aligned} \quad (14)$$

From all possible orientations and combinations, a κ^2 - distribution and its corresponding arithmetic mean ($\langle \kappa^2 \rangle$) can be determined and compared to the assumed $\kappa^2 = 2/3$ (Fig. 3).

Error propagation

To estimate the statistical uncertainty in the experimental distances, we used an error propagation rule, which considered the uncertainty associated with the minimization of χ^2 , ($\Delta\chi^2$), along with the uncertainty associated with the unknown orientation of the dyes, ($\Delta\kappa^2$). For $\Delta\chi^2$, we first calculate the maximum allowed $\chi_{r,max}^2$ for a given confidence-level (P ; e.g. for 2σ or $P = 0.95$) as

$$\chi_{r,max}^2(P) = \chi_{r,min}^2 \cdot \left[1 + \frac{n}{v} \cdot cdf^{-1}(F(n, v, P)) \right], \quad (15)$$

given the minimum $\chi_{r,min}^2$ obtained by the fit and identified the distances R_{DA} at these particular values of χ^2 . We then obtain $\Delta\kappa^2$ by determining the arithmetic mean of the calculated κ^2 distribution. Then, the overall uncertainty is determined as

$$\varepsilon(\kappa^2, k_{FRET}) = \Delta R_{DA}^2(\kappa^2) + \Delta R_{DA}^2(k_{FRET}), \quad (16)$$

which is presented in Supplementary Table 1.

Global Analysis and goodness of fit.

To test the goodness of the fit for the fluorescence decays, we rely on i) the visual inspection of residuals ii) the autocorrelation of the residuals; and ii) the F-test^(10,11). The F-test compares the ratio of two independent χ^2 variables, each with their own numbers of degrees of freedom with a similar equation as the one shown in Eq. 15.

The global analysis improved χ^2 from 3- to 2-FRET distributed states by 0.15. This change, given the changes in the number of degrees of freedom, corresponds to a confidence of 99.99% that the 3-FRET distributed states better describes the experimental observations. Particularly because the number of degrees of freedom is significantly reduced from an independent fit to a global analysis. Additionally, the global analysis allows for an unbiased assignment of distances to particular configuration states, based on the population fractions.

References

1. Sindbert, S. et al. *Journal of the American Chemical Society* **133** (8), 2463 (2011).
2. Cai, Q. et al. *Biophys J* **93** (6), 2110 (2007).
3. Muschielok, A. et al. *Nat Methods* **5** (11), 965 (2008).
4. Kalinin, S. P., T.; Sindbert, S.; Rothwell, P. J.; Berger, S.; Restle, T.; Goody, R. S.; Gohlke, H.; Seidel, C. A. M. *Nature Methods* (2012).
5. Kühnemuth, R. and Seidel, C. A. M. *Single Molecules* **2** (4), 251 (2001).
6. Schaffer, J. et al. *Journal of Physical Chemistry A* **103** (3), 331 (1999).
7. Maus, M. et al. *Anal Chem* **73** (9), 2078 (2001).
8. Woźniak, A. K. et al. *Proc.Natl.Acad.Sci.USA.* **105**, 18337 (2008).
9. Dale, R. E., Eisinger, J., and Blumberg, W. E. *Biophys J* **26** (2), 161 (1979).
10. Box George, E. P. *Annals of the New York Academy of Sciences* **86** (3), 792 (1960).
11. Lakowicz, J. R., (Springer US, 2006).
12. Sisamakos, E. et al. *Methods in enzymology* **475**, 455 (2010).