Multi-domain structure and correlated dynamics determined by self-consistent FRET networks

B. Hellenkamp[1], P. Wortmann[1], F. Kandzia[2], M. Zacharias[2], T. Hugel[1*]

[1]*Institute of Physical Chemistry, University of Freiburg, Germany*

[2]*Physics Department, Technische Universität München, Germany*

**Corresponding author: th@physchem.uni-freiburg.de*

Supplementary Notes

Supplementary Note 1: Static domain arrangement

Testing for dimer asymmetry

The standard procedure for static domain arrangement is described in **Online Methods**. In addition to the standard χ^2-minimization for symmetric structures, we tested for an average asymmetric dimer structure with the following criterion:

$$\chi^2_{dimer} = \sum_{i=1}^{n}\sum_{k=1}^{2}\left(\frac{R_{k,i}^{mod} - R_i^{exp} + \sigma R_i^{exp}\cdot 0.5(-1)^k\rho_i}{\Delta R_i^{exp}}\right)^2;\ \rho_i = [0;1] \qquad (1)$$

The criterion virtually sets a spring between the average dye positions of each FRET pair. The stiffness of the spring is the reciprocal distance error ΔR_i^{exp}. The right summand enables dimer asymmetry with left and right distances for k=2 monomers. The fraction of asymmetry ρ_i is a free fit parameter that becomes larger than 0 for asymmetric dye pairs. The maximum difference between left and right distance is the standard deviation of the experimental distance distribution σ_i^{exp}. This procedure resulted in two symmetric structures with χ^2=0.2 and ρ=0 for all distance distributions and in an asymmetric structure with χ^2=0.5 favoring the symmetric solution with an RMSD of 2 Å. Thus, the average symmetric structure is more likely than an average asymmetric structure; but asymmetric intermediates are possible.

Efficiency average and distance average

The translational diffusion of the dyes within their accessible volume normally occurs on a timescale slower than the fluorescence lifetime and faster than the photon binning time. Therefore, the efficiencies of all inter-dye distance combinations are averaged. This average can deviate from the distance average or from the distance between the center positions of the accessible volumes. This averaging artifact is significant for large accessible volumes.
For a specific model structure the deviation between efficiency average and distance average can be calculated from the dye accessible volumes. However, the relative orientations of the accessible volumes are unknown before the minimization procedure of the domain arrangement. A recalculation for the accessible volumes for every possible model structure during minimization would be computationally extremely expensive. Thus, we calculate the orientation-independent effective radii of the accessible volumes of donor and acceptor dye for each FRET pair. With the average radius $R_{AV}=(R_{AV,D}+R_{AV,A})/2$, the distance-averaged distance $R_{<R>}$ can be transformed to the efficiency-averaged distance $R_{<E>}$:

$$R_{\langle E\rangle} = f\left(R_{\langle R\rangle}\right) = R_0\left(\left(\sum_{i=-1}^{1}\left(1+\left(\frac{R_{\langle R\rangle}+iR_{AV}}{R_0}\right)^6\right)^{-1}\right)^{-1}-1\right)^{1/6} \qquad (2)$$

Now, the model distances calculated between the center positions of the accessible volumes can be transformed into efficiency-averaged model distances. But, this would be still computationally expensive because the distance had to be transformed for each model structure. More conveniently is the transformation of experimental distances to distance-averaged distances prior to the minimization procedure with the inverse function:

$$R_{\langle R\rangle} = f^{-1}\left(R_{\langle E\rangle}\right) \qquad (3)$$

The inverse function is analytically not solvable and therefore realized with a numeric look-up table calculated prior to the minimization procedure. **Figure N1** shows different transformations for different average dye radii. The small remaining deviations between this simplified model and the real average for the complete accessible volumes is verified after the minimization procedure for the final model structure.

If the average radii are not known, the length of the dye linker can be used instead as a maximal possible radius for the accessible volume and then the deviation $R_{<E>}$-$R_{<R>}$ can be used as an additional distance-dependent uncertainty which is geometrically added to the uncertainties described in **Online Methods**.

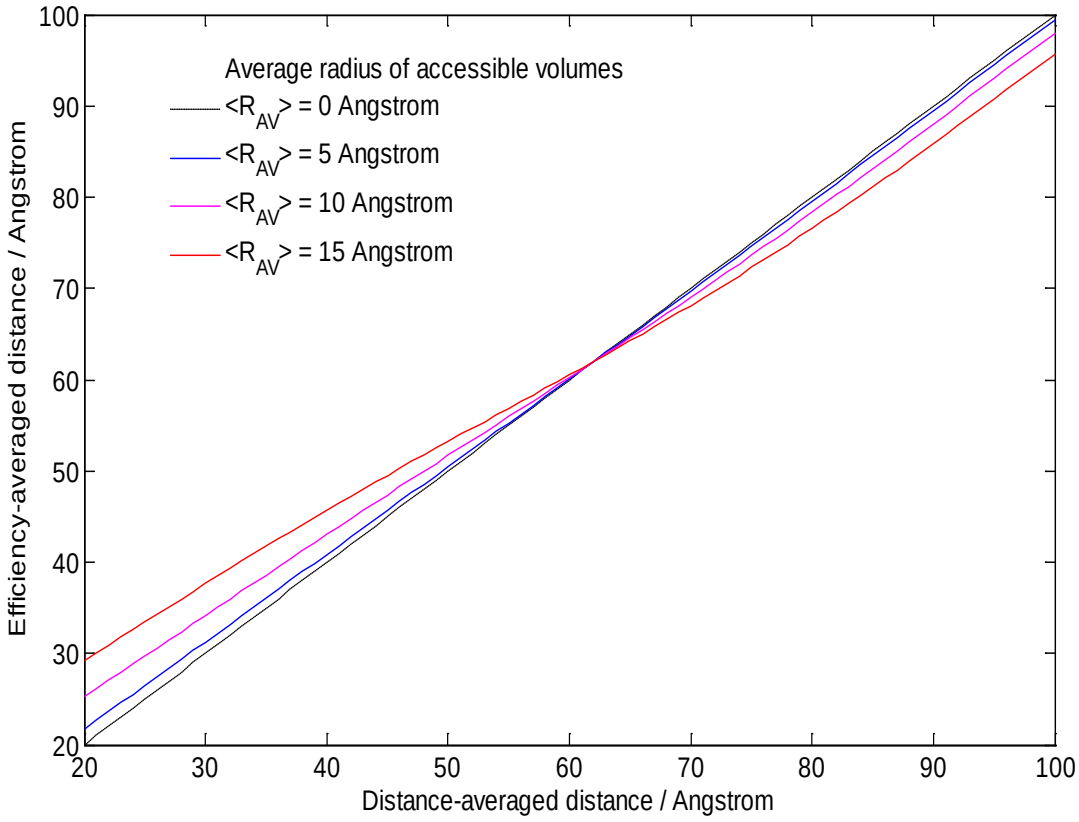

Figure N1 | Transformation between different distance-averaged and efficiency-averaged distances
The function between distance-averaged and efficiency-averaged distance is plotted for different average dye radii $<R_{AV}>$ using equation (2) and a Förster radius of R_0=65 Å.

Supplementary Note 2: Determination of precise distances.

FRET efficiencies *E* and stoichiometries *S* of single molecules labeled with FRET dyes are correlated with each other in a scatter plot (**Fig. N2a**) to exclude homo dimers and obtain the following photophysical parameters[1]: the fraction *d* of the direct excitation of the acceptor dye with the donor laser, fraction *c* of the leakage of the donor fluorescence into the acceptor channel, the gamma factor $\gamma=Q_A\eta_A/(Q_D\eta_D)$ – describing relative quantum yields *Q* of donor D and acceptor A and the relative detection efficiencies η of donor channel D and acceptor channel A – and the beta factor $\beta=(\sigma_A I_A)/(\sigma_D I_D)$ – describing the relative absorption cross sections σ of donor D and acceptor A and the relative intensities *I* of donor excitation D and acceptor excitation A.

The macro-time data of the detected fluorescence photons is binned for certain time intervals (normally 1 ms) to obtain intensities *N* from the donor channel (D in subscript) and the acceptor channel (A in subscript) separately for donor excitation (no superscript) and for acceptor excitation (A in superscript). The corrected FRET efficiency then is:

$$E = \left(\frac{\gamma N_D}{N_A - dN_A^A - cN_D} + 1\right)^{-1} \qquad (4)$$

The corrected stoichiometry is:

$$S = \left(\frac{N_A^A/\beta}{N_A - dN_A^A - cN_D + \gamma N_D} + 1\right)^{-1} \qquad (5)$$

We pre-correct the binned intensities to obtain corrected efficiency and stoichiometry histograms. This is a quite robust procedure, because the corrected scatter plots directly visualize correctly chosen parameters and possible deviations due to measurement artifacts. The manipulated photon statistics are taken into account if using the probability distribution analysis for pre-corrected intensities (**Supplementary Note 5**). Before applying the correction factors to the burst intensities, the average background counts per channel and per binning time window should be subtracted from the respective channels. This is especially important to determine correct direct excitation and leakage factors.

Homo dimers are excluded by selecting only molecule events that hold for: $0.3 > S > 0.7$.

The inter-dye distances are obtained from the corrected efficiencies by applying the Förster theory:

$$R = R_0(1/E - 1)^{1/6} \qquad (6)$$

For each single FRET pair an individual Förster radius is calculated by measuring the donor quantum efficiency Q_D, the spectral overlap *J* and the relative dipole orientation factor κ^2:

$$R_0{}^6 \sim \langle\kappa^2\rangle \cdot Q_D \cdot J \qquad (7)$$

A multi-parameter fluorescence lifetime analysis reveals the position-specific donor quantum yield for a quantum yield Q_{D0} and lifetime τ_{D0} specified by the manufacturer:

$$Q_D = Q_{D0} \cdot \tau_D / \tau_{D0} \qquad (8)$$

Time-resolved anisotropies of donor and acceptor dyes are used to estimate κ^2 (see **Supplementary Note 4**).

If the position-specific quantum yield of the donor cannot be calculated, for example, because of a missing donor only population, it can be canceled out in the following way:

$$R^6 = R_0^{*6} \frac{Q_A \eta_{AD} N_D}{N_A - dN_A^A - cN_D}, \qquad with\ R_0^{*6} = R_0^6 / Q_D \qquad (9)$$

Instead, the acceptor quantum yield must be determined which can be done using the non-quenched acceptor fluorescence after acceptor excitation. Additionally, the relative detection efficiency η_{AD} has to be determined separately which is not recommended.

The fraction of dark states of the acceptor dye has to be checked. In our experiments, the excitation powers were set to 50 µW (about a third of the power at fluorescence intensity saturation for Atto550 and Atto647N) to guarantee low dark state fractions of less than 3% for Atto550 and Atto647N (**Fig. N2b**). In contrast, Alexa647 exhibited a large dark state fraction of 30 % and was therefore not used for distance determination (**Fig. N2c**).

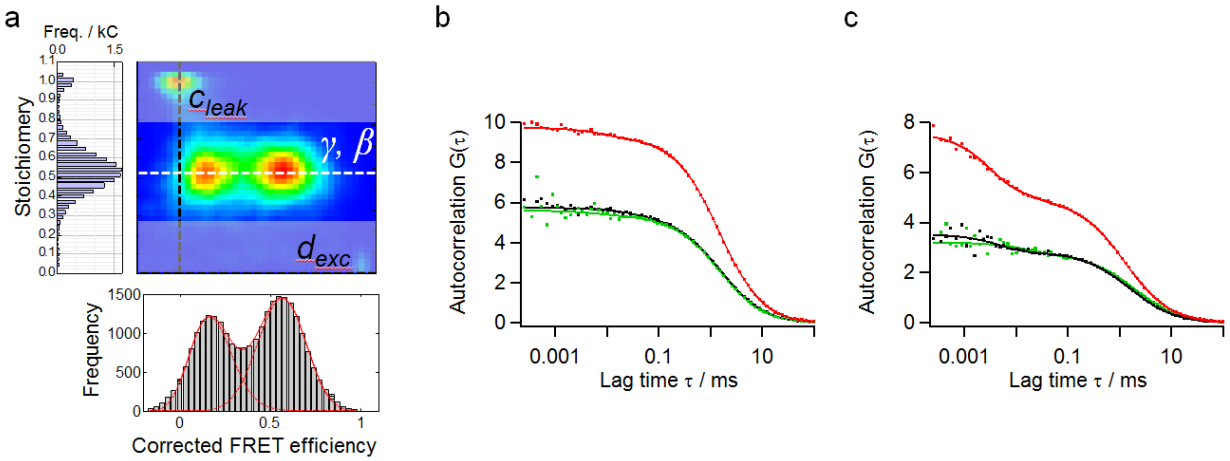

Figure N2 | Global comparison of photophysical correction factors.
(a) FRET efficiencies are correlated with stoichiometries in a scatter plot to exclude homo dimers (shaded part) and to correct FRET efficiencies for all relevant photophysical parameters (see **Online Methods**). **(b)** Autocorrelations of acceptor (Atto647N) photons after donor (Atto550) excitation for two samples with equal acceptor positions, but different donor positions (green and black) and autocorrelation of acceptor photons after acceptor excitation (red). The autocorrelations were fitted with diffusion and blinking terms. The fits revealed dark state fractions of 3 %. **(c)** The same as in (b), but with Alexa647 instead of Atto647N resulting in a large dark state fraction for the acceptor fluorescence of 30 %.

Supplementary Note 3: Generation of self-consistent parameter and distance networks

For each single measurement photophysical correction parameters were calculated. Their consistency was verified by generating a parameter network with redundancies. About 10 distances (FRET pairs) are linked to each position and each distance was measured at least four times.
First, redundancy is given by calculating a factor from different parameters. For example, the gamma factor was calculated from stoichiometry and FRET efficiency for a specific FRET pair, but also from donor and acceptor lifetimes assuming constant detection efficiencies. A disagreement of both values could originate from setup misalignment, other wrong correction factors, blinking or bleaching. As these things are difficult to correct after the measurement, such measurements were simply discarded.
Second, the FRET network connects several correction parameters. For example, the gamma factor (γ) for the positions D1 and A1 (donor positon and acceptor positon of distance 1, respectively) has to be described by a group of other gamma factors for the positions Dx and Ax (see also **Fig. N3a**):

$$\gamma_{A1D1} = \gamma_{A1D2}\gamma_{A2D1}/\gamma_{A2D2}\,, \qquad with\ \gamma_{AxDx} = \eta_{AD}Q_{Ax}/Q_{Dx} \qquad (10)$$

where η_{AD} is the ratio of the detection efficiencies and Q the quantum yield of donor D or acceptor A.

The tolerance for a gamma factor deviation was set to 20%. In this simple case, it is difficult to locate the wrong measurement in case of disagreement. However, within the complete network a disagreement can be directly assigned to one single measurement. The respective measurement is not considered for further analysis as the reasons for such an outlier are normally setup misalignment, wrong labeling or contaminated buffer, which cannot be corrected afterwards.
Further correction factors like direct acceptor excitation or spectral crosstalk were also compared within the distance network. Normally, these factors have to be the same for each measurement. However, in few cases they slightly differ due to spectral shifts of absorption and emission spectra.

The position specific dye anisotropy is also checked within the network. The anisotropy of each dye positions is measured several times with different dye partners. Deviating low donor anisotropy can occur if the donor dye undergoes different dynamic states and if the energy transfer to the nearby acceptor is selective. Deviating low acceptor anisotropy is observed for some average inter-dye distances below 4 nm and may arise from strong coupling.

Local rearrangements within the solution structure that are not resolved in the crystal structures can be detected by analyzing site-specific and state-specific dye anisotropies and their connectivity within the FRET networks.
Figure N3b shows an expected average anti-correlation (long diagonal line) between accessible volume and residual anisotropy for several dye positions within the closed state (blue squares) and the open state (red circles) of Hsp90. The accessible volume is normalized to a complete sphere containing about 7500 simulation points. An accessible volume of 50% already corresponds to a nearly 100% isotropic distribution of dipole orientations. The residual anisotropies are determined as described in **Supplementary Note 4**. Significant deviation (red

area) from the average anti-correlation indicates transient sticking of the dye or local structural rearrangements in solution. The latter is more likely as the deviating positions correlate with lower fit quality to the electron densities of the x-ray structure (normalized real-space R-values > 2).
Importantly, **Figure N3b** indicates correlations between global and local conformational changes. The global states are connected by their labeling positions and indicated here as green lines (positions at domain interfaces) or black lines (other). The accessible volume and residual anisotropy of a dye at a domain interface are mainly affected by the global domain rearrangement. In contrast, a long black connecting line indicates a local structural rearrangement that correlates with the global conformation.
Remarkably, after equilibration of the open structure with unrestrained MD simulations in explicit solvent (see **Online Methods**), the accessible volumes of positions with indicated local rearrangements change towards the expected anti-correlation between accessible volume and anisotropy (red crosses in **Fig. N3b**).

Potential local rearrangements within the domains are screened by excluding single dye positions from the distance network followed by checking the self-consistency of the remaining sub-networks. **Figure N3c** compares the remaining RMSD with the remaining mean combined anisotropies of the sub-networks. A deviation from the correlated relation (red line) towards larger anisotropies and at the same time from the total RMSD (dashed black line) towards smaller RMSDs is an indicator for a local structural rearrangement.

It has to be noted that the distance deviations are recognized best if prior accessible volumes according to the x-ray crystal structure are used. A dynamic adaption of the accessible volumes as done in other publications hinders the recognition of real systematic deviations.

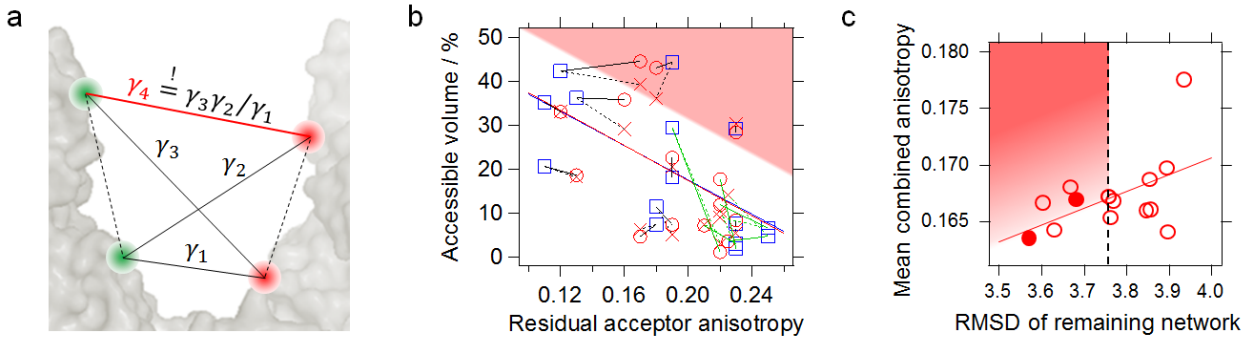

Figure N3 | Network features.
(a) Some correction factors such as the gamma factor γ have to be described by a group of other correction factors. Deviations are directly monitored in larger FRET networks. **(b)** Correlation between accessible volumes and residual acceptor anisotropies for dye positions within the closed structure of Hsp90 (blue squares), the arranged average open structure (red circles) and an open structure refined by MD simulations in explicit solvent (red crosses). The lines indicate corresponding positions, whereat the green lines are located at domain interfaces and the black lines are not. The refinement with MD simulations leads to local rearrangements (red crosses) towards the expected average anti-correlation between accessible volume and anisotropy (long diagonal line). The change of the anisotropy for positon 452 (highlighted with the bold black line in fig. 2c) from the closed to the open state is clearly indicated. Such a change is not expected for positions that do not rearrange upon the global conformational change. In

contrast, for position 298 the anisotropy change (highlighted with the bold green line in fig. 2c) is most likely caused by the reduction of the accessible volume. This position is located at the interface between the M- and N- domain that rearrange upon the global conformational change. That is why the loop position 298 does not need to rearrange in order to explain the anisotropy change. **(c)** Potential local rearrangements within the domains are screened by excluding single dye positions from the distance network and comparing the remaining RMSD with the remaining mean combined anisotropies. A deviation from the correlated relation (red line) towards larger anisotropies and at the same time from the RMSD of the complete network (dashed black line) towards smaller RMSDs is an indicator for a local structural rearrangement.

Supplementary Note 4: Combined residual anisotropy as a selection criterion

The following discussion holds for single distances in the absence of the FRET network.

An average orientation factor of κ^2=2/3 is correct if the relative dye dipole orientations sample all possible combinations on a timescale that is faster than the energy transfer. But, the dye motions are often restricted on different timescales. We experimentally derived a value for which the dye dipole orientations are sufficiently dynamically averaged.

Donor and acceptor anisotropies were calculated from parallel and perpendicularly polarized fluorescence intensities $N_{||}$ and $N_{\perp}$ after excitation with the linearly polarized donor laser or acceptor laser, respectively:

$$r = \frac{N_{||} - GN_{\perp}}{N_{||} + 2GN_{\perp}} \qquad (11)$$

The *G* factor corrects for relative detection efficiencies. For objectives with a high numerical aperture depolarization factors have to be considered (*l1*=0.0308 and *l2*=0.0368 for *NA*=1.2).

We determined the residual dye anisotropies at the timescale of energy transfer using time-resolved anisotropies. Most dye positions revealed time-resolved anisotropies with two relaxation times, τ_1 for fast rotational relaxation of the dye dipole axis (mainly contributing to dynamic averaging of dye dipole orientations) and τ_2 for slow dye diffusion along the protein surface:

$$r(t) = r_1 e^{-t/\tau_1} + r_2 e^{-t/\tau_2} + r_\infty \qquad (12)$$

The sum of the remaining anisotropies r_1 and r_2 and the residual anisotropy r_∞ must equal the fundamental anisotropy r_0 which is estimated to be 0.38. See **Figure N4a** for an example.

The residual anisotropy r_∞ (the lowest anisotropy we can measure) is in most cases lower than the anisotropy value at the time of energy transfer. That means the residual anisotropy considers both dynamic and static averaging. Normally, Förster law accounts only for dynamic averaging; however, static averaging can be included in case of sufficient dynamic averaging, because then the difference between the different averaging domains becomes negligible:

$$\sqrt[6]{\langle \kappa^2(1 \pm 0.5)\rangle} \approx \langle \sqrt[6]{\kappa^2(1 \pm 0.5)}\rangle \qquad (13)$$

To guarantee sufficient dynamic averaging, we use relative small dyes with fast rotational relaxation times, relative large lifetimes and dye dipoles oriented perpendicular to the linker (Atto550 and Atto647N).
The isotropy of a dye can to a large extent compensate the anisotropy of the other FRET dye for an isotropic dipole coupling. Therefore we determined the geometric mean of the time-resolved donor and acceptor anisotropies – i.e. the combined anisotropy – for each FRET pair:

$$r_C = \sqrt{r_{A,\infty}}\sqrt{r_{D,\infty}} \qquad (14)$$

The combined anisotropy directly reflects the transfer anisotropy according to the Soleillet theorem. The transfer anisotropy is not directly measurable as often claimed, because the measured acceptor anisotropy after donor excitation consists of the transfer anisotropy and the acceptor anisotropy. Only acceptors with a very short lifetime would give a good approximation.
With a decreasing upper threshold for the combined anisotropy the root mean square deviation (RMSD) between the remaining experimental distances and the corresponding model distances from the published x-ray crystal structure of the closed conformation decreases (**Fig. N4b**). For an upper threshold of 0.22, the RMSD for the 132 remaining distances converges to 2.8 Å. The steep decrease of the RMSD in dependence on the upper threshold for the combined anisotropy is caused by the dye geometries. Sufficient dye accessibility means both a fast rotational relaxation and slow dye diffusion, but a restricted dye can hinder both motions at the same time.

The deviations of the remaining distances not rejected by the threshold are not significantly correlated with any measured anisotropy value. These deviations probably arise from structural differences between the model structure and the solution structure. Four of the 132 remaining distances were deviating by about 10% despite very low anisotropies for both dyes. For example, the distance 517-298 deviated with 12% from the crystal structure, although the respective combined anisotropy is 0.16. Simulated κ^2 values for the accessible volumes at the positons 517 and 298 resulted in deviations up to +2% or -3.5% when screening all possible geometries between dipole axis and linker. The deviation lie within the theoretical worst case errors for the dye anisotropy values of 0.13 and 0.2.
Following the derivations from[2,3] the theoretical maximum distance error is 18% for anisotropy values of 0.2 for both dyes. However, the probability for such a worst case constellation depends on the geometries of the dyes. Numeric calculations resulted in much less probability for worst case scenarios and a small average error if assuming a freely rotating dipole axis that is oriented perpendicular to the dye linker. Indeed, comparison of experimental distances and anisotropies with model distances reflects this prediction when using indicated dyes. Most distances with a low combined anisotropy showed deviations much below the theoretical worst case errors.

It seems that the rotational relaxation leads to sufficient dynamic averaging and thus to accurate distances in most cases. However, if a dye is completely restricted, free rotational motion can also be hindered. Indeed, the largest combined anisotropy values have been observed for dyes located at domain interfaces and are about 0.27-0.3. The estimation of dye accessibility can be quite erroneous in case of a very small accessible volume or in case of large accessible volume despite high dye anisotropy. In these cases the relative distance uncertainty was increased by an offset of ±5 Å for corresponding dye pairs.

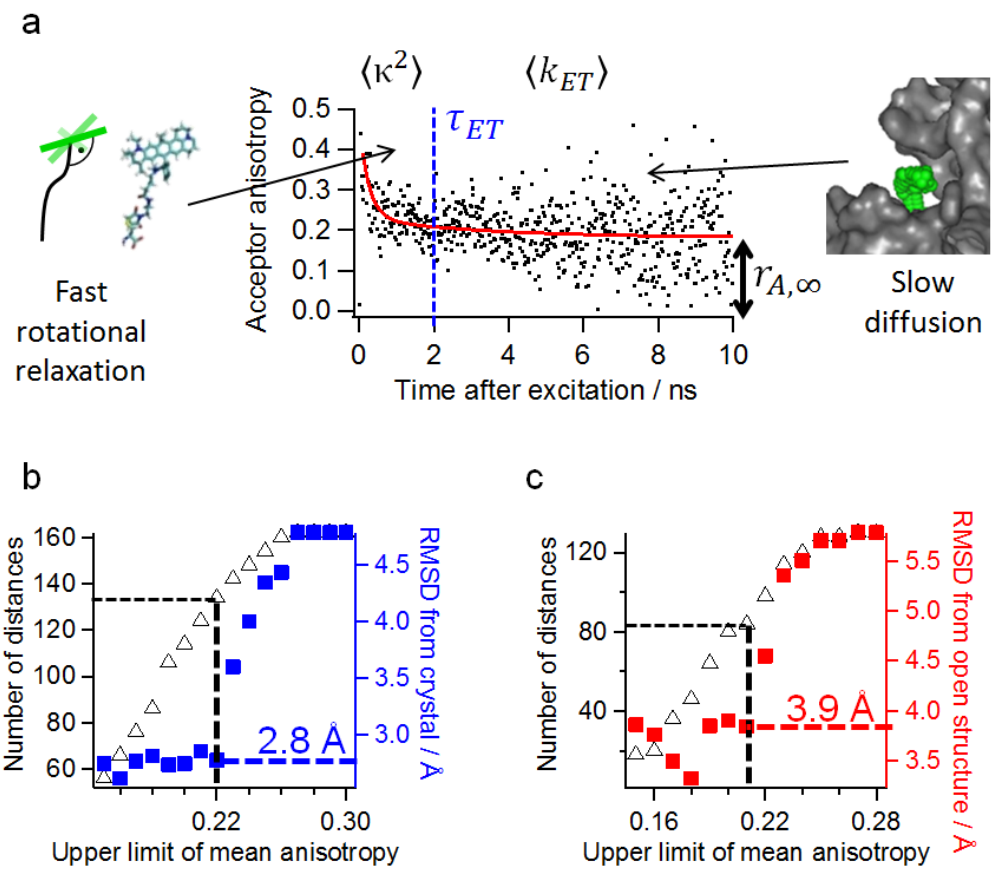

Figure N4 | Time-resolved dye anisotropies.
(a) The time-resolved anisotropies of dyes bound to a larger protein normally consist of a fast decay from rotational relaxation of the dipole axis and of a slow decay from dye diffusion along the protein surface. The residual anisotropies r_∞ of both dyes of a FRET pair is used as part of the selection criterion. **(b)** A threshold value for nearly isotropy dipole coupling is determined for the closed state of Hsp90 (see text). **(c)** The same as in (b), but for the open state of Hps90.

Supplementary Note 5: Determination of distance fluctuations

In order to separate structural heterogeneities from other heterogeneities such as shot noise, a probability distribution analysis has been derived accounting for the manipulated statistics of pre-corrected photon traces. An advantage of this pre-correction is an equal weighting of low and high efficiencies and therefore of the efficiency populations and independence from the intensity threshold necessary for burst searching. Pre-correction of intensities and pre-calculation of shot noise limited efficiency histograms (before minimization) allows a fast deconvolution of distance distributions. This enables a fast global analysis of distances and thus nearly arbitrary distance models. Moreover, the corrected efficiencies and the underlying fluctuations can be directly correlated with stoichiometries to distinguish whether the fluctuations arise from structural heterogeneities or from photophysical heterogeneities such as intersystem crossing.

In the work of Antonik et al.[4] and Kalinin et al.[5] uncorrected FRET efficiency histograms were fitted with a probability distribution considering photophysical parameters from separate measurements. In contrast, here, the binned photon traces were directly corrected for quantum efficiencies, relative detection efficiencies, relative extinction coefficients, crosstalk, direct excitation and donor- and acceptor-only populations (see **Supplementary Note 2**). The excessive photons, for example from crosstalk, are then considered as a source of additional shot noise broadening.

Other sources of broadening have been considered, too. Background photons have been neglected as we have a signal to noise ratio of at least 50. Fluctuations of the dipole orientation factor κ^2 or fast distance fluctuations due to translational diffusion of the dyes have been ignored, because they are many orders of magnitudes faster than the photon count rate.
Instead, bleaching and blinking of the dyes occur on the times scale of the photon count rate and can increase the efficiency broadening beyond shot noise. However, bleaching and blinking are reduced to a minimum by fast laser scanning, by using moderate laser intensities and by using stable dyes. Under these conditions the triplet state populations were smaller than 3% for Atto550 and Atto647N as checked for several distances with FCS (see **Supplementary Note 2**). Furthermore, the observed efficiency broadening strongly depends on the conformational state of Hsp90. For these reasons, photophysics can be excluded as a source for the large fluctuations of Hsp90's open state.

The photon counts per channel and time bin are Poisson distributed. If consecutive photons are not correlated, the shot noise limited FRET efficiency distribution is a combination of independent Poisson processes and can be described by the joint probability[6]:

$$P\left(E=\frac{N_A}{N_A+N_D}\middle|\langle E\rangle,n\right)=\frac{(\langle E\rangle n)^{N_A}}{N_A!}\frac{\left((1-\langle E\rangle)n\right)^{N_D}}{N_D!}e^{-n} \qquad (15)$$

Here, $<E>$ is the expected FRET efficiency, $n=<N_A+N_D>=<N_{AD}>$ the expected number of total photons and N_D and N_A are the number of detected donor or acceptor photons, respectively. Considering every combination of N_A and N_D for different n is computationally extensive, especially when the distributions are convoluted with numerous distance distributions. Extending the joint distribution (15) with $N_{AD}!$ leads to the following equation:

$$P\left(E=\frac{N_A}{N_A+N_D}\middle|\langle E\rangle,n\right)=\langle E\rangle^{N_A}(1-\langle E\rangle)^{N_D}\frac{N_{AD}!}{N_A!\,N_D!}\cdot\frac{n^{N_{AD}}}{N_{AD}!}e^{-n} \qquad (16)$$

with a binomial term on the left side and a Poisson term on the right side. Now, with the restriction $n=N_{AD}$, the Poisson term is constant, so that equation (16) results in a binomial distribution. This restriction is not valid for low intensities, because in reality donor and acceptor intensities are independently Poisson distributed. Notably, Monte Carlo simulations show a perfect agreement between two independent Poisson distributed random variables and a binomial distribution when the mean intensity *n* is the same as or larger than the number of bins chosen for the efficiency histogram. The binomial distribution is generated for the mean total intensity *n* and then interpolated to the number of FRET efficiency bins n_{bins} to overcome the discontinuity of the binomial distribution.
On the experimental side, a random number is added to each intensity value before calculating the efficiencies:

$$r=\left[-\frac{n_{bins}}{2n};\frac{n_{bins}}{2n}\right] \qquad (17)$$

That way, discretization artifacts (i.e. binning artifacts) due to rational numbers are overcome and an optimal bin number as supposed by [7] is not necessary any more. Also in the Monte Carlo simulations, random numbers *r* are added to the Poison distributed intensities, in order to be most closely to the experiment. They resulted in perfect agreement of the simulated efficiency histograms with the binomial distributions.
Now, additional acceptor photons due to spectral crosstalk *c* and direct excitation *d* of the acceptor from the green laser are considered. The expectation value of the sum of independent Poisson random variables equals the Poisson distribution of the sum of the expectation values: $n_A^*=n_A+n_c+n_d$. Thus, the corrected FRET sensitized acceptor intensities can be calculated with $N_A=N_A^*-n_c-n_d$. However, after subtraction of absolute values from Poisson distributed intensities, the mean and the variance have not the expected relation anymore. This discrepancy was overcome by deriving corrected probability density distributions. In short, expected corrected efficiencies $\langle E\rangle$ are transferred to uncorrected ones $\langle E^*\rangle$ followed by a coordinate transformation from E^* to E.

The uncorrected FRET sensitized acceptor intensity:

$$n_A^*=n(\langle E\rangle+d+(1-\langle E\rangle)c/\gamma) \qquad (18)$$

depends on the corrected acceptor intensity $n_A=n\langle E\rangle$, the correction factor $\gamma=Q_A\eta_A/(Q_D\eta_D)$ for different quantum efficiencies *Q* and detection efficiencies η, the correction factor *c* for cross talk and the correction factor *d* for direct excitation of the acceptor by the donor laser. With the donor intensity $n_D^*=n(1-\langle E\rangle)/\gamma$ and $\langle E^*\rangle=n_A^*/(n_A^*+n_D^*)$ one obtains the expected uncorrected efficiency:

$$\langle E^* \rangle = \frac{n(\langle E \rangle(1 - c/\gamma) + d + c/\gamma)}{n(\langle E \rangle(1 - c/\gamma - 1/\gamma) + d + c/\gamma + 1/\gamma)} \qquad (19)$$

and the expected uncorrected total intensity:

$$n^* = n(\langle E \rangle(1 - c/\gamma - 1/\gamma) + d + c/\gamma + 1/\gamma). \qquad (20)$$

This leads to the probability density distribution of uncorrected shot noise limited FRET efficiencies, which considers the variances of the original uncorrected photon traces:

$$P_{E^*}(N_A^*|n^*) = \langle E^* \rangle^{N_A^*}(1 - \langle E^* \rangle)^{n^* - N_A^*} \frac{n^*!}{N_A^*!\,(n^* - N_A^*)!} \qquad (21)$$

With the control variable N_A^* the distribution is directly transformed to the corrected efficiencies:

$$E(N_A^*) = \frac{N_A^* - n(d + (1 - \langle E \rangle)c/\gamma)}{n} \qquad (22)$$

to finally obtain the corrected shot noise limited FRET efficiency probability density distribution:

$$P_{E(N_A^*)} = P_{E^*}(N_A^*) \qquad (23)$$

To obtain the convolution between distance fluctuations and shot noise $P_E(R,n)$, the normalized corrected shot noise limited FRET efficiency distributions P_E are numerically superimposed with the distance probability distribution $P\big(\langle \mathrm{E} \rangle(\mathrm{R}_{DA})\big)$:

$$P_E(R,n) = \sum_{k=1}^{n} P\left(\int_{\langle E \rangle=(k-1)/n}^{k/n} \langle E \rangle(R_{DA}) d\langle E \rangle\right) P_E(E|n, \langle E \rangle = \frac{2k-1}{2n}) \qquad (24)$$

The integration over all distance probabilities per FRET efficiency bin $(k-1)/n < \langle E \rangle < k/n$ is especially important for low and high efficiencies:

$$\int_{\langle E \rangle=(k-1)/n}^{k/n} \langle E \rangle(R_{DA}) d\langle E \rangle = \frac{1}{2}\left[1 + erf\left(\frac{(1/\langle E \rangle - 1)^{1/6} R_0 - R_{DA}}{\sigma\sqrt{2}}\right)\right]_{(k-1)/n}^{k/n} \qquad (25)$$

Here, R_0 is the Förster radius, R_{DA} the mean distance between both dyes and σ_{DA} the standard deviation of the Gaussian distribution. The probability density distribution for n bins is interpolated to the number of bins used in the experiment n_{bin}.

A gradient-descent algorithm[8] with stochastically varied starting parameters is used to find the global minimum of the χ^2 error between experimental and theoretical efficiencies. For a faster

minimization procedure, the shot noise limited distributions P_E are calculated for each bin $\langle E \rangle(k)$ before minimization. The only free parameters are the ones searched for: mean distance R_{DAi}, distance distribution σ_{DAi} and the relative fraction for i states. In this way it is possible to find multiple distance populations (i.e. states) for one FRET pair. This approach is especially useful for low efficiency distributions as required for the large open structures of Hsp90.

Distance distributions of Hsp90

The experimental efficiency distribution obtained from an inter-monomer FRET pair of the dynamic open state of Hsp90 (**Fig. N5.1**, bars) is fitted with (24) and then verified with Monte Carlo simulations. The assumption of Gaussian distributed distances describes the experiment surprisingly well. When assuming quasi-rigid states instead, the distributions cannot be described with one or two states. Three or more states have to be assumed, but then they are weighted by a Gaussian distribution in the distance space, again.

The same histogram was fitted with the global distance distribution analysis derived herein (**Online Methods**) assuming four states (**Fig. N5.2** left, dashed red curves). The global analysis of all FRET pairs reveals the underlying Gaussian distance distributions (**Fig. N5.2** right, dashed red curves). Hereby, the efficiency histograms of all FRET pairs are simultaneously fitted while generating structure models. In this way, there is only one most likely solution for the given number of states. A standard non-global distance distribution analysis for multiple nearby structures would result in ambiguous solutions.

The efficiency distribution for the closed state of Hsp90 are nearly shot noise limited and are perfectly described with single Gaussian distance distributions. In this case, the average standard deviation is 4 Å. The efficiencies of the open state are much broader distributed. Therefore, three model structures were generated for the open state with the global distance distribution analysis. The best solution revealed symmetrically distributed distances. That is why single Gaussian distance distributions were extracted from all measured efficiency histograms.
The distance distributions were calculated for FRET pairs between numerous positions of chain A and B of yeast Hsp90 and the results are listed in **Supplementary Table 1**. The table indicates inter-monomer distances (A-B or B-A) or intra-monomer distances (A-A or B-B). Mean distances $\langle R \rangle$ and standard deviations σ_R of Gaussian distance distributions were extracted from the measured efficiency histograms with the distance distribution analysis. The deviation between the measured mean distance R and respective model distance R_M takes into account the accessible volumes of the dyes. The x-ray crystal structure (pdb: 2cg9) and the average domain arrangement of the open state served as model structures. The combined anisotropy r_C was used as a selection criterion and was calculated to be 0.22 for the closed structure and 0.21 for the open structure (**Supplementary Note 4**). The distances that were excluded are marked with an A for high anisotropy, L for relative low ATPase activity (**Supplementary Methods**), D for a position-specific deviation of the anisotropy (**Supplementary Note 3**) or G for a deviating global gamma-factor (**Supplementary Note 3**).

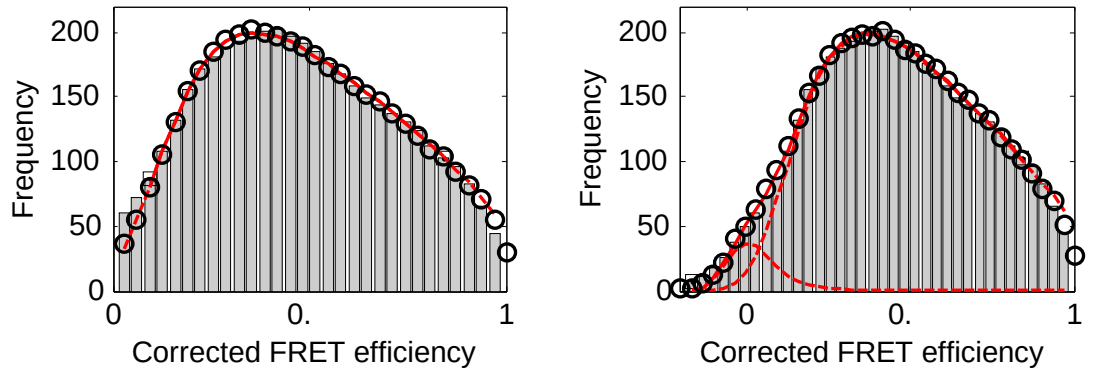

Figure N5.1 | Corrected FRET efficiency histogram of the dynamic open state of Hsp90 for a FRET pair between the middle domains and a photon binning time of 1 ms. The histogram was fitted with the probability distribution analysis derived herein (red line) and verified with a Monte Carlo Simulation (black circles). The underlying model was either a single Gaussian distance distribution (left) or two Gaussian distance distributions (right).

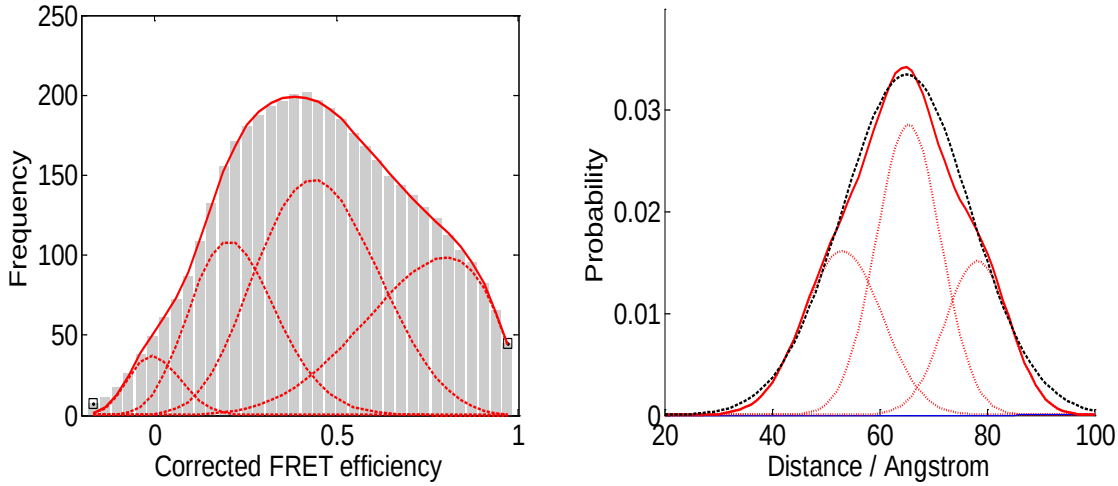

Figure N5.2 | Corrected FRET efficiency histogram of the dynamic open state of Hsp90 for a FRET pair between the middle domains and a photon binning time of 1 ms. The histogram was fitted with the global distance distribution analysis derived herein assuming four states (left, dashed red curves). The global analysis of all FRET pairs reveals the underlying Gaussian distance distributions (right, dashed red curves). For this example, the sum of these distributions (right, solid red curve) is similar to the single Gaussian distance distribution extracted from the same efficiency histogram (right, dashed black curve) with a standard distance distribution analysis.

Supplementary Note 6: Determination of dynamics at different timescales

The multi-domain protein in solution can be seen as an ensemble of conformations with individual dwell times. Conformational changes happen on a broad range of timescales. There are slow conformational changes of global domain arrangements (here: open or closed state) and there are fast conformational changes of small elements and inter-domain fluctuations.
As FRET occurs within nanoseconds it is especially suited to describe the dynamics of the conformational ensemble of a protein. A limitation is signal-to-noise ratio or maximal observation time. Therefore different methods must be applied. The rates of each method can be determined independently as long as they differ at least by an order of magnitude.

Slow conformational changes of freely diffusing FRET labeled molecules can be monitored by disequilibrating the ensemble, e.g. by the addition of nucleotides. In contrast to ensemble fluorescence microscopy, with a confocal microscope changes of small populations can be tracked, because snapshots of single molecules are recorded. Here the number of events linearly increases with time, therefore only dwell times of more than 10 seconds can be observed. The expectation time of yeast Hsp90 to reach the closed state was (8±4) minutes with AMPPNP and (1.5±1.0) minutes with ATP.

Faster dynamics within a certain global state can be gained from photon bursts of single freely diffusing molecules, whereat at least 1000 single molecules are usually needed. The actual state of each single molecule when passing through the confocal laser volumes can be identified structurally via efficiency, stoichiometry and anisotropy. In case of Hsp90 each single molecule was assigned to the open or closed state.
Burst variance analysis (BVA) indicates dynamics faster than the diffusion time[9]. This method is insensitive to small subpopulations, with the advantage that linker dynamics or background photons are negligible. On the other hand, it is sensitive to small differences in efficiencies as required for the analysis of small fluctuations. The closed state of Hsp90 was found to be quasi-static at that timescale and the open state revealed dynamics faster than 4 ms.
Based on the idea of the burst variance analysis, we developed a more detailed analysis that not only reveals the timescale of the fluctuations, but also the distance amplitudes. First, we separated structural heterogeneities from shot noise with a probability distribution analysis (see **Supplementary Note 5**) and described the underlying distance fluctuations with different models whereat Gaussian distributions resulted in optimal fits. Then, we revealed the timescale of the fluctuations by varying the time binning of the photon traces between 0.1 and 10 ms and plotting the standard deviation of the Gaussian distance distributions versus observation time. The closed state of Hsp90 was found to be quasi-static, while the fluctuations in the open state strongly increased with a decreasing time binning. Efficiency histograms from nearly every FRET pair are well described with single Gaussian distance distributions on all observed timescales. Additionally, the fluctuations vary with observation time in a manner suggesting multiple energetically equivalent conformations within the open state. On the one hand, most of the FRET efficiency histograms of the open state cannot be fitted with a single (shot-noise limited) state. On the other hand, there are no substates visible (no separated FRET efficiency peaks) and the distance deviations do continuously increase with decreasing binning times even at the smallest binning times.

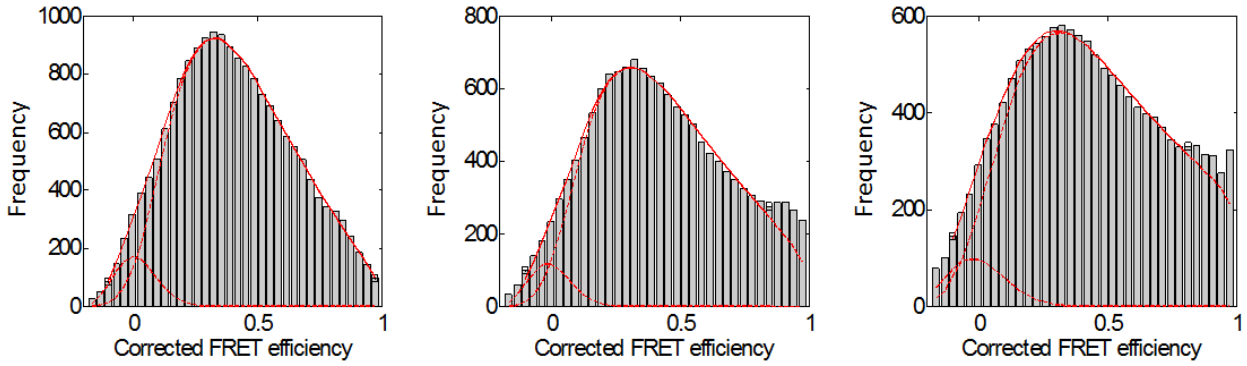

Figure N6.1 | Distance distribution analysis of a FRET pair in the open ADP-bound state of Hps90 for different binning times: 2000 µs (left), 600 µs (middle) and 200 µs (right) resulting in Gaussian distance distributions with standard deviations of 10.5 Å (left), 12 Å (middle) and 14 Å (right).

Fluorescence lifetime analysis reflects dynamics occurring between 10 ns and 1 ms[10,11]. This method is especially sensitive for small dynamic populations when the difference in FRET efficiency is high. Thus it can be seen as complementary to the burst variance analysis and to our binning variation analysis. In the case of Hsp90, dynamics faster than 1 ms were observed for nearly each position in the open state. Fast dynamics in the 100 µs range were observed for intra-monomer distances between N- and M-domain.

From the above-mentioned slow and fast dynamics, Gibbs free energies were estimated by the Arrhenius equation:

$$k = A \cdot e^{-\frac{\Delta G^{\ddagger}}{R \cdot T}} \qquad (26)$$

with the transition rate k from one conformation to the next and the Arrhenius prefactor A, which can be estimated at 10^8 s^{-1} for proteins[12]. The observed fluctuations of Hsp90's open state were much slower than expected for free diffusion and would equal a free energy of about 9 k_BT in a two-state Arrhenius model.

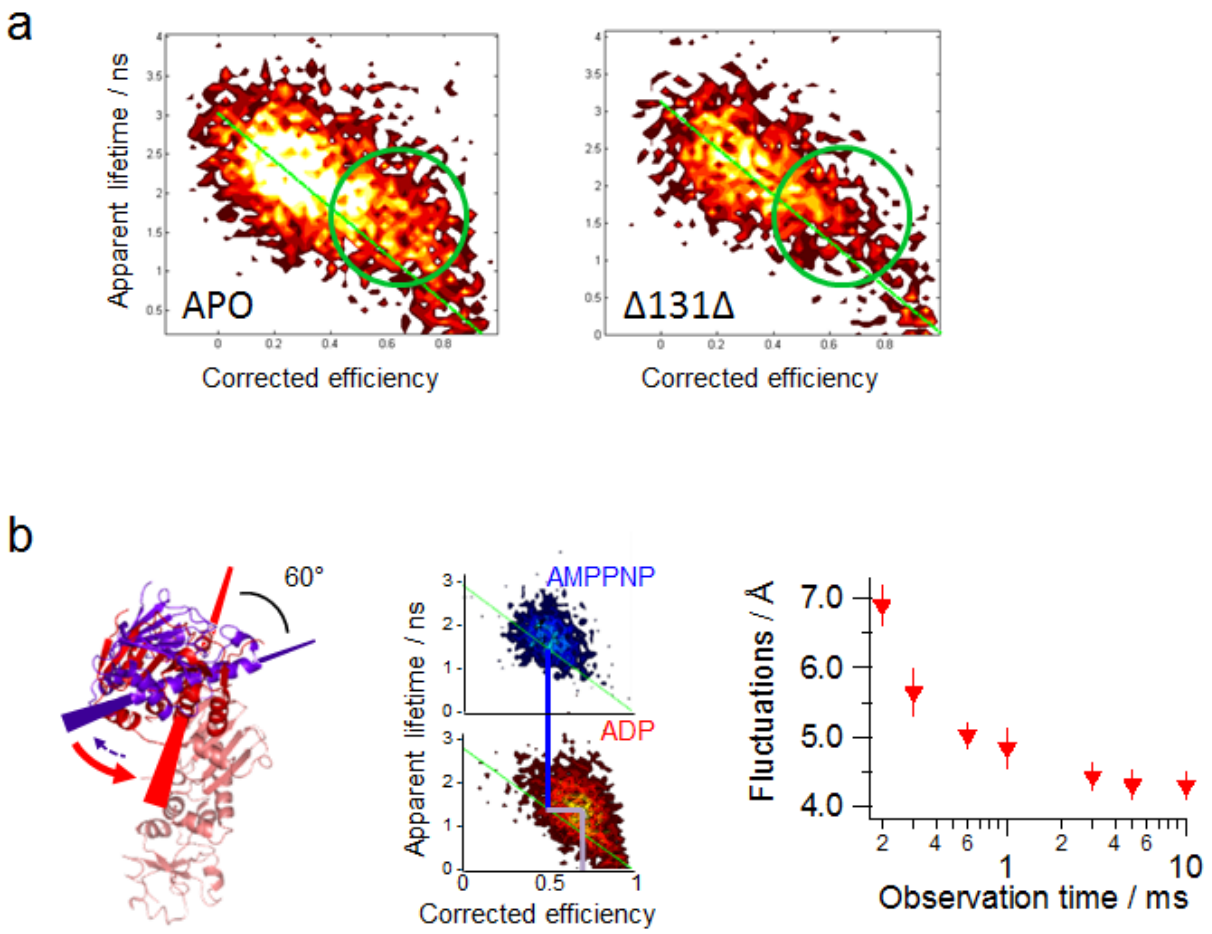

Figure N6.2 | Local and global dynamics.
(a) The fast inter-monomer dynamics of the Hsp90 dimer are suppressed in the presence of the model client Δ131Δ. Correlation of burst wise fluorescence lifetimes with efficiencies reveals a dynamic population (deviating from the static line), which is not present in the presence of the client protein. This is consistent with the results from the time-correlated distance distribution analysis presented herein. The lifetime analysis[10,11] can be seen as a complementary analysis, but is not suited for quantifying the timescale of the dynamics and for its unambiguous assignment to the global state. **(b)** Analysis of a dynamic intermediate state of the N-domain (purple). Fluorescence lifetime vs. FRET efficiency plots for a FRET pair spanning between the N- and M-domain within one monomer reveal a dynamic ADP-bound open state. The projection of the dynamic population of the ADP-bound open state (left, red) to the static line (left, green) results in a FRET efficiency that is similar to the one of the AMPPNP-bound closed conformation (left, blue). This suggests a similar, but transiently populated conformation. Variation of the photon binning times reveals kinetics in the range of 100 µs between the average open conformation and a small populated intermediate state. The inter-domain distances for the dynamic intermediate state are determined for several FRET pairs resulting in a conformation with the N-domain being rotated relatively to the M-domain similar to the conformation within the closed dimer state.

References and Notes:

1 Lee, N. K. *et al.* Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys J* **88**, 2939-2953 (2005).

2 Dale, R. E., Eisinger, J. & Blumberg, W. E. The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer. *Biophys J* **26**, 161-193 (1979).

3 Ivanov, V., Li, M. & Mizuuchi, K. Impact of emission anisotropy on fluorescence spectroscopy and FRET distance measurements. *Biophysical journal* **97**, 922-929 (2009).

4 Antonik, M., Felekyan, S., Gaiduk, A. & Seidel, C. A. Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. *The Journal of Physical Chemistry B* **110**, 6970-6978 (2006).

5 Kalinin, S., Felekyan, S., Valeri, A. & Seidel, C. A. Characterizing multiple molecular states in single-molecule multiparameter fluorescence detection by probability distribution analysis. *The Journal of Physical Chemistry B* **112**, 8361-8374 (2008).

6 Gopich, I. & Szabo, A. Theory of photon statistics in single-molecule Forster resonance energy transfer. *The Journal of chemical physics* **122**, 14707 (2005).

7 Nir, E. *et al.* Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. *J Phys Chem B* **110**, 22103-22124 (2006).

8 Arfken, G. The method of steepest descents. *Mathematical methods for physicists* **3**, 428-436 (1985).

9 Torella, J. P., Holden, S. J., Santoso, Y., Hohlbein, J. & Kapanidis, A. N. Identifying molecular dynamics in single-molecule FRET experiments with burst variance analysis. *Biophys J* **100**, 1568-1577 (2011).

10 Kalinin, S., Valeri, A., Antonik, M., Felekyan, S. & Seidel, C. A. 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).

11 Gopich, I. V. & Szabo, A. Theory of the energy transfer efficiency and fluorescence lifetime distribution in single-molecule FRET. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 7747-7752 (2012).

12 Lapidus, L. J., Eaton, W. A. & Hofrichter, J. Measuring the rate of intramolecular contact formation in polypeptides. *Proc Natl Acad Sci U S A* **97** (2010).