

This protocol was added to the paper after the formal review process.

1. Title: Determination of structural ensembles and correlated dynamics of a multi-domain protein

2. Authors: P. Wortmann¹, F. Kandzia², M. Zacharias², T. Hugel^{1*}, B. Hellenkamp^{1*},

¹Institute of Physical Chemistry, University of Freiburg, Germany

²Physics Department, Technische Universität München, Germany

*Corresponding author: b.hellenkamp@tum.de; thorsten.hugel@pc.uni-freiburg.de

3. Abstract

This protocol describes how structural ensembles and state-dependent local dynamics of a multi-domain protein can be determined. The hybrid approach combines static structures of sub-domains with self-consistent FRET networks derived from of time-resolved fluorescence parameters and distance distributions. The method has successfully been applied on yeast heat shock protein 90, Hsp90. It reproduced Hsp90's closed conformation as determined earlier by x-ray crystallography and resolved the previously unknown dynamic open conformation.

4. Introduction

Proteins often undergo conformational changes on many timescales. In multi-domain proteins global conformational changes normally occur on slow timescales of milliseconds to seconds, while local conformational changes normally occur on timescales from microseconds to milliseconds. The following procedure uniquely covers these timescales. It is demonstrated on the molecular chaperone Hsp90.

5. Reagents

- a. Piranha: 3:1 (vol:vol) H₂SO₄ (97 %):H₂O₂ (30 %)
- b. RCA 1:1:5 (vol:vol:vol) H₂O₂ (30 %):NH₄OH (33 %):ddH₂O
- c. 2 % Hellamex (Hellma) solution with ddH₂O
- d. 100 mM NaHCO₃ in ddH₂O at pH 8.5
- e. Pure acetone
- f. Vectrabond (Axxora)
- g. 5 kDA Polyethylenglycol (PEG, Rapp Polymere)
- h. Nescofilm (Nesco)
- i. LB-Medium
- j. Appropriate antibiotics (Kanamycin)
- k. Induction reagent (Isopropyl β-D-1-thiogalactopyranoside, IPTG)
- l. Gene in expression plasmid (*yhsp90* in pET28a)
- m. Buffer and reagents for protein biochemistry
 - Phosphate buffered saline (PBS)
 - Purification buffer for chromatography:
 - Strep W (100 mM Tris/HCl, 150 mM NaCl, 1 mM EDTA, pH 8.0)
 - Strep EI (100 mM Tris/HCl, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin, pH 8.0)

IEX A (40 mM Hepes, 50 mM NaCl, 1 mM DTT, pH 7.5)

IEX B (40 mM Hepes, 1 M NaCl, 1 mM DTT, pH 7.5)

SEX (40 mM Hepes, 200 mM KCl, pH 7.4)

Measurement buffer (40 mM Hepes, 150 mM KCl, 10 mM MgCl₂, pH 7.4)

Tris(2-carboxyethyl)phosphine (TCEP)

Dithiothreitol (DTT)

- n. Fluorescent dyes in form of reactive derivate as FRET donor and FRET acceptor
 - Atto532-Maleimide, AttoTec
 - Atto550-Maleimide, AttoTec
 - Atto647N-Maleimide, AttoTec
 - Alexa647-Maleimide, Thermo Fisher Scientific

6. Prerequisites

- a. This method requires basic structural input. Either x-ray crystal structures or cryoEM structures from (sub-) domains that shows secondary structure elements must be known.
- b. In case the protein populates more than one state, it is very helpful to shift the state populations, e.g. by a (pseudo-) substrate. Otherwise the states have to be assigned during global domain arrangement and therefore many more distances are necessary.
- c. Protein labeling by cysteine-maleimide chemistry is the simplest site-specific labeling method. It requires the protein to contain no native cysteines. If the protein contains cysteines, a mutation towards structurally similar residues (serine, alanine) will be the best option, unless the cysteines are essential. In that case, a more comprehensive strategy for site-specific labeling by unnatural amino acids would be necessary.

7. Equipment

- a. Cloning
 - i. Site directed mutagenesis Kit (QuikChange Lightning Kit, Stratagene) and appropriate mutagenesis primers according to manufacturer's instructions
 - ii. DNA purification kit (QIAprep Spin Miniprep Kit, Qiagen)
 - iii. Competent bacteria for plasmid preparation (*e.coli* XL10-Gold, Agilent)
 - iv. 37 °C incubator
- b. Protein Expression
 - i. Competent bacteria for protein expression (*e.coli* BL21DE3, Thermo Fisher Scientific)
 - ii. Incubator shaker (37°C for recombinant expression in *e.coli*)
 - iii. Centrifuge to harvest the bacteria, for ~1 l sample volumes and > 3000 g (Hettich Rotanta 460R, Hettich)
 - iv. Cell disrupter (Constant Systems)
 - v. Centrifuge for removal of cell debris (Avanti JXN-26, Beckman Coulter)
- c. Protein Biochemistry
 - i. SDS-PAGE gels with appropriate instrumentation for monitoring of the protein purification
 - ii. Äkta (GE Healthcare) or a similar automated chromatography system with a sample pump
 - iii. Purification columns: An affinity column (Strep-Tactin Superflo, IBA), an Ion exchange column (HiTrap Q, GE Healthcare), a Size Exclusion column (HiLoad 16/60 Superdex200, GE Healthcare)
 - iv. Desalting columns (PD10, GE Healthcare)

- v. Microcentrifuge with temperature control (5424R, Eppendorf) and Vivaspin500 centrifugal concentrators (Sartorius Stedim Biotech)
 - vi. Block heater (ThermoMixer C, Eppendorf) to heat samples to 47°C
 - d. Spectrometer
 - i. Fluorescence spectrometer to measure fluorescence spectra (Fluoromax-4, Horiba) and a quartz cuvette (105.250-QS, Hellma analytics)
 - ii. UV/VIS spectrometer to measure absorbance spectra and ATPase activity by a time course measurement in case of Hsp90 (Lambda 35, PerkinElmer), with temperature control unit and a quartz cuvette (105.202-QS, Hellma analytics)
 - iii. A confocal setup for single molecule fluorescence detection (see Figure 1): The setup has to be color- and polarization-sensitive and detected photons must be time-correlated to the laser pulse with picosecond resolution. The setup used in our study is described in the following:
 A linearly polarized green diode laser (532 nm LDH-P-FA-530, PicoQuant) and a linearly polarized red laser (640 nm LDH-D-C-640, PicoQuant) are spatially aligned and focused with an Apo 60x water immersion objective (Nikon). This enables separate excitation of FRET donor and FRET acceptor in order to derive photophysical correction factors from the measured intensities. The lasers are pulsed interleaved with repetition rates of 20 MHz enabling detection of fast conformational relaxation rates. Fluorescence photons from the FRET donor and acceptor are spectrally and polarization-sensitively separated by dichroic mirrors and polarization beam splitters, respectively. They are detected with single photon avalanche photodiodes (PDM-50µm and τ-SPAD, PicoQuant) at picosecond resolution utilizing a commercial data acquisition system (HydraHarp 400, PicoQuant). The detected single photons are time-correlated to the laser pulse with picosecond resolution (micro-time). In a separate time channel, the photons are time-correlated to the start of the measurement (macro-time) to obtain photon bursts from single diffusing molecules. By combing micro-time data (fluorescence lifetimes) and macro-time data (efficiencies and stoichiometries), sub-ensemble fluorescence lifetimes and sub-ensemble time-resolved anisotropies can be calculated. The laser intensities were adjusted such that on average 50 photon counts per millisecond are obtained. With a threshold of 100 counts per millisecond a compromise was found between a sufficient signal to noise ratio of 10 (for distinguishing several sample specific efficiencies) and a low probability of simultaneous excitation of donor and acceptor with the green laser.
 - e. For a surface functionalized measurement chamber
 - i. Precision cover slips (170µm, Carl Roth)
 - ii. Teflon block (1 mm thick, width of the cover slip)
 - iii. Ultrasonic bath
 - iv. Heat plate
8. Software
- a. The FPS tool (available at <http://mpc.hhu.de/software/fps.html>) (2) is mainly used to generate the dye accessible volumes and calculate the mean dye positions.
 - i.
 - b. With the following Matlab tools (available at <http://www.single-molecule.uni-freiburg.de>) (see associated publication) self-consistent networks of distance distributions and state-

dependent structural ensembles are generated (see the Readme file for detailed information):

- i. Generate distance distributions from listed efficiency histograms with “Out_Dist = MDAGENDistances (In_NrOfStates, In_MinWidth, In_MaxWidth, In_ParameterTable, In_PositionTable, Out_DistanceTable)”
- ii. Generate global minimum structures and local minima structures with “Out_Minima = MDAGENStructures (In_PositionTable, In_DistanceTable, Out_NewPositionTable, Out_PyMolStructures, In_MaxChi)”
- iii. Generate structures from random distance subsets and compare them with each other with “Out_Subsets = MDAGENSubsets (In_PositionTable, In_DistanceTable, Out_PyMolStructures, In_NrOfReducedDistances, In_NrOfSubsets)”
- iv. Generate sigma structures with “Out_Sigma = MDAGENSigmaStructures (In_PositionTable, In_DistanceTable, Out_PyMolStructures)”

9. Procedure

a. Generate stable homodimers

It is essential to ensure a dimeric protein will not dissociate under the experimental conditions ($c < 200$ pM). This can be achieved by e.g. artificial insertion of a zipper motif (by mutagenesis) close to the native dimerization site. The zipper motif used in our study is made of the amino acid sequence PG-EQKAEWKRRYEKEKEKN ARLKKGKVEKLEIELARWR-PG and derived from the kinesin heavy chain of *D. melanogaster*, dmKHC (1). It did not show strong effects on the proteins conformational dynamics and activity in our hands.

b. Create Mutant Library

- i. Select possible labeling positions in each domain based on conservation among homologues (low conservation preferred). Based on known secondary structure, choose positions within small loops that are evenly distributed among the domains and preferably not located at domain interfaces.
- ii. Create point mutants by site-directed mutagenesis following the manufacturer's protocol, purify plasmid DNA of mutants and verify mutants by sequencing.
- iii. Transform plasmid with verified single point mutant gene into bacteria for expression. Express protein mutants and purify them by chromatography to > 90 % purity, as checked by SDS-PAGE. Final protein concentration should be ~ 50 μ M.

c. Sample preparation: Dye labelling

- i. Label cysteine mutants with maleimide dyes. Reduce the cysteine by addition of 10 mM (TCEP) and incubation at room temperature for 30 minutes. Remove TCEP and change buffer to PBS pH 6.7 with a desalting column. The concentration of the protein during labeling should not drop below 30 μ M. In case, the concentration is lower, concentrate the protein with vivaspin500 centricons. Add a two-fold excess of maleimide dye to the protein and incubate in the dark for one hour at room temperature. Remove dye excess and change buffer to measurement buffer with a desalting column.
In case of multimeric protein complexes, label one monomer with one dye at a certain position, and exchange one donor with one acceptor labeled mutant to obtain a certain

amount of FRET species that have exactly one acceptor dye at one monomer and one donor dye at another monomer. In case of Hsp90, the exchange was done by incubating the mixed sample for about 20 minutes at 47°C in order to accelerate the opening of the coiled coil zipper. Then, remove possible aggregates by extensive centrifugation at 14000 g for one hour. Correlate measured FRET efficiencies with stoichiometries in a scatter plot, to sort out homo dimers.

In contrast, intra-monomer mutants (two labels within one monomer) are exchanged with a 200-fold excess of wild type monomers to exclude inter-monomer FRET. In this case, equimolar dye concentrations are added at the same time. In our study, FRET between these positions did not significantly depend on which dye reacted to which site, indicated by dye swapping in the heterodimer.

d. Sample preparation: Protein activity

If enzymatic activity of the protein is known, test the purified and labeled mutants for their activity compared to the wild-type protein. Any labeled mutant that loses its activity completely has to be excluded immediately from further experiments. In case a heterodimeric protein is studied and FRET distributions of mixed homodimers will be measured, check their activity as well (see Online Methods and Supplementary Figure 5 in the associated publication).

e. Production of a measurement chamber

i. First step, cleaning of glass slides (170µm, Carl Roth):

1. Put slides in 2% Hellmanex in an ultrasonic bath for 10 min.
2. Put slides in ddH₂O in ultrasonic bath two times for 10 min.
3. Clean slides for 2 h at 55°C in piranha solution.
4. Clean slides with ddH₂O.
5. Put slides in 2% Hellmanex in ultrasonic bath two times for 10 min.
6. Clean slides with ddH₂O.
7. Put slides in ddH₂O in ultrasonic bath two times for 10 min.
8. Clean slides for 2 h at 55°C in RCA solution.
9. Clean slides with ddH₂O.
10. Put slides in 2% Hellmanex in ultrasonic bath.
11. Put slides in ddH₂O in ultrasonic bath for 10 min.
12. Storage in ddH₂O.

ii. Second step, functionalization of cleaned glass slides:

1. Solve 80 mg PEG in 600 µL 100 mM NaHCO₃ at pH 8.5; shake it and keep it cool.
2. Dry slides with compressed pure nitrogen and put them in pure acetone.
3. Incubate slides 5-10 min in 200 ml acetone with 1 ml Vectabond.
4. Swivel slides in ddH₂O and place each one on a petri dish.
5. Bring 70 µL of PEG solution on each slide and cover it with an untreated slide.
6. Keep everything dark, cold (4°C) and humid with water-soaked wipes.
7. After 3 hours: Swivel slides in ddH₂O to wash off unbound PEG.
8. Dry slides with nitrogen; storage them dark, cold, ideally in a vacuum chamber.

- iii. Third step, chamber construction:
1. Cut 5mm-holes into a Nescofilm with a laser cutter (3 rows, 8 columns).
 2. Drill holes into a 1 mm-thick teflon block.
 3. Bring the Nescofilm between a functionalized slide and the teflon block.
 4. Burden the chamber on a heat plate (70°C) for 20 minutes with the slide side down.
 5. Finally, use each hole for 1-2 measurement days.
 6. Reuse the Teflon block after cleaning with 2% Hellmanex and ddH₂O.
- f. Data Acquisition
- i. Determine the normalized fluorescence spectrum $\bar{F}_D(\lambda)$ for each donor labeled mutant and absorbance spectrum for each acceptor labeled mutant, all in measurement buffer.
 - ii. Record micro- and macro-time data with the described confocal setup to obtain FRET efficiencies, stoichiometries and time-resolved anisotropies. For each FRET pair and condition (nucleotide, co-factor, salt concentration...) record at least 1000 single molecule events (at least 30 minutes per measurement at a sample concentration of about 50 pM). Measure each pair under the conditions that shift the equilibrium among the state populations. Start with measuring 10-16 FRET pairs between two large domains (each with 4 labeling positions) followed by the following steps. Then, measure further FRET pairs and add further labeling positions if necessary.
- g. Data Evaluation: Preliminary distance determination
- i. Calculate the quantum yield for each donor position from the measured fluorescence lifetime (micro-time data) in the donor-only population τ_{D0} and the given fluorescence lifetime τ'_{D0} and quantum yield Q'_D specified by the manufacturer:

$$Q_D = Q'_D \cdot \tau_{D0} / \tau'_{D0}$$
 - ii. Calculate the overlap integral $J(\lambda)$ for each FRET pair:

$$J_{DA}(\lambda) = \int \bar{F}_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda$$

Assume the extinction coefficient in the absorbance spectra maximum to remain constant and calculate the extinction spectrum $\varepsilon_A(\lambda)$.
 - iii. Calculate the Förster radius R_0 for each FRET pair from the determined overlap spectra J and donor quantum yield Q_D . Assume $\kappa^2 = 2/3$ (see section 'Verification' for checking if this assumption holds).
 - iv. Correct FRET efficiencies for leakage, direct excitation, β - and γ -factors iteratively by checking for self-consistency within the network (see Supplementary Note 2 and 3 in the associated publication).
 - v. Calculate preliminary distances from corrected efficiencies with a probability distribution analysis (see Supplementary Note 5 in the associated publication).
- h. Data Evaluation: State Assignment
- Compare the data to existing structural data: In case a structure of a functional state exists, screen it with the FPS tool and compare the experimental distances to the model for verification and state assignment. With no existing global structure, states will later be assigned during the global domain arrangement.

i. Data Evaluation: Verification

- i. Check the fraction of dark states for donor and acceptor dye with fluorescence correlation spectroscopy. Use the equation of (2) to fit the autocorrelation of the donor fluorescence intensity after donor excitation and to fit the autocorrelation of the acceptor fluorescence intensity after acceptor excitation (macro-time data). Choose moderate excitation powers (about one third of the power at fluorescence intensity saturation) to guarantee low dark state fractions. For the described setup, an average laser power of 50 μ W was chosen and resulted in dark state fractions of less than 3% for Atto550 and Atto647N.
- ii. Calculate the time-resolved dye anisotropies from parallel and perpendicularly polarized fluorescence intensities $N_{||}$ and N_{\perp} (micro-time data) after excitation with the linearly polarized donor laser or acceptor laser, respectively (see Figure 2a):

$$r = \frac{N_{||} - GN_{\perp}}{N_{||} + 2GN_{\perp}}$$

The G factor corrects for relative detection efficiencies. For objectives with a high numerical aperture depolarization factors have to be considered ($I_1=0.0308$ and $I_2=0.0368$ for NA=1.2). Determine the combined anisotropy, if possible for each state separately:

$$r_C = \sqrt{r_{A,\infty}} \sqrt{r_{D,\infty}}$$

- iii. Determine a threshold for the combined anisotropy with distance network and apply it to the data set (see Figure 2b). A reasonable threshold value found in our study using Atto550 and Atto647N as FRET pair is $r_C = 0.22$. For further information see Supplementary Note 4 in the associated publication.
- iv. Test for state-dependent local conformational changes. First, compare the position-specific anisotropies between the different states. Therefore, correlate the time-resolved anisotropies with the respective accessible volumes (calculated with the FPS software (2)) of acceptor or donor dye for each labeling position and for each state and check for changes between the states. Such changes hint local rearrangements (see Supplementary Note 3 in the associated publication). Second, use intra-domain and intra-monomer FRET pairs to test for rearrangements within.

j. Data Evaluation: Domain Arrangement

- i. For each state, arrange the domains with the assigned distances. Use the FPS software (2) for a first average domain arrangement. Then, use this software to calculate mean dye positions from accessible volumes. Then, run the software in the associated publication.
- ii. Analyze the fluorescence intensity traces (macro-time data) with different time bins and generate one efficiency histograms for each time interval. Then, run the software tool 'MDAgenDistances' to generate distances, their standard deviations of a Gaussian distribution and their uncertainties.
- iii. Run the software tool 'MDAgenStructures' to generate average structures with different χ^2 values.
- iv. Generate structures from distance subsets with the software tool 'MDAgenSubsets'.

- v. In case of ambiguous solutions, substitute domain structures with other crystal structures if known. In case the solution remains ambiguous, go back to step (a) or (d) and measure further FRET pairs.
 - vi. Generate an ensemble of structures with the software tool 'MDAgenEnsembles'.
- k. Data Evaluation: Molecular refinement
- i. Refine the ensemble structures from the domain arrangement with MD simulations. Screen the results with the FPS software (3). In case deviations occur, repeat the global distance distribution analysis and subsequent domain arrangement.
 - ii. Test the stability and fast local dynamics of the optimal solution with unrestrained MD simulations in explicit solvent.

10. References

1. Mickler, M., Hessling, M., Ratzke, C., Buchner, J. & Hugel, T. The large conformational changes of Hsp90 are only weakly coupled to ATP hydrolysis. *Nature structural & molecular biology* **16**, 281–286 (2009)
2. J. Widengren, R. Rigler, U. Mets, Triplet-state monitoring by fluorescence correlation spectroscopy. *J Fluoresc* **4**, 255-258 (1994)
3. Kalinin, S. *et al.* A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nature methods* **9**, 1218–1225 (2012)

11. Associated Publication

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

B. Hellenkamp, P. Wortmann, F. Kandzia, M. Zacharias, T. Hugel; *Nature Methods* (2016)

12. Competing financial interests

The authors declare no competing financial interests.

13. Figures

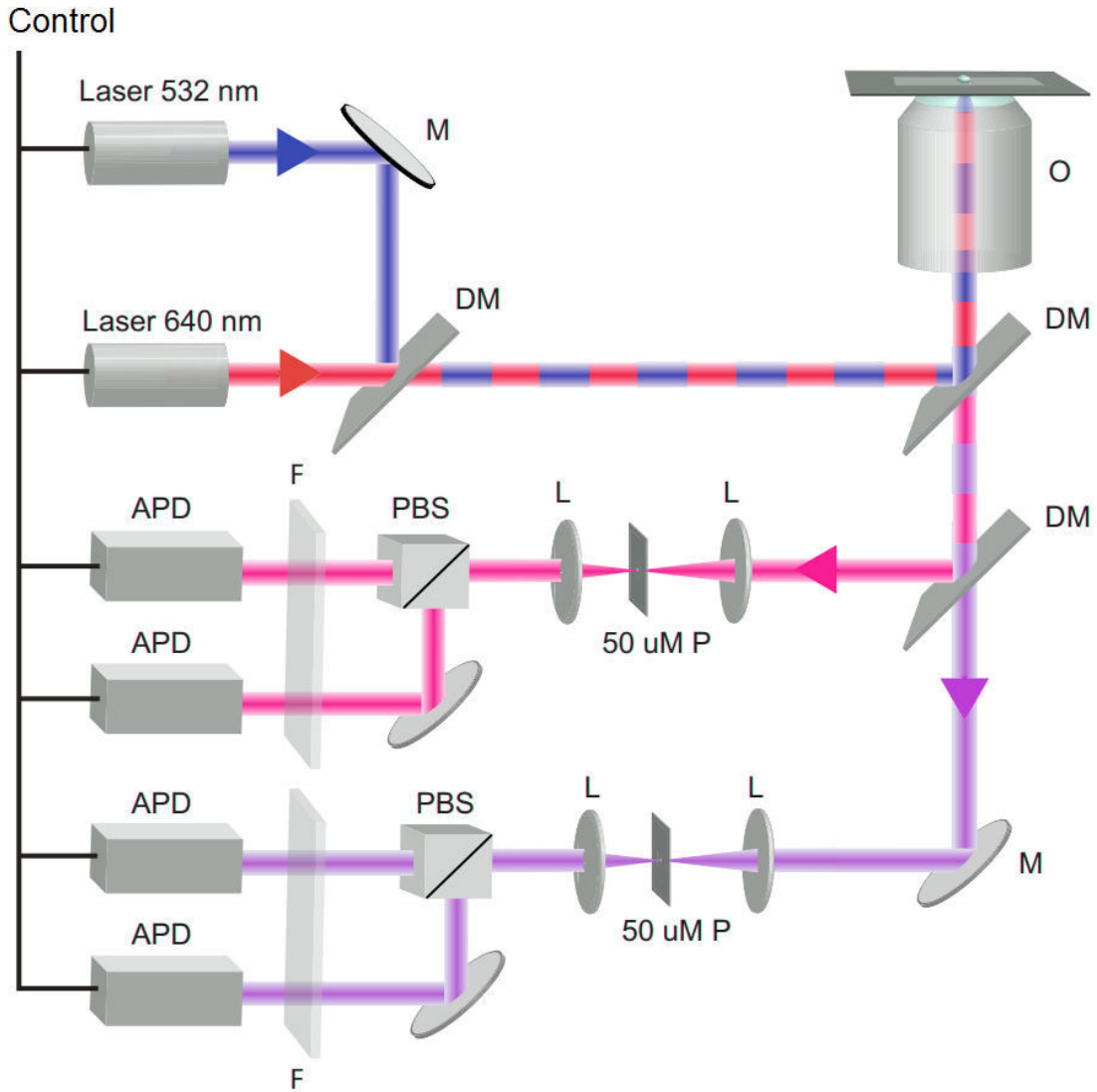


Figure 1: A confocal setup for single molecule detection with pulsed interleaved excitation (PIE). M: mirror; DM: dichroic mirror; O: objective; L: lens; P: pinhole; PBS: polarization beam splitter; F: band pass filter; APD: avalanche photo diode detector.

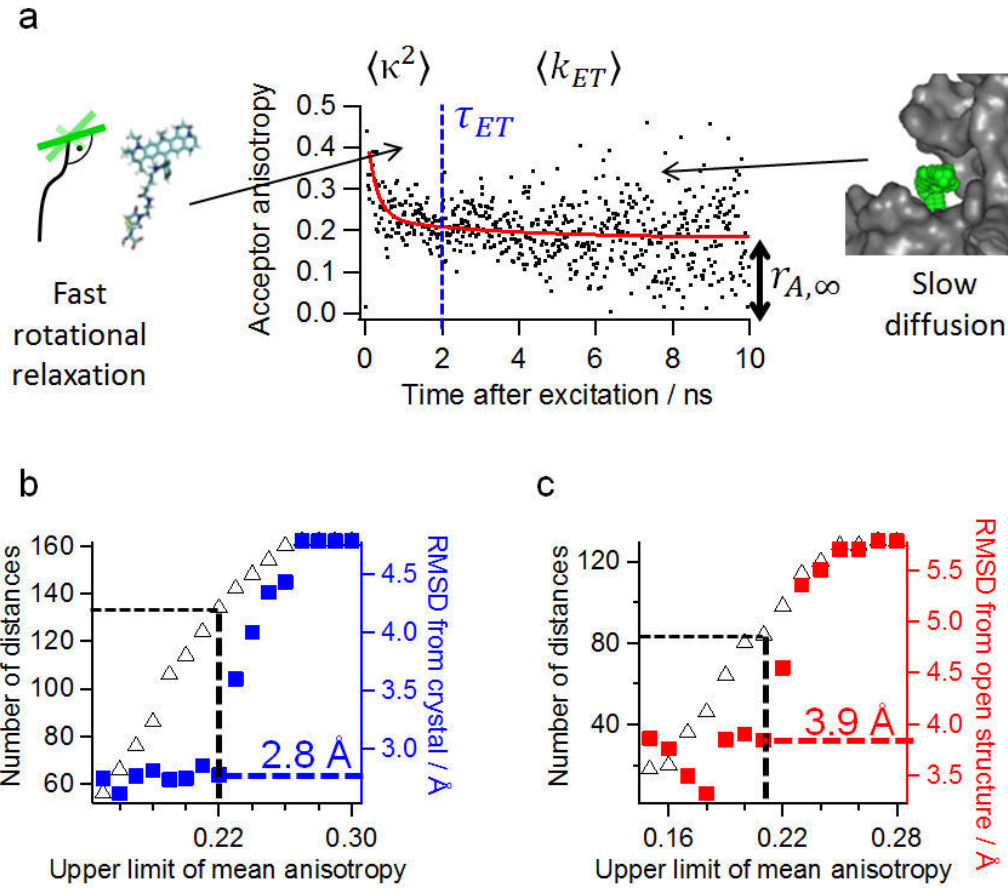


Figure 2: (a) The time-resolved anisotropies of dyes bound to a larger protein normally consist of a fast decay from rotational relaxation of the dipole axis and of a slow decay from dye diffusion along the protein surface. The residual anisotropies r_{∞} of both dyes of a FRET pair is used as part of the selection criterion (see associated publication). (b and c) A threshold value for nearly isotropy dipole coupling is determined for the closed (b) and open (c) state of Hsp90 (see associated publication).

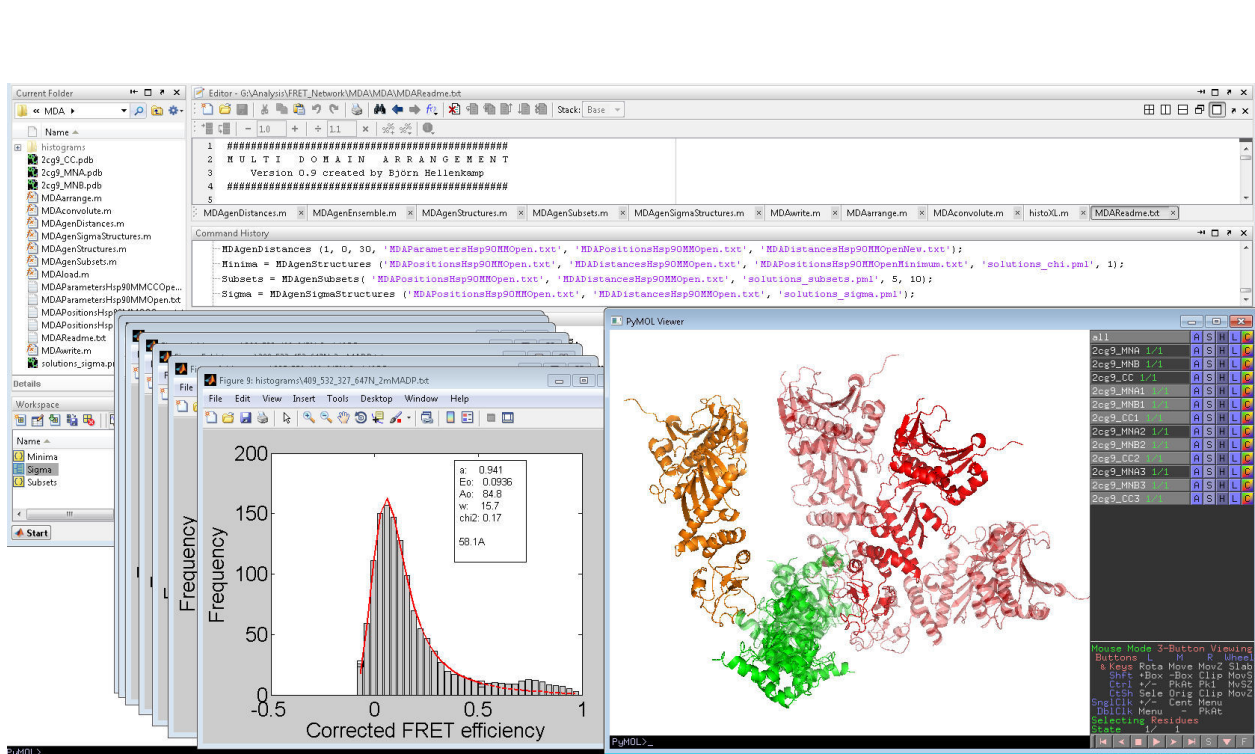


Figure 3: Software tool for generating self-consistent distance distribution networks and dynamic domain arrangements (see associated publication).