Electronic Supplementary Material (ESI) for Molecular BioSystems

Intramolecular Three-Colour Single Pair FRET of Intrinsically Disordered Proteins with Increased Dynamic Range

Sigrid Milles [a], Christine Koehler [a], Yann Gambin,[b] Ashok A Deniz[c*] and Edward A Lemke [a*]

Supplementary Methods

Triple labelling

A 1:1 mixture of Alexa594 and Alexa488 was given at fivefold excess to freshly reduced triple yNup49FG mutant and allowed to react for 2h at room temperature. The ketone group of yNup49FG harboring *p*-acetylphenylalanine was labeled under denaturing conditions (4 M GdmCl) at pH 4 in sodiumacetate buffer with a hydroxylamine derivative of Alexa647 as previously reported.[5, 9, 27]

Three-color single molecule fluorescence spectroscopy

For smFRET experiments, the triply labeled proteins were diluted to a final concentration of 50–200 pM in XM (X=0, 1, 2, 3 or 4) GdmCl in PBS, pH 7.4 with 2 mM Magnesium acetate and 2 mM dithiothreitol and measured on a custom built spectrometer centred around an Olympus IX81 microscope (Hamburg, Germany) equipped with a high numerical water immersion objective (60x, 1.2NA). A green laser diode LDH 485, a red laser diode LDH 660 (Picoquant, Berlin, Germany) and a white light laser (SuperK Extreme, NKT Photonics) filtered through a 572/15 excitation filter (Semrock, New York, NY), were pulsed sequentially at 80 MHz total. Using quarter wave plates the light was circularly polarized and used to excite freely diffusing labelled proteins. Fluorescence originating from single molecules was first spatially filtered by a 100 µm pinhole and then split into green donor (D, detected on an MPD, Picoquant), orange (A1, tauSPAD Picoquant) and red acceptor (A2, Perkin Elmer, Ma, USA) fluorescence emission light.

The data were recorded in a time correlated single photon counting format using a HydraHarp 400 (Picoquant) and analyzed with a custom-written code in IgorPro (Wavemetrics, Lake Oswego, OR). The time trace was binned to 1 ms bins and photons resulting upon green and orange excitation were subjected to a threshold of 50.

Labeling stoichiometries were calculated according to

$$S_{GR} = \frac{I_{A2} + I_D}{I_D + I_{A2} + I_{A2}^{dir} + I_{A1}}, S_{OR} = \frac{I_{A1} + I_D}{I_D + I_{A1} + I_{A1}^{dir} + I_{A2}}.$$

For identification and selection of the triple labelled population, S_{OR} was plotted against S_{GR} in a two-dimensional histogram and the middle population was selected with $0.3<S_{OR/GR}<0.7$. The corresponding bins were then evaluated for Forster resonance energy transfer (FRET) and FRET efficiencies were calculated via

$$E_{GO} = \frac{I_{A1} - le_{GO} \cdot I_D}{\gamma_{GO} \cdot I_D + I_{A1}}, E_{GR} = \frac{I_{A2} - le_{GR} \cdot I_D}{\gamma_{GR} \cdot I_D + I_{A2}}, E_{OR} = \frac{I_{A2} - le_{OR} \cdot I_{A2}}{\gamma_{OR} \cdot I_{A1} + I_{A2}}$$

with E_{XX} representing the FRET efficiencies between the different colours. Leakage le_{xx} was determined for every measurement condition (0…4 M GdmCl). γ_{XX} were determined experimentally for each dye pair in PBS and 4 M GdmCl, and approximated to remain constant across the entire GdmCl concentrations used, a valid assumption since no difference in γ_{XX} could be observed between measurements in PBS and 4 M GdmCl.

For identification of the two populations within E_{GO} and E_{OR} corresponding to the long and the short labelling distance, we plotted E_{GO} against E_{OR} in a two-dimensional (2D) histogram. In these histograms, the two populations could be clearly identified and were fit with a 2D double Gaussian according to

$$f(x,y) = A_1 \cdot \exp\left(\frac{1}{2\cdot(1-cor_1^2)} \cdot \left(\left(\frac{x-x_1}{xw_1}\right)^2 + \left(\frac{y-y_1}{yw_1}\right)^2 - \frac{2\cdot cor_1 \cdot (x-x_1)\cdot(y-y_1)}{xw_1 \cdot yw_1}\right)\right) +$$
$$A_2 \cdot \exp\left(\frac{1}{2\cdot(1-cor_2^2)} \cdot \left(\left(\frac{x-x_2}{xw_2}\right)^2 + \left(\frac{y-y_2}{yw_2}\right)^2 - \frac{2\cdot cor_2 \cdot (x-x_2)\cdot(y-y_2)}{xw_2 \cdot yw_2}\right)\right).$$

A_1 and A_2 are the amplitudes of the two populations, $x_{1/2}$ and $y_{1/2}$ the x and y positions (corresponding to S_{GR} and S_{OR}) of the two populations and $xw_{1/2}$ and $yw_{1/2}$ are the corresponding widths of the fit Gaussians. cor_1 and cor_2 allow ellipticity in the 2D

fit.
Starting values were estimated visually from the 2D-histogram and the fit converged always to a distinct solution.

Forster distance (R_0) for the GR and the OR dye-pair were determined according to

$$R_0 = \sqrt[6]{\frac{9(\ln 10)\kappa^2 J(\lambda)\Phi_{D(0)}}{128\pi^5 n^4 N_A}} \ .$$

It contains the orientation factor $\kappa^2 = 2/3$ between donor emission dipole and acceptor excitation dipole, the overlap integral $J(\lambda)$ between the donor emission spectrum and the acceptor absorbance spectrum, the refractive index n of the buffer, the Avogadro number N_A and the fluorescence quantum yield of the donor in the absence of acceptor. Quantum yields and the overlap integral were measured from singly labelled protein (G-only, O-only and R-only) following standard protocols.