Electronic Supplementary Material (ESI) for Chemical Science.

Electronic Supplementary Information

Stable End-Sealed DNA as Robust Nano-rulers for In Vivo Single-Molecule Fluorescence

Anne Plochowietz,[a†] Afaf H. El-Sagheer,[b,c] Tom Brown[b] and Achillefs N. Kapanidis[a,†]

a A. Plochowietz, Prof. A. N. Kapanidis
Biological Physics Research Group, Department of Physics
University of Oxford
Clarendon Laboratory, Parks Road, Oxford, OX1 3PU, UK.
E-mail: anne.plochowietz@gmail.com or a.kapanidis1@physics.ox.ac.uk

b A. H. El-Sagheer, Prof. T. Brown
Chemistry Research Laboratory, Department of Chemistry
University of Oxford
Mansfield Road, Oxford, OX1 3TA, UK.

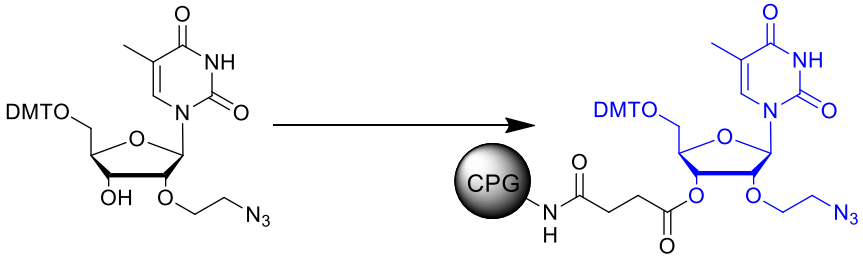
c A. H. El-Sagheer
Chemistry Branch, Department of Chemistry
Faculty of Petroleum and Mining Engineering
Suez University
Suez 43721, Egypt

Electronic Supplementary Information

General method for oligonucleotide synthesis and purification

Standard DNA phosphoramidites, solid supports and additional reagents were purchased from Link Technologies and Applied Biosystems. Oligonucleotides were synthesized on an Applied Biosystems 394 automated DNA/ RNA synthesizer using a standard 1.0 *μ*mole phosphoramidite cycle of acid-catalyzed detritylation, coupling, capping, and iodine oxidation. Stepwise coupling efficiencies and overall yields were determined by the automated trityl cation conductivity monitoring facility and in all cases were >98.0%. All β-cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C, and T monomers was 60 sec, whereas the coupling time for the modified phosphoramidite monomers were coupled for 600 sec. Cleavage of oligonucleotides from the solid support and deprotection was achieved by exposure to concentrated aqueous ammonia solution for 60 min at room temperature followed by heating in a sealed tube for 5 hr at 55 °C. The oligonucleotides were purified by reversed-phase HPLC on a Gilson system using an XBridgeTM BEH300 Prep C18 10 μM 10x250 mm column (Waters) with a gradient of acetonitrile in ammonium acetate (0% to 50% buffer B over 30 min, flow rate 4 mL/min), buffer A: 0.1 M ammonium acetate, pH 7.0, buffer B: 0.1 M ammonium acetate, pH 7.0, with 50% acetonitrile. Elution was monitored by UV absorption at 305 or 295 nm. After HPLC purification, oligonucleotides were desalted using NAP-10 columns (GE Healthcare), analyzed by gel electrophoresis and characterised by mass spectrometry.

Preparation of 5´-*O*-(4,4´-Dimethoxytrityl)- 2´-*O*-(2-azidoethyl)-5-methyluridine on solid support



Solid support Amino SynBase (Link technologies, 1000/110, 59 μmol/g) (250 mg) was activated in 3% trichloroacetic acid (TCA) in dichloromethane for 1 hour in a stoppered

glass vessel fitted with a sinter and tap. After filtration, the support was washed with triethylamine:diisopropylethylamine (9:1), dichloromethane and diethyl ether. The support was dried under vacuum for 1 hour then soaked in dry pyridine (10 mL) for 10 min. A solution of succinic anhydride (112 mg, 1.12 mmol) and 4-dimethylaminopyridine (23 mg, 0.19 mmol) in dry pyridine (3 mL) was added and the vessel was rotated for 20 hours. The support was washed with pyridine, dichloromethane and diethyl ether, dried and soaked in pyridine for 10 min. *N*-(3-Dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride (EDC.HCl) (48 mg, 0.25 mmol), 4-dimethylaminopyridine(2 mg, 0.013 mmol), triethylamine (10 µL) and 5´-*O*-(4,4´-Dimethoxytrityl)- 2´-*O*-(2-azidoethyl)-5-methyluridine[1] (64 mg, 0.100 mmol) were dissolved in pyridine (2 mL) and added to the solid support in the vessel, which was rotated for 20 hours at room temperature. Pentachlorophenol (17 mg, 0.064 mmol) was added and the reaction vessel was left to rotate for 1 h. The solvent was then removed by filtration and the support was washed with pyridine, dichloromethane and diethyl ether. Piperidine (10% in DMF, 5 mL) was added and after rotating for 1 min the solid support was washed with DMF, dichloromethane and diethyl ether. Capping reagent (oligonucleotide synthesis grade, acetic anhydride/pyridine/tetrahydrofuran: N-methyl imidazole in tetrahydrofuran, 1:1, 5 mL) was added and the vessel was rotated for 1 h after which the support was washed with THF, pyridine, dichloromethane and diethyl ether then left to dry under vacuum overnight. The loading of 5´-*O*-(4,4´-dimethoxytrityl)- 2´-*O*-(2-azidoethyl)-5-methyluridine on the support was 30 µmol/g which was determined from the cleaved dimethoxytrityl group.

Labeling of amino modified oligonucleotides with Azido hexanoic or dyes active ester

Oligonucleotides containing amino-C6dT, 3'-amino linker or 5'-amino linker were purified by HPLC as described above in the general method, freeze dried and dissolved in 80 µL of 0.5 M Na_2CO_3/$NaHCO_3$ buffer at pH 8.75. To this solution was added Azido hexanoic, Cy3B or ATTO647N NHS ester (1 mg) in DMSO (80 µL) and the mixture was left at room temperature for 5 h. The labeled oligonucleotides were then gel-filtered through a NAP-25 column and purified by reversed-phase HPLC as explained above in the general method.

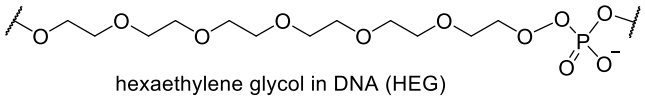
Synthesis and Purification of end-sealed duplexes

To a solution of tris-hydroxypropyl triazole ligand[2] (7.0 µmol in 70 µL 200 mM NaCl) under argon was added sodium ascorbate (10.0 µmol in 20.0 µL 200 mM NaCl) followed by $CuSO_4.5H_2O$ (1.0 µmol in 10 µL 200 mM NaCl). The oligonucleotides mixture (S1+S2, S1+S3 or S4+S5) (10.0 nmol each in 900 µL 200 mM NaCl) were heated at 85 °C for 5 min, cooled down slowly and added to the above Cu^I solution. The reaction mixture was kept under argon at room temperature for 2 h and a disposable NAP-25 gel-filtration column was used to remove reagents (GE Healthcare). End sealed duplexes were analysed and purified using 10% polyacrylamide gel electrophoresis. Protected DNA standards were stored in ddH_2O at -20°C.

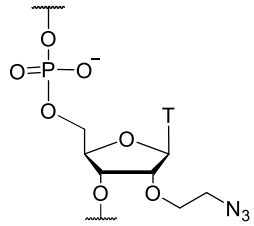
Table S1. Oligonucleotide sequences used in this study

Code	5´ - 3´ Oligonucleotide sequence	Calc.	Found
S1	**K**-**Y**AAATCTAAAGTAACATAAGGTAACATAACGTAAGCTCATTCGCG-**H**-**Tz**	**15444**	**15443**
S2	**K**-CGCGAATGAGCTTACGTTATGTTACCT**X**ATGTTACTTTAGATTTA-**H**-**Tz**	**15500**	**15497**
S3	**K**-CGCGAATGAGCTTACGTTATGTTACCTTATGTTACTT**X**AGATTTA-**H**-**Tz**	**15500**	**15497**
S4	**K**-**Y**AAATCTAAAG**Y**AACATAAGG**Y**AACATAACG**Y**AAGCTCATTCGCG-**K_1**	**17134**	**17135**
S5	**Z_1**CGCGAATGAGCTTACGTTATGTTACCTTATGTTACTTTAGATTTA-**Z_2**	**14485**	**14487**

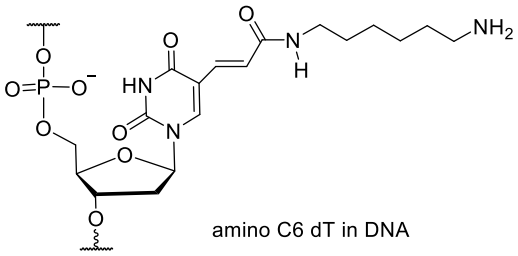
K = 5´-Hexynol. K_1 = 3´-alkyne-modifier serinol. Y = amino-C6dT labelled with Cy3B. X = amino-C6dT labelled with ATTO647N. Tz = 2´-azido ethoxy dT. Z_1 = 5'-amino linker labelled with azido hexanoic NHS ester. Z_2 = 3'-amino linker labelled with azido hexanoic NHS ester. H = hexaethylene glycol linker (HEG). Mass spectra were recorded on a Bruker micrOTOF™ II focus ESI-TOF MS instrument in ES^- mode. HEG, aminoC6dT and alkyn structures are below.



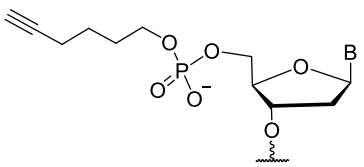
hexaethylene glycol in DNA (HEG)



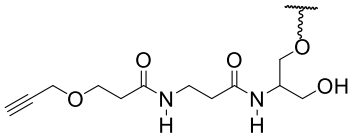
2'-azidoethyl T in DNA



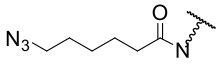
amino C6 dT in DNA



hexynol attached to 5'-nucleotide



3'-Alkyn-modifier serinol in DNA



Azido hexanoic acide part
conected to 3'-or 5'- amino linker
in DNA

In vitro single-molecule FRET analysis

Information on photon arrival times, excitation cycle, and detection channel, enables the calculation of fluorescence bursts (donor or acceptor) arising due to molecules traversing the confocal spot under either donor or acceptor excitation[3]. FRET efficiency (E*, uncorrected), and relative probe stoichiometry (S) were calculated as follows:

$$E^*_{in\ vitro} = \frac{F^{Aem}_{Dex}}{F^{Dem}_{Dex} + F^{Aem}_{Dex}}$$

$$S^*_{in\ vitro} = \frac{F^{Dem}_{Dex} + F^{Aem}_{Dex}}{F^{Dem}_{Dex} + F^{Aem}_{Dex} + F^{Aem}_{Aex}}$$

where F^{Yem}_{Xex} represents the fluorescence of a burst arising in the *Y* emission channel under *X* excitation.

In vitro fluorescence correlation spectroscopy (FCS) analysis

The autocorrelation of FRET bursts was calculated using a computer-implemented correlator (Flex 02-01D, Correlator.com) and autocorrelation curves were fitted using PyCorrFit, open-source software[4] to an autocorrelation function $G(\tau)$ describing a single diffusive species (3D Brownian motion) and one triplet state of the fluorophore:

$$G(\tau) = A_0 + \frac{1}{N}\frac{1}{1 + {\tau}/{\tau_D}}\frac{1}{\sqrt{1 + {\tau}/{(E^2\tau_D)}}}\left(1 + \frac{Te^{-\tau/\tau_T}}{1-T}\right)$$

with A_0: offset, N: effective number of molecules in confocal volume, τ_D: characteristic diffusion time through confocal volume, E: elongation factor of the confocal volume, T: fraction of molecules in triplet state, τ_T: characteristic triplet state lifetime.

In vivo single-molecule FRET and single-particle tracking analysis

Single-molecule FRET analysis was performed using custom-written MATLAB software. First, single-molecule localization analysis was performed as described in Refs.[3, 5]; PSFs in DD, DA and AA-channel in each movie frame were fitted by a 2D elliptical Gaussian (free fitting parameters: x/y position, x/y width, elliptical rotation angle, amplitude, background) using initial position guesses from applying a fixed localization-intensity

threshold on the bandpass filtered fluorescence image[6]. Upon localization of the single molecule in the DA-channel the fitting routine was performed in the DD channel, which was mapped onto the FRET channel using a transformation matrix, and the FRET efficiencies were calculated as described below. For measurements using an ALEX scheme, PSFs were also fitted in the AA channel. Single-molecule localizations were filtered for PSF-shape (width of the 2D Gaussian ranging from 50-400nm), a total photon count of (DD+DA)-intensity>400ph/frame, the existence of an Acceptor molecule, i.e. AA-intensity>130ph/frame, and autofluorescence in the Donor-channel, DD>40ph/frame (**Figure S5**). Such background corrected single-molecule FRET and stoichiometry values were calculated from photon counts (phC, integral of fitted 2D elliptical Gaussians) for each filtered single molecule in the respective emission channel:

$$E^*_{in\ vivo} = \frac{phC^{Aem}_{Dex}}{phC^{Dem}_{Dex} + phC^{Aem}_{Dex}}$$

$$S^*_{in\ vivo} = \frac{phC^{Dem}_{Dex} + phC^{Aem}_{Dex}}{phC^{Dem}_{Dex} + phC^{Aem}_{Dex} + phC^{Aem}_{Aex}}$$

where phC^{Yem}_{Xex} denotes the photon counts of a single molecule in the *Y* emission channel under *X* excitation.

Single-particle tracking was performed in the red emission channel by adapting the MATLAB script based on a published algorithm Ref.[7]. Localized PSFs in the DA-channel (gCW), and DA and AA-channel (ALEX) were linked to a track if they appeared in consecutive frames within a window of 7 pixels (0.69µm). This window size ensures that 98% of steps are correctly linked for an apparent diffusion coefficient of 1.0µm^2/s and 20ms exposure time[8]. To account for PSF disappearance due to blinking or missed localization, a memory parameter of 1 frame was used. To eliminate noise, only molecules appearing in 5 consecutive frames were included in the analysis.

Accurate single-molecule FRET measurements *in vitro* and *in vivo*

In vitro and *in vivo* single-molecule FRET values were initially obtained as described above and for generality we describe *in vitro* and *in vivo* background corrected FRET (*E**) and Stoichiometry (*S**) values:

$$E^* = \frac{DA}{DA + DD}$$

$$S^* = \frac{DD + DA}{DD + DA + AA}$$

with *DA,* photon counts in the acceptor channel after donor excitation, *DD,* photon counts in the donor channel after donor excitation, and *AA,* photon counts in the acceptor channel after acceptor excitation. To obtain accurate FRET and Stoichiometry values, the E* and S* values were corrected for cross-talk contribution of donor and acceptor dyes to the FRET signal.

$$FRET = DA - Lk - Dir = DA - l \cdot DD - d \cdot AA$$

with *Lk* is the leakage contribution to the *DA* signal, and *Dir* is the direct-excitation contribution to the *DA* signal[9, 10]. Donor-only species (present due to incomplete labeling or acceptor photobleaching) give direct access to the leakage contribution, and acceptor-only species (present due to incomplete labeling or donor photobleaching) give direct access to the direct excitation contribution[9, 10]. Thus, we obtained cross-talk corrected FRET (E_{cc}) and Stoichiometry (S_{cc}) values

$$E_{cc} = \frac{FRET}{FRET + DD} \quad (1)$$

$$S_{cc} = \frac{DD + FRET}{DD + FRET + AA} \quad (2)$$

Finally, the correction factor γ was determined for *in vitro* and *in vivo* studies, which represents the quantum yields of donor and acceptor dye, and the detection efficiencies of the optical setups. To determine γ, we used the peak position of S_{cc} and E_{cc} for the protected intermediate and high FRET standards and plotted $1/S_{cc}$ versus E_{cc} (**Figure S6**). We obtained the intercept (b) and slope (m) from a linear fit and calculated γ = (b-1)/(b+m-1), Ref.[10]. For the *in vitro* confocal single-molecule FRET spectroscopy studies, we obtained γ~0.71, and for the *in vivo* single-molecule FRET TIRF microscopy studies, we obtained γ~0.75, see **Figure S6**. The FRET and Stoichiometry values were then corrected for the γ-factor

$$E_\gamma = \frac{FRET}{FRET + \gamma \cdot DD} \quad (1)$$

$$S_\gamma = \frac{\gamma \cdot DD + FRET}{\gamma \cdot DD + FRET + AA} \quad (2)$$

The correction of *in vitro* and *in vivo* single-molecule FRET signals is shown in **Figure S5** and **Figures S8**, respectively. The different FRET and Stoichiometry values uncorrected and γ-corrected are shown in **Table S2** below and *in vitro* and *in vivo* single-molecule FRET values for both DNA standards are in very good agreement.

	Protected intermediate FRET standard, P18		Protected high FRET standards, P8	
	In vitro	*In vivo*	*In vitro*	*In vivo*
Uncorrected FRET (E*)	0.43±0.07	0.42±0.12	0.86±0.06	0.85±0.06
Uncorrected Stoichiometry (S*)	0.68±0.05	0.63±0.12	0.63±0.06	0.58±0.12
γ-corrected FRET (E_γ)	0.42±0.10	0.40±0.18	0.89±0.06	0.87±0.05
γ-corrected Stoichiometry (S_γ)	0.60±0.06	0.54±0.12	0.59±0.06	0.55±0.11

Table S2: FRET and Stoichiometry values uncorrected and γ-corrected for intermediate and high FRET DNA standards P18 and P8 obtained *in vitro* and *in vivo*.

HMM fitting of single-molecule fluorescence time-traces

Single-molecule fluorescence time-traces of protected 4xCy3B-DNA were fitted with a hidden Markov model (HMM) described in[3] to obtain the number of Cy3B labels per DNA molecule. HMM is a stochastic model that maps measured values to unobserved (or hidden) states. Here, the single-molecule fluorescence time trajectories (4xCy3B protected DNA data) were modeled as a sequence of up to 5 hidden states (different single-molecule intensity levels corresponding to number of Cy3B fluorophores) and transitions between these states (photobleaching events). A custom-written MATLAB script was run recursively; keeping only the state values for the last photobleaching step and removing this data before the next iteration step. Each iteration allowed up to 5 hidden states to be fitted. Such a method took advantage of the exponential

photobleaching kinetics: the last photobleaching step was likely to last significantly longer than previous steps, and thus fitting the last step was likely to have the greatest data support.

References

1. J. A. Richardson, M. Gerowska, M. Shelbourne, D. French and T. Brown, *Chembiochem*, 2010, **11**, 2530-2533.
2. T. R. Chan, R. Hilgraf, K. B. Sharpless and V. V. Fokin, *Org. Lett.*, 2004, **6**, 2853-2855.
3. R. Crawford, J. P. Torella, L. Aigrain, A. Plochowietz, K. Gryte, S. Uphoff and A. N. Kapanidis, *Biophys. J.*, 2013, **105**, 2439-2450.
4. P. Muller, P. Schwille and T. Weidemann, *Bioinformatics*, 2014, **30**, 2532-2533.
5. S. Uphoff, R. Reyes-Lamothe, F. Garza de Leon, D. J. Sherratt and A. N. Kapanidis, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 8063-8068.
6. S. J. Holden, S. Uphoff, J. Hohlbein, D. Yadin, L. Le Reste, O. J. Britton and A. N. Kapanidis, *Biophys. J.*, 2010, **99**, 3102-3111.
7. J. C. Crocker and D. G. Grier, *J. Colloid Interf. Sci.*, 1996, **179**, 298-310.
8. X. Michalet and A. J. Berglund, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, 2012, **85**, 061916.
9. N. K. Lee, A. N. Kapanidis, Y. Wang, X. Michalet, J. Mukhopadhyay, R. H. Ebright and S. Weiss, *Biophys. J.*, 2005, **88**, 2939-2953.
10. J. Hohlbein, T. D. Craggs and T. Cordes, *Chem. Soc. Rev.*, 2014, **43**, 1156-1171.
11. M. Tokunaga, N. Imamoto and K. Sakata-Sogawa, *Nat. Methods*, 2008, **5**, 159-161.

Supporting Figures

Figure S1

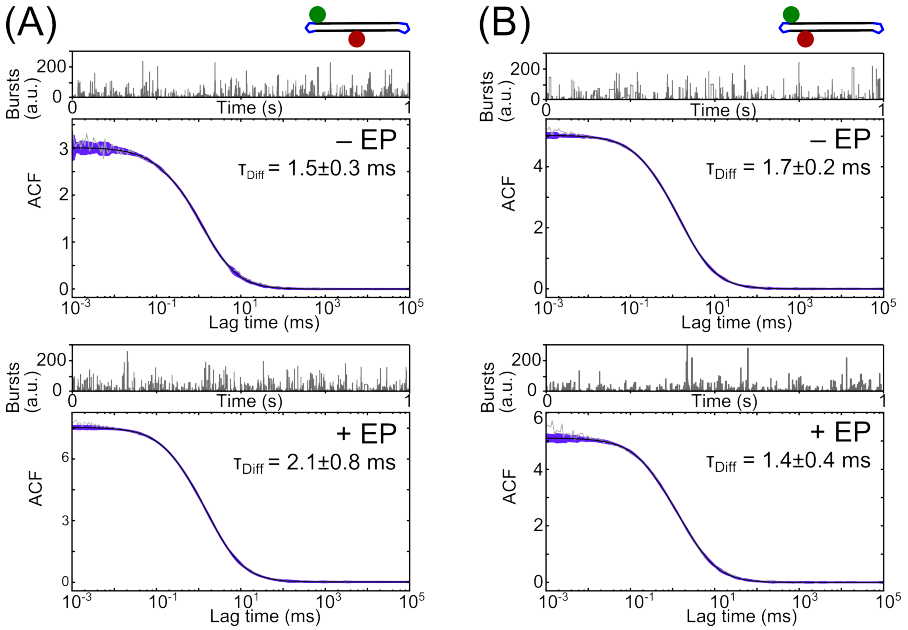


Figure S1. *In vitro* characterization of protected DNA A) intermediate- and B) high-FRET standards using confocal single-molecule fluorescence microscopy. Single-molecule fluorescence time-traces (DA-channel) do not show long dwell bursts (>50 ms, compare to blunt-ended DNAs, **Figure S2**, and the auto-correlation functions (ACF) exhibit a similar characteristic diffusion time (τ_{Diff}, ± standard deviation) before (-EP) and after electroporation (+EP).

Figure S2

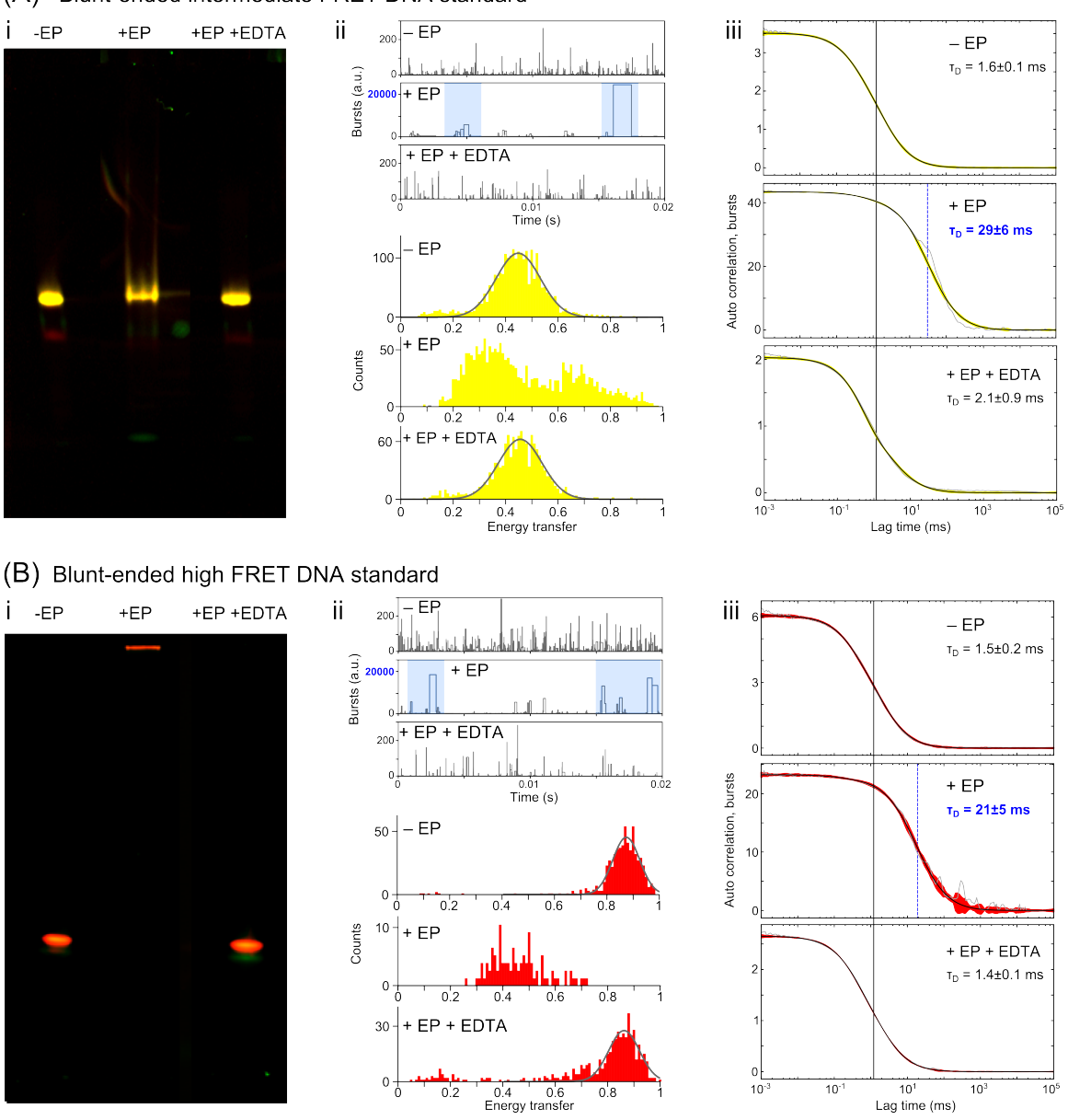


Figure S2. Aggregation of blunt-ended DNA FRET standards after electroporation and addition of EDTA to circumvent aggregation**.** A) Intermediate and B) high FRET DNA standards, non-electroporated (–EP), electroporated (+EP, 1.4kV), and electroporated in presence of 1mM EDTA (+EP+EDTA, 1.4kV) samples were analyzed *in vitro* using biochemical and single-molecule fluorescence assays. i) Colored native gels of +EP sample show smeary band or DNA stuck in the well in comparison to –EP sample and +EP+EDTA sample. ii) Burst time traces (gray) and intermediate and high FRET histograms from smFRET confocal data show long dwell times for +EP sample and heterogeneous smFRET species, respectively, which were not present in the –EP sample and +EP+EDTA sample. iii) Auto-correlation analysis of FRET bursts show more than 10-fold larger molecule complexes crossing the confocal volume for the +EP sample and very similar decay half-time for –EP and +EP+EDTA samples.

Figure S3

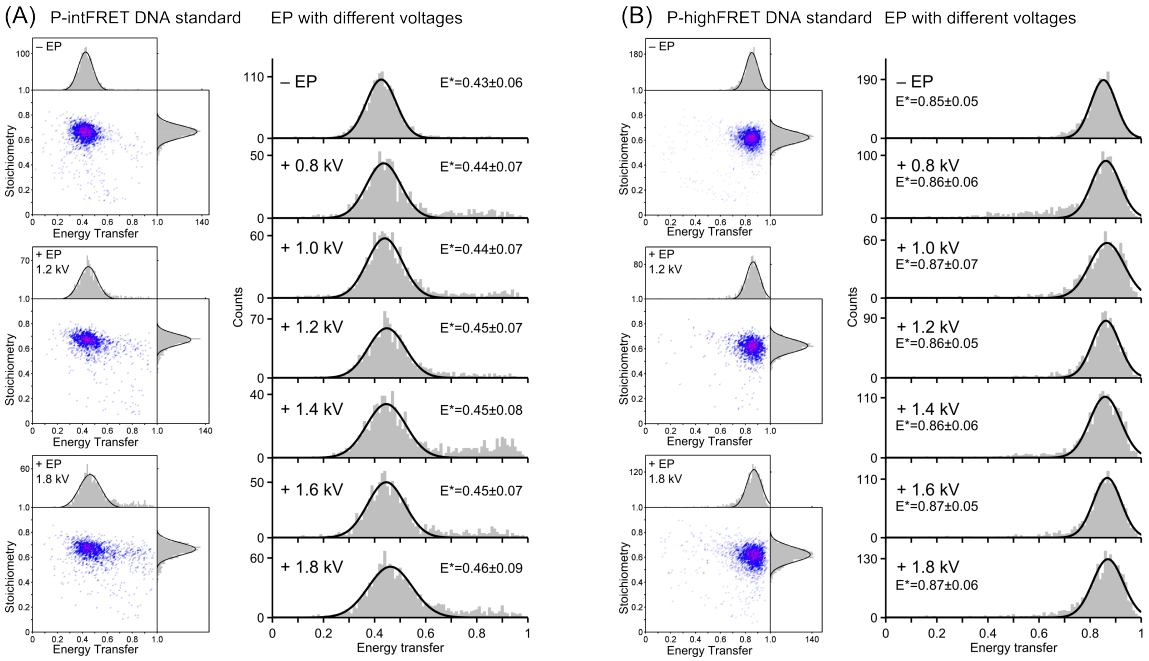


Figure S3. Protected DNA FRET standards withstand electroporation. FRET histograms for A) intermediate and B) high FRET protected DNA standards do not show a significant change in peak shape and position between non-electroporated sample (– EP) and electroporated sample at different electroporation voltages (0.8-1.8 kV). Example FRET-Stoichiometry histograms for non-electroporated sample, and samples electroporated at 1.2 kV and 1.8 kV show very similar FRET-Stoichiometry distributions.

Figure S4

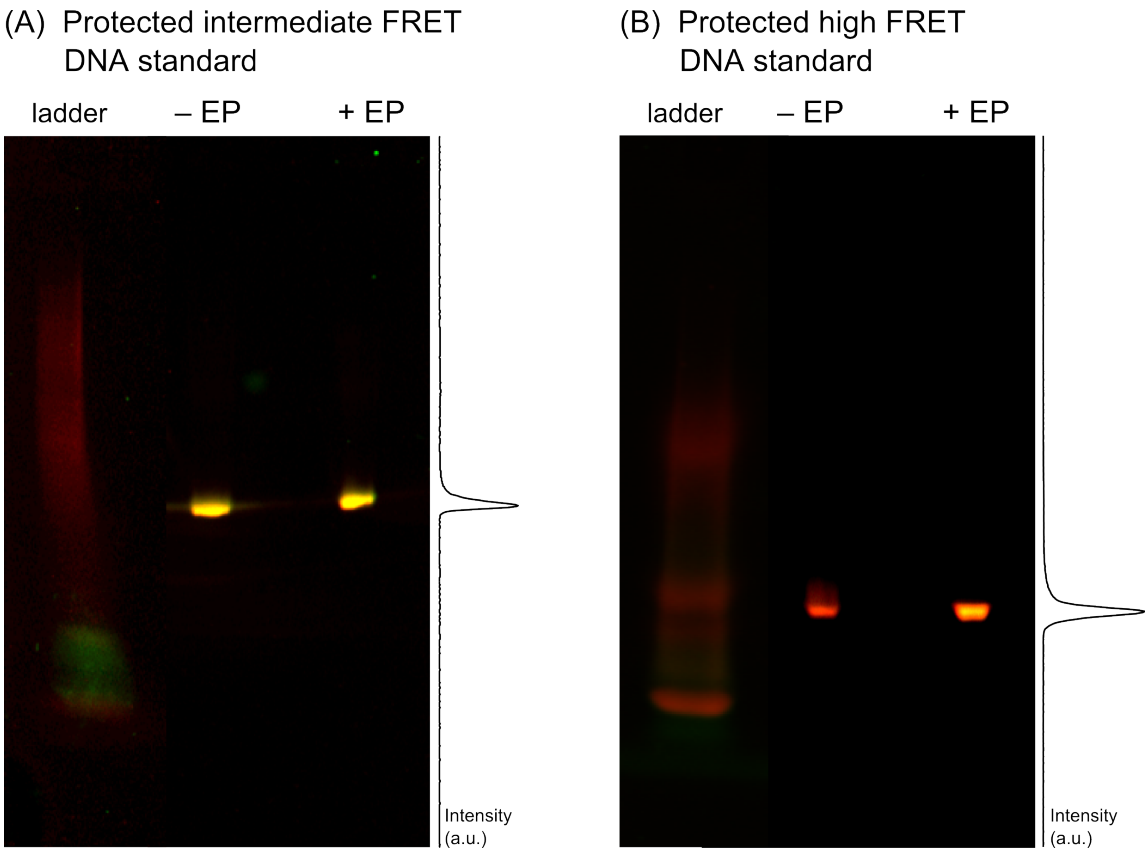


Figure S4. Electroporation has no impact on molecule integrity of protected DNA FRET standards. A) Protected intermediate and B) high FRET DNA standards, non-electroporated (– EP) and electroporated (+ EP, 1.4 kV), were analyzed on native gels showing no significant difference in band size and position. The intensity plot of the electroporated sample along the gel lane shows a single peak corresponding to labeled protected DNAs; no free dye front (ladder) was present in the gels.

Figure S5

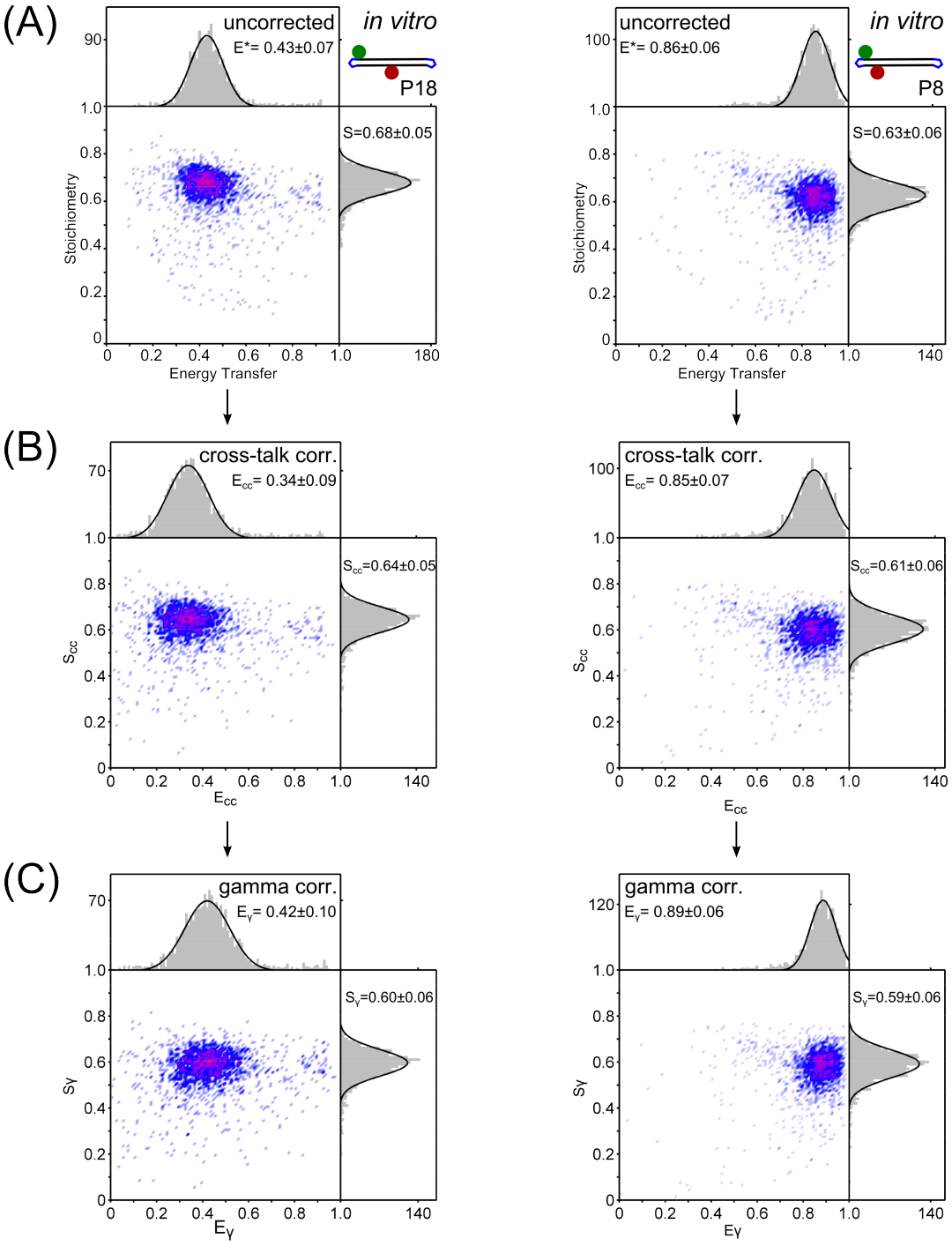


Figure S5. *In vitro* single-molecule FRET correction for intermediate (left) and high FRET (right) protected DNA standards. A) Uncorrected FRET-Stoichiometry histograms for protected DNA standards (see main **Figure 1**). B) Cross-talk corrected FRET and Stoichiometry values (see ESI eq. 1 and 2 with l=0.147±0.004, and d=0.127±0.003). C) Gamma corrected FRET and Stoichiometry values (see ESI eq. 3 and 4 with γ=0.71; for gamma-factor see **FigureS6-A**).

Figure S6

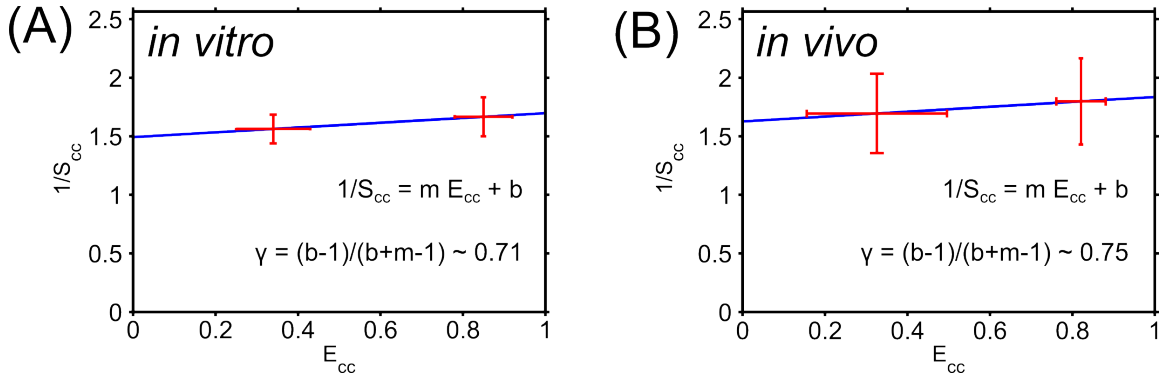


Figure S6. Estimation of gamma-factor for A) *in vitro* confocal single-molecule FRET spectroscopy study and B) *in vivo* single-molecule FRET microscopy study using HILO illumination[11]. The gamma-factor was obtained from a linear regression of $1/S_{cc}$ over E_{cc} (see equation inset, and Ref.[10]) to γ~0.71 *in vitro* and γ~0.75 *in vivo*.

Figure S7

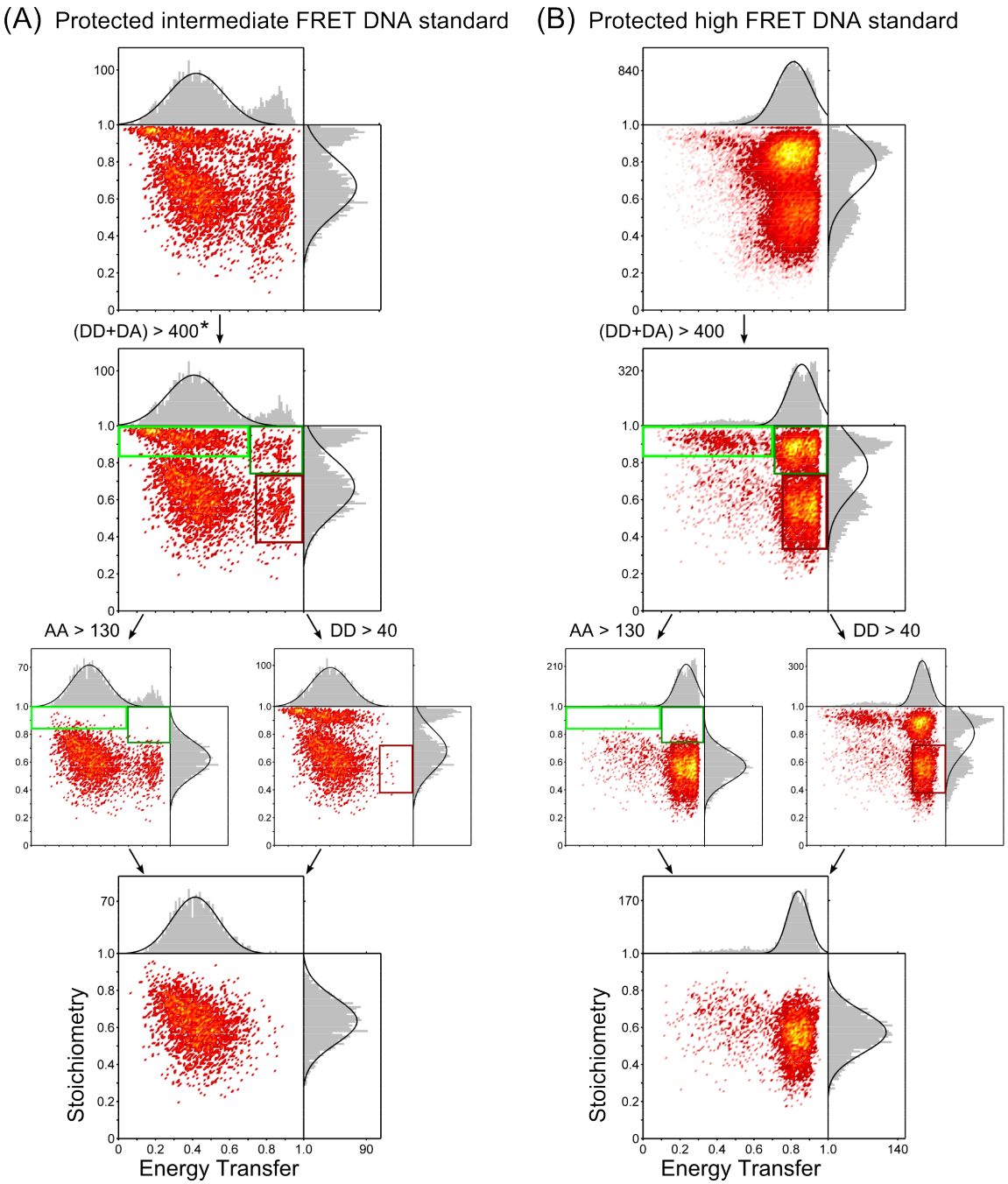


Figure S7. *In vivo* single-molecule FRET-Stoichiometry histograms for A) intermediate and B) high FRET protected DNA standards. Single-molecule data was filtered for PSF-shape (full-width at half-maximum of fitted 2D Gaussian ranging from 50-400 nm), a total fluorophore intensity of (DD+DA)>400 ph/frame, the existence of an acceptor molecule AA>130 ph/frame (green boxes), and autofluorescence in the Donor-channel DD>40 ph/frame (red box). *Values are given in photon counts per frame-time (ph / 20 ms).

Figure S8

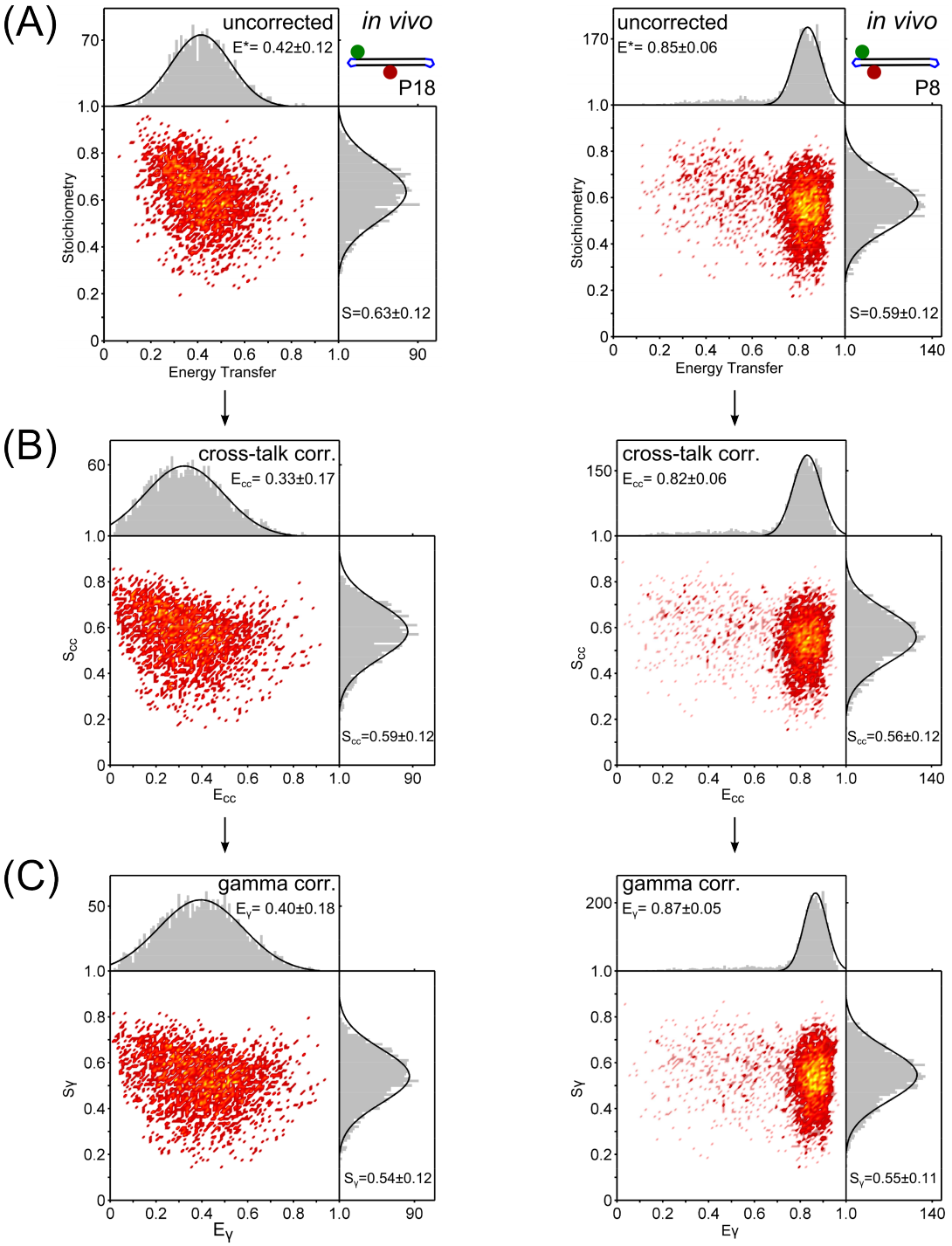


Figure S8. *In vivo* single-molecule FRET correction for intermediate (left) and high FRET (right) protected DNA standards. A) Uncorrected FRET-Stoichiometry histograms for protected DNA standards (see main **Figure 2**). B) Cross-talk corrected FRET and Stoichiometry values (see ESI eq. 1 and 2 with l=0.20±0.02, and d=0.06±0.01). C) Gamma corrected FRET and Stoichiometry values (see ESI eq. 3 and 4 with γ=0.75; for gamma-factor see **FigureS6-B**).

Figure S9

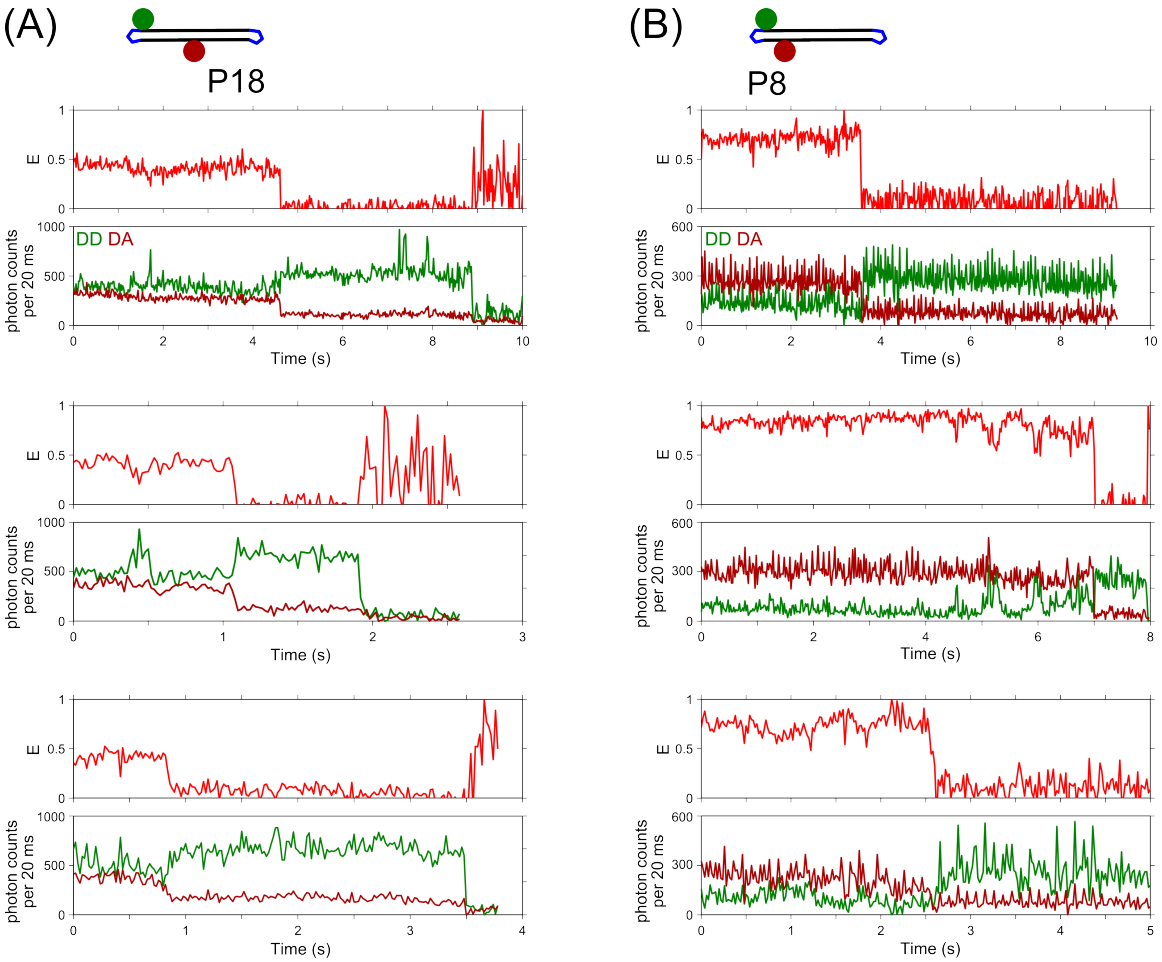


Figure S9. *In vivo* single-molecule FRET time-traces showing single photobleaching events for A) protected intermediate FRET DNA standard (P18) and B) protected high FRET DNA standard (P8). Single-molecule FRET time-traces were corrected for cross-talk contribution and γ-factor as described in the ESI and shown in **Figure S8**. Single-molecule FRET time-traces showed an increase in DD-signal after the photobleaching of the acceptor dye (anti-correlated DD and DA signal, a signature of single-molecule FRET), which was often followed by the photobleaching of the donor dye (all time-traces in (A) and mid time-trace in (B)). Spikes in the DD-signal are caused due to nearby D-only molecules.