Supporting Information

Iqbal *et al.* *10.1073/pnas.0801707105*

SI Materials

DNA Sequences Used to Construct the Duplex Series

10bp
Cy3-5'-CCACTCTAGG-3'
Cy5-5'-CCTAGAGTGG-3'

11bp
Cy3-5'-CCACTGCTAGG-3'
Cy5-5'-CCTAGCAGTGG-3'

12bp
Cy3-5'-CCACTGGCTAGG-3'
Cy5-5'-CCTAGCCAGTGG-3'

13bp
Cy3-5'-CCACTGCGCTAGG-3'
Cy5-5'-CCTAGCGCAGTGG-3'

14bp
Cy3-5'-CCACTGCTGCTAGG-3'
Cy5-5'-CCTAGCAGCAGTGG-3'

15bp
Cy3-5'-CCACTGCCTGCTAGG-3'
Cy5-5'-CCTAGCAGGCAGTGG-3'

16bp
Cy3-5'-CCACTGCACTGCTAGG-3'
Cy5-5'-CCTAGCAGTGCAGTGG-3'

17bp
Cy3-5'-CCACTGCACCTGCTAGG-3'
Cy5-5'-CCTAGCAGGTGCAGTGG-3'

18bp
Cy3-5'-CCACTGCACGCTGCTAGG-3'
Cy5-5'-CCTAGCAGCGTGCAGTGG-3'

19bp
Cy3-5'-CCACTGCACCGCTGCTAGG-3'
Cy5-5'-CCTAGCAGCGGTGCAGTGG-3'

20bp
Cy3-5'-CCACTGCACTCGCTGCTAGG-3'
Cy5-5'-CCTAGCAGCGAGTGCAGTGG-3'

22bp
Cy3-5'-CCACTGCACTGCCGCTGCTAGG-3'
Cy5-5'- CCTAGCAGCGGCAGTGCAGTGG-3'

24bp
Cy3-5'-CCACTGCACTGCTGCGCTGCTAGG-3'
Cy5-5'-CCTAGCAGCGCAGCAGTGCAGTGG-3'

The RNA/DNA duplexes had the same sequence, where the Cy3-labeled strand was exchanged for RNA, for example, the 10-nt strand was Cy3–5′-CCACUCUAGG-3′. These were hybridized with the same Cy5-labeled DNA strands.

Modeling the Dependence of FRET Efficiency on Duplex Length for Rigidly Stacked Fluorophores. We calculate the interfluorophore distance for each duplex R as

$$R = ((L - 1)^*H) + D$$

where L is the length of the helix (bp), H is the helical rise per bp step, and D is the additional axial separation for the two fluorophores.

The angle between the transition moments A is calculated as

$$A = ((L - 1)^*T) + C3A + C5A$$

where T is the twist angle for each base pair, and $C3A$ and $C5A$ are the rotations of Cy3 and Cy5 relative to the terminal base pairs.

For B form DNA $H = 3.6$ Å and $T = 36°$, and $D = 8$ Å, $C3A + C5A = 62°$.

The orientation factor is calculated as.

$$\kappa^2 = (\sin\Theta_D{}^*\sin\Theta_A{}^*\cos A\text{-}2{}^*\cos\Theta_D{}^*\cos\Theta_A)^2$$

For the case where the fluorophore planes are parallel and close to coaxial (i.e., Θ_D, $\Theta_A \sim 90°$), this reduces to

$$\kappa^2 = \cos^2 A$$

κ^2 and hence R_0 are calculated for each length, and thus the value for E_{FRET}. This calculation has been performed over 141 steps from 10 to 24 bp in Kalaidagraph, and interpolated (Fig. 2).

Modeling the Dependence of FRET Efficiency on Duplex Length with Lateral Mobility of Fluorophores. As before, the interfluorophore distance for each duplex is given by

$$R = ((L - 1)^*H) + D$$

and the mean angle between the transition moments A is given by

$$A = ((L - 1)^*T) + C3A + C5A$$

For each species, we set up an array of angles (AA) over a ± 100° range about the mean, and calculate κ^2 and hence R_0 for each.

The value of E_{FRET} is calculated for each angular position (E_{AA}), and the resulting distribution summed, weighted by its distance from the mean angle by using a Gaussian distribution, that is,

$$E_{\mathrm{FRET}} = \Sigma\ (E_{AA}{}^*P)$$

where

$$P = \exp\text{-}(AA^2/1.44H^2)$$

where H is the half-width. The sum of P is normalized to unity.

This simulation procedure was implemented in a MATLAB program.

Modeling the Dependence of FRET Efficiency on Duplex Length with Lateral Mobility of Fluorophores, Together with an Unstacked Fraction. Where a fraction of unstacked fluorophores is allowed, that fraction is calculated by using $\kappa^2 = 2/3$, together with an increment to the interfluorophore distance (7 Å was found to give the best agreement with experimental data). The resulting E_{FRET} is calculated as a linear combination of the contributions from the stacked fluorophore (calculated as above) and the freely mobile fluorophore.

Time-Resolved Duplexes Data Analysis. From the short time traces average donor (I_D) and acceptor (I_A) intensities are calculated over a period of 300 to 500 ms. These values are recorded for up to 40,000 molecules for further processing. Using a scatter plot of total intensity vs. E_{FRET}, lower and higher thresholds are determined so that E_{FRET} populations of Cy3-Cy5 pairs are selected. All of the molecules outside the threshold limits are discarded. Because the time traces are short we usually don't see the photobleached flat lines in time traces to be used as the background value. We employ two techniques to determine the background values (I_{Do} and I_{Ao}) as shown in supporting information (SI) Fig. S4. In the first method, we make histograms of I_D and I_A that are measured from the last 3–5 data points of each time trace so that fluorophores are more likely to photobleach (Fig. S4*C*). The flat lined I_D and I_A will appear as a peak near 0. Peak positions are then used as the background values. The second method involves the dual histogram of I_D and I_A pairs from all of the molecules of the duplex specie in question. A color density plot of such a distribution is shown in Fig. S4*B*. Because, in theory, Cy3 and Cy5 intensity ratio is constant, E_{FRET} populations in the dual histogram can be fitted to a line. The E_{FRET} of Cy3-Cy5 pairs shows up at a higher I_A value, whereas Cy3 only duplexes populates lower I_A values. Fitting each of these populations to a line and finding the intersection point provides us with the background values, I_{Do} and I_{Ao}. Sometimes some I_D and I_A population is found near the origin because of nonactive Cy3-Cy5 pairs, the center of which should also coincide with the intersection. By using these guidelines, we determine the background values and subtract them from the mean intensities to calculate the apparent E_{FRET} value for each molecule.

$$E_{\mathrm{FRET}} = (1 + [I_D - I_{Do}]/[I_A - I_{Ao}])^{-1}$$

From these values we prepare the initial FRET histograms, which usually comprise two dominant peaks. The first peak is the donor-only molecules with nonactive Cy5. It is not at zero, because some of the donor emission leaks into the acceptor channel, yielding a donor-only peak. The peak position of this value, p, is then used to perform crosstalk correction. Another factor in our apparatus arises from the long-pass dichroic mirror in the dual-view emission pathway, which reflects a fraction of the Cy5 signal into the Cy3 channel. We measured the ratio of this backreflection by using a construct where the Cy3 and Cy5 are so close to each other that all of the emission comes from Cy5, and Cy3 emission should be zero. E_{FRET} from that substrate should be 1, but we only measure ≈0.88 indicating a 12% reflection ratio. The experiment was repeated with a number of substrates where the fluorophores are at very close proximity; the average value for the reflection ratio is calculated to be r = 0.12 with a very narrow distribution.

True FRET efficiency (E_{FRET}) of each molecule is then calculated from averaged donor and acceptor emissions along with the crosstalk correction factor, p, and backreflection ratio, r, using the formula:

$$E^*_{\mathrm{FRET}} = (E_{\mathrm{FRET}} - p)/(1 - p - r)$$

The gamma factor in our system (1) was measured by looking at Cy5 photobleaching or blinking traces, where the decrease in the acceptor intensity should be equal to the increase in donor intensity in theory. If they are not equal, then the ratio of donor-to-acceptor emissions should be scaled accordingly, which is called gamma correction. This value was measured and found to be 1 in our system, and therefore we did not apply any additional correction.

E^*_{FRET} values from each molecule were then used to generate the histograms for DNA and RNA/DNA duplexes. The peak positions of Gaussian fits were insensitive to the bin size in a range from ≈0.005 to ≈0.05; we chose 0.02. Histograms clearly show the peaks for Cy3-Cy5 pairs and the donor-only signal $E^*_{\mathrm{FRET}} = 0$. E^*_{FRET} values less than that of the single pair populate the intermediate E_{FRET} region; these arise because we use all of the molecules indiscriminately within the threshold boundaries, and because of multiple duplexes in single vesicles with at least one photobleached Cy5. The magnitude of intermediate peaks depends on the relative population of the vesicles with multiple duplexes. We accounted for the intermediate peaks by adding an additional Gaussian peak in the fits. As the E^*_{FRET} value gets lower, and the two major peaks starts to overlap, the presence or absence of the intermediate Gaussian peak does not change the higher E^*_{FRET} peak; so, for the longest duplex species, the intermediate peak is omitted.

1. Ha T, *et al.* (1999) Single molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proc Natl Acad Sci USA* 96:893–898.

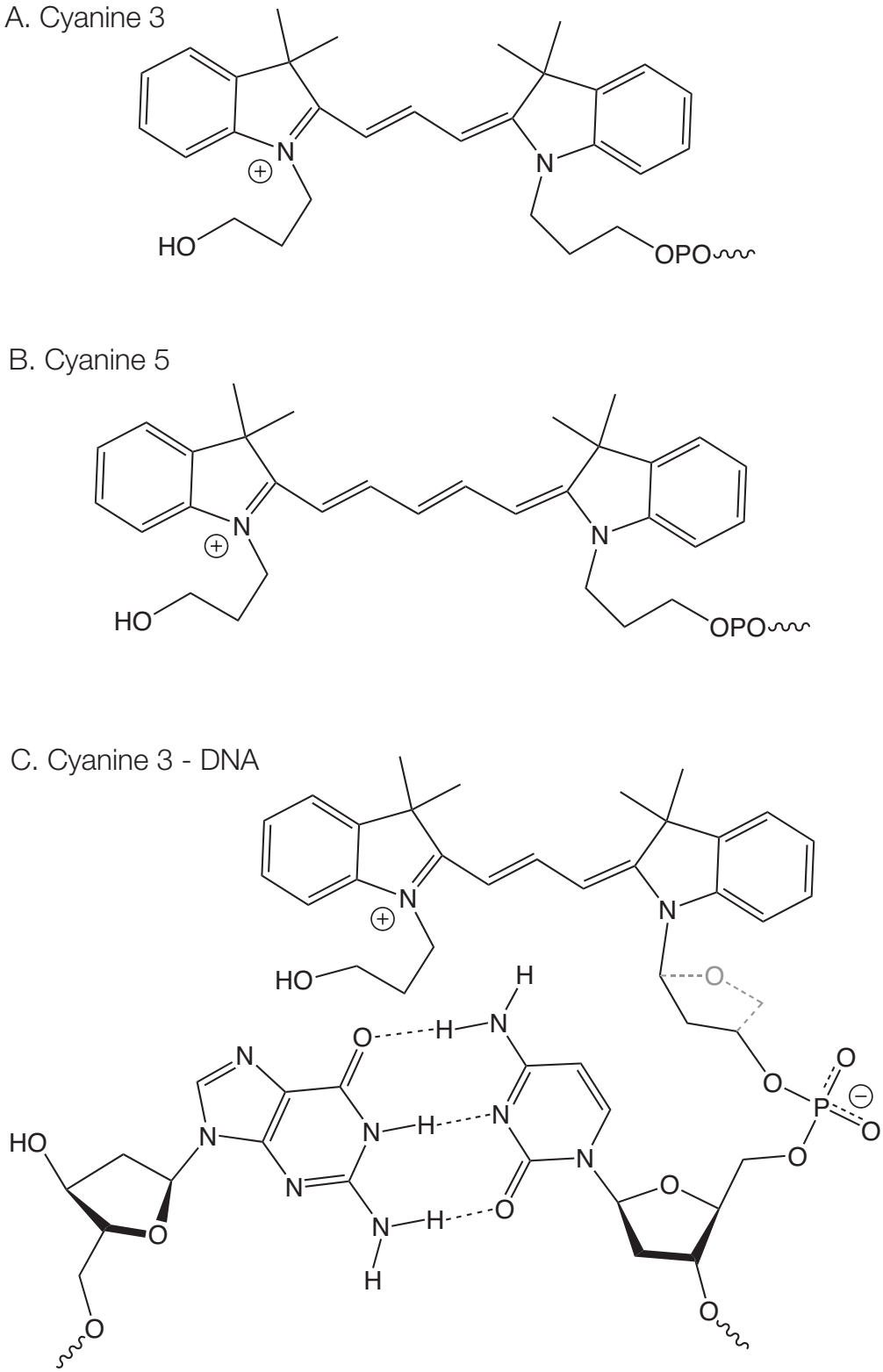

Fig. S1. The chemical structures of the cyanine fluorophores. (*A* and *B*) The structures of Cy3 and Cy5, respectively. (*C*) The structure of Cy3 attached to a terminal GC base pair of DNA. The broken line shows how the 3-carbon tether attaching Cy3 to the 5′-phosphate can be drawn equivalently to an additional deoxyribose group.

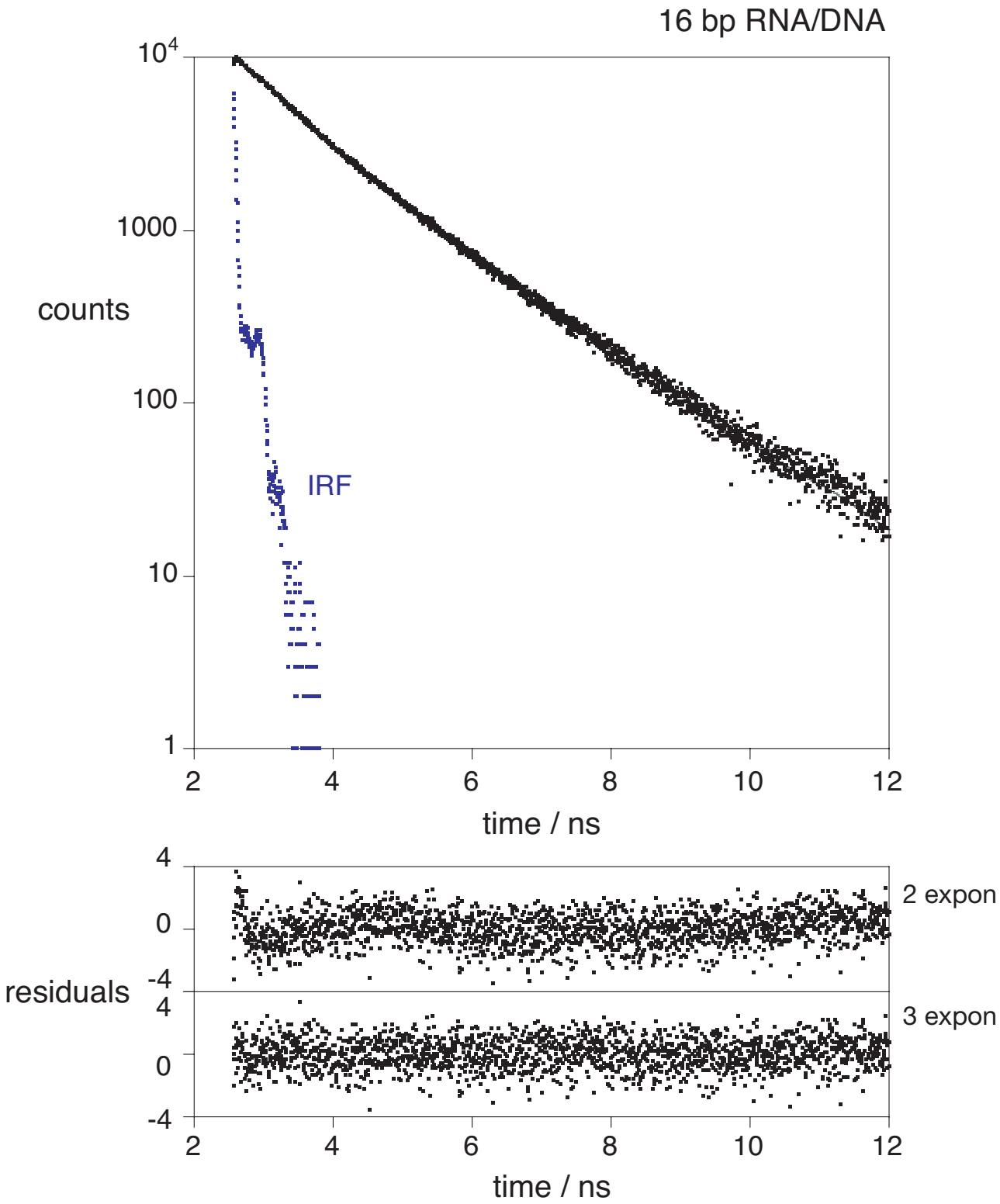

Fig. S2. Time-resolved fluorescent lifetime analysis of Cy3 attached to double-stranded RNA/DNA hybrid duplex. Fluorescent decay curve for Cy3 attached to a 16-bp RNA/DNA duplex, showing the experimental data and the instrument response function (IRF), and the fit to three exponential functions (line). Plots of weighted residuals for fitting the data to two ($\chi^2 = 1.113$) or three ($\chi^2 = 0.986$) exponential functions are shown below.

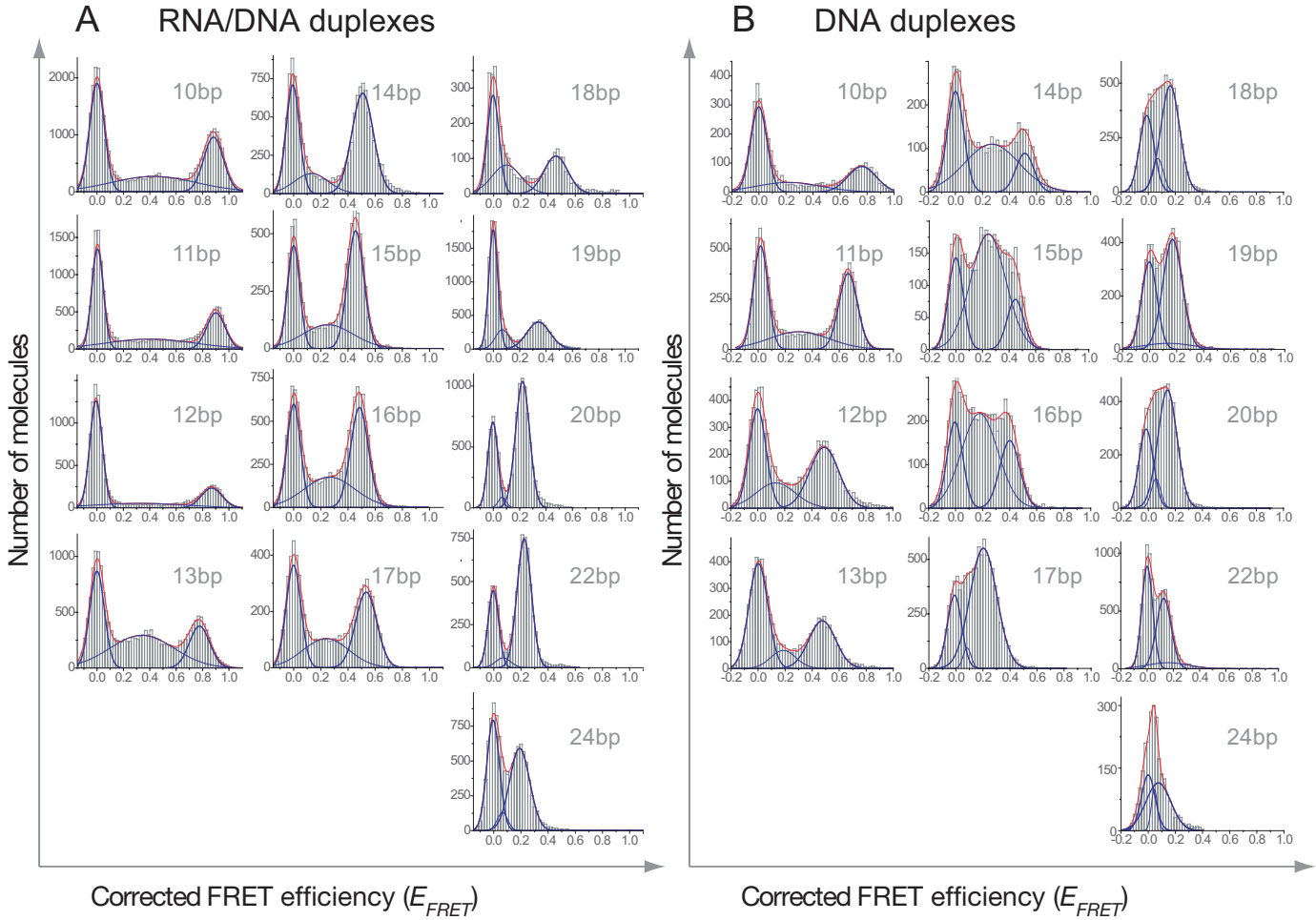

Fig. S3. Histograms of FRET efficiency (E_{FRET}) for phospholipid vesicle-encapsulated DNA and RNA/DNA duplex species. The *y* axes represent the number of molecules in each FRET efficiency bin. Mean E_{FRET} value for each molecule have been calculated from Cy3 and Cy5 fluorescent intensities that are corrected for the crosstalk between Cy3 and Cy5 channels, backreflection from the dichroic mirror surface in the dual-view emission pathway, and the background. Each histogram contains two major peaks (thick blue curves) and an intermediate peak (thin blue curves). The first major peak at $E_{FRET} = 0$ corresponds to duplexes with active Cy3 only, and the second major peak at higher E_{FRET} resulting from duplex molecules with an active Cy3-Cy5 pair. The latter was used to calculate the E_{FRET} value for each species. Intermediate peaks are due to the multiple Cy3-Cy5 pairs in vesicles, some of which have photobleached Cy5, yielding average E_{FRET} values that are less than E_{FRET} of the active Cy3-Cy5 pairs.

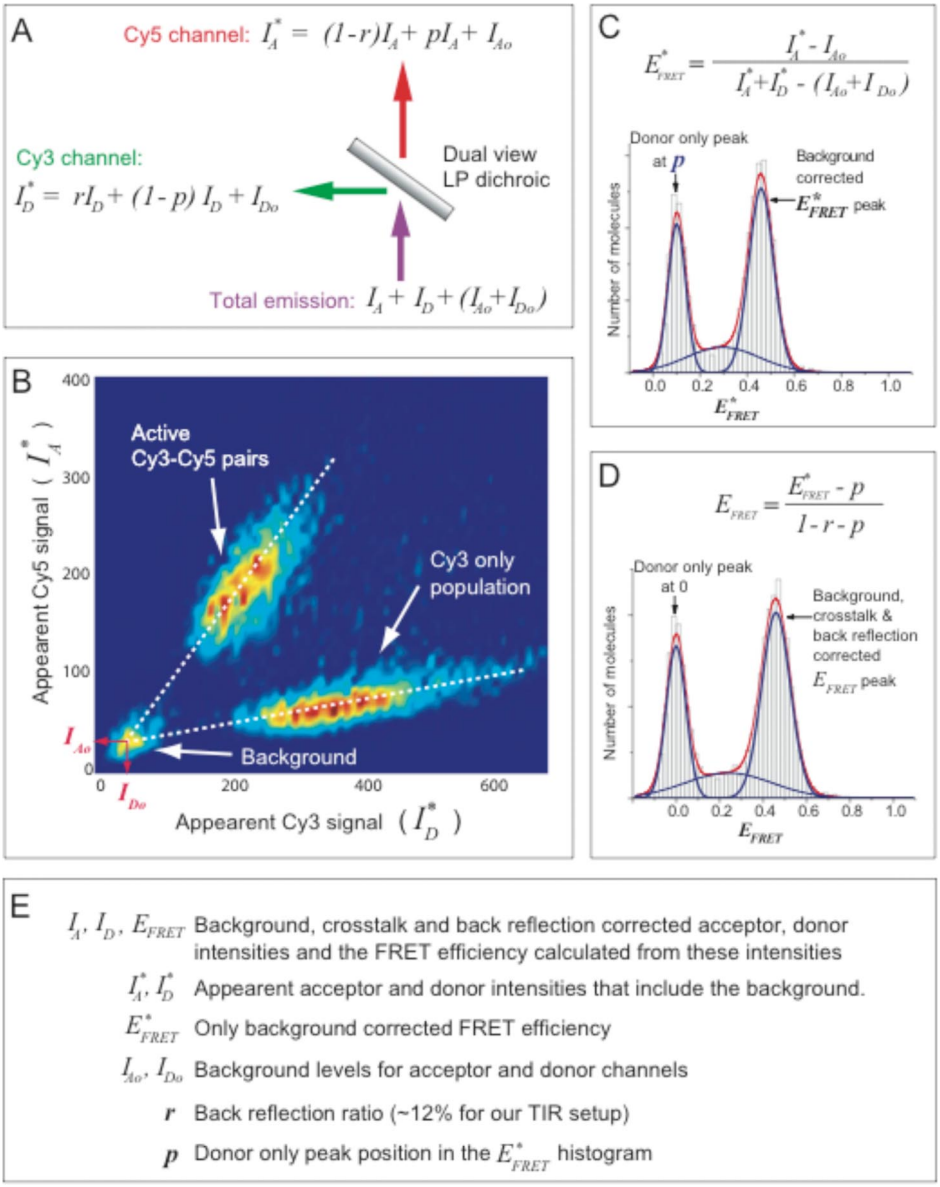

Fig. S4. Data analysis method for single-molecule experiments. (*A*) How the total emission from the Cy3 and Cy5 molecules is separated to their respective channels at the 650-nm long-pass dichroic mirror. The explanation for all of the terms in the formulas is provided in *E*. (*B*) The background for the donor and acceptor channel is determined by using a dual histogram of apparent Cy3 and Cy5 signal pairs. The particular density map was generated from 8,800 molecules of 15-bp DNA/RNA duplexes. The color represents the density of the molecules, red being the highest density and the blue being the lowest. Because E_{FRET} is determined by the ratio of the acceptor signal-to-donor signal, a population of certain E_{FRET} pairs should be distributed along a straight line in this plot depending on the total intensity of the pair. Two major populations belong to active Cy3-Cy5 pairs and only Cy3 active duplexes as shown. A small third population appears close to origin and arises from duplexes with inactive Cy3 or Cy5. Because we have a background signal in both channels, this third population is not at the origin, so the center of this population gives the background values in the respective channels as shown with red arrows. The intersection point of the lines that are fitted to the two major populations should also give this background value. Both techniques are used. (*C*) E_{FRET} histogram after the background correction is shown for the 8,800 15-mer RNA/DNA molecules. The donor-only peak position (crosstalk correction parameter) p is determined at this stage via Gaussian fitting. The intermediate broad peak arises mainly from vesicles with multiple duplexes of active and only Cy3 active pairs. (*D*) The same histogram after the crosstalk (p) and the backreflection ($r = 0.12$) corrections are performed. The true E_{FRET} value of the Cy3-Cy5 pairs on this duplex is given by the center of the second major peak. (*E*) The explanation for the symbols used in the formulae used above.

Table S1. Fluorescent lifetime data for 5-Cy3-labeled DNA and RNA/DNA duplexes

	Three components				Two components			
	τ	α	f	χ^2	τ	α	f	χ^2
DNA 22 bp								
	0.39	0.019	16.5	0.957	0.61	0.032	42.6	1.238
	1.05	0.026	61.3		1.53	0.017	57.4	
	1.97	0.005	22.2					
DNA 16 bp								
	0.39	0.018	16.9	1.062	0.59	0.029	40.6	1.326
	1.04	0.025	61.6		1.47	0.017	59.4	
	1.91	0.005	21.4					
RNA/DNA 22 bp								
	0.36	0.011	7.0	1.060	0.71	0.026	33.0	1.493
	1.03	0.026	49.1		1.64	0.023	67.0	
	1.84	0.013	43.9					
RNA/DNA 16 bp								
	0.31	0.009	5.3	0.986	0.69	0.025	31.2	1.113
	0.96	0.025	43.7		1.63	0.023	68.8	
	1.77	0.016	51.1					

The results of fitting to two and three exponential functions are presented. τ (ns), α, and f are the fluorescent lifetime, relative amplitude, and steady-state fractional intensity for each component, and χ^2 is the chi-squared statistic.