Biophysical Journal, Volume *117*

Supplemental Information

Single-Molecule Observation of Ligand Binding and Conformational Changes in FeuA

Marijn de Boer, Giorgos Gouridis, Yusran Abdillah Muthahari, and Thorben Cordes

Appendix S1

The donor and acceptor fluorophore distance ratio of state 1 and 2, denoted by r_1 and r_2, respectively, satisfies (see Materials and Methods section):

$$\left(\frac{r_1}{r_2}\right)^6 = \frac{\phi_{1\cdot A}}{\phi_{2\cdot A}} \frac{n_{1\cdot DD}}{n_{1\cdot DA}} \frac{n_{2\cdot DA}}{n_{2\cdot DD}} \tag{S1}$$

where $n_{i\cdot DA}$ and $n_{i\cdot DD}$ are the (background- and spectral crosstalk-corrected) acceptor and donor count rates upon donor excitation when being in state i ($i = 1{,}2$), respectively, and $\phi_{1\cdot A}$ and $\phi_{2\cdot A}$ are the acceptor quantum yields of state 1 and 2, respectively. Eq. S1 holds when the refractive index of the medium, the dipole orientation factor κ^2, the molar extinction coefficient of the acceptor and the normalized donor emission spectra are the same for state 1 and 2.

Here, we will consider how the distance ratio $(r_1/r_2)^6$ can be estimated from the data. We use the following notation: $N_{i\cdot XY}$ represents the measured count rate of $n_{i\cdot XY}$, that is, the background- and spectral crosstalk-corrected count rate of Y emission (D̲onor, A̲cceptor) upon X excitation (D̲onor, A̲cceptor) when being in state i ($i = 1{,}2$) and R_i and r_i are the measured and true distance of state i ($i = 1{,}2$), respectively. In other words, R_i and $N_{i\cdot XY}$ are the *point estimators* for r_i and $n_{i\cdot XY}$, respectively. In the derivation below we assume that the relaxation times of the excited states of the fluorophores are short compared to the time between two consecutively detected photons, so that there is no correlation between the photons and the distribution of $N_{i\cdot XY}$ can be approximated by a Poisson distribution with parameter $n_{i\cdot XY}$ (1). Then,

$$\left(\frac{R_1}{R_2}\right)^6 = \left\langle\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right\rangle \left\langle\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right\rangle \left\langle\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right\rangle \tag{S2}$$

with

$$\begin{aligned} \left\langle\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right\rangle &= \frac{1}{k}\sum \frac{N_{1\cdot AA}}{N_{2\cdot AA}} \\ \left\langle\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right\rangle &= \frac{1}{p}\sum \frac{N_{1\cdot DD}}{N_{1\cdot DA}} \end{aligned} \tag{S3}$$

$$\left\langle \frac{N_{2\cdot DA}}{N_{2\cdot DD}} \right\rangle = \frac{1}{w} \sum \frac{N_{2\cdot DA}}{N_{2\cdot DD}}$$

where k, p and w denote the number of observations, is an unbiased and consistent estimator. The sum in Eq. S3 extends over all observations, i.e., the total number of traces or time-bins. Noteworthy, in the absence of additional fluorophore quenching we have $\phi_{1\cdot A} = \phi_{2\cdot A}$, so that

$$\left(\frac{r_1}{r_2}\right)^6 = \frac{n_{1\cdot DD}}{n_{1\cdot DA}} \frac{n_{2\cdot DA}}{n_{2\cdot DD}} \quad \text{(S4)}$$

and can be estimated from the data by using the estimator

$$\left(\frac{R_1}{R_2}\right)^6 = \left\langle \frac{N_{1\cdot DD}}{N_{1\cdot DA}} \right\rangle \left\langle \frac{N_{2\cdot DA}}{N_{2\cdot DD}} \right\rangle \quad \text{(S5)}$$

Estimation of the interprobe distance ratio does not require the determination of the gamma factor γ or the Föster radius R_0. Below we will focus on the more general scenario as given by Eq. S1 and S2 and note that the results also apply to the more specific case of Eq. S4 and S5.

First, we will show that $(R_1/R_2)^6$ is an unbiased estimator for $(r_1/r_2)^6$, that is, $\mathbb{E}[(R_1/R_2)^6] = (r_1/r_2)^6$, where $\mathbb{E}[X]$ is the expectation value of the random variable X. Each term in the product of Eq. S2 is independent of each other, so that

$$\mathbb{E}\left[\left\langle \frac{N_{1\cdot AA}}{N_{2\cdot AA}} \right\rangle \left\langle \frac{N_{1\cdot DD}}{N_{1\cdot DA}} \right\rangle \left\langle \frac{N_{2\cdot DA}}{N_{2\cdot DD}} \right\rangle\right] = \mathbb{E}\left[\left\langle \frac{N_{1\cdot AA}}{N_{2\cdot AA}} \right\rangle\right] \mathbb{E}\left[\left\langle \frac{N_{1\cdot DD}}{N_{1\cdot DA}} \right\rangle\right] \mathbb{E}\left[\left\langle \frac{N_{2\cdot DA}}{N_{2\cdot DD}} \right\rangle\right] \quad \text{(S6)}$$

Furthermore, it holds that

$$\mathbb{E}\left[\left\langle \frac{N_{1\cdot AA}}{N_{2\cdot AA}} \right\rangle\right] = \mathbb{E}\left[\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right]$$

$$\mathbb{E}\left[\left\langle \frac{N_{1\cdot DD}}{N_{1\cdot DA}} \right\rangle\right] = \mathbb{E}\left[\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right] \quad \text{(S7)}$$

$$\mathbb{E}\left[\left\langle \frac{N_{2\cdot DA}}{N_{2\cdot DD}} \right\rangle\right] = \mathbb{E}\left[\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right]$$

as the terms in the sum of Eq. S3 are independent and have the same distribution. By combining Eq. S6 and S7 we have,

$$\mathbb{E}\left[\left\langle\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right\rangle\left\langle\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right\rangle\left\langle\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right\rangle\right] = \mathbb{E}\left[\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right]\mathbb{E}\left[\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right]\mathbb{E}\left[\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right] \tag{S8}$$

We can approximate each term in Eq. S8 further by approximating it to second-order,

$$\mathbb{E}\left[\frac{X}{Y}\right] \cong \frac{\mathbb{E}[X]}{\mathbb{E}[Y]}\left(1 - \frac{\mathrm{Cov}(X,Y)}{\mathbb{E}[X]\mathbb{E}[Y]} + \frac{\mathrm{Var}(Y)}{\mathbb{E}[Y]^2}\right) \tag{S9}$$

The covariances between $N_{1\cdot AA}$ and $N_{2\cdot AA}$, $N_{1\cdot DD}$ and $N_{1\cdot DA}$ and of $N_{2\cdot DA}$ and $N_{2\cdot DD}$ are zero(2). Further, under our assumption that $N_{i\cdot XY}$ has a Poisson distribution it holds that $\mathrm{Var}(N_{i\cdot XY})/\mathbb{E}[N_{i\cdot XY}]^2 = n_{i\cdot XY}{}^{-1}$ and is thus is negligible when $n_{i\cdot XY} \gg 1$. Hence, we can safely make the approximation that

$$\begin{aligned}
\mathbb{E}\left[\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right] &= \frac{\mathbb{E}[N_{1\cdot AA}]}{\mathbb{E}[N_{2\cdot AA}]} = \frac{n_{1\cdot AA}}{n_{2\cdot AA}} \\
\mathbb{E}\left[\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right] &= \frac{\mathbb{E}[N_{1\cdot DD}]}{\mathbb{E}[N_{1\cdot DA}]} = \frac{n_{1\cdot DD}}{n_{1\cdot DA}} \\
\mathbb{E}\left[\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right] &= \frac{\mathbb{E}[N_{2\cdot DA}]}{\mathbb{E}[N_{2\cdot DD}]} = \frac{n_{2\cdot DA}}{n_{2\cdot DD}}
\end{aligned} \tag{S10}$$

The count rate $n_{i\cdot AA}$ is the product of the probabilities that (i) the acceptor is excited by the laser (p_{EX}), (ii) the acceptor decays to its ground state by photon emission ($\phi_{i\cdot A}$) and (iii) the emitted photon is detected (η_A) (2):

$$n_{i\cdot AA} = p_{EX}\phi_{i\cdot A}\eta_A \tag{S11}$$

By using Eq. S10 and S11 we have

$$\mathbb{E}\left[\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right] = \frac{\phi_{1\cdot A}}{\phi_{2\cdot A}} \tag{S12}$$

when p_{EX} and η_A remain the same. By combing Eq. S8, S10 and S12 it follows that

$$\mathbb{E}\left[\left(\frac{R_1}{R_2}\right)^6\right] = \frac{\phi_{1\cdot A}}{\phi_{2\cdot A}}\frac{n_{1\cdot DD}}{n_{1\cdot DA}}\frac{n_{2\cdot DA}}{n_{2\cdot DD}} \tag{S13}$$

From Eq. S1 and S13 we have,

$$\mathbb{E}\left[\left(\frac{R_1}{R_2}\right)^6\right] = \left(\frac{r_1}{r_2}\right)^6 \tag{S14}$$

and shows that $(R_1/R_2)^6$ is an unbiased estimator for $(r_1/r_2)^6$.

If the random variables $X_i \cdots X_n$ are independent, then it can be shown that

$$\mathrm{Var}(X_i \cdots X_n) = \prod_{i=1}^{n}(\mathrm{Var}(X_i) + \mathbb{E}[X_i]^2) - \prod_{i=1}^{n}\mathbb{E}[X_i]^2. \tag{S15}$$

where $\mathrm{Var}(X_i)$ is the variance of X_i. The terms in the product of Eq. S2 are independent so by using Eq. S15 we find that

$$\begin{aligned}
&\mathrm{Var}\left[\left\langle\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right\rangle\left\langle\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right\rangle\left\langle\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right\rangle\right] \\
&\quad = \left(\mathrm{Var}\left(\left\langle\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right\rangle\right) + \mathbb{E}\left[\left\langle\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right\rangle\right]^2\right)\left(\mathrm{Var}\left(\left\langle\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right\rangle\right)\right. \\
&\quad \left. + \mathbb{E}\left[\left\langle\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right\rangle\right]^2\right)\left(\mathrm{Var}\left(\left\langle\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right\rangle\right) + \mathbb{E}\left[\left\langle\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right\rangle\right]^2\right) \\
&\quad - \mathbb{E}\left[\left\langle\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right\rangle\right]^2 \mathbb{E}\left[\left\langle\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right\rangle\right]^2 \mathbb{E}\left[\left\langle\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right\rangle\right]^2
\end{aligned} \tag{S16}$$

As before, each term in the sum of Eq. S3 are also independent and have the same distribution, so

$$\begin{aligned}
\mathrm{Var}\left[\left\langle\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right\rangle\right] &= \frac{1}{k}\mathrm{Var}\left[\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right] \\
\mathrm{Var}\left[\left\langle\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right\rangle\right] &= \frac{1}{p}\mathrm{Var}\left[\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right] \\
\mathrm{Var}\left[\left\langle\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right\rangle\right] &= \frac{1}{w}\mathrm{Var}\left[\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right]
\end{aligned} \tag{S17}$$

By combining Eq. S7, S16 and S17 we obtain,

$$\mathrm{Var}\left[\left(\frac{R_1}{R_2}\right)^6\right] = \left(k^{-1}\mathrm{Var}\left(\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right) + \mathbb{E}\left[\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right]^2\right)\left(p^{-1}\mathrm{Var}\left(\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right) + \mathbb{E}\left[\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right]^2\right)\left(w^{-1}\mathrm{Var}\left(\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right) + \mathbb{E}\left[\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right]^2\right) - \mathbb{E}\left[\frac{N_{1\cdot AA}}{N_{2\cdot AA}}\right]^2 \mathbb{E}\left[\frac{N_{1\cdot DD}}{N_{1\cdot DA}}\right]^2 \mathbb{E}\left[\frac{N_{2\cdot DA}}{N_{2\cdot DD}}\right]^2 \quad \text{(S18)}$$

To show that $(R_1/R_2)^6$ is a consistent estimator, we need to show that $(R_1/R_2)^6$ converges in probability to $(r_1/r_2)^6$. We define $\boldsymbol{n} = \{k, p, l\}$, where $(R_1/R_2)^6$ depends implicitly on $\boldsymbol{n}$. We should proof that for any $\varepsilon > 0$ it holds that,

$$\lim_{\boldsymbol{n}\to\infty} P(|(R_1/R_2)^6 - (r_1/r_2)^6| > \varepsilon) = 0 \quad \text{(S19)}$$

where $\boldsymbol{n} \to \infty$ should be understood as $k \to \infty$, $p \to \infty$ and $w \to \infty$. By using Chebyshev's inequality and $\mathbb{E}[(R_1/R_2)^6] = (r_1/r_2)^6$ we can obtain an upper bound for $P(|(R_1/R_2)^6 - (r_1/r_2)^6| > \varepsilon)$,

$$P(|(R_1/R_2)^6 - (r_1/r_2)^6| > \varepsilon) \leq \frac{\mathrm{Var}((R_1/R_2)^6)}{\varepsilon^2} \quad \text{(S20)}$$

From Eq. S18 it follows that

$$\lim_{\boldsymbol{n}\to\infty} \mathrm{Var}((R_1/R_2)^6) = 0 \quad \text{(S21)}$$

thereby proving that for any $\varepsilon > 0$ Eq. S19 is true. In conclusion, $(R_1/R_2)^6$ is an unbiased and consistent estimator for $(r_1/r_2)^6$.

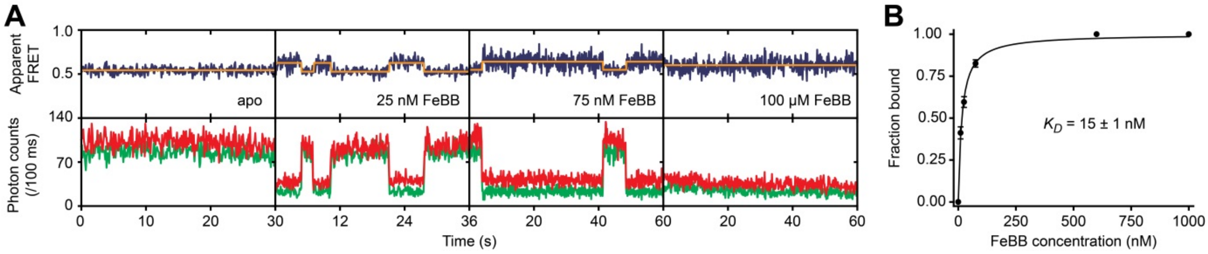

FIGURE S1. Ligand dependence on the protein conformational equilibrium. (A) Representative fluorescence trajectories of FeuA in the absence and presence of varying concentrations of FeBB as indicated. In all fluorescence trajectories presented: top panel shows calculated apparent FRET efficiency (blue) from the donor (green) and acceptor (red) photon counts as shown in the bottom panels. Orange lines indicates most probable state-trajectory of the Hidden Markov Model (HMM). (B) Fraction of time FeuA is in the bound (low intensity) level. The points denote the fraction of time the molecules are in the low intensity level, relative to the total observation time. The error denotes the s.d. of 10^5 bootstrapping steps on all traces recorded at the same ligand concentration. The continuous line denotes the fit to the Hill-Langmuir equation (see Eq. 15), with a 95% confidence interval for K_D indicated.

Table S1. Primers used in this study

Primer	Primer sequence (5’ to 3’)
Forward primer isolation feuA from gDNA (inserts NdeI site)	GGGAATTCCATATGGGCAGTAAAAATGAATCAACTGCCAGCAAG
Reverse primer isolation feuA from gDNA (inserts HindIII site)	GACCCGAAGCTTGTTTTGTGTCAATTTTTCAGCAGCCGCTTT
Forward primer to abolish internal feuA NdeI site	GAATACCTTGATAAAACATACGAAGTAACTGTACCGACA
Reverse primer to abolish internal feuA NdeI site	TGTCGGTACAGTTACTTCGTATGTTTTATCAAGGTATTC
Forward FeuA(Q112C)	TTTCCGGAAAAAACGCTGTGCAAAATCAGCACAGCAGGC
Reverse FeuA(Q112C)	GCCTGCTGTGCTGATTTTGCACAGCGTTTTTTCCGGAAA
Forward FeuA(I255C)	GATTTAGAGAAA AATCCATGCTGGAAAAGCCTTAAAGCA
Reverse FeuA(I255C)	TGCTTTAAGGCTTTTCCAGCATGGATTTTTCTCTAAATC

Table S2. Steady-state anisotropy values

	Anisotropy	
	Alexa555	Alexa647
Free dye	0.182 ± 0.005	0.133 ± 0.006
FeuA(Q112C/I255C)	0.206 ± 0.007	0.175 ± 0.010

Data correspondents to mean ± s.d. of 3-4 measurements. See the Materials and Methods section for details.

Table S3. Number of analysed molecules

Solution-based smFRET

Condition	Number of analysed molecules
Apo	1572
100 µM FeBB	1362

Surface-based smFRET

Condition	Number of analysed molecules
Apo (5 ms)	459
Apo (100 ms)	50
10 nM FeBB (100 ms)	66
25 nM FeBB (100 ms)	73
25 nM FeBB (5 ms)	60
75 nM FeBB (100 ms)	63
100 µM FeBB (100 ms)	83
100 µM FeBB (4 ms)	94

Surface-based single-fluorophore assay

Condition	Number of analysed molecules
FeuA(Q112C)-Alexa647 + 40 nM FeBB	50
FeuA(I255C)-Alexa647 + 40 nM FeBB	87
FeuA(Q112C)-Alexa555 + 25 nM FeBB	48
FeuA(I255C)-Alexa555 + 25 nM FeBB	50

SUPPORTING REFERENCES

1. Gopich, I. and A. Szabo 2005. Theory of photon statistics in single-molecule forster resonance energy transfer. J. Chem. Phys. 122, 14707.

2. Gopich, I. and A. Szabo 2012. Theory of the energy transfer efficiency and fluorescence lifetime distribution in single-molecule FRET. Proc. Natl. Acad. Sci. U. S. A. 109, 7747-7752.