Supplementary Data

E. coli elongation factor Tu bound to a GTP analogue displays an open conformation equivalent to the GDP-bound form

Jesper S. Johansen, Darius Kavaliauskas, Shawn H. Pfeil, Mickaël Blaise, Barry S. Cooperman,

Yale E. Goldman, Søren S. Thirup and Charlotte R. Knudsen

SUPPLEMENTAARY MATERIALS AND METHODS

Expression and purification of *E. coli* EF-Tu for crystallographic studies. *E. coli* BL21 (DE3) Rosetta cells (Novagen) were incubated in Luria-Bertani (LB) medium for 4.5 h at 37°C with shaking at 200 rpm. The samples were pelleted by centrifugation at 4200 rpm for 45 min at 4°C. The cell pellet was resuspended in 4 volumes of Lysis Buffer (50 mM KCl, 50 mM Hepes pH 7.5 (NaOH), 10 mM $MgCl_2$, 3 mM DTT 0.5 mM PMSF, 1 mM benzamidine, 15 µM GDP, 10% glycerol) and lysed by sonication at 200W for 15 min. The lysate was clarified by centrifugation at 50,000 rpm for 45 min at 4°C and the supernatant was filtered through a 0.8 µm filter. The sample was applied to a 20-ml Source 30Q column equilibrated in 5 column volumes (CV) of Lysis Buffer. The column was washed in 5 CV Lysis Buffer and trapped protein was eluted by running a 100-ml linear gradient from 50–300 mM KCl. Eluted fractions containing EF-Tu were precipitated in 2.8 M ammonium sulphate. The precipitated protein was collected by centrifugation for 20 min at 4000 rpm at 4°C. The protein pellet was isolated from the supernatant and dissolved in buffer Iso-A (1.5 M ammonium sulfate, 50 mM Hepes pH 7.5 (NaOH), 5 mM $MgCl_2$, 1 mM DTT, 1 mM benzamidine, 15 µM GDP and 5% glycerol) and applied to a Source Iso 15 column equilibrated in 5 CV Iso-A buffer. The Source Iso 15 column was washed in 5 CV of buffer Iso-A before trapped protein was eluted on a 100- ml linear gradient from 1.5 M–0.75 M ammonium sulphate. The eluted fractions containing EF-Tu were dialysed against buffer QFF-A (50 mM KCl, 50 mM hepes pH 7.5 (NaOH), 5 mM $MgCl_2$, 1 mM DTT, 1 mM benzamidine, 0.5 mM EDTA and 15 µM GDP) overnight at 4°C. The protein sample was applied to a 5 ml Resource Q column equilibrated in 5 CV of buffer QFF-A. The Resource Q column was washed in 5 CV of buffer QFF-A and EF-Tu was eluted on a 110-ml linear gradient from 50–300 mM KCl. The eluted fractions containing EF-Tu was concentrated to 8 mg/ml in a Vivaspin centrifugal concentrator (Sartorius) and glycerol was added to a final concentration of 20%. Samples were stored at -80 °C.

Nucleotide exchange and crystallisation. Purified EF-Tu was thawed on ice and dialysed against a nucleotide exchange buffer (200 mM ammonium sulphate, 30 mM Tris pH 8.0, 1 mM DTT and 10% glycerol). $MgCl_2$ was added to a final concentration of 1 mM, plus GDPNP at a molecular ratio of 5:1 relative to EF-Tu and 2 units of alkaline phosphatase (Roche) per mg EF-Tu. The solution was incubated for 30 min on ice and clarified by centrifugation for 15 min at 13,000 rpm at 4°C. The EF-Tu·GDPNP complex was purified from the solution by gel filtration on a 24 ml Superdex 75 column (GE Healthcare) equilibrated in GF buffer (30 mM KCl, 20 mM Hepes 7.5 (NaOH), 5 mM $MgCl_2$ and 1 mM DTT). *E. coli* EF-Tu·GDPNP was concentrated to 8 mg/ml prior to crystallisation experiments.

Crystallisation trials were carried out at 293°K utilising the sitting drop vapour diffusion technique. The initial crystallisation conditions were established through commercial crystal screens (Hampton Research and Molecular Dimension) by mixing 1µl EF-Tu·GDPNP (8 mg/ml) with 1 µl reservoir solution. Small bundles of crystals appeared after 1-3 days. To improve the initial crystals the crystallisation buffer was optimised to 2 M ammonium sulphate, 0.1 M Hepes pH 7.5, 2% (v/v) PEG 400 and 1 mM DTT and single needle crystals were obtained by streak-seeding the drops with serial dilutions of crushed crystals.

The streak-seeded crystals reached their maximum size of approximately 5 x 10 x 100 µm within 7-10 days. Single crystals of uniform appearance were isolated from the mother liquor and rapidly transferred to cryoprotection buffer (2 M ammonium sulphate, 0.1 M Hepes 7.5 (NaOH) and 20 % glycerol) and immediately flash frozen. Data were collected at beamline I911-2 (Maxlab, Lund).

Data processing & structure determination. Data processing and scaling was performed with the XDS package (1). Phases were determined by molecular replacement with the Phaser application of Phenix (2). Search models were prepared by extracting domains I (residues 8-200), II (residues 209-293) and III (residues 303-393) from the structure of E. coli EF-Tu bound to GDP (PDB ID 1EFC) (3). The guanine nucleotide and all water molecules were removed and the individual domains were

separated and used in combination as search models. Model building was performed in Coot (4) and the fitted model was refined with Phenix refine (2). The structure geometry, bond lengths and intramolecular non-bonded atomic distances of the final model was validated with Procheck (5) and molprobity (6).

Preparation of EF-Tu for solution FRET studies. Structural transitions between the open and closed conformation of EF-Tu can be detected by introducing a FRET reporter pair in each of the domains I and III, which are predicted to move relative to each other. Sites for introducing Cys residues (labeling sites for maleimide dyes) on the two domains were selected based on the analysis of the following structures: 70S with stalled *Thermus thermophilus* EF-Tu·GDP·kirromycin·aa-tRNA (PDB codes 1WRN, 1WRO), *Escherichia coli* EF-Tu·GDPNP·kirromycin·aa-tRNA (PDB code 1OB2) and EF-Tu·GDP (PDB code 1DG1). Three pairs of amino acid residues, T33/M351, D47/D314 and D165/D314 (Figure S1), were chosen for labeling, and cysteines were introduced at these positions in a variant of EF-Tu, denoted EF-Tu^{AV}, which contains only one native cysteine, as described in the accompanying paper (Kavaliauskas *et al.*). The resulting mutants, denoted EF-Tu$^{AV\text{-}33/351}$, EF-Tu$^{AV\text{-}47/314}$ and EF-Tu$^{AV\text{-}165/314}$, were expressed, purified, labeled and characterized as described (Kavaliauskas *et al.*).

FRET states of dual-labeled EF-Tu observed by single-molecule confocal fluorescence burst detection. Fluorescence from molecules diffusing through a tightly-focused laser spot (514 nm Argon-ion, ~100 μW prior to the excitation dichroic) was collected by a high-N.A. water immersion objective (Nikon 60X, N.A=1.2) mounted on an inverted microscope (Nikon Eclipse TE 300). Fluorescence emission was focused onto a 50 μm confocal pinhole by a tube lens, split by a dichroic mirror and 570 ± 15 nm and 670 ± 15 nm band-pass filters into Cy3 and Cy5 channels, and projected onto the active areas of avalanche photo-diodes (PerkinElmer, NJ).

The GDP conformation of EF-Tu was analyzed in a sample containing 50 pM of dual-labeled EF-Tu$^{AV\text{-}33/351}$ in W buffer (50 mM Tris pH 7.5, 7 mM $MgCl_2$, 30 mM KCl, 70 mM NH_4Cl, 1 mM DTT) containing 1 mM GDP. In the case of the GTP conformation, EF-Tu·GDP was diluted to 50 pM in W buffer containing 1 mM GTP, 3 mM PEP and 0.005 mg/ml pyruvate kinase. Where relevant, 2 μM (final concentration) Phe-tRNA^{Phe} was added into the solution containing 50 pM EF-Tu·GTP. EF-Tu·GDP·kirromycin·Phe-tRNA^{Phe}·70SIC complexes were formed by mixing the EF-Tu·GDP·kirromycin·Phe-tRNA^{Phe} complex with 5-fold excess of 70S initiation complex (70SIC) followed by sedimentation of the mixture through a 1.1 M sucrose cushion by centrifugation at 100k rpm for 12 min. The pellet was dissolved in W buffer supplemented with 30 μM kirromycin and used in burst detection measurements. The 70S initiation complex was prepared as previously described (7). Prior to each experiment, control measurements performed on the buffers yielded burst detection rates of 5% or less compared to EF-Tu samples, and an apparent FRET efficiency confined to E_{App}<0.4.

Detected pulses were time stamped and binned with an integration time of 500 μs (Flex 03-LQ-01 correlator card) into donor (Cy3) and acceptor (Cy5) channels. A threshold was applied to the sum of the two channels for each time bin to distinguish fluorescence bursts from background counts. Adjacent time bins above the threshold were pooled into single events, and the appropriate background was subtracted. The counts on the donor and acceptor channels, n^{Cy3} and n^{Cy5}, were used to calculate apparent FRET efficiencies, $E_{App} = (n^{Cy5} - \chi n^{Cy3}) / (n^{Cy5} - \chi n^{Cy3} + n^{Cy5})$, for each event. The bleed-through of donor photon counts into the acceptor channel, χ, was estimated using data taken on molecules labeled with a donor only. The data were not corrected for differential detector sensitivity or quantum yields, so FRET values obtained from the free-diffusion experiments indicate relative but not absolute distances. Since experimental FRET data takes on values slightly smaller than zero and greater than 1.0 due to noise and photon counting statistics, we transformed the

distributions from the interval [0,1] onto the interval [E_{min}, E_{max}], minimum and maximum FRET values collected.

For dual-labeled EF-Tu$^{AV\text{-}33/351}$, all FRET data, E, were fit by a linear combination three subpopulations modelled by beta distributions (8).

$$P(E) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \cdot E^{(p-1)} \cdot (1-E)^{(q-1)}$$

with amplitudes A_1 - A_3 using maximum likelihood estimation (9). p and q are shape parameters of each beta distribution component. The mean FRET efficiency of the component is given by $E_{App} = p / (p + q)$. The standard deviation width, w, of each component is given by

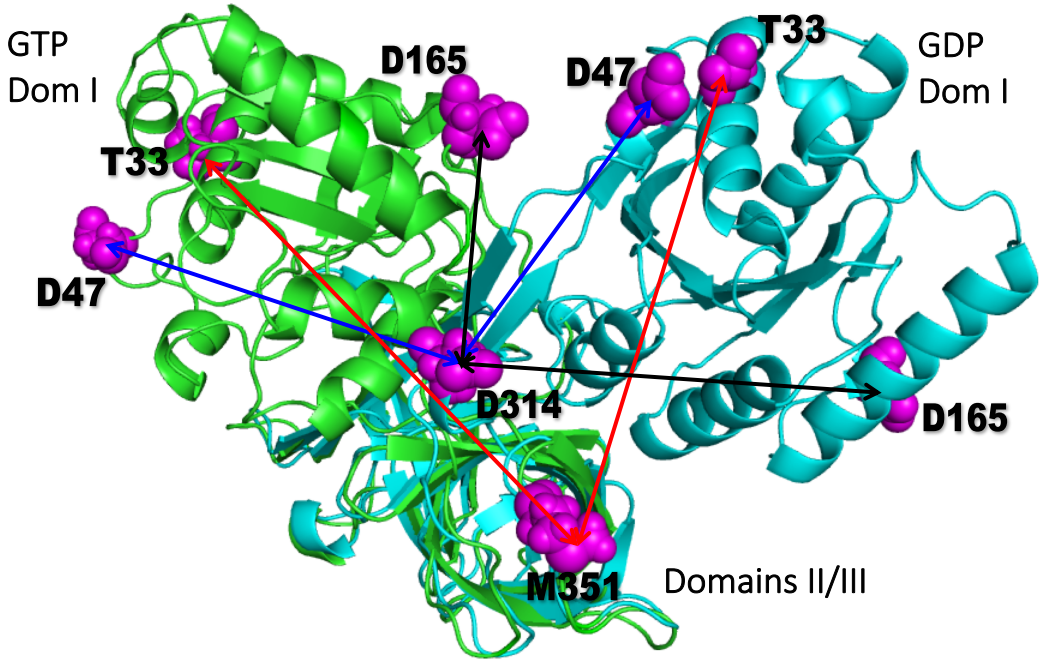
$$w = \sqrt{\int_0^1 P(E)\left(E - \frac{p}{p+q}\right)^2 dE} = \sqrt{\frac{p \cdot q}{(p+q)^2 \cdot (p+q+1)}}$$

When the width, w, of the distribution of FRET values can be explained due solely to photon counting statistics, the average number of photons in a burst $N = p + q$ can be used to constrain the fit (8). Applying this constraint to the data here led to good fits to the donor-only and high-FRET peaks, but gave much too narrow a central peak. Relief of the burst number constraint led to much better fits suggesting that the large width of the central peak is due to actual variability or dynamic disorder in the samples, not just counting statistics. Computer code for this procedure is available upon request.

The majority of samples were fit best by a three-component model including a donor/auto fluorescence peak, intermediate-FRET value, and high-FRET value. As an extra check of the validity of the high-FRET state, control measurements were performed on the EF-Tu mutant Q159C labeled with Cy5 only, which demonstrated that in the absence of FRET, only an auto fluorescence peak is detected.

Uncertainties in fitted parameters were determined by bootstrapping. The measured FRET values were sampled with replacement creating a set of 200 replica data sets. Each data set was fit and the 95% confidence ranges of each fitted parameter calculated (Table S1).

SUPPLEMENTARY FIGURES AND TABLES



Mutant	GTP,nm	GDP,nm	FRET(GTP)	FRET(GDP)
EF-Tu$^{AV-165/314}$	2.2	4.3	0.99	0.71
EF-Tu$^{AV-47/314}$	4.1	2.7	0.77	0.98
EF-Tu$^{AV-33/351}$	5.3	4.2	0.41	0.74

Figure S1. Design of double-cysteine mutants of EF-Tu for dual-labeling with Cy3 and Cy5. The chosen pairs of cysteines are shown in magenta on the structures of the GDP- and GTP-bound forms of EF-Tu (Protein Data Bank entries 1DG1 and 1OB2, respectively) represented in cyan and green, respectively. Domains II and III of the two structures are superimposed to illustrate the expected distance change between the Cα positions of the two cysteines in each pair. The arrows illustrate the expected distances for the three double-mutants EF-Tu$^{AV-165/314}$ (black), EF-Tu$^{AV-47/314}$ (blue) and EF-Tu$^{AV-33/351}$ (red). The table provides the calculated distances in nm along with the corresponding, theoretical FRET values calculated based on R_o=5.0 nm and the assumption that κ^2=2/3, which is standard for relative distance measurements (10). The actual distance changes between the Cy dyes could be different due to linker flexibility. The effector region is disordered in

1OB2, rendering the prediction of distances in the GTP-bound form less trustworthy for EF-Tu$^{AV-33/351}$ and EF-Tu$^{AV-47/314}$.

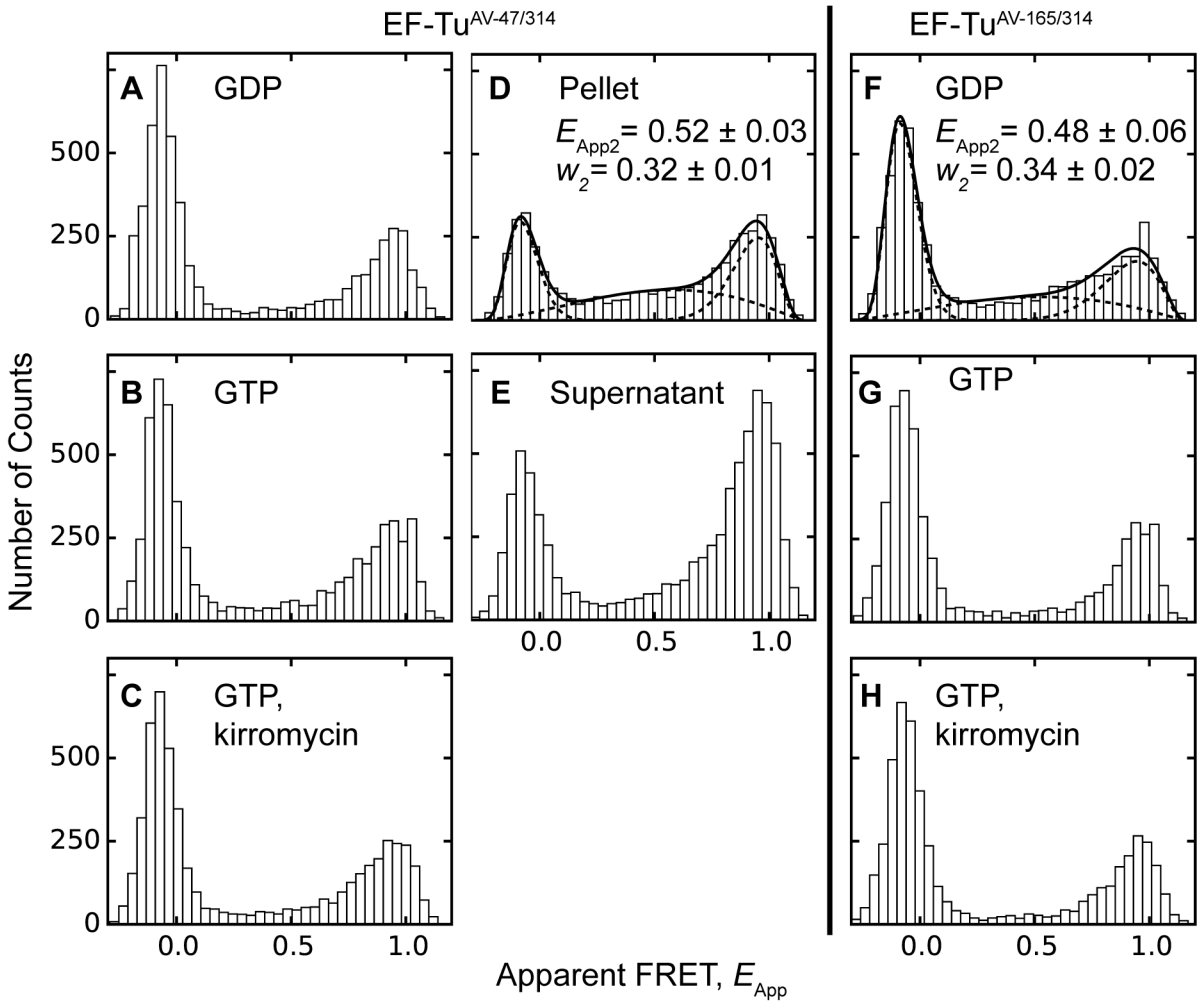


Figure S2. FRET distributions of dual-labeled EF-Tu$^{AV-47/314}$ and EF-Tu$^{AV-165/314}$ on and off the ribosome. Dual-labeled mutants were observed for 10 minutes under a confocal microscope at 50 pM concentration to detect fluorescence bursts for EF-Tu$^{AV-47/314}$ (panels A-E) and EF-Tu$^{AV-165/314}$ (panels F-H) under different conditions: (A and F) 1 mM GDP; (B and G) 1 mM GTP, 3 mM PEP, 0.005 mg/ml pyruvate kinase; (C and H) 1 mM GTP, 3 mM PEP, 0.005 mg/ml pyruvate kinase, 30 µM kirromycin; (D) pellet of EF-Tu$^{AV-47/314}$·GDP·kirromycin·Phe-tRNA^{Phe}·70SIC obtained after ultracentrifugation through a sucrose cushion; (E) supernatant from (D). The common peak below E_{App} = 0.2 is an artefact resulting from molecules with no or inactive acceptor fluorophores.

Only the data in panels D and F could be fit to a model of three beta distributions (same procedure as for the data in figure 3). The apparent FRET value, EApp2, and the width, w2, of the middle distribution are provided along with their S. E. M. uncertainties for these panels.

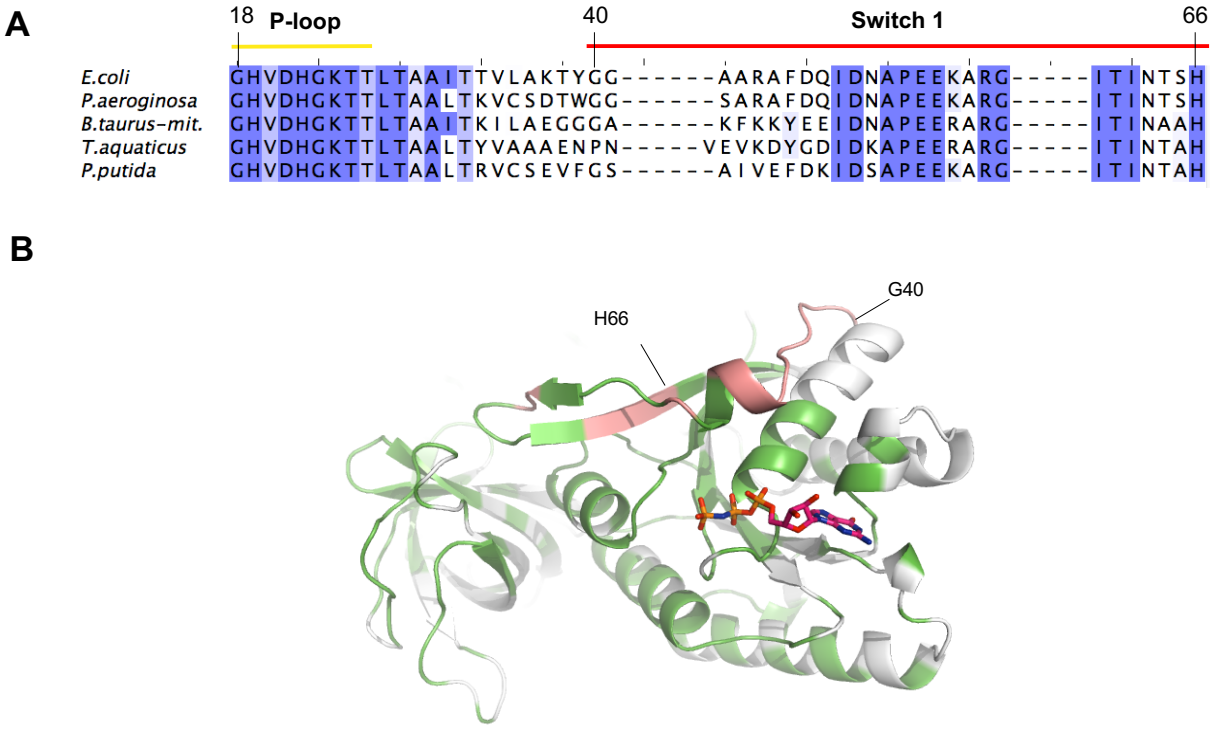


Figure S3. The switch 1 region of EF-Tu is poorly conserved. (A) Sequence alignment of EF-Tu from 897 different species showing the segment covering the P-loop (yellow bar) and the Switch 1 region (red bar). The numbers of select *E. coli* EF-Tu positions are indicated. Only the sequences of the five EF-Tu species with known GDP- or GDPNP-bound structures are shown. Conserved positions in 897 sequences from different species (UNIREF90 data base) are shaded dark blue (100% conserved) to light blue (50% conserved). The alignment was generated using JalView (11). (B) The degree of sequence conservation mapped onto the structure of *E. coli* EF-Tu·GDPNP. Positions that are more than 70% conserved are coloured green. Positions less than 70% conserved are white except in the Switch 1 region (residues 40-66; these boundaries are indicated), where these residues are shown in red. GDPNP is shown as sticks.

Figure [Condition]	A_1 [C.I.]	E_{App1} [C.I.]	w_1 [C.I]	A_2 [C.I.]	E_{App2} [C.I.]	w_2 [C.I.]	A_3 [C.I.]	E_{App3} [C.I.]	w_3 [C.I]	$f = A_2/(A_2+A_3)$ [C.I.]
3A **[GDP]**	2520 [2450,2590]	-0.064 [-0.067, -0.060]	0.081 [0.077,0.085]	2220 [2090,2320]	0.783 [0.769,0.794]	0.186 [0.177,0.193]	350 [260,450]	0.980 [0.971,0.988]	0.034 [0.030,0.041]	0.86 [0.82,0.90]
3B **[GTP]**	2990 [2900,3070]	-0.068 [-0.072,-0.064]	0.077 [0.074-0.081]	2420 [2300,2520]	0.694 [0.678,0.706]	0.224 [0.214,0.232]	700 [617,800]	0.994 [0.988,0.999]	0.043 [0.038,0.047]	0.78 [0.76,0.80]
3C **[TC]**	2700 [2675,2800]	-0.066 [-0.068,-0.063]	0.076 [0.073,0.077]	1460 [1450,1520]	0.609 [0.575,0.640]	0.223 [0.205,0.239]	763 [597,907]	0.965 [0.938,0.981]	0.085 [0.072,0.101]	0.66 [0.62,0.70]
3D **[GTP-kirro]**	2690 [2640,2680]	-0.065 [-0.069,-0.061]	0.081 [0.074,0.086]	1570 [1440,1729]	0.528 [0.507,0.546]	0.208 [0.187,0.234]	940 [810,103]	0.982 [0.966,0.995]	0.070 [0.062,0.082]	0.63 [0.59,0.67]
3E **[GDPNP]**	2800 [2710,2880]	-0.070 [-0.075,-0.064]	0.092 [0.087,0.097]	1980 [1820,2110]	0.672 [0.639,0.697]	0.257 [0.243,0.270]	614 [474,777]	0.977 [0.964,0.990]	0.054 [0.041,0.067]	0.76 [0.72,0.80]
3F **[TC-kirro-70SIC]**	2260 [2170,2340]	-0.064 [-0.069,-0.060]	0.080 [0.076,0.085]	1270 [1160,1450]	0.469 [0.448,0.493]	0.203 [0.179,0.229]	1220 [1080,1310]	0.961 [0.949,0.974]	0.089 [0.079,0.099]	0.51 [0.47,0.55]
3G **[Pellet from 3F]**	3060 [2970,3146]	-0.028 [-0.034,-0.024]	0.081 [0.077,0.085]	2180 [2080,2290]	0.446 [0.435,0.455]	0.154 [0.146,0.163]	369 [320,409]	0.943 [0.920,0.960]	0.107 [0.093,0.124]	0.855 [0.839,0.871]
3H **[Suprnt from 3F]**	2770 [2420,3120]	-0.048 [-0.052,-0.044]	0.086 [0.081,0.091]	3040 [2870,3210]	0.669 [0.647,0.689]	0.253 [0.244,0.267]	1890 [1710,2070]	0.985 [0.979,0.991]	0.060 [0.055,0.066]	0.62 [0.58,0.66]
3I **[Q159C-Cy5]**	661 [492,743]	-0.025 [-0.051,-0.007]	0.109 [0.090,0.121]	111 [29,269]	0.245 [0.096,0.409]	0.156 [0.083,0.205]	4 [0,9]	0.937 [0.854,0.981]	0.047 [0.025,0.099]	**X**

Table S1: Confocal fluorescence burst analysis of freely diffusing, dual-labeled EF-Tu$^{AV\text{-}33/351}$. Fluorescence bursts of freely-diffusing dual-labeled EF-Tu$^{AV\text{-}33/351}$ were detected at different conditions: (A) GDP; (B) GTP, PEP, and pyruvate kinase; (C) Phe-tRNA^{Phe}, GTP, PEP, and pyruvate kinase; (D) GTP, kirromycin, PEP, and pyruvate kinase; (E) GDPNP; (F) Phe-tRNA^{Phe}, 70SIC, kirromycin, GTP, PEP, and pyruvate kinase; (G) pellet of EF-Tu$^{AV\text{-}33/351}$·GDP·kirromycin·Phe-tRNA^{Phe}·70SIC obtained after ultracentrifugation through a sucrose cushion; (H) supernatant resulting after ultracentrifugation of EF-Tu$^{AV\text{-}33/351}$·GDP·kirromycin·Phe-tRNA^{Phe}·70SIC through a sucrose cushion. See legend to Figure S2 for reagent concentrations. (I) is a negative control containing the Cy5-labeled EF-Tu mutant Q159C. The distribution of apparent FRET values was fitted by a weighted sum of three beta distributions to obtain the mean apparent FRET values, E_{App1}-E_{App3}, total counts for each population, A_1-A_3, and beta component widths, w_1-w_3, as described in the text. Bootstrapping, with 200 replicas, was used to obtain 95% confidence intervals, [C.I.], for each fit parameter. We also report $f = \frac{A_2}{A_2+A_3}$, the ratio of events in the mid-FRET peak to the total number of events in the mid- and high-FRET peaks.

Figure	**Condition**	3A	3B	3C	3D	3E	3F	3G	3H
3A	GDP								
3B	GTP	<0.001 / <0.001							
3C	TC	<0.001 / <0.001	<0.001 / n.s.						
3D	GTP-kirro	<0.001 / <0.1	<0.001 / n.s.	<0.001 / n.s.					
3E	GDPNP	<0.001 / <0.001	n.s. / <0.001	<0.01 / <0.01	<0.001 / <0.001				
3F	TC-kirro-70SIC	<0.001 / n.s.	<0.001 / n.s.	<0.001 / n.s.	<0.001 / n.s.	<0.001 / <0.001			
3G	Pellet from F	<0.001 / <0.001	<0.001 / <0.001	<0.001 / <0.001	<0.001 / <0.001	<0.001 / <0.001	<0.1 / <0.001		
3H	Supernatant from F	<0.001 / <0.001	<0.1 / <0.001	<0.01 / <0.01	<0.001 / <0.001	n.s. / n.s.	<0.001 / <0.001	<0.001 / <0.001	

Table S2. Statistical comparison of E_{App2} and w_2 values for freely diffusing dual-labeled EF-Tu$^{AV\text{-}33/351}$. A Student's t test was performed to determine if the fitted values of the intermediate FRET peak, E_{App2} and w_2, suggested significantly different populations. Where relevant, the resulting statistical significance is provided as P-values, while "n.s." means that no significant difference was observed between the compared populations. P-values for E_{App2} and w_2 are provided above and below the diagonal in each entry, respectively.

SUPPLEMENTARY REFERENCES

1. Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography*, **26**, 795-800.
2. Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K. and Terwilliger, T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta crystallographica. Section D, Biological crystallography*, **58**, 1948-1954.
3. Song, H., Parsons, M., Rowsell, S., Leonard, G. and Philips, S. (1999) Crystal structure of intact elongation factor EF-Tu from Escherichia coli in GDP conformation at 2.05 Å resolution. *J Mol Biol*, **285**, 1245-1256.
4. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography*, **60**, 2126-2132.
5. Laskowski, R., Macarthur, M., Moss, D. and Thornton, J. (1993) Procheck - a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, **26**, 283-291.
6. Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S. *et al.* (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res*, **35**, W375-383.
7. Chen, C., Stevens, B., Kaur, J., Cabral, D., Liu, H., Wang, Y., Zhang, H., Rosenblum, G., Smilansky, Z., Goldman, Y.E. *et al.* (2011) Single-molecule fluorescence measurements of ribosomal translocation dynamics. *Mol Cell*, **42**, 367-377.
8. Gopich, I.V. and Szabo, A. (2006) Theory of the statistics of kinetic transitions with application to single-molecule enzyme catalysis. *J Chem Phys*, **124**, 154712.
9. Woody, M.S., Lewis, J.H., Greenberg, M.J., Goldman, Y.E. and Ostap, E.M. (2016) MEMLET: An Easy-to-Use Tool for Data Fitting and Model Comparison Using Maximum-Likelihood Estimation. *Biophys J*, **111**, 273-282.
10. Dale, R., Eisinger, J. and Blumberg, W. (1979) The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer. *Biophys J*, **26**, 161-193.
11. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. and Barton, G.J. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics (Oxford, England)*, **25**, 1189-1191.