SUPPLEMENTARY INFORMATION

Conformational selection of translation initiation factor 3 signals proper substrate selection

Margaret M. Elvekrog[1,2] and Ruben L. Gonzalez, Jr.[1]

[1]*Department of Chemistry, Columbia University, New York, NY, USA*
[2]*Present address: Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA, USA*

Correspondence should be addressed to R.L.G. (rlg2118@columbia.edu)

SUPPLEMENTARY FIGURES

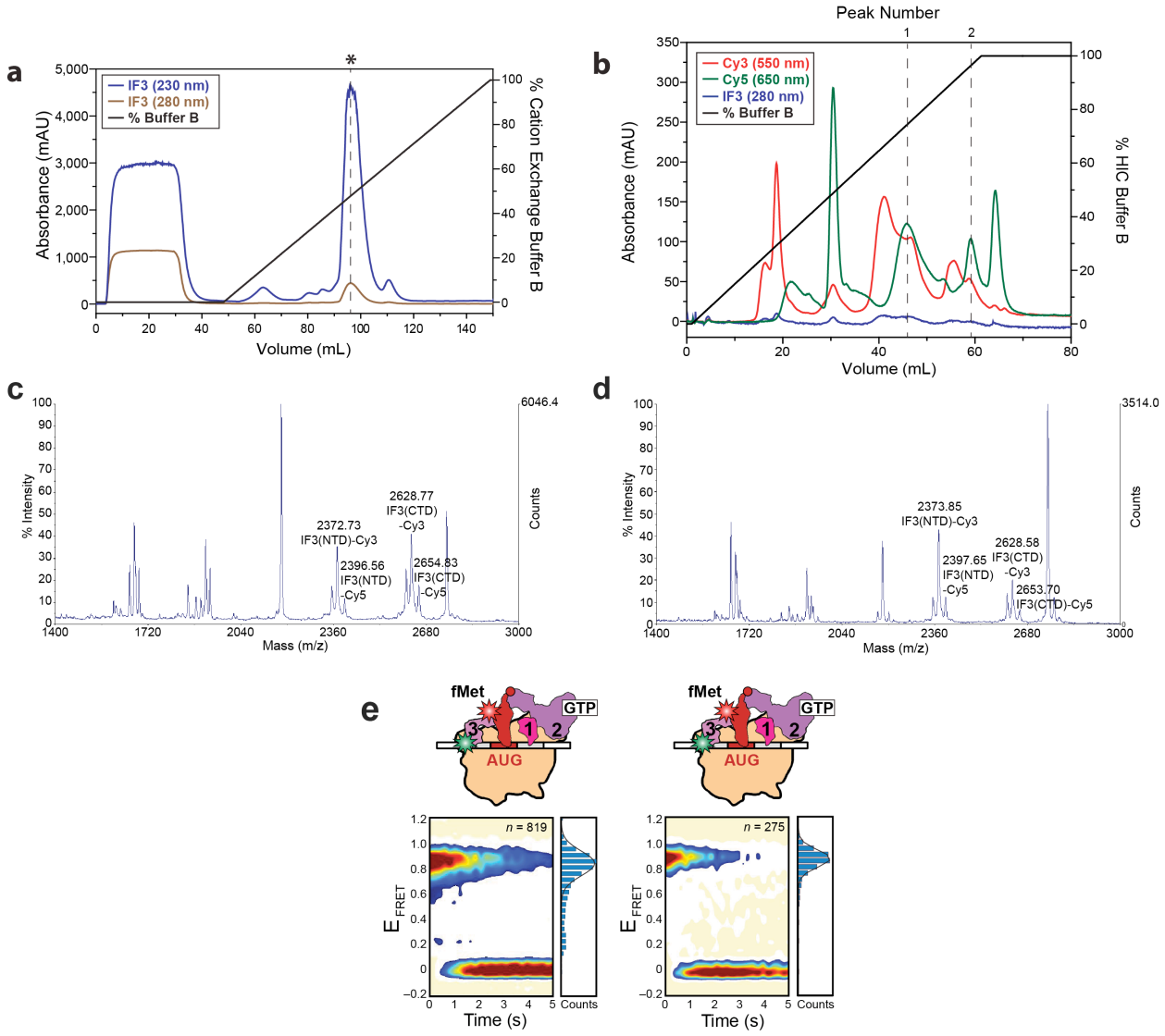

Figure 1. Purification of IF3$^{C65S\ S38C\ K97C}$ and Cy3- and Cy5-labeled IF3$^{C65S\ S38C\ K97C}$. (a) Purification of IF3$^{C65S\ S38C\ K97C}$ using cation exchange chromatography (HiTrap SP HP Sepharose, GE Healthcare). The major peak (denoted with an asterisk) represents pure IF3$_{C65S\ S38C\ K97C}$, as confirmed by Tris-tricine-SDS polyacrylamide gel electrophoresis (data not shown). **(b)** Purification of Cy3- and Cy5-labeled IF3$^{C65S\ S38C\ K97C}$ using hydrophobic interaction chromatography (HIC) (TSKgel Phenyl-5PW, Tosoh Bioscience). The peaks denoted 1 and 2 in the chromatogram both contain IF3$^{C65S\ S38C\ K97C}$(Cy3-Cy5), hereafter IF3(Cy3-Cy5), as confirmed by MALDI-TOF mass spectrometry of tryptic digests (**c-d**) and smFRET imaging (**e**) of the sample fractions collected from under peaks 1 and 2. (**c-d**) MALDI-TOF mass spectrometry of Cy3- and Cy5-labeled, HIC-purified, and trypsin-digested IF3$^{C65S\ S38C\ K97C}$. MALDI-TOF mass spectrometry of tryptic digests of the sample fractions collected from under peaks **(a)** 1 and **(b)**

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

2 of the chromatogram resulting from the HIC purification of Cy3- and Cy5-labeled IF3$^{C65S\ S38C\ K97C}$ (See **b**) Trypsin digestion and mass spectrometry analysis was carried out as described in the Supplementary Note. The peaks in the mass spectra labeled IF3(NTD)-Cy3 and IF3(NTD)-Cy5 have average masses corresponding to tryptic peptides containing the Cy3- and Cy5-labeled N-terminal domain (amino acid residues 27-41) of IF3$^{C65S\ S38C\ K97C}$, respectively. Similarly, the peaks in the mass spectra labeled IF3(CTD)-Cy3 and IF3(CTD)-Cy5 have average masses corresponding to tryptic peptides containing the Cy3- and Cy5-labeled C-terminal domain (amino acid residues 96-111) of IF3$^{C65S\ S38C\ K97C}$, respectively. Comparison of the two mass spectra indicates that the sample fractions collected from under peaks 1 and 2 of the HIC chromatogram both contain IF3(Cy3-Cy5) that is composed of a mixture of IF3$^{C65S\ S38C}$(Cy3)^{K97C}(Cy5) and IF3$^{C65S\ S38C}$(Cy5)^{K97C}(Cy3). **(e)** smFRET measurements of 30S IC^{fMet} prepared using Cy3- and Cy5-labeled and HIC-purified IF3$^{C65S\ S38C\ K97C}$. Two-dimensional surface contour plots of the time evolution of population FRET for 30S IC^{fMet} prepared using the sample fractions collected from under peaks 1 (left panel) and 2 (right panel) of the chromatogram resulting from the HIC purification of Cy3- and Cy5-labeled IF3$^{C65S\ S38C\ K97C}$. "*n*" represents the total number of E_{FRET} *versus* time trajectories that were used to construct each contour plot. To the right of each contour plot is a normalized one-dimensional E_{FRET} histogram for the first 0.5 seconds of the E_{FRET} *versus* time trajectories comprising each dataset. Each histogram was fit with one, two, or three Gaussian distributions (black lines) using Origin 7.0 or 8.0 (OriginLab Corporation), with initial estimates for the mean of each Gaussian distribution obtained from visual inspection of the E_{FRET} distribution peak centers. The cartoon above each contour plot and E_{FRET} histogram depicts the composition of the 30S IC corresponding to the contour plot and E_{FRET} histogram. Comparison of the two sets of smFRET measurements reveals that the sample fractions collected from under peaks 1 and 2 of the HIC chromatogram both contain IF3(Cy3-Cy5) with very similar thermodynamic and kinetic properties. Unless otherwise noted, all experiments described in this article were performed using the sample fraction collected from under peak 1 of the HIC chromatogram.

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

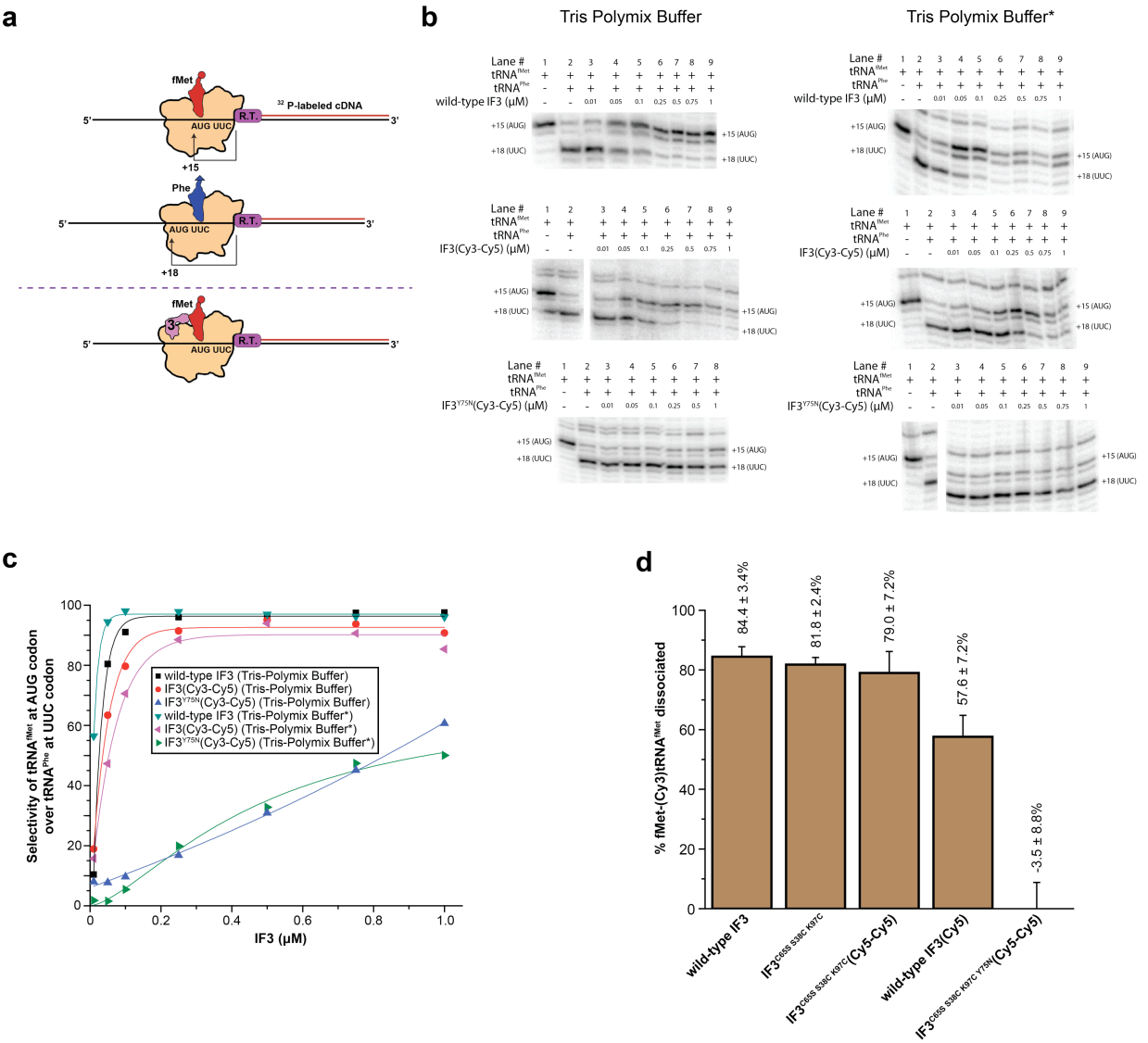

Figure 2. Biochemical activity assays of mutagenized and fluorescently labeled IF3 variants. (a-c) Primer extension inhibition, or toeprinting, activity assay to test IF3's ability to promote selection of $tRNA^{fMet}$ over $tRNA^{Phe}$ on 30S subunits. **(a)** Cartoon depicting the principles of the assay. "R.T." is reverse transcriptase. The 30S subunit exhibits a strong preference for $tRNA^{fMet}$ over $tRNA^{Phe}$ only in the presence of IF3. Reactions were performed as described in Fei et al.[1], Hartz et al.[2], and Maar et al.[3]. **(b)** Lane 1 in each row of gels contains $tRNA^{fMet}$ in the absence of IF3; the strong stop at the +15 mRNA nucleotide position (where the "A" in the AUG start codon is nucleotide position +1) indicates that $tRNA^{fMet}$ is bound at the AUG start codon in the P site of the 30S subunit. Lane 2 in each row of gels contains $tRNA^{fMet}$, and $tRNA^{Phe}$ (in a 10-fold excess over $tRNA^{fMet}$) in the absence of IF3; the strong stop at nucleotide position +18 indicates that $tRNA^{Phe}$ is bound at the UUC codon in the P site of the 30S subunit. Increasing concentrations of wild-type IF3 (Lanes 3-9, top row of gels) shift the preference of the 30S

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

subunit from $tRNA^{Phe}$ at the UUC codon (+18) to $tRNA^{fMet}$ at the AUG codon (+15). Similar behavior is seen with increasing concentrations of unlabeled IF3(Cy3-Cy5) (second row of gels). Consistent with previous work on $IF3^{Y75N}$, the $IF3^{Y75N}$(Cy3-Cy5) negative control (third row of gels) exhibits a significantly reduced ability to promote selection of $tRNA^{fMet}$ over $tRNA^{Phe}$ on 30S subunits relative to wild-type IF3 or IF3(Cy3-Cy5) (ref. 3). **(c)** Plot of the selectivity of $tRNA^{fMet}$ at AUG codon over $tRNA^{Phe}$ at UUC codon for each IF3 variant tested. The selectivity of $tRNA^{fMet}$ at AUG codon over $tRNA^{Phe}$ at UUC codon was defined as described in the Supplementary Note. **(d)** TIRF microscopy-based activity assay to test IF3's ability to promote dissociation of tRNAs from the P site of the 30S subunit. See the Supplementary Note for a description of the assay. Error bars represent the standard deviation of the mean obtained from three independent datasets. As can be seen from the bar graph, each 30S IC_{-2} exhibits varying degrees of fMet-(Cy3)$tRNA^{fMet}$ dissociation from the P site of the 30S IC_{-2} depending on the identity of the IF3 that was used to prepare the 30S IC_{-2}. $IF3^{C65S\ S38C\ K97C}$ and $IF3^{C65S\ S38C\ K97C}$(Cy5-Cy5) exhibit fMet-(Cy3)$tRNA^{fMet}$ dissociation activity that is very similar to that of wild-type IF3; wild-type IF3-Cy5 (labeled at its native Cys65 residue) exhibited fMet-(Cy3)$tRNA^{fMet}$ dissociation activity that was slightly impaired relative to wild-type IF3; and $IF3^{C65S\ S38C\ K97C\ Y75N}$(Cy5-Cy5) exhibited no fMet-(Cy3)$tRNA^{fMet}$ dissociation activity relative to wild-type IF3, consistent with previous work demonstrating that, although it does not affect the binding of IF3 to 30S subunits, the Y75N mutation does affect IF3's ability to promote fMet-$tRNA^{fMet}$ selection[3,4].

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

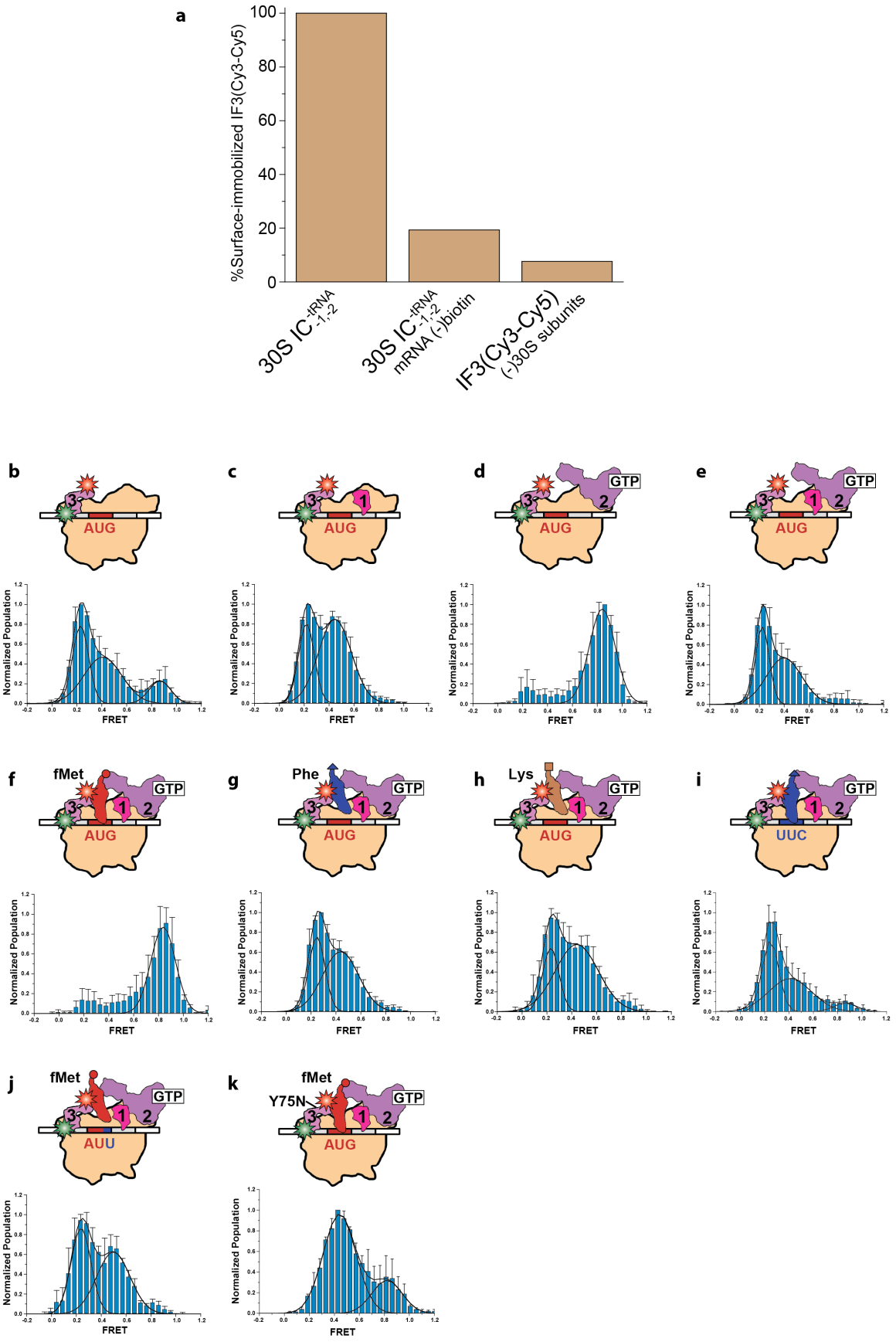

Figure 3. The specificity of IF3(Cy3-Cy5) binding to surface-immobilized 30S ICs and the

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

day-to-day reproducibility of smFRET results. (a) Specificity of 30S $IC^{-tRNA}_{-1/-2}$-bound IF3(Cy3-Cy5) and free IF3(Cy3-Cy5) binding to the surface of microfluidic flowcells. A flowcell-to-flowcell comparison within the same microfluidic device was performed to investigate the specificity of: (i) 30S $IC^{-tRNA}_{-1/-2}$-bound IF3(Cy3-Cy5) prepared using a 5'-biotinylated mRNA, (ii) 30S $IC^{-tRNA}_{-1/-2}$-bound IF3(Cy3-Cy5) prepared using a non-biotinylated mRNA, and (iii) a free IF3(Cy3-Cy5) solution prepared in the absence of 30S subunits or any other initiation components binding to the polyethylene (PEG)- and biotinylated-PEG-passivated and streptavidin-derivatized surface of the flowcells. The two 30S $IC^{-tRNA}_{-1/-2}$ samples and the free IF3(Cy3-Cy5) sample were prepared in parallel and loaded into separate flowcells on the same microfludic device such that the concentration of IF3(Cy3-Cy5) that was introduced into each flowcell was the same for all three samples. The flowcells were incubated at room temperature for 5 min, washed using Tris-Polymix Buffer containing an enzymatic oxygen scavenger system and a triplet state quencher cocktail, and imaged by recording 10 randomly selected fields-of-view from each flowcell using TIRF microscopy. The % surface-localized IF3(Cy3-Cy5) for the 30S $IC^{-tRNA}_{-1/-2}$ sample and the free IF3(Cy3-Cy5) sample were quantified by counting the average number of diffraction-limited Cy3 spots observed over the 10 fields-of-view of each sample and normalizing the resulting average number of Cy3 spots to the average number of Cy3 spots observed over 10 fields-of-view of 30S $IC^{-tRNA}_{-1/-2}$ prepared using a 5'-biotinylated mRNA (assuming that the average number of Cy3 spots observed over 10 fields-of-view of 30S $IC^{-tRNA}_{-1/-2}$ prepared using a 5'-biotinylated mRNA represents 100% surface-localization of IF3(Cy3-Cy5)). As can be seen from the bar graph, 30S $IC^{-tRNA}_{-1/-2}$ containing a non-biotinylated mRNA (middle bar) and free IF3(Cy3-Cy5) (right bar) exhibit 20% and 5% localization to the surface of the flowcell relative to 30S $IC^{-tRNA}_{-1/-2}$ containing a biotinylated mRNA (left bar), respectively. We conclude that, under our experimental conditions, 80-95% of individual, surface-localized IF3(Cy3-Cy5)s were bound to the flowcell surface *via* their interaction with a 30S IC carrying a biotinylated mRNA. **(b–k)** Reproducibility of E_{FRET} histograms obtained from the analysis of three independently collected datasets for each 30S IC. Three independent datasets consisting of 12-15 movies each were collected on separate days using independently prepared samples and microfluidic devices for each 30S IC. Each blue bin represents the mean population at that range of E_{FRET} values from the three independently collected datasets and the error bars on each blue bin represents the standard deviation from the mean population. The bin size in each histogram was set to 0.047. The histograms were fit with one, two, or three Gaussian distributions (black lines), with initial estimates for the mean of each Gaussian distribution obtained from visual inspection of the E_{FRET} distribution peak centers. The cartoon above each E_{FRET} histogram depicts the composition of the 30S IC corresponding to the E_{FRET} histogram: **(b)** 30S $IC^{-tRNA}_{-1/-2}$, **(c)** 30S IC^{-tRNA}_{-2}, **(d)** 30S IC^{-tRNA}_{-1}, **(e)** 30S IC^{-tRNA}, **(f)** 30S IC^{fMet}, **(g)** 30S IC^{Phe}, **(h)** 30S IC^{Lys}, **(i)** 30S $IC^{Phe,UUC}$, **(j)** 30S $IC^{fMet,AUU}$ and **(k)** 30S $(IF3_{Y75N})$ IC^{fMet}. Analysis of the means and standard deviations of these E_{FRET} histograms demonstrates that the reproducibility of smFRET measurements for each 30S IC was excellent.

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

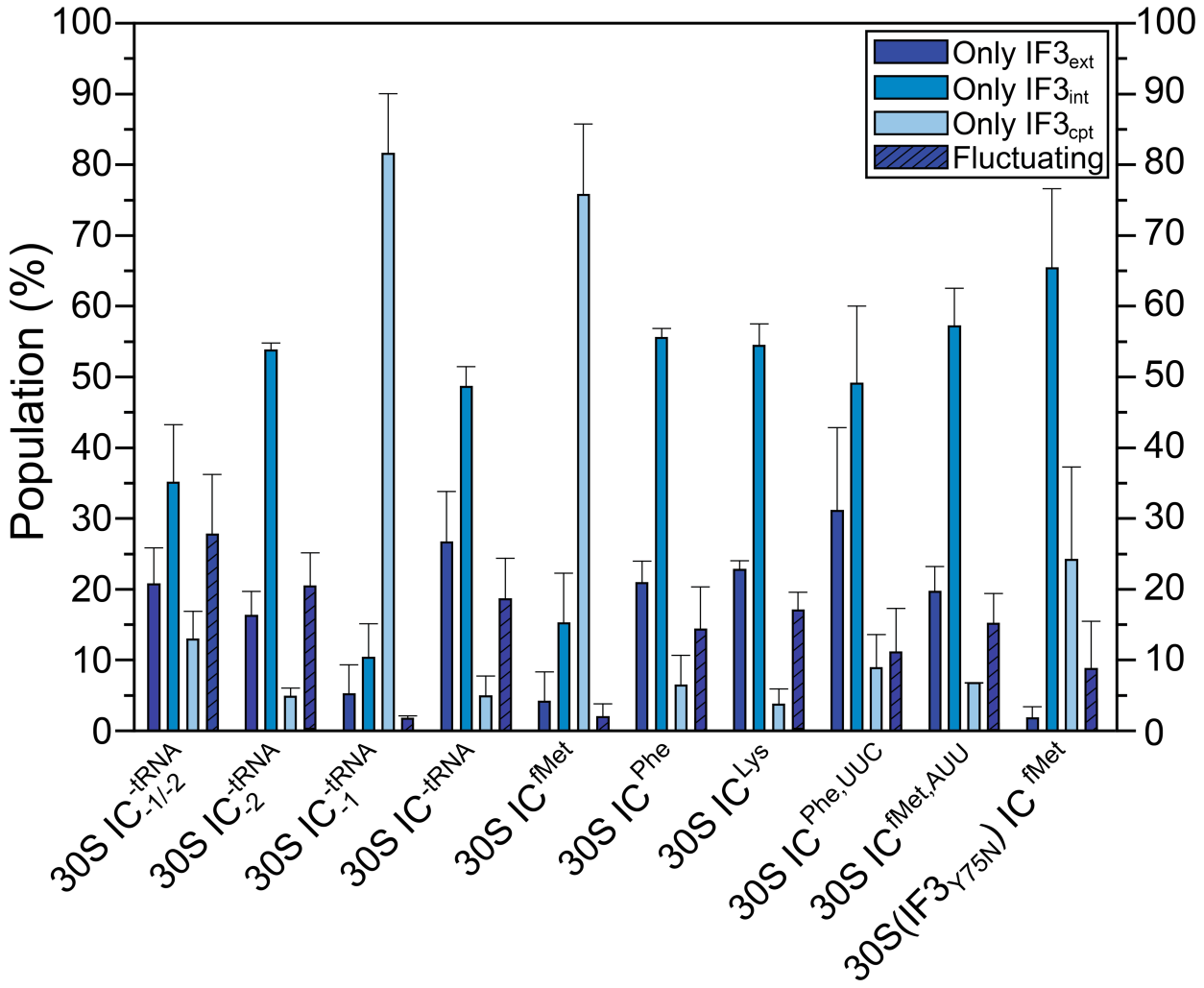

Figure 4. Classification of E_{FRET} *versus* time trajectories. The E_{FRET} axes of the idealized E_{FRET} *versus* time trajectories obtained from hidden Markov modeling of the raw E_{FRET} *versus* time trajectories (see Supplementary Methods) acquired from all of the 30S ICs reported in this work were divided into a single set of E_{FRET} value ranges corresponding to the $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ conformations of IF3. Thresholds to define the E_{FRET} value ranges corresponding to $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ were determined globally by combining all of the idealized E_{FRET} *versus* time trajectories from all three independently collected datasets of all ten 30S ICs reported in this work (see Supplementary Fig. 3) into a global E_{FRET} histogram and fitting this global E_{FRET} histogram with three Gaussian distributions using Origin 7.0 or 8.0 (OriginLab Corporation), with initial estimates for the mean of each Gaussian distribution obtained from visual inspection of the E_{FRET} distribution peak centers. The window of high and low E_{FRET} efficiencies defining each IF3 conformational state were defined using the center and full width at half maximum (FWHM) of each Gaussian distribution. Using this approach, E_{FRET} values between 0.10–0.29 were assigned to the $IF3_{ext}$ conformation of IF3, E_{FRET} values between 0.30–0.69 were assigned to the $IF3_{int}$ conformation of IF3, and E_{FRET} values between 0.70–1.0 were assigned to the $IF3_{cpt}$

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

conformation of IF3. Using these E_{FRET} value ranges for each IF3 conformation, individual E_{FRET} *versus* time trajectories were visually inspected and classified according to whether they: (i) sampled only the $IF3_{ext}$ conformation of IF3 prior to photobleaching; (ii) sampled only $IF3_{int}$ conformation of IF3 prior to photobleaching; (iii) sampled only the $IF3_{cpt}$ conformation of IF3 prior to photobleaching; or (iv) fluctuated between at least two conformations of IF3 prior to photobleaching. As can be seen from the bar graph, relative to $30S\,IC_{-1/-2}^{-tRNA}$, $30S\,IC_{-1}^{-tRNA}$ exhibits a dramatic increase in the number E_{FRET} *versus* time trajectories that sample only the $IF3_{cpt}$ conformation of IF3 prior to photobleaching, consistent with the fact that, relative to $30S\,IC_{-1/-2}^{-tRNA}$, $30S\,IC_{-1}^{-tRNA}$ exhibits a strong shift in the conformational equilibrium of IF3 towards $IF3_{cpt}$ (Fig 2b). Likewise, relative to 30S IC^{-tRNA}, 30S IC^{fMet} exhibits a dramatic increase in the number E_{FRET} *versus* time trajectories that sample only the $IF3_{cpt}$ conformation of IF3 prior to photobleaching, consistent with the fact that, relative to 30S IC^{-tRNA}, 30S IC^{fMet} exhibits a strong shift in the conformational equilibrium of IF3 towards $IF3_{cpt}$ (Fig 3).

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

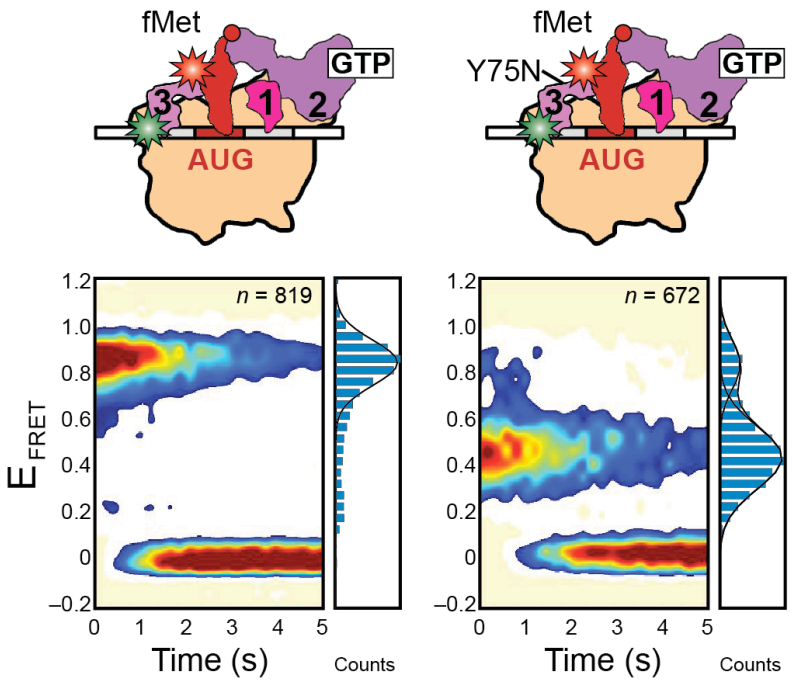

Figure 5. smFRET measurements of 30S IC^{fMet} and 30S ($IF3_{Y75N}$) IC^{fMet}. Two-dimensional surface contour plots of the time evolution of population FRET for 30S IC^{fMet} and 30S ($IF3_{Y75N}$) IC^{fMet}. "*n*" represents the total number of E_{FRET} *versus* time trajectories that were used to construct each contour plot. To the right of each contour plot is a normalized one-dimensional E_{FRET} histogram for the first 0.5 seconds of the E_{FRET} *versus* time trajectories comprising each dataset. Each histogram was fit with one, two, or three Gaussian distributions (black lines) using Origin 7.0 or 8.0 (OriginLab Corporation), with initial estimates for the mean of each Gaussian distribution obtained from visual inspection of the E_{FRET} distribution peak centers. The cartoon above each contour plot and E_{FRET} histogram depicts the composition of the 30S IC corresponding to the contour plot and E_{FRET} histogram. Comparison of the two sets of smFRET measurements reveals that a single amino acid substitution mutation to the interdomain linker of IF3, Y75N, alters the conformational equilibrium of IF3.

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

SUPPLEMENTARY TABLES

Table 1. Means of the Gaussian distributions defining the FRET states that correspond to the $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ conformations of IF3 in each 30S IC. The E_{FRET} values for the first 0.5 sec of the raw E_{FRET} *versus* time trajectories obtained from each of the three independently collected datasets for each 30S IC were used to generate a one-dimensional E_{FRET} histogram. Each E_{FRET} histogram was fit with one, two, or three Gaussian distributions to define the FRET states corresponding to the $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ conformations of IF3 using Origin 7.0 or 8.0 (OriginLab Corporation), with initial estimates for the Gaussian means obtained from visual inspection of the E_{FRET} distribution peak centers. The Gaussian means from the E_{FRET} histograms obtained from each of the three independently collected datasets for each 30S IC were used to calculate the average and standard deviation of the mean of the Gaussian distribution defining the FRET states that correspond to the $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ conformations of IF3 in each 30S IC.

30S IC	$IF3_{ext}$ (E_{FRET} values)	$IF3_{int}$ (E_{FRET} values)	$IF3_{cpt}$ (E_{FRET} values)
30S $IC_{-1/-2}^{-tRNA}$	0.23 ± 0.01	0.42 ± 0.01*	0.87 ± 0.01
30S IC_{-2}^{-tRNA}	0.22 ± 0.01*	0.45 ± 0.03	0.74 ± 0.06
30S IC_{-1}^{-tRNA}	N.D.	N.D.	0.84 ± 0.02
30S IC^{-tRNA}	0.23 ± 0.01	0.42 ± 0.01	0.82 ± 0.06
30S IC^{fMet}	N.D.	N.D.	0.84 ± 0.04
30S IC^{Phe}	0.25 ± 0.02	0.45 ± 0.03	N.D.
30S IC^{Lys}	0.23 ± 0.01	0.47 ± 0.04	N.D.
30S $IC^{Phe,UUC}$	0.25 ± 0.03	0.45 ± 0.05	N.D.
30S $IC^{fMet,AUU}$	0.24 ± 0.01*	0.47 ± 0.04	N.D.
30S ($IF3_{Y75N}$) IC^{fMet}	N.D.	0.43 ± 0.01	0.76 ± 0.1

N.D., could not be determined
*, 0.01 represents an upper limit on the standard deviation

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

Table 2. Equilibrium fractional occupancies of the $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ conformations of IF3. The E_{FRET} value ranges assigned to the $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ conformations of IF3 were used to divide each idealized E_{FRET} *versus* time trajectory into regions corresponding to the $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ conformations of IF3. To calculate the fractional occupancies of $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ in each dataset, the total number of time points spent in each of the three conformational states of IF3 summed over all of the idealized E_{FRET} *versus* time trajectories in a dataset was divided by the total number of time points summed over all of the idealized E_{FRET} *versus* time trajectories in the dataset. The means and standard deviations of the fractional occupancies of the $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ conformations of IF3 for each 30S IC were calculated using the fractional occupancies obtained from the analysis of the three independently collected datasets for each 30S IC.

30S IC	$IF3_{ext}$ (%)	$IF3_{int}$ (%)	$IF3_{cpt}$ (%)
30S $IC_{-1/-2}^{-tRNA}$	54 ± 9	40 ± 10	6 ± 2
30S IC_{-2}^{-tRNA}	45 ± 3	52 ± 3	3 ± 1
30S IC_{-1}^{-tRNA}	23 ± 17	11 ± 4	66 ± 17
30S IC^{-tRNA}	56 ± 7	42 ± 6	2 ± 1
30S IC^{fMet}	15 ± 13	17 ± 7	68 ± 17
30S IC^{Phe}	45 ± 3	53 ± 3	2 ± 1
30S IC^{Lys}	48 ± 7	47 ± 6	5 ± 4
30S $IC^{Phe,UUC}$	58 ± 4	37 ± 9	5 ± 5
30S $IC^{fMet,AUU}$	47 ± 5	50 ± 4	3 ± 1
30S ($IF3_{Y75N}$) IC^{fMet}	13 ± 3	70 ± 7	17 ± 6

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

Table 3. Rates of interconversions between the $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ conformations of IF3. (See next page.) The rates of transitions between $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ in each 30S IC that was investigated were much slower than the experimental data acquisition rate (10 s^{-1}). Given this, the transition rates for each 30S IC were estimated by using the idealized E_{FRET} *versus* time trajectories from the three independently collected datasets for each 30S IC to construct three transition probability matrices for each 30S IC and then multiplying the off-diagonal elements of each transition probability matrix by the experimental data acquisition rate (10 s^{-1}) (ref. 5). The individual transition rates obtained from each of the three independently collected datasets were then used to calculate the average and standard deviation of the transition rates observed in each 30S IC.

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

Table 3. Rates of interconversions between the $IF3_{ext}$, $IF3_{int}$, and $IF3_{cpt}$ conformations of IF3.

30S IC	$IF3_{ext \rightarrow int}$ (sec^{-1})	$IF3_{ext \rightarrow cpt}$ (sec^{-1})	$IF3_{int \rightarrow ext}$ (sec^{-1})	$IF3_{int \rightarrow cpt}$ (sec^{-1})	$IF3_{cpt \rightarrow ext}$ (sec^{-1})	$IF3_{cpt \rightarrow int}$ (sec^{-1})
30S $IC^{-tRNA}_{-1/-2}$	0.08 ± 0.01	0.002 ± 0.002	0.11 ± 0.03	0.01 ± 0.01	0.03 ± 0.02	0.07 ± 0.03
30S IC^{-tRNA}_{-2}	0.09 ± 0.001	0.003 ± 0.001	0.07 ± 0.01	0.01 ± 0.002	0.05 ± 0.04	0.1 ± 0.03
30S IC^{-tRNA}_{-1}	0.05 ± 0.04	0.03 ± 0.02	0.09 ± 0.05	0.17 ± 0.05	0.04 ± 0.01	0.09 ± 0.03
30S IC^{-tRNA}	0.06 ± 0.02	0.002 ± 0.001	0.08 ± 0.03	0.003 ± 0.001	0.1 ± 0.2	0.12 ± 0.06
30S IC^{fMet}	0.09 ± 0.04	0.05 ± 0.05	0.04 ± 0.02	0.11 ± 0.04	0.03 ± 0.02	0.10 ± 0.02
30S IC^{Phe}	0.07 ± 0.01	0.002 ± 0.001	0.06 ± 0.008	0.005 ± 0.002	0.09 ± 0.03	0.08 ± 0.03
30S IC^{Lys}	0.06 ± 0.01	0.004 ± 0.004	0.06 ± 0.005	0.01 ± 0.01	0.06 ± 0.05	0.09 ± 0.03
30S $IC^{Phe,UUC}$	0.04 ± 0.01	0.001 ± 0.001	0.06 ± 0.01	0.01 ± 0.01	0.03 ± 0.02	0.10 ± 0.08
30S $IC^{fMet,AUU}$	0.08 ± 0.008	0.002 ± 0.001	0.08 ± 0.01	0.01 ± 0.01	0.05 ± 0.02	0.2 ± 0.1
30S ($IF3_{Y75N}$) IC^{fMet}	0.05 ± 0.06	0.004 ± 0.004	0.02 ± 0.003	0.03 ± 0.02	0.05 ± 0.02	0.2 ± 0.2

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

Table 4. Composition of all buffers used in this study.

Buffer Name	Components
IF3 Labeling Buffer	100 mM Tris-acetate, pH_{25^oC} = 7.0 50 mM Potassium chloride
Cation Exchange Buffer A	40 mM Tris-hydrochloride, pH_{4^oC} = 7.5 30 mM Sodium chloride 40 mM Ammonium chloride 5 mM Magnesium chloride 2 mM 2-mercaptoethanol
Cation Exchange Buffer B	40 mM Tris-hydrochloride, pH_{4^oC} = 7.5 750 mM Sodium chloride 40 mM Ammonium chloride 5 mM Magnesium chloride 2 mM 2-mercaptoethanol
HIC Buffer A	1 M Ammonium sulfate 100 mM Sodium phosphate, dibasic, pH 7.0
HIC Buffer B	100 mM Sodium phosphate, dibasic, pH 7.0
IF3 Storage Buffer	50 mM Potassium chloride 10 mM Tris-acetate, pH_{25^oC}=7.0 10 mM Magnesium acetate 6 mM 2-mercaptoethanol 50% Glycerol
Tris-Polymix Buffer	10 mM Tris-acetate, pH_{25^oC}=7.0 20 mM Potassium chloride 1 mM Ammonium chloride 0.1 mM Calcium acetate 0.1 mM Ethylenediamine tetraacetic acid (EDTA) 6 mM 2-mercaptoethanol 1 mM Putrescine dihydrochloride 0.2 mM Spermidine, free base 1% ß-D-glucose 5 mM Magnesium acetate
Oxygen Scavenging System	165 U/ml Glucose oxidase from *Aspergillus niger* 2170 U/ml Catalase from bovine liver (1% ß-D-glucose – included in Tris-Polymix Buffer)
Triplet State Quencher Cocktail	1 mM 1,3,5,7-cyclooctatetraene (Aldrich) 1 mM *p*-nitrobenzyl alcohol (Fluka)

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

SUPPLEMENTARY NOTE

Reagents

See Supplementary Table 4 for the composition of all buffers used. All reagents were of molecular biology grade or better. Buffers used for protein purifications are described in detail in reference 1. The Tris-Polymix Buffer used in this study differs from that used in our previous work on translation elongation[6] in that it has been optimized to maximize the stability of IF2 binding to the set of 30S ICs that comprise the current study (unpublished data). Inclusion of a 1 μM concentration of IF2 in this IF2-binding-optimized Tris-Polymix Buffer during total internal reflection fluorescence (TIRF) microscopy imaging thus ensures that all of the 30S ICs that were studied using smFRET were saturated with IF2, an important stipulation for interpreting the resulting smFRET data. Using a standard primer extension inhibition assay, we have demonstrated that IF3 retains full fMet-tRNA^{fMet}-selection activity in this IF2-optimized Tris-Polymix Buffer that is indistinguishable from that which is observed in the Tris-Polymix Buffer which we have historically used in our studies of translation elongation[6].

Preparation of ribosomes, translation factors, mRNAs, and tRNAs

E. coli ribosomes and translation factors were purified as previously described[1]. 5'-biotinylated mRNA with a sequence derived from the mRNA encoding gene product 32 from T4 bacteriophage was purchased from Dharmacon, Inc. The sequence of this mRNA was:

5'-Biotin-
CAACCUAAAACUUACACAAAUUAAA**AAGGAAAU***AGAC*AUGUUCAAAGUCGAAAAAUCUACUGCU-3',

where the nucleotides in bold denote the Shine-Dalgarno sequence, the nucleotides in italics denote the spacer between the Shine-Dalgarno sequence and the start codon, and the underlined nucleotides denote the start codon. Two analogous 5'-biotinylated mRNAs were purchased containing the same sequence described above with the exception that the AUG start codon was mutagenized to UUC or AUU. In addition, a non-biotinylated control mRNA was used for the control experiments to test the specificity of 30S IC-bound IF3(Cy3-Cy5) and free IF3(Cy3-Cy5) binding to the flowcell surface (Supplementary Fig. 3) whose sequence was: 5'-GCAACCUAAAACUUACACAGGGCCC**UAAGGAAAU***AAAA*AUGUUUAAA-3', where the nucleotides in bold again denote the Shine-Dalgarno sequence, the nucleotides in italics denote the spacer between the Shine-Dalgarno sequence and the start codon, and the underlined nucleotides denote the start codon. tRNA^{fMet} was purchased from MP Biomedicals and tRNA^{Phe} and tRNA^{Lys} were purchased from Sigma. All tRNAs were aminoacylated and, in the case of tRNA^{fMet}, formylated as previously described[1].

Preparation, fluorescence labeling, and purification of mutagenized and fluorescently labeled IF3 mutants

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

The gene for *E. coli* IF3 was cloned into the pProEx-HTb plasmid vector (Invitrogen), which encodes a six-histidine (6×His) affinity purification tag and a TEV protease cleavage site at the N-terminal end of the gene encoding IF3. Mutagenesis of IF3 in the pProEx-HTb plasmid vector was performed using the QuickChange Site-Directed Mutagenesis System (Stratagene). DNA primers for mutagenesis were designed following the recommendations provided by the QuickChange Site-Directed Mutagenesis System and were purchased from Integrated DNA Technologies. No further purification of the DNA primers was performed. Using this approach, the single Cys65 in wild-type IF3 was mutagenized to Ser, Ser38 in the NTD was mutagenized to Cys, and Lys97 in the CTD was mutagenized to Cys, yielding a triple-mutant IF3 variant ($IF3^{C65S\ S38C\ K97}$). Note that the amino acid numbering for IF3 used in this study is based on wild-type *E. coli* IF3 numbering. Mutations were verified by DNA sequencing of the plasmid purified from an ampicillin-resistant clone (Genewiz). The pProEx-HTb plasmids encoding all of the IF3 variants used in this study were transformed into BL21-DE3 cells for protein overexpression and the overexpressed 6×His-tagged IF3 variants were purified using Ni^{2+}-nitrilotriacetic acid affinity purification, treated with TEV protease to remove the 6×His tags, and further purified using cation exchange chromatography. Further details regarding the cloning, overexpression, and purification of the IF3 variants used in this study can be found in reference 1. The N-terminus of the IF3 variants used in this study consists of a Gly-Ala-Met-Ala-Lys2 sequence, where Gly-Ala-Met-Ala denotes four non-wild-type amino acids resulting from the cloning strategy and Lys2 denotes the beginning of the wild-type *E. coli* IF3 sequence.

To label $IF3^{C65S\ S38C\ K97C}$, the protein was buffer exchanged into IF3 Labeling Buffer (Supplementary Table 4) and incubated with a 10-fold molar excess of tris(2-carboxyethyl)phosphine hydrochloride for one hour at room temperature. An equimolar mixture of Cy3- and Cy5-maleimide (GE Healthcare) dissolved in anhydrous dimethyl sulfoxide (DMSO) was added to the reaction mixture such that the Cy3- and Cy5-maleimide were both in a 20-fold molar excess over IF3 and the reaction contained a final DMSO concentration of ~5% v/v. The reaction was incubated for 1 hour in the dark at room temperature and subsequently incubated overnight in the dark at 4°C. The reaction was quenched by adding 2-mercaptoethanol to a final concentration of 100 mM and any remaining, unreacted Cy3-and Cy5-maleimide was removed by extensive dialysis at 4°C in HIC Buffer A (Supplementary Table 4).

The unlabeled, mono-labeled, and dual-labeled $IF3^{C65S\ S38C\ K97C}$ products were separated using a TSKgel Phenyl-5PW hydrophobic interaction chromatography (HIC) column (Tosoh Bioscience) that had been pre-equilibrated with HIC Buffer A (Supplementary Table 4). A 0-100% linear gradient of HIC Buffer B applied over 20 column volumes enabled separation of the various unlabeled and labeled IF3 species (Supplementary Fig. 1). Incomplete separation of the mono- and dual-labeled species resulted in overlapping chromatographic peaks and some heterogeneity in the dual Cy3-Cy5 labeled IF3 sample; however, only those IF3s labeled with both Cy3 and Cy5 are able to undergo FRET. Thus, further “purification” occurs at the level of smFRET data acquisition and processing.

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

Mass spectrometric analysis of IF3(Cy3-Cy5)

MALDI-TOF mass spectrometry of trypic digests of IF3(Cy3-Cy5) was performed at the Columbia University Medical Center Protein Core Facility using the following protocol: 1 μL of an approximately 2 μg μL^{-1} solution of the sample fraction collected from under peak 1 or 2 of the HIC chromatogram was diluted to a final volume of 10 μL using 25 mM Tris-HCl ($pH_{4°C}$ = 8.5). Subsequently, 0.08 μg of trypsin (modified sequencing grade, Roche Applied Science) was added to the sample fraction and the digestion reaction was incubated overnight at 32 ºC. The digestion reaction was then desalted using a Millipore C18 ZipTip and the eluted tryptic peptides were dried completely using a Speed-Vac concentrator. The dried tryptic peptides were dissolved in 3 μL of Matrix Buffer (10 mg mL^{-1} 4-hydroxy-α-cyanocinnamic acid, 193 fmol mL^{-1} angiotensin internal standard, and 340 fmol mL^{-1} ACTH 7-38 peptide internal standard in a 50% acetonitrile/49.1% NanoPure water/0.1% trifluoroacetic acid solution). The resulting mass spectrometry sample was spotted onto a MALDI target plate and allowed to air dry. The air-dried mass spectrometry sample was then analyzed on an Applied Biosystems Voyager DE-PRO mass spectrometer in linear mode with the following settings: Accelerating Voltage = 21000; Grid Voltage = 95%; Guide Wire Voltage = 0.050%; Delay Time = 200 nsec; Laser Power = 1800-2000. The resulting average masses were smoothed using 19-point Gaussian smoothing and manually calibrated using the angiotensin and ACTH 7-38 peptide internal standards.

Primer extension inhibition, or toeprinting, assays of mutagenized and fluorescently labeled IF3 variants

The biochemical activities of mutagenized and fluorescently labeled $IF3^{C65S\ S38C\ K97C}$ (IF3(Cy3-Cy5)) and $IF3^{C65S\ S38C\ K97C\ Y75N}$(Cy3-Cy5) were tested using a primer extension inhibition, or toeprinting, assay (Supplementary Fig. 2). Toeprinting reactions were performed as described in Fei et al.[1], Hartz et al.[2], and Maar et al.[3]. Briefly, reactions were performed by mixing 0.1 μM 30S subunits, 0.05 μM ^{32}P-primer-annealed T7 gp32 mRNA, 1 μM $tRNA^{fMet}$, and 10 μM $tRNA^{Phe}$ when present (note that neither tRNA is aminoacylated in this assay) in a final volume of 4.5 μL in either Tris-Polymix Buffer (Supplementary Table 4) or Tris-Polymix Buffer* (50 mM Tris-acetate, $pH_{25°C}$=7.0, 100 mM potassium chloride, 5 mM ammonium acetate, 0.5 mM calcium acetate, 0.1 mM ethylenediamine tetraacetic acid (EDTA), 6 mM 2-mercaptoethanol, 1 mM putrescine dihydrochloride, 0.2 mM spermidine free base, and 5 mM magnesium acetate) and incubating this mixture for 10 minutes at 37 °C. Either Tris-Polymix Buffer (for those reactions performed in the absence of IF3) or the specified concentration of IF3 (for those reactions performed in the presence of IF3) was then added to the mixture in a final volume of 5 μL in Tris-Polymix Buffer and the reaction was subsequently incubated for 10 minutes at 37 °C and an additional 10 minutes on ice. Reverse transcription was initiated by adding 31.3 nmol ATP, 15.6 nmol each of dGTP, dATP, dCTP, and dTTP, and 15 U AMV

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

reverse transcriptase (Promega or New England Biolabs) to the reaction in a final volume of 20 µL in Tris-Polymix Buffer and allowing reverse transcription to proceed for 15 minutes at 37 °C. The resulting cDNA products were purified by phenol extracting and chloroform extracting the reaction and precipitating the cDNA products through the addition of 0.1× reaction volume of 3 M sodium acetate and 3× reaction volume of 100% ethanol to the reaction, incubating for 10 min at room temperature and centrifuging on a table top microfuge for 10 min at 14,000 ×g. The precipitated cDNA products were resuspended in Gel Loading Buffer (23 M formamide, 0.09% bromophenol blue, and 0.09% xylene cyanol) and were separated by running them on a 9% denaturing polyacrylamide gel for 1.5 hours at a constant power of 55 W. The gel was dried and exposed to a PhosphorImager screen (GE Life Sciences). The screen was scanned with a Typhoon FLA 7000 PhosphorImager (GE Healthcare) and analysis was performed with ImageQuant (Molecular Dynamics). The selectivity of $tRNA^{fMet}$ at AUG codon over $tRNA^{Phe}$ at UUC codon was defined as follows:

$$\text{Selectivity of } tRNA^{fMet} \text{ at AUG over } tRNA^{Phe} \text{ at UUC (\%)} = \left(1 - \left(\left(\frac{UUC}{AUG}\right)_{\text{lane x}} \div \left(\frac{UUC}{AUG}\right)_{(-)IF3}\right)\right) \times 100$$

where UUC is the phosphor intensity at the +18 band, AUG is the phosphor intensity at the +15 band, lane x is any one of lanes 3-9 in the gels, and (–)IF3 is lane 2.

Preparation of 30S ICs

30S ICs for smFRET studies were prepared by incubating 1.8 µM 5'-biotinylated mRNA, 0.9 µM IF1, 0.9 µM IF2, 0.9 µM tRNA, 0.6 µM 30S subunits, and 0.6 µM IF3(Cy3-Cy5) at 37 °C for 10 min in Tris-Polymix Buffer (Supplementary Table 4). 30S ICs were then aliquoted, flash frozen in liquid nitrogen, and stored at –80 °C until further use.

Total internal reflection fluorescence (TIRF) microscopy

30S ICs for imaging by TIRF microscopy were thawed, diluted to ~200 pM in Tris-Polymix Buffer (Supplementary Table 4), introduced into a microfluidic flowcell that had been passivated with a mixture of polyethylene glycol and biotinylated polyethylene glycol (PEG) and derivatized with streptavidin as previously described[7], and incubated at room temperature for 5 min. 30S ICs that failed to tether to the surface of the flowcell at the end of the 5 min incubation were removed by flushing the flowcell with Tris-Polymix Buffer containing an enzymatic oxygen scavenger system, a triplet state quencher cocktail, and, as specified in individual experiments, mixtures of IFs and aminoacyl-tRNAs (Supplementary Table 4).

A previously described, laboratory-built, wide-field, prism-based TIRF microscope[6] was used to image the flowcells containing the surface-tethered 30S ICs. Briefly, a diode-pumped, solid-state, 532 nm laser (CrystaLaser) operating at a power of 7 mW (measured just prior to striking the prism) was used to directly excite Cy3 and a diode-pumped, solid-state, 643 nm

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

laser (CrystaLaser) operating at a power of 18 mW (measured just prior to striking the prism) was used to directly excite Cy5. Fluorescence emission from Cy3 and/or Cy5 was collected through a high numerical aperture objective (Nikon), wavelength separated into individual Cy3 and Cy5 fields-of-view using a Dual-View simultaneous imaging system (Photometrics, Inc.), and simultaneously imaged using the two halves of a back-thinned, 512 x 512 pixel electron-multiplying charged-coupled device (EMCCD) camera (Cascade II 512:B; Photometrics, Inc.) operating with 2 x 2 pixel binning and a frame rate of 10 frames sec^{-1}.

200-400 spatially well-separated 30S ICs were imaged within a 60 x 120 μm^2 field-of-view. Direct excitation of Cy5 using the 643 nm laser during the first frame of each movie was used to record the spatial location of each Cy5 fluorophore in the field-of-view. The 643 nm laser was subsequently switched off and, simultaneously, the 532 nm laser was switched on in order to directly excite Cy3 and perform smFRET imaging starting with the second frame of each movie. Imaging continued until >95% of the Cy3 fluorophores had photobleached. Three independent datasets consisting of 12-15 movies each were collected on separate days using independently prepared samples and microfluidic devices for each 30S IC.

TIRF microscopy-based tRNA dissociation assays of mutagenized and fluorescently labeled IF3 variants

The biochemical activities of mutagenized and fluorescently labeled $IF3^{C65S\ S38C\ K97C}$ (IF3(Cy3-Cy5)) and $IF3^{C65S\ S38C\ K97C\ Y75N}$(Cy5-Cy5) were further tested using a TIRF microscopy-based tRNA dissociation assay (Supplementary Fig. 2) that were performed as follows. 30S ICs lacking IF2 (30S IC_{-2}) were assembled using a 5'-biotinylated-mRNA, 30S subunits, IF1, an fMet-$tRNA^{fMet}$ that is labeled at the 4-thiouridine at nucleotide position 8 (fMet-(Cy3)$tRNA^{fMet}$) (ref. 1), and wild-type, mutant, and/or Cy5-labeled variants of IF3. A control 30S IC_{-2} additionally lacking IF3 (30S $IC_{-2/-3}$) was prepared in a manner identical to that used to prepare the five experimental 30S IC_{-2}s with the exception that IF3 was not included in the reaction. The same ~200 pM concentrations of each experimental 30S IC_{-2} and the control 30S $IC_{-2/-3}$ were loaded into separate flowcells on the same microfluidic device for direct side-by-side comparison and, after a 5 minute incubation at room temperature, any 30S ICs that remained untethered to the surface of the flowcell were removed by flushing the flowcell with Tris-Polymix Buffer containing an enzymatic oxygen scavenger system and a triplet state quencher cocktail (Supplementary Table 4). The flowcells were then imaged by recording 10 randomly selected fields-of-view from each flowcell using TIRF microscopy. The % fMet-(Cy3)$tRNA^{fMet}$ bound to the P site of each surface-tethered 30S IC_{-2} was quantified by counting the average number of diffraction-limited Cy3 spots observed over the 10 fields-of-view of each 30S IC_{-2} and normalizing the resulting average number of Cy3 spots to the average number of Cy3 spots observed over the 10 fields-of-view of 30S $IC_{-2/-3}$ (assuming that the average number of Cy3 spots observed over 10 fields-of-view of 30S $IC_{-2/-3}$ represents 100% binding of fMet-(Cy3)$tRNA^{fMet}$ to a 30S IC_{-2} or a 30S $IC_{-2/-3}$). The % fMet-(Cy3)$tRNA^{fMet}$ dissociated from the P site of each 30S IC_{-2} was then calculated

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

by subtracting the % fMet-(Cy3)tRNA^{fMet} bound to the P site of each 30S IC_{-2} from 100%. (See Supplementary Fig. 2d)

smFRET data analysis

Generation and selection of single-molecule E_{FRET} *versus* time trajectories from each movie was performed as previously described[7-9]. Briefly, the first frame of each movie, which was collected using direct excitation of Cy5 with a 643 nm laser, was used to identify single, diffraction-limited Cy5 spots. The locations of these spots were transferred to the Cy3 field-of-view in order to align the Cy5 field-of-view with the subsequent 532 nm-directly excited Cy3 field-of-view. The aligned Cy3 and Cy5 fields-of-view were used to identify pairs of Cy3 and Cy5 spots corresponding to single, surface-tethered, 30S ICs carrying dual Cy3-Cy5 labeled IF3s and MetaMorph (Molecular Devices), Excel (Microsoft), Origin (OriginLab Corporation), or Matlab (The MathWorks) were used to plot Cy3 and Cy5 intensity *versus* time trajectories for each IF3. Trajectories exhibiting: (i) time-averaged Cy3 and Cy5 intensity values characteristic of single Cy3 and Cy5 fluorophores, respectively, as determined by visual inspection; (ii) single-step photobleaching of Cy3 and/or Cy5 fluorophores as determined by visual inspection; (iii) anticorrelated changes in Cy3 and Cy5 intensities as determined by visual inspection and (iv) Cy5 fluorescence lasting longer than one second prior to photobleaching as determined by visual inspection were kept for further analysis (see Fig. 2a for representative Cy3 and Cy5 *versus* time trajectories). In addition to these selection criteria, trajectories in which FRET could not be confirmed due to the simultaneous, single-step drop of both Cy3 and Cy5 intensities to baseline prior to undergoing an anticorrelated change in Cy3 and Cy5 intensities (<10 % of the total number of trajectories per independently collected dataset) were omitted from further analysis. Each of the Cy3 and Cy5 *versus* time trajectories selected for further analysis was baseline corrected by subtracting the average EMCCD readout over the last ten Cy3 time points (i.e. after photobleaching of the Cy3 fluorophore) from each Cy3 time point and subtracting the average EMCCD readout over the last ten Cy5 time points (i.e. after photobleaching of the Cy5 fluorophore) from each Cy5 time point. In addition, each Cy5 time point was corrected for bleed-through of Cy3 intensity into the Cy5 field-of-view, which arises from the imperfect performance of emission filters, by subtracting 7% of the total Cy3 intensity (i.e. the experimentally determined average amount of Cy3 intensity that bleeds through into the Cy5 field-of-view in our TIRF microscope system) at each time point from the Cy5 intensity at the same time point. Each pair of baseline- and bleed-through-corrected Cy3 and Cy5 *versus* time trajectories were converted to a single, raw E_{FRET} *versus* time trajectory using the equation $E_{FRET}=I_{Cy5}/(I_{Cy3} + I_{Cy5})$, where E_{FRET} is the FRET efficiency at each time point and I_{Cy3} and I_{Cy5} are the baseline- and bleed-through-corrected Cy3 and Cy5 intensities at each time point, respectively. The raw E_{FRET} *versus* time trajectories were idealized by hidden Markov modeling using the vbFRET software package[10] and further analyzed as described in Fig. 2 and Supplementary Tables 1-3.

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554

REFERENCES

1. Fei, J. et al. A highly purified, fluorescently labeled in vitro translation system for single-molecule studies of protein synthesis. *Methods Enzymol* **472**, 221-59 (2010).

2. Hartz, D., McPheeters, D.S. & Gold, L. Selection of the initiator tRNA by Escherichia coli initiation factors. *Genes Dev* **3**, 1899-912 (1989).

3. Maar, D. et al. A single mutation in the IF3 N-terminal domain perturbs the fidelity of translation initiation at three levels. *J Mol Biol* **383**, 937-44 (2008).

4. Sussman, J.K., Simons, E.L. & Simons, R.W. Escherichia coli translation initiation factor 3 discriminates the initiation codon in vivo. *Mol Microbiol* **21**, 347-60 (1996).

5. Greenfeld, M., Pavlichin, D.S., Mabuchi, H. & Herschlag, D. Single Molecule Analysis Research Tool (SMART): an integrated approach for analyzing single molecule data. *PLoS One* **7**, e30024 (2012).

6. Blanchard, S.C., Kim, H.D., Gonzalez Jr., R.L., Puglisi, J.D. & Chu, S. tRNA dynamics on the ribosome during translation. *Proc Natl Acad Sci U S A* **101**, 12893-8 (2004).

7. Fei, J., Kosuri, P., MacDougall, D.D. & Gonzalez Jr., R.L. Coupling of ribosomal L1 stalk and tRNA dynamics during translation elongation. *Mol Cell* **30**, 348-59 (2008).

8. Fei, J. et al. Allosteric collaboration between elongation factor G and the ribosomal L1 stalk directs tRNA movements during translation. *Proc Natl Acad Sci U S A* **106**, 15702-7 (2009).

9. Sternberg, S.H., Fei, J., Prywes, N., McGrath, K.A. & Gonzalez Jr., R.L. Translation factors direct intrinsic ribosome dynamics during translation termination and ribosome recycling. *Nat Struct Mol Biol* **16**, 861-8 (2009).

10. Bronson, J.E., Fei, J., Hofman, J.M., Gonzalez Jr., R.L. & Wiggins, C.H. Learning rates and states from biophysical time series: a Bayesian approach to model selection and single-molecule FRET data. *Biophys J* **97**, 3196-205 (2009).

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2554