Protein-guided RNA dynamics during early ribosome assembly

Hajin Kim1,2*, Sanjaya C. Abeyasinghawarden3*, Ke Chen4,5, Megan Mayerle6†, Kaushik Ragunathan4†, Zaida Luthey-Schulten1,4,5, Taekjip Ha1,2,4,5 & Sarah A. Woodson3,6

The assembly of 30S ribosomes requires the precise addition of 20 proteins to the 16S ribosomal RNA. How early binding proteins change the ribosomal RNA structure so that later proteins may join the complex is poorly understood. Here we use single–molecule fluorescence resonance energy transfer (FRET) to observe real–time encounters between Escherichia coli ribosomal protein S4 and the 16S 5′ domain RNA at an early stage of 30S assembly. Dynamic initial S4–RNA complexes pass through a stable non–native intermediate before converting to the native complex, showing that non–native structures can offer a low free–energy path to protein–RNA recognition. Three–colour FRET and molecular dynamics simulations reveal how S4 changes the frequency and direction of RNA helix motions, guiding a conformational switch that enforces the hierarchy of protein addition. These protein–guided dynamics offer an alternative explanation for induced fit in RNA–protein complexes.

1Department of Physics, Center for the Physics of Living Cells and Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. 2Howard Hughes Medical Institute, Urbana, Illinois 61801, USA. 3T. C. Jenkins Department of Biophysics, Johns Hopkins University, 3400 N. Charles Street, Baltimore, Maryland 21218, USA. 4Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. 5Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. 6CMDB Program, Johns Hopkins University, 3400 N. Charles Street, Baltimore, Maryland 21218, USA. †Present addresses: School of Nano-Bioscience and Chemical Engineering, Ulsan National Institute of Science and Technology, Ulsan 689-798, Republic of Korea (H.K.); Department of Biochemistry and Biophysics, University of California at San Francisco, 600 16th Street, San Francisco, California 94143-2200, USA (MM); Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, LHRBB-517, Boston, Massachusetts 02115-5730, USA (K.R.).

*These authors contributed equally to this work.

RNA helix dynamics in S4–rRNA complexes

To observe internal motions in the S4–rRNA complex, we labelled S4 with a donor fluorophore, cyanine 3 (Cy3), via an engineered single cysteine (Methods). We also labelled the 5′ domain RNA by annealing a Cy5-conjugated oligonucleotide to a 3′ extension of 16S h3 helix (5′-dom-h3). We labelled h3 because it docks under the h18 pseudoknot and contacts S4 in the mature 30S ribosome (Fig. 1b), yet was proposed to point away from h18 and S4 in an assembly intermediate21. Therefore, a label on h3 was likely to capture the dynamics of early assembly intermediates. Chemical footprinting and ensemble binding assays established that these modifications did not noticeably change the folding of the 5′ domain RNA or its affinity for S4 (Extended Data Figs 1–3), which is similar to that of the natural 16S RNA21.

Complexes of S4–Cy3 and 5′-dom-h3–Cy5 were immobilized on a polymer-passivated quartz surface via biotin on the oligonucleotide extending from h3, and imaged by total internal reflection fluorescence microscopy. Single complexes over time displayed discrete transitions between two stable FRET states (Fig. 1c). Analysis of the dwell times showed that the low- and high-FRET states interconvert in 1–10 s in 20 mM Mg2+ (Extended Data Fig. 4). In 20 mM Mg2+, S4 remained bound to the RNA and the complex stayed mostly in the high-FRET state (FRET efficiency E ∼ 0.7) (Fig. 1c, d). Because native interactions bring the Cy5 acceptor on h3 close to the Cy3 donor on S4 (ref. 22), we inferred that the high-FRET state represents the docked conformation of h3 observed in the 30S ribosome, which we take to be the native state of this complex. At 4 mM Mg2+, we observed frequent dissociation and re-binding of S4 (Fig. 1c) and greater occupancy of the low-FRET state (E ∼ 0.2) (Fig. 1d). We assigned this low-FRET state to a ‘flipped’ assembly intermediate in which h3 has swung away from the body of the complex, in agreement with footprinting of the S4–rRNA complex21.

334 | NATURE | VOL 506 | 20 FEBRUARY 2014

©2014 Macmillan Publishers Limited. All rights reserved
traces to the moment of binding (Fig. 2c). Although the high-FRET state
the FRET distribution over time by synchronizing single-molecule
An abrupt increase in fluorescence defined the moment of initial con-
rearranged in 20 mM Mg\textsuperscript{2+} broadley distributed and included a mid-FRET state (Fig. 1d and Extended Data Fig. 4).
These structural assignments were validated by rRNA mutations pre-
ticted to destabilize the docked conformation of h3 (Fig. 1e). The mutation A397C removes a tertiary interaction between h3 and h4, whereas the mutation A499U disrupts adenine stacking that maintains a 90° angle between h3 and h18 (ref. 23). These mutations did not reduce S4 binding, but they prevented stable docking of h3 in 20 mM Mg\textsuperscript{2+}, as revealed by a higher population of low-FRET structures and a broader distribution of mid-to-high FRET values (Fig. 1d and Extended Data Fig. 4). In addition, a base mismatch in the h18 pseudoknot (C507G; Fig. 1b), or deletion of the h18 loop (Ah18loop; teal in Fig. 1b) reduced the high-FRET population, showing that interactions with the folded h18 loop stabilize the docked (native) state of h3 (Fig. 1d and Extended Data Fig. 4).
A dynamic encounter complex
We next asked how S4 initially binds the rRNA. Time-resolved foot-
printing experiments showed that S4 contacts the 16S rRNA in multiple stages spanning 10 ms to 100 s (ref. 12). These stretched binding dynamics may arise from conformational changes in the 16S rRNA\textsuperscript{22} or co-folding of the S4 amino terminus, which interacts extensively with h16 and h18 in the ribosome yet is mostly unstructured in free S4 (ref. 24).
We measured the dynamics of S4 binding by flowing S4–Cy3 over
immobilized 5' dom–h3–Cy5 RNA pre-folded in 20 mM Mg\textsuperscript{2+} (Fig. 2a).
An abrupt increase in fluorescence defined the moment of initial con-
tact with S4–Cy3 (Fig. 2b and Extended Data Fig. 5). We built a map of the FRET distribution over time by synchronizing single-molecule traces to the moment of binding (Fig. 2c). Although the high-FRET state was favoured at equilibrium in 20 mM Mg\textsuperscript{2+}, the FRET distribution at the moment of binding (histogram in Fig. 2c) showed that more than 80% of complexes passed through the low-FRET state before switching to the high-FRET state in 1 s.
In 4 mM Mg\textsuperscript{2+}, closer to the physiological condition, ≥40% of complexes showed transient spikes of fluorescence, meaning that S4 dissociated before the complexes could progress to the native state (Fig. 2d and Extended Data Fig. 5). A synchronized map of successful binding trajectories showed that the FRET values at the initial encounter were broadly distributed and included a mid-FRET state (E ≈ 0.6) that was rarely observed in 20 mM Mg\textsuperscript{2+} (Fig. 2e). Lowering [Mg\textsuperscript{2+}] to 2 mM shifted the initial FRET values further towards ~0.6 (Fig. 2f). Analysis of individual trajectories showed that these diverse initial complexes first transitioned to the low-FRET (flipped) state before reaching the high-FRET (Docked) state (Fig. 2d and Extended Data Fig. 5). This sequential pathway was evident in the synchronized map that showed a transient build-up of low-FRET population in the first 1–2 s after initial association (Fig. 2e, f and Extended Data Fig. 6).
To better visualize initial S4 encounters, we made additional observations with 10-ms resolution. These experiments revealed that initial complexes fluctuate rapidly between different structures (Fig. 2g; see also Extended Data Fig. 7) before converging to the low-FRET state within ~0.2 s (Fig. 2h, orange rectangle). Although the S4–rRNA complex starts from a disordered ensemble, it does not directly jump to its native structure but progresses through the low-FRET intermediate. Thus, S4 interactions steer the system from a heterogeneously fluctuating encounter complex towards a slower dynamic equilibrium between two RNA conformations.
Kinetic model of S4–rRNA binding pathway
Our observations of individual binding events revealed a minimal kinetic pathway for S4–rRNA recognition that accounts for the stretched time-frame of RNP assembly (Fig. 3a). In our model, a rapidly fluctuating, heterogeneous S4–5' domain encounter complex (EC) diffuses to a less-
dynamic, low-FRET flipped intermediate (FI), in which h3 swings away from S4 and h18. Finally, h3 docks against S4 and underneath the h18 pseudoknot to form the high-FRET native complex (NC).
The lifetimes of these states in different [Mg\textsuperscript{2+}] provide additional information on the S4 binding mechanism (Fig. 3b–g). First, we estimated the binding rate constant (k\textsubscript{on}) by observing the delay between the addition of S4–Cy3 and the appearance of a fluorescent signal on the rRNA. The maximum k\textsubscript{on} ~5 × 10\textsuperscript{7} M\textsuperscript{−1}s\textsuperscript{−1} agreed well with time-
resolved footprinting\textsuperscript{12}, showing that initial binding occurs near the diffusion limit for macromolecules. S4 bound the 5' domain RNA fast-
est in 10 mM Mg\textsuperscript{2+}, which in footprinting experiments\textsuperscript{21} favours the flipped conformation of h3, but 16 times slower in 20 mM Mg\textsuperscript{2+}, which favours the docked conformation of h3 (Fig. 3b). This supports our observation that the non-native flipped conformation offers a low fre-
energy path for S4 binding, and contrasts with the usual assumption that proteins are more likely to recognize the natively folded RNA.
S4 guides the RNA dynamics

We next investigated how S4 changes pre-existing motions in the free RNA. To compare the conformational dynamics of the RNA before and after S4 binding, we introduced a second label in the rRNA at the tip of h16 (5′-dom-h3-h16). Previous smFRET measurements of a minimal 5WJ revealed the complex folding energy landscape of the RNA in the absence of S4 (ref. 28). We observed similarly heterogeneous dynamics between h3 and h16 in the larger 5′ domain RNA, with varied transition rates between low- and high-FRET states in different molecules (Fig. 4a and Extended Data Fig. 8). These changes in RNA structure are reversible as a subset of molecules switched their dynamics within a single trajectory (Extended Data Fig. 9), as observed in other RNAs\textsuperscript{33–34}.

To see how S4 modulated RNA dynamics, we performed three-colour smFRET\textsuperscript{39} experiments between S4–Cy3, h16–Cy5 and h3–Cy7. By alternating Cy3 and Cy5 excitation, we measured all pairwise distances pseudo-simultaneously, yielding information on the direction as well as frequency of motion\textsuperscript{33,34}. S4 binding suppressed relative motions of h16, as inferred from the lack of discernible fluctuations in the S4–h16 and h16–h3 FRET efficiencies (Fig. 4b). The FRET from S4 to h3 still fluctuated between high and low values, consistent with two-colour data showing that h3 continues to move after S4 binds (Fig. 1c).

We used the observed FRET efficiencies to construct a geometric model of how S4 modulates the motions of the RNA helices. Without S4, the distance between h3 and h16 fluctuates heterogeneously (black arrows in Fig. 4c) as expected for the loosely folded free RNA\textsuperscript{37}. When S4 binds, persistent low FRET between h16 and S4 shows that the motion of h16 is restricted and agrees with the expected distance between these labels in the 30S ribosome (Fig. 4d). Conversely, steady low FRET between h3 and h16, combined with fluctuations in the S4–h3 distance, can be explained by movement of h3 in a plane perpendicular to h16, such that its distance to h16 does not change substantially (blue arrows in Fig. 4c). These dynamics are not likely due to motions in S4, because solution NMR showed that the Cy3-labelled domain of S4 is stably folded\textsuperscript{35}.

Molecular dynamics simulations

To gain further insight into how S4 modulates the dynamics of the RNA, we performed all-atom molecular dynamics simulations of a minimal RNA containing just the 5WJ (Fig. 4e). We pictorialized the global motions of the helices by tracing the movements of their tips (Methods). The helix motions in these and other simulations of the 5WJ in 0–30 mM \( \text{Mg}^{2+} \) (ref. 28) agreed well with the experimentally observed S4–h3–h16 distances. Without S4, h16 explored two distinct regions of space whereas h3 swept out a wide cone (Fig. 4c, top panel), consistent with the fluctuations in FRET between h3 and h16 in the free RNA (Fig. 4a). S4 fixed h16, allowing it to sample only a small region around its native structure (Fig. 4c, bottom panel), consistent with the stable S4–h16 FRET signal (Fig. 4b). More interestingly, S4 constrained h3 to an in-plane bending motion towards and away from S4, maintaining a nearly constant distance between h3 and h16.
Researchers have observed that protein S4, when bound, alters the free RNA through induced motions that are confined to a specific direction. In the absence of S4, the free RNA exhibits anisotropic motions, whereas upon S4 binding, it becomes more restricted, affecting the recruitment of other proteins to the 16S rRNA. This change in dynamics facilitates the search for high-affinity configurations.

**Discussion**

Protein-induced remodelling of RNA structures is widespread and assists the hierarchical assembly of RNA-protein complexes. Unexpectedly, the low free-energy path for S4–rRNA recognition passes from a mobile encounter complex through a non-native intermediate before reaching the newly folded complex, S4 not only slows rRNA fluctuations but also induces anisotropic motions between the intermediate and the native complex. These protein-induced changes in the rRNA dynamics are reminiscent of the classical model of substrate-induced fit in enzymes. Induced fit has been proposed as a universal feature in RNA–protein recognition, as both protein and RNA structures often change upon binding. Our results suggest an alternative picture of induced fit, in which S4 changes the ensemble of thermally accessible RNA conformations by modulating the RNA dynamics. This differs from binding of small ligands to RNA helix junctions, which trap conformations accessed by bending motions of the unliganded RNA. The ability of S4 to change the RNA dynamics after it binds is probably due to its larger interface with the 16S rRNA, and to the plasticity of its N-terminal domain. Many ribosomal proteins contain flexible segments that change structure upon RNA binding. Transient protein–RNA interactions may facilitate the search for a high-affinity configuration.

We anticipate that the observed S4-guided dynamics of the rRNA contribute to subsequent steps of 30S assembly. First, the conserved h18 loop interacts with transfer RNA in the mature ribosome, and must fold correctly for normal protein synthesis. h18 folds correctly only after S4 binds the 16S rRNA in a temperature-dependent step. Our experiments show that folded h18 stabilizes the docked form of h3 (Extended Data Fig. 4). Native interactions between h18 and h3 enable folding of h1 (Fig. 1a) and the central pseudoknot at the end of 30S assembly. S4 mutations that inhibit remodelling of h18 and h3 docking impair 30S assembly in *E. coli*.

Second, h3 dynamics affect the recruitment of other proteins to the 16S rRNA. Protein S16, which adds to 30S complexes only when S4 is present, stabilizes the docked conformation of h3 (ref. 21). Later in 30S assembly, protein S12 binds the docked h3 on the side opposite S4 (Fig. 1a). The restricted in-plane motions of h3 between the flipped and docked conformations may reject S12 until the h18 loop has folded correctly and S16 has bound, ensuring the proper hierarchy of protein addition. Thus, protein-guided RNA dynamics creates additional checkpoints for molecular interactions, thereby improving the fidelity of RNA folding and ribosome assembly.

**METHODS SUMMARY**

The 5′ domain of *E. coli* 16S rRNA (nucleotides (nt) 21–556) was labelled by hybridizing dye-labelled oligonucleotides to an extension of h3 (5′-dom-h3h16) (Extended Data Table 1). *E. coli* protein S4(C32S, S189C) labelled with Cy3 was incubated with the RNA in 80 mM K-HEPES, pH 7.6, 300 mM KCl plus the desired [MgCl₂] for 5 min at 37 °C, or flowed into slide chambers containing immobilized RNA for real-time binding experiments. Complexes were immobilized on quartz slides for imaging with a total internal reflection fluorescence microscope (100 ms per frame except where stated otherwise).

Two-colour FRET efficiencies were defined with leakage correction as $E_{\text{FRET}} = (I_q - 0.06 \times I_h)/I_q + I_h$, where $I_q$ and $I_h$ are the apparent fluorescence intensities of the donor and acceptor, respectively. Three-colour FRET efficiencies were corrected for the leakage, direct excitation of acceptor dyes by lasers, and the wavelength-dependence detection efficiency. See full Methods for data statistics. Molecular dynamics simulations (150 ns) using NAMD2 and CHARMM27 force fields started from 3WJ RNA and S4 models based on the *E. coli* ribosome structure (PDB 2I2P), neutralized with Mg²⁺ and Na⁺ ions, and solvated with a periodic boundary condition. Hybrid MD-Gō simulations of 100 S4 binding trajectories started from an extended RNA structure and included a Lennard-Jones Gō potential to drive protein–RNA association.

**Online Content**

Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 September 2013; accepted 20 January 2014.
Published online 12 February 2014.

1. Held, W. A., Ballou, B., Mizushima, S. & Nomura, M. Assembly mapping of 30S ribosomal proteins from *Escherichia coli*. Further studies. *J. Biol. Chem.* 249, 3103–3111 (1974).
METHODS

Oligonucleotides and design of modified 5′-domain RNA for smFRET. Two extensions in the 5′ domain of the E. coli 165 RNA (nucleotides (nt) 21–556, E. coli numbering) created binding sites for fluorescently labelled oligonucleotides as described previously50,51. First, the 3′-end (5′-dom-h3) was extended by 38 nucleotides (5′- AGGACGACACUUUGGAGAGGACACACAGGACACA GG-3′) to base-pair with DNA oligonucleotide h3p5 derivatized on its 3′-end with the acceptor dye (Cy3 or Cy5) and 5′-end with a 15-nt RNA, h16P3, modified with the donor dye (Cy3 or Cy5) of h16 (nt 420–423) was replaced by a 26-nt extension (5′- CPh3; 5′-end). Secondary DNA strand (CPh3; 5′-end) with the acceptor dye (Cy5 or Cy7) and 5′-end with the donor dye was removed by ion exchange chromatography followed by overnight dialysis against 50 mM K-HEPES, pH 7.6, 2 mM EDTA, 0.6% glucose, 0.1 mg ml−1 turboMEF (GE Healthcare) at 4°C for 2 h. The reaction was quenched by adding 20 mM Tris-HCl, pH 7.0, 6 mM urea, 2 mM mercaptoethanol. Unreacted dye was removed by ion exchange chromatography followed by overnight dialysis against 80 mM K-HEPES, pH 7.6, 1 mM KCl, 1 mM TCEP, 3 M urea) was incubated at 20°C for 30 min, and then treated with sifoloxin oxene Cy3-maleimide (GE Healthcare) at 20°C for 2 h. The protein and Cy3 concentrations were determined from the absorbance at 280 nm and 490 nm, respectively. Post-synchronized density plots were constructed from the traces that eventually progressed into the high-FRET state before the dyes photobleached, to exclude non-specific binding events. The traces were overlaid at the same FRET radius (55 Å) used in the analysis of experimental data. 61 out of 100 trajectories showed successful binding of S4 within the simulation time (2,250,000 steps) and were taken for analysis. These trajectories were further synchronized such that the step time where the first native contact between S4 and the 5WJ was established in each trajectory was set as time 0 to generate the density map.

S4 protein. A single cytinese for fluorescent labelling of E. coli protein S4 was created by site-directed mutagenesis (QuikChange) of the PET24b/pS4 gene plasmid52. The natural cyteinse at position 32 was replaced with serine, and serine 189 was replaced with cytinese (E. coli numbering). S4(C32S, I189C) was overexpressed and purified using cation exchange chromatography53. Isolated proteins were dialysed against 80 mM K-HEPES, pH 7.6, 1 mM KCl, 1 mM TCEP (2-carboxyethyl) phosphine HCl (TCEP) and stored at −80°C in 500-μl aliquots.

smFRET measurements. RNA 5′-dom-h3 (10 mM), biotinylated h3 oligonucleotide (h3p5, 5 mM), and the complementary strand (CPh3, 10 mM) were mixed in 80 mM K-HEPES, pH 7.6, 330 mM KCl. In case of 5′-dom-h3h16, h16 oligonucleotide (h16P2-2 (5′-r(AGGGCACAGACAGGTG)-Cy3-Cy5-3′), 60 mM) was added. The solution was heated to 75°C for 5 min and slowly cooled to 37°C by leaving the microcentrifuge tube in the heating block. The desired amount of MgCl2 was added and the solution was heated at 37°C for 5 min to fold the RNA. For equilibrium measurements, the reaction was diluted fivefold in the same buffer and incubated with 5 mM S4–Cy3 for 5 min at 37°C. The complex was immobilized on quartz slides covered with a 30:1 mixture of PEG and biotinylated PEG and treated with neutravidin54. For flow-in-measurements, the complex was prepared without S4–Cy3 and immobilized on the surface. S4–Cy3 was added using the syringe attached to the micro-chamber while the movie was recorded. smFRET traces were measured using a total internal reflection fluorescence microscope as described previously55. The imaging solution contained HK buffer, the stated amount of MgCl2, 2 mM Trolox, 0.6% glucose, 0.1 mg ml−1 glucose oxidase (Sigma) and 0.02 mg ml−1 catalase (Sigma). For two-colour FRET, we calculated the FRET efficiency defined with leakage correction as $E_{\text{FRET}} = (I_o - 0.06 I_d)/(I_d + I_o)$, where $I_o$ and $I_d$ are the measured intensities of the donor and acceptor, respectively. For three-colour FRET, we further considered the direct excitation of Cy5 by the red laser (for exciting Cy5) and corrected for the lower detection efficiency for Cy7. Details of the corrections needed to calculate multicolour FRET efficiencies are described elsewhere56.

smFRET data analysis. FRET histograms were made from FRET values at each time frame (100 or 10 ms). In Fig. 1d, traces of 224, 64, 39, 611, 139 and 253 molecules were collected for 20 mM Mg2+, 4 mM Mg2+, A397C, A499U, C507G and Δh18loop, respectively. Post-synchronized density plots were constructed from the traces that eventually progressed into the high-FRET state before the dyes photobleached, to exclude non-specific binding events. The traces were synchronized at the moment of the increase of total fluorescence above a threshold, which was determined as half of the total fluorescence of the bound state. The order of analysis of the smFRET data was the same FRET radius (55 Å) used in the analysis of smFRET data. For equilibrium measurements, the reaction was diluted fivefold in the same buffer and incubated with 5 mM S4–Cy3 for 5 min at 37°C. The complex was immobilized on quartz slides covered with a 30:1 mixture of PEG and biotinylated PEG and treated with neutravidin54. For flow-in-measurements, the complex was prepared without S4–Cy3 and immobilized on the surface. S4–Cy3 was added using the syringe attached to the micro-chamber while the movie was recorded. smFRET traces were measured using a total internal reflection fluorescence microscope as described previously55. The imaging solution contained HK buffer, the stated amount of MgCl2, 2 mM Trolox, 0.6% glucose, 0.1 mg ml−1 glucose oxidase (Sigma) and 0.02 mg ml−1 catalase (Sigma). For two-colour FRET, we calculated the FRET efficiency defined with leakage correction as $E_{\text{FRET}} = (I_o - 0.06 I_d)/(I_d + I_o)$, where $I_o$ and $I_d$ are the measured intensities of the donor and acceptor, respectively. For three-colour FRET, we further considered the direct excitation of Cy5 by the red laser (for exciting Cy5) and corrected for the lower detection efficiency for Cy7. Details of the corrections needed to calculate multicolour FRET efficiencies are described elsewhere56.

REFERENCES

©2014 Macmillan Publishers Limited. All rights reserved
56. Lee, J. et al. Single-molecule four-color FRET. Angew. Chem. Int. Edn Engl. 49, 9922–9925 (2010).
57. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. J. Comput. Chem. 26, 1781–1802 (2005).
58. MacKerell, A. D., Jr, Banavali, N. & Foloppe, N. Development and current status of the CHARMM force field for nucleic acids. Biopolymers 56, 257–265 (2000).
59. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. J. Mol. Graph. 14, 33–38, 27–28 (1996).
60. Wilkinson, K. A., Merino, E. J. & Weeks, K. M. Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. Nature Protocols 1, 1610–1616 (2006).
61. Adilakshmi, T., Ramaswamy, P. & Woodson, S. A. Protein-independent folding pathway of the 16S rRNA 5` domain. J. Mol. Biol. 351, 508–519 (2005).
Extended Data Figure 1 | Modification of the 5′ domain RNA preserves its structure. a, The secondary structures of the wild-type and extended 5′ domain RNAs were probed by selective 2′-hydroxyl acylation analysed by primer extension (SHAPE)<sup>19</sup>. The 5′ domain RNA (2 pmol) was annealed to unlabelled oligonucleotides and folded in HKM20 buffer (80 mM K-HEPES, pH 7.6, 300 mM KCl, 6 mM 2-mercaptoethanol, 20 mM MgCl<sub>2</sub>) before treatment with 3 mM N-methylisatoic acid (NMIA) at 42 °C for 26 min. Modifications were detected by primer extension and quantified as described previously.<sup>19</sup> Results of SHAPE chemical probing of the free RNA structure for the wild-type 5′ domain (left) and the 5′ domain with h16 and h3 extensions after annealing with oligonucleotides (right). Saturation of grey indicates reactivity with NMIA. Dashed circles indicate nucleotides that were not detected in our primer extension assay. The results show that the extensions added for the fluorescent labelling of the rRNA do not significantly perturb the rRNA folding. b, Native PAGE folding assay of 5′ dom-h3h16. Fluorescently labelled oligonucleotide (h3P5–Cy5; 25 nM) was annealed to an equimolar concentration of extended 5′ domain RNA in 10 μl CE buffer (20 mM Na-cacodylate, 0.5 mM Na<sub>2</sub>EDTA) for 5 min at 70 °C and 5 min at 25 °C. The RNA–oligonucleotide complex was then folded at 37 °C for 30 min in varying [MgCl<sub>2</sub>] (0–20 mM) before electrophoresis on a native 8% polyacrylamide gel containing 10 mM MgCl<sub>2</sub>. The folding midpoint was 0.5 ± 0.1 mM MgCl<sub>2</sub>, similar to that of the wild-type 5′ domain RNA (1.4 ± 0.2 mM) reported previously.<sup>41</sup>
Extended Data Figure 2 | Annealing of labelled oligonucleotides to the 5’ domain RNA. a, b, 32P-labelled oligonucleotides were annealed to the 3’ extension of h3 of 5’dom-h3 (h3P5, DNA) (a) or to the extended loop of h16 of 5’dom-h3h16 (h16P2-2, RNA) (b). Annealing reactions were performed in 80 mM K-HEPES pH 7.6, 330 mM KCl, 6 mM 2-mercaptoethanol at 25 °C. Binding data were fit to the quadratic form of a two-state binding isotherm. Apparent dissociation constants were ≤0.4 nM and 2.6 ± 0.2 nM for h3 and h16 oligonucleotides, respectively. Equilibrium constants are the average and s.d. of two or more independent trials. The lengths of the labelled oligonucleotides were varied to optimize affinity with the extended 5’ domain RNA, while avoiding perturbations to S4 binding (see Extended Data Fig. 1).
**Extended Data Figure 3 | S4 labelling and its binding to the rRNA.** *E. coli* ribosomal protein S4 was overexpressed, purified and labelled with Cy3 or Cy5 fluorescent dyes as described in Methods. a, SDS–PAGE of unlabelled protein stained with Coomassie (left) or labelled with Cy5 (right). b, c, Ensemble titration of the modified 5′ domain RNAs in HKM20 buffer shows that S4–Cy5 binds with similar affinity as the wild-type S4–rRNA complex. Extended 5′ domain RNAs annealed with h3P5–Cy3 and/or h16 oligonucleotides were titrated with S4–Cy5 in a 500-μl cuvette, and the fluorescence emission was recorded from 550 to 700 nm with 540-nm excitation (Fluorolog-3, Horiba). Excitation and emission slits were fixed at 2 nm and 5 nm, respectively. The sample was incubated at 37 °C for 1 min after each addition. Two or more independent measurements were averaged and titration curves were fitted to a quadratic binding expression. Equilibrium dissociation constants were 5′dom-h3, $0.11 \pm 0.02 \text{nM}$ (statistical error of the fit parameter) and 5′dom-h3h16, $0.2 \pm 0.1 \text{nM}$, at 37 °C, and were comparable to that of the 5′ domain RNA with wild-type *E. coli* S4 (0.9 nM)\(^1\).
Extended Data Figure 4 | Exchange kinetics of docked and flipped complexes. a, Sample FRET traces are shown for mutant rRNAs in 20 mM MgCl₂. b, The cumulative histograms of the dwell times in the high- and low-FRET conformations were calculated for wild-type 5' domain RNA, C507G mutant and Δh18loop mutant, and fit with both mono-exponential and bi-exponential decay functions. One of the triplicate sets of data is demonstrated for each (refer to Methods for number of traces). Significantly lower χ² values suggest that the data are best fit with two exponential terms, except for transitions from the low-FRET to the high-FRET state of the wild-type complex, which was well fit by a single exponential decay function. c, d, All of the dwell-time histograms were fit with bi-exponential decay and the fitting parameters were compared between the wild-type and the mutants. Both components of the transition from the high- to low-FRET state were faster in the mutants than in the wild-type complex. The lifetime of the wild-type low-FRET state had a single component; for the mutants we observed an additional slow component in the lifetime.
Extended Data Figure 5 | Sample traces of S4 binding and binding trials.

a, In 20 mM Mg^{2+}, the binding occurred mostly at the low-FRET state. Infrequently, we observed the dissociation and secondary binding of S4.

b, At 4 mM Mg^{2+}, we often observed unsuccessful and transient binding of S4. Arrows indicate the transient fluorescence signals from unstable binding. These traces also exhibit the mid-FRET spike at the beginning of successful binding events.

c, The portion of the molecules that form a stable complex on the first try to the total molecules that form stable complex within 5 min was plotted with varying [Mg^{2+}]. The error bars represent the 95% confidence interval assuming a binary distribution. The number of molecules used was 76, 81, 318 and 170 for 2, 4, 10 and 20 mM Mg^{2+}, respectively.
Extended Data Figure 6 | Progression of FRET population at different [Mg\(^{2+}\)]. From the synchronized maps of FRET distribution as shown in Fig. 2, the relative populations of low (0–0.35)-, mid (0.35–0.55)- and high (0.55–0.9)-FRET states were plotted at different [Mg\(^{2+}\)]. In 20 and 10 mM Mg\(^{2+}\), the bound complexes started with large low-FRET population that converted to high-FRET population within 5–10 s. At 4 and 2 mM Mg\(^{2+}\), there were considerable mid- and high-FRET populations in the beginning, reflecting the broad initial FRET distribution (Fig. 2e, f). This quickly converted to the low-FRET population, which was then followed by slow conversion to the high-FRET population within several seconds. The number of molecules used was 112, 239, 116 and 275 for 20, 10, 4 and 2 mM Mg\(^{2+}\), respectively.
Extended Data Figure 7 | Sample traces at 10-ms time resolution.
Single-molecule traces at higher time resolution demonstrate the heterogeneous and fluctuating behaviour of the encounter complex (S4–Cy3 and 5’dom-h3–Cy5). The change of FRET in different molecules cannot be described as a single behaviour. In general, the initial FRET distribution over the complexes is broad and the FRET signal converges to the relatively stable low FRET before the transition to the high FRET.
Extended Data Figure 8 | Dynamics of free 5’ domain-h3h16. 

a, Schematic of extensions in h16 and h3 with labelled oligonucleotides. b, Sample FRET traces showing the fluctuation between two distinct states. The frequency and distribution of these fluctuations varied between molecules. c, Mg\(^{2+}\) dependence of the molecular heterogeneity. Histograms show the relative high-FRET population for each molecule (157, 162 and 74 traces for 20, 5 and 1 mM Mg\(^{2+}\), respectively). At higher [Mg\(^{2+}\)], more molecules stay in the high-FRET state for longer periods of time. High FRET between h16 and h3 does not necessarily correspond to the native structure of the 5’ domain RNA in complex with protein S4 that is represented by high FRET between S4 and h3.
Extended Data Figure 9 | Examples of switching between different h16–h3 dynamics. Arrows indicate when the fluctuation dynamics switch between stable high FRET, stable low FRET and alternating high- and low-FRET behaviours. The RNA was labelled as in Extended Data Fig. 8.
Extended Data Figure 10 | S4 binding trajectories from hybrid MD-Go simulations. a, Simulated FRET between S4 and h3 from representative binding trajectories displays various binding pathways. b, Density map constructed from 61 successful binding trajectories. The trajectories were synchronized at the moment when the first native contact between S4 and the 5WJ is established (dotted lines in a). c, Sample trajectories of successful binding, showing how folding of h16 and h3 is induced by S4 binding.