Methods xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Structural dynamics of protein S1 on the 70S ribosome visualized by ensemble cryo-EM

Anna B. Loveland, Andrei A. Korostelev *

RNA Therapeutics Institute, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, 368 Plantation St., Worcester, MA 01605, USA

ARTICLE INFO

Article history:
Received 19 October 2017
Accepted 6 December 2017
Available online xxxx

Keywords:
Rps1
Ribosome
Electron cryo-microscopy
3D maximum-likelihood classification

ABSTRACT

Bacterial ribosomal protein S1 is the largest and highly flexible protein of the 30S subunit, and one of a few core ribosomal proteins for which a complete structure is lacking. S1 is thought to participate in transcription and translation. Best understood is the role of S1 in facilitating translation of mRNAs with structured 5′ UTRs. Here, we present cryo-EM analyses of the 70S ribosome that reveal multiple conformations of S1. Based on comparison of several 3D maximum likelihood classification approaches in Frealign, we propose a streamlined strategy for visualizing a highly dynamic component of a large macromolecular assembly that itself exhibits high compositional and conformational heterogeneity. The resulting maps show how S1 docks at the ribosomal protein S2 near the mRNA exit channel. The globular OB-fold domains sample a wide area around the mRNA exit channel and interact with mobile tails of proteins S6 and S18. S1 also interacts with the mRNA entrance channel, where an OB-fold domain can be localized near S3 and S5. Our analyses suggest that S1 cooperates with other ribosomal proteins to form a dynamic mesh near the mRNA exit and entrance channels to modulate the binding, folding and movement of mRNA.

1. Introduction

Advances in electron cryo-microscopy (cryo-EM), including direct electron detectors and improved computational techniques [65,52,18,57], allow structural ensembles of biological complexes to be visualized at high resolution. Unlike in X-ray crystallography where a macromolecular complex is usually found in one or two conformational states, computational separation using 3D maximum likelihood classification allows determination of numerous states in a sample studied by cryo-EM [66,65,52]. We have recently applied these techniques using Frealign 9 [52,35] to visualize several steps of tRNA selection by the ribosome [51], activation of stringent factor RelA [50], viral translation initiation by IRES-containing mRNA [1], ribosome rescue by ArfA [24] and tRNA translocation [9]. In these studies, we separated classes of ribosomes based on global features, including composition (presence or absence of a protein or RNA cofactor) and conformational differences of the ribosome or cofactors, to yield a set of functionally related ribosome structures or "structural ensembles". While cryo-EM data classification often allows interpretation of structural differences at nearly atomic resolution within the ribosome core [30,29,39,49,25], the low resolution of the ribosome surface usually prevents detailed interpretation of mobile peripheral components [7,10,50].

S1 is the largest and one of the least understood ribosomal protein due to its structural dynamics. In *E. coli*, the protein has a mass of 61 kDa and comprises six β-barrel-shaped domains that each resemble the oligonucleotide/oligosaccharide-binding (OB)-fold implicated in nucleic acid binding [11,64]. Biochemical studies suggest that in solution and on the ribosome, S1 can have an elongated shape stretching over 200 Å long [34,46,44,69]. S1 interacts with the 30S subunit and 70S ribosomes [26], with *E. coli* RNA polymerase [72,73,23], and is an integral part of the Qβ phage replisome [74]. In translation, S1 may play multiple roles, including those in translation initiation [42,76,70,22,27], elongation [42,61,70], tmRNA-mediated ribosome rescue [55,63] and possibly ribosome recycling [22]. The best understood role for S1 is in translation initiation where S1 is required for translation of mRNAs with structured 5′ ends or with weak or no Shine-Dalgarno sequences [27]. Unlike most other ribosomal proteins, S1 is thought to exchange off and on to the ribosome [55,22]. Biochemical studies showed that S1 interacts with the ribosome using its N-terminal domains 1 and 2 (D1 and D2), while its C-terminal domains 3–6 interact with mRNA [71,6,64,13,27]. Intermediate-resolution cryo-EM structures show that the major binding site of S1 is near

* Corresponding author.
E-mail address: andrei.korostelev@umassmed.edu (A.A. Korostelev).

https://doi.org/10.1016/j.ymeth.2017.12.004
1046-2023/

the mRNA exit channel and protein S2 [67,58,12,5,25]. The N-terminal alpha helix and D1 have been crystallized as a fusion protein connected with ribosomal protein S2 [12] suggesting specific interactions between these proteins. Whether the first domain is stably bound or loosely associated with S2 on the ribosome is unclear. Other than weak density likely representing D2 near the mRNA exit channel [12], no visualization beyond the first domain is available on the ribosome. However, NMR structures of the individual domains 2, 4, and 6 have been reported [64,33]. Crosslinking suggests that both the N-terminal and C-terminal domains interact with the mRNA exit channel, while the C-terminal domains also extend to distant locations up to 200 Å apart [41,19,45].

Cryo-EM studies that employ "focusing" on a large region of a macromolecular complex, such as a ribosome subunit or domains, have been successful in separating distinct ribosome conformations and improving the resolution of the region [28,35]. S1 is only ~2.5% of the ribosome's mass, binds at the poorly resolved (solvent) side of the ribosome [13,12] and contains 6 domains [64], which are similar to each other and have overlapping binding sites on the ribosome [45]. These factors make S1, the "mRNA catching arm" of the ribosome [71], a challenging subject for structural investigation. In this work, we employ deep classification in Frealign 9 [52,35] to study the dynamics of S1 bound to the ribosome. We propose a workflow for rapid assessment of the conformational variability of mobile components of a macromolecular complex.

2. Materials and methods

We have analyzed our recently described cryo-EM dataset of the decoding *E. coli* 70S ribosome in the presence of ternary complex EF-Tu·GDPCP·aminoacyl-tRNA, comprising 800,367 particles and yielding ribosome structures at up to 3.2 A resolution [51]. A brief description of the sample and cryo-EM conditions is followed by a detailed description of the classification and modeling. We compare four classification strategies, including unsupervised global and focus-based local approaches. We conclude that using a global classification followed by a "dual-mask"-focus classification results in the identification of the most distinct classes of S1.

2.1. Preparation of E. coli 70S ribosome bound with the cognate ternary complex

The ribosome complex was assembled as described [51] and contained the following concentrations of components in the final sample used for cryo-EM grid preparation: 250 nM 50S subunits; 250 nM 30S subunits; 1.25 μM mRNA; 500 nM fMet-tRNA^{fMet}; 1 μM Phe-tRNA^{Phe}; 1 μM EF-Tu and 500 μM GDPCP, in Reaction Buffer (20 mM HEPES·KOH, pH 7.5, 20 mM magnesium chloride, 150 mM ammonium chloride, 2 mM spermidine, 0.1 mM spermine, 6 mM β-mercaptoethanol). The synthetic mRNA contained the Shine-Dalgarno sequence and a linker to place the AUG codon in the P site and the Phe codon (underlined) in the A site (IDT): GGC AAG GAG GUA AAA AUG UUC AAA AAA.

2.2. Grid preparation

Holey-carbon grids (C-flat 1.2–1.3, Protochips) were coated with a thin layer of carbon then glow discharged with 20 mA with negative polarity for 45 s in an EMITECH K100X unit. 2 μl of ribosome complex was applied and, after a 10-s incubation, the grids were blotted for 4 s at 4 °C and ~95% humidity, then plunged into liquid ethane using a CP3 cryo plunger (Gatan Inc.).

2.3. Electron microscopy and image processing

A dataset of 3028 movies was collected automatically using SerialEM [54] on a Titan Krios electron microscope (FEI) operating at 300 kV and equipped with K2 Summit direct electron detector (Gatan Inc.) using 0.5–2.2-μm underfocus. 50 frames per movie were collected at 1 $e^-/Å^2$ per frame for a total dose of 50 $e^-/Å^2$ on the sample. The super-resolution pixel size was 0.82 Å on the sample.

800,367 particles were automatically picked using a ribosome reference and extracted from aligned movie sums as previously described [51].

2.4. Map refinement and reconstruction

Frealign (versions 9.07-9.11) was used for all steps of refinement and reconstruction [52] on a 12-core desktop workstation (Dell Tower 7910 equipped with Dual E5-2670 processor and 128 Gb of RAM). A 6×-binned image stack was initially aligned to a ribosome reference (EM databank map 1003, [32] using 10 rounds of mode 3 (global search) alignment, originally including data in the resolution range from 300 Å to 60 Å and gradually dropping to 30 Å. Next, the 2×-binned, and later the unbinned image stacks were successively aligned against the common reference using mode 1 (local refinement), gradually including data up to a high-resolution limit of 8 Å, whereupon the resolution of the "common reference" stopped improving. During the alignment's reconstruction steps, thresh_reconst was limited to 0.6 so that only 60% of particles were used for reconstruction. Subsequently, the refined parameters were used for classification of 4×-binned image stack into 6 classes in 50 rounds using a spherical (60-Å radius) focus mask around EF-Tu and A/T tRNA, including resolutions from 300 to 8 Å during classification. This procedure yielded three EF-Tu-containing classes, a class with two tRNAs and the non-rotated (classical) ribosome state (Empty A), a class with three tRNAs and the non-rotated ribosome state (Full A), and a class with 1 tRNA and the rotated (hybrid) ribosome state. The particle alignment parameters reported in [51] were used in this work without further refinement.

2.4.1. Global classification

To identify areas of heterogeneity attributable to S1, we classified the full 4×-binned image stack into 4, 8, 16 or 32 classes using a resolution range of 300 to 8 Å for classification without any masks. 50 rounds of classification were used. Maps of the final round of the classification were prepared from the unbinned image stack without further parameter refinement.

2.4.2. Classification with 3D mask around proteins known to crosslink to S1

To limit classification to parts of the ribosome known to interact with S1, we created a binary mask for classification around the 30S ribosomal subunit proteins: S2, S3, S6, S7, S9, S18, S19 and S21, which were previously shown to crosslink to S1 [45] and their close neighbors: S4, S5, S10, S11, S14 and S15. To create the mask, the protein coordinates were extracted from PDB: 5UYM [51] and a density volume was calculated using Bsoft (bgex command) [36] using the 6×-binned voxel size. The volume was low-pass filtered to 30 Å in EMAN2 [75]. The low-pass filtered volume was viewed in UCSF Chimera [59] over PDB: 5UYM and a level was found such that the protein backbone of modeled proteins was ~15–20 Å away from the edge of the mask. The level was 0.1 and was slightly lower than the mean of the low-pass filtered volume (−0.07 sigma). EMAN2 [75] was used to create a binary mask at this level.