Single-particle cryo-electron microscopy studies of ribosomes with fragmented 28S rRNA

Cristina Gutierrez-Vargas

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

under the Executive Committee

of the Graduate School of Arts and Sciences

Columbia University

2020

Abstract

Single-particle cryo-electron microscopy studies of ribosomes with fragmented 28S rRNA

Cristina Gutierrez-Vargas

In the past five years, single-particle cryo-electron microscopy (cryo-EM) has revolutionized structural biology. Recent advances in detector technology and powerful computational methods now allow images of unprecedented detail to be recorded and structures to be determined at near-atomic resolution. For my PhD studies, I took advantage of this technique to study the structure of uniquely fragmented ribosomes.

Ribosomes, are large macromolecular complexes that translate genetic information carried by messenger RNAs (mRNAs) into polypeptide chains. They are the protein production factories of all living cells and are thus involved in virtually all aspects of cellular development and maintenance. By virtue of their core role in the cell, ribosomes share a highly evolutionarily conserved core that carries out the fundamental processes of protein synthesis [1]. However, outside of this core, ribosome composition varies considerably. The main differences among eukaryotic ribosomes are due to ribosomal RNA (rRNA) expansion segments (ESs) and variations of ribosomal proteins (r-proteins) [1, 2]. Further, rRNA fragmentation occurring in regions of high variability has been reported in several organisms from bacteria to protozoa, insects, helminths, fish, and surprisingly mammals [3-15]. Recently, the naked mole-rat (*Heterocephalus glaber*) was discovered to have unusual cleavage sites in its 28S rRNA resulting in the deletion of the major part of the D6 variable region (ES15L) and leaving the two

rRNA fragments disconnected [14]. The cleaved 28S rRNA has been associated with the naked mole rat's increased translational fidelity [14]. The only other known mammals having fragmented rRNA are the tuco-tuco rodent (of the genus *Ctenomys*) and the degu (in the related genus *Octodontomys*) [13]. Here we present the high-resolution structures of the naked mole-rat, tuco-tuco, and guinea pig (*Cavia porcellus*) ribosomes. Guinea pig, which has canonical (non-fragmented) 28S rRNA is used as a rodent model for comparisons to the naked mole-rat and tuco-tuco ribosomes.

During my PhD studies, I also looked at another uniquely fragmented ribosome, that of the protozoan parasite, *Trypanosoma cruzi* (*T. cruzi*), the causative agent of Chagas disease. The *T. cruzi* large subunit rRNA is assembled from 8 pieces-5S, 5.8S, and six pieces forming jointly the 28S rRNA. Together with my colleagues from Joachim Frank's and Liang Tong's research groups, we solved the structure of the *T. Cruzi* large subunit and identified distinctive trypanosome interactions, which allowed us to propose a tentative model for assembly of the large ribosomal subunit [16, 17].

Table of Contents

Acknowledgments

This thesis would not have been accomplished without the support and encouragement of a large number of people. First and foremost, I would like to express my sincere thanks to Dr. Joachim Frank, my graduate advisor, for his guidance and mentorship throughout my studies at Columbia University. He is a dedicated scientist that has patience for his students and grants many opportunities to learn and interact with the cryo-EM community. It has been an immense privilege to work under Dr. Joachim Frank's mentorship.

I would also like to express my deep gratitude to all my committee members: Dr. Ruben L. Gonzalez, Dr. Liang Tong, Dr. Vera Gorbunova and Dr. Israel S. Fernandez, for their help and discussions on my work as well as their continuous support throughout my Ph.D. studies.

I owe much of my PhD to Dr. Zheng Liu, who directly mentored me, I am immensely grateful for his training of the single-particle cryo-EM pipeline from grid preparation to image processing. I would also like to thank Martina Nieß for her work on model building. Further I would like to thank Bob Grassucci for his training on the FEI Tecnai microscopes, Harry Kao for data storage and hardware support, and Zhonghe Ke from Vera Gorbunova's lab for providing samples for the naked mole-rat project. I also would like to thank all current and former members of the Frank lab for their helpful discussions and support. I would also like to thank the National Science Foundation (NSF GRFP: DGE - 1644869) for their support in my graduate studies. Studies were also supported by NIH grant R01GM29169 (to J.F.).

On a more personal note, I would like to thank my sisters- Dominique Gutierrez and Stephanie K. Gutierrez. And last but not least, I would like to thank my mother and father, I am more grateful than I can express for their enduring love and support.

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Preface

This thesis will present the methods and work carried out during my PhD studies, **Chapter 1** addresses the basic principles of single-particle cryo-EM including recent technological advances in hardware and the software implemented in pre-processing and image-processing pipelines. **Chapter 2** provides an overview of the basic structure and functions of the ribosome. **Chapter 3** presents and compares the high-resolution reconstructions of the naked mole-rat, tuco-tuco, and guinea pig ribosomes. **Chapter 4** focuses on the reconstructions of two conformational states from the naked mole-rat ribosome dataset and provides an overview of the observed dynamics. The work presented in Chapters 3 and 4 remains to be finalized for publication. Atomic coordinates obtained from the naked mole-rat, tuco-tuco, and guinea pig ribosomes will be validated and deposited within two months of submission of this dissertation. For a complete record of my graduate work, **Chapter 5** documents the published collaborative work, Structure and assembly model for the *Trypanosoma cruzi* 60S ribosomal subunit [16]. My contributions included the purification of the *T. cruzi* ribosome, structural analysis and interpretation, as well as making figures and writing the manuscript resulting in a co-first authorship [16, 17]. **Chapter 6** documents a review article I coauthored, detailing cryo-EM procedures for high-resolution structure determination entitled, Determination of the ribosome structure to a resolution of 2.5 Å by single-particle cryo-EM [16, 17]. And **Chapter 7** documents the book chapter, New Insights into Ribosome Structure and Function to which I contributed to the Parasitic Protozoan subsection for a Cold Spring Harbor Laboratory (CSHL) Translational Control publication [18].

Chapter 1: An introduction to cryo-electron microscopy and single-particle reconstruction

Structural biology is the study of the molecular structure and dynamics of biological macromolecules, allowing us to gain valuable insights into their function and roles in cellular processes. The main techniques (Figure 1) used in structural biology are x-ray crystallography (XRC), nuclear magnetic resonance (NMR), and cryo-electron microscopy (cryo-EM). XRC is an extremely powerful technique for determining the structure of individual proteins, for many years this has been the method of choice for structural studies. However XRC requires proteins to be crystalized which in many cases can be very difficult or impossible to achieve and may require considerable protein engineering and/or sample quantity to identify suitable conditions for crystallization. Further, crystal packing may lead to steric constraints, limiting the functional states that can be studied. NMR is a structural technique with the advantage of analyzing proteins in solution however large amounts of pure sample are needed to achieve a reasonable signal to noise level and only the structures of small proteins (~ 50 kDa) can typically be studied as it remains very difficult to interpret the NMR spectrum of large biomolecules.

The term Cryo-EM can refer to three very different yet closely related techniques whose common principle is the imaging of radiation-sensitive specimens in a transmission electron microscope (TEM) under cryogenic conditions: electron crystallography, cryo-electron tomography (CET), and single-particle cryo-EM. Both CET and single-particle cryo-EM overcome the need for crystallization. CET can provide 3D reconstructions of organelles, cells, and macromolecules *in situ*. CET combined with subtomogram averaging (STA), which are

image processing methods similar to those used in single particle analysis (subjecting repetitive molecules from tomograms to averaging and classification), has become a very promising technique to bridge knowledge from molecules to cells. In parallel with improvements in cryo-EM hardware and software, CET's potential has improved vastly but few structures have been solved by CET with STA at resolutions better than 8Å, and there are still great margins for routine sample preparation as well as hardware and software improvements [19].

Single-particle cryo-EM, a technique that has been developing for nearly 4 decades, has recently emerged as a leading method for elucidating macromolecular structures at near-atomic resolution. This includes structures of large macromolecular assemblies and membrane proteins which have been difficult to solve with traditional methods. In single-particle cryo-EM, macromolecules are applied to a grid and rapidly frozen in a thin vitreous ice layer at liquid nitrogen temperature. The grids are then imaged with a transmission electron microscope (TEM). Ideally, random views of particles from many angles are captured. However, images obtained from the TEM are noisy due to the low electron dose required to reduce radiation damage of the biological specimen within acceptable bounds; this results in images with a low signal-to-noise ratio (SNR). In addition, cryo-EM images also suffer from the effect of the microscope's extended point-spread function (PSF) due to defects in the imaging system (see section 1.2.1 for a description of the PSF). Fortunately, thousands of images of identical particles can be averaged to improve the SNR and the effects of the PSF can be deconvoluted allowing correction of inherent TEM defects and the imaging conditions. This chapter provides an overview of single-particle cryo-EM principles and procedures.

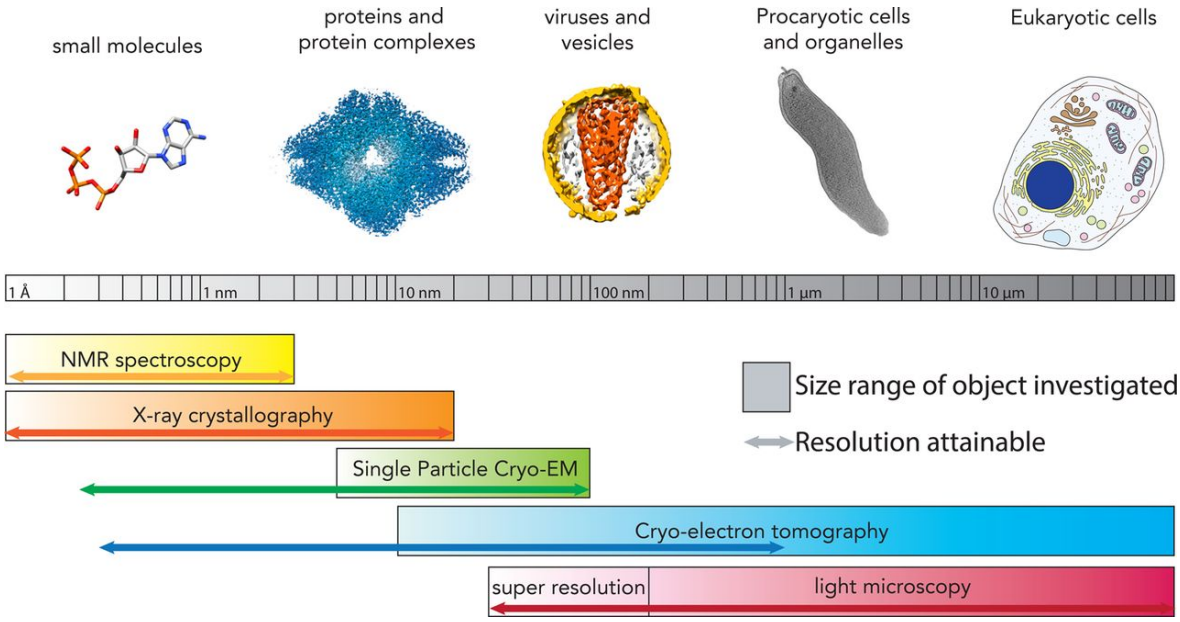

Figure 1. Overview of structural biology techniques and the objects they investigate. Joshua Hutchings and Giulia Zanetti Biochm. Soc. Trans.

1.1 Specimen preparation for single-particle cryo-EM

A single-particle cryo-EM experiment begins with a purified sample in aqueous solution that is applied to an EM grid. Grids are typically 3 mm in diameter and consist of a perforated amorphous carbon or "holey carbon" film supported by a metal frame. These holey carbon grids can be either home-made with irregular holes, or nanofabricated with a regular repeating array of circular holes to facilitate manual and automated EM data collection (e.g. Quantifoil grids from Quantifoil Micro Tools GmbH, and C-flat grids from Protochips Inc). After application of the sample, excessive solution is removed with filter paper under controlled environmental conditions (typically at 4°C and 100% humidity), then the grid is plunged rapidly into liquid ethane (-180°C) precooled with liquid nitrogen (-196°C) . This rapid freezing results in a thin layer of aqueous suspension frozen as amorphous ice, preserving the biological molecules in a native-like state without ice crystallization. The cryogenic temperature of the sample also

reduces the extent of radiation damage to the specimen during data collection. Ideally, the amorphous ice layer is sufficiently thick enough to preserve the integrity of the structure but thin enough so that the particles are still clearly visible. Particles should also be well distributed and adopt a wide range of orientations.

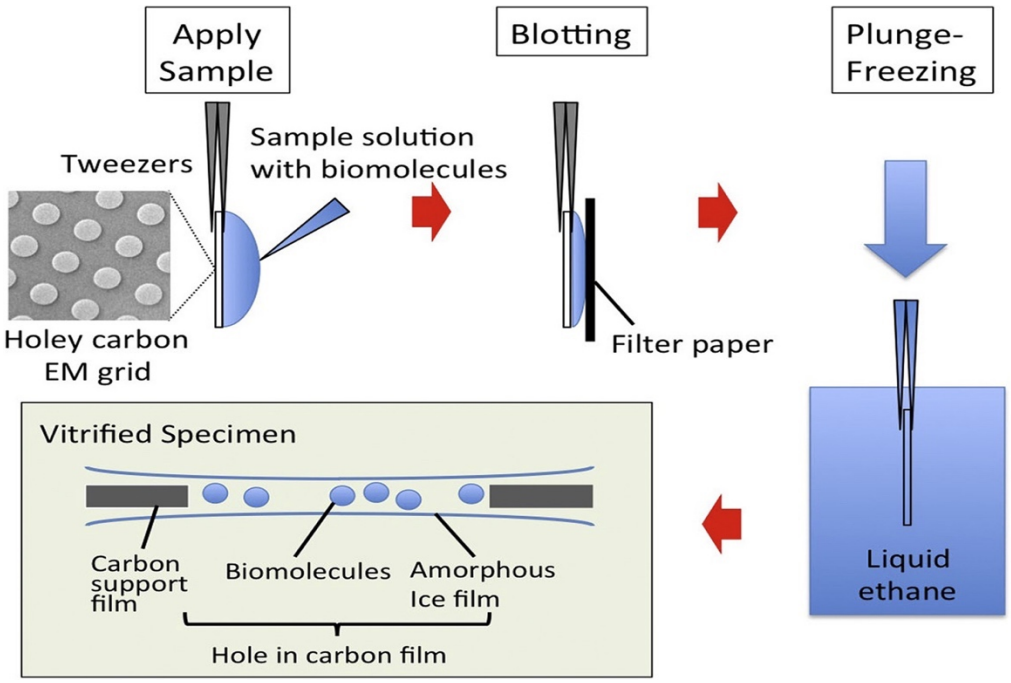

Figure 2. Specimen preparation for single-particle cryo-EM. Aqueous sample solution is applied onto glow-discharged (hydrophilic) holey carbon film. Excess solution is removed with filter paper. Blotted grid is rapidly plunged into a cryogen (ethane) pre-cooled at liquid nitrogen temperature. Biomolecules embedded in thin amorphous ice film are observed in cryo-electron microscope. Kazuyoshi Murata and Matthias Wolf Biochim. et Biophys. Acta.

In practice, obtaining cryo-EM grids of sufficient quality for high-resolution single-particle analysis may require the careful optimization of many variables. The first consideration is in the choice of support and film for the grid. A copper support with amorphous carbon film is

most commonly used as these grids are relatively inexpensive, conductive, and stable in the beam. However, these traditional grids (copper/carbon or sometimes gold/carbon) when cooled to liquid nitrogen temperature result in “cryo-crinkling” where the metal support shrinks more than the carbon film due to the different thermal expansion properties between the metal support and the film [20, 21]. Recently, alternative grid designs have been explored, such as replacing carbon film with graphene [22, 23], and replacing both grid support and carbon film with a single material, gold [21, 24]. Graphene or graphene oxide films have been successfully used to tune particle distribution of various specimens including ß-galactosidase particles [25]. While the use of an all-gold support prevents cryo-crinkling (maintaining rigidity/flatness), it also significantly reduces beam induced movement and image blurring movement compared to conventional carbon foil substrates [24, 26]. Although the problem of specimen movement during imaging had been known, the roles of the support and film on this movement could only be recently characterized, with the advent of direct electron detector cameras.

Another critical step in the sample preparation stage is obtaining a thin, evenly distributed ice layer. The ice thickness should be as thin as possible to achieve high contrast between the particles and the surrounding ice layer while still thick enough to accommodate the particles. A sufficiently thin layer will also minimize defocus spread that arises from particles found at different heights in the ice layer. Key parameters that affect ice thickness are the blotting time and the subsequent time the grid is allowed to dry before it is plunge frozen. These parameters can be controlled with commercially available plungers, such as the Vitrobot (FEI) and Cryo-plunge (Gatan). Typically a few grids are first prepared to screen for ice thickness and particle distribution on the microscope, a heuristic approach is then employed by the user to identify optimal parameters for data collection. Another factor that affects ice thickness is the

hydrophilicity of the support film, which can be optimized with glow-discharge parameters. To homogenize ice thickness and increase particle concentration, the grid may also be coated with an additional layer of thin amorphous carbon (~ 5 nm). For this, a strip of mica and grids are placed inside the bell jar of an evaporator apparatus with a carbon rod suspended above. The bell jar is then placed under vacuum and a voltage is applied across the carbon rod in order to evaporate and deposit a small amount of carbon onto the mica and grids. The carbon deposited onto the mica, which characteristically has a smooth surface, is then floated onto the grids in a water chamber apparatus. Once dry, carbon-coated grids undergo glow discharging, in which they are placed under vacuum and exposed to plasma—typically of hydrogen and oxygen—this removes contaminants from the surface of the grid that may impede its hydrophilicity. Hydrophilic grid surfaces allow the sample to evenly spread, whereas hydrophobic surfaces induce high-curvature droplets, resulting in undesirable and uneven sample thickness. Once the particle density and ice thickness/quality has been assessed and optimized a data collection can be performed.

1.2 Imaging and data acquisition

1.2.1 Image formation in the transmission electron microscope (TEM)

A typical TEM consists of an electron source, a series of electromagnetic lenses, electron deflection coils, stigmator coils, and apertures all operating under high vacuum. For convenience, the TEM can be divided into four essential components- the illumination system, the objective lens/stage, the imaging system, and the image recording system. *The illumination system* comprises the electron source and condenser lenses operated in parallel beam mode. A widely used *electron source* in an EM is a field emission gun (FEG) with a tungsten crystal and

two anodes, the first anode provides the extraction voltage to pull electrons out of the tip of tungsten crystal and the second anode accelerates the electrons in the range of 100—300 kilovolts (kV), this electron beam then emerges from the gun through the EM column. The emerging electron beam is condensed into a nearly parallel beam by the *condenser lenses*, situated below this are then the *condenser lens apertures*, which remove electrons with a large angular spread, improving beam coherence before reaching the specimen.

Electrons in the beam that encounter the specimen may become scattered without loss of energy (elastic scattering), scattered with loss of energy (inelastic scattering), or they may pass through the specimen unimpeded (no interaction with an atom). Those electrons that are elastically scattered (by the Coulomb potential of the specimen) contain the high-resolution structural information. The objective lens then focuses the beam that comes through the sample and initially magnifies the image. The objective lens aperture, situated below the specimen, excludes widely scattered electrons thus increasing amplitude contrast of the final image. Below the objective lens/stage components, the intermediate and projector lens are found which are used to further magnify the image before the electrons arrive to the image recording system. The recorded EM images are effectively 2D projections or line integrals of the 3D Coulomb potential distribution of the specimen convoluted with the point-spread function (PSF) of the microscope. This convolution (Image = PSF $*$ Object) mathematically describes the relationship between the specimen and its image where the PSF represents the blurring of a point source/object through the optical system. Blurring is a result of inherent TEM defects and the imaging conditions (i.e. lens aberrations). The PSF is not determined directly from the imaging system, instead the blurred image is used to estimate blur parameters (i.e. astigmatism and defocus). Estimation of

these parameters, which allows elimination of blurring and restoration of the image, is performed during the pre-processing steps (see section 1.3.2).

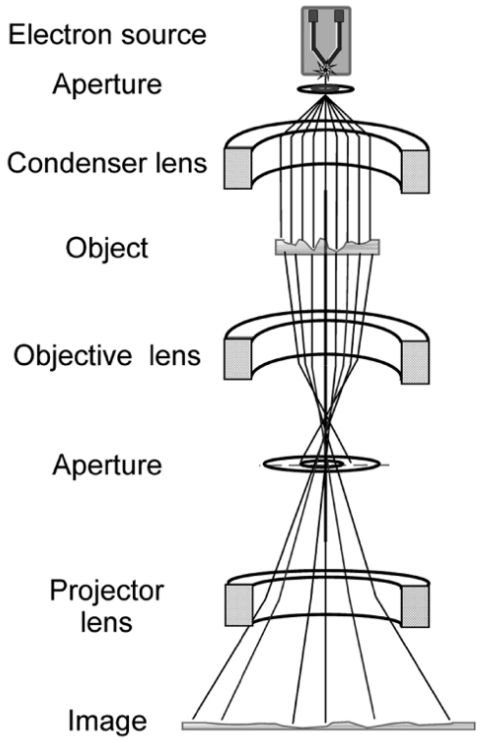

Figure 3. Image formation in the TEM. Elena V. Orlova and Helen R. Saibil Chem. Rev.

1.2.2 Electron beam-specimen interactions

The image formed in the electron microscope is a result of the various electron-beam specimen interactions that occur (Figure 4). The majority of incident electrons pass through the specimen unimpeded as most matter is empty space. Of the interacting electrons, the vast majority are deflected by the electron cloud of a specimen's atom at a specific scattering angle with no energy loss (elastic scattering). A small proportion of the electron beam may directly collide with electrons or the nucleus, resulting in energy loss or transfer of energy to the

specimen (inelastic scattering). Inelastically scattered electrons make no contribution to high-resolution images. As for inelastic scattering events resulting in energy transfer, these may cause ionization of atoms, X-ray emission, chemical bond rearrangement, or induce secondary electron scattering, all of which result in changes to the specimen structure (radiation damage). In order to minimize this damage, a low electron dose is necessary for data collection; however, this results in images with very low signal-to-noise ratio (SNR), requiring many images of similar particles to be combined to determine the 3D reconstruction.

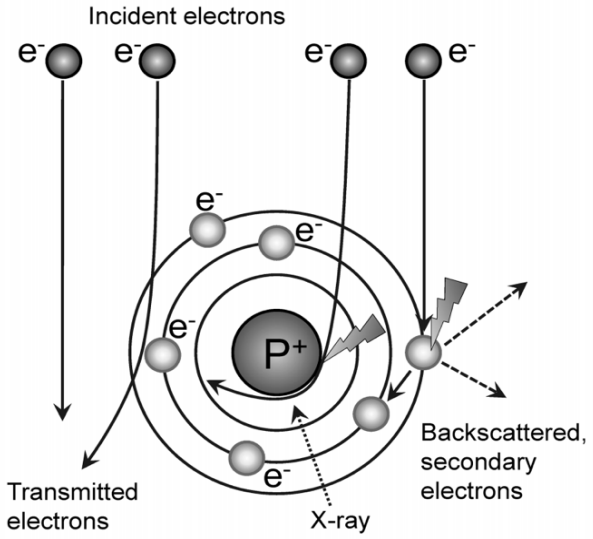

Figure 4. Interactions of electrons with the specimen. Elena V. Orlova and Helen R. Saibil Chem. Rev.

1.2.3 Phase contrast

The electron waves passing through a specimen in the TEM are modulated in their amplitude and phase, resulting in amplitude contrast and phase contrast, respectively. Amplitude contrast arises primarily from removal, by the objective aperture, of electrons elastically scattered into wide angles (for high Z-number elements). Imaging of biomacromolecules and/or

thin biological specimens which are usually composed of low Z-number elements (e.g. C, H, N, and O) results in very low amplitude contrast. Contrast for such specimens (phase contrast) arises primarily from phase shifts exp(iφ) of the elastically scattered electrons during their transition through the sample and is manifested by interference between the exit wave of these electrons and the transmitted unscattered electrons. In the weak-phase approximation, exp(iφ) ~ 1+iφ, which is a wave shifted by $\frac{\pi}{2}$ relative to the unscattered wave. Observable phase contrast relies on the introduction of an additional phase shift (ideally +/- $\frac{\pi}{2}$) [27]. Conventionally, this is achieved by defocusing the objective lens, whereby images are acquired slightly out of focus (-1.0 μm to -3.0 μm). This approach (defocus phase contrast), however, results in a spatial frequency-dependent distortion or loss of information. Recently, phase plates (e.g. the Volta phase plate) have been developed, that introduce an additional phase shift to generate sufficient contrast and that allow imaging at (or close to) focus and thereby enable retaining the undistorted information across a wide band of spatial frequencies [28].

The two contributors to phase contrast are the wave aberrations that occur due to objective lens imperfections (with the main component being spherical aberration) and the defocus at which an image is collected [27]. The way the information is transferred from object projection to image is expressed by the contrast transfer function (CTF). For a given electron microscope with a given defocus setting the CTF describes the relationship between the Fourier transforms of the image contrast and projected object potential:

$$\boldsymbol{T(k) = -\sin\left[\frac{\pi}{2} C_s \lambda^3 k^4 + \pi \Delta f \lambda k^2\right]},$$

Equation 1

where $\boldsymbol{C_s}$ is spherical aberration coefficient (a value reflecting the quality of the objective lens), $\boldsymbol{\lambda}$ is the wave-length defined by accelerating voltage, $\boldsymbol{\Delta f}$ **is** the defocus value,

and $\boldsymbol{k}$ is the spatial frequency. The CTF thus describes the contrast in the image as a function of spatial frequency and accounts for aberrations in a TEM.

Equation 1 above, represents the sine-type CTF for conventional TEM (defocus phase contrast) where low spatial frequency components are greatly suppressed, generating images with overall low contrast for a particle. Users must collect images at a range of defoci rather than at a single defocus value in order to obtain a complete complement of phase contrast information over a wide range of spatial frequencies. This allows recovery of phase contrast information at points where the CTF for a single defocus crosses the x-axis which occurs several times, representing zero information transfer. Collecting images at a whole range of defoci varies the x-coordinates of these crossings, thus providing full converge of phase contrast information across the entire range of spatial frequencies (Figure 5).

For phase plates, the CTF is a cosine-type function with enhanced contrast of low frequency components. From a theoretical point of view, imaging with a phase plate (e.g. Volta phase plate) at zero focus is ideal but has a few practical disadvantages- it requires very accurate focusing and is limited by the spherical aberration of the objective lens to ~3 Å [29]. Thus currently, a small amount of defocus is applied when imaging with phase plates to simplify the data acquisition (accuracy of focusing) and to solve the resolution limit issue (i.e. spherical aberration) by enabling CTF fitting and correction (as in defocus phase contrast, section 1.3.2).

An ideal CTF would result in all of the information for all the components being transferred to the image; however in practice, the CTF is dampened for increasing spatial frequency by the loss of coherence resulting from chromatic aberrations, focal, and energy spread (temporal partial coherency) or from the finite source size (spatial partial coherency). This results in a CTF function modified by envelope functions which can be represented as:

$$\boldsymbol{T(k)_{eff} = T(k)E_cE_a},$$

Equation 2

where $\boldsymbol{E_c}$ is the temporal partial coherency envelope function and $\boldsymbol{E_a}$ is the spatial partial coherency envelope function. The effect of spatial partial coherency can be significantly reduced by a smaller illuminating aperture (condensor lens 3, or C3) found in dedicated high-resolution TEM. For example, the Titan Krios microscope has a three-condensor system, where C1 controls the spot size of the beam, C2 controls the beam intensity and C3 allows the convergence angle of the beam to be adjusted so that the specimen can be illuminated with a parallel beam. Furthermore, the effect of total beam tilt which introduces phase shift to the images can be completely eliminated by use of a spherical aberration (Cs) corrector, leaving only the partial temporal coherence contributing to dampening of the CTF [30]. Correction of the modified CTF is required to obtain the true image and is critical for high-resolution reconstruction.

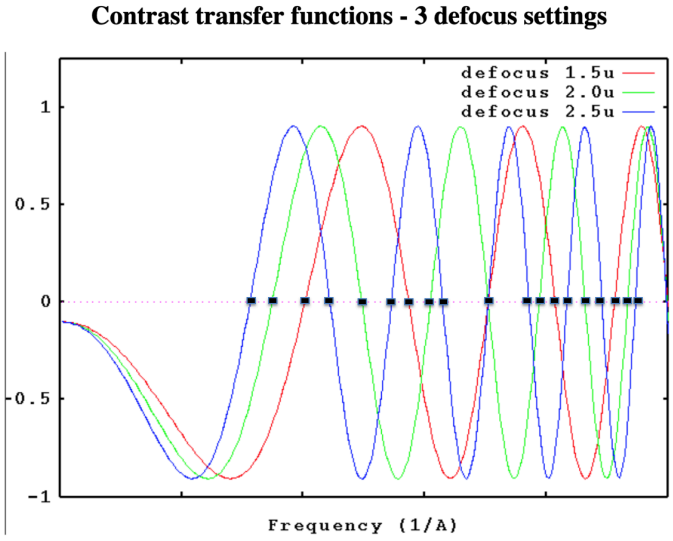

Figure 5. CTFs - 3 defocus settings. A range of defoci varies the location of zero crossings (marked with black box), thus providing full converge of phase contrast information across the entire range of spatial frequencies. Adapted from **https://spider.wadsworth.org/spider_doc/spider/docs/techs/ctf/ctf.html**

1.2.4 Lens aberrations

In a perfect lens electrons that are parallel to the principle axis of the lens will converge at the focal point, F. However, electromagnetic lenses (e.g. the objective lens) bear several defects such as –spherical aberration, chromatic aberration, and astigmatism which limit microscope performance (Figure 6). ***Spherical aberration*** is most significant in defining the performance of the objective lens and is due to the dependence of the focal length on the distance of the ray from the optical axis. The further off-axis the electron beam is, the more strongly it is bent back toward the axis and results in the image of a point becoming a blurred disk. As expressed in equation 1, for T(k), underfocusing the lens can partially compensate the effect of this spherical aberration.

Chromatic aberration arises from variations in the electrons energy (elastic scattering, voltage fluctuation, etc) and results in electrons of different wavelengths being focused closer or father way from the image. ***Axial astigmatism*** can be understood as an "unroundness" of the lens where the lens is stronger in one direction than the perpendicular direction, resulting in the image of a point becoming stretched or the shape of an ellipse. Fortunately, axial astigmatism can be compensated with special corrector elements (stigmators) that are present in both the illumination system (condenser lenses) and in the imaging system (objective lens).

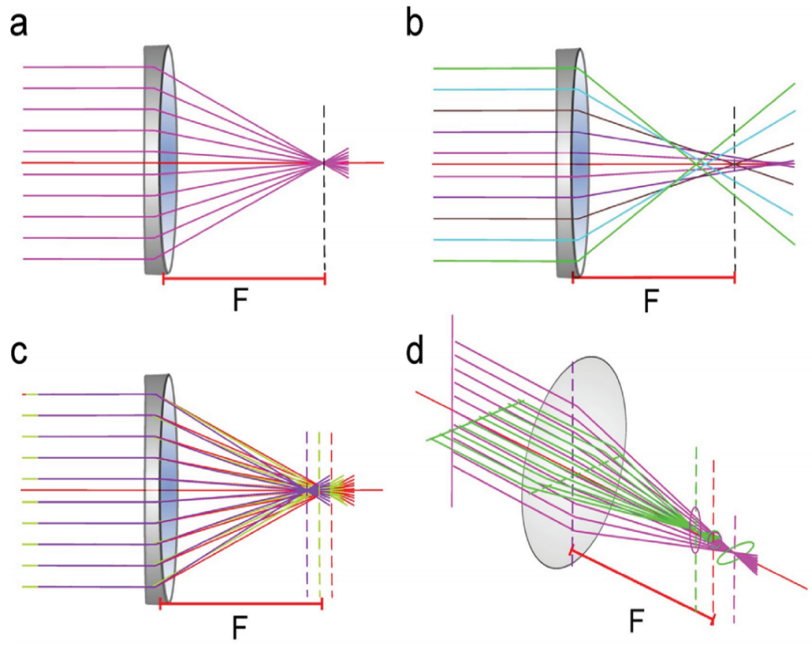

Figure 6. Ray diagrams showing aberrations in lenses. (A) Example of a perfect lens (B) Spherical (C) Chromatic and (D) Astigmatic aberrations. Elena V. Orlova and Helen R. Saibil Chem. Rev.

1.2.5 Projection slice-theorem

Several methods are available to determine the 3D reconstruction from its 2D projection images [27], independent of whether the reconstruction is performed in real space or Fourier space, the projection slice-theorem is fundamental to all 3D reconstruction algorithms. The theorem states that in Fourier space, each 2D projection image represents a central slice of the 3D Fourier transform of the object. It follows from the theorem that one can obtain a reconstruction by a 3D inverse Fourier transform from the Fourier domain which is filled by the 2D Fourier slices [27, 31]

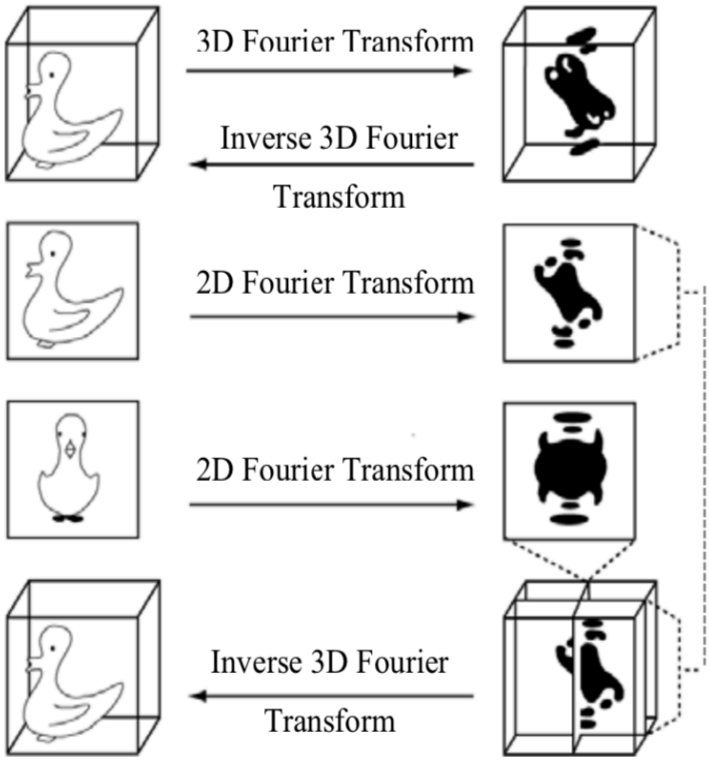

Figure 7. Projection slice theorem and its use in 3D reconstruction. Joachim Frank Three-dimensional electron microscopy of macromolecular assemblies: visualization of biological molecules in their native state. 2006, New York: Oxford University Press

1.2.6 Detectors

Recent advances in single-particle cryo-EM were largely facilitated by direct electron detector devices (DDD) which have significantly higher detective quantum efficiency (DQE) at high frequency compared to previous recording devices. The DQE is defined by the ratio of the squared output SNR to the squared input SNR of the imaging detector, thus it describes how much the detector degrades the original signal in the image.

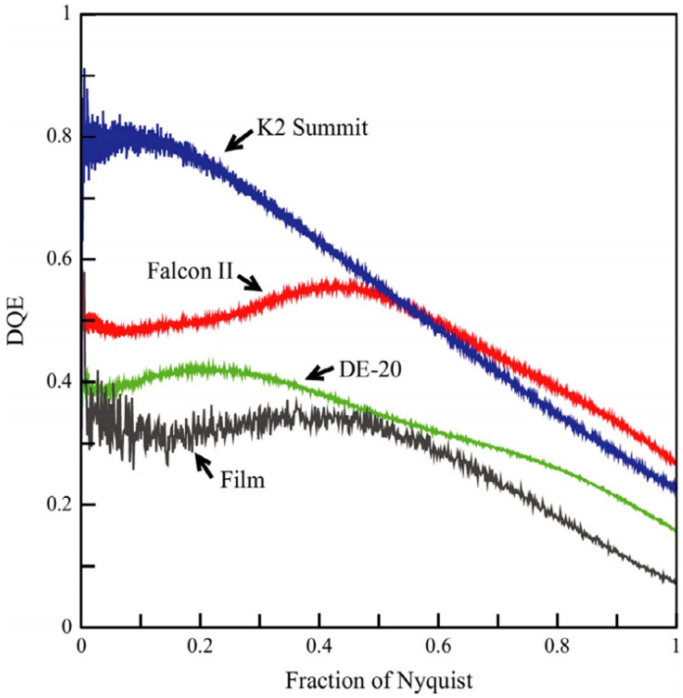

Figure 8. Measured DQE as a function of spatial frequency for detectors. The DQE is the ratio of input to output SNR, where a DQE of unity (DQE=1) implies a perfect detector adding no noise. Recent direct detector devices have nearly doubled the low-resolution DQE and have significantly higher DQE at high frequency. G. McMullan et al. 2014 Ultramicroscopy.

Traditionally, cryo-EM images of biological molecules were recorded on photographic film or charge-coupled devices in a given exposure time. Photographic film is comprised of an emulsion containing microscopic sized grains of light-sensitive silver halide crystals, which can absorb electrons and be subsequently ionized forming a latent image. This recorded image can then be chemically developed, making it visible to the eye and ready for further digitization. Film has the advantage of an extremely small pixel size (the grain size) and large image detection area however recording large data sets is inconvenient and time consuming. And the data is not immediately available requiring development of the film in the darkroom and digitization of optical densities by a scanner.

The charge-coupled device (CCD) was the first digital detector available, enabling automated data collection. CCD cameras typically use a phosphor coated scintillator to convert primary electrons to photons, before being detected by the imaging sensor. This conversion is essential as CCDs are destroyed by direct illumination with electrons. This imaging sensor is made up of pixels, each of which is a metal-oxide semiconductor (MOS) capacitor which then converts these photons to electrons as they strike each pixel due to the photoelectric effect. These electrons are stored in the pixel's capacitor and readout by a "charge-coupled" mechanism in which charges are transferred between neighboring pixels to a readout register, and then amplified and converted to a digital signal. The main problem of CCDs in TEM applications is the signal conversion process which degrades the spatial resolution and reduces sensitivity as light undergoes multiple scattering within the scintillator or at the optical interfaces [32].

Detectors avoiding intermediate light conversions were recently developed, these direct electron detector devices (DDD) can be directly exposed to the high energy electron beam. Popular commercially available DDD cameras are the K2 Summit from Gatan, Inc. (Pleasanton, CA), and Falcon camera from FEI (Hillsboro, OR). DDD cameras may be operated in integration or counting mode. In integration mode, charges generated by a single primary electron event distribute over a small region (PSF), generally a few pixels. Similar integration is used by indirect detection cameras (e.g. CCDs), however scattering introduced in the scintillator worsens the PSF.

In counting mode individual electron events are identified at the time they reach the detector, further reducing the PSF to a single pixel and removing read noise or the variation in the signal that is dependent on the electron interaction with the detector. Counting mode improves the DQE, particularly at low resolution but requires a relatively low electron dose rate

and higher frame rate to minimize coincidence loss (two primary electron events overlap resulting in only one event being counted). The K2 also has a super-resolution mode available which allows an individual electron event to be localized (centroiding) to sub-pixel precision (a quarter pixel).

DDD cameras also enable the recording of dose-fractionated image stacks (movies) due to their high frame rate. This allows for correction of drift and beam-induced motion and provides a way to deal with rapid radiation damage where later frames may be removed or down-weighted during image-processing.

1.3 Pre-processing pipeline

1.3.1 Frame alignment

Raw cryo-EM images are first subjected to pre-processing steps- frame alignment, CTF correction, and particle selection. Whole frame alignment corrects image blurring caused by specimen drift during data acquisition. Motioncorr [33], MotionCor2 [34] and Unblur [35] can perform global, whole frame alignment. Beam-induced motion whether caused by movement of the particles or localized deformation of the ice results in non-uniform local motion that vary across the particles or image. Available beam-induced motion correction software can be divided into those that perform per-particle motion correction or patch-level correction. Per-particle correction also termed “particle polishing”, is implemented in RELION [36], Unblur, and alignparts_LMBFGS [37] and essentially involves calculating smooth estimates of particle trajectories while weighing averages of nearby trajectories (ensuring local correlation). Per-patch correction implemented in MotionCor2 [34] divides the frame into patches (subframes) where motions within each patch are then iteratively determined, subsequently each subframe is

remapped at each individual pixel. The latter is less computationally intensive and allows for local motion correction at preprocessing steps. Early frames may also be removed as strong beam-induced movement has been observed during initial sample irradiation (1- 2 e/Å of exposure), making the first movie frames unusable. These software also implement dose weighting to account for increasing damage in later frames. This allows collection of movies having a higher total dose to improve the SNR (important for particle picking and orientation determination) as well as the preferential selection of high-resolution information from the remaining early frames having less accumulated radiation damage.

1.3.2 CTF correction

Cryo-EM images are extremely noisy and suffer from effects of the CTF of the microscope, The parameters of the CTF must be first determined so that the images can be corrected to recover undistorted information buried in the noise [27]. The CTFs modulate the Fourier transform of the true projections in a defocus-dependent way where the CTFs oscillate rapidly (with many zero crossings) and decay in the high frequency domain at defocus values typically used for image acquisition. To handle the CTF effects, information from multiple images (of different defocus groups) is combined to recover the information for reconstruction.

To estimate the CTF parameters, CTF correction software typically fit CTFs to the Thon rings visible in the power spectrum of an entire electron micrograph to measure defocus and astigmatism. Recently, with the increase of high-resolution maps more localized approaches have also been implemented: stage tilt, uneven ice, or a warped supporting film can all lead to defocus variations among particles within a single cryo-EM image. Software such as CTFFIND

[38] and GCTF [39] have now incorporated both global and local (per-particle) CTF correction to improve estimation of the CTF parameters for all particles for subsequent image processing.

1.3.3 Particle selection and extraction

To acquire individual images for further processing, particle are semi-automatically or automatically selected and extracted from the image. The particle images are also normalized, where the average pixel intensity outside of the particle region is subtracted from the particle image to bring all particles to the same intensity level, independent of experimental factors (variations in ice thickness, exposure, etc.). These procedures are available in several packages including EMAN [40], RELION [41], and Gautomatch (unpublished, http://www.mrc-lmb.cam.ac.uk/kzhang/).

1.4 Determination of particle orientations

The orientations of the particles need to be known to determine the 3D reconstruction, however in single-particle cryo-EM the particles are randomly distributed with unknown orientations - this presents a chicken-and-egg problem of determining the 3D structure without knowing the orientations and vice versa [42]. The approach used in major software packages is to start with an initial guess of the map (alternatively *ab initio* methods are available [43, 44]), here a filtered or low-resolution 3D map of a similar object (e.g. a ribosome) is given to assign projection vectors (reference projections) to all particles. This is termed, projection-matching where a single-particle image (experimental projection) is compared with each reference projection and the projection vector (orientation) of the best-matching reference is the one assigned to this particle. This generates a new model via back projection and reference

projections are again generated. The single-particle images are now matched to the new reference projections and a new model is generated. This process is iterated a few rounds until assigned orientations assignments cannot be improved. Conventionally one was required to find the unique optimal parameters (in-plane rotations and translation) and determine “hard” orientation assignments. Now a maximum likelihood (ML) approach allows for “soft” orientation assignments in which each rotation is assigned a distribution rather than just the best match. Conceptually, this is as though each image is added into the reconstruction for each possible orientation weighted according to its probability. The ML approach was first introduced to the cryo-EM field by Sigworth for 2D alignment, and the idea was extended for 3D reconstruction [45, 46].

More recently, a variant of the ML approach was introduced with a Bayesian interpretation of cryo-EM structure determination [46]. The main difference between this interpretation and earlier ML approaches is in the introduction of a regularization term in the likelihood function. This regularization imposes prior distributions on model parameters in which the ML optimization target may be augmented with such prior distributions. The resulting regularized optimization algorithm, termed a *maximum posteriori* (MAP) estimation has proven useful for high-resolution reconstruction and 2D or 3D classification of heterogenous data.

The optimization of the regularized likelihood target (MAP estimates of parameters) is performed by the standard expectation-maximization algorithm which is an iterative method involving 2 steps. The first step, the Expectation or E-step (E step: alignment) is where the calculated projections of the structure (determined by the current estimated model parameters) are compared with the experimental images, to calculate probability distributions for the relative orientations of all particles. In the subsequent maximization or M-step (M step: Reconstruction)

one computes model parameters with the expected log-likelihood found in the prior E-step and updates the estimate of the 3D map, which is used for the next E-step. This is guaranteed to increase the likelihood of the 3D map given the data and the prior at every iteration until converging onto the nearest local minimum. The relative contributions of the experimental data and the prior to the reconstruction are dictated by Bayes' law and depend on the noise and signal in the data.

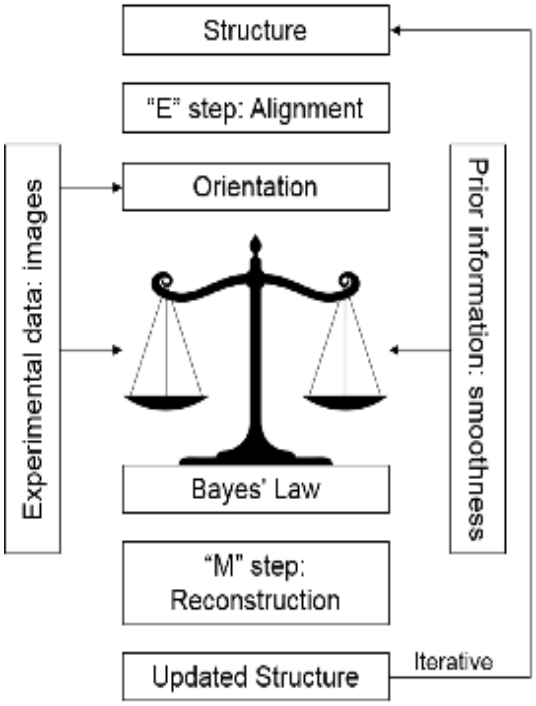

Figure 9. Schematic illustration of the Bayesian approach. S.H Scheres 2012 J Mol Biol

1.5 Image processing pipeline

The low SNR of individual particle images makes it difficult to evaluate the raw projection images, clustering of similar particle images and calculation of class-average images, amplifies the SNR by averaging noisy images of similar view angles [47, 48]. This clustering and calculation of class-average images allows for an initial evaluation and assessment of data quality, such as the angular distribution of the particle views and removal of "bad" particles

After particle extraction, individual particle images are subjected to reference-free 2D class averaging where they are aligned (rotationally and translationally) with one another and classified, As described previously RELION implements a regularized term to the likelihood target function, in 2D classification only orientations are marginalized.

Next 3D classification is performed to continue the selection of suitable particles for high-resolution structure determination. 3D classification begins by initializing multireference refinements from a single, low-resolution reference (40-60 Å) and assigning a random class to each particle in the first iteration. This initial 3D reference may be generated from the 2D particles themselves using EMAN or more recently RELION. Alternatively, maps of similar complexes from EMDB, or maps generated from PDB entries, or in some cases (known point-group symmetry) a spherical blob can be used. The heavy filtering of the initial reference minimizes reference bias and the "Einstein from noise" problem. Apart from this reference, the user provides the number of desired classes (i.e. k=10).

RELION implements a regularized term to the likelihood target function (section 1.4), this allows for efficient simultaneous alignment and classification of the particles without the need for prior knowledge of the differences between the structures in the dataset. In particular without any knowledge of the relative orientations of the particles, in 3D classification one needs to marginalize over both orientational and class assignments. This requires the need to perform exhaustive angular searches to limit the associated computational costs. The first rounds of 3D classification are typically performed using a relatively coarse angular sampling (i.e. 7.5 degrees). This proceeds iteratively where upon convergence estimates of the orientations no longer improve, convergence is monitored by the user in 2D and 3D classification. The angular sampling can be set to a finer increment to improve these estimates, (i.e. 3.75 degrees). Local

angular searches are implemented in later stages after exhaustive searches with coarser samplings have provided orientations relatively close to the correct ones. The particles from the reconstructions deemed suitable are selected for further processing, thus removing the junk particles. Multiple rounds may be performed to remove junk particles and separate particles into homogenous classes which can be refined independent for high-resolution structure determination. This heterogeneity is due to compositional and/or conformational differences present in the sample.

Refinement of the homogenous classes can then be performed. RELION's 3D auto-refinement protocol is similar to its 3D classification protocol however its output is a single volume and it calculates more precise angular assignments. Convergence no longer relies on user innervation but is instead automatically monitored. This is implemented in RELION's 3D auto-refinement procedure by using the estimated angular accuracies to automatically increase the sampling rates during refinement until the noise in the data prevents distinguishing smaller angular differences and the angular assignments can no longer be improved. Further, RELION's 3D auto-refinement procedure implements a so-called gold-standard approach where the data is split into two halves and refined independently. Then the Fourier shell correlation (FSC) between the two independent reconstructions is used to yield a reliable resolution estimate. This approach prevents overfitting and over-estimation of resolution. RELION's 3D auto-refinement procedure may also be used prior to 3D classification to get a consensus model with particles having close to optimal angular assignments. This resulting model can then be used to generate random seeds and classify the dataset, often this gives better results. Subsequent 3D classification can then be performed using only local searches of the angles around the angles from the consensus refinement or alternatively one keeps the orientations from the consensus refinement and only

marginalizes over the class assignments. These approaches can be used for the initial classification of heterogenous data from a consensus model and may also be employed for further classification in which a homogenous class was independently refined and small conformational and/or compositional differences still remain to be separated.

Additional procedures such as masked 3D auto-refinement (focused refinement), masked 3D refinement with partial signal subtraction, and multibody refinement may also be implemented to improve the reconstruction of heterogenous datasets. In masked 3D auto-refinement one applies a 3D mask to the structure to mask out part of the complex and align only with respect to the part of the structure that lies within the mask.

In the signal subtraction approach, one designs two masks. The first mask is the region of interest on which one will perform a masked 3D refinement or a masked classification. The second mask is the entire complex except for region of interest. This second mask is applied to the reconstruction from a consensus refinement, and projections of this masked reconstruction are subtracted from all experimental particle images in the dataset. This creates a new dataset of experimental images with part of the signal subtracted which are used as input images to calculate a 3D reconstruction. This often results in better reconstruction compared to masked 3D-auto refinement, as the part of the complex that lies outside of the mask is removed instead of acting as noise which may deteriorate the angular or class assignments.

The multibody refinement procedure is an automated approach in which user defined regions are subjected in parallel to masked 3D auto-refinement with partial signal subtraction. The masked alignment of the independent regions leads to every particle having optimal orientations assigned for each user defined region. These optimal orientations are used to perform a better signal subtraction in the next iteration. Currently this approach suffers from

strong artifacts at the boundaries between regions but is promising to automatically refine multiple independently moving regions with minimal user intervention.

Following refinement procedures, a RELION postprocessing procedure further improves the quality of the reconstruction. This procedure includes automasking, modulation transfer function (MTF) correction, and B-factor amplitude correction. RELION implements high-resolution noise substitution [49] which consists of randomizing the phases of the half maps (from refinement), masking them, calculating the FSC between them, and comparing that FSC to that of the FSC between two masked, non-randomized half maps.

The MTF describes resolution-limiting effects related to the detector, particularly it describes how much contrast is transferred from the object to the image at each resolution. This can be easily deconvoluted from the map and a suitable MTF curve of the detector provided by the manufacturer.

The B-factor describes contrast loss at high-resolution in a reconstruction. The effective B-factor is used to model the combined effects of molecular drifting due to charging effects, molecular flexibility, errors in the detection process, errors in image processing, etc., into a Gaussian envelope function that describes the signal falloff [50]. This B-factor value may be estimated automatically for reconstructions that extend beyond 10 Å, and applied to restore the high-resolution structure amplitudes. This corrected FSC curve results in a sharpened map with the high-resolution details.

1.6 Resolution

For single-particle cryo-EM reconstructions, resolution is measured in Fourier space as a function of spatial frequency. Specifically, resolution is estimated from the Fourier Shell

Correlation (FSC) between two half-maps (from the refinement), each reconstructed from half of the particles in the dataset. In terms of signal and noise, the half-maps should contain identical signal and correlate very well whereas noise in the half maps should correlate very little across all spatial frequencies. For gold-standard FSCs, the spatial frequency at which the FSC equals 0.143 is the cut-off value used to describe the global resolution of a given map. The highest achievable resolution for a sampled image is limited as described by the Nyquist-Shannon theorem [51]. In the case of cryo-EM data, the resolution is limited by the pixel size (the sampling distance) where its resolution, in real space, cannot exceed the width of two pixels. This resolution ceiling, in Fourier space is termed the Nyquist limit.

1.7 Resolution assessment and validation

The estimated resolution (section 1.6) is not an absolute measurement but rather is akin to a consistency score and does not guarantee that structural details in the map will be reflected in the estimated resolution [52]. This calls for the need to re-evaluate and validate the resolution of a cryo-EM map. Structural features at an a resolution ~4-10Å should reveal secondary structures such as α-helices and the relative arrangements of domains and/or subunits in a complex. At near-atomic resolution <4Å, β-strands allowing polypeptide backbone tracing should be clearly resolved. Recently several tools for resolution evaluation and/or validation have implemented in several software [53-55]. Phenix software (mtriage), has four available metrics – d_FSC, d99, d_model, and d_FSC_model [55]. The d_FSC corresponds to the gold-standard FSC. The d99, is the resolution at which Fourier map coefficients (corresponding to the real space map) can be omitted before the map starts changing significantly. The d_model, and d_FSC_model metrics both require an atomic model to estimate map resolution. The d_model is the resolution cutoff at

which the model map is the most similar to the experimental map. For this metric, a series of model-calculated maps at different resolutions are generated and then each is compared with the experimental map; the resolution of the model-calculated map that maximizes the correlation is considered as the most representative of the map resolution. The d_FSC_model is the resolution cutoff up to which the model and map Fourier coefficients are similar (also a map-model correlation). An FSC curve between the experimental cryo-EM map and model is also calculated. For further validation, half maps are provided for calculation of their respective map-model FSC curves. Ideally, the FSC curve of the full map and the 2 half maps should overlap. Another available metric is the EMRinger score, which describes the local fit of individual amino-acid residue side chains to the map [53].

Chapter 2: Structure and functions of the ribosome

Proteins, encoded by individual genes in DNA, are responsible for nearly every task of cellular life. Thus protein synthesis is crucial for cellular maintenance and development. The production of proteins requires two major processes -transcription and translation. Transcription is the process by which a DNA strand is copied into a type of RNA molecule. In particular genes encoding proteins are transcribed into messenger RNA (mRNA), which carry the information needed for protein synthesis (triplet code or codon). The transcription process is followed by translation whereby the genetic message encoded in the mRNA is translated into a sequence of amino acids. Translation, the final step of gene expression (DNA→ RNA →Protein) is mediated by ribosomes in concert with transfer RNAs (tRNAs) and other proteins. Below is a brief overview on eukaryotic translation:

2.1 Translation

Translation consists of four distinct steps: initiation, elongation, termination, and recycling. Most major control checkpoints are exerted in the initiation stage, before translation of the polypeptide chain begins. Canonical translation, begins with formation of the ternary complex (TC) where GTP-bound eIF2 binds the initiator methionyl-tRNA (tRNAiM). The TC then interacts with the small ribosomal subunit (40S) and a number of initiation factors – eIF1, eIF1A, eIF5, and the eIF3 complex, together forming the 43S preinitiation complex (PIC). The 43S PIC is then recruited to the 7-methylguanosine (m7G) cap (5’ end of mRNA) by eIF4F complex. Mammalian eIF4F consists of eIF4E (the cap-binding protein), eIF4A (an RNA helicase),and eIF4G (a scaffold protein). EIF4G binds PABP at the 3’end of mRNA thus forming a nascent 48S complex. The PIC then scans in search of a start codon and once the anti-codon of

tRNAiMet base-pairs with the start codon in the peptidyl site (P-site), GTP is hydrolyzed by eIF2 (eIF5) and Pi is released. Next, displacement of factors- eIF1, eIF1A, eIF2•GDP, eiF3, and eIF5 and an additional GTP hydrolysis event (eIF5B) allows the 60S ribosomal subunit to join, forming an 80S ribosome.

Following initiation, protein synthesis occurs through progressive elongation cycles [1, 56]. An elongation cycle, can be divided into three principle steps, (1) tRNA selection or decoding, (2) peptide bond formation, and (3) translocation of the mRNA-tRNA complex. In elongation, aa-tRNAs are delivered to the ribosomal A site instead of the P site as in initiation. The aa-tRNA is delivered to the ribosomal A site in the form of a ternary complex (EF1A•GTP•aa-tRNA). Upon a cognate anticodon-codon binding interaction (tRNA selection), a conformational change of EF1A occurs, triggering GTP hydrolysis. This enables the tRNA to become fully accommodated into the A site. At the peptidyl transferase center (PTC) of the large subunit, the acceptor ends of the A- and P-tRNAs are stabilized through interaction with the conserved A- and P- loops of the large subunit rRNA. The amino acid bound to the A-site tRNA can now undergo a favorable peptidyl transfer reaction. The amino acid's free nucleophilic amino group attacks the carboxyl end of the amino acid in the P-site, thereby joining the two amino acids with a peptide bond. This is accompanied by a conformational change, in which the ribosomal subunits rotate with respect to one another (intersubunit rotation) and the tRNAS adopt an altered conformation termed, the 'hybrid state'. Essentially, the anticodon end of the tRNAs remain positioned in the P- and A- sites of the small ribosomal subunit (SSU), while the acceptor ends of the tRNA are positioned in the E- site and P- sites of the large subunit (P/E and A/P). Next, the rotated state of the ribosome binds another translocation GTPase, eEF2, this translocates the mRNA-tRNA complex relative to the ribosome and returns the tRNAs to their

‘classical states’ (E/E and P/P) [57-59]. The tRNA movement into the ‘classical states’ is accompanied by rotation of the small subunit relative to the large subunit (a counterclockwise rotation) and is expected to involve head swiveling to facilitate tRNA movements [59, 60]. The A-site is now free and ready for the next aa-tRNA to bind. This completes a cycle of elongation, which will iterate until the ribosome reaches a stop codon (UAG, UGA, or UAA) in the A site. Termination in eukaryotes requires two proteins -eRF1, responsible for stop codon recognition and eRF3, a translational GTPase which seems to facilitate termination. Recycling then takes place once the polypeptide chain is released. Recycling involves the dissociation of the ribosomal subunits and release of the mRNA and tRNA to regenerate necessary components for subsequent rounds of translation [61].

2.2 Ribosome structure

Ribosomes were first observed in EM images as granules dotting the outer surface of the endoplasmic reticulum (ER) by Palade in 1955 [62]. In 1976, Jim Lake described the main topology of the ribosome and produced a 3D model by interference from a few views [63]. And in 1990, the first cryo-EM reconstruction of the 70S was obtained at 40 Å [64]. By 1995, a cryo-EM reconstruction of the ribosome was resolved at 25 Å allowing for characterization of topological features- bridges joining the two distinct subunits, the binding sites of tRNA, and the peptide exit tunnel [65].

In 2000, the first atomic structures of the 30S subunit from *Thermus thermophilis* [66] and the 50S from *Haloarcula marismortui* [67] were obtained, followed by the complete 70S from *Thermus thermophilis* [68]. These studies revolutionized the field of ribosomal research, providing extensive information on RNA and protein structure, protein-RNA interactions, and

characterizing the functional sites of the two subunits. The atomic structures provided irrefutable evidence that the ribosome was a ribozyme, with the RNA performing the primary functions of decoding and peptide bond formation on the 30S and 50S, respectively.

The 30S subunit consists of 21 ribosomal proteins (r-proteins) and the 16S rRNA molecule (~1500 nucleotides). The secondary structure of the 16S, is made up of over 50 RNA helices, connected by single stranded internal and terminal loops whereby the internal loops are involved in long-range interactions (non-Watson-Crick base pairing) to stabilize the helices and ribosome structure. The stacking and/or packing of the 16S RNA results in three compact domains (the 5' domain, central domain, and the 3' major domain), which correspond to the landmark features of the 30S- the body, platform, and head; and one extended domain (3' minor domain) which consists of two helices at the subunit interface (H45 and H45) and available to interact with the large ribosomal subunit (Figure 10).

The r-proteins are dispersed throughout the structure, asymmetrically and concentrated on the surface of 30S. The r-proteins are located between different RNA elements and stabilize the tertiary structure, which can be indicated by their topology. Globular proteins often bind three- or four- helix junctions tying together distant regions of the RNA. R-proteins which resemble a "screw" in topology have a globular head which remains on the surface of the ribosome and an extension or insertion (of a loop, a C-terminal, or N-terminal) which penetrates the interior of the ribosome. These insertions or extensions are fairly basic and engage in electrostatic interactions, allowing the RNA to adopt a more compact conformation by reducing the negative interactions between the phosphates. R-protein extensions (N or C-terminal) may also be found extending out to the exterior, stabilizing the rRNA.

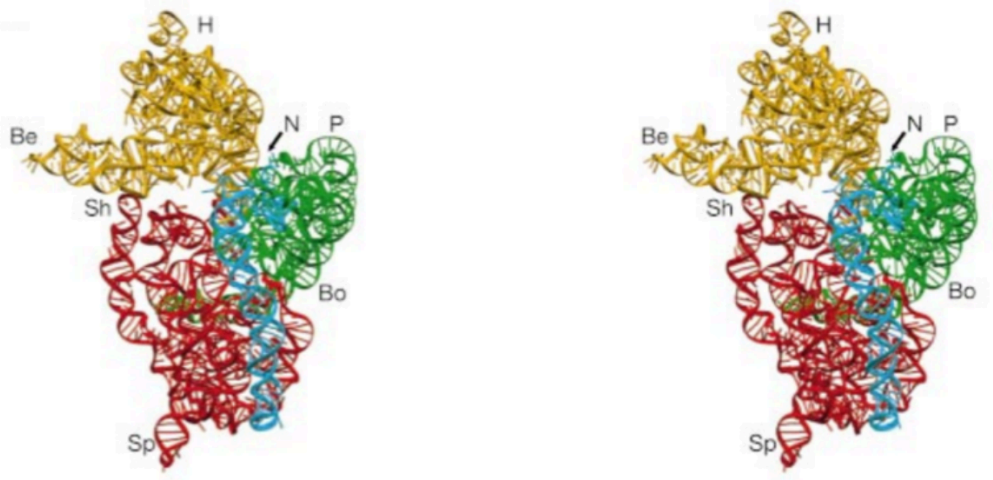

Figure 10. Tertiary structure of the 16S RNA showing the 50S or 'front' view. H head; Be, beak; N, neck; P, platform; Sh shoulder; Sp, spur; Bo, body. B. Wimberly et al. 2000 Nature

The 50S from *Haloarcula marismortui*, is composed 23S rRNA (~3000 nucleotides), 5S rRNA (~120 nucleotides), and 31 r-proteins. The 23S rRNA consists of 6 domains (domain I-VI), each of which has a highly asymmetric tertiary structure. Domain I, lies in the back, behind and below the L1 region. Domain II, accounts for most of the back of the particle and has three protrusions that reach toward the subunit interface side. The first protrusion includes the RNA portion of the L7/L12 stalk (helix 42 and helix 44) which interact with elongation factors. The second protrusion (of domain II) is helix 38, which starts in the back of the particle and extends towards the small subunit (30S) and is found between domain V and the 5S rRNA. The third protrusion (helix 32-helix 35) points directly toward the small subunit. With the loop of stem-loop 34, emerging at the subunit interface between domains III and IV, and interacting directly with the small ribosomal subunit. Domain III, occupies the bottom left region of the large ribosomal subunit in the crown view, and looks like a four-pointed star (stem-loop 48, 52, 57, and 58). This domain unlike all other domains, hardly interacts with domain V (peptidyl

transferase activity). Domain IV accounts for the 50S interface side which contacts the 30S subunit, with helices 67-71 constituting the most prominent features. Domain V, is found between domains IV and II and encompasses the peptidyl transferase activity of the ribosome. Helices 80-88 form the bulk of the central protuberance region and interacts with domain II and the 5S rRNA. Helices 89-93 of this domain extend toward domain VI and help stabilize the elongation factor binding region of the ribosome. Domain VI, found immediately below the L7/L12 stalk consists of the sarcin-ricin loop (SRL or stem-loop 95), essential for factor binding (Figure 11). As for the r-proteins, they appear on the surface of the 50S subunit and are absent from the interface or active site regions. Similarly to the 30S, the r-proteins have N- or C-terminal extensions penetrating the interior of the ribosome and stabilizing the rRNA.

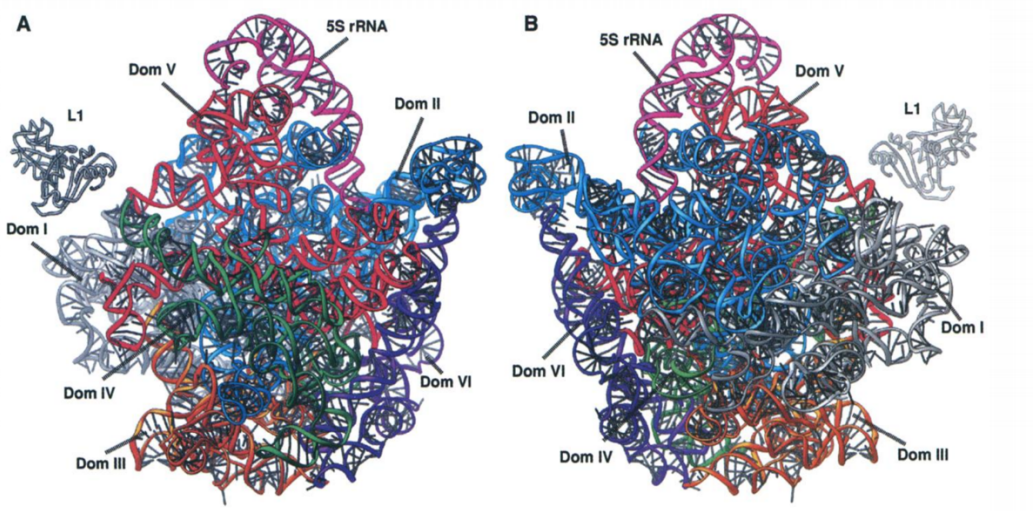

Figure 11. The tertiary structures of the RNA in *H. marismortui* large ribosomal subunit and its domains. Ban et al 2000 Science.

As for eukaryotic ribosomes, the first structures were determined from cryo-EM maps and the fitting of atomic structures from the bacterial and archaeal small and large ribosomal

subunits, respectively [69]. These studies identified 46 eukaryotic r-proteins with bacterial and/or archaeal homologs as well as eukaryotic RNA expansion segments. Subsequent cryo-EM reconstructions enabled the identification of additional r-proteins as well as eukaryotic r-protein extensions [70-72]. In 2011, the atomic structures of *Tetrahymena thermophila* and *Saccharomyces cerevisiae* solved by x-ray crystallography enabled a full assignment of r-protein extensions and rRNA expansion segments in these species [2, 73, 74]. More recently, with the advent of new technology (direct detectors) and software, single-particle cryo-EM has reached high resolution (< 3.5 Å) allowing the study of diverse ribosomes and ribosomal complexes capturing a range of conformational states. This includes ribosomes from a variety of organisms ranging from pathogenic bacteria, yeast, parasitic protozoans, to humans. The increased diversity of ribosome structures reveals species-specific features, providing insights into unique mechanisms of translation and the evolution of ribosomes [75].

Eukaryotic ribosomes are composed of a large ribosomal subunit, the 60S and a small ribosomal subunit, the 40S. The eukaryotic large subunit (60S) is typically composed of 3 RNA molecules- the 5S, 5.8S, and a 28S rRNA. And the eukaryotic small subunit (40S) is composed of a single RNA molecule- the 18S rRNA. Both 28S and 18S rRNA have multiple insertions (with a total length of 900 bases or more) to the core rRNA fold of their prokaryotic counterparts (ESs). And the size of the 80S ribosomes varies within ~ 1-MDA range, mainly owing to insertions in 4 rRNA expansion segments (ESs)- ES7L, ES15L, ES27L, and ES39L [1]. These ESs interrupt the conserved core rRNA along the same sites (variable regions) across different species however their length and actual sequence vary significantly between different species. These ESs are found and constrained to the periphery, where they do not perturb the common core and its functions. And most resemble tentacle-like arms extending outwards and are highly

flexible making it difficult to visualize them. Little is known of ESs biological functions but they accessible from the solvent side for potential interactions [1]. R-protein extensions (N or C-terminal) are also be found extending out to the solvent exterior, stabilizing the expanding rRNA ESs, supporting the idea that they coevolve [71]. Altogether, the majority of changes in eukaryotic ribosomes (rRNA expansion segments and eukaryotic r-protein extensions) compared to their prokaryotic counterparts occur in the outer shell of the ribosome.

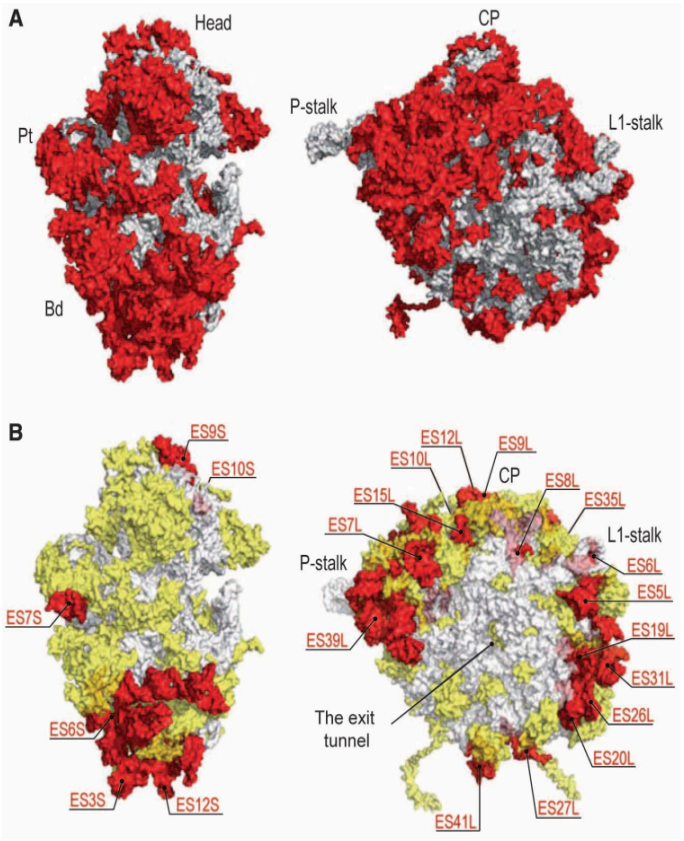

Figure 12. Characteristics of the 80S ribosome surface. (A) Eukaryotic-specific elements (in red) on the solvent side of both subunits. Conserved core elements are in gray. Most of the ribosome surface is eukaryotic-specific. (B) Distribution of eukaryotic-specific moieties (yellow) and rRNA expansion segments (red) around the conserved core. *A. Ben-Shem 2011 Science*

Chapter 3: Single-particle cryo-EM structures of the naked mole-rat, tuco-tuco, and guinea pig ribosomes

Ribosomes share an evolutionarily highly conserved core consisting of the major functional centers -- the decoding site, peptidyl transferase center, and tRNA-binding sites. However, outside of this core, ribosome composition varies considerably. The main differences among eukaryotic ribosomes are due to domain-specific moieties (e.g. insertions or extensions) in their ribosomal proteins (r-proteins) and/or rRNA expansion segments (ESs) [1]. Further, rRNA fragmentation occurring in regions of high variability has been reported in several organisms from bacteria to protozoa, insects, helminths, fish, and, surprisingly, some mammals [3-15]. Recently, the naked mole-rat (*Heterocephalus glaber*) was discovered to have unusual cleavage sites in its 28S rRNA resulting in the deletion of the major part of the D6 variable region and leaving the two rRNA fragments disconnected [14]. The cleaved nature of the 28S rRNA has been suggested to be associated with the naked mole rat's increased translational fidelity [14]. The only other known mammals having fragmented rRNA are the tuco-tuco rodent (of the genus *Ctenomys)* and the degu (in the related genus *Octodontomys*) [13]. Interestingly, the tuco-tuco rodent has also shown to have increased translational fidelity [76]. Here we present the high-resolution structures of the naked mole-rat, tuco-tuco, and guinea pig (*Cavia porcellus*) ribosomes. Guinea pig, which has canonical (non-fragmented) 28S rRNA, is used as a rodent control model for comparisons to the naked mole-rat and tuco-tuco ribosomes.

3.1 Overall architecture of the naked mole-rat, tuco-tuco, and guinea pig large subunits

In general outline, the structures of large ribosomal subunits (60S) from the naked mole-rat and tuco-tuco resemble those of other eukaryotes; particularly, there is a high accordance with the rabbit ribosome, which was used for initial model building (PDB 5LZS [77]). For our studies we elected to compare the fragmented naked mole-rat and tuco-tuco ribosome structures to that of guinea pig, the most closely related organism. The guinea pig ribosome has a conventional 28S rRNA built from a single rRNA molecule. Guinea pig was found to have a lower translational fidelity compared to naked mole-rat and tuco-tuco [76].

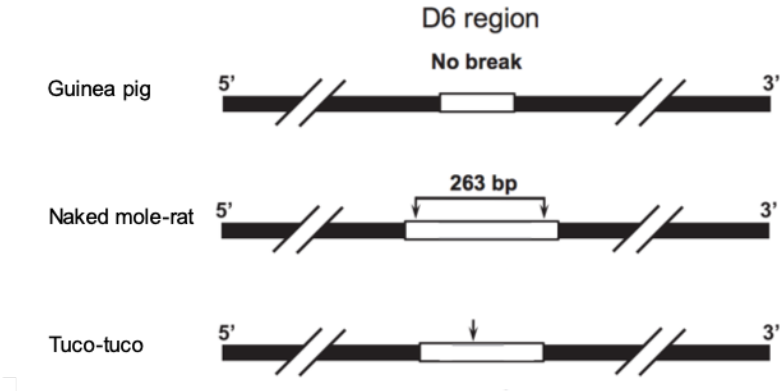

Figure 13. 28S rDNA of the Guinea pig, Naked mole-rat and Tuco-tuco rodent species. The guinea pig 28S molecule has no cleavage sites. The D6 region of the naked mole-rat contains two cleavage sites that excise a 263-nt fragment and results in a 28S fragmented molecule. The D6 region of tuco-tuco contains a single-cut site also resulting in a fragmented 28S molecule.

Fragmentation of 28S rRNA in both the naked mole-rat and tuco-tuco results from cleavage sites corresponding to the 28S rDNA divergent region 6 (D6), which has a high degree of sequence conservation between the two species [13, 78]. However in the naked mole-rat the D6 region contains two cleavage sites that excise a 263-nt fragment resulting in a 28S rRNA

molecule being split into two pieces, whereas in tuco-tuco, the D6 region contains a single cut site within a unique 106-bp insertion and no fragment is excised [14].

In both cases the 28S rRNA of the naked mole-rat and tuco-tuco is processed into two smaller pieces of unequal size that are no longer covalently linked. These two pieces, analogous to the nomenclature for 28S rRNA, will be referred to as LSU-α and LSU-β following a naming convention introduced by Hashem et al [79]. LSU-α is primarily located on the solvent side of the particle with helices H38 (also known as the A-site finger) and H34 inserted into the interface. LSU-β is primarily located in the interface of the particle (Figure 14).

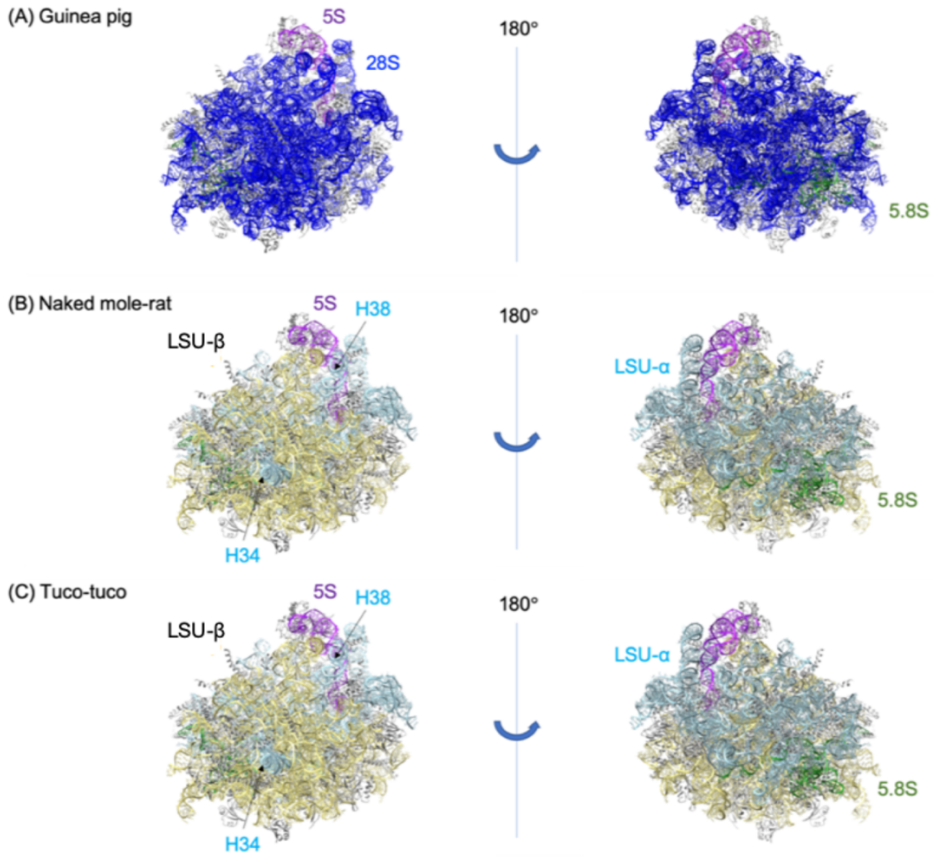

Figure 14. Overview of Guinea pig, Naked mole-rat, and Tuco-tuco large subunit (60S) architecture. (A) Guinea pig (B) Naked mole-rat (C) Tuco-tuco. Left, interface view. Right, solvent view.

3.2 Fragmentation site of the naked mole-rat and tuco-tuco large subunits

The reconstruction of the naked mole-rat large subunit (60S) reveals high-resolution details of the fragmentation site found in ES15L, which caps helix 45 (H45). The LSU- α is sealed by LSU-β and r-protein, uL4. The LSU-β is sealed off by a base-stacking interaction with ES7L from LSU-α. In the unsharpened map a disc-shaped mass of density is observed sealing the gap between the fragments. This mass of density interacts with r-protein uL30 and most likely corresponds to flexible rRNA from ES7L.

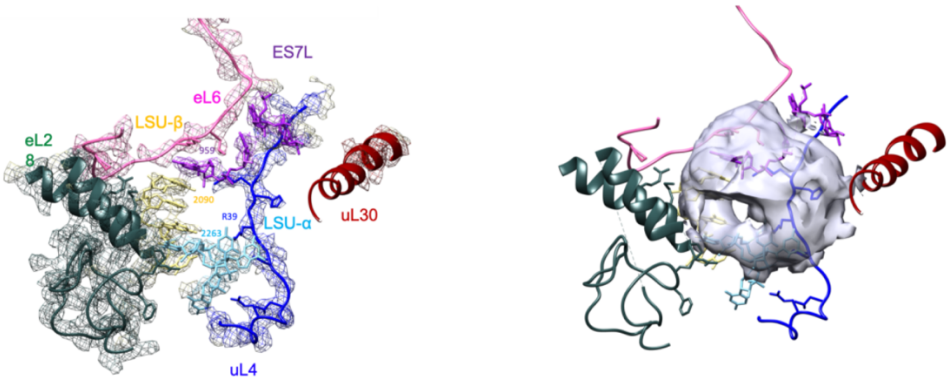

Figure 15. Fragmentation site of the Naked mole-rat ribosome. (Left) postprocess map and model (Right) a disc like density is seen in the refinement map (not modeled) sealing off the fragmentation site.

The tuco-tuco ribosome reveals a similar structure and rRNA-rRNA and rRNA-protein interactions as the naked mole-rat ribosome in this region. However, studies mapping the fragmentation site of tuco-tuco indicate that ES15L is longer as no fragment has been excised. (Figure 13). The density map of the tuco-tuco ribosome does reveal that ES15L further extends to the surface of the large subunit (60S) ribosome; however, this region is apparently highly flexible so that the density disappears as we move toward the periphery. The post-processed map

of the ribosome from tucu-tuco can only visualize the stable region of ES15L, upstream of the fragmentation site. This stable region corresponds to the naked mole-rat fragmentation site.

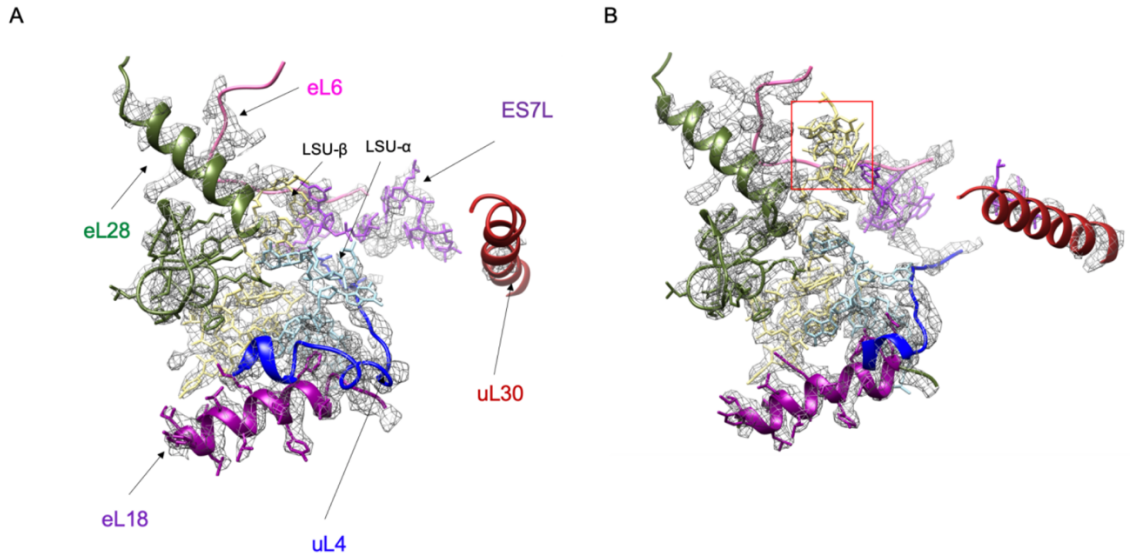

Figure 16. Fragmentation site of the Naked mole-rat ribosome and comparison of the corresponding region in the Tuco-tuco ribosome. (A) Naked mole-rat (B) Tuco-tuco, red box shows density for a portion of the longer ES15L.

3.3 Comparison of the naked mole-rat, tuco-tuco, and guinea pig large subunits

The predominant differences between the three rodent species are all identifiable in flexible regions at low resolution in the refinement maps. By contrast, these flexible regions are invisible in the high-resolution post-process (sharpened) maps

In particular, the largest differences occur on the solvent side and are found among five major expansion segments (ES7L, ES9L, ES15L, ES10L, and ES12L). These expansion segments are found on the top and back of the particle and form a network that extend towards the interface region of the particle (Figure 17A). ES7L is the largest expansion segment, which is found on the top of the particle and made up of three long helices, namely ES7La, ES7Lb, and

ES7Lc. These helices extend to the L1 stalk, the P stalk, and CP regions, respectively. ES7L also merges with ES9L and ES15L (fragmentation site of naked mole-rat and tuco-tuco ribosome) to form a knot-like structure on the back of the large subunit. ES9L starts near this pseudoknot and is found stacked onto ES12L, an insertion of H38 or the A-site finger. ES10L is found behind ES9L and is another insertion of H38.

Three major differences were found between the naked mole-rat and the guinea pig large ribosomal subunit. The first difference is due to the fragmentation of the naked mole-rat ribosome which results in a shorter ES15L that no longer interacts with ES9L on the solvent and back side of the particle. In the guinea pig reconstruction, ES15L is extended and stacks onto ES9L. This stacking is a result of the extended ES15L base-pairing with the terminal loop of H30 within ES9L, as described for the human and rabbit ribosome structures [77, 80]. However the region is not modelled in our guinea pig structure due to its high flexibility. As for the tuco-tuco ribosome, ES15L is longer compared to the naked mole-rat ribosome and extends further out to the surface, where it forms contacts with eL6 and ES7L. And although the naked mole-rat and tuco-tuco are cleaved in a different way, the resulting fragmentations both result in their ES15L no longer interacting with ES9L (H30).

The second difference between the ribosomes from the three rodent species is found in the internal loop of ES10L. In particular, extra density is observed in both the guinea pig and tuco-tuco ribosomes compared to the naked mole-rat. This non-helical insertion of ES10L links the three expansion segments – ES9L, ES10L, and ES12L. In the guinea pig ribosome this cluster of three expansion segments contacts ES15L, forming an extensive network from the top of the particle to the central protuberance (CP) of the ribosome.

The third difference among the three rodents is the appearance of an additional density found on the solvent surface below ES7Lc and above the CP. This density was assigned to r-protein eL29. This assignment was determined from mammalian structural assignments found in the PDB, in particular rabbit [77] and human [81]. The extra density in the guinea pig and tuco-tuco ribosome originates from a c-terminal portion of eL29 which bridges ES7L and ES9L. The naked mole-rat lacks this bridge, suggesting ES7L has increased flexibility compared to tuco-tuco and the guinea pig ribosomes. The C-terminal eL29 density in tuco-tuco is weaker compared to guinea pig, indicating the tuco-tuco ES7L may have intermediate flexibility. In the three rodent structures, an N-terminal portion of eL29 (amino acids 2-76) is found on the solvent surface where it interacts with ES12L and ES10L before tethering into the rRNA core, near the peptidyl transferase center (PTC). Thus eL29 links distal sites (ES10L/ES12L and ES7L/ES9L) spanning the solvent surface in the guinea pig and tuco-tuco ribosomes.

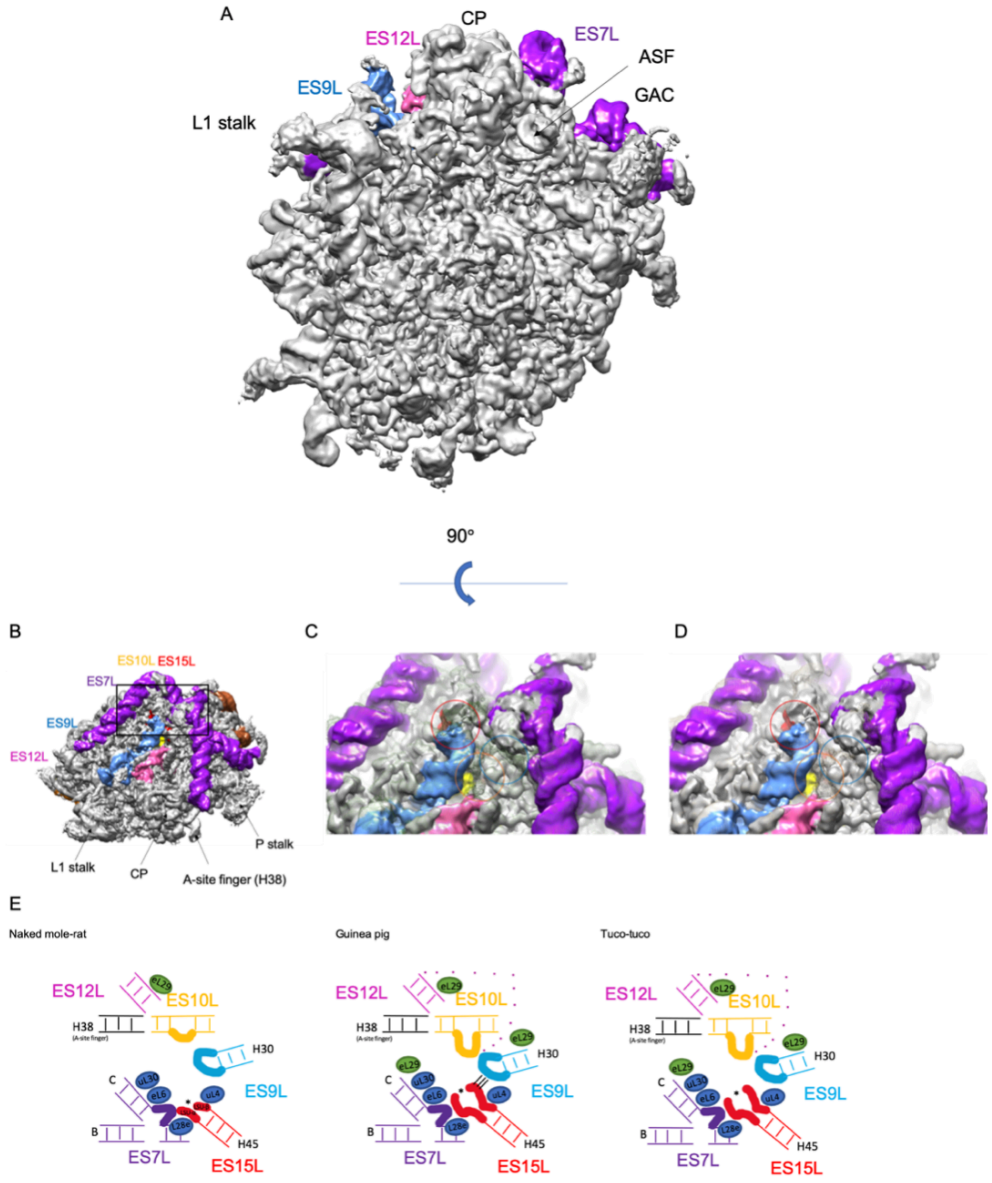

Figure 17. Comparison of the Naked mole-rat large subunit (60S) to the Guinea pig and Tuco-tuco ribosomes. (A) Crown view of the naked mole-rat density with color zone of prominent expansion segments (ESs). (B) Naked mole-rat density is turned on X axis 90° in chimera to show ESs. (C) Overlay of guinea pig density in mesh (D) Overlay of tuco-tuco density in mesh. Three main differences are seen in the density, colored in red (ES15L fragmentation site), orange (ES10L), and blue (eL29). (E) Cartoon representation of the five major ESs (ES15L, ES9L, ES7L, ES10L, ES12) and r-proteins found near the fragmentation site (uL30, eL6, eL28, uL4) and eL29 which bridges ES9L (H30) and ES7L (ES7Lc). Asterisk* corresponds to the fragmentation site (naked mole-rat ribosome) or endpoint of model in ES15L due to flexibility (guinea pig and tuco-tuco ribosomes).

3.4 Methods

Sample preparation

Fibroblast cells from naked mole-rat and tuco-tuco rodents were resuspended and incubated on ice for 20min in buffer containing 50mM Tris (pH7.4), 300mM K-acetate, 7mM Mg^{+2} acetate, 380mM sucrose, 2mM DTT, 0.14% v/v Triton X-100, and Roche protease inhibitor tablet. To complete lysis, cells were freezed and thawed 2x's. The lysate was clarified by centrifugation at 11700 g until no pellet was detected. Ribosomes in the supernatant were pelleted through a 1 M sucrose cushion in 20mM HEPES (pH 7.4), 300mM K-acetate, 7mM Mg^{+2} acetate, 2mM DTT, and 20% glycerol for 16 hours at 40,000 rpm in a 70.1 Ti fixed angle rotor. Sucrose was removed from the ribosome enriched pellet, which was then resuspended in 200 µl of the same buffer. This suspension was clarified by centrifugation at 11700 *g* for 10 min until no pellet was detected. The sample was then loaded onto a two-layer sucrose cushion formed with a bottom layer of 40% w/v sucrose and a top layer of 20% w/v sucrose in buffer containing 20mM HEPES (pH 7.4), 100mM K-Ac, 10mM Mg^{+2} acetate, and 2mM DTT. The resulting translucent ribosome pellet was resuspended in the same buffer and applied onto a grid for imaging. For the guinea pig rodent sample which was purified at a later date, a similar procedure was used, however the two-layer sucrose cushion step was exchanged for a sucrose gradient (10-50%) as this setup was now successfully implemented and resulted in a ribosomal peak (Figure 18). The sucrose gradient improves the quality and/or purification of particles for data collection. The gradient was prepared with a BioComp Gradient Master (BioComp Instruments, Fredericton, NB) and spun down for 15 hours at 19,000 rpm at 4°C. The generated gradient was fractioned and the fractions corresponding to 80S ribosomes were dialysed (Spectra/Por7 Dialysis Membrane pre-treated RC tubing, MWCO 10 kDa) for 16 h to

remove the sucrose. The final sample containing ribosomes was concentrated using Corning Spin-X UF 6 concentrator with a MWCO of 10 kDa. Final concentration of the sample was ~600 nM.

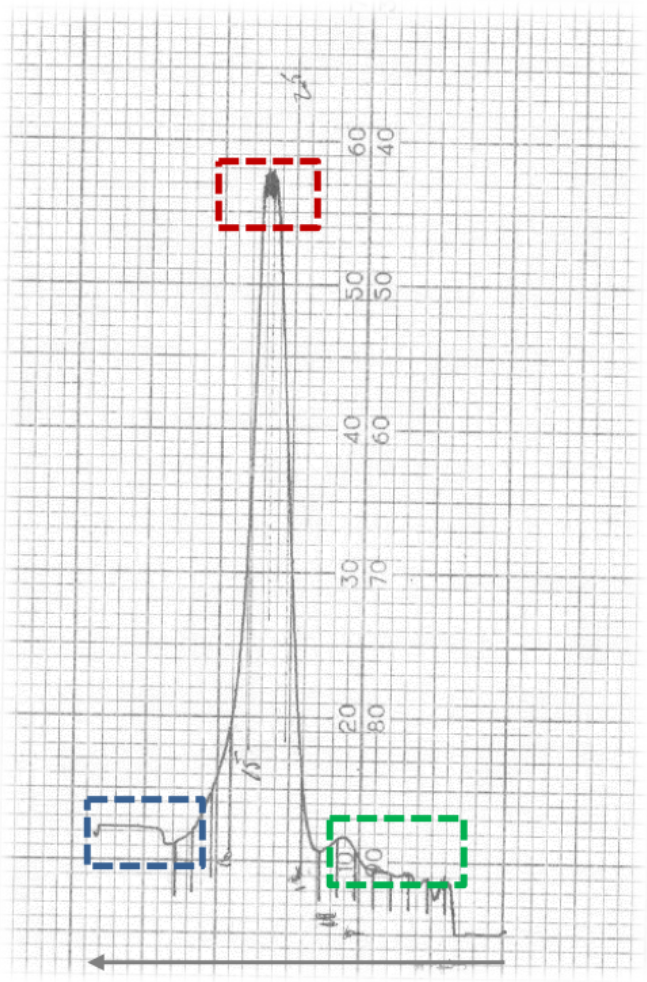

Figure 18. Sucrose gradient and fractionation of 80S Guinea pig ribosomes. Each kind of particle sediments at a zonal level through the gradient at a rate dependent on its size, shape, and density. The red box represents the main peak containing ribosomes, the green box represents heavier species which sediment faster (i.e. polysomes) and the blue box represents species that sediment slower (i.e. subunits, truncated and/or broken particles)

The purified ribosomes (3 microliters) were then applied onto copper/holey carbon grids (carbon-coated Quantifoil R2/2 grids, Quantifoil Micro Tools, GmbH, Großlöbichau, Germany) that were coated with an additional continuous thin layer of carbon applied with an Edwards Auto 306 evaporator, and then glow-discharged using the Gatan Solarus 950. Grids were blotted for 3 seconds at 4°C at 100% humidity and plunge frozen into liquid ethane cooled with liquid nitrogen to -180°C. Gold grids were exchanged for copper/holey carbon grids for the later guinea pig sample (Quantifoil R 1.2/1.3 300 mesh Au holey carbon grids, Quantifoil Micro Tools, GmbH, Großlöbichau, Germany).

The naked mole-rat and tuco-tuco ribosomes were imaged with the FEI Falcon 2 camera installed on the Titan Krios microscope with a Cs corrector and EPU software (HHMI Janelia Farm Research Campus). Aiming to obtain high resolution, a high-magnification setting was used (after calibration 133,970 X on the Falcon 2 camera), yielding a pixel size of 1.045 Å/pixel, which corresponds to a Nyquist limit of 2.09 Å. For the naked mole-rat dataset each image collected was composed of 25 frames with a total dose of 40 e/Å^2 in integrating mode. The guinea pig sample, was imaged on the FEI Tecani F30 Polara microscope equipped with a K2 Summit camera (Gatan) now available in house.

Sample	EM	Voltage	Frames	Dose	Pixel Size
Naked mole-rat	Krios Falcon II	300 kv	25	40 e/Å^2	1.045
Tuco-tuco	Krios Falcon II	300 kv	16	30 e/Å^2	1.045
Guinea pig	Polara K2	300 kv	40	40 e/Å^2	0.98

Figure 19. Data collection of the Naked mole-rat, Tuco-tuco, and Guinea pig ribosomes.

Image processing

Micrographs were examined visually to remove those with contamination by pieces of ice and ethane and those exposed to the shifted beam, before movie processing and particle picking. Next the pre-processing of the images was performed; this includes frame alignment, CTF estimation, and particle picking procedures described in Chapter 1, section 1.3. Frame alignment was originally performed with Motioncorr (whole-frame alignment; [33]) only but with the release of new motion correction software (described in section 1.3), particle polishing was implemented in later steps as well as the new MotionCor2 (patch-level correction, [34]). Micrographs showing thick ice were excluded in a semi-automatic particle selection step using *e2boxer.py*. The power spectra of micrographs generated for CTF determination after movie processing were also examined to exclude those with uncorrected drift and astigmatism.

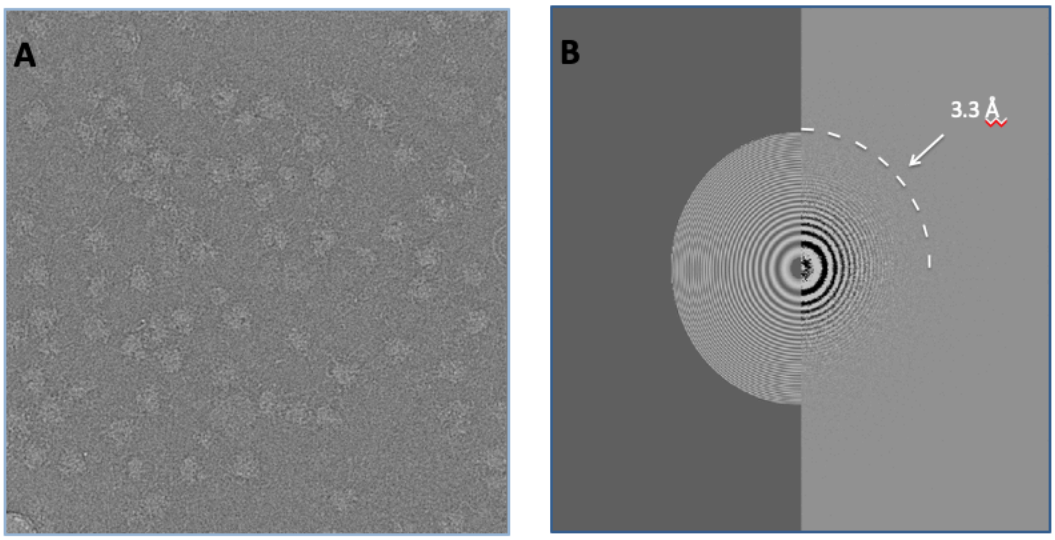

Figure 20. Micrograph and Power spectrum (A) Micrograph collected on FEI Titan Krios at Janelia Farms. (B) Power spectrum shows Thon rings with good fit up to 3.3Å

The naked mole-rat dataset was processed first. Approximately 400k particles were selected for subsequent image processing. The particles were subjected to RELION's (RELION 1.4 or RELION 2.0 [46] versions as made available were used 2D classification, 3D classification, 3D auto-refinement, and post-processing procedures described in Chapter 1, section 5. The first processing scheme adapted from Li W et al (2015) is described below and resulted in a reconstruction at 4.1 Å (Chapter 3, section 2.1). This scheme was optimized and new software releases improving the resolution were incorporated resulting in an improved reconstruction of 3.2 Å for the 80S ribosome (section 2.2 in this chapter) and a 3.0-Å reconstruction for the large ribosomal subunit. Since classification progresses from coarse to fine details, it can be done on binned versions of the data first to save time and computational resources. For this reason, 4x and 2x binning was done before proceeding to the processing of unbinned data.

Image processing scheme 1

On the 4x binned dataset, approximately 400K particles were subjected to an initial round of 3D classification to eliminate impurities and "bad" particles. All ribosome-like particles (~ 200K) were pooled for a second round of classification. In this round, various states of the ribosome could be recognized, including ribosomes with and without bound components (tRNAs, eEF2), in rotated and nonrotated configurations. Homogenous classes (three distinct conformations were identified in Naked mole-rat data) were subjected to independent rounds of 3D classification (now with a 2x binned dataset). The idea for these subsequent rounds of classifications was to sort out new classes or conformations not sorted out in previous rounds to improve homogeneity and thus the quality of the reconstruction (Figure 21). The homogenous class resulting from these further classifications was then subjected to RELION's 3D auto-

refinement procedure to form an improved consensus refinement, which was then subjected to 2D classification and re-refinement again through RELION's 3D auto-refinement procedure.

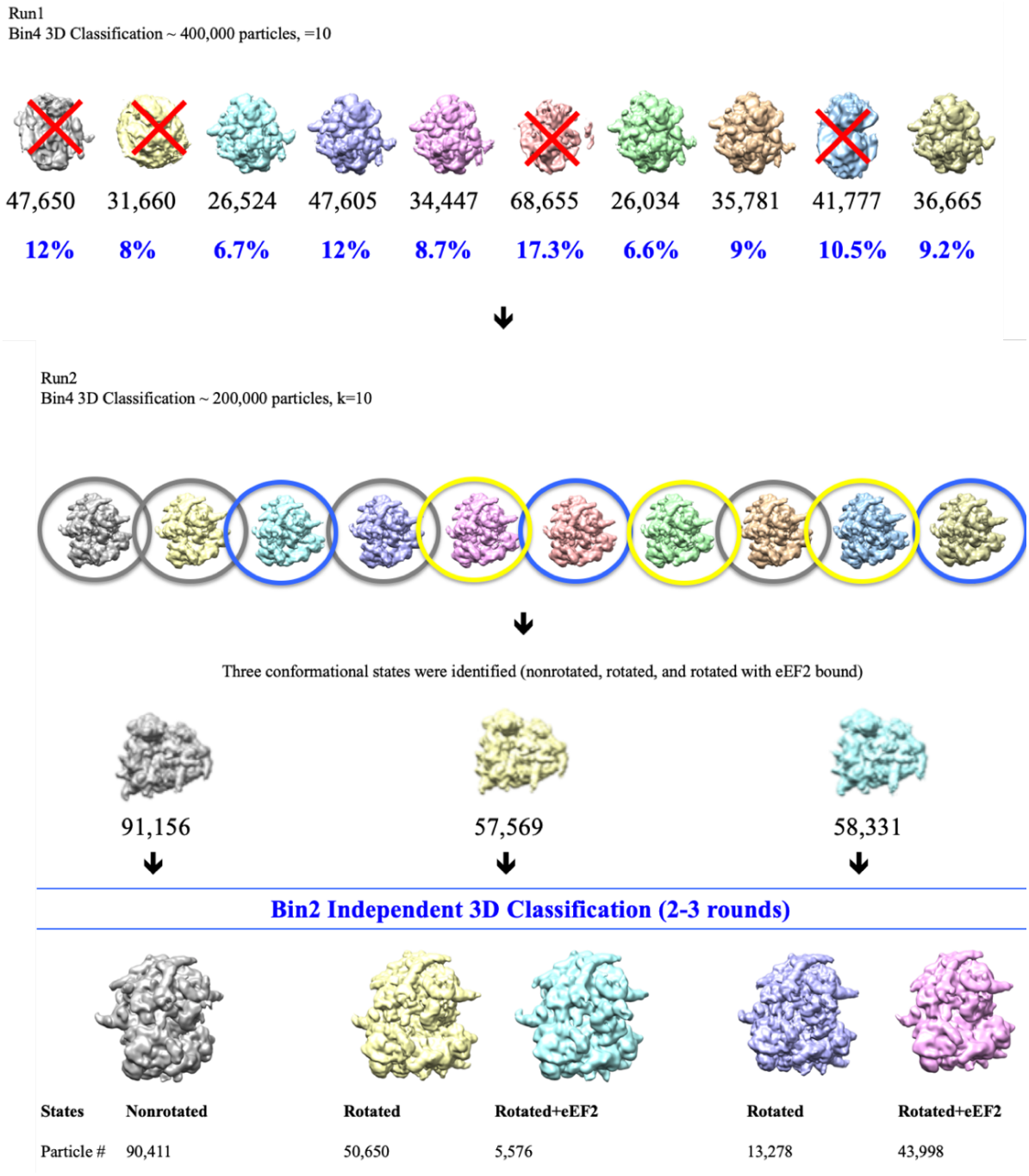

Figure 21. Image processing of the Naked mole-rat ribosome.

Image processing scheme 2

The initial processing described above allowed me to obtain practical experience and build an intuitive understanding of image processing. This then allowed me to optimize the procedures for high-resolution. Further, new software (i.e. particle polishing) was becoming available. For this image-processing pipeline, 2D classification (on 4x binned data) was now performed first on the initially selected 400k particles to remove "bad" particles and gain an overview of the dataset. Next, the remaining particles (285K) were subjected to RELION's 3D auto-refinement procedure prior to subsequent 3D classification to form a consensus reconstruction with orientations close to the correct ones. The first round of 3D classification was used to remove bad particles and a large ribosomal subunit class (~20K particles) from the 80S data. The 80S ribosome-like particles (~215 K) were then subjected to two rounds of 3D classification (on 2x binned data). Three conformations were identified which were then independently subjected to RELION's 3D auto-refinement procedure and further rounds of independent (2D and 3D) classification to improve homogeneity.

3D auto-refinement as well as particle polishing (alignparts_LMBFGS) [37] and then re-refinement (the 3D auto-refinement procedure implemented again but with polished particles) was then performed on the unbinned data to obtain the final reconstructions of the ribosomes in the three conformations found in the naked mole-rat dataset. At this point the newly released Motioncor2 [34] was implemented to compare the performances of different motion correction methods. For my dataset, I found they resulted in similar reconstructions -- the nonrotated class reached 3.2 Å via particle polishing and 3.3 Å with Motioncor2. However, upon subjection to per-particle CTF refinement (RELION) they both reached 3.1 Å.

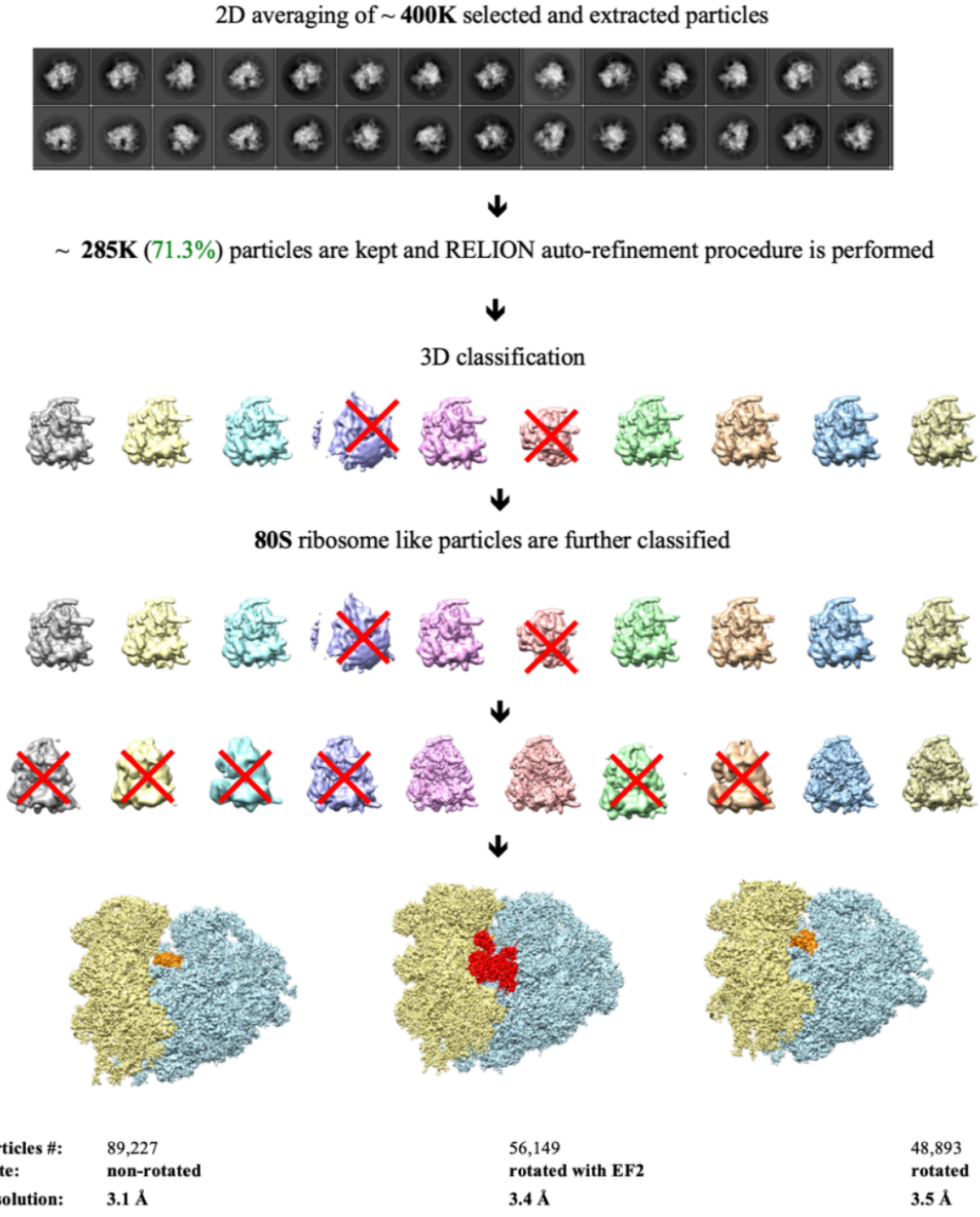

Figure 22. Optimized image processing of the Naked mole-rat ribosome

A focused 3D refinement with partial signal subtraction was performed focusing on the large subunit of an independent class (nonrotated), representing a single conformation. This resulted in a large subunit map with a global resolution of 3.0Å, which was used for model building. Modeling of the naked mole-rat, guinea pig, and tuco-tuco ribosomes was performed in Coot [82], optimizing for fit to the density using rigid body fitting and jiggle fit followed by real-space refinement in Phenix [83]. For the naked mole-rat and guinea pig ribosomes, sequences were obtained from the National Center for Biotechnology Information protein databases (www.ncbi.nlm.nih.gov) and RNAcentral (https://rnacentral.org/) for the ribosomal proteins and the rRNA, respectively. For the tuco-tuco ribosome, which has no available genome sequence, model building was performed using the naked mole-rat structure as a starting point. Protein side chains could then be adjusted from the high-resolution map directly. And an available sequence for a portion of rRNA near the fragmentation site was used [14]. The density of the small subunit was still worse even with partial signal subtraction; particularly the head region. The guinea pig and tuco-tuco ribosomes were processed similarly to the naked mole-rat; this resulted in all high-resolution reconstructions of their large ribosomal subunits (<3.5Å).

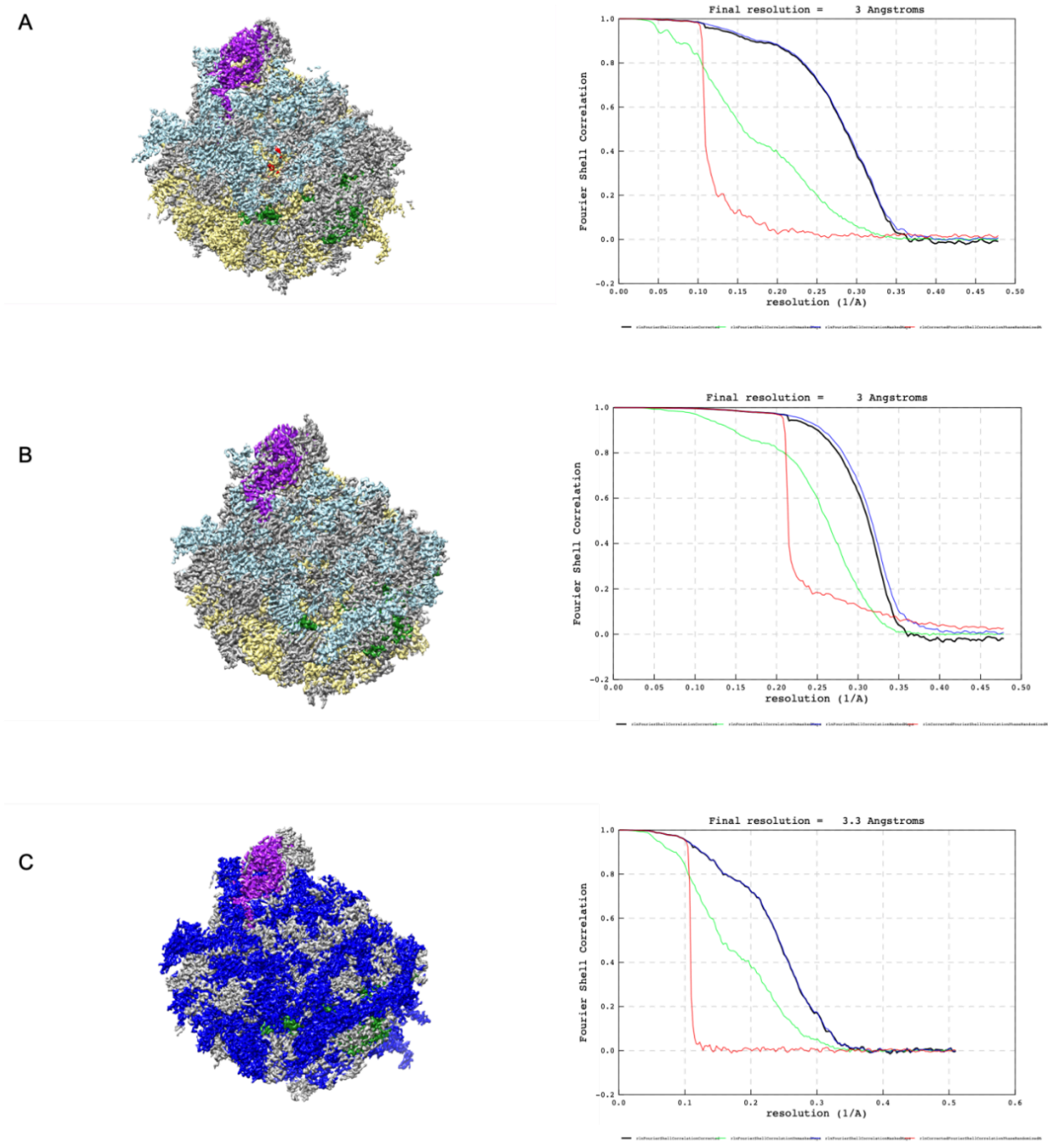

Figure 23. Refinement and global resolution of the Naked mole-rat, Tuco-tuco, and Guinea pig large subunit (60S). (A) Naked mole-rat (B) Tuco-tuco (C) Guinea pig

Protein	Chain ID	NMR #AA	Sequence ID	Range modeled	GP %ID
uL2	A	257	**XP_004837716.1**	2-245	100
uL3	B	403	**XP_004845639.1**	2-395	99
uL4	C	420	**XP_004855657.1**	2-363	95
uL18	D	297	**XP_004841318.1**	3-295	99
eL6	E	287	**XP_004843823.1**	38-287	93
uL30	F	248	**XP_004842110.1**	24-248	98
eL8	G	266	**XP_004849005.1**	27-266	99
uL6	H	192	**XP_004870352.1**	1-90	99
uL16	I	253	**XP_004875050.3**	41-253	100
uL5	J	178	**XP_004850617.1**	8-177	100
eL13	K	211	**XP_004842693.1**	2-211	99
eL14	L	216	**XP_021113684.1**	2-139	81
eL15	M	204	**XP_004857820.1**	2-204	100
uL13	N	203	**XP_004866976.1**	5-203	99
uL22	O	184	**XP_012931439.1**	2-154	100
eL18	P	188	**XP_004867046.1**	2-188	98
eL19	Q	196	**XP_004859572.1**	2-181	100
eL20	R	212	**XP_004873387.2**	37-212	99
eL21	S	160	**XP_004854871.1**	2-160	100
eL22	T	128	**XP_004863788.1**	17-115	99
uL14	U	140	**XP_004859587.1**	2-140	100
eL24	V	157	**XP_004858048.1**	1-63	100
uL23	W	156	**P_004857066.1**	39-156	100
uL24	X	145	**XP_004857553.1**	1-134	100
eL27	Y	136	**XP_004859346.1**	2-136	100
uL15	Z	148	**XP_004851440.1**	2-148	100
eL29	a	227	**XP_004834249.1**	2-74	74
eL30	b	115	**XP_004861199.1**	10-107	97
eL31	c	125	**XP_004844503.1**	18-124	100
eL32	d	135	**XP_004870505.1**	2-129	100
eL33	e	110	**XP_004834722.1**	2-110	100
eL34	f	117	**XP_004866273.1**	2-115	100
uL29	g	123	**XP_004849447.1**	2-123	98
eL36	h	105	**XP_004865734.1**	2-103	100
eL37	i	97	**XP_004921160.2**	2-87	100
eL38	j	99	**XP_021096759.1**	31-99	100
eL39	k	88	**XP_004839730.1**	2-51	98
eL40	l	128	**XP_004873436.1**	77-128	100
eL42	n	237	**XP_004835652.1**	2-105	100
eL43	o	92	**NP_001254782.1**	2-92	100

Figure 24. Modelled naked mole-rat (NMR) ribosomal proteins (r-proteins) and their sequence identity (ID%) to guinea pig (GP) r-proteins. Sequences are from the NCBI database and are corroborated with the cryo-EM density maps when possible (N- or C-terminal extensions found in solvent surface are unable to be modelled due to flexibility) Most r-proteins are highly conserved across mammals (≥ 98% sequence identity).

Chapter 4: Dynamical features of the naked mole-rat ribosome

Translation requires the ribosome to undergo both large- and small-scale conformational rearrangements as it moves along the mRNA and interacts with translational factors and tRNA substrates. Biochemical, structural and biophysical studies over the last 30 years have revealed the extent of the ribosome's dynamics and flexibility during translation. Large-scale molecular movements in the ribosome were long anticipated and inferred from the nature of the translation elongation cycle [84, 85]. And in the late 1980s, biochemical and foot printing experiments first suggested intersubunit movements allow translocation to occur in two steps [86]. In the first step, the acceptor end of the tRNA moves relative to the large subunit resulting in 'hybrid states' of binding. During the second step, which is promoted by elongation factor, EF-G, the anticodon of tRNA, along with the messenger RNA, moves relative to the small subunit. Subsequent ensemble FRET experiments also demonstrated that peptide bond formation results in the spontaneous rearrangement of the ribosome bound tRNAs from their "classical" P/P and A/A configurations, into intermediate "hybrid" P/E and A/P configurations (denoting to the small subunit/ large subunit tRNA sites) [87].

A decade later, cryo-EM studies with reconstructions of pretranslocation (PRE) complex analogs containing vacant A sites and stabilized through the binding of EF-G in the presence of GDPNP, a nonhydrolyzable GTP analog, allowed visualization of the P/E tRNA configuration and observation of large-scale conformational rearrangements during translocation [88, 89]. These studies revealed three major conformational changes (1) the aforementioned movement of the deacylated P-site tRNA from the P/P to the P/E configuration [89] (2) the ~20Å movement of the L1 stalk, from an open to a closed conformation such that it establishes a direct interaction

with the elbow of the P/E-configured tRNA [89] and (3) the counter-clockwise, ratchet-like rotation of the 30S (small subunit) with respect to the 50S subunit (large subunit) from a nonrotated to a rotated subunit orientation [88, 89]. This gave rise to the proposal that translocation is driven by a ratchet-like mechanism that is coupled to intersubunit rotation. Studies restricting potential intersubunit movement via a disulfide cross-link resulted in ribosomes that were unable to carry out translation [90]. Specifically, cross-linked ribosomes were blocked in EF-G-dependent translocation, supporting the requirement of intersubunit movement for ribosomal translocation.

This intersubunit rotation is not a simple rigid-body movement, but one that includes both large- and small-scale structural rearrangements within both subunits. One of these is a movement of the head of the small subunit, an autonomous structural domain that is connected to the rest of the subunit by a single helix (h28) of the rRNA (16S in prokaryotes, 18S in eukaryotes). Cryo-EM reconstructions of EF-G-containing ribosomal complexes have revealed that the aforementioned intersubunit rotation is accompanied by a rotation of the head domain of the small subunit by up to 21° relative to the rest of the small subunit [89, 91]. Eukaryotic ribosomes bound with eEF2 also show a comparable rotational movement of the small subunit head [57]. This rotation of the head, also termed head swivel, is required to maintain proper contact between the tRNAs and the two ribosomal subunits upon transition to the hybrid state [57]. In particular, head swivel allows movement of the anticodon stem-loops (ASLs) of tRNAs and their associated mRNA codons along the direction of translocation through the ribosome [86]. Furthermore, head swivel is thought to be essential for ribosomal translocation because it opens a wide (over 20 Å) path for tRNA translocation between the P and E sites on the small subunit that is otherwise constricted by the rRNA residues of the head and platform of the small

subunit [92]. In eukaryotes and SecM-stalled postranslocational (POST) state complexes from *Escherichia coli*, a "rolling" of the small subunit has also been observed [93, 94]. The rolling motion in eukaryotes is described as a ~6° rotation of the small subunit (40S) subunit toward the L1 stalk around the long axis of the small subunit and demonstrated to occur during tRNA selection [93]. This small subunit (40S) body/platform rotation (rolling) is measured relative to the large subunit, while head swivel motion is measured relative to the small subunit body/platform [93].

Large-scale conformational rearrangements also occur in the large subunit. The largest movement is found in the L1 stalk, which comprises helices 76, 77, and 78 (H78 is absent in most eukaryotes) and protein uL1 [95]. The L1 stalk interacts with deacylated tRNA in the E site during translocation and regulates its release from the ribosome [96]. Cryo-EM and X-ray studies have shown the L1 stalk in at least three different conformations. In X-ray structures of ribosomes with a vacant E site, the L1 stalk is observed in an "open" conformation leaning away from the body of the subunit [92, 96]. When deacylated tRNA is bound in the classical E/E state (nonrotated), formation of the contact between the L1 stalk and the elbow of the tRNA requires the stalk to move inward by 30-40 Å relative to its open conformation (i.e. half-closed conformation) [97-99]. In rotated, hybrid-state complexes having EF-G, the L1 stalk moves by an additional 15-20 Å, relative to its position in the E/E state complex (i.e. closed conformation) [91]. This enables the L1 stalk to contact the elbow of hybrid-state P/E tRNA. In the absence of EF-G, the L1 stalk spontaneously fluctuates between the open and closed positions with movements coupled to the fluctuations of tRNAs between classical (nonrotated) and hybrid (rotated) states, respectively [100]. Thus movements of the L1 stalk allow remodeling of the ribosomal E site for different states of the elongation cycle [101].

Another highly dynamic domain of the large subunit is the L11 stalk (p stalk in eukaryotes), which contains the so-called "GTPase-associated center" (GAC) and is located opposite of the L1 stalk on the large subunit. The L11 stalk is formed from helices 43 and 44 as well as protein uL11. At the base of the stalk is the binding site for protein uL10, to which are bound two dimers of L7/L12 (P1/P2 in eukaryotes). The L11 or P-stalk interacts with translational GTPases. GTPases have several roles in translation, including delivering tRNA to the P site of the ribosome during initiation (e.g. eIF2) or the A site during elongation (e.g. eEF1A), and catalyzing translocation during elongation (e.g. eEF2). For all these GTPases, the GAC acts in a manner similar to GTPase-activating proteins to trigger GTP hydrolysis. X-ray structures of vacant and t-RNA-containing ribosome complexes from *Escherichia coli* and *Thermus thermophilus* showed that the L11 moves by more than 15 Å toward the A site [92, 96, 97]. Various degrees of inward or outward movement of the L11 stalk have been seen in cryo-EM reconstructions of ribosomal complexes containing EF-G [102], EF-TU [103], and RF3 [104] and eukaryotic translational factors (eEF2 and eEF1A) [58, 93].

Here we have solved structures for two conformational states of the 80S naked mole-rat ribosome, a nonrotated state harboring a single tRNA (E/E) and a rotated state harboring two hybrid-state tRNAs (P/E and A/P). By comparing these structures we identified dynamic components of the naked mole rat's large subunit (60S) whose motions may be correlated to the functional dynamics of the small subunit (40S) during translation (intersubunit rotation, head swiveling, etc) in an attempt to reveal a role for the 28S rRNA fragmentation and its possible link to the high translational fidelity reported for the naked mole rat.

4.1 Predominant motions of the naked mole-rat 80S ribosome

Predominant motions for the naked mole-rat ribosome were characterized by performing a consensus refinement of two conformational states in the naked mole-rat dataset followed by a multibody refinement. The consensus refinement of the nonrotated and rotated ribosomes yielded a global resolution estimate of 4.1 Å after standard RELION post-processing. The consensus map showed fuzzy density for the small subunit compared to the rest of the ribosome. In particular the head region of the small subunit exhibited poor density. For multibody refinement in RELION [105], the ribosome was split into three bodies, (1) the large subunit (2) the small subunit without the head, and (3) the head of the small subunit.

Application of the principal component analysis in the relion_flex_analyse program revealed that approximately 50% of the variance in the rotations and translations of the three bodies is explained by the first two eigenvectors (Figure 25). Motion along the first eigenvector alone accounts for almost 40% of the variance and is reminiscent of a ratchet-like motion of the small subunit with respect to the large subunit. The motion along the second eigenvector accounts for approximately 10% of the variance and seemingly corresponds to the rotation of the head or head swiveling relative to the small subunit.

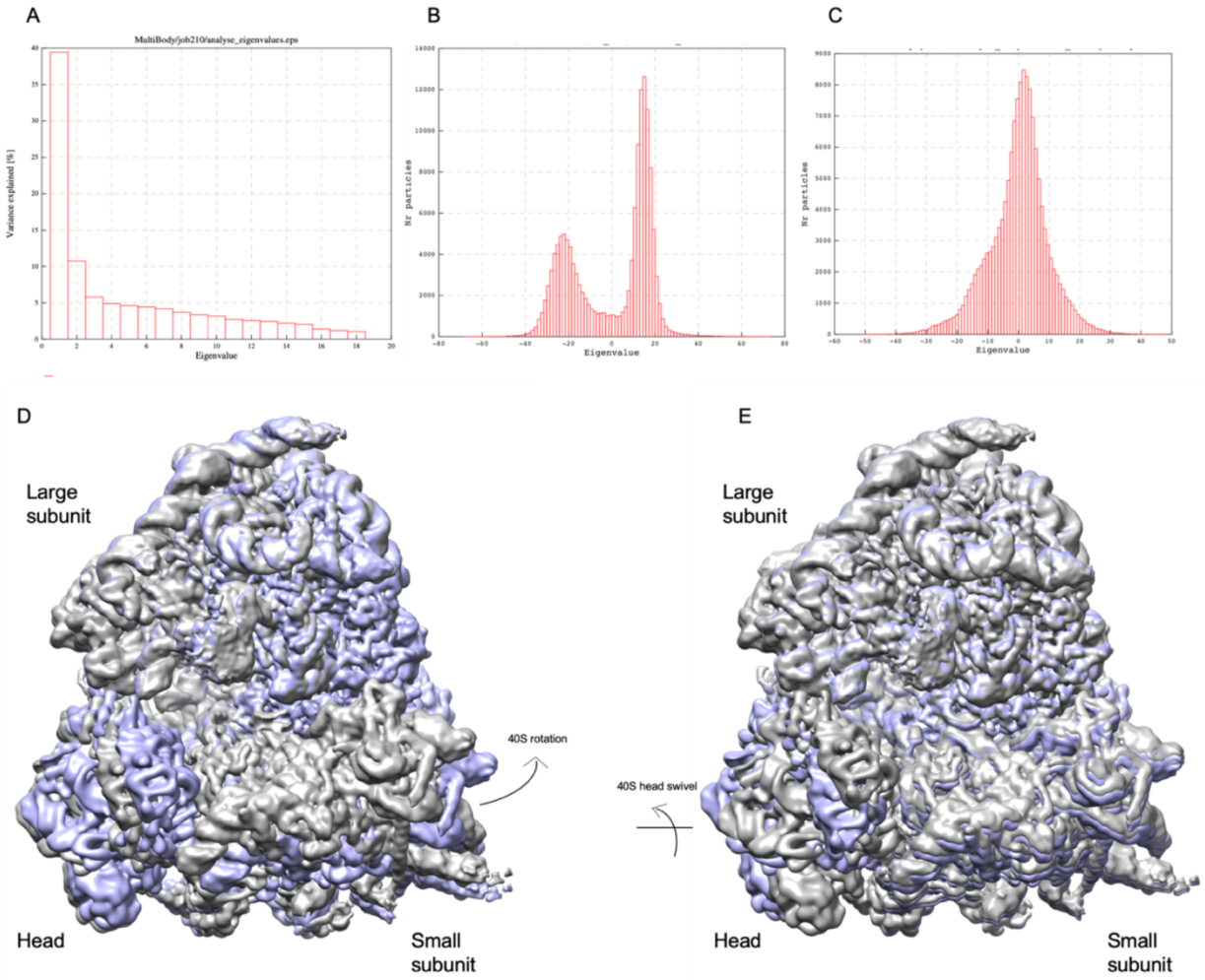

Figure 25. Multibody refinement of the Naked mole-rat ribosome. (A) The contributions of all eigenvectors to the variance. Approximately 50% of the variance in the rotations and translations of the three bodies is explained by the first two eigenvectors. (B) The histogram of the amplitudes along the first eigenvector shows a bimodal distribution. (C) The histogram of the amplitudes along the second eigenvector. (D) Maps at the extremes shown for the first eigen vector (E) Maps at the extremes shown for the second eigen vector.

4.2 Overview of the global conformational changes in the naked mole-rat 80S structures

Models for the two conformational states (nonrotated and rotated), primarily related by a ratchet-like motion, were built into their respective cryo-EM density maps. To identify dynamic elements we then calculated and visualized a pairwise Root-mean-square deviation (RMSD, in Å) in PyMOL [106] and Chimera [107], respectively.

The small subunit is highly dynamic having the largest displacements of the two ribosomal subunits (Figure 26). This includes, the interface region of the small subunit head. In contrast the large subunit has displacements restricted to only a few Ångstroms (Figure 27). The largest displacements (>4Å) in the large subunit are elements of the central protuberance, Helix 38 (the A-site Finger), Helix 69, and the GTPase-associated center.

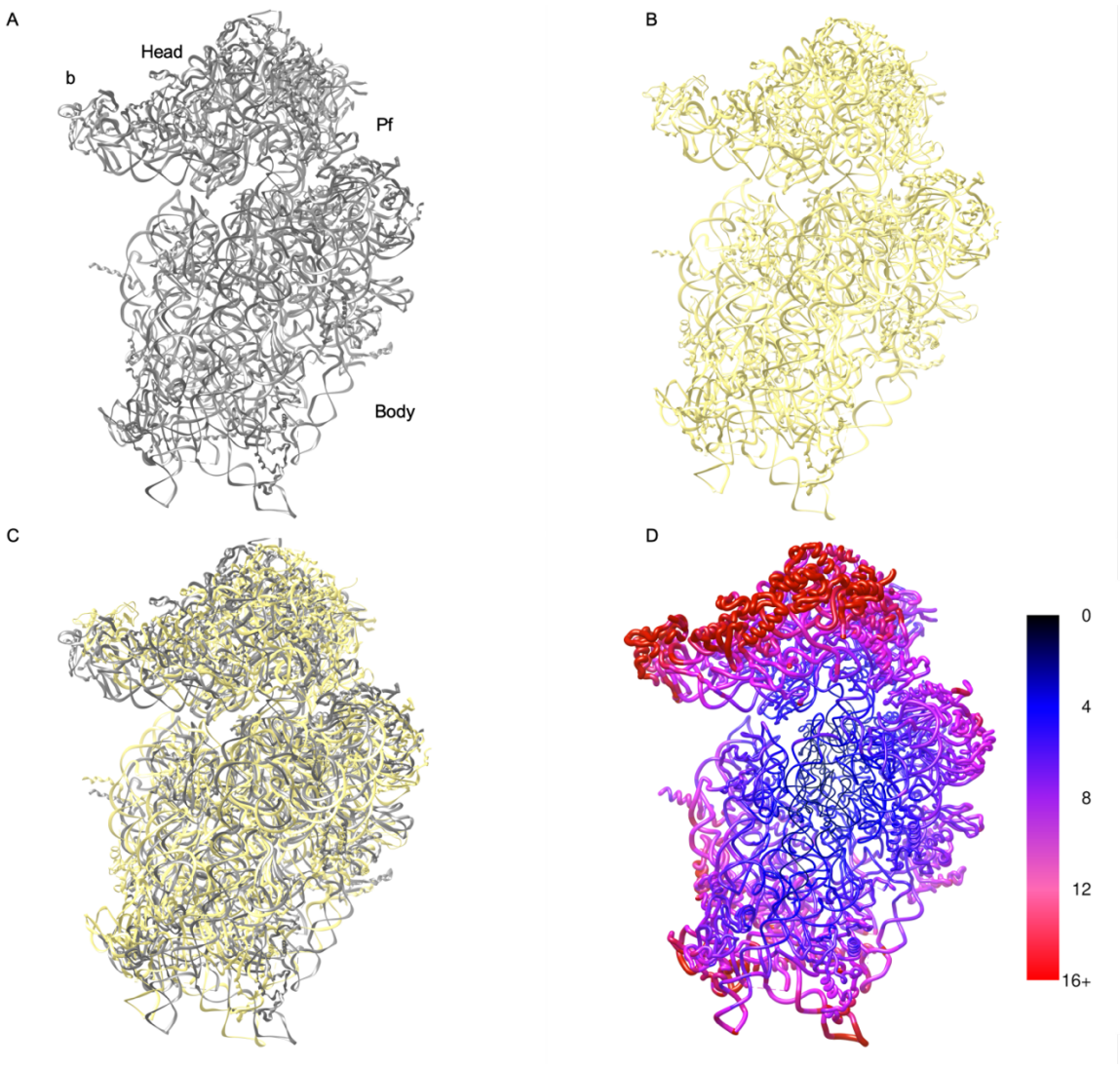

Figure 26. Root-mean-square deviation (RMSD) of the small subunit (SSU). *Interface View* (A) Ribbon diagram of the nonrotated small subunit (b, beak; Pf, platform) (B) Ribbon diagram of the rotated small subunit (C) Overlay of the two conformational states (D) Worm ribbon diagram of the small subunit colored by pairwise root-mean-square deviation displacement of the two conformational states.

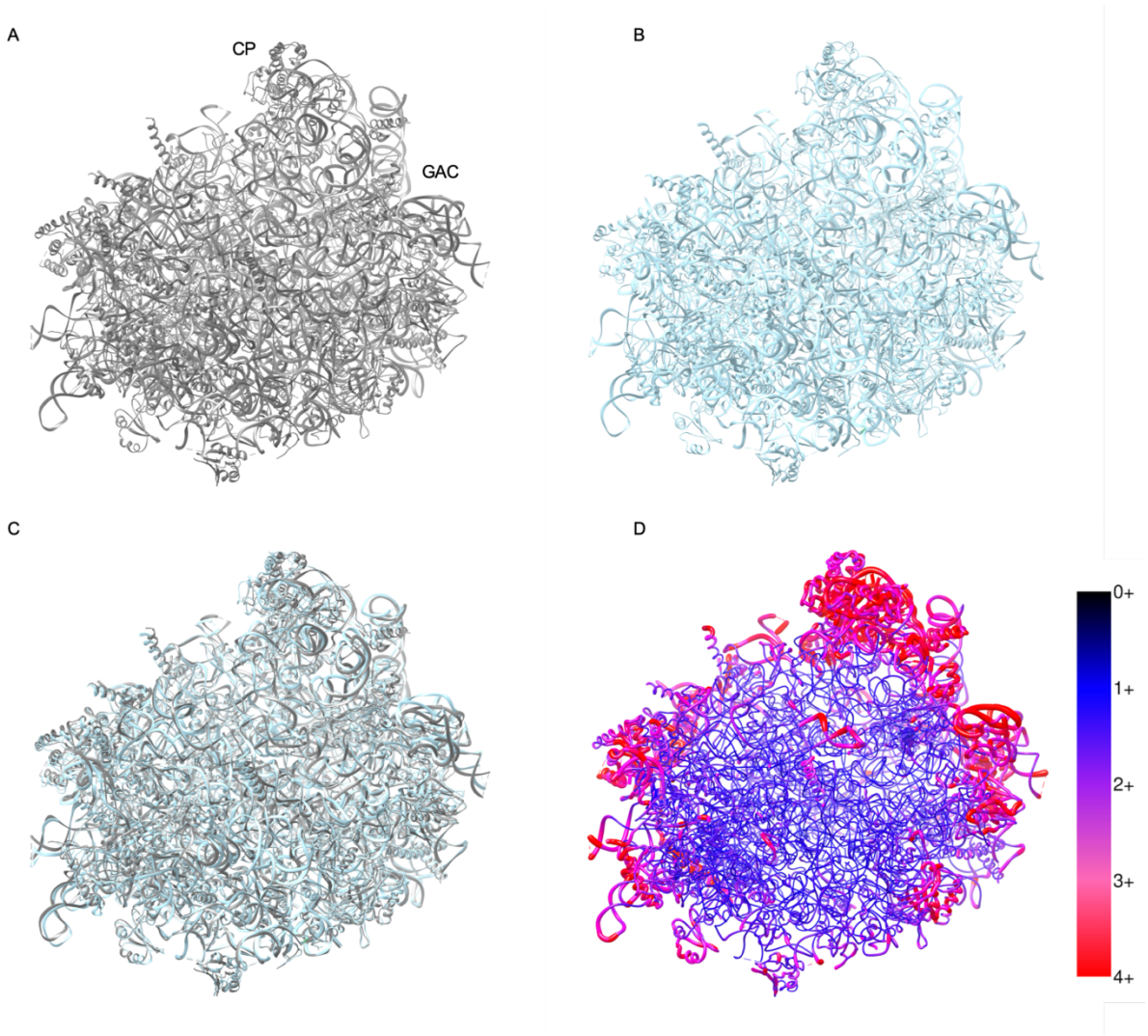

Figure 27. Root-mean-square deviation (RMSD) of the large subunit (LSU). *Interface View* (A) Ribbon diagram of the nonrotated large subunit (B) Ribbon diagram of the rotated large subunit (C) Overlay of the two conformational states (D) Worm ribbon diagram of the large subunit colored by pairwise root-mean-square deviation displacement of the two conformational states.

4.3 Dynamic elements of the fragmented naked mole-rat large ribosomal subunit

The central protuberance, the A-site finger, and H69 all form contacts termed intersubunit bridges with the small subunit that are rearranged and/or broken/reformed during translation. These intersubunit bridges keep the ribosomal subunits in proper register and facilitate precise global conformational rearrangements of the ribosome which are required for tRNA-mRNA movement (translocation).

Components of the central protuberance and the A-site finger form contacts with the highly dynamic head of the small subunit. The central protuberance is composed of the 5S rRNA sandwiched between r-protein uL18 on the solvent side of the large subunit and uL5 together with H84 on the interior region (interface). R-protein uL5 forms the B1b/c bridge with the small subunit r-protein, uS13. R-protein uL5 contacts the elbow of hybrid (A/P) tRNA in the rotated state of the ribosome.

The A-site finger is located adjacent to uL5 and is a long rRNA helix (H38) reaching from the base of the CP into the intersubunit space. The A-site finger forms the B1a intersubunit bridge with the small subunit r-protein, uS19. Interestingly, the C-terminal end of uS19 extends towards the decoding center of the small subunit, thus linking the A-site finger to the decoding center. During translation elongation, bridge B1a remains intact in the nonrotated state. However, when ribosome becomes fully rotated, the A-site finger and uS19 separate, breaking this bridge. This facilitates translocation, allowing the elbow of an A-site tRNA to reach its P-site position on the large subunit. Our structures show the naked mole-rat similarly has an intact and a broken bridge in the nonrotated state and rotated states, respectively.

Helix 69 (H69) contacts helix 44 (h44) of the small subunit, forming the B2a intersubunit bridge next to the decoding center of the small subunit. H69 also interacts with both A and P site

tRNAs during translation, suggesting a role in tRNA translocation [108]. Another dynamic region in the large subunit is the GTPase associated center (GAC) formed by helices 43-44 and linked to the main body of the large subunit by H42. However, the GAC is highly dynamic and only weakly resolved in our cryo-EM densities.

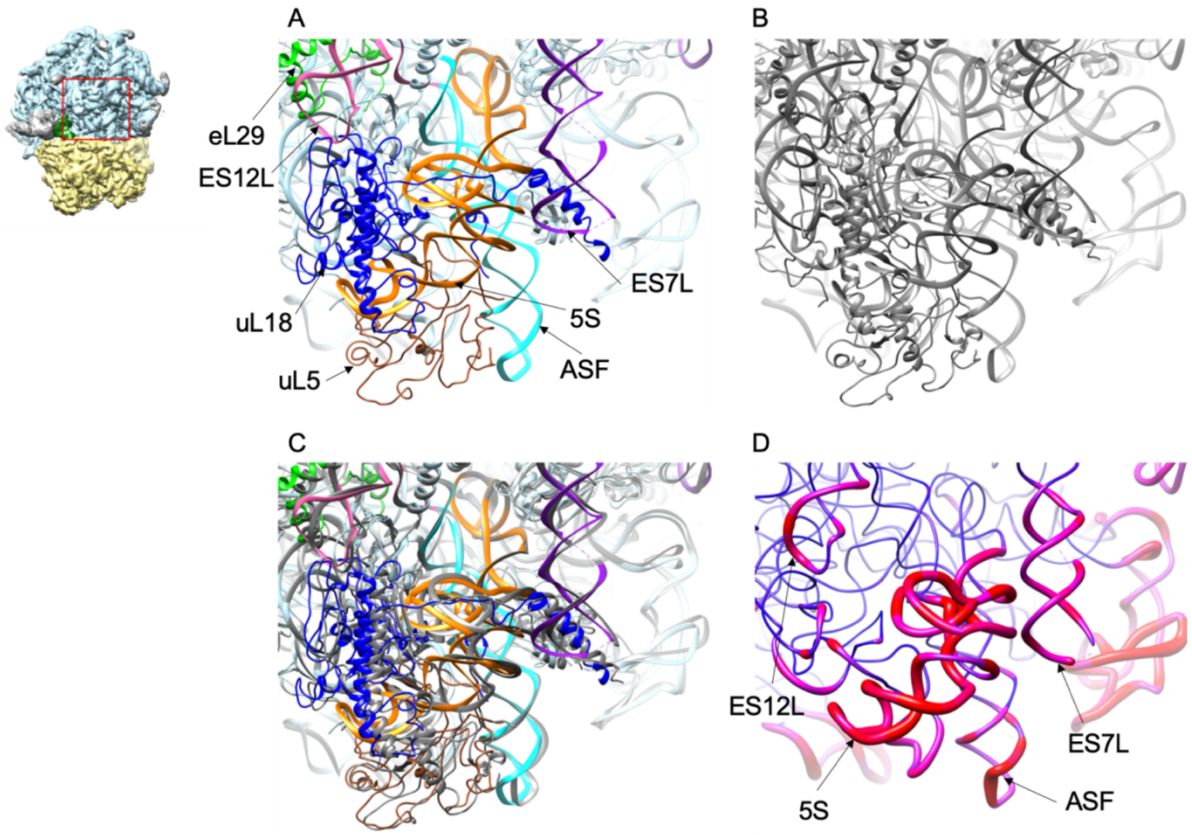

Figure 28. Root-mean-square deviation (RMSD) of Central Protuberance and A-site finger. (A) Ribbon diagram of the rotated state (B) Ribbon diagram of the nonrotated state (C) Overlay of the two conformational states (D) Worm ribbon diagram (only rRNA is shown for clarity) colored by pairwise root-mean-square deviation displacement of the two conformational states.

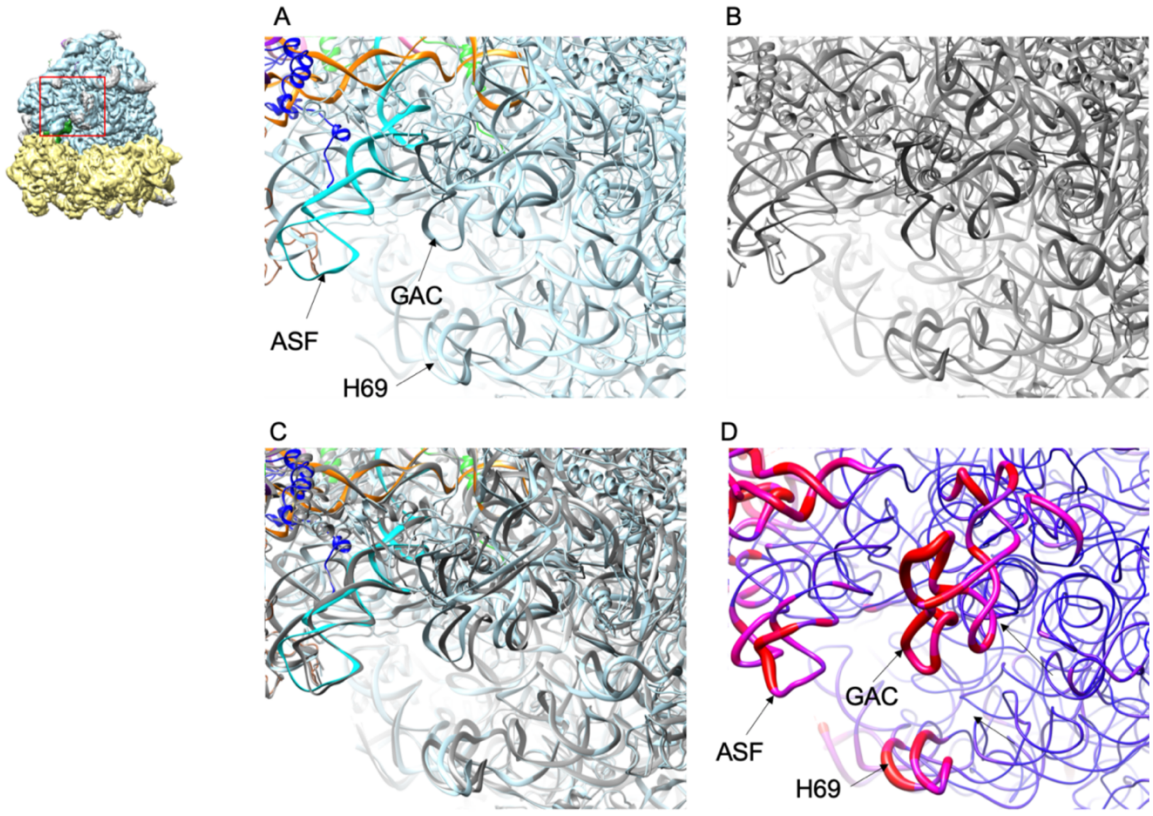

Figure 29. Root-mean-square deviation (RMSD) of the A-site finger (ASF), GTPase Associated Center (GAC), and H69. (A) Ribbon diagram of the rotated state (B) Ribbon diagram of the nonrotated state (C) Overlay of the two conformational states (D) Worm ribbon diagram (only rRNA is shown for clarity) colored by pairwise root-mean-square deviation displacement of the two conformational states.

4.4 Discussion

By using cryo-EM and the tools of 3D classification we identified dynamic elements of the fragmented naked mole-rat large subunit (60S) in our sample, primarily related to the functional dynamics of intersubunit rotation and small subunit (40S) head swiveling. Regions having high flexibility become blurred or even invisible in the reconstruction. Dynamic elements were spotted by identifying faded out or missing density in part of the naked mole-rat structure compared to the control, the guinea pig structure. The extended internal loop of ES10L

and a C-terminal portion of r-protein, eL29 is missing in the density map which may correspond to highly flexible parts of the naked mole-rat ribosome (Chapter 3). These elements are found among a cluster of ESs located in the top of the particle that extend towards the interface region of the two ribosomal subunits (Figure 30). Further ES15L in the naked mole-rat is fragmented, resulting in a shortened ES15L which no longer interacts with the terminal loop of H30 from ES9L. In an attempt to reveal a possible role for the 28S rRNA fragmentation and these observed differences in the naked mole-rat ribosome density we aligned non-rotated and rotated models of the naked mole-rat ribosome to the large subunit and proceeded to calculate the root-mean-squared deviation (RMSD) between the two structures. Dynamic elements in the naked mole-rat large subunit include the central protuberance, the A-site finger, H69, and the GTPase associated center. These elements have been previously described in the literature to be dynamic throughout translational processes such as translocation [90-105]. Of these the central protuberance and A-site finger are adjacent to structural features that are different among the three rodents compared in these studies – the naked mole rat, tuco-tuco, and guinea pig (see Chapter 3).

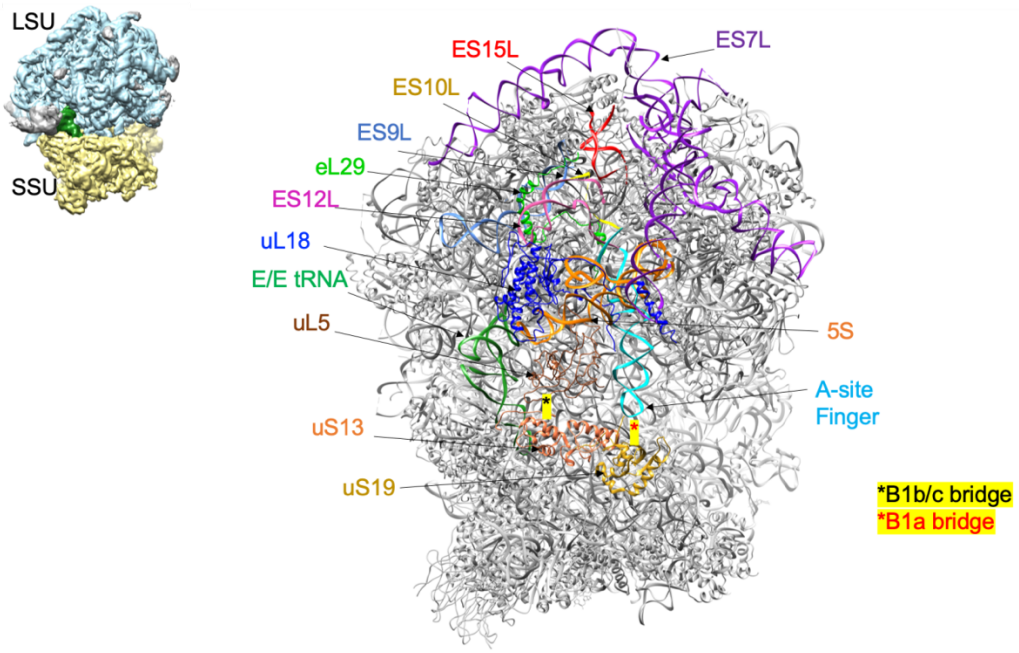

Figure 30. The naked mole-rat nonrotated 80S ribosome. Network of ESs and r-protein eL29 which starts from the top of the particle, near the fragmentation site, and extends to the A-site finger and central protuberance. Small subunit r-proteins uS13 and uS19 form intersubunit bridges B1a and B1b/c with the A-site finger and uL5, respectively.

As previously mentioned, the naked mole-rat has missing density corresponding to an extended internal loop of ES10L. This density of ES10L extends towards the surface and contacts ES9L and ES12L in the guinea pig ribosome. ES12L which emerges from the base of H38 or the A-site finger also contacts the 5S rRNA (a major component of the central protuberance). Conceivably dynamic differences in ES10L observed in the density may affect the dynamics of nearby components of the central protuberance and the A-site finger.

Another missing density in the naked mole-rat ribosome corresponds to a C-terminal portion of eL29. In all three rodents compared, the N-terminal of eL29 is found in the rRNA core, in the vicinity of the peptidyl transferase center; it then tethers to the solvent side and is found sandwiched between ES12L and ES9L. However in the guinea pig and tuco-tuco

ribosomes, a density for the C-terminal portion of eL29 extends out towards the distal r-protein, uL30 and ES7L. This density of eL29 stacks against the N-terminal end of uL30 and seemingly stabilizes the peripheral rRNA segment, ES7L (helix c). ES7L (helix b) is also found above the GTPase associated center, which was noted as another mobile element during translation.

Altogether, these rRNA expansion segments seem to form a network starting from the top of the particle and extend into the vicinity of mobile elements (the A-site finger, central protuberance, and the GTPase associated center) which interact with tRNAs and translational factors. Further, ES10L and ES12L are insertions to the base of the A-site finger and thus are closely linked to the dynamic behavior of the A-site finger. Conceivably the fragmentation and differences found among ES10L, ES12L, ES7L may alter the dynamics and/or function of the A-site finger.

The role of the A-site finger has been investigated in several biochemical studies [109-114]. Mutations in *Escherichia coli* resulting in ribosomes with truncated A-site finger, thus a severed B1a bridge, resulted in acceleration of mRNA-tRNA translocation but also enhanced frameshift activity [112]. Thus, it appears the A-site finger performs complex fine-tuning of the ribosomal activity [112]. Another study performed mutations on the A-site finger which affected formation of the B1a bridge or A-site tRNA interactions [113]; these mutations caused increased ribosomal affinity for the A-site tRNA and had variable effects on A-site-specific drugs and on the suppression of nonsense codons. The authors also observed that these mutations induced structural changes that propagate into distant rRNA regions. Thus they proposed the B1a bridge may be involved in fine-tuning of ribosomal activity and also may function as part of an allosteric communication pathway between the large and small subunits. Mutational studies deleting individual intersubunit bridges (B1a, B4, B7a, and B8) involved in constraining head

movement (B1a) and intersubunit rotation (B4, B7a, B8) also resulted in the acceleration of the maximal rate of translocation [115] . Altogether this reinforces the view of the B1a bridge as a fine-tuning translational control element. The structural differences to insertions of the A-site finger and adjacent regions among the three rodent structures compared in our study may result in an altered (species-specific) fine-tuning mechanism. This could explain the higher translational fidelity observed in the naked mole-rat and tuco-tuco compared to other rodents such as guinea pig and mouse, although the detailed mechanism by which this might be accomplished cannot be ascertained from our study.

RMSD measurements of the nonrotated and rotated states of the naked mole-rat ribosome revealed dynamic elements, related by intersubunit rotation and/or head swiveling seen in our dataset. Accordingly, the tip of the A-site finger (forming one side of the B1a bridge) is dynamic and so are components of the central protuberance including the 5S rRNA, uL18 and uL5. Further ES10L, ES12L, ES7L, and r-protein eL29, which have altered structures and/or interactions among the three rodent ribosomes are also similarly mobile. This suggests communication may occur among these elements of the large subunit. This also supports the notion that these ESs may regulate functions of the A-site finger and central protuberance, such as rearrangement of intersubunit bridges, hybrid-state formation, and fine-tuning of translation.

4.5 Methods

Image Processing

The naked mole-rat dataset was further processed using RELION 3.0 [116]. A map of the intersubunit space, including ligands (tRNAs and EF-G) was made from PDB 4WPO [117] using the molmap feature in chimera. A mask of the intersubunit space was then prepared in RELION

and used for further 3D classification on the previously identified (Chapter 3) nonrotated and rotated ribosomes harboring E/E tRNA and P/E tRNA, respectively. Typically, in 3D classifications a regularization parameter of T=4 is used; however, higher values (10-40) are required in focused classification of a smaller subregion. A higher T value will more heavily weigh the data during assignment of orientations and classification. Classification of our dataset was optimized with a T value of 8. Following focused classification, a subclass of nonrotated ribosomes with E/E tRNA was re-refined and re-classified to establish a homogenous class. The rotated class was similarly further processed. Following, focused classification of the intersubunit space, a subclass having two tRNAs (P/E and A/P) was identified (Figure 31). Standard post-processing in RELION yielded a global resolution estimate of 3.3 Å and 3.5 Å for nonrotated (E/E tRNA) and rotated (P/E and A/P) conformational states of the naked mole-rat ribosome, respectively (Figure 32).

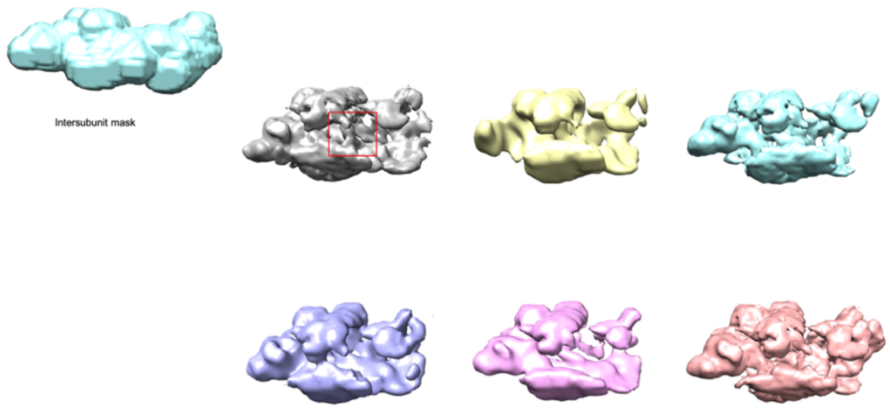

Figure 31. Focused 3D Classification on rotated 80S ribosome. A subclass harboring two tRNAs (P/E and A/P) was identified using regularization parameters of T=4 and K=6 (class number).

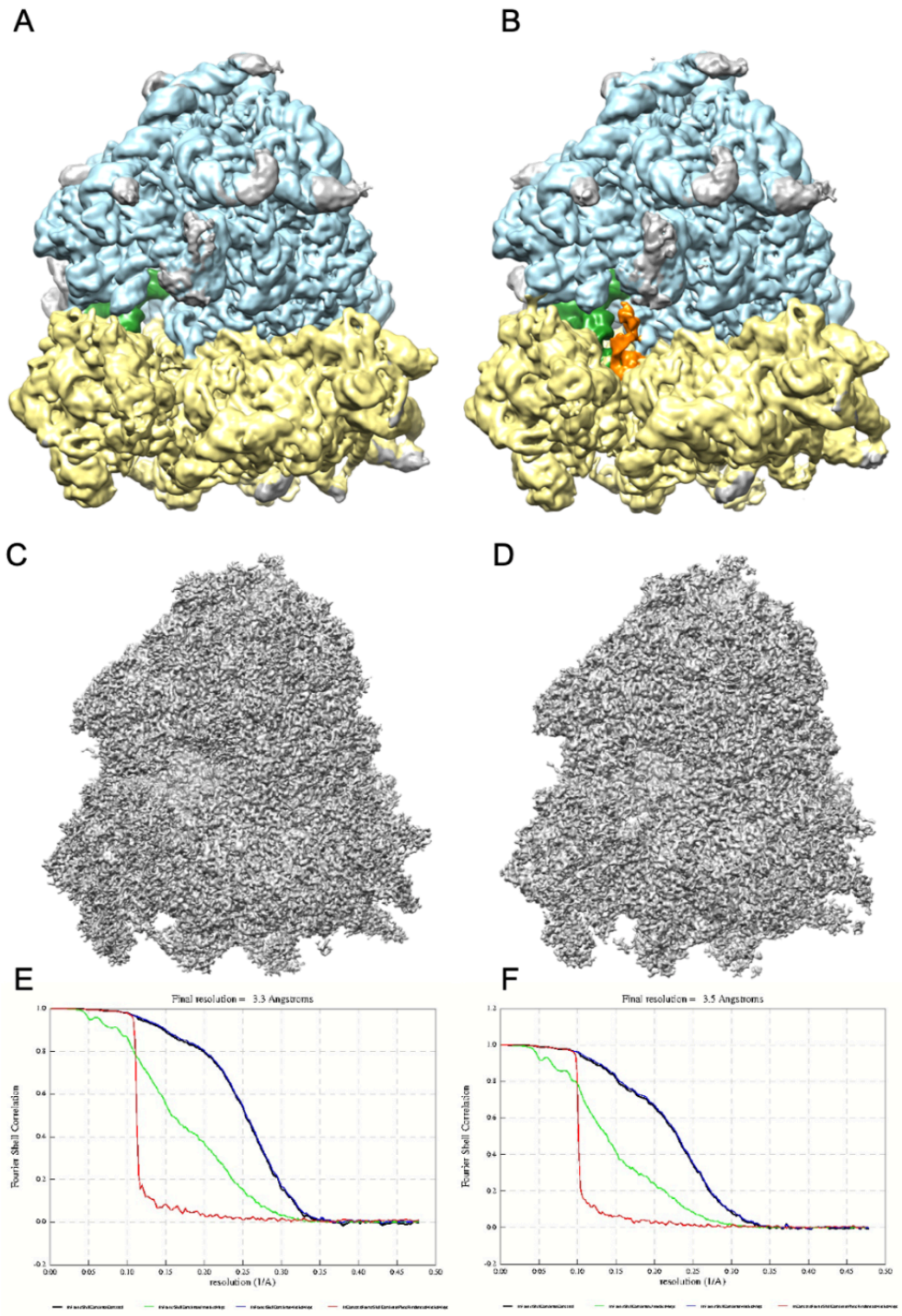

Figure 32. Naked mole-rat 80S nonrotated and rotated ribosomes. (A) Nonrotated map from RELION Refinement (B) Rotated map from RELION Refinement (C) Postprocess nonrotated map (D) Postprocess rotated map (E) Global resolution estimate for the nonrotated map (F) Global resolution estimate for the rotated map.

Modeling of the naked mole-rat 80S ribosome

The previously modelled large subunit (60S) of the naked mole-rat (Chapter 3) was refined against the maps of nonrotated and rotated conformational states of the ribosome in PHENIX [83]. The small subunit (40S) was then built using the rabbit PDB 5LZS [77] as an initial starting point. Sequences were obtained from the National Center for Biotechnology Information protein databases (www.ncbi.nlm.nih.gov) and RNAcentral (https://rnacentral.org/) for the proteins and rRNA of the naked mole-rat, respectively. The proteins and rRNA were computationally mutated to match their sequences and manually adjusted to the fit the density map in Coot [82]. The high-resolution density map and the base-pairing principle allowed us to identify r-protein side-chains and nucleotide bases, respectively. This small subunit (40S) model was refined against the maps for nonrotated and rotated conformational states of the ribosome by real-space refinement (phenix.real_space_refine) [83]. The independently refined models of the ribosomal subunits were than merged and re-refined into the naked mole-rat 80S cryo-EM densities. Ligands (E/E and P/E configured tRNAs) were then subsequently modelled. The 80S models with ligands was then re-refined into the cryo-EM densities. The A/P tRNA remains to be modelled in the rotated state of the ribosome. The two conformational states of the ribosome underwent further altering rounds of manual model adjustment using COOT and model refinement in PHENIX. The models have been evaluated using MolProbity [118].

Parameter	Nonrotated 80S	Rotated 80S
All-atom Clashscore	8.5%	12%
Ramachandran plot	(Below)	(Below)
Outliers	2.5%	3.2%
Allowed	5.6%	9.6%
Favored	93.3 %	90.4%

Figure 33. MolProbity statistics of the naked mole-rat 80S structures.

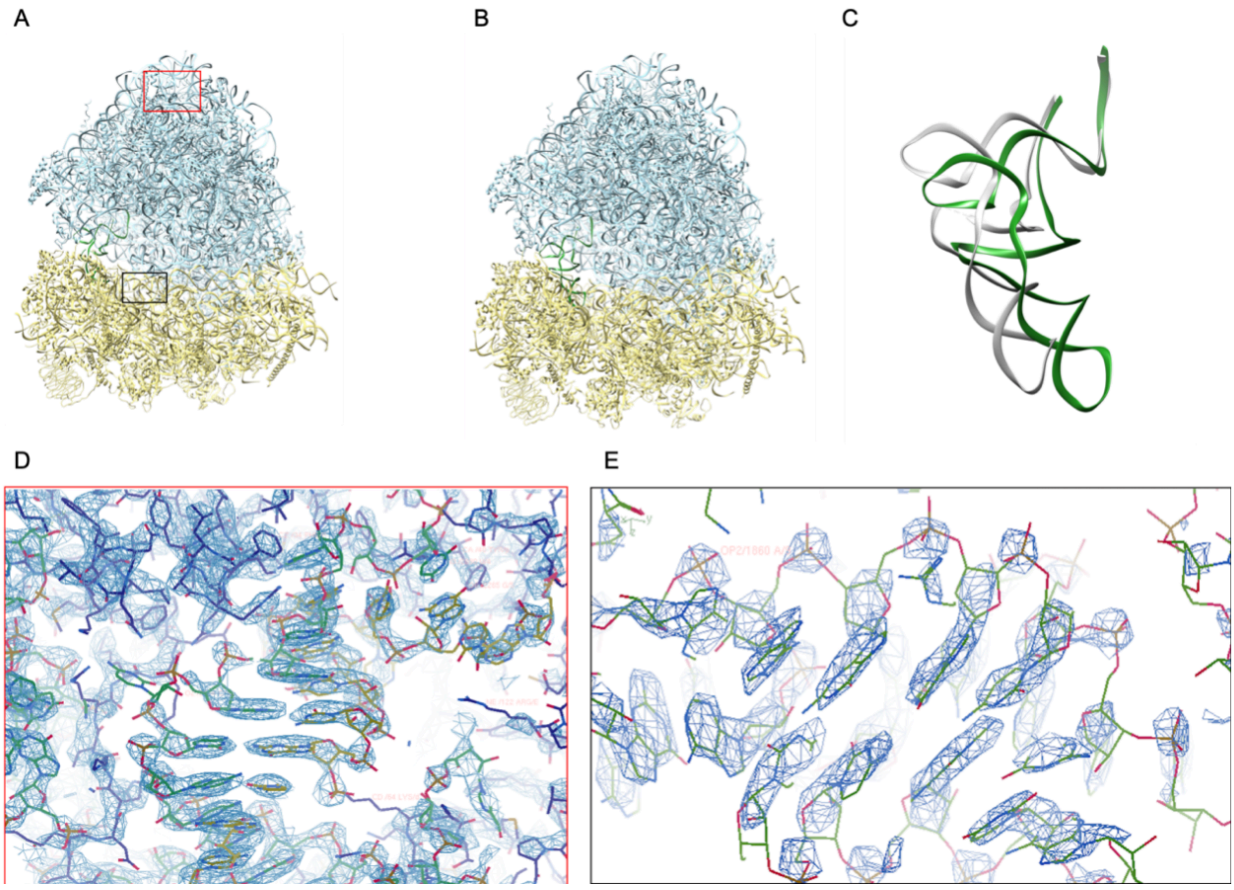

Figure 34. Naked mole-rat 80S ribosomes. (A) Nonrotated 80S model (B) Rotated 80S model (C) Overlay of E/E and E/P configured tRNA from the two conformational states, fitted to the large ribosomal subunit (60S) (D) High-resolution features (red box) of the large subunit (60S), RMSD 5 (E) High-resolution features (black box) of the small subunit (40S), RMSD 4.

Chapter 5: Structure and assembly model for the *Trypanosoma cruzi* 60S ribosomal subunit

This chapter reproduces a paper published in Proceedings of the National Academy of Sciences (PNAS) by Zheng Liu, Cristina Gutierrez-Vargas, Jia Wei, Robert A. Grassucci, Madhumitha Ramesh, Noel Espina, Ming Sun, Beril Tutuncuoglu, Susan Madison-Antenucci, John L. Woolford, Liang Tong, and Joachim Frank. My contributions to the project were the purification of the *T.cruzi* ribosome, structural analysis and interpretation, making figures and tables, as well as writing the manuscript.

Structure and assembly model for the *Trypanosoma cruzi* 60S ribosomal subunit

Zheng Liu[a,1], Cristina Gutierrez-Vargas[b,1], Jia Wei[b,1], Robert A. Grassucci[a,c], Madhumitha Ramesh[d], Noel Espina[e], Ming Sun[b], Beril Tutuncuoglu[d], Susan Madison-Antenucci[e], John L. Woolford Jr.[d], Liang Tong[b], and Joachim Frank[a,b,c,2]

[a]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032; [b]Department of Biological Sciences, Columbia University, New York, NY 10027; [c]Howard Hughes Medical Institute, Columbia University, New York, NY 10027; [d]Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA 15213; and [e]Parasitology Laboratory, Wadsworth Center, New York State Department of Health, Albany, NY 12208

Contributed by Joachim Frank, September 1, 2016 (sent for review July 12, 2016; reviewed by Shulamit Michaeli and James R. Williamson)

Ribosomes of trypanosomatids, a family of protozoan parasites causing debilitating human diseases, possess multiply fragmented rRNAs that together are analogous to 28S rRNA, unusually large rRNA expansion segments, and r-protein variations compared with other eukaryotic ribosomes. To investigate the architecture of the trypanosomatid ribosomes, we determined the 2.5-Å structure of the *Trypanosoma cruzi* ribosome large subunit by single-particle cryo-EM. Examination of this structure and comparative analysis of the yeast ribosomal assembly pathway allowed us to develop a stepwise assembly model for the eight pieces of the large subunit rRNAs and a number of ancillary "glue" proteins. This model can be applied to the characterization of *Trypanosoma brucei* and *Leishmania* spp. ribosomes as well. Together with other details, our atomic-level structure may provide a foundation for structure-based design of antitrypanosome drugs.

ribosome structure | *Trypanosoma cruzi* | biogenesis | multiply fragmented rRNA | antitrypanosome drug design

Ribosomes share a universally conserved core that carries out the fundamental processes of protein synthesis (1). Outside of this core, ribosome composition varies considerably. The main differences among eukaryotic ribosomes are due to rRNA expansion segments (ESs) and variations of ribosomal proteins (r-proteins) (2). An extreme case is presented by the ribosome of trypanosomatids, a family of kinetoplastid protozoans the members of which cause human diseases such as Chagas disease (*Trypanosoma cruzi*), sleeping sickness (*Trypanosoma brucei*), and Leishmaniasis (*Leishmania* spp.) (3–5). Trypanosomatid ribosomes have unusually large ESs and r-protein variations (6). Moreover, their large subunit contains eight pieces of rRNA, six of which result from unique cleavages of a precursor rRNA. Upon assembly, these rRNAs fulfill the functions of the 28S rRNA found in other eukaryotes (7, 8). This unique multiple fragmentation strongly suggests that trypanosomatid ribosomes have pronounced differences that could be exploited for drug design.

Results and Discussion

Here, we report the structure of the 60S large ribosomal subunit of *T. cruzi* at a 2.5-Å resolution (Fig. 1; *SI Appendix*, Fig. S2). Our large subunit structure represents a significant improvement over earlier trypanosomatid ribosome models (6, 9) (*SI Appendix*, Fig. S14) and provides extensive details of the constitutive molecular interactions that stabilize the eight pieces of rRNA (Fig. 1 *C* and *D*; *SI Appendix*, Fig. S4). The quality of the EM map is indicated by the clearly resolved density of the nucleotides and the base-pairing and base-stacking in the helical segments of rRNA (Fig. 1*B*; *SI Appendix*, Fig. S3 *A*–*D*). Purine bases can be readily distinguished from pyrimidine bases, and it is also routinely possible to distinguish between G and A (Fig. 1*B*; *SI Appendix*, Fig. S3). Well-defined EM density is also observed for most protein side chains (*SI Appendix*, Fig. S7 and Table S7). In addition, the high resolution allowed us to identify magnesium ions as well as water molecules (Fig. 1*B*; *SI Appendix*, Fig. S3). Importantly, we located many rRNA covalent modifications (Fig. 1*B*). The number of 2′-O methylations (66 sites) remarkably exceeds that in other eukaryotes, which is consistent with bioinformatics analysis (9) and the modifications on the *Leishmania* ribosome structure recently revealed (10).

In general outline, the 60S ribosomal subunit of *T. cruzi* resembles that of yeast. However, the total size of the six rRNA pieces is about one-fourth larger than the 25S rRNA in yeast (4,225 vs. 3,392 nts) (11, 12). This increase is accounted for by the larger size of ESs and the addition of trypanosome-specific ESs. We refer to the six pieces analogous to 28S rRNA as LSU-α, LSU-β, and srRNAs1–4 (13). The corresponding six pieces across the three pathogenic trypanosomatids are shown in *SI Appendix*, Fig. S15. LSU-α roughly corresponds to domains I and II of the yeast 25S RNA, srRNA1 to domain III, LSU-β to domains IV and V, and srRNA2–4 to domain VI (*SI Appendix*, Fig. S6). LSU-α and LSU-β are located on the solvent and interface side, respectively, whereas srRNA1 is located at the bottom of the large subunit (solvent-side view) (Fig. 1*D*; *SI Appendix*, Fig. S4). The srRNAs2–4 are in close mutual contact and situated on the left side of the large subunit under the P stalk (solvent-side view) (Fig. 1; *SI Appendix*, Fig. S4). Most r-proteins are conserved between yeast and *T. cruzi* with exceptions typically occurring near the srRNAs. Comparative analysis of trypanosome-specific rRNA–rRNA and r-protein–rRNA interactions with those of other eukaryotes, especially yeast, has allowed us to identify the

Significance

The pathogenic trypanosomatids—*Trypanosoma cruzi*, *Trypanosoma brucei*, and *Leishmania* spp.—are the causative agents of Chagas disease, African trypanosomiasis, and leishmaniasis, respectively. These diseases, with high morbidity and mortality rates, affect millions of people worldwide. Current treatments typically use drugs with high toxicity and marginal efficacy. Here we present, a 2.5-Å structure of the *T. cruzi* ribosome large subunit by single-particle cryo-EM. Our structure highlights distinctive trypanosome interactions and has allowed us to propose a tentative model for assembly of the 60S large ribosomal subunit. These atomic details highlighting trypanosome-specific interactions and the differences between *T. cruzi* and the human ribosome can be used directly for structure-based drug design of antitrypanosome drugs.

Author contributions: Z.L. and J.F. designed research; Z.L., C.G.-V., J.W., R.A.G., and L.T. performed research; N.E. and S.M.-A. contributed new reagents/analytic tools; Z.L., C.G.-V., J.W., M.R., M.S., B.T., J.L.W., L.T., and J.F. analyzed data; and Z.L., C.G.-V., J.W., M.R., B.T., S.M.-A., J.L.W., L.T., and J.F. wrote the paper.

Reviewers: S.M., Faculty of Life Sciences, University Ramat-Gan Israel; and J.R.W., The Scripps Research Institute.

The authors declare no conflict of interest.

Data deposition: The data reported in this article have been deposited in The Electron Microscopy Data Bank (accession no. EMD-8361) and the Protein Data Bank (PDB ID code: 5T5H).

[1]Z.L., C.G.-V., and J.W. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: jf2192@cumc.columbia.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1614594113/-/DCSupplemental.

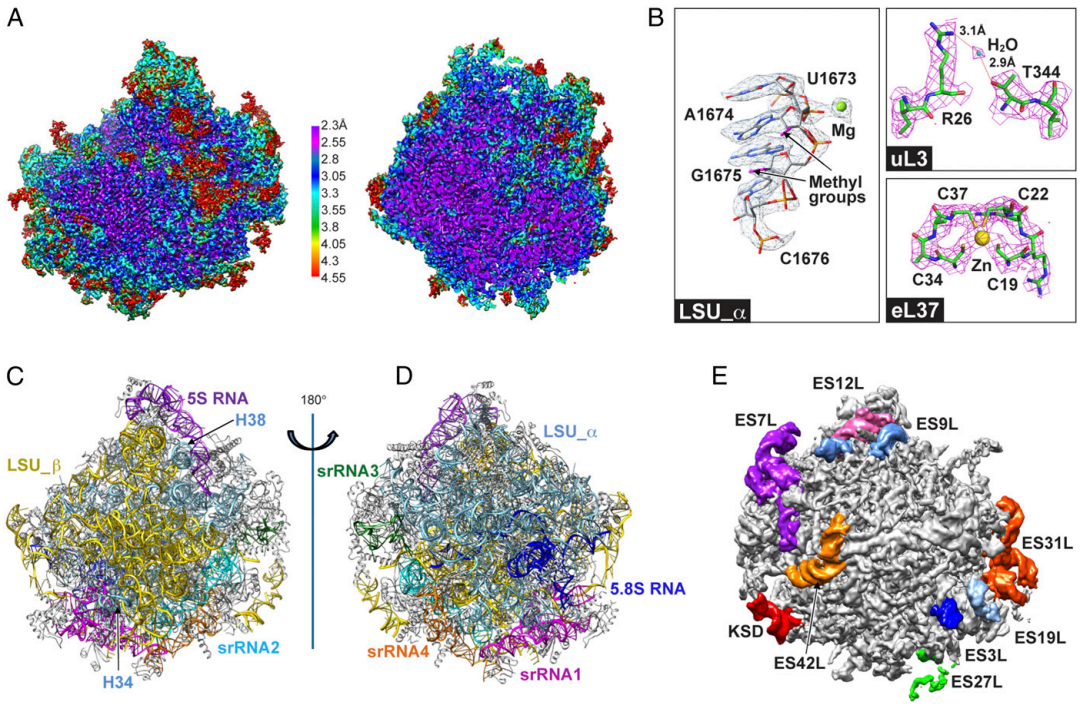

Fig. 1. Structure of the *T. cruzi* large subunit ribosome. (*A*) Cryo-EM map of the 60S subunit after sharpening, colored by local resolution and viewed from the subunit interface. (*Left*) Surface view. (*Right*) Central cut-away view. (*B*) Selected views of density for rRNA and proteins with associated ions and water molecule. (*C*) Some expansion segments in the unsharpened map of a large subunit, viewed from the solvent side. (*C* and *D*) rRNA architecture of the large subunit: interface (*C*) and solvent (*D*) view. (*E*) Some expansion segments in the unsharpened map of a large subunit, viewed from the solvent side.

interactions stabilizing the eight pieces of rRNA to form a mature large subunit. Using our structure as the endpoint of assembly in combination with this comparative analysis has allowed us to arrive at a plausible model for assembly of the eight trypanosomal rRNA pieces.

The rRNA Scaffold. Examination of rRNA–rRNA interactions in our structure reveals that the 5.8S rRNA and the conserved LSU-α and LSU-β domains form the backbone of the rRNA scaffold with two rigid blocks acting as stabilizing core. The first block is formed by the whole 5.8S rRNA and nucleotides in the 5′-end of LSU-α. The second rigid block is constructed by nucleotides in the 3′-end of LSU-α and the 5′-end of LSU-β. Both of the rigid blocks are conserved in yeast; however, in trypanosomes they interact with each other via ES3L of 5.8S rRNA (Fig. 2*A*). Clearly, the 5.8S serves as a critical organizing center for rRNA folding and assembly. The stability of the LSU-α–LSU-β junction is further reinforced by the insertion of two helices, H38 and H34 from LSU-α, into the interface from the solvent side anchoring LSU-α to LSU-β. Mg^{2+} ions and modified nucleotides also mediate contacts in the scaffold (*SI Appendix*, Fig. S11). Along with rRNA–rRNA interactions, r-proteins play important roles in stabilizing this scaffold (*SI Appendix*, Fig. S7).

Structure and Interactions Stabilizing the Small rRNAs. The srRNA1 consists of five helices corresponding to yeast H55-59 and a trypanosome-specific ES, named srRNA1-ES, which adopts a similar position in *T. brucei* (13) and shows highly dynamic behavior, preventing us from modeling it (*SI Appendix*, Fig. S9*A*). The srRNA1 is anchored to the scaffold with the help of srRNA2 via base-pairing interactions. In addition to the contacts with srRNA2, two anchoring proteins, eL19 and eL34, pin H58 and H55 of srRNA1 to the scaffold (Fig. 2*E*). Protein eL34 has adopted an insertion and protein eL19 has a C-terminal extension of about 170 aa, crossing through to the SSU, as already revealed in the structure of *T. brucei*, which suggests that the small subunit bears a role in further stabilizing srRNA1 (13).

The srRNA2 has a scissor-shape architecture and is composed of four helices corresponding to H94-97 in yeast (*SI Appendix*, Fig. S9*B*). H94-95 and H96-97 line up to form the two prongs of the scissor. We found in our structure that srRNA2 has five contact sites with LSU-β. In addition, srRNA2 forms various contacts with LSU-α (Fig. 2*B*). Apart from direct interactions with the scaffold, we found srRNA2 to be further anchored by uL3 onto the scaffold. Specifically, protein uL3 forms a C-shaped cavity, which facilitates the stable assembly of srRNA2 onto the scaffold by accommodating H94 and H96 (Fig. 2*B*; *SI Appendix*, Figs. S8*B* and S9*B*).

The srRNA3 of *T. cruzi* corresponds to yeast ES39L, which is made up of three helices, ES39L1–3, threaded together by single rRNA stretches (*SI Appendix*, Fig. S9*C*). In contrast, srRNA3 is composed of a stretch of single-stranded rRNA plus a helix corresponding to helix ES39L-h2 (*SI Appendix*, Fig. S12*A*). The vacancy resulting from the missing ES39L-h1 adjacent to the sRNA2–4 cleavage sites is filled partially by ES42L, a trypanosome-specific ES. The absence of ES39L-h3 makes room for the kinetoplastid-specific domain (KSD) stretch, the extension of the C-terminal portion of eL14, and the N-terminal domain of eL33. Protein eL33 anchors srRNA3 to the scaffold and contains an

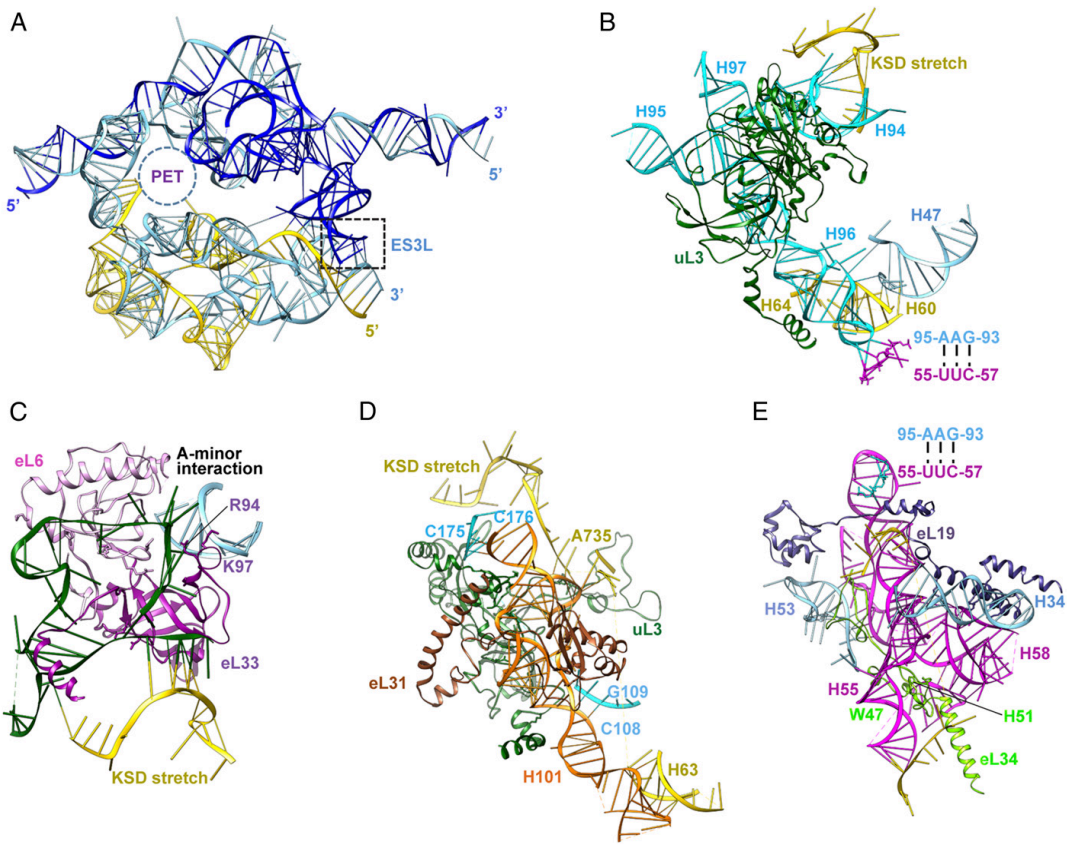

Fig. 2. Interactions stabilizing the srRNAs. (*A*) Scaffold formed by 5.8S rRNA, LSU-α, and LSU-β. PET, peptide exit tunnel. (*B*) srRNA2 (cyan). The components interacting with srRNA2 include the KSD stretch, H60, H64, H47, and uL3. (*C*) srRNA3 (forest green). The rRNA-contacting residues of the anchoring protein, eL6 and eL33, are shown. R94 and K97 are in the trypanosome-specific insertion of eL33. (*D*) srRNA4 (orange). The interacting residues are from the KSD stretch, H63, srRNA2, eL31, and uL3. (*E*) srRNA1 (magenta). W47 and H51 are from the trypanosome-specific insertion of eL34.

insertion in its globular domain that provides extra binding sites for both srRNA3 and LSU-α (Fig. 2*C*; *SI Appendix*, Figs. S8*D*, S9*C*, and S11*A*). Moreover, srRNA3 forms an A-minor interaction with ES7L (Fig. 2*C*) (14), indicating that ES7L also facilitates the assembly of srRNA3 onto the scaffold.

Finally, the srRNA4 is composed of a long helix, corresponding to H101 in yeast, and two short helices (H99 and H100 in yeast) located on one end of H101, as well as an ES, named srRN4-ES, emerging from H101 (*SI Appendix*, Fig. S9*D*). Nucleotides 106–108 of srRNA2 and the exposed positively charged residues of uL3 establish a binding surface for srRNA4 to stack onto (Fig. 2*D*). Furthermore, several residues of the N-terminal extension of eL31 cover H101 of srRNA4. These additions on the N-terminal portion of eL31 exhibit more flexibility compared with the C-terminal region of eL31, which results in only C-A modeling for this region. Also, *T. cruzi* srRNA4 adopts a shorter loop at the base of H100 than its counterpart in yeast, thus avoiding conflict with the N terminus of eL31 (*SI Appendix*, Fig. S9*D*). In addition, residues 1561–1564 of the KSD stretch provide two binding sites with H99 from srRNA4 (Fig. 2*D*).

From the arrangement of the srRNAs revealed in our structure, we see that both srRNA1 and srRNA4 require srRNA2 to be in its final position to provide binding sites for their stable assembly onto the scaffold (Fig. 2 *B*, *D*, and *E*). Interestingly, the known relative time points of assembly of the r-proteins anchoring these small pieces correlates with this arrangement. Studies in yeast have shown that r-proteins, localized in specific positions within the mature 60S subunit, can be classified as early, middle-, or late-acting proteins, thus providing a temporal hierarchy of 60S assembly (15–17). As described earlier, we found in our structure that srRNA2 is anchored by uL3 and srRNA3 is anchored by eL6 and eL33 (Fig. 2 *B* and *C*), all of which are early acting proteins whereas proteins eL19, eL34, and eL31 (Fig. 2 *D* and *E*; *SI Appendix*, Fig. S8 *E*–*G*), which are identified as middle-acting proteins in yeast, interact with srRNA1 and srRNA4 (Fig. 2 *D* and *E*) (17, 18). Thus, the arrangement of the srRNAs as well as their anchoring proteins help define the likely order of assembly of the srRNAs, such that stable association of srRNA2 and srRNA3 occurs before that of srRNA1 and srRNA4. Indeed, deletion of the counterpart of srRNA3 in yeast (ES39L) leads to an early processing defect (19). Strikingly, although srRNA1 is the first small rRNA transcribed, it is stably assembled after srRNA2 and srRNA3, suggesting that its RNA processing occurs during or after assembly of the ribosome instead of the 5′ to 3′ direction following transcription. Furthermore, this

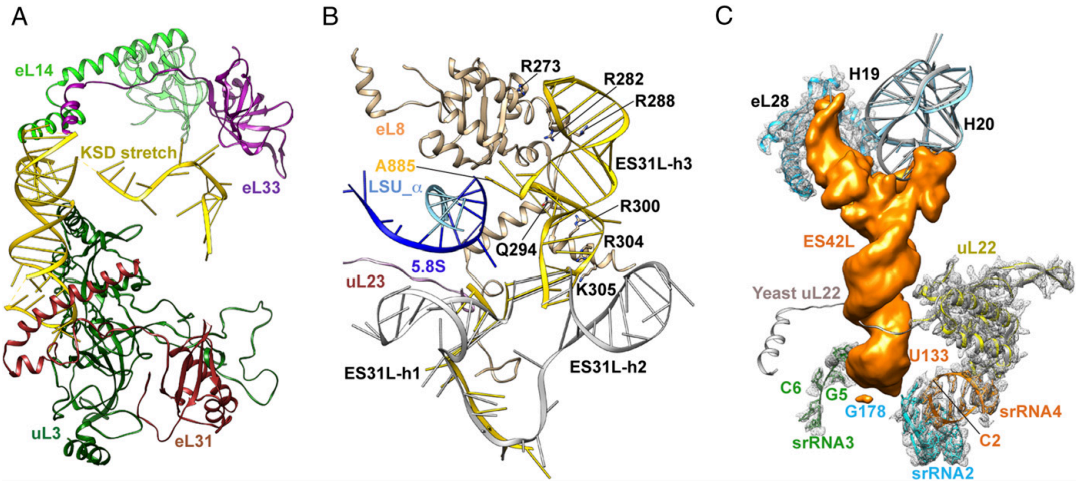

Fig. 3. *T. cruzi*-specific expansion segments. (*A*) KSD contacts uL3, eL31, eL33, and eL14. (*B*) ES31L (yellow) in *T. cruzi* has the specific helix ES31L-h3 contacting eL8 and the 5′ end of 5.8S rRNA. ES31L-h1 and -h2 are only partially modeled. For comparison, ES31L from yeast is also shown here (in gray). (*C*) ES42L, emerging from H19, covers the cleavage sites of srRNA2–4 and blocks the extension of uL22.

assembly order is also consistent with the proposed order of assembly of the three focal points on the *Leishmania* ribosome structure (10) (*SI Appendix*, Fig. S6).

Prominent Expansion Segments and the 5S rRNA. The trypanosome-specific ESs, ES42L and the KSD, both located on the surface of the ribosome, bundle srRNAs2–4. The KSD is a binding partner of the sRNA2- and srRNA4-anchoring proteins, uL3 and eL31, respectively (Fig. 3*A*). This places KSD in a pivotal position interacting with ribosomal functional centers (*SI Appendix*, Table S6) (20, 21). ES42L covers the hub region encompassing the cleavage sites of srRNAs2–4 (Fig. 3*C*; *SI Appendix*, Fig. S12), where the extension of uL22, a component of the peptide exit tunnel (PET), would reach toward ES39L in yeast. The location of ES42L suggests that it may also play a role in the remote regulation of PET by uL22 (22) (Fig. 3*C*; *SI Appendix*, Fig. S12). Both ES42L and KSD envelope small rRNAs2–4 and make additional contacts to stabilize these small rRNA pieces (Figs. 1*E* and 3 *A* and *C*), which suggests that their assembly occurs after that of srRNAs2–4.

Another prominent ES, ES31L, emerges from the base of the L1 stalk and contains an additional helix compared with yeast (Fig. 3*B*). This helix contacts extensions of eL8 and interacts with the 3′-end of the 5.8S rRNA. In yeast, deleting ES31L or the N-terminal extension of eL8 blocks ribosome assembly at later stages, suggesting that ES31L may be involved in subsequent stabilization steps of the previously established scaffold (23). Also associated with the later stages is the final positioning of the 5S rRNA. As in other eukaryotes, the *T. cruzi* 5S rRNA is independently transcribed and associates with uL5 and uL18 forming the 5S ribonucleoprotein particle (RNP) (24). In yeast, this 5S RNP associates with the body of the large subunit during early steps of biogenesis, but is essential only for much later pre-rRNA processing. During this late nuclear stage, the 5S RNP undergoes a 180° rotation before nuclear export (25, 26). The conservation of 5S RNA and its associated proteins uL5 and uL18 (*SI Appendix*, Figs. S10 and S13), as well as identification of trypanosome homologs for 5S RNP assembly factors (27), lead us to hypothesize that similar pathways may exist for maturation of the 5S RNP in trypanosomes at late stages.

Implications for Trypanosomatid Ribosome Assembly. Ribosome biogenesis is a complex, multistep process, requiring the actions of hundreds of assembly factors and small nucleolar RNAs (26, 28). The multiply fragmented rRNAs increase the complexity of *T. cruzi* ribosome biogenesis (27). Based on the above analysis of interactions of individual rRNA components with r-proteins, and on a comparison with the hierarchy of yeast ribosome biogenesis, we propose here a step-wise assembly model for rRNAs of the large subunit (Fig. 4). First, the earlier transcribed components, 5.8S rRNA, LSU-α (domains I, II and part of domain III in yeast), and LSU-β (domains IV, V, and part of domain III), form two large intermingled blocks, and with the addition of the associated r-proteins, the scaffold is constructed. Next, srRNA2 and srRNA3 (both part of domain VI) are positioned in part by their anchoring r-proteins, which are already bound to the scaffold. ES7L is stably assembled concurrently with srRNA3. Subsequently, srRNA1 (the

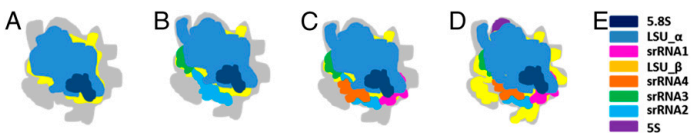

Fig. 4. A model of the proposed *T. cruzi* large subunit rRNAs assembly pathway. (*A*) 5.8S, LSU-α, and LSU-β form a scaffold. The gray color indicates the outline of the completely assembled large subunit. (*B*) srRNA2, -3, and ES7L are assembled. (*C*) srRNA4 and srRNA1 are assembled next. (*D*) 5S rRNA, ES42L, ES31L, and the KSD properly assemble to complete the large subunit. (*E*) Color key for the eight pieces of rRNA.

rest of domain III) and srRNA4 (the rest of domain VI) are stabilized with the aid of strong binding interactions with srRNA2. ES31L stabilizes the previously established scaffold, and the trypanosome-specific ESs (ES42L and the KSD) stabilize srRNAs2–4. The 5S rRNP containing 5S RNA, having assembled with the body of the large subunit earlier, is then rotated to its final position, enabling subsequent maturation steps, concurrent with proper folding of the large ESs on the periphery.

Conclusion. Our high-resolution cryo-EM density map has revealed the structure of the 60S large ribosomal subunit in *T. cruzi* in unprecedented detail, including the rRNA nucleotides, amino acid side chains, and rRNA modifications as well as small molecules (metal ions, and water molecules). We present an analysis of distinctive trypanosome interactions and propose a tentative model for assembly of the 60S large ribosomal subunit. These results from trypanosomatid ribosomes can be used to further understand how different rRNA domains are compacted as assembly proceeds in yeast and other eukaryotes. The multiple fragmentation of the *T. cruzi* ribosome is conserved within the whole trypanosomatid family (*SI Appendix*, Fig. S15 and Table S1), and thus our assembly model can be applied to the characterization of *T. brucei* and *Leishmania* spp. ribosomes as well. Moreover, trypanosomatids, as a group of protozoa early diverged from the eukaryotic lineage, are a compelling model to study the variability of highly conserved processes such as protein translation. Our structure thus serves as a basis for future studies of trypanosome-specific features in protein synthesis. Most importantly, the atomic details highlighting the differences between the *T. cruzi* and the human ribosome can be directly used for structure-based design of antitrypanosome drugs (*SI Appendix*, Tables S9 and S10).

Materials and Methods

***T. cruzi* Culture and Ribosome Purification.** *T. cruzi* Y-strain epimastigote cultures were grown at 27 °C in American Type Culture Collection 1029 liver infusion trypsose medium (9 mg/ml liver infusion broth, 5 mg/ml tryptose, 17 mM NaCl, 56 mM Na_2HPO_4, 5 mM KCl, 5 mM glucose, 10% heat inactivated fetal bovine serum, 0.01 mg/ml hemin) to a concentration of 3 × 107/mL in 1-L glass bottles. A total of 12–18 L of culture was chilled on ice, and cells were harvested by centrifugation at 1,000 × *g* for 20 min. Cell pellets from 5 L of culture were pooled and washed three times in ice-cold buffer containing 0.25 M sucrose and 5 mM KCl. Washed pellets were resuspended in 30 mL of buffer containing 0.25 M sucrose, 5 mM KCl, 5 mM EGTA, 1× protease inhibitor mixture (Sigma P2714), and 0.5% Triton X-100. Lysis was achieved by five freeze–thaw cycles using a dry ice/ethanol bath and warm water. Following lysis the sample was frozen at −80 °C.

Ribosomes were prepared based on previous purification protocols (29) with slight modifications. Cell lysate was loaded onto a 30% (wt/vol) sucrose cushion prepared in buffer I (containing 20 mM Hepes-KOH, pH 7.5), 100 mM KOAc, 4 mM Mg(OAc)2, 20% glycerol, 2 mM DTT, and Roch EDTA-free protease inhibitor mixture) and was centrifuged at 33,800 × *g* for 18 h at 4 °C. The resulting ribosome-enriched pellets were further purified using the 20-K PEG precipitation methods described previously (29). The final pellets were suspended in buffer II (10 mM Hepes-KOH, pH 7.4, 50 mM KOAc, 10 mM NH4Cl, 5 mM Mg(OAc)2, 2 mM DTT, and protease inhibitor mixture) and kept at −80 °C for further use.

Electron Microscopy. Three microliters of ∼80 nM (0.3 mg/mL) purified ribosomes were applied onto a holey carbon grid (carbon-coated Quantifoil 2/2 grid, Quantifoil Micro Tools). The grids were coated by an additional pre-floated continuous thin layer of carbon and glow-discharged before use. The sample-loaded grids were blotted for 3 s at 4 °C in 100% humidity and vitrified by plunging rapidly into liquid ethane cooled by liquid nitrogen at −180 °C with a Vitrobot (FEI). Prepared grids were quickly transferred to a storage box and stored in liquid nitrogen for imaging.

Micrographs were acquired with a back-thinned FEI Falcon II detector on a TITAN KRIOS (FEI) operating at 300 kV with the automatic imaging software FEI EPU. The defocus ranged from −1.0 to −3.0 μm, and the nominal magnification was 143,000× (133,970 after calibration), resulting in a final pixel size of 1.045 Å. Sixteen frames per second were collected with a total dose of about 32e-/Å2 in integrating mode—the only mode available for this camera. To select the grid for imaging, four grids were loaded into the microscope column and subjected to careful screening. Only grids presenting thin ice and good particle distribution were used for data collection. Before data collection, the microscope was carefully aligned including the Cs corrector to minimize electron optical aberrations, linear geometrical distortions, and coma. During the imaging session, micrographs were monitored for evidence of bad pixels, unusually large drift, beam shift, or astigmatism, in which case corrective actions were taken including regaining reference, skipping the squares exhibiting drift, and realigning the microscope. About 11,000 micrographs were collected from a total of two grids during 5 consecutive days.

Image Processing. Dose-fractionated frame stacks were aligned using the program dosefgpu_driftcorr (30). The averaged micrographs were first screened with Arachnid (31), a python version of SPIDER (32), to exclude micrographs with obvious drift, ice contamination, or sparse coverage with particles. The contrast transfer function (CTF) parameters of the remaining good micrographs were calculated using CTFFIND3 (33). Particle selection was performed using the program e2boxer.py (34). About 700,000 selected ribosome-like single partides were subjected to 3D classifications with RELION (35) using a 70S ribosome map (EMD-6315) (36) filtered to 60 Å as the starting reference (*SI Appendix*, Fig. S1). About 400,000 particles were kept after excluding bad particles. One further classification with 10 classes was performed to separate ribosomes in different binding states in the dataset, yielding 60S subunits, E-tRNA–bound ribosomes, and empty ribosomes as the only relevant distinctions. No other tRNAs or factors such as eEF2 or eEF1A were found bound to the ribosomes. One resulting dominant dass of 80S-E-tRNA including 79,000 particles was auto-refined and yielded a 3.1-Å resolution map showing a rigid 60S subunit and a less ordered 40S subunit, which indicated residual heterogeneity in the dataset.

To improve the resolution of the reconstruction, we used several image-processing strategies. First, all of the 60S subunit-containing particles from the first classification were recombined into a pool of about 345,000 particles. These were subjected to focused classification using a large-subunit mask. Before the refinement, the particle dataset was subjected to 2D reference-free classification to eliminate heterogeneous particles. Focused classification and reconstruction were used to overcome heterogeneity in the dataset and even in the classified subset (37). To further improve the homogeneity, three sets of partide positional parameters were obtained by refining each particle with three independent references, which were generated from 3D classification, and only the particles having consistent views were kept (38). In this way, about 4% of the particles were excluded. Second, the micrographs and their power spectra were visually reassessed to exclude those with contamination by ice and ethane, uncorrected drift, and astigmatism, ensuring that only partides extracted from high-quality micrographs were kept. Third, CTF values were recalculated using another program, *S2focus* (39), and the resultant values were cross-validated with those previously determined by CTFFIND3. Micrographs with discrepancies in defocus larger than 500 Å were excluded—in our case, a total of 480 micrographs (4.4%). Fourth, particles for final reconstruction were extracted from good micrographs averaged with weighting from frames 3–10, taking into account the fact that the earlier frames have more drift and the later frames experience more radiation damage.

In this manner, about 235,000 particles were used to reconstruct the large subunit to 2.5 Å. Resolution was estimated using the 0.143 Fourier shell correlation criterion computed from two independently reconstructed maps, following the "gold standard" protocol (40) (*SI Appendix*, Fig. S2). In addition, the two independently reconstructed maps were also used to estimate the distribution of local resolution using Resmap (41) (Fig. 1*A* and *SI Appendix*, Fig. S2 *B–D*). The resolution estimates were in agreement with visual assessment of map quality for secondary structure elements (e.g., appearance of α-helices and β-sheets), individual amino acid residues and nucleotides, small molecules such as water, and methyl groups (*SI Appendix*, Fig. S2 *B–D*). The accuracies of rotation and translation from final reconstruction reported by RELION were 0.52° and 0.49 Å, respectively. The resolution of the final reconstruction without masking was 3.0 Å. RELION postprocessing with auto_mask, auto_bfac, and the inimask_threshold 0.035 (which yields a masked map of the 60S subunit approximately enclosing the atomic model) determined the B-factor as 58.5 $Å^2$ and the resolution of the masked region as 2.5 Å. The resolution of the whole map of the 80S ribosome including the disordered small subunit was determined as 2.8 Å (*SI Appendix*, Fig. S2).

Modeling. A multimap modeling strategy was used in model building (42). The crystal structure of the yeast large ribosomal subunit (PDB: 4V88) (12) and the high-resolution cryo-EM model of the *T. brucei* large ribosomal subunit (PDB: 4V8M) (13) were used as references for building atomic models of rRNA and ribosomal proteins, respectively. The references were initially fitted as rigid

bodies into the negative B-factor–sharpened density map using Chimera (43). The rRNA sequences were obtained from TriTrypDB Kinetoplastid Genomics Resource (tritrypdb.org/tritrypdb/) (11). Regions of rRNA conserved between *T. cruzi* and yeast were used as starting points, followed by extensive manual building using Coot (44). The high-resolution density map and the base-pairing principle allowed us to identify most of the bases. The polypeptide chains in *T. brucei* were mutated to match their sequences in *T. cruzi* taken from the National Center for Biotechnology Information protein databases (www.ncbi.nlm.nih.gov) and manually adjusted to fit the density map. The atomic model of the large ribosomal subunit was refined in real space using PHENIX (45). In addition to Mg^{2+} and Zn^{2+} ions, a number of water molecules were identified as density globules in the distance range of 2.8–3.4 Å from neighboring amino acids, a typical range for a hydrogen bond, occurring reproducibly in two independent reconstructions. The statistics for reconstruction and model refinement are summarized in *SI Appendix*, Table S2. The modeled parts and unmodeled residues are summarized in *SI Appendix*, Tables S3–S5.

ACKNOWLEDGMENTS. We thank Amédée des Georges for helpful suggestions on image processing; Harry Kao for assistance with computer hardware; Nadia Severina for help in ribosome purification; Melissa Thomas-Baum (Buckyball Design) for assistance with the preparation of figures; and Z. Yu, C. Hong, R. Huang, and H. He for their assistance in the data collection at the Janelia Farm Research Campus of Howard Hughes Medical Institute (HHMI). This study was supported by HHMI and NIH Grants R01 GM29169 (to J.F.), R35GM118093 (to L.T.), and R01 GM028301 (to J.L.W.) and by National Science Foundaton Grant GRFP DGE-11-44155 (to C.G.-V).

1. Schmeing TM, Ramakrishnan V (2009) What recent ribosome structures have revealed about the mechanism of translation. *Nature* 461(7268):1234–1242.
2. Yusupova G, Yusupov M (2014) High-resolution structure of the eukaryotic 80S ribosome. *Annu Rev Biochem* 83:467–486.
3. El-Sayed NM, et al. (2005) The genome sequence of Trypanosoma cruzi, etiologic agent of Chagas disease. *Science* 309(5733):409–415.
4. El-Sayed NM, et al. (2005) Comparative genomics of trypanosomatid parasitic protozoa. *Science* 309(5733):404–409.
5. Ivens AC, et al. (2005) The genome of the kinetoplastid parasite, Leishmania major. *Science* 309(5733):436–442.
6. Gao H, Ayub MJ, Levin MJ, Frank J (2005) The structure of the 80S ribosome from Trypanosoma cruzi reveals unique rRNA components. *Proc Natl Acad Sci USA* 102(29):10206–10211.
7. Gray MW (1981) Unusual pattern of ribonucleic acid components in the ribosome of Crithidia fasciculata, a trypanosomatid protozoan. *Mol Cell Biol* 1(4):347–357.
8. White TC, Rudenko G, Borst P (1986) Three small RNAs within the 10 kb trypanosome rRNA transcription unit are analogous to domain VII of other eukaryotic 28S rRNAs. *Nucleic Acids Res* 14(23):9471–9489.
9. Liang XH, et al. (2005) A genome-wide analysis of C/D and H/ACA-like small nucleolar RNAs in Trypanosoma brucei reveals a trypanosome-specific pattern of rRNA modification. *RNA* 11(5):619–645.
10. Shalev-Benami M, et al. (2016) 2.8-Å Cryo-EM ctructure of the large ribosomal subunit from the eukaryotic parasite Leishmania. *Cell Reports* 16(2):288–294.
11. Aslett M, et al. (2010) TriTrypDB: A functional genomic resource for the Trypanosomatidae. *Nucleic Acids Res* 38(Database issue):D457–D462.
12. Ben-Shem A, et al. (2011) The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science* 334(6062):1524–1529.
13. Hashem Y, et al. (2013) High-resolution cryo-electron microscopy structure of the Trypanosoma brucei ribosome. *Nature* 494(7437):385–389.
14. Nissen P, Ippolito JA, Ban N, Moore PB, Steitz TA (2001) RNA tertiary interactions in the large ribosomal subunit: The A-minor motif. *Proc Natl Acad Sci USA* 98(9):4899–4903.
15. Woolford JL, Jr, Baserga SJ (2013) Ribosome biogenesis in the yeast Saccharomyces cerevisiae. *Genetics* 195(3):643–681.
16. de la Cruz J, Karbstein K, Woolford JL, Jr (2015) Functions of ribosomal proteins in assembly of eukaryotic ribosomes in vivo. *Annu Rev Biochem* 84:93–129.
17. Gamalinda M, et al. (2014) A hierarchical model for assembly of eukaryotic 60S ribosomal subunit domains. *Genes Dev* 28(2):198–210.
18. Pöll G, et al. (2009) rRNA maturation in yeast cells depleted of large ribosomal subunit proteins. *PLoS One* 4(12):e8249.
19. Ramesh M, Woolford JL, Jr (2016) Eukaryote-specific rRNA expansion segments function in ribosome biogenesis. *RNA* 22(8):1153–1162.
20. García-Ortega L, Alvarez-García E, Gavilanes JG, Martínez-del-Pozo A, Joseph S (2010) Cleavage of the sarcin-ricin loop of 23S rRNA differentially affects EF-G and EF-Tu binding. *Nucleic Acids Res* 38(12):4108–4119.
21. Mailliot J, et al. (2016) Crystal structures of the uL3 mutant ribosome: Illustration of the importance of ribosomal proteins for translation efficiency. *J Mol Biol* 428(10 Pt B):2195–2202.
22. Nilsson OB, et al. (2015) Cotranslational protein folding inside the ribosome exit tunnel. *Cell Reports* 12(10):1533–1540.
23. Ramesh M, Woolford JL, Jr (2016) Eukaryote-specific rRNA expansion segments function in ribosome biogenesis. RNA 22(8):1153–1162.
24. Michaeli S, Agabian N (1990) A Trypanosoma brucei small RNP particle containing the 5S rRNA. *Mol Biochem Parasitol* 41(1):7–15.
25. Leidig C, et al. (2014) 60S ribosome biogenesis requires rotation of the 5S ribonucleoprotein particle. *Nat Commun* 5:3491.
26. Wu S, et al. (2016) Diverse roles of assembly factors revealed by structures of late nuclear pre-60S ribosomes. *Nature* 534(7605):133–137.
27. Michaeli S (2011) rRNA biogenesis in Trypanosomes. *RNA Metabolism in Trypanosomes (Nucleic Acids and Molecular Biology)*, ed Binderelf A (Springer, Berlin), Vol 28, pp123–148.
28. Gamalinda M, Woolford JL, Jr (2015) Paradigms of ribosome synthesis: Lessons learned from ribosomal proteins. *Translation (Austin)* 3(1):e975018.
29. Khatter H, et al. (2014) Purification, characterization and crystallization of the human 80S ribosome. *Nucleic Acids Res* 42(6):e49.
30. Li X, et al. (2013) Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods* 10(6):584–590.
31. Langlois R, et al. (2014) Automated particle picking for low-contrast macromolecules in cryo-electron microscopy. *J Struct Biol* 186(1):1–7.
32. Frank J, et al. (1996) SPIDER and WEB: Processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* 116(1):190–199.
33. Mindell JA, Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* 142(3):334–347.
34. Tang G, et al. (2007) EMAN2: An extensible image processing suite for electron microscopy. *J Struct Biol* 157(1):38–46.
35. Scheres SH (2015) Semi-automated selection of cryo-EM particles in RELION-1.3. *J Struct Biol* 189(2):114–122.
36. Li W, et al. (2015) Activation of GTP hydrolysis in mRNA-tRNA translocation by elongation factor G. *Sci Adv* 1(4):e1500169.
37. Brown A, et al. (2014) Structure of the large ribosomal subunit from human mitochondria. *Science* 346(6210):718–722.
38. Liu Z, Guo F, Wang F, Li TC, Jiang W (2016) 2.9 Å resolution Cryo-EM 3D reconstruction of close-packed virus particles. *Structure* 24(2):319–328.
39. Jiang W, Guo F, Liu Z (2012) A graph theory method for determination of cryo-EM image focuses. *J Struct Biol* 180(2):343–351.
40. Henderson R, et al. (2012) Outcome of the first electron microscopy validation task force meeting. *Structure* 20(2):205–214.
41. Kucukelbir A, Sigworth FJ, Tagare HD (2014) Quantifying the local resolution of cryo-EM density maps. *Nat Methods* 11(1):63–65.
42. Brown A, et al. (2015) Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr D Biol Crystallogr* 71(Pt 1):136–153.
43. Pettersen EF, et al. (2004) UCSF chimera: A visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
44. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
45. Adams PD, et al. (2002) PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 58(Pt 11):1948–1954.

SUPPORTING INFORMATION (SI) APPENDIX

Structure and assembly model for the *Trypanosoma cruzi* 60S ribosomal subunit

Zheng Liu[1]*, Cristina Gutierrez-Vargas[2]*, Jia Wei[2]*, Robert A. Grassucci[1,3], Madhumitha Ramesh[4], Noel Espina[5], Ming Sun[2], Beril Tutuncuoglu[4], Susan Madison-Antenucci[5], John L. Woolford, Jr.[4], Liang Tong[2], and Joachim Frank[1,2,3#]

1. Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA
2. Department of Biological Sciences, Columbia University, New York, NY 10027, USA
3. Howard Hughes Medical Institute, Columbia University, New York, NY 10032, USA
4. Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA
5. Parasitology Laboratory, Wadsworth Center, New York State Department of Health, Albany, NY, 12208, USA

*These authors contributed equally to this work.
#To whom correspondence should be addressed. Email: jf2192@cumc.columbia.edu

Author Contributions: J.F. and Z.L. designed the experiments. J.F. supervised the project. S.M.-A. and N.E. provided the *T. cruzi* cell extract. Z.L. and C.G.-V. purified the ribosomes. Z.L. performed the cryo-EM experiments and the image data analysis. R.A.G. assisted in the screening of the samples. J.W., Z.L., M.S., and L.T. performed the structural modeling. Z.L., C.G.-V., J.W., L.T., M.S., M.R., B.T. and J.L.W. analyzed the structure. Z.L., C.G.-V., J.W., L.T., M.S., S.M.-A., M.R., B.T., J.L.W and J.F. wrote the manuscript.

The authors declare no conflict of interest.

Data deposition: All cryo-EM map and model described in this manuscript have been deposited in the Electron Microscopy Data Bank (EMDB) (accession nos. EMD-8361, PDB code: 5T5H)

Supporting Information (SI Appendix)

SI Text: Structure and possible role of Expansion segments

Our structure shows various trypanosome-specific ESs forming contacts with functionally significant proteins that are extended or truncated in *T. cruzi*, when compared to their yeast counterparts. These ESs include ES9L, ES27L, ES7L, ES12L, and ES31L as well as those described in the main text, ES42L and the KSD.

ES7L in *T. cruzi* is composed of two long bent helices while in yeast this ES is made up of only one long helix and a short helix (Fig. S16). The tip of the trypanosome-unique long helix interacts with proteins uL30 and uL4. Protein uL30 is conserved with respect to its yeast counterpart and uL4, situated between the two long bent helices, bears a C-terminal extension (Fig. S16). Considering that uL4 is a component of the protein exit tunnel, this unusual configuration of ES7L has a possible function in sensing or regulating the passage of nascent polypeptides through the tunnel.

T. cruzi ES9L is composed of three helices instead of the two helices seen in yeast ES9L (named ES9L-h1-3). Helix ES9L-h3, with no homolog in yeast, surrounds one side of eL29, which bears an extended C-terminal portion exposed on the surface (Fig. S17). The other sides of eL29 are surrounded by ES9L-h1 and ES12. Together these three helices position the C-terminal of eL29 to be exposed stably on the surface, while the protein's N-terminal domain threads through the rRNA core in the vicinity of the PTC.

ES27L (located below srRNA1, as viewed from solvent side) and ES31L, which emerges from the base of the L1 stalk, have both increased in size compared to their yeast counterparts. The dynamic behavior of ES27L makes it visible only at low thresholds on the unsharpened density map (Fig. 1E). ES27L forms a novel intern subunit bridge as revealed in the *T. brucei* structure (1), and also provides a platform for the S-domain of the signal recognition particle (SRP) (2). Thus, it has been proposed to regulate access of factors to the peptide exit tunnel. ES31L, whose assembly is described in the main text, is anchored at the base of the L1 stalk and is a dynamic element instrumental in evacuating the exit-site tRNA (3). ES31L is composed of three helices instead of two as seen in yeast (Fig. 3B). The third, trypanosome-specific helix is in contact with protein eL8 and runs parallel with the L1 stalk. This helix also interacts with the 3'-end of the 5.8S rRNA via residue 865 (Fig. 3B). Moreover, ES31L contacts the N-terminal extension of uL23, a signaling protein for SRP (4).

Altogether, these trypanosome-specific rRNA ESs form contacts with functionally significant extended or truncated r-proteins (summarized in Table S6), suggesting the architecture of the *T. cruzi* ribosome may be specialized to modulate nascent peptide conducting, protein targeting, and other translational events such as GTPase activation by the sarcin-ricin loop.

Additional References

1. Hashem Y, *et al.* (2013) High-resolution cryo-electron microscopy structure of the Trypanosoma brucei ribosome. *Nature* 494(7437):385-389.
2. Anger AM, *et al.* (2013) Structures of the human and Drosophila 80S ribosome. *Nature* 497(7447):80-85.
3. Korostelev A, Ermolenko DN, & Noller HF (2008) Structural dynamics of the ribosome. *Current opinion in chemical biology* 12(6):674-683.
4. Park E, *et al.* (2014) Structure of the SecY channel during initiation of protein translocation. *Nature* 506(7486):102-106.

Fig. S1

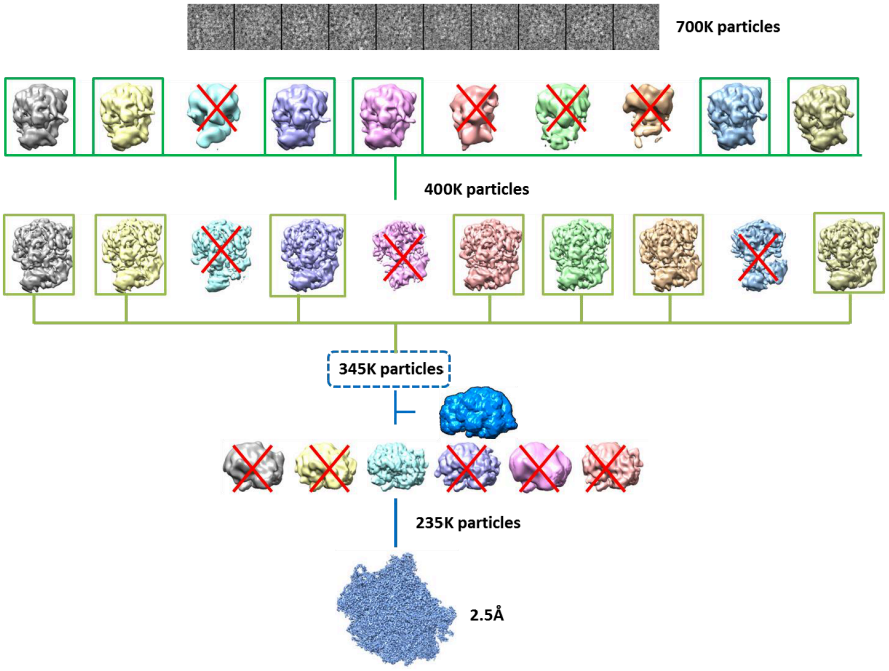

Fig. S1. Classification of the *Trypanosoma cruzi* ribosome cryo-EM data set. First row, ten particles selected from 700,000. Second row, primary classification with K= 10 classes, generating six ribosome-like and four bad classes. Third row, further classification of ribosome-like particles, again with K=10, yielding seven fully formed ribosome classes and three with broken 40S subunit. Fourth row, focused classification using K=6, with 60S subunit mask placed on all particles (345,000), yielding one dominant class. Fifth row, resulting focused reconstruction of the 60S subunit from 235,000 particles.

Fig.S2

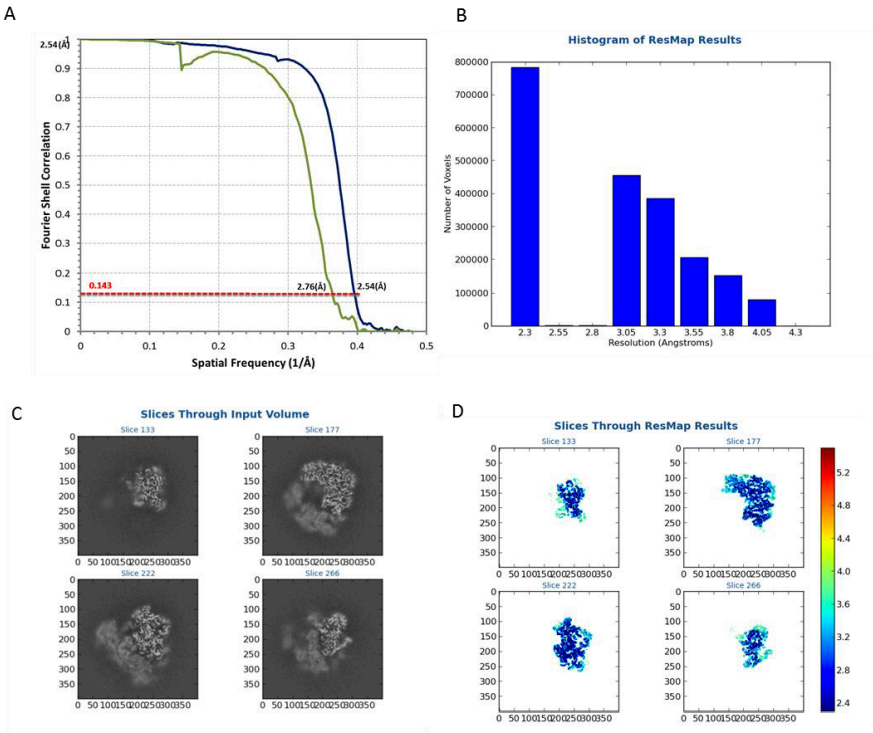

Fig S2. **Resolution estimation of the 3D-reconstruction** (A) FSC of the 60S subunit reconstruction (dark blue: model region; green: whole map). (B-D) Local resolution evaluation using *resmap*. Shown are outputs of the *resmap* software: (B) histogram distribution of resolution values; (C) density sections and (D) resolution distributions in the same set of sections shown in (C).

Fig S3

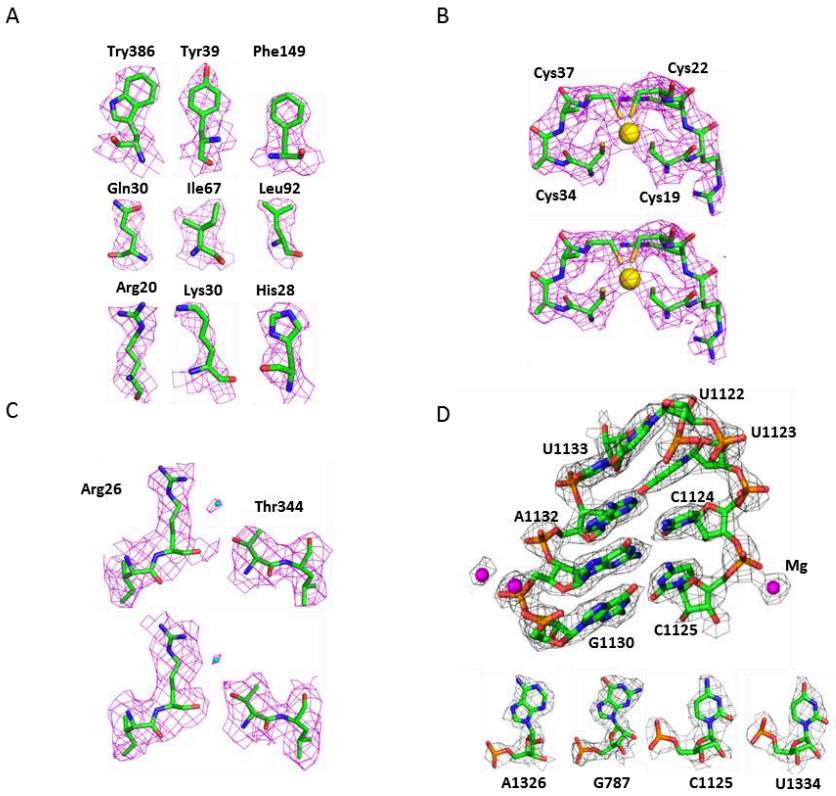

Fig. S3. Side chain features of 60S subunit of *Trypanosoma cruzi* ribosome. (A) Densities for some of the amino acids. Top, three aromatic residues; middle, three hydrophobic residues; bottom, three basic residues. These residues are from the proteins uL3, eL37 and uL15. (B,C) Reproducibility of water molecule (B), and zinc ion (C) assignments through comparison of half-maps. These regions are also displayed in Fig 1.b. (D) Cryo-EM density of a highlighted rRNA region of LSU-α docked with atomic model. Bottom panel, examples for the four nucleotides in rRNA.

Fig.S4

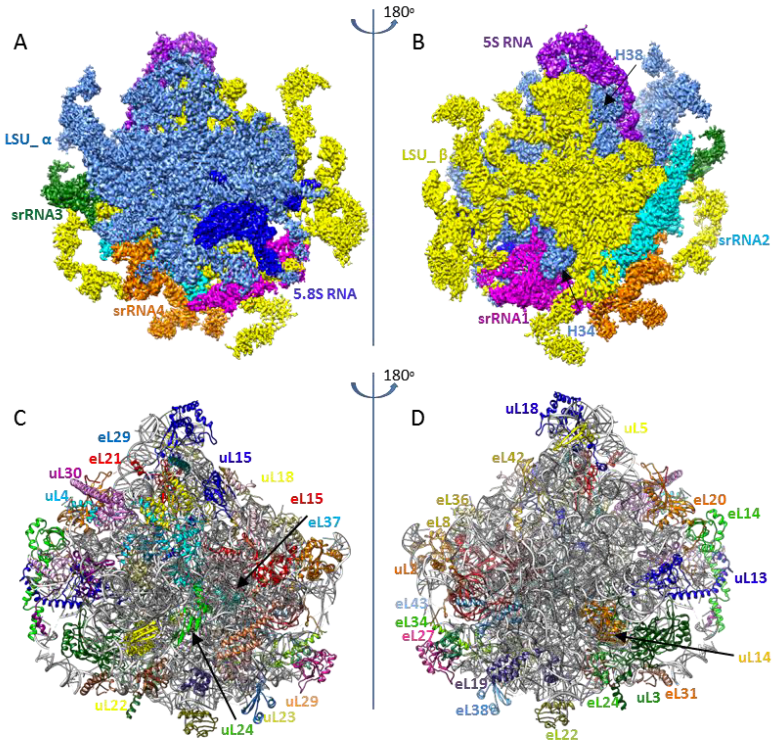

Fig. S4. rRNAs arrangement and atomic model of the 60S subunit. (A, B) Arrangement of the eight pieces of rRNA in the large subunit within the electron density map; r-proteins are omitted for clear display of rRNA. (A) Solvent and (B) interface view. (C, D) architecture of proteins in the large subunit, in solvent (C) and interface (D) view.

Fig.S5

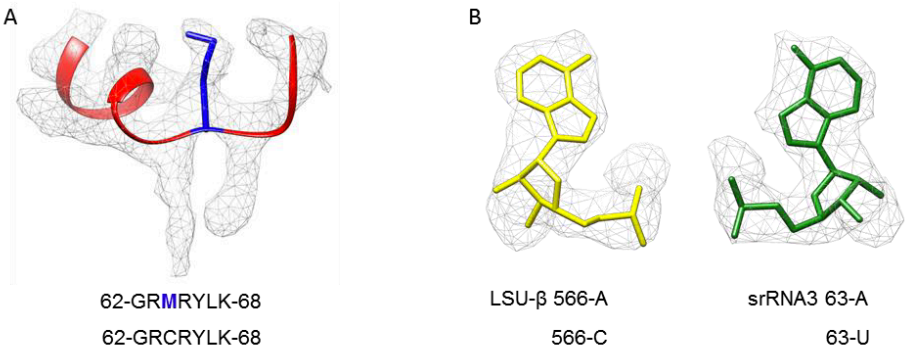

Fig. S5. Cryo-EM allows identification of variations in rRNA and proteins. (A) Density map (gray mesh) of the fragment of a.a. 62-68 on eL37. Met-64 (blue) is clearly represented in the density although the sequence XP_807547.1. has a Cysteine residue (B) Density map (gray mesh) of the A-566 on LSU-β and A-63 on srRNA3, in conflict with the sequences which contain C and U at these locations, respectively.

Fig.S6

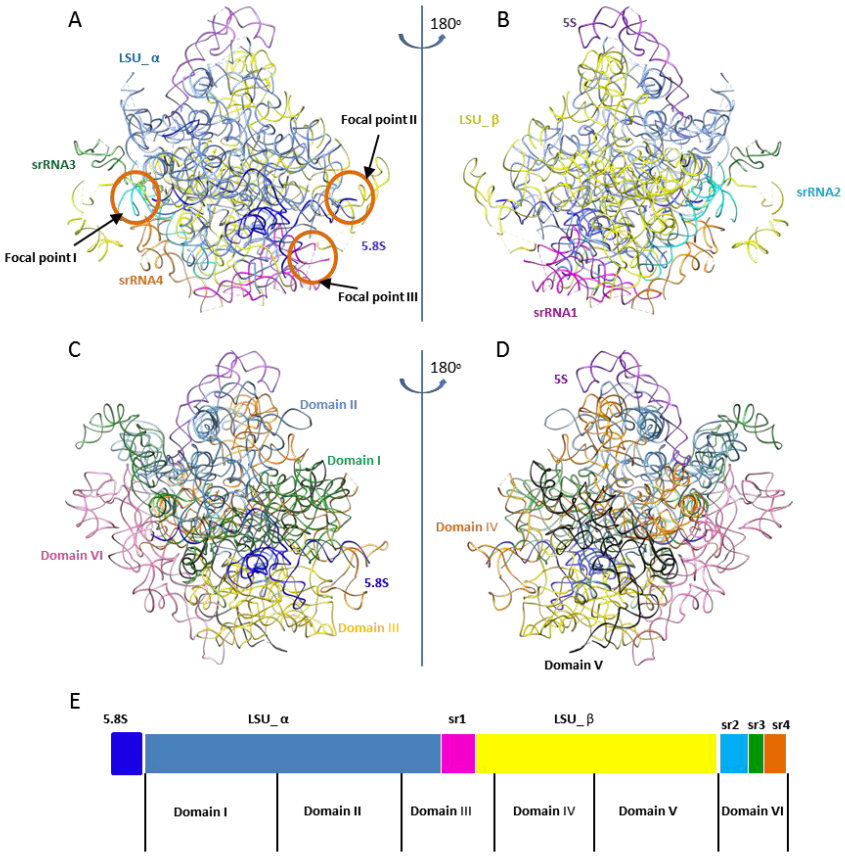

Fig. S6. ***T. cruzi*** **large subunit rRNA pieces and the corresponding rRNA domains in yeast.** (A, B) The eight pieces of rRNA in the *T. cruzi* ribosomal large subunit, in solvent (A) and interface (B) view. (C, D) The domains of rRNA in yeast 25S rRNA, in solvent (c) and interface (D) view. (E) Correspondence between the different pieces of *T. cruzi* rRNA and the domains from yeast 25S rRNA. srRNA1-4 are denoted by sr1-4.

Fig.S7

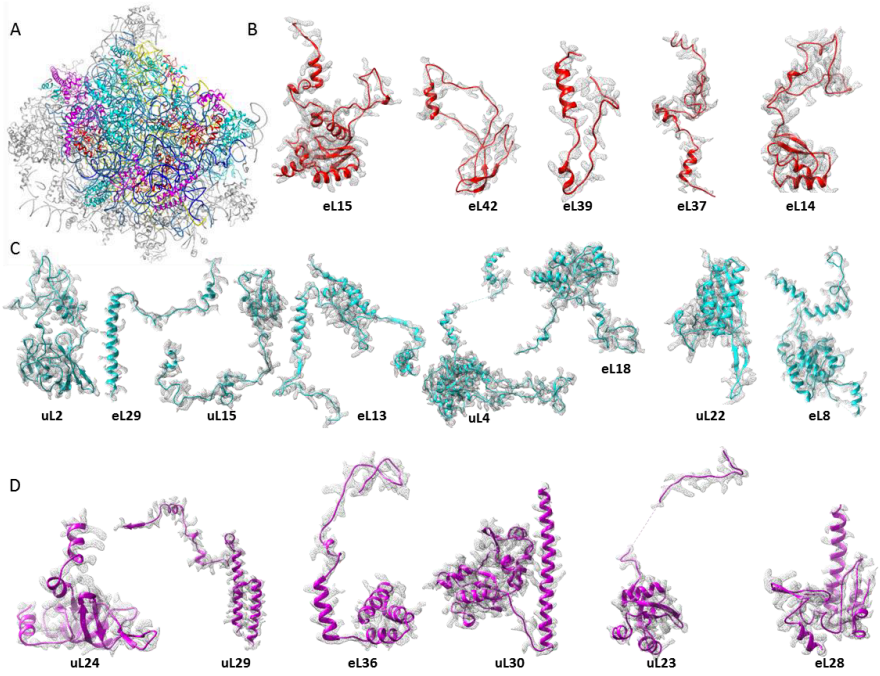

Fig. S7. Proteins for scaffold stabilization. (A) The scaffold encompasses the conserved region of LSU-α (sky blue) and LSU-B (yellow), the whole 5.8S rRNA (blue), and its associated proteins. (B) Group-one proteins, which are buried in the scaffold. (C) Group-two proteins, bearing large spanning domains. (D) Group-three proteins, located on the surface.

10

Fig.S8

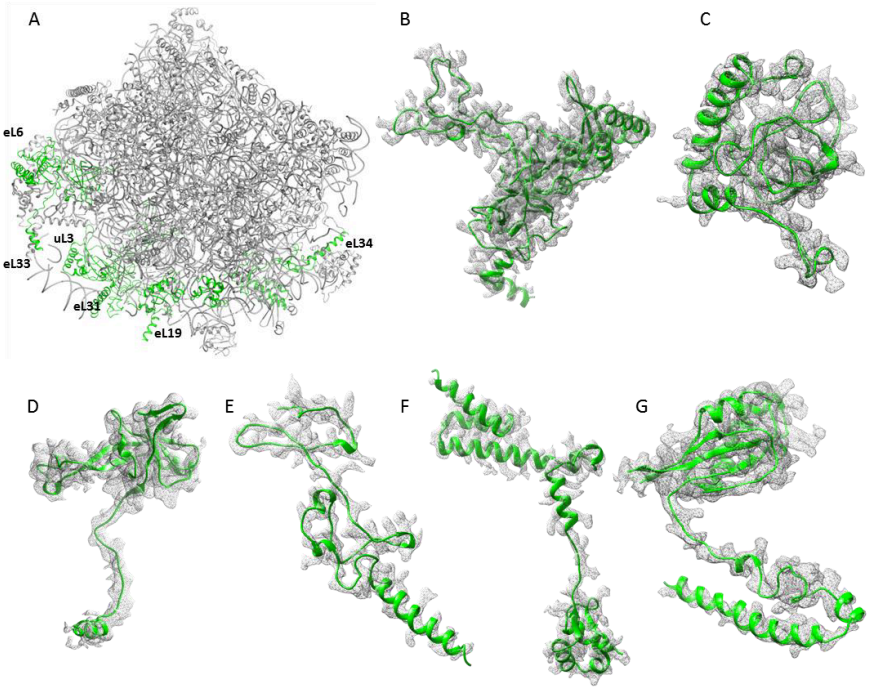

Fig. S8. Anchoring proteins for small-rRNAs assembly. (A) Locations of the anchoring proteins connecting the scaffold and small rRNAs. (B) uL3. (C) eL6. (D) eL33. (E) eL34. (F) eL19. (G) eL31.

Fig.S9

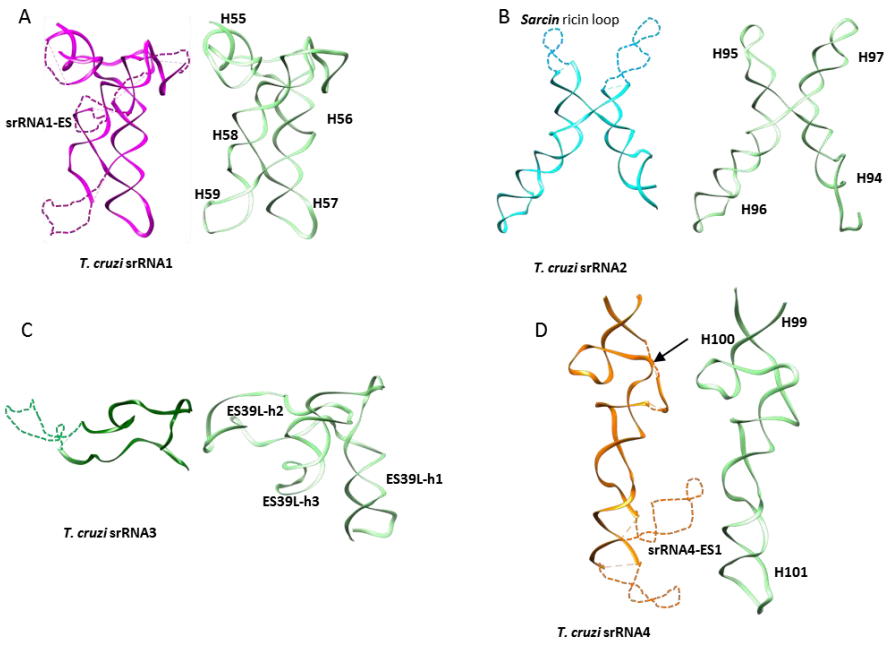

Fig. S9. Small rRNAs. (A) srRNA1 (magenta, left) and its homologous region in yeast (light green). (B) *T. cruzi* srRNA2 (cyan, left) and its homologous region, H94-97 of yeast (light green). (C), *T. cruzi* srRNA3 (forest green) and its homologous region ES39L of yeast (light green). (D) *T. cruzi* srRNA4 (orange) and its homologous region in yeast (light green). The arrow marks the shorter *T. cruzi* loop compared with yeast, which avoids conflict with the N-terminus of eL31. The un-modelled regions of the small rRNAs are connected by a dotted line according to the *T. brucei* model (4V8M.pdb) and our unsharpened density map.

Fig.S10

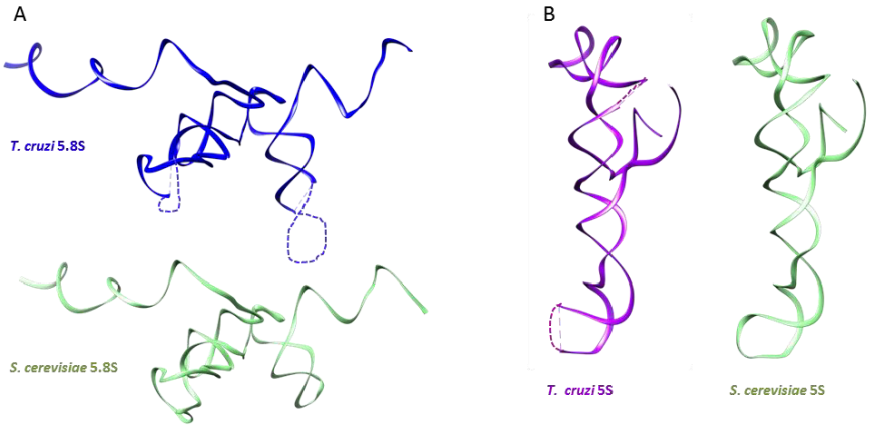

Fig. S10. ***T. cruzi*** **and yeast 5.8S and 5S rRNAs**. (A) 5.8S rRNA, upper is from our *T. cruzi* structure, bottom (light green) is from the yeast ribosome structure. (B) 5S rRNA, left is from our *T. cruzi* structure, right (light green) is from yeast. The un-modelled regions of the small rRNAs are connected by a dotted line according to the *T. brucei* model (4V8M.pdb) and our unsharpened density map.

Fig.S11

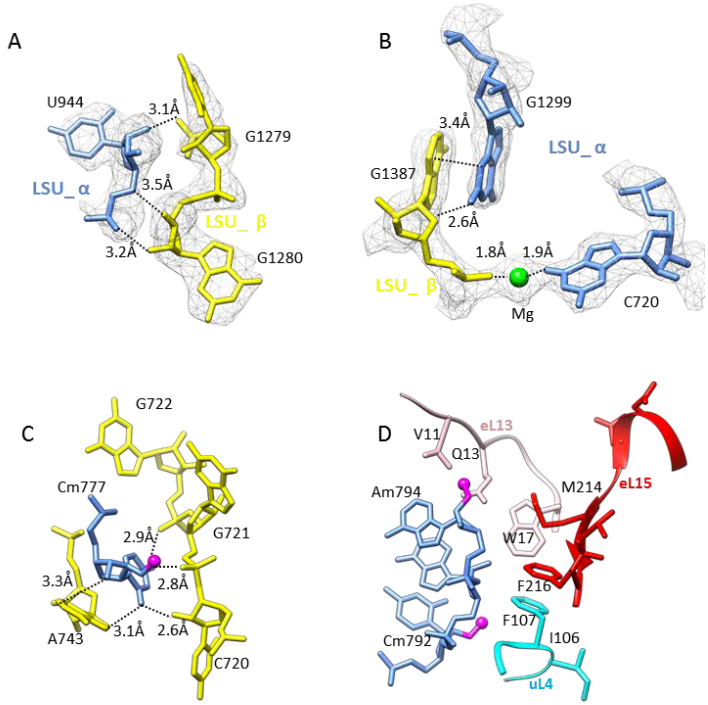

Fig. S11. rRNA and r-protein associations on the large subunit. (A) Interactions between U944 of LSU-α and C1279 and G1280 of LSU-β. (B) Interactions between A1387 of LSU-β and A1299 and G720 of LSU-α; Mg^{2+} takes part in forming the LSU-α - LSU-β interaction. (C) Modified rRNA nucleotides participate in forming interactions between LSU-α and LSU-β. Cm777 is a methyl-modified residue of LSU-α. (D) A hydrophobic region formed by modified nucleotides and r-proteins. The magenta spheres in rRNA model denote methyl groups.

Fig.S12

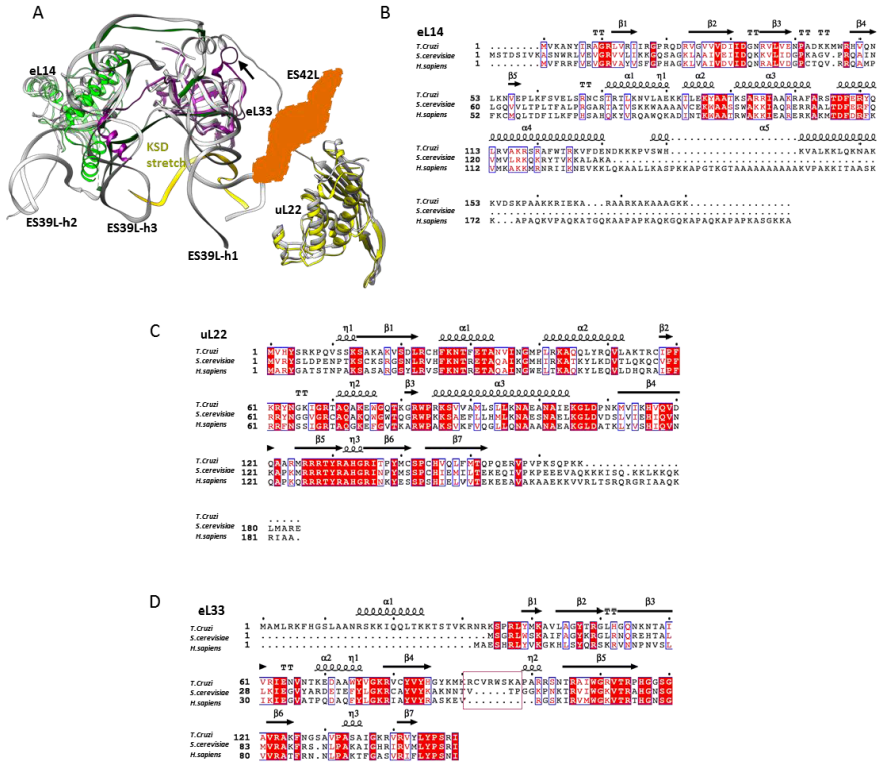

Fig. S12. srRNA3 and its homologous rRNA ES39L in yeast surroundings. (A), Overlay of srRNA3 and its surrounding region, which includes eL14, eL33, and uL22 of *T. cruzi*. For comparison, the ES39L region of yeast is also shown in grey. ES42L has been added by manual drawing. The arrows indicate the insertion of the eL33 (marked by magenta box in (D)), which occupies the position of the C-terminal of uL22 in the yeast ribosome structure. (B) Sequence alignment of r-protein eL14. (C) Sequence alignment of r-protein uL22. (D) Sequence alignment of r-protein eL33.

Fig. S13

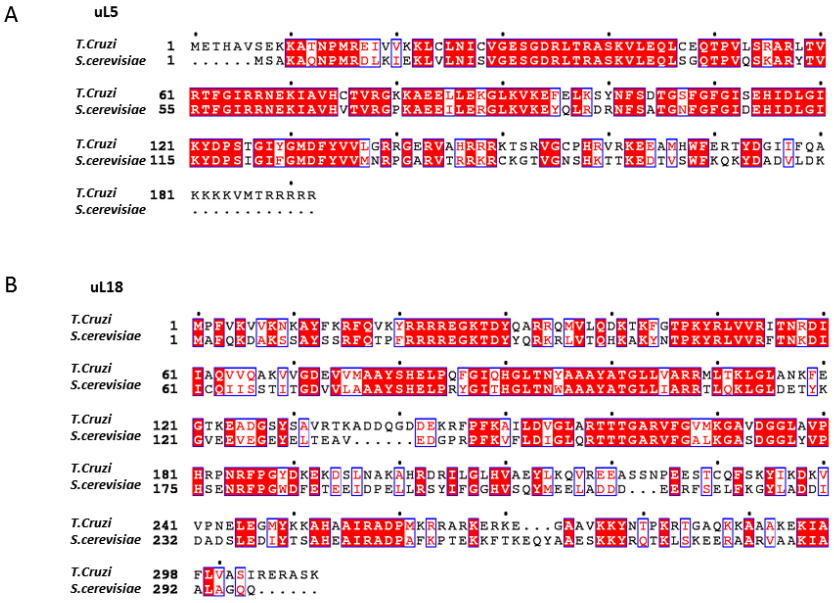

Fig. S13. Sequence alignment of 5S rRNA-associated proteins. (A) Sequence alignment of protein uL5, whose identity between the two sequences is 69%. (B) Sequence alignment of protein uL18, whose identity between the two sequences is 49%.

Fig. S14

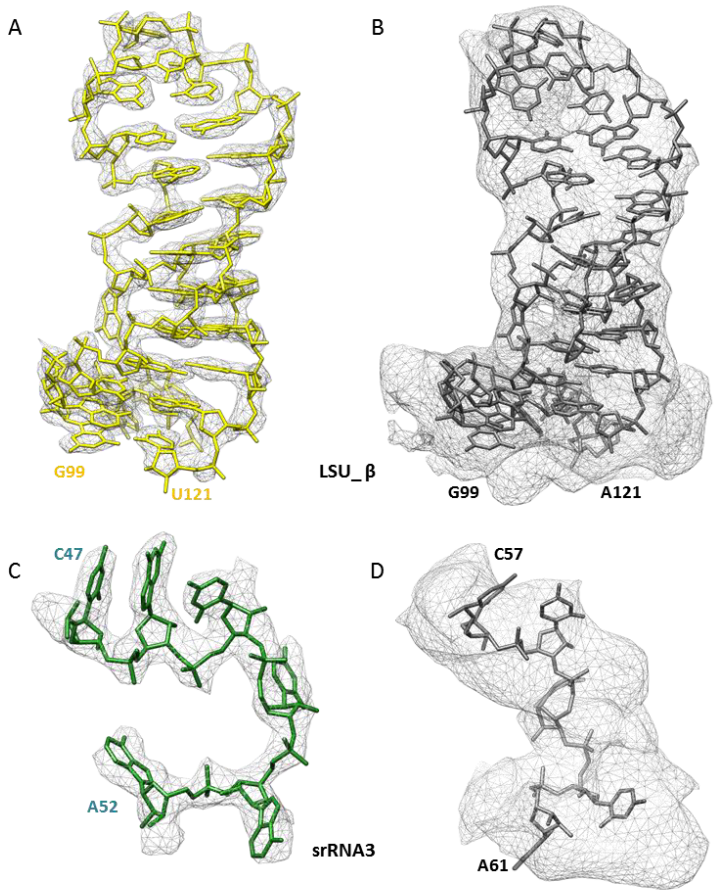

Fig. S14. Comparison of some regions of 2.5 Å *T. cruzi* and 5.5 Å *T. brucei* ribosome density maps, superimposed with their corresponding atomic models. (A) A double-helical region of LSU-β from the *T. cruzi* map and model. (B) The same region from the *T. brucei* map and model as in (A). (C) A single-stranded fragment of srRNA3 from the *T. cruzi* map and model. (D) The same region from the *T. brucei* map and model as in (C).

Fig. S15

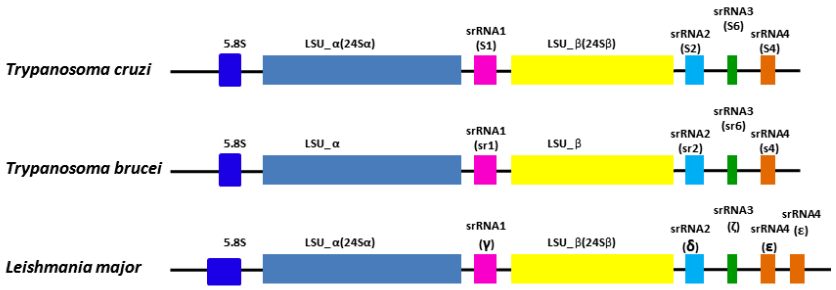

Fig. S15. Schematic representation of the large subunit rRNA genes (except 5S rRNA) from trypanosomatids. The names in the bracket are their respective old nomenclatures.

Fig. S16

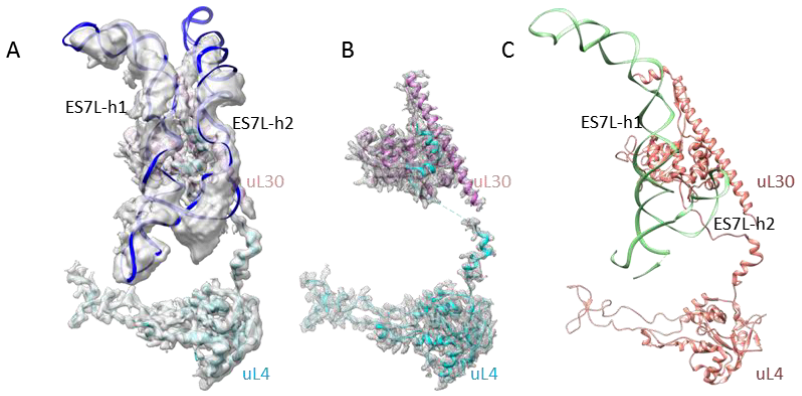

Fig. S16. ES7L and its contacting proteins uL4 and uL30. (A) ES7L (blue) taken from the *T. brucei* model (4V8M.pdb), and uL4 and uL30 from our model were fitted to the *T. cruzi* unsharpened map. (B) uL4 and uL30 superimposed onto the sharpened map, in the same view as in (A). The extension of uL4 is gripped by the long helixes of ES7L, as shown by combination of (A) and (B). (C) ES7L (light green), composed of one long bent helix and one short helix, and uL4 and uL30 (brown) from yeast shown in the same view as (A) and (B).

Fig. S17

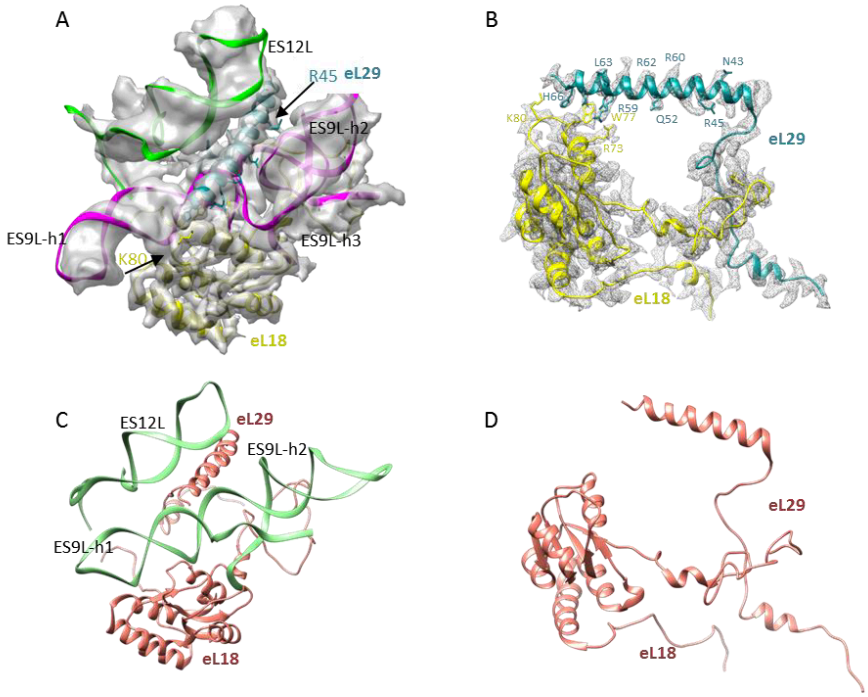

Fig. S17. C-terminal extension of eL29 contacts eL18 and is surrounded by ES9L and ES12L. (A) ES9L (magenta) and ES12L (green) taken from the *T. brucei* ribosome model (4V8M.pdb), and eL18 and eL29 from our model were fitted into the *T. cruzi* unsharpened map. (B) Interactions of eL29 extension with eL18. The models are fitted to the sharpened map. (C) ES9L, ES12L, eL18, and eL29 in yeast, in the same view as in (A). (D) eL29 and eL18 of yeast in the same view as in (B).

Table. S1

	5.8S	LSU- α	srRNA1	LSU- β	srRNA2	srRNA3	srRNA4	Ref.
T. cruzi	172	1969	209	1661	178	72	136	[11]
T. brucei	115	1840	220	1570	180	70	140	[8]
L. major	283	1781	213	1525	183	73	129	[5]

Table S1. The length of different pieces of rRNA (except 5S rRNA) for three trypanosomatids

Table. S2

Parameters	Value
3D reconstruction	
Particle mumbles	235 348
accuracy of rotation (°)	0.43
accuracy of translation (Å)	0.42
Resolution (unmasked, Å)	2.9
Resolution (masked, Å)	2.5
Map sharpening B-factor (Å^2)	-51.16
Model Refinement	
Cell dimensions a=b=c (Å)	240.35
Cell dimensions α =β=γ (°)	90
Map CC	0.58
Resoluion(Å)	2.5
r.m.s. deviations Bond length (Å)	0.016
r.m.s. deviations Bond angle (°)	1.5
Clash score	18.8
Ramachandran plot	
Favored (%)	93.4
Allowed (%)	5.4
Outliers (%)	1.2
MolProbity Score	2.2
Favored rotatomers (%)	98.0
Corrected sugar puckers (%)	97.6
Good backbone conformation (%)	75.2
EM Ringer Score	4.32

Table S2. 3D reconstruction and model statistics of *T. cruzi* large ribosomal subunit

Table S3

Components	Number
Proteins	38
rRNA	8 pieces
H20	84
Mg	105
Zn	3
Methylation	66 2'-O methylation
	10 nucleobase methylation

Table S3. Summary of the *T. cruzi* ribosomal large subunit model

Table S4

RNA	Length	Built
5S	118	1-87,91-118
5.8S	172	1-79,89-125,139-169
LSU-α	1969	3-154,175-198,205-219,232-281,304-519,585-604,744-853,866-880,916-1171,1221-1384,1491-1545,1569-1691,1709-1805,1918-1964
srRNA1	209	2-12,21-74,95-152,175-185,191-209
LSU-β	1661	6-136,446-594,616-789,852-913,1105-1113,1134-1232,1240-1330,1343-1397,1443-1580, KSD (partially).
srRNA2	178	21-44,57-128,154-178
srRNA3	72	1-17,44-72
srRNA4	136	2-47,62-68,87-112,120-125,128-133

Table S4. Summary of the modeled ribosomal rRNA

Table S5

LSU-α	Location	LSU-β	Location
154-175	H16	**137-445**	ES27L
198-232	ES42L	**595-615**	H69
281-304	H17	**790-851**	H76
519-584	ES7L	**914-1104,1114-1133**	ES31L
853-866,880-916	ES9L	**1233-1249**	H84
1171-1221	H38	**1331-1342**	H87
1385-1491	H42-43	**1398-1442**	H89
1545-1569	ES15L		
1691-1709	*T.c.* specific fragment		
1805-1917	ES19L		

Table S5. The unmolded ribosomal rRNA of LSU-α /β location

Table S6

ES	Coordinates	Function	Uniqueness compared with yeast ribosome	Description
ES3L	5.8S (121-139)	Structural stabilization	Increased size	Contacts the 3′ end of LSU-α
ES7L	LSU-α (484-750)	Regulation	two long bent helices	Surrounds uL4 linking ES7L to PET
ES9L	LSU-α(848-953)	Regulation	One additional helix	Surrounds eL29 linking ES9L to PTC
ES27L	LSU-β(128-451)	Structure stabilization and regulation	Longer	Contacts S-domain of SRP and forms a bridge with SSU
ES31L	LSU-β(866-1146)	Structure stabilization and regulation	One additional helix	Contacts the 3′end of 5.8S and uL23, which contacts SRP
ES42L	LSU-α(198-239)	Structure stabilization and regulation	Trypanosome specific Occupies the C-termini of yeast uL22	Contacts srRNA2-4 cleavage sites and contacts uL22(part of PET)
KSD	LSU-β(1585-1661)	Regulation	Trypanosome specific	Binds to uL3 which "wiggle" to PTC (19) and contacts srRNA2 containing sasin-ricin loop

Table S6. The expansion segments with unusual size and Trypanosome-specific ES

Table S7

Protein names	Old nomenclature	Protein ID	Range modeled	a.a. size	Modeled percentage	ChainID
uL2	L2	XP_816366.1	2-246	260	94%	e
uL3	L3	XP_814673.1	2-404	428	94%	f
uL4	L4	XP_804516.1	6-302,315-342	374	87%	r
uL5	L11	XP_803249.1	17-27,73-78,105-109,130-172	192	34%	L
eL6	L6	XP_806123.1	23-104,144-193	193	68%	v
eL8	L7a	XP_821284.1	88-169,185-310	319	65%	x
uL13	L13a	XP_810252.1	20-222	222	91%	O
eL13	L13	XP_810966.1	2-134,168-206	218	79%	N
uL14	L23	XP_821640.1	13-139	139	91%	W
eL14	L14	XP_804656.1	5-160	180	87%	P
uL15	L27a	XP_821642.1	2-145	145	99%	b
eL15	L15	XP_814692.1	2-204	204	100%	Q
uL18	L5	XP_805667.1	6-113, 146-259	309	62%	u
eL18	L18	XP_819826.1	2-193	193	99%	I
eL19	L19	XP_820996.1	3-189	372	50%	T
eL20	L18a	XP_819631.1	2-178	179	99%	S
eL21	L21e	XP_812750.1	3-113,134-148	159	79%	U
uL22	L17	XP_804874.1	3-154	166	92%	R
eL22	L22	XP_819834.1	24-33,46-72,75-122	130	65%	V
uL23	L23a	XP_820993.1	78-94,102-193	194	56%	X
uL24	L26	XP_806321.1	7-119	143	79%	Z
eL24	L24	XP_804713.1	3-63	125	49%	Y
eL27	L27	XP_809804.1	2-50,58-86,107-133	133	79%	a
eL28	L28	XP_820211.1	5-82,88-128	146	82%	c
uL29	L35	XP_811059.1	7-38,46-126	127	89%	k
eL29	L29	XP_803994.1	6-68.	71	89%	d
uL30	L7	XP_816099.1	28-242	242	89%	w
eL30	L30	XP_810701.1	14-72,91-97	105	63%	g
eL31	L31	XP_808028.1	13-156,164-186	188	89%	h
eL32	L32	XP_817439.1	11-123	132	86%	i
eL33	L35a	XP_821812.1	8-149	149	95%	l
eL34	L34	XP_811573.1	4-107	171	61%	j
eL36	L36	XP_805417.1	14-108	114	83%	m
eL37	L37	XP_807547.1	2-82	84	96%	n
eL38	L38	XP_816521.1	2-14,21-59	82	63%	p
eL39	L39	XP_808953.1	2-51	51	98%	q
eL42	L44	XP_806035.1	2-94	106	88%	t
eL43	L37a	XP_808231.1	3-87	90	94%	o

Table S7. Modeled proteins of the *T. cruzi* 60S ribosomal subunit

Table S8

srRNAs	Anchoring proteins	Contacting sites
srRNA1	eL19	92,103,110,124,128
	eL34	W47, H51
srRNA2	uL3	26,27,31,50,92,99,104,161,130,131,133,185,229,236,339,
srRNA3	eL6	47,73,81,76-78,170
	eL33	85,87,139,39,141,143,109,79,79,145,31-33
srRNA4	eL31	91,115,146

Table S8. Residues of anchoring proteins involved in assembly of small pieces rRNA

Table S9

Expansion Segment	Comparison between *T. cruzi* 60S and human ribosome
ES3L	A couple of nucleotides longer in *T. cruzi*
ES4L	Conserved
ES5L	A couple of nucleotides shorter in *T. cruzi*
ES7L	Truncated in *T. cruzi*
ES8L	A couple of nucleotides shorter in *T. cruzi*
ES9L	Conserved, both composed of three helices but have different configurations
ES10L	Truncated in *T. cruzi*
ES12L	Conserved
ES15L	Truncated in *T. cruzi*
ES19L	Expanded in *T. cruzi*
ES20L	Missing in trypanosomes/location at cleavage site LSU-α and srRNA1
ES26L	Missing in trypanosomes/location at cleavage site srRNA1 and LSU-β
ES27L	Truncated in *T. cruzi*
ES30L	Truncated in *T. cruzi*
ES31L	Expanded in *T. cruzi*
ES35L	Conserved
ES39L	Truncated in *T. cruzi*
ES41L	Expanded in *T. cruzi*
ES42L	Trypanosome specific

Table S9. Expansion segments compared with human ribosome counterparts

Table S10

Protein names	a.a. size of *T. cruzi*	a.a. of *S. cerevisiae*	a.a. size of *H. sapiens*	Description
uL2	260	254	257	Very conserved
uL3	428	387	403	Insertion of a.a 209-211 and extension on C-terminal contacts sRNA4
uL4	374	362	427	C-terminal Extension
uL5	192	174	178	N- and C- terminal extensions
eL6	193	176	192	Inserted residues 128-143 are located on the less ordered region.
eL8	319	256	288	N-terminal extension
uL13	222	199	203	Contacts srRNA3
eL13	218	199	211	Conserved
uL14	139	137	140	Conserved
eL14	180	138	215	Shortened at N-terminal and C-terminal extension compared to S.c
uL15	145	149	148	Conserved
eL15	204	204	204	Conserved
uL18	309	297	297	Insertion a.a 136-141, a.a 222-224, deletion of residues 163-165 in yeast, and C-terminal extension
eL18	193	186	188	Insertion of a.a 72-77
eL19	372	189	196	Conserved, C-terminal extension contacts SSU
eL20	179	174	176	Conserved
eL21	159	160	160	Conserved
uL22	166	184	184	C-terminal is shortened, resulting in absence of contacts with ES39L homolog, srRNA3
eL22	130	121	128	Conserved
uL23	194	142	156	Extension at N-terminal
uL24	143	127	145	Extension at N-terminal
eL24	125	155	145	Shortened at C-terminal
eL27	133	136	136	Conserved
eL28	146	/	137	Insertion of a.a 118-132
uL29	127	120	123	Insertion of a.a 81--84
eL29	71	59	159	Extension at C-terminal compared to yeast
uL30	242	244	248	Conserved
eL30	105	105	115	Conserved
eL31	188	113	125	Extension at C-terminal. Contacts KSD
eL32	132	130	135	Conserved
eL33	149	107	110	Extension at N-terminal and insertion of a.a 89-98
eL34	171	121	117	Extension at C-terminal and insertion of a.a 45-50 that contact srRNA1
eL36	114	100	105	Conserved
eL37	84	88	97	Conserved
eL38	82	78	70	Contacts srRNA1
eL39	51	51	51	Conserved
eL42	106	106	106	Conserved
eL43	90	92	92	Conserved

Table S10. Proteins compared with human and yeast ribosome counterparts

Chapter 6: Determination of the ribosome structure to a resolution of 2.5 Å by single-particle cryo-EM

This chapter reproduces a review article published in Protein Science by Zheng Liu, Cristina Gutierrez-Vargas, Jia Wei, Robert A. Grassucci, Ming Sun, Noel Espina, Susan Madison-Antenucci, Liang Tong, and Joachim Frank. My contributions were coauthoring the manuscript to provide an overview on the methods utilized to achieve high-resolution of the *Trypanosoma cruzi* ribosome.

REVIEW

Determination of the ribosome structure to a resolution of 2.5 Å by single-particle cryo-EM

Zheng Liu,[1] Cristina Gutierrez-Vargas,[2] Jia Wei,[2] Robert A. Grassucci,[1] Ming Sun,[2] Noel Espina,[3] Susan Madison-Antenucci,[3] Liang Tong,[2] and Joachim Frank[1,2,4]*

[1]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032

[2]Department of Biochemistry and Molecular Biophysics, Howard Hughes Medical Institute, Columbia University, New York, New York 10032

[3]Division of Infectious Diseases, New York State Department of Health, Wadsworth Center, Albany, New York 12201

[4]Department of Biological Sciences, Columbia University, New York, New York 10027

Received 26 September 2016; Accepted 14 October 2016
DOI: 10.1002/pro.3068
Published online 17 October 2016 proteinscience.org

Abstract: With the advance of new instruments and algorithms, and the accumulation of experience over decades, single-particle cryo-EM has become a pivotal part of structural biology. Recently, we determined the structure of a eukaryotic ribosome at 2.5 Å for the large subunit. The ribosome was derived from *Trypanosoma cruzi*, the protozoan pathogen of Chagas disease. The high-resolution density map allowed us to discern a large number of unprecedented details including rRNA modifications, water molecules, and ions such as Mg^{2+} and Zn^{2+}. In this paper, we focus on the procedures for data collection, image processing, and modeling, with particular emphasis on factors that contributed to the attainment of high resolution. The methods described here are readily applicable to other macromolecules for high-resolution reconstruction by single-particle cryo-EM.

Keywords: high resolution cryo-EM; single particle analysis; ribosome structure; 2.5 Å resolution; *Trypanosoma cruzi*

Introduction

In the past three years, single-particle cryo-EM, a technique developed in the last three decades, has revolutionized structural biology and has gained popularity in academic and, increasingly, industrial research. The revolution is manifested in three aspects: the high resolution, comparable with that obtained in X-ray crystallography; the capacity to determine multiple structures co-existing in the same sample; and the proven capability of solving small membrane proteins and large supramacromolecular complexes, which are both difficult to solve by X-ray crystallography.

With traditional recording devices, the best resolution for asymmetric molecules such as ribosomes was not better than 5 Å.[1,2] It has now moved into the range of 2–3 Å, permitting *de novo* atomic modeling.[3–6]

This work was supported by HHMI and NIH grants R01 GM29169 and GM55440 (to J.F.), R35GM118093 (to L.T.), and NSF GRFP DGE-11-44155 (to C.G).

*Correspondence to: Joachim Frank, Biochemistry and Molecular Biophysics, Columbia University, 650 West 168th Street, New York, New York 10032. E-mail: jf2192@cumc.columbia.edu

Improvements in the data quality of cryo-EM have been provided primarily by advances in instrumentation. First of all, the direct electron detector device (DDD) dramatically improves the Detection Quantum Efficiency (DQE) and allows fractionating the electron dose of each micrograph.[7] In this way, radiation damage can be reduced a posteriori by selecting a subset of frames, and the beam-induced movement of the sample can be compensated by correcting the displacements among the frames. In addition to the introduction of direct electron detectors, spherical aberration (Cs) correction,[5] energy filtration[8] and the addition of a phase plate are all measures that have the potential to improve image quality and resolution.[9]

The ribosome has served as a benchmark sample for decades in the development of single-particle cryo-EM.[10] Sub-3 Å resolution of ribosomes, in the range of 2.5–2.9 Å, has now been achieved by four groups.[3–6] Our own 2.5-Å cryo-EM reconstruction[6] is of the ribosome extracted from *Trypanosome cruzi*, a protozoan pathogen causing Chagas disease in humans. This ribosome, as those from other trypanosomatids, possesses a 28S rRNA composed of six fragments. The high-resolution structure not only reveals the precise interactions stabilizing the RNA but also provides insights that may allow the design of safe trypanosome-specific drugs. Our work makes judicious use of recent advances in cryo-EM methodology including instrumentation, image acquisition, and image processing algorithms, modeling tools and computing resources. Achievement of the highest resolution, beyond 3 Å, requires bottlenecks to be eliminated that are not present at lower resolution. The following sections will describe each of these steps and discuss the rationales for the choices of conditions and parameters in optimizing the results.

Data Collection

In our experiments for the *T. cruzi* project, we made use of the FEI Falcon 2 camera installed on the Titan Krios microscope (FEI, Eindhoven) with a Cs corrector and EPU software for imaging our sample, in preference to an FEI Tecani F30 Polara microscope equipped with a K2 Summit (Gatan, Pleasanton) available in house. The whole dataset was collected in a single pass of five consecutive days, yielding about 11,000 micrographs with a speed of about 100 micrographs per hour. Thus, high throughput of imaging allowed collection of a big dataset with sufficient numbers of particles for high-resolution reconstruction, even covering different states of the ribosome.

The electron microscope

For high-resolution single-particle cryo-EM, transmission electron microscopes operating at 300 kV acceleration voltage show better performance over 200 kV voltage as the use of higher acceleration voltage guarantees larger depth of field, extends the contrast transfer function (CTF) into a higher resolution range, and improves sample penetration depth.[11] Even before the advent of DDD cameras, most, if not all, of the published cryo-EM works achieving resolution beyond 4 Å were done on 300-kV microscopes on icosahedral viruses.[12–14] Three widely used 300-kV microscopes are JEM-3200 FSC (JEOL, Tokyo, Japan), FEI Tecani F30 "Polara" and FEI Titan Krios (both FEI, Eindhoven, The Netherlands).

With the collimating C3 condenser lens, which ensures illumination with a highly coherent beam,[15] and a convenient sample auto-loading system, the Titan Krios, the instrument of our choice, currently generates most of the high-resolution structures with resolutions beyond 3.5 Å. For our experiments, we had the choice between two differently equipped Titan Krios microscopes. We chose the one with a spherical aberration (Cs) corrector since Cs correction minimizes geometrical distortions due to coma. Use of the corrector removes the effect of beam tilt, which is thought to be one of the main limitations preventing achievement of high resolution beyond 3 Å.[5]

Choice of camera

After the advent of DDD cameras, films and CCD cameras are now rarely used for high-resolution imaging.[16] Currently, three commercially available DDD cameras have demonstrated good performance for high resolution, each with its own advantages and disadvantages. For example, the camera made by DE (Direct Electron, San Diego) provides a larger field of view than both the FEI Falcon (FEI, Eindhoven) and K2 Summit (Gatan, Pleasanton) cameras. The K2 Summit camera has shown particularly good performance for small molecules since it is able to record single incoming electrons in a sub-pixel, "super-resolution" counting mode. We chose to use the Falcon 2 camera in spite of the K2 Summit cameras being available on the same microscope, because it takes a shorter exposure time than the typical counting camera K2 and still has good performance, as proven by its precursor Falcon 1, which achieved high resolution to 3.6 Å[17] and even to 2.9 Å[5] for 70S ribosomes. With shorter exposure time, within a fixed session, this camera generates more data (i.e., more particles), which is a significant factor in obtaining high resolution by single-particle cryo-EM.

Software for data acquisition

A set of software packages have been developed to control CCD cameras and microscopes for fully automated data acquisition. These packages include *AutoEM*,[18] *JADAS*,[19] *Leginon*,[20] and *SerialEM*.[21] Some of these continue to serve data acquisition on new generation cameras such as *SerialEM* and *Leginon*. New algorithms developed either by academic labs, such as *UCSFImage*,[22] or by commercial companies, such as EPU (FEI), have been recently added to this list for

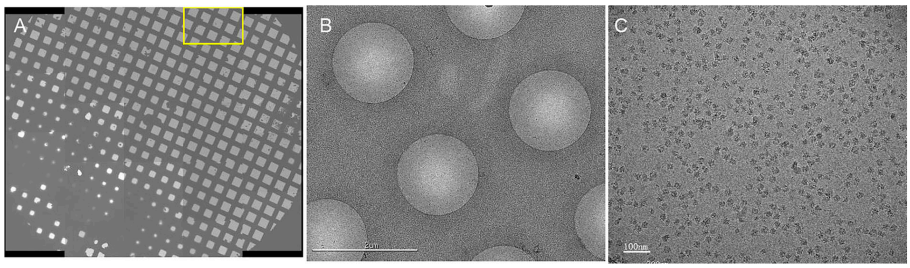

Figure 1. Cryo-EM of *T. cruzi* ribosome. (A) Overview of the grid indicates that the grid is good, containing thin ice. (B) Hole views, which is from the region marked by the yellow box in A, reveal thin vitreous ice; the black spots within the hole are the ribosome particles. (C) Ribosome particle distribution on a micrograph.

single-particle data collection. Considering the compatibility of the software with the electron microscope and cameras, we chose EPU, a dedicated program for single-particle data collection closely modeled after the *Leginon* system[16] by FEI, as the data acquisition software.

Imaging conditions

Magnification directly defines the pixel size of the object on the recording device. According to the Nyquist–Shannon theorem, the highest resolution (in terms of smallest resolved distance) achievable by a recording medium is twice the sampling distance (Nyquist limit). For direct electron detectors possessing high DQE, the resolution of the signal extracted can reach 2.5 times pixel size or even closer to the Nyquist limit. For resolutions in the range of 3–4 Å, pixel sizes in the range of 1.3–1.7 Å/pixel have been previously used.[15,23–25] In our data collection, with the aim to obtain resolutions beyond 3 Å, we used a high-magnification setting (after calibration 133,970 X on the Falcon 2 camera), yielding a pixel size of 1.045 Å/pixel, which corresponds to a Nyquist limit of 2.09 Å. The choice of such high magnification inevitably reduces field size and hence the number of particles captured per micrograph; to compensate for this effect, we adopted the strategy of collecting multiple exposures per hole, setting four exposure targets in each hole. In this way, the throughput was increased compared with the conventional strategy of one exposure per hole. Also for highest throughput, the total exposure time was set to 1 second, similar to that used in films and shorter than used with K2 Summit camera in counting mode. Sixteen frames per movie were collected, with a total dose of 32 e/Å^2. The overview of the grid, typical hole views and one micrograph of this dataset is showed in Figure 1.

Image Processing

Overview

The complete image processing workflow for single-particle cryo-EM usually contains particle selection, determination of the parameters of the CTF, determination of angles and positions (3D projection matching), 3D classification, and 3D reconstruction. Since the advent of the first-single particle analysis software package—*SPIDER*[26] in the early 80s – a number of software packages have been developed, including *IMAGIC*,[27] *EMAN*,[28] *XMIPP*,[29] *BSOFT*,[30] *SPARX*,[31] *Frealign*,[32] and *RELION*[33] (For an exhaustive review, see https://en.wikibooks.org/wiki/Software_Tools_For_Molecular_Microscopy). Furthermore, two entire software platforms have been developed from which individual packages can be accessed interactively: *Appion*[34] and *Scipion*.[35] Even though most popular packages provide standard and common pipelines to complete processing, each project may in fact require different strategies, depending on the properties of the sample. In the image processing of our *T. cruzi* ribosome dataset, multiple available packages such as *SPIDER*, *EMAN*, *RELION*, and so forth, have been utilized, as will be detailed below. In giving an overview, we distinguish two sections, pre-processing conducted on a single desktop, and processing on a computer cluster (Fig. 2).

Pre-processing

Movie processing. A raw movie output from the electron microscope comprises multiple exposures (frames) of a single region of the sample grid. It can be recorded either in integrating mode by current versions of DE and Falcon or in counting mode by the K2 Summit camera. Whatever nature of the movies, they are amenable for correction of drift induced either by interacting electrons or by motions of the specimen stage. At present, the available drift correction algorithms include *motioncorr* (*dosefgpu_driftcorr*)[7] and its successor *motioncorr2*, *alignframes_lmbfgs* and *alignparts_lmbfgs*,[36] the *optical Flow* method in *Xmipp*,[37] *Unblur*,[38] and the particle polishing procedure in *RELION*.[39] The micrographs or particles resulting from movie processing are subject to the traditional image processing pipeline.

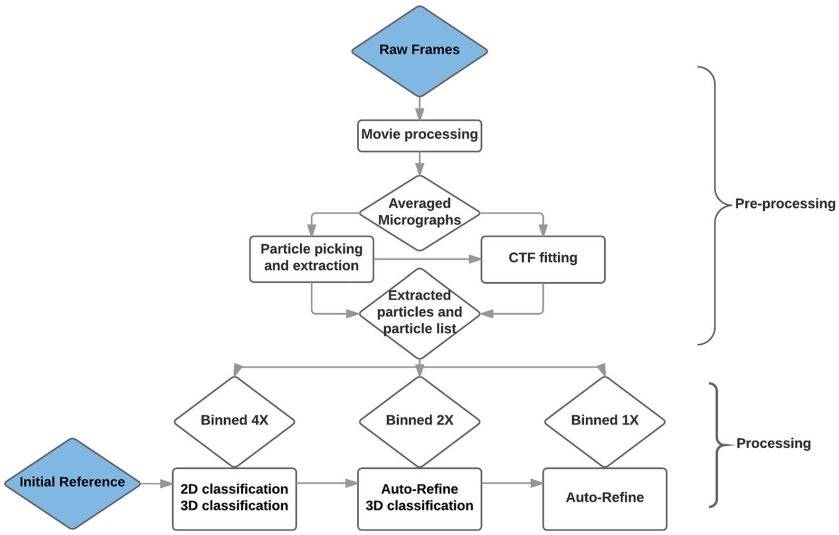

Figure 2. Overview of the image processing workflow for the *T. cruzi* data set.

The principle of these movie processing algorithms is that they first calculate the movements between the frames, compensate the offsets and sum over the shifted frames to yield the drift-corrected dataset. Drift calculation and image summing may be done at different levels of the image: whole frame, sub-frame, and particle. *Motioncorr*,[7] the first popular movie processing software, relies on pair-wise cross-correlations to calculate translations at the whole-frame level. The programs *alignframes_lmbfgs* and *alignparts_lmbfgs*[36] align either whole movie frames or individual particles, respectively. The *Optical Flow* approach[37] estimates translations of frames or individual particles based on the representation of local motion by a vector field. *Unblur*[38] uses cross-correlation iteratively to align each individual frame to an average of frames that does not include the frame being aligned. Particle polishing[39] is an approach that corrects for particle movement using the 2D projections of a 3D reference map to align individual particles and thereby refine their positions.

In the image processing for the *T. cruzi* ribosome, movie processing was done on movies that were screened visually in advance. In our strategy, we first used *motioncorr1* initially for averaging all frames of each micrograph and then ran the complete image processing workflow as outlined in Figure 2. At a later stage, described further below, the movie processing was revisited with different frame averaging combinations as it is known that earlier frames have more drift and later frames experience more radiation damage.[7]

CTF-fitting. Bright-field images obtained in the transmission electron microscope (TEM) are affected by the phase contrast transfer function (CTF).[40] More specifically, the Fourier transform of the image is modulated in amplitude and flipped in phase in a spatial frequency-dependent manner. This distorting effect can be corrected by restoring the actual Fourier amplitudes and phases for each micrograph. Among the parameters (defocus, B-factor, astigmatism, etc.), the defocus value is the only one that can be controlled in the experiment, but its actual value always differs from the nominal one for several reasons, including precise sample height. As the most critical parameter to be determined, the defocus is estimated by fitting of a simulated power spectrum to the power spectrum of the electron micrograph. The fitting is facilitated by the characteristic signature of the CTF visible in the Fourier transform of the micrograph, known as Thon rings. The power spectrum is usually calculated by averaging absolute-squared Fourier transforms of boxed particles or of selected sub-regions of a micrograph.

To precisely determine defocus values, we used two methods, *CTFFIND3*,[41] which uses the average of power spectra from boxed regions of the micrograph, and *CTFIT2*,[42] which instead uses the average of power spectra of selected particles. The values obtained from both methods were cross-validated. Only micrographs with matching defocus values (i.e., with a difference smaller than 50 nm) were kept for subsequent processing, while the remaining micrographs were subjected to visual verification and manual CTF fitting.

Only those with well-fitting CTF curves were returned to the processing workflow while the rest were discarded.

Data screening. Automatic imaging brings with it convenience in data collection but also an increase in the amount of contaminated or suboptimal data in the output dataset. Inevitably such a dataset requires more careful screening compared with data collected manually. Altogether we used four different strategies to screen the data at different steps in the workflow of image processing.

First, micrographs were examined visually to remove those with contamination by pieces of ice and ethane and those exposed to the shifted beam, before movie processing and particle picking. Second, micrographs with too few particles or showing areas of thick ice were excluded in a semi-auto particle selection step using *e2boxer.py*.[28] Third, power spectra of micrographs generated for CTF determination after movie processing were examined to exclude those with uncorrected drift and astigmatism. Fourth, as already noted above, inconsistencies among defocus values calculated by different CTF algorithms were used to reject some micrographs. In this manner, we ensured that only high-quality micrographs proceeded to the image processing.

Processing

The processing step usually comprises 2D image alignment, 3D classification and reconstruction from hundreds of thousands to millions particle images, which requires a substantial amount of computational resources in terms of CPU time and memory. In our case, after pre-screening, we set out with a total number of 700 k particles. This section in our work is divided into three successive steps performed on the 4× binned, 2× binned and 1× binned (i.e., un-binned) dataset, respectively (Fig. 3).

Initial reference. To determine view angles and positions of the selected particles, an initial 3D low-resolution reference is routinely used. From this reference a set of projections is generated for comparison with the raw (i.e., unprocessed) particles, and the angles and positions from the best matching projections are assigned to the particle. These initial parameters are further iteratively refined by comparing each particle with projections of the map reconstructed from the previous iteration, and so forth. Since strong low-pass filtration is used to mitigate the effect of reference bias, the choice of initial reference is not critical; it may be derived from the cryo-EM map of a loosely related structure, or from coordinates of a related structure deposited in the PDB database (http://www.rcsb.org/). In processing our *T. cruzi* dataset, we used a previously reconstructed cryo-EM 70S ribosome structure[17] filtered to 60 Å as the initial reference. The fact that our reconstruction displays no similarity to the 70S ribosome in its high-resolution features indicates absence of reference bias.

Classifications on binned datasets. Because of the large degree of heterogeneity in single-particle cryo-EM data, classification is a hierarchical, multi-stage procedure that requires careful consideration. It is a step that is difficult to automate, in part because strategic decisions that are dependent on the outcome at every level must be made in the process. Since the sorting progresses from coarse to fine details, it can be done on binned versions of the data first, to save time. In our case, we chose two stages of binning, 4× and 2×, before proceeding to the processing of un-binned data. The entire workflow of classification at these three stages is schematically displayed in Figure 3.

On the 4× binned dataset, an initial step of 2D classification was used to eliminate impurities and "bad" particles in large part. This was followed by 3D *RELION*-based classification run in a hierarchical way. The first round aimed at further cleaning the dataset resulting from screening by 2D classification, and obtain an inventory of existing conformations and compositions in the dataset, with K = 10 chosen as number of classes. As the sample had been obtained by purification from a cell extract without further intervention, is was uncertain at this stage which, if any, work cycle state of the ribosome would dominate, and could therefore be singled out in the subsequent image processing for high-resolution structure determination.

For the second round of classification, all ribosome-like particles (400 k) were pooled, and the number of classes was chosen to be higher (K = 10) than the number of ribosome-like classes (K = 7) revealed in the first-round classification. In this round, various states of the ribosome could be recognized, including whole ribosomes with/without bound components (tRNAs, eEF2), in rotated and un-rotated configurations, and the two disassociated subunits. Two sizable classes of ribosomes without P-site tRNA and GTPase factors were found, showing occupation by E-site tRNA as the only difference.

The third-round classification was applied to various sub-groups, which were pooled based on similarity in the second-round classification. The idea was that in this way, possible new classes of conformations not sorted out in the previous round of classification could be established, or, alternatively, the previous classification could be confirmed.

On the 2× binned data, refinement and reconstruction (Auto-refine in *RELION*) was first performed on the two 80S classes with most particles obtained in the third- and last-round classification of the 4× binned data. In one (87 k particles) the ribosome

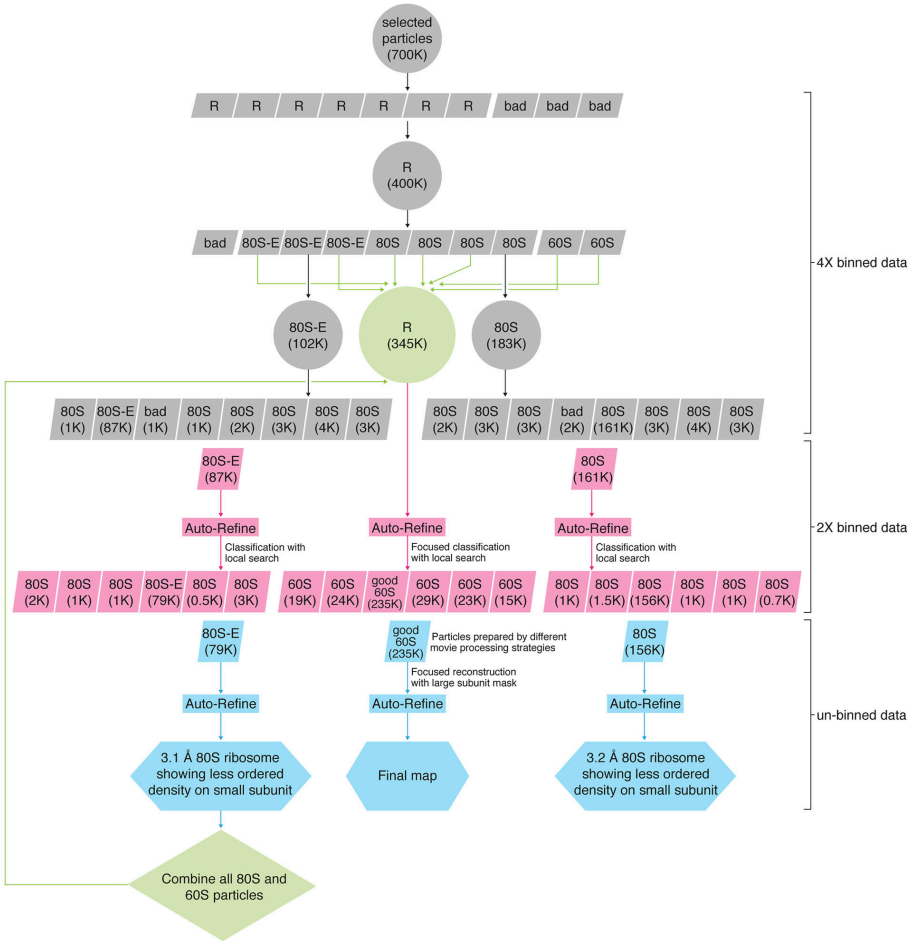

Figure 3. The entire workflow of the classification and reconstruction for *T. cruzi* data set. R, ribosome like class including both 80S and 60S; bad, bad class; 80S, class of empty 80S ribosomes; 80S-E, class of 80S ribosomes with E-site tRNA; 60S, class of 60S subunits.

contained the E-site tRNA; in the other (161 k particles) it was empty. The resulting reconstructed maps were then taken as references for 3D classification, this time using a finer local angular and translational search, of the same particles subset which generated this reconstruction. Here the rationale was that in this way, any deviant particles might be relegated to a less populated class and a more homogeneous subset would be obtained. This strategy worked: the two classes were whittled down to 79 k and 156 k, respectively.

Refinement and reconstruction of unbinned data. At this stage of the analysis, following a strategy that is now routine, auto-refine in *RELION* may be used to generate a final 3D reconstruction from a homogeneous subpopulation of particles. However, there are many cases where a tailored, individual strategy is required to handle difficulties during imaging processing. Such difficulties arise when the structure in parts of the molecule is stable and reproducible, but flexible in other parts, and can be solved by application of a mask on the 3D map to single out regions of interest for further classification or refinement ("focused classification" or "focused reconstruction," respectively[43]). In our case, the residual heterogeneity of the small subunit in the 80S ribosome classes compromised the resolution that could be obtained for the large subunit. Application of Auto-refine on the two fixed subpopulations produced

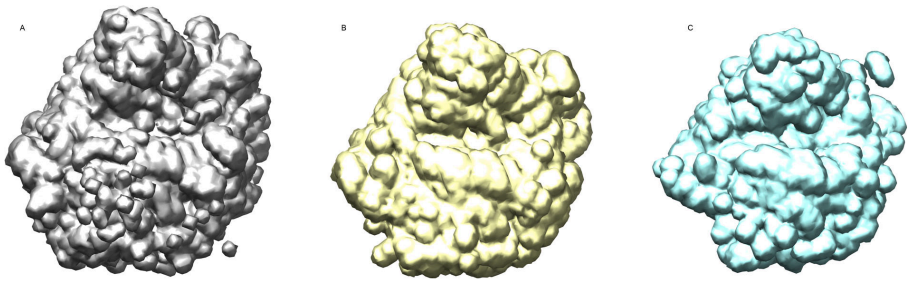

Figure 4. Masks for resolution estimation of *T. cruzi* cryo-EM reconstruction. (A) A mask including all the components of the 60S subunit and some residual density from the small subunit. (B) A mask used in focus refinement, encompassing only the 60S subunit. (C) The mask used for resolution estimation, eliminating flexible components such as the P stalk.

maps of the 80S ribosome with the overall resolutions of 3.1 Å and 3.2 Å, respectively.

At the next stage we made the decision to focus on the large subunit only in the subsequent refinement. We went back to the entire pool of 345 k particles identified as 80S and 60S in the second round of the classification of 4× binned data, and used the 3.1 Å 80S reconstruction with mask passing the 60S subunit as reference in a focused classification of this pooled dataset, yielding 2.9 Å resolution. Next, to obtain a refined reconstruction at highest resolution, the reconstruction for the pooled subset was made to revisit the particles extracted from micrographs generated for different choices frame averaging. We tried the following combinations: 2–10, 3–10, 4–10, 2–11, 3–11, and 4–11, and obtained the best result (2.5 Å) with frames 3–10.

Resolution Estimation and Modeling

Resolution estimation

In X-ray or electron crystallography, the extent of the diffraction pattern (i.e., the radius in reciprocal space up to which diffraction peaks are detected) is an indicator of resolution. No equivalent criterion exists that would indicate the resolution of signal contained in the raw data for the single-particle cryo-EM method. Instead, resolution is estimated based on reproducibility of the reconstruction from independent datasets. The Fourier shell correlation (FSC) between the reconstructions of two half-sets that have been refined independently eliminates in large part the bias from over-fitted noise. When independence is thus ensured, the spatial frequency at the cutoff of FSC = 0.143 is regarded as a measure of resolution.[44]

In the use of FSC to estimate resolution from two maps, a mask is applied to the two maps before calculation of the FSC. Purpose of the mask is to screen off irreproducible peripheral density and noise[45] that, if admitted, would lead to underestimation of resolution. In our case, flexible expansion segments (ES) and some extended proteins in the periphery needed to be masked off for a realistic resolution estimation of the reconstruction. We used a *RELION*-tailored mask choosing a threshold such that well-ordered regions of the ribosome were included, with a smooth 3-pixel wide edge falloff. Specifically, our mask excludes the L1 and P stalk proteins, as well as some long ribosomal RNA expansion segments (Fig. 4).

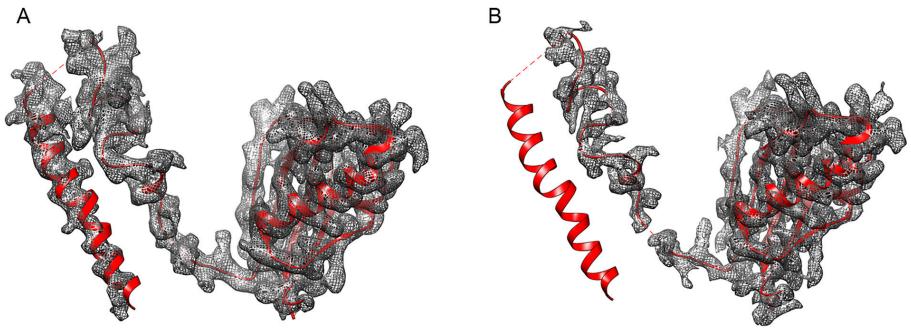

Figure 5. Protein eL31 after sharpening with different negative B factors. (A) B = −30 Å^2. (B) B = −75 Å^2. The missing density on the left helix of eL31 is caused by sharpening with a larger B factor, −75 Å^2. The display of the figures is set at 3 × σ.

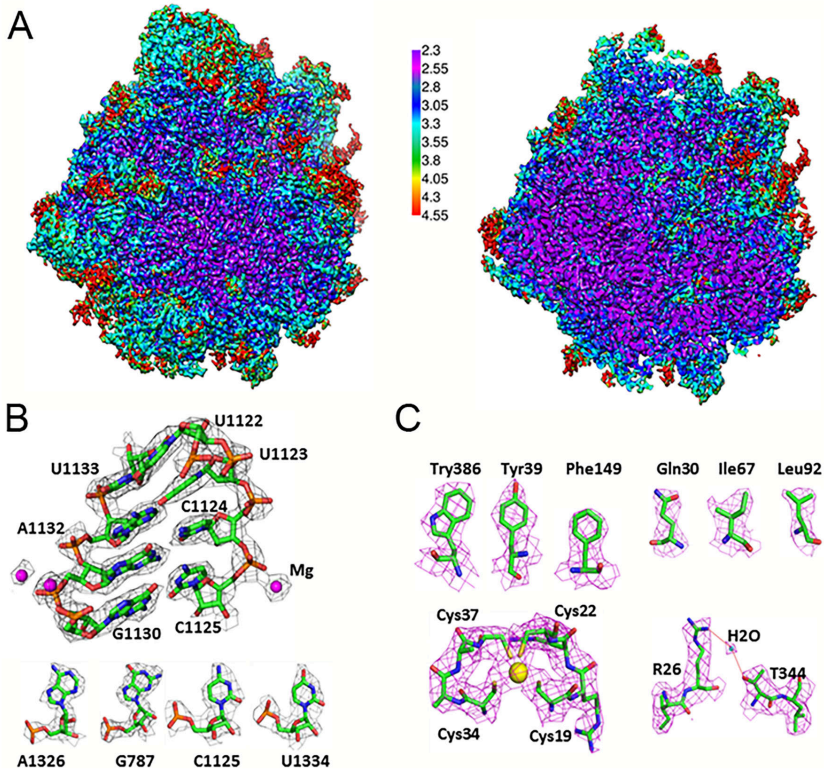

Figure 6. High-resolution structure of the *T. cruzi* 60S ribosomal subunit. (A) Cryo-EM map of the 60S subunit after sharpening, colored by local resolution and viewed from the subunit interface. Left, surface view; right, central cut-away view. (B) Cryo-EM density of a highlighted rRNA region of LSU-α docked with atomic model. Bottom panel, examples for the four nucleotides in rRNA. (C) Selected views of the density maps of proteins. Top, densities for some of the amino acids. Bottom left, density of a zinc ion; Bottom right, density of a water molecule.

In addition to the fixed number provided by the FSC method, local resolution estimation is quite informative since the definition of features varies substantially across the map. In our work, we estimated local resolution of our large subunit map using the software *Resmap*,[46] and found it to range from 2.3 Å in the core region to 4.5 Å to in the periphery (Fig. 6).

Modeling

Typically, the cryo-EM map of a macromolecular complex is a chimera composed of regions with highly variable resolution, and a single overall resolution figure cannot do justice to this situation. Therefore, unlike X-ray crystallography where model-building and refinement strategies are chosen on the basis of the overall resolution,[47] cryo-EM often requires multi-map modeling and refinement. In the case of the *T. cruzi* ribosome structure, such a multi-map modeling strategy was also used (Fig. 5). Three density maps were generated by sharpening the initial map with negative B-factors of -30 Å^2, -50 Å^2, and -75 Å^2. The resulting maps were combined, taking advantage of the fact that they complemented one another in connectivity and definition of high-resolution features. Before being loaded into the Coot modeling software,[43] the maps (originally $400 \times 400 \times 400$) were cropped to a size ($230 \times 230 \times 230$, with the spacing at 1.045 Å/voxel) just large enough to contain the whole large subunit, so as to keep within the bounds of the computer memory.

The high-resolution features of the density map and the sequence information of the *T. cruzi* ribosome allowed us to register the rRNA and protein residues with well-ordered blocks of density (Fig. 6B,C). In most of the density regions (approximately

85%) our map was of sufficient quality to allow *de novo* modeling as routinely done in X-ray crystallography. However, to simplify and accelerate our modeling work, we took advantage of the extent of sequence identity and structural conservation between the *T. cruzi* ribosome and the previously studied yeast and *T. brucei* ribosomes. First the crystal structure of the yeast large ribosomal subunit (PDB: 4V88)[48] and the structure of *T. brucei* large ribosomal subunit (PDB: 4V8M)[1] were fitted into the density map as a rigid bodies using *UCSF Chimera*.[49] The conserved regions matched the density map well after real space refinement. Regions of rRNA conserved between *T. cruzi* and yeast were used as starting points, followed by extensive manual building using *Coot*.[50] In the non-conserved rRNA regions, which account for more than half of the rRNA of the large subunit, the distinguishable side chain features for the bases, application of the base-pairing principle, and the sequence information of the *T. cruzi* strain used allowed us to register each of the well-resolved bases. The rRNA sequences were obtained from TriTrypDB Kinetoplastid Genomics Resource (http://tritrypdb.org/tritrypdb/).[51] For the modeling of the non-conserved regions of proteins, the polypeptide chains in *T. brucei* were mutated to match their sequences in *T. cruzi* taken from the National Center for Biotechnology Information (NCBI) protein databases (http://www.ncbi.nlm.nih.gov). On account of the fact that the fitted PDB of *T. brucei* was derived from a 5.5-Å cryo-EM map (EMD-2239), some assignments of amino acid residues were found to be offset and inaccurate. In our density map, the high-resolution features of the residue side-chain density map allowed us to correct these assignments and also extend the peptide chains for the non-conserved regions of proteins. Interestingly, some residue densities reveal inconsistencies with the residues from the *T. cruzi* sequence in the database, a reflection of natural mutations. For example, in our structure, residue 64 in protein eL37 is a Cys in the sequence of the database, but clearly a Met based on the density.[6] The resultant atomic model of the large ribosomal subunit was subjected to real-space refinement using *PHENIX*[52] against the map generated by sharpening with the negative B factor of -50 Å^2.

Perspective

The foregoing outline of methods and choices of processing paths that we used in the determination of *T. cruzi* ribosome structure will make it clear that even with a given high-quality dataset, achievement of highest resolution by single-particle cryo-EM is far from routine; that it requires a good knowledge of the principles underlying data processing and a good deal of intuition. Heterogeneity is a particularly vexing problem that often requires trial and error approaches that rarely make it into the formal Methods descriptions of scientific articles.

The last three years have witnessed spectacular achievements of biological macromolecules structure determination by single-particle cryo-EM. Among novel biological insights that have been gained are the mechanism of transcription initiation[53] and the activation and gating of the calcium release channel,[23,54] which are refractory to structure characterization by traditional methods. Certain samples, such as ribosomes and viruses, are now rarely pursued using X-ray crystallography as cryo-EM needs fewer samples and works with higher efficiency. Even more promise lies in the near future since sample preparation, instrumentation, and computer software are all under development by the cryo-EM community. In the study of ribosomes, usage of high magnification with finer pixel size, electron counting cameras, and anisotropic scale correction in image processing might further enhance the resolution to beyond 2 Å, which has already been realized on a small molecule.[55]

Acknowledgments

We would like to thank Amédée des Georges (CUNY) for helpful suggestions in image processing, Harry Kao for assistance with computer hardware, Nadia Severina for help in ribosome purification, and Melissa Thomas-Baum (Buckyball Design) for assistance with the preparation of figures. We also thank Z. Yu, C. Hong, R. Huang, and H. He (FEI) for their assistance with the data collection at the Janelia Farm Research Campus of HHMI.

References

1. Hashem Y, des Georges A, Fu J, Buss SN, Jossinet F, Jobe A, Zhang Q, Liao HY, Grassucci RA, Bajaj C, Westhof E, Madison-Antenucci S, Frank J (2013) High-resolution cryo-electron microscopy structure of the *Trypanosoma brucei* ribosome. Nature 494:385–389.
2. Armache JP, Jarasch A, Anger AM, Villa E, Becker T, Bhushan S, Jossinet F, Habeck M, Dindar G, Franckenberg S, Marquez V, Mielke T, Thomm M, Berninghausen O, Beatrix B, Soding J, Westhof E, Wilson DN, Beckmann R (2010) Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-A resolution. Proc Natl Acad Sci USA 107:19748–19753.
3. Shalev-Benami M, Zhang Y, Matzov D, Halfon Y, Zackay A, Rozenberg H, Zimmerman E, Bashan A, Jaffe CL, Yonath A, Skiniotis G (2016) 2.8-Å cryo-EM structure of the large ribosomal subunit from the eukaryotic parasite Leishmania. Cell Rep 16:288–294.
4. Passos DO, Lyumkis D (2015) Single-particle cryoEM analysis at near-atomic resolution from several thousand asymmetric subunits. J Struct Biol 192:235–244.
5. Fischer N, Neumann P, Konevega AL, Bock LV, Ficner R, Rodnina MV, Stark H (2015) Structure of the *E. coli* ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. Nature 520:567–570.

6. Liu Z, Gutierrez-Vargas C, Wei J, Grassucci RA, Ramesh M, Espina N, Sun M, Tutuncuoglu B, Madison-Antenucci S, Woolford JL, Tong L, Frank J (2016) Structure and assembly model for the *Trypanosoma cruzi* 60S ribosomal subunit. Proc Natl Acad Sci USA doi:10.1073/pnas.1614594113.
7. Li X, Mooney P, Zheng S, Booth CR, Braunfeld MB, Gubbens S, Agard DA, Cheng Y (2013) Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat Methods 10:584–590.
8. Yonekura K, Braunfeld MB, Maki-Yonekura S, Agard DA (2006) Electron energy filtering significantly improves amplitude contrast of frozen-hydrated protein at 300 kV. J Struct Biol 156:524–536.
9. Danev R, Baumeister W (2016) Cryo-EM single particle analysis with the Volta phase plate. Elife 5:e13046.
10. Frank J (2009) Single-particle reconstruction of biological macromolecules in electron microscopy-30 years. Q Rev Biophys 42:139–158.
11. Kudryashev M, Castano-Diez D, Stahlberg H (2012) Limiting factors in single particle cryo electron tomography. Comput Struct Biotechnol J 1:e201207002.
12. Grigorieff N, Harrison SC (2011) Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy. Curr Opin Struct Biol 21:265–273.
13. Liu Z, Guo F, Wang F, Li TC, Jiang W (2016) 9 A resolution cryo-EM 3D reconstruction of close-packed virus particles. Structure 24:319–328.
14. Guo F, Liu Z, Fang PA, Zhang Q, Wright ET, Wu W, Zhang C, Vago F, Ren Y, Jakana J, Chiu W, Serwer P, Jiang W (2014) Capsid expansion mechanism of bacteriophage T7 revealed by multistate atomic models derived from cryo-EM reconstructions. Proc Natl Acad Sci USA 111:E4606–E4614.
15. Zhang X, Zhou ZH (2011) Limiting factors in atomic resolution cryo electron microscopy: no simple tricks. J Struct Biol 175:253–263.
16. Liu Z, Zhang J-J (2014) Revolutionary breakthrough of structure determination—recent advances of electron direct detection device application in cryo-EM. Acta Biophysica Sinica 30:1–12.
17. Li W, Liu Z, Koripella RK, Langlois R, Sanyal S, Frank J (2015) Activation of GTP hydrolysis in mRNA-tRNA translocation by elongation factor G. Sci Adv 1. Vol. 1, no. 4, e1500169:1–7.
18. Lei JL, Frank J (2005) Automated acquisition of cryo-electron micrographs for single particle reconstruction on an FEI Tecnai electron microscope. J Struct Biol 150:69–80.
19. Zhang JJ, Nakamura N, Shimizu Y, Liang N, Liu X, Jakana J, Marsh MP, Booth CR, Shinkawa T, Nakata M, Chiu W (2009) JADAS: a customizable automated data acquisition system and its application to ice-embedded single particles. J Struct Biol 165:1–9.
20. Suloway C, Shi J, Cheng A, Pulokas J, Carragher B, Potter CS, Zheng SQ, Agard DA, Jensen GJ (2009) Fully automated, sequential tilt-series acquisition with Leginon. J Struct Biol 167:11–18.
21. Mastronarde DN (2005) Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol 152:36–51.
22. Li XM, Zheng S, Agard DA, Cheng YF (2015) Asynchronous data acquisition and on-the-fly analysis of dose fractionated cryoEM images by UCSFImage. J Struct Biol 192:174–178.
23. Zalk R, Clarke OB, des Georges A, Grassucci RA, Reiken S, Mancia F, Hendrickson WA, Frank J, Marks AR (2015) Structure of a mammalian ryanodine receptor. Nature 517:44–U49.
24. Bai XC, Rajendra E, Yang G, Shi Y, Scheres SH (2015) Sampling the conformational space of the catalytic subunit of human gamma-secretase. Elife 4:e11182.
25. Sun M, Li W, Blomqvist K, Das S, Hashem Y, Dvorin JD, Frank J (2015) Dynamical features of the *Plasmodium falciparum* ribosome during translation. Nucleic Acids Res 43:10515–10524.
26. Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M, Leith A (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. J Struct Biol 116:190–199.
27. van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M (1996) A new generation of the IMAGIC image processing system. J Struct Biol 116:17–24.
28. Tang G, Peng L, Baldwin PR, Mann DS, Jiang W, Rees I, Ludtke SJ (2007) EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol 157: 38–46.
29. Sorzano COS, Marabini R, Velazquez-Muriel J, Bilbao-Castro JR, Scheres SHW, Carazo JM, Pascual-Montano A (2004) XMIPP: a new generation of an open-source image processing package for electron microscopy. J Struct Biol 148:194–204.
30. Heymann JB (2001) Bsoft: image and molecular processing in electron microscopy. J Struct Biol 133:156–169.
31. Hohn M, Tang G, Goodyear G, Baldwin PR, Huang Z, Penczek PA, Yang C, Glaeser RM, Adams PD, Ludtke SJ (2007) SPARX, a new environment for Cryo-EM image processing. J Struct Biol 157:47–55.
32. Grigorieff N (2007) FREALIGN: high-resolution refinement of single particle structures. J Struct Biol 157: 117–125.
33. Scheres SH (2012) A Bayesian view on cryo-EM structure determination. J Mol Biol 415:406–418.
34. Lander GC, Stagg SM, Voss NR, Cheng A, Fellmann D, Pulokas J, Yoshioka C, Irving C, Mulder A, Lau PW, Lyumkis D, Potter CS, Carragher B (2009) Appion: an integrated, database-driven pipeline to facilitate EM image processing. J Struct Biol 166:95–102.
35. de la Rosa-Trevin JM, Quintana A, Del Cano L, Zaldivar A, Foche I, Gutierrez J, Gomez-Blanco J, Burguet-Castell J, Cuenca-Alba J, Abrishami V, Vargas J, Oton J, Sharov G, Vilas JL, Navas J, Conesa P, Kazemi M, Marabini R, Sorzano CO, Carazo JM (2016) Scipion: a software framework toward integration, reproducibility and validation in 3D electron microscopy. J Struct Biol 195:93–99.
36. Rubinstein JL, Brubaker MA (2015) Alignment of cryo-EM movies of individual particles by optimization of image translations. J Struct Biol 192:188–195.
37. Abrishami V, Vargas J, Li XM, Cheng YF, Marabini R, Sorzano COS, Carazo JM (2015) Alignment of direct detection device micrographs using a robust Optical Flow approach. J Struct Biol 189:163–176.
38. Grant T, Grigorieff N (2015) Measuring the optimal exposure for single particle cryo-EM using a 2.6 angstrom reconstruction of rotavirus VP6. Elife 4:e06980.
39. Scheres SH (2014) Beam-induced motion correction for sub-megadalton cryo-EM particles. Elife 3:e03665.
40. Wade RH, Frank J (1977) Electron-microscope transfer-functions for partially coherent axial illumination and chromatic defocus spread. Optik 49:81–92.
41. Mindell JA, Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. J Struct Biol 142:334–347.

42. Jiang W, Guo F, Liu Z (2012) A graph theory method for determination of cryo-EM image focuses. J Struct Biol 180:343–351.
43. Brown A, Amunts A, Bai XC, Sugimoto Y, Edwards PC, Murshudov G, Scheres SH, Ramakrishnan V (2014) Structure of the large ribosomal subunit from human mitochondria. Science 346:718–722.
44. Henderson R, Sali A, Baker ML, Carragher B, Devkota B, Downing KH, Egelman EH, Feng Z, Frank J, Grigorieff N, Jiang W, Ludtke SJ, Medalia O, Penczek PA, Rosenthal PB, Rossmann MG, Schmid MF, Schroder GF, Steven AC, Stokes DL, Westbrook JD, Wriggers W, Yang H, Young J, Berman HM, Chiu W, Kleywegt GJ, Lawson CL (2012) Outcome of the first electron microscopy validation task force meeting. Structure 20:205–214.
45. LeBarron J, Grassucci RA, Shaikh TR, Baxter WT, Sengupta J, Frank J (2008) Exploration of parameters in cryo-EM leading to an improved density map of the *E. coli* ribosome. J Struct Biol 164:24–32.
46. Kucukelbir A, Sigworth FJ, Tagare HD (2014) Quantifying the local resolution of cryo-EM density maps. Nat Methods 11:63–65.
47. Brown A, Long F, Nicholls RA, Toots J, Emsley P, Murshudov G (2015) Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. Acta Cryst D71:136–153.
48. Ben-Shem A, de Loubresse NG, Melnikov S, Jenner L, Yusupova G, Yusupov M (2011) The structure of the eukaryotic ribosome at 3.0 angstrom resolution. Science 334:1524–1529.
49. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF chimera—a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612.
50. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Cryst D60:2126–2132.
51. Aslett M, Aurrecoechea C, Berriman M, Brestelli J, Brunk BP, Carrington M, Depledge DP, Fischer S, Gajria B, Gao X, Gardner MJ, Gingle A, Grant G, Harb OS, Heiges M, Hertz-Fowler C, Houston R, Innamorato F, Iodice J, Kissinger JC, Kraemer E, Li W, Logan FJ, Miller JA, Mitra S, Myler PJ, Nayak V, Pennington C, Phan I, Pinney DF, Ramasamy G, Rogers MB, Roos DS, Ross C, Sivam D, Smith DF, Srinivasamoorthy G, Stoeckert CJ Jr., Subramanian S, Thibodeau R, Tivey A, Treatman C, Velarde G, Wang H (2010) TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic Acids Res 38:D457–D462. 4
52. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Cryst D58:1948–1954.
53. Yan C, Wan R, Bai R, Huang G, Shi Y (2016) Structure of a yeast activated spliceosome at 3.5 Å resolution. Science 353:904–911.
54. des Georges A, Clarke OB, Zalk R, Yuan Q, Condon KJ, Grassucci RA, Hendrickson WA, Marks AR, Frank J (2016) Structural basis for gating and activation of RyR1. Cell 167:145–157 e117.
55. Merk A, Bartesaghi A, Banerjee S, Falconieri V, Rao P, Davis MI, Pragani R, Boxer MB, Earl LA, Milne JLS, Subramaniam S (2016) Breaking cryo-EM resolution barriers to facilitate drug discovery. Cell 165:1698–1707.

Chapter 7: New insights into Ribosome Structure and Function

This chapter reproduces the manuscript of a book chapter I coauthored with Amy Jobe, Zheng Liu, and Joachim Frank. In particular, my contribution was writing the Parasitic Protozoan subsection. This chapter provides an overview of ribosome structure and function with an emphasis on recent knowledge obtained by single-particle cryo-EM methods.

New Insights into Ribosome Structure and Function

Amy Jobe,[1] Zheng Liu,[1] Cristina Gutierrez-Vargas,[2] and Joachim Frank[1,2]

[1]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032

[2]Department of Biological Sciences, Columbia University, New York, New York 10032

Correspondence: jf2192@cumc.columbia.edu

Cold Spring Harbor Perspectives in Biology
www.cshperspectives.org

In the past 4 years, because of the advent of new cameras, many ribosome structures have been solved by cryoelectron microscopy (cryo-EM) at high, often near-atomic resolution, bringing new mechanistic insights into the processes of translation initiation, peptide elongation, termination, and recycling. Thus, cryo-EM has joined X-ray crystallography as a powerful technique in structural studies of translation. The significance of this new development is that structures of ribosomes in complex with their functional binding partners can now be determined to high resolution in multiple states as they perform their work. The aim of this article is to provide an overview of these new studies and assess the contributions they have made toward an understanding of translation and translational control.

To understand translation and translational control, an understanding of structure as well as dynamics of the various processes is required. The ability to see interactions at the functional sites in atomic detail, facilitated by much of the recent work, has made a large difference in this respect. One mechanism can be singled out that is ubiquitous, with different ramifications in initiation, translocation, decoding, termination, and recycling; in each of these steps of translation, the action of a GTPase is required, and typically the engagement of a factor creates an atomic constellation that triggers guanosine triphosphate (GTP) hydrolysis, followed by changes in both the factor and the ribosome.

There are tantalizing questions on how this action unfolds in each case, and the reason so much is still unanswered is that it occurs quite rapidly, and that its structural characterization therefore requires some kind of interference, such as the use of a nonhydrolyzable GTP analog or mutation at a key site, each time invoking questions about the authenticity of the state trapped. Thus, the nature of the decisive power-stroke that brings us from point A to point B in the processive machinery is uncertain.

Among all steps of translation, eukaryotic initiation is probably the least understood because of the multitude of steps and the complexity of the scanning mechanism leading to the placement of the AUG codon at the P site of the small subunit. Next there come the variants of translation initiation through the viral internal ribosome entry site (IRES) mechanism; here it is necessary to understand how the IRES is able to "push the right buttons" on the ribosome to take over control. Another highlight of this review is the recent progress in the structural

Editors: Michael B. Mathews, Nahum Sonenberg, and John W.B. Hershey
Additional Perspectives on Translation Mechanisms and Control available at www.cshperspectives.org

 doi: 10.1101/cshperspect.a032615
Cite this article as *Cold Spring Harb Perspect Biol* 2019;11:a032615

characterization of ribosomes of parasitic protozoans; here, the reason little is known thus far is related to the prevalence of the diseases that they cause in underdeveloped countries and the politics of science funding.

As translation of the genetic code into polypeptides proceeds through its main phases of initiation, peptide elongation, termination, and recycling, the ribosome undergoes numerous conformational changes and ligand binding/unbinding events. For a long time, steps toward the goal of obtaining information on the structural basis of these events, critical for understanding the mechanistic basis of biological function, have been handicapped by shortcomings of the two main structural visualization techniques, X-ray crystallography and single-particle cryo-electron microscopy (cryo-EM). Single-particle cryo-EM, recently highlighted by the 2017 Nobel Prize in Chemistry, is the visualization of molecules in a thin layer of vitreous ice following rapid plunge-freezing of the sample into a cryogen, or a freezing agent, at liquid-nitrogen temperature (McDowall et al. 1983). Raw data collected on the electron microscope comes in the form of two-dimensional "snapshots" called micrographs, each of which include a few to over 100 copies of a biological molecule. Because molecules are randomly oriented, all view directions are sampled when a sufficient number of micrographs are taken. As the electron dose is spread over a large number of molecules, radiation damage is avoided or strongly reduced. Raw data are subjected to a specialized series of computational techniques, during which individual copies of the specimen are identified and the two-dimensional view represented by each copy is assigned a viewing angle, and a three-dimensional reconstruction of the Coulomb potential distribution (colloquially called "density map") can ultimately be computed.

While X-ray crystallography, starting with publications in 2000 (Ban et al. 2000; Schluenzen et al. 2000; Wimberly et al. 2000), has been the source of many high-resolution atomic models of the ribosome structure, it is intrinsically limited as a technique by the conformational selection taking place during the formation of the crystal, such that many interesting functional states will escape visualization. Conversely, states that are only transiently visited may receive inordinate attention because they happen to be captured in crystals diffracting to high resolution. Cryo-EM, on the other hand, apart from the fact that it does not require crystals, has the advantage that it is capable of displaying all states coexisting in the sample without restriction. The much lower sample volume required is another benefit. However, until recently, this technique has been limited in resolution to 5–6 Ångströms (Å, 10^{-10} meters) for asymmetric molecules (e.g., Hashem et al. 2013a) because of the low quality of the recording medium. This situation has been radically changed with the commercial introduction, in 2012, of direct electron detecting cameras (see McMullan et al. 2009), which allow near-atomic resolution to be reached. For the ribosome field, this has been a game changer as it is now possible to study ligand-binding events (e.g., with transfer RNA [tRNA] or various translation factors) in minute detail. Even ab initio structural modeling is now possible over large stretches of the reconstructed density map; magnesium ions can be located; and ribosomal RNA (rRNA) modifications can be directly identified by the positions of extra density in the map (see Fischer et al. 2015; Liu et al. 2016). This greatly expanded potential of cryo-EM has led an increasing number of crystallographers to adopt the new technique, helped by the fact that both techniques of structure research share many mathematical concepts and modeling tools.

The aim of this article is to give an overview of the recent gain in fundamental knowledge on ribosome structure and function, now that cryo-EM has joined X-ray crystallography as a powerful alternative technique of high-resolution structure research. However, the ensuing large proliferation of structures precludes an exhaustive coverage, and some areas receive particular attention.

Readers unfamiliar with the ribosome may benefit from a brief introduction to its structure (Fig. 1). In all organisms, the ribosome is a 2- to 4.5-megadalton (MDa) structure made up of a small subunit (SSU) and a large subunit (LSU). Both subunits are composed of rRNA surround-

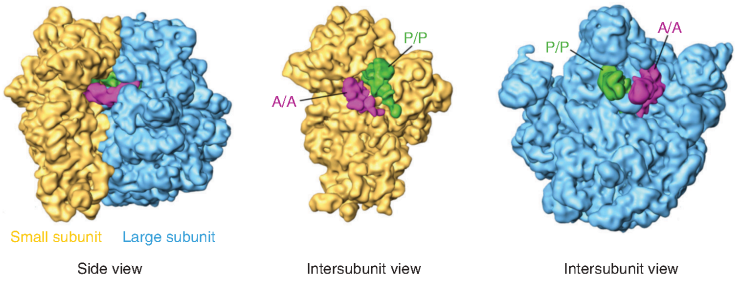

Figure 1. General features of ribosome structure and positions of transfer RNAs (tRNAs). The large subunit appears in blue and the small subunit in yellow. A-site (A/A) tRNA is pink, and P-site (P/P) tRNA is green. Note that the E site is to the *right* of the P site as shown from the intersubunit view of the small subunit and simultaneously to the *left* of the P site as shown from the view of the large subunit. A hybrid-state A/P tRNA would appear as an occupied A site on the small subunit and an occupied P site on the large subunit; in the same way, a P/E tRNA occupies the P site on the small subunit and the E site on the large subunit. (From Agirrezabala et al. 2008; adapted, with permission, from Elsevier © 2008.)

ed by many smaller ribosomal proteins (r-proteins). From a side view, the solvent side of the SSU appears relatively flat, whereas that of the LSU is rounded. A view of the SSU's intersubunit side displays its head region, complete with a leftward-facing beak, above a thin neck region that separates the head from the body of the subunit. The body features a large platform on the right and a smaller shoulder on the left, narrowing somewhat toward a pointed left foot and right foot in eukaryotes, or to a single spur in prokaryotes. Meanwhile, the solvent side of the large subunit appears roughly round with the central protuberance at the top. The messenger RNA (mRNA) passes through the mRNA channel created by the groove between the head and the body of the SSU at its intersubunit face. The mRNA and tRNAs both pass through the three adjacent A (aminoacyl), P (peptidyl), and E (exit) sites, which are located on both the SSU and LSU intersubunit faces.

EUBACTERIAL TRANSLATION AND TRANSLATIONAL CONTROL

In previous studies, before 2013, cryo-EM reconstructions could only be used to study large conformational changes such as intersubunit rotation or subunit head swivel, at intermediate (~5–9 Å) resolutions. At high (2.5–3.5 Å) resolutions readily attainable now, where domain motions and molecular contacts are much better defined, insights have been gained into key events during translation, and into diverse mechanisms of translational control. Additionally, the capability of single-particle cryo-EM to visualize multiple states in a single sample has particularly benefited the understanding of multistep processes such as mRNA-tRNA translocation.

Detailed structural studies have focused on two pivotal stages of the polypeptide elongation cycle (see reviews by Voorhees and Ramakrishnan 2013; Rodnina and Wintermeyer 2016): elongation factor G (EF-G)-assisted mRNA-tRNA translocation and tRNA selection/decoding. As fast and highly dynamic events involving a multitude of domain motions, these are very hard to capture by cryo-EM because of the high degree of ensuing heterogeneity. Particular difficulties are posed by structural states that differ in structural constellations (e.g., base flipping) in a local region only, whereas remaining virtually unchanged elsewhere. In these cases, a technique called local classification can still be used, by singling out the region of interest through imposition of an appropriate mask in the application of the classification algorithm (Penczek et al. 2006). In a variant of the tech-

nique introduced by Scheres (2016), the projected average density is subtracted from the data, thereby increasing the sensitivity of the local classification.

Research on the structural basis of mRNA-tRNA translocation was recently reviewed by Ling and Ermolenko (2016) and Frank (2017). Despite a long history of structural studies, the role of EF-G as it interacts with the ribosome during mRNA-tRNA translocation is still not well understood, even though there is basic agreement on the factor's catalytic action leading to a large (10^4- to 10^6-fold) acceleration of the translocation rate (Katunin et al. 2002). Brownian intersubunit motion of the ribosome has been observed in the absence of factors (Agirrezabala et al. 2008, 2012; Cornish et al. 2008; Julián et al. 2008), and the role of EF-G as providing a "pawl" in the promotion of translocation is generally accepted, implying the existence of some type of power stroke. Specifically, this pawl is manifest in the tip of EF-G's domain IV, which disrupts the codon–anticodon interaction (Frank et al. 2007; Taylor et al. 2007), a notion that has received renewed support in a recent mutation study (Liu et al. 2014).

Intermediate states in factor-free intersubunit rotation (Fischer et al. 2010; Agirrezabala et al. 2012), also observed by X-ray crystallography (Zhang et al. 2009; Tourigny et al. 2013; Zhou et al. 2014), reflect a trajectory across the ribosome's "metastable" free-energy landscape accessible through ambient thermal energy (Munro et al. 2009). Brilot et al. (2013) used cryo-EM to determine the structure of the 70S ribosome bound with EF-G, trapped in the pre-translocation state using the antibiotic viomycin. The structure shows a normally transient state where the A-site tRNA is still bound to the ribosome as well. A surprising discovery of the Steitz group made by X-ray crystallography (Lin et al. 2015) is a large conformational change of EF-G that apparently occurs on the binding of this factor to the ribosome. It is still unclear how this result fits into the overall picture of translocation. In their cryo-EM study, Li et al. (2015) were able to trap EF-G on the pre-translocational ribosome in two different binding configurations, one bound to the rotated and the other to the unrotated ribosome, making use of the mutation H94A. The results of the study allowed the investigators to formulate necessary and sufficient conditions for GTPase activation. Accordingly, stabilization of the factor by contacts with protein S12 and the L11-lobe is required, a condition that is only fulfilled in the rotated configuration.

In the area of tRNA selection, the most interesting question is the discrimination between cognate and near-cognate codon–anticodon pairing, essential for the fidelity of translation. Here it is clear from previous work (Ogle et al. 2001) that discrimination is achieved as a result of at least two successive steps of probing the short codon–anticodon helix with the help of three bases: A1492, A1493, and G650. It has been a challenge to image the ribosome-A/T-tRNA complex for a programmed ribosome and bring order into the sequence of events. Here a comparison between two cryo-EM studies, Agirrezabala et al. (2011) and Loveland et al. (2017) shows the progress recently achieved in the examination of the near-cognate case in a striking way, as the former did not achieve resolution sufficient to delineate the positions of the bases with certainty, whereas the latter shows the complex at close to 3 Å resolution in multiple states of intermediate engagement. In passing, it should also be noted that Fischer et al. (2015), working with cryo-EM, achieved 2.9 Å resolution with the cognate aminoacylated tRNA-EF-Tu•GDP-ribosome complex stabilized with kirromycin, a big step forward from the previous reconstruction of the same complex by Villa et al. (2009), while shedding no new light upon dynamics of tRNA selection. The main gain of the study was the identification and enumeration of rRNA modifications.

It is increasingly recognized that structural snapshots are not sufficient to understand translocation (or any of the steps of translation, for that matter) and that information on real-time dynamics is required. A whole series of recent studies uses single-molecule fluorescence resonance energy transfer (FRET) to study the timing of binding, domain motions on EF-G, intersubunit rotation, small subunit head swivel, and mRNA-tRNA translocation with respect to the

small subunit (Chen et al. 2011, 2014, 2016; Adio et al. 2015; Salsi et al. 2015; Sharma et al. 2016; Wasserman et al. 2016; Kim and Tinoco 2017). A detailed model for the sequence of events inferred from structural and single-molecule FRET (smFRET) studies is described by Wasserman et al. (2016). A new method called single-molecule polarized total internal reflection fluorescence (polTIRF) microscopy provides information not just on distance but on relative orientation between domains. Application of this method to the ribosome indeed allowed the investigators to observe a power stroke of domain IV of EF-G during translocation (Chen et al. 2016).

Important inroads have also been made into the understanding of translational control in eubacteria involving the binding of specific protein factors, an advance helped by the improvement in spatial resolution. Here the emerging common theme is the existence of very specific mechanisms for structural recognition of an abnormal state of the translation process, often going hand-in-hand with stress conditions in the cell, which requires intervention and rescue. EttA is a newly found ATP-binding cassette (ABC) transporter protein, which is found to bind at the ribosomal E site and suppresses translation when the energy supply in the cell falls below a critical level (Boel et al. 2014; Chen et al. 2014). EttA has a "feeler" domain that is apparently able to sense the presence or absence of a peptide on the P-site tRNA and controls the action of the factor through a mechanism of conformational signaling accordingly. RelA is a protein factor active in the stringent response, a response to conditions of amino acid starvation. In *Escherichia coli*, RelA synthesizes alarmones on binding to the ribosome and encountering a deacylated tRNA at the A site (Agirrezabala et al. 2013). Recent cryo-EM studies have uncovered the detailed molecular mechanism, which involves sensing the absence of the aminoacyl group on the CCA end of the strongly distorted, A/T-shaped A-site tRNA (Fig. 2) (Brown et al. 2016; Loveland et al. 2016). BipA is a GTPase closely related to EF-G and engaged at the same factor-binding site of the ribosome, but only if a tRNA is bound at the A site. Curiously, the ribosome was found in the rotated position (Kumar et al. 2015), again implying specific recognition of a local conformational anomaly.

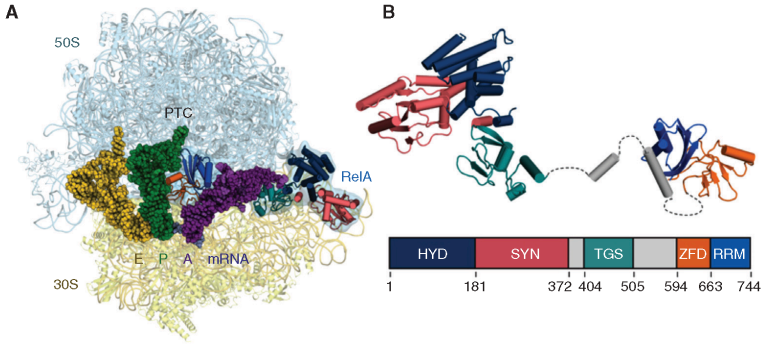

Figure 2. RelA in action on the ribosome, as revealed by cryoelectron microscopy (cryo-EM). (*A*) RelA monitors the CCA end of an incoming putative aminoacyl-transfer RNA (tRNA) and binds to the ribosome on recognizing the absence of an amino acid. (*B*) Model of ribosome-bound form of RelA oriented from amino to carboxyl terminus, with the domain organization shown below; hydrolase (HYD), synthetase (SYN), TGS (threonyl-tRNA synthetase, GTPase and SpoT), Zinc-finger (ZFD) and RNA recognition motif (RRM) domains. Flexible elements between ordered RelA domains are shown in dashed lines. (From Brown et al. 2016; adapted, with permission, from Springer Nature © 2016.)

BipA is known to be essential to survival at low temperature, nutrient depletion, and various other stress conditions, but the molecular mechanism of its action is as yet unknown. Finally, ArfA is a factor that recruits release factor RF2 on encountering a ribosome containing truncated mRNA, a condition characterized by the presence of a peptidyl-tRNA in the P site and an empty A site. Thus ArfA rescues translation in bacteria lacking the trans-translation mechanism. In their recent work, Demo et al. (2017) were able to elucidate the structure of the ArfA-RF2-bound ribosome and describe its likely mechanism of action. Interestingly, this capability adds to the versatility of RF2, the protein factor already known from its classical role as a class I release factor in stop codon recognition and, more recently, in translation quality control (Zaher and Green 2009).

EUKARYOTIC RIBOSOME STRUCTURE AND FUNCTION

In the past few years, studies of ribosome structure for eukaryotes have led to a better appreciation of their diversity, particularly because of the great variation in rRNA expansion segments and the addition of a variety of eukaryotic-specific proteins. As we describe below, special cases are the ribosomes of parasitic protozoans, which for the first time reached resolutions that bring them within reach of drug design, as resolutions in the 2.5 to 3.5 Å range allow bound drug molecules to be directly located and visualized.

Yeast and Higher Eukaryotes

The first X-ray structures of eukaryotic ribosomes, denoted as 80S ribosomes, each comprising one 40S small subunit and one 60S large subunit, were from *Tetrahymena thermophila* (a 40S•eIF1 complex [Rabl et al. 2011] and a 60S•eIF6 complex [Klinge et al. 2011]) and yeast (80S [Ben-Shem et al. 2011]), and were reviewed by Wilson and Doudna Cate (2012) and Klinge et al. (2012). Specific progress with the ribosome structure of yeast was discussed by Yusupova and Yusupov (2014). Yusupov's group also published 16 X-ray structures of the drug-bound yeast ribosome (de Loubresse et al. 2014; Yusupova and Yusupov 2017). 80S ribosomes from a growing number of species were solved by cryo-EM in various states of translation, at close to atomic resolutions: canine (Voorhees and Hegde 2016), porcine (Voorhees et al. 2013, 2014), and human (Khatter et al. 2015).

A number of studies focused on translation mediated by IRESs from viruses on ribosomes from yeast (Fernández et al. 2014; Abeyrathne et al. 2016; Murray et al. 2016), rabbit (Hashem et al. 2013c; Yamamoto et al. 2014, 2015; Muhs et al. 2015), and human (Quade et al. 2015). Besides adding to our rapidly growing structural knowledge base, these reconstructions offer novel insights into the way the eukaryotic ribosome is hijacked by viral mRNA. These works are reviewed in further detail below.

Parasitic Protozoans

In recent cryo-EM studies, several structures of ribosomes from a number of protozoan parasites have been solved at near-atomic resolutions (2.5 to 3 Å), elucidating their unique features, which are distinct from those of other eukaryotes. The resulting structures provide a basis for future functional studies of the translational machinery of protozoan parasites and have established the ribosome as a promising drug target against these human pathogens. These include a number of trypanosomatids and *Plasmodium falciparum*, all causing debilitating and often-fatal diseases. Current treatments against these parasites have poor efficacy, high toxicity, and increasing levels of drug resistance (Andrews et al. 2014). The importance of these parasites in epidemic diseases warrants a special highlight in our review.

Among pathogenic protozoans, the trypanosomatids, *Trypanosoma cruzi*, *Trypanosoma brucei*, and *Leishmania* spp., are a unique family causing insect-borne diseases: Chagas disease, African trypanosomiasis (sleeping sickness), and Leishmaniasis, respectively. Altogether, an estimated 37 million people are infected with these parasites worldwide. The first cryo-EM structures of ribosomes from this family (for

T. cruzi at 12 Å [Gao et al. 2005] and for *T. brucei* at 5.5 Å [Hashem et al. 2013a]) provided a first glimpse and overview of their architecture. More detailed insights have only recently emerged with the advent of the new direct-detector technology. Cryo-EM reconstructions of native 60S subunits from *T. cruzi* at 2.5 Å (Liu et al. 2016, 2017) and *Leishmania donovani* at 2.8 Å (Shalev-Benami et al. 2016), as well as 80S ribosomes of *L. donovani* at 2.9 Å (Zhang et al. 2016) have revealed the ribosome structure in unprecedented detail, making it possible to discern rRNA nucleotides, amino acid side chains, rRNA modifications, solvent molecules, and metal ions. As a result, atomic models are now available for the ribosomes of all three species.

A very striking feature of trypanosomatid ribosomes is the fragmentation of the 28S rRNA in the large subunit, discovered by use of nucleic acid electrophoresis (Campbell et al. 1987). Following the terminology of Hashem et al. (2013a), we refer to the six fragments as LSU-α, LSU-β, and srRNA1 through srRNA4. The two largest pieces, LSU-α and LSU-β, are located on the solvent and interface side, respectively, and provide the scaffold for the large subunit. Locations of these rRNA components and the unique ways they are knitted together in the mature ribosome were described in various depths (Fig. 3) (Hashem et al. 2013a; Liu et al. 2016; Shalev-Benami et al. 2016; Zhang et al. 2016). The srRNA1 piece is at the bottom of the large subunit, as viewed from the solvent side, opposite the 5S rRNA. It lies close to the 5.8S rRNA, without being associated with it. The remaining three pieces srRNA2-4 are in close contact with one another and situated mainly on the left side of the large subunit under the P

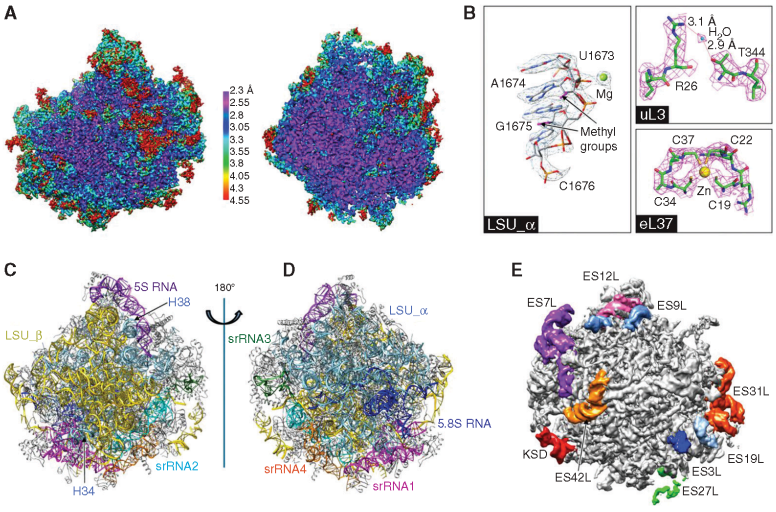

Figure 3. Structure of the large ribosomal subunit of *Trypanosoma cruzi*. (*A*) Sharpened cryoelectron microscopy (cryo-EM) map of the 60S subunit, colored by local resolution and viewed from the subunit interface. (*Left*) Surface view. (*Right*) Central cutaway view. (*B*) Selected views of density for ribosomal RNA (rRNA) and proteins, with associated ions and water molecule. (*C*,*D*) rRNA architecture of the large subunit: interface (*C*) and solvent (*D*) view. (*E*) Some rRNA expansion segments in the unsharpened map of a large subunit, viewed from the solvent side. (From Liu et al. 2016; adapted, with permission, from the National Academy of Sciences in conjunction with Creative Commons licensing.)

stalk (solvent-side view). Strikingly, despite the fragmented nature of the rRNA, the ribosomal functional cores adopt the complete 3D architecture that is common to other eukaryotic ribosomes. Although the exact mechanism of rRNA fragmentation in trypanosomatids is unknown, the high-resolution structures of the *T. cruzi* and *L. donovani* 60S large ribosomal subunit suggest that cleavage occurs during or postassembly (Liu et al. 2016; Shalev-Benami et al. 2016). Further, these structures contain clues on how these rRNA fragments are most likely assembled (Liu et al. 2016). In fact, this "archeological" approach to the interpretation of structure has resulted in a proposal for the chronology of biosynthesis of trypanosomatid ribosomes (Liu et al. 2016).

Compared with other eukaryotic ribosomes, in particular with the yeast ribosome, the cytosolic ribosome in trypanosomatids shows a set of trypanosome-specific components both in RNAs and in proteins. In the large subunit, a number of rRNA kinetoplastid-specific expansion segments (ESs) are located in the periphery and are unusually large. Many r-proteins have carboxy-terminal or amino-terminal extensions. Most remarkably, the r-protein eL19 contains an extremely long carboxy-terminal extension, doubling its size compared to its yeast counterpart. In addition, kinetoplastid ribosomes lack the r-protein eL41, resulting in the absence of the eukaryote/archaea-specific intersubunit bridges eB12 and eB14. In contrast to the large differences seen in the large subunit rRNA, the small subunit is more conserved.

P. falciparum, an apicomplexan parasite transmitted by the bite of an infected female mosquito (*Anopheles* spp.), is the causative agent of the most severe form of human malaria. Malaria constitutes an immense global public health burden; in 2015, it caused an estimated 214 million clinical episodes and 438,000 deaths (www.who.int/mediacentre/factsheets/fs094/en). Using cryo-EM, the parasite-specific features and structural dynamics of the *P. falciparum* 80S (Pf80S) ribosome were recently characterized (Wong et al. 2014, 2017; Sun et al. 2015). In the first of these studies, Wong et al. (2014) solved the structure of the Pf80S ribosome bound to emetine, a eukaryotic protein synthesis inhibitor, at a resolution of 3.2 Å. The Pf80S ribosome displays extensive differences in the lengths of its rRNA expansion segments and protein extensions in comparison to the human ribosome.

As for the r-proteins, 15 of these are extended in *P. falciparum* when compared to yeast ribosomes, which are commonly used as a reference against other eukaryotic ribosomes. Importantly, these extensions result in unique interactions not observed in the ribosomes of the human host (Wong et al. 2014; Sun et al. 2015). For instance, r-protein eL41 has a 14-residue amino-terminal extension that reaches into a pocket formed by 18S rRNA, bridging the large and small subunits. An additional small bridge is formed between the platform of the small subunit and the region around the L1 stalk via interactions between the carboxy-terminal helix extension of eL8 and the carboxy-terminal helix eS1. Stabilizing interactions are also observed near the L1 and P stalks, likely with functional implications on translation in *P. falciparum*. Sun et al. (2015), with a sample purified from a schizont-stage *P. falciparum* cell extract, obtained five reconstructions showing the ribosome either empty or in four different tRNA-binding states, two of which (an 80S with tRNAs in the P/P and E/E sites, and a rotated 80S bound by hybrid-state A/P and P/E tRNAs) are recognizable as part of the canonical elongation cycle. This study is an example for the capability of single-particle cryo-EM to provide multiple functionally relevant structures from the same sample.

Altogether, the newly determined structures of ribosomes from human parasites provide a platform for functional studies of parasite-specific features (rRNA expansion segments, r-protein extensions, dynamics, assembly, etc.) and the development of new therapeutics. Future studies are expected to explore the binding of translation initiation, elongation, and release factors in these parasites, as they are incompletely understood. Thus, there is vast potential for high-resolution cryo-EM to characterize the functions of parasite-specific elements as well as complexes to guide the design of drugs targeting their translational machinery.

Canonical Eukaryotic Translation Initiation

Eukaryotic translation initiation occurs in four stages: 43S preinitiation complex formation, mRNA attachment, scanning and start codon recognition, and large ribosomal subunit joining. This process requires the concerted action of over a dozen protein factors at different times, with concomitant conformational changes in those factors, the ribosome, tRNA, and mRNA (Fig. 4).

In the field of eukaryotic initiation, structural knowledge has lagged far behind the knowledge in other steps of translation. This has been the result mainly of the low stability and structural variability of the various complexes engaged in preinitiation and their pathway toward assembly. The resulting compositional and conformational heterogeneity of such complexes poses challenges in visualization. A breakthrough was achieved with the dis-

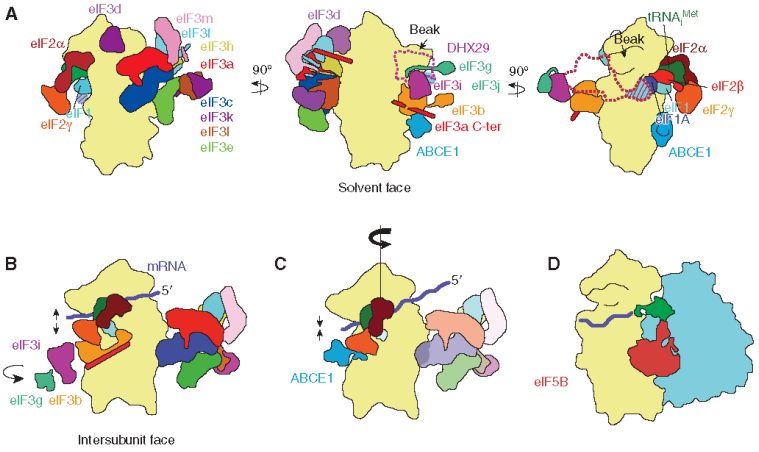

Figure 4. A model of eukaryotic translation initiation. (Note that eukaryotic initiation factor [eIF]4 subunits and eIF5 are not shown.) (*A*) 43S complex. Factors eIF1, eIF1A, ternary complex, and eIF3 bind to the 40S subunit. DHX29 is required only for initiation on highly structured messenger RNAs (mRNAs) in mammals, and is shown as a dotted outline. The binding site of the substoichiometric subunit eIF3j overlaps with part of DHX29's binding site; the area of overlap is indicated with diagonal lines. The mammalian homolog of factor eIF3 is shown; yeast eIF3 comprises only core subunits a and c and peripheral subunits b, i, g, and j. Factor ABCE1 likely binds to the 40S subunit interface, at a conserved GTPase binding site. (*B*) 48S-open complex. mRNA is delivered to the complex, and start codon scanning begins. The latch between rRNA helices 18 and 24 on the 40S subunit opens to allow mRNA entry. Subunits eIF3b, i, and g move from the solvent face to the intersubunit face of the 40S subunit, likely on mRNA binding. (*C*) 48S-closed complex. Scanning concludes on start codon recognition in the 40S subunit P site. The 40S subunit head rotates toward the solvent side of the subunit, aiding the closure of the latch. Subunits eIF3b, i, and g have dissociated from the 40S subunit. Factor ABCE1 binds the 40S subunit interface. Initiator tRNA is accommodated fully in the P site, caused, in part, by eIF2β's releasing its contacts with eIF1 and eIF1A in favor of contacting the 40S subunit head, moving the ternary complex toward the 40S subunit head away from the body, whereas the anticodon remains securely in the P site. Factors eIF1 (shown transparently) and the carboxy-terminal tail (CTT) of eIF1A (not shown) begin to dissociate. Factor eIF5 (not shown) catalyzes GTP hydrolysis on eIF2, after which both factors will dissociate. (*D*) 80S complex. Factor eIF5B catalyzes 60S subunit joining. Upon subunit joining, eIF5B will dissociate with eIF1A (not shown), leaving an elongation-competent 80S complex. (Figure based on data in des Georges et al. 2015 and Simonetti et al. 2016.)

covery that DHX29, a newly discovered initiation factor (Pisareva et al. 2008; Parsyan et al. 2009), stabilizes the preinitiation complex to some extent, sufficient for visualization by cryo-EM (Hashem et al. 2013b). It is worth noting, however, that only 4% of the data showed the complete complex (with the exception of eIF1), and that all other classes presented complexes where some of the factors were missing. It is only because of the power of today's classification methods that such a small portion of the data can be identified. The reason for the low yield of intact complexes is unknown, but it may be related to the low binding affinity of some of the factors and to the forces that act on the specimen during the preparation of the EM grid (Glaeser 2016). This low yield explains the relatively low resolution of the reconstruction. Nevertheless, in this work, the structure and position of eIF3 on the ribosome were well defined for the first time.

43S Preinitiation Complex

Helicase DHX29, first visualized by Hashem et al. (2013b) as part of the 43S preinitiation complex, is required to translate highly structured mRNAs bound to the mammalian 43S complex (Pisareva et al. 2008). DHX29 is positioned on the 40S subunit shoulder near the mRNA entry channel latch, consistent with a role in unwinding double-stranded stems of incoming mRNA. Additionally, the resolution of this 43S complex, 11.6Å, was sufficient to clarify the orientation of the eIF3 core octamer bound to the 40S subunit, which differed significantly from the previous understanding of its orientation (Siridechadilok et al. 2005); in its actual orientation, the eIF3 core contacts much less of the 40S subunit than previously thought.

des Georges et al. (2015) presented the first set of cryo-EM reconstructions that collectively show all but eIF3j of the 13 subunits of mammalian eIF3 in the context of the 43S complex. The PCI/MPN octameric core is attached to the platform of the 40S subunit via subunits eIF3a and eIF3c; the eIF3 core in mammals differs substantially from that in yeast, which comprises only subunits eIF3a and eIF3c. In a separate reconstruction, the investigators resolved peripheral subunits eIF3bgi on the solvent side of the 40S subunit near the shoulder. In a third reconstruction, a density seen previously (Hashem et al. 2013b) at low resolution is now visible in des Georges et al. (2015) at intermediate resolution and tentatively assigned to eIF3d, the first time a cryo-EM density is attributed to this peripheral subunit.

Subunit eIF3j was finally visualized by Aylett et al. (2015) as part of a yeast 40S•eIF1•eIF1A•eIF3 complex. These investigators enriched for complete complexes using a lysine cross-linking technique that tethered components to one another (Erzberger et al. 2014). Subunit eIF3j is a substoichiometric, nonessential component of eIF3, conserved between yeast and mammals (Valásek et al. 2001; see Browning et al. 2001); in the 40S•eIF1•eIF1A•eIF3 complex, eIF3j is situated at the mRNA entry channel on the 40S subunit shoulder at the intersubunit face, near the mRNA entry channel (Fraser et al. 2007). This binding site overlaps with that of DHX29, indicating that their binding associations in the 43S complex are mutually exclusive. Subunit eIF3j contacts eIF1A and, because eIF1 and eIF1A bind cooperatively (Maag and Lorsch 2003), the cryo-EM study by Aylett et al. (2015) is the first evidence that links binding of eIF3 to that of eIF1 and eIF1A. eIF3 makes relatively few contacts with the 40S subunit, so its interactions with eIF1 and eIF1A are particularly important.

Also visible for the first time in a cryo-EM reconstruction is the helical, spectrin-like carboxy-terminal domain of subunit eIF3a, which extends along the solvent side of the 40S subunit, from the eIF3ac subunit core situated on the platform to the peripheral eIF3bgi subunit module near the shoulder.

48S Translation Initiation Complex

For some time, it was difficult to directly assess structural differences between the stages of scanning and start codon recognition during initiation, because the two were thought to differ only by relatively small conformational changes and by start codon placement. Hussain et al. (2014) broke ground in this area, using a tRNA variant

to stabilize start codon recognition to yield a reconstruction of a partial yeast 48S complex (py48S), containing the 40S subunit, eIF1, eIF1A, ternary complex (i.e., eIF2•GTP-Met-$tRNA_i^{Met}$ or nonhydrolyzable GTP analog), and mRNA, with partial putative eIF5 occupancy. The 4 Å structure, which represents an initiation complex just after start codon recognition, presents initiator tRNA, eIF1, and the amino-terminal tail (NTT) of eIF1A together for the first time. According to this structural mapping, the initiator tRNA interacts with the mRNA start codon in the P site; the eIF1A NTT is ordered and stabilizes the codon–anticodon duplex; and eIF1 is seen in a novel conformation in which its two β-hairpins have shifted to avoid a clash with the tRNA in its current close proximity to the mRNA. Additionally, the mRNA entry channel latch between h18 on the 40S subunit head and h34 on its body (Spahn et al. 2001) is in a closed state, as expected on start codon recognition. Furthermore, the reconstruction reveals a 13° clockwise swivel of the 40S subunit head, with the beak having moved toward the solvent face of the subunit, about the axis of h28, the "neck" between the 40S subunit head and body. This swivel motion brings h29 of the 18S rRNA from the 40S subunit head into contact with the anticodon stem-loop (ASL) of initiator tRNA, thereby stabilizing the tRNA in its start codon-bound state.

Llácer et al. (2015) built on this result by introducing at 6 Å the first open-state py48S complex, omitting eIF5 and using an mRNA variant to stabilize a scanning complex (with 40S subunit, eIF1, eIF1A, ternary complex, mRNA, and parts of eIF3 visible) in addition to another py48S-closed complex at 4.9 Å similar to the one in Hussain et al. (2014) (with the same components as the open complex in addition to eIF5). Here "open" and "closed" refer to the state of the mRNA entry channel latch. The 40S head tilts away from the body to open the latch during scanning, which permits mRNA movement through the now widened channel between the 40S subunit head and neck; the latch closes after start codon recognition, to "lock" the mRNA in position.

Specifically, Llácer et al. (2015) showed that on start codon recognition, h28 is compressed, bringing the 40S subunit head closer to the body to close the latch. The 40S subunit body component of the P site is brought into contact with the initiator tRNA's ASL, helping to accommodate $tRNA_i$ fully in the P site. Concomitantly, subunit eIF2β releases its contacts with eIF1 and eIF1A in favor of contacting the 40S subunit head. As a result, the bulk of the ternary complex moves toward the 40S subunit head and away from the body, whereas the ASL is stationary in the 40S subunit P site (Llácer et al. 2015; Simonetti et al. 2016). In conjunction with rearrangements of other initiation factors, including eIF1, eIF1A, and eIF2α, the 40S subunit head positions the initiator tRNA in the P site in a closed, scanning-arrested state.

Simonetti et al. (2016), via a late-stage mammalian 48S subunit reconstruction (containing the 40S subunit, ternary complex, mRNA, eIF3i, and eIF3g), reinterpreted Llácer et al.'s results by swapping the attributions of eIF3b and eIF3i, both of which bear WD40 domains of similar size. The reassigned factor eIF3b appears on the intersubunit face of the 40S subunit in both open and closed complexes, which suggests that it has moved from the 40S subunit solvent face before scanning and start codon recognition, and likely on mRNA attachment. The relocated eIF3b contacts eIF2γ, maintaining the ternary complex in a scanning-competent conformation, as opposed to the 40S head-proximal conformation mentioned above on start codon recognition. A follow-up analysis of the same 48S subunit reconstruction (Mancera-Martínez et al. 2017) indicates that eIF3i and eIF3g likely relocate with eIF3b from the 40S subunit solvent face to the intersubunit face as a module, although whether eIF3i+g remain bound to the 40S subunit is unclear. Importantly, the investigators additionally reassign a density on the 40S subunit intersubunit face to the NTPase ABCE1, positing that ABCE1 binds the interface on start codon recognition, after dissociation of the eIF3b (along with i+g). Moreover, ABCE1 may cycle on and off the 40S subunit throughout ribosome recycling and subsequent translation initiation, possibly acting as an anti-association factor and a gatekeeper to multiple steps of translation initiation.

Fernández et al. (2013) visualized a complex representing the result of the final step of eukaryotic initiation, eIF5B-mediated subunit joining. Their 6.6 Å structure of factor eIF5B in complex with the yeast 80S ribosome and the initiator aminoacylated tRNA in the P/I state—a conformation in the P site with the large-subunit end bound to an initiation factor—was achieved by optimizing the sample in a stepwise fashion in response to shortcomings of early reconstructions of the complex, because of partial occupancy of some components and tRNA heterogeneity. The resultant reconstruction shows that eIF5B binds tightly to ribosomal protein eL40 at the base of the P stalk, such that there would be a steric clash if eL40 were ubiquitinated. In fact, as eL40 is ubiquitinated through the late stages of 60S subunit maturation (Fernández-Pevida et al. 2012), eIF5B promotes subunit joining strictly with mature 60S subunits. Furthermore, some aspects of this complex that are conserved in bacteria—the P/I conformation of the initiator tRNA, salient intersubunit rotation, and the general conformation of ribosome-bound eIF5B—corroborate an earlier low-resolution cryo-EM reconstruction in bacteria (Allen et al. 2005) that featured eIF5B's prokaryotic homolog IF2. This suggests that eIF5B's role in initiation is universally conserved, which would explain why eIF5B is required even when most initiation factors are not during many types of noncanonical 5′-cap-independent initiation.

With all these pieces in place, it is possible now to put a composite model together for the stepwise progression of canonical translation initiation in eukaryotes (Fig. 4).

IRES-Mediated Initiation and Translocation

Cryo-EM has advanced the understanding of noncanonical translation initiation via IRESs. In viruses, IRESs are classified into four types, according in part to the subset of initiation factors that they require. Type 3 IRESs, which require very few initiation factors, and type 4 IRESs, which require none, depart the most from canonical requirements, and recent cryo-EM studies have helped to elucidate the mechanisms of types 3 and 4 IRES function (Kwan and Thompson 2018) and of IRES-bound ribosome behavior.

Type 3 IRESs, such as the classical swine fever virus (CSFV) IRES, bind to the 40S subunit in a way that clashes with the mammalian eIF3 core's ribosomal contacts, eS1, eS26, and eS27. Hashem et al. (2013c) observed that the IRES resolves the clash by displacing the eIF3 core entirely from the ribosome and instead binding eIF3 via the IRES's own domain III, holding eIF3 in a position 55 Å shifted and 60° rotated from its canonical position on the 40S subunit platform.

Initiation with the type 3 IRES appears to rely on 40S subunit *rolling*, a movement usually confined to the elongation stage during canonical translation. Specific to eukaryotes, rolling is a ~6° rotation about the longest axis of the 40S subunit, orthogonal to intersubunit rotation (Budkevich et al. 2011) with the beak moving closer to the 60S subunit. Canonically, rolling occurs in the early pre-translocation state where it serves to shift an incoming tRNA in the A/T conformation, in which the 3′ acceptor end of the tRNA is not yet situated in the 60S A site, into the A/A conformation, which is fully accommodated in the A site. Reverse rolling takes place concurrently with intersubunit reverse rotation to yield the posttranslocation complex (Budkevich et al. 2014). However, Yamamoto et al. (2014) showed that rolling occurs in the final steps of hepatitis C virus (HCV) IRES-mediated initiation. The subunit-joining GTPase eIF5B recruits the 60S subunit to a 40S•HCV-IRES•Met-tRNA$_i^{Met}$ complex in the rolled state, yielding a pre-translocation-like 80S complex presented at 8.2 Å. eIF5B•GTP hydrolysis induces reverse rolling as eIF5B begins to disengage from the complex; this posttranslocation-like state is captured in an 8.6 Å reconstruction. Yamamoto et al. propose that subsequent eIF5B dissociation yields an elongation-competent complex. Thus, the HCV IRES uses canonical-elongation-like rolling to complete initiation.

Fernández et al. (2014) obtained two reconstructions of a model type 4 IRES, the cricket paralysis virus (CrPV) intergenic IRES, bound to a yeast 80S ribosome in rotated and nonrotated states, at sufficiently high resolution (3.7 Å

and 3.8 Å, respectively) to build an atomic model of the full-length IRES. They took advantage of the fact that the 80S ribosome from the yeast *Kluyveromyces lactis* is stable (Rodicio and Heinisch 2013) and does not aggregate at pH 6, and used it to build complexes with the CrPV IRES RNA, as RNA and especially RNA pseudoknots, which are found in the CrPV IRES, are stable at pH 6 (Li and Breaker 1999; Nixon and Giedroc 2000). Importantly, the resultant high-resolution structures revealed that the CrPV IRES binds the ribosome with its start codon-initiator tRNA mimic (Costantino et al. 2008), called pseudoknot I (PKI), in the A site rather than in the P site as the canonical start codon-initiator tRNA complex does; the 80S ribosome•CrPV IRES assembly therefore requires translocation to occur before elongation proceeds, and represents a pre-translocation complex rather than an initiation complex. In fact, the CrPV IRES-bound 80S ribosome is in equilibrium between rotated and nonrotated states (Fernández et al. 2014).

Type 3 and 4 IRESs both bind between the head and body of the 40S subunit in a way that constrains the flexibility of the 40S head position, tilting the head away from the body and thereby forcing the latch into its open position (Quade et al. 2015; Yamamoto et al. 2015; Murray et al. 2016). In this way, the IRESs mimic the induction of head tilt and latch opening by factors eIF1, eIF1A, and eIF3 during canonical initiation (Aylett et al. 2015; Llácer et al. 2015). In the case of the HCV IRES, the model type 3 IRES head tilt of about 17° persists until P-site tRNA binding occurs, which takes place after 60S subunit joining in HCV IRES-driven initiation (Yamamoto et al. 2015). The incoming tRNA displaces the HCV IRES domain II from its position as a wedge between the 40S subunit head and body, and domain II undergoes a 55° rotation resulting in contact with the 28S rRNA on the 60S subunit. In type 4 IRES-driven initiation, stem loops IV and V of the IRES insert themselves between the head and body of the 40S subunit (Fernández et al. 2014; Koh et al. 2014).

During eEF2-dependent translocation, the type 4 IRES must shift through the intersubunit space as a single flexible body. Two groups (Muhs et al. 2015; Murray et al. 2016) elucidated individual mid- and posttranslocation states, respectively, of ribosome-bound type 4 CrPV IRES. In a tour-de-force of cryo-EM data collection and classification, Abeyrathne et al. (2016) captured the entire arc of IRES translocational motion by five high-resolution (3.5 Å to 4.2 Å) cryo-EM reconstructions of the type 4 IRES from Taura syndrome virus (TSV) bound to the 80S ribosome, which was bound in turn by eEF2. The TSV IRES, in an extended conformation before translocation, compresses and then extends again in a stepwise fashion, whereas its PK1 moves from the A site to the P site. Concerted intersubunit rotation and 40S head swivel accompany the movement of the IRES. As the ribosome proceeds from the rotated state (3° more rotated than when eEF2 is absent [Murray et al. 2016]) to the nonrotated state, 40S subunit head swivel progresses from moderate (12°) to high (17°) to minimal (1°). The swiveling action toward the 60S subunit appears to align the 40S subunit head A-site region with the body P-site region to permit the IRES PK1 to advance from the A site to the P site (Abeyrathne et al. 2016) in addition to a previously known role in which swiveling permits P-site tRNA's shift to the E site (Spahn et al. 2004; Schuwirth et al. 2005).

CONCLUSION

When compared to the knowledge of only 4 years ago, the structural information available on ribosomes from eubacteria and eukaryotes is now much richer and more detailed. Especially, the addition of information on multiple states and dynamics obtained by single-particle cryo-EM has helped to broaden the functional knowledge base in eubacterial and eukaryotic translation. As we have shown, there are beautiful examples for the elucidation of detailed mechanisms in the recent structural literature on translational control and tRNA selection in eukaryotes, and IRES-mediated initiation and translocation in eukaryotes. In all, we can look forward to a much-enriched integration of biochemical, structural, and single-molecule

FRET data relating to translation than previously thought possible.

REFERENCES

*Reference is also in this collection.

Abeyrathne PD, Koh CS, Grant T, Grigorieff N, Korostelev AA. 2016. Ensemble cryo-EM uncovers inchworm-like translocation of a viral IRES through the ribosome. *eLife* **5:** e14874.

Adio S, Senyushkina T, Peske F, Fischer N, Wintermeyer W, Rodnina MV. 2015. Fluctuations between multiple EF-G-induced chimeric tRNA states during translocation on the ribosome. *Nat Commun* **6:** 7442.

Agirrezabala X, Lei J, Brunelle JL, Ortiz-meoz RF, Green R, Frank J. 2008. Visualization of the hybrid state of tRNA binding promoted by spontaneous ratcheting of the ribosome. *Mol Cell* **32:** 190–197.

Agirrezabala X, Schreiner E, Trabuco LG, Lei J, Ortiz-Meoz RF, Schulten K, Green R, Frank J. 2011. Structural insights into cognate vs. near-cognate discrimination during decoding. *EMBO J* **30:** 1497–1507.

Agirrezabala X, Liao HY, Schreiner E, Fu J, Ortiz-Meoz RF, Schulten K, Green R, Frank J. 2012. Structural characterization of mRNA-tRNA translocation intermediates. *Proc Natl Acad Sci* **109:** 6094–6099.

Agirrezabala X, Fernandez IS, Kelley AC, Carton DG, Ramakrishnan V, Valle M. 2013. The ribosome triggers the stringent response by RelA via a highly distorted tRNA. *EMBO Rep* **14:** 811–816.

Allen GS, Zavialov A, Gursky R, Ehrenberg M, Frank J. 2005. The cryo-EM structure of a translation initiation complex from *Escherichia coli*. *Cell* **121:** 703–712.

Andrews KT, Fisher G, Skinner-Adams TS. 2014. Drug repurposing and human parasitic protozoan diseases. *Int J Parasitol Drugs Drug Resist* **4:** 95–111.

Aylett CHS, Boehringer D, Erzberger JP, Schaefer T, Nenad Ban N. 2015. Structure of a Yeast 40S–eIF1–eIF1A–eIF3–eIF3j initiation complex. *Nat Struct Mol Biol* **22:** 269–271.

Ban N, Nissen P, Hansen JJ, Moore PB, Steitz TA. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289:** 905–920.

Ben-Shem A, de Loubresse NG, Melnikov S, Jenner L, Yusupova G, Yusupov M. 2011. The structure of the Eukaryotic ribosome at 3.0 Å resolution. *Science* **334:** 1524–1529.

Boel G, Smith PC, Ning W, Englander MT, Chen B, Hashem Y, Testa AJ, Fischer JJ, Wieden HJ, Frank J, et al. 2014. The ABC-F protein EttA gates ribosome entry into the translation elongation cycle. *Nat Struct Mol Biol* **21:** 143–151.

Brilot AF, Korostelev AA, Ermolenko DN, Grigorieff N. 2013. Structure of the ribosome with elongation factor G trapped in the pretranslocation state. *Proc Natl Acad Sci* **110:** 20994–20999.

Brown A, Fernández IS, Gordiyenko Y, Ramakrishnan V. 2016. Ribosome-dependent activation of stringent control. *Nature* **534:** 277–280.

Browning KS, Gallie DR, Hershey JW, Hinnebusch AG, Maitra U, Merrick WC, Norbury C. 2001. Unified nomenclature for the subunits of eukaryotic initiation factor 3. *Trends Biochem Sci* **26:** 284.

Budkevich T, Giesebrecht J, Altman RB, Munro JB, Mielke T, Nierhaus KH, Blanchard SC, Spahn CMT. 2011. Structure and dynamics of the mammalian ribosomal pretranslocation complex. *Mol Cell* **44:** 214–224.

Budkevich TV, Giesebrecht J, Behrmann E, Loerke J, Ramrath DJF, Mielke T, Ismer J, Hildebrand PW, Tung C-S, Nierhaus KH, et al. 2014. Regulation of the mammalian elongation cycle by subunit rolling: A eukaryotic-specific ribosome rearrangement. *Cell* **158:** 121–131.

Campbell DA, Kubo K, Graham Clark C, Boothroyd JC. 1987. Precise identification of cleavage sites involved in the unusual processing of trypanosome ribosomal RNA. *J Mol Biol* **196:** 113–124.

Chen C, Stevens B, Kaur J, Cabral D, Liu H, Wang Y, Zhang H, Rosenblum G, Smilansky Z, Goldman YE, et al. 2011. Single-molecule fluorescence measurements of ribosomal translocation dynamics. *Mol Cell* **42:** 367–377.

Chen B, Boel G, Hashem Y, Ning W, Fei J, Wang C, Gonzalez RL Jr, Hunt JF, Frank J. 2014. EttA regulates translation by binding the ribosomal E site and restricting ribosome-tRNA dynamics. *Nat Struct Mol Biol* **21:** 152–159.

Chen C, Cui X, Beausang JF, Zhang H, Farrell I, Cooperman BS, Goldman YE. 2016. Elongation factor G initiates translocation through a power stroke. *Proc Natl Acad* **113:** 7515–7520.

Cornish PV, Ermolenko DN, Noller HF, Ha T. 2008. Spontaneous intersubunit rotation in single ribosomes. *Mol Cell* **30:** 578–588.

Costantino DA, Pfingsten JS, Rambo RP, Kieft JS. 2008. tRNA-mRNA mimicry drives translation initiation from a viral IRES. *Nat Struct Mol Biol* **15:** 57–64.

De Loubresse NG, Prokhorova I, Holtkamp W, Rodnina MV, Yusupova G, Yusupov M. 2014. Structural basis for the inhibition of the eukaryotic ribosome. *Nature* **513:** 517–522.

Demo G, Svidritskiy E, Madireddy R, Diaz-Avalos R, Grant T, Grigorieff N, Sousa D, Korostelev AA. 2017. Mechanism of ribosome rescue by ArfA and RF2. *eLife* doi: 10.7554/eLife.23687.

des Georges A, Dhote V, Kuhn L, Hellen CU, Pestova TV, Frank J, Hashem Y. 2015. Structure of mammalian eIF3 in the context of the 43S preinitiation complex. *Nature* **525:** 491–495.

Erzberger JP, Stengel F, Pellarin R, Zhang S, Schaefer T, Aylett CHS, Cimermančič P, Boehringer D, Sali A, Aebersold R, et al. 2014. Molecular architecture of the 40S•eIF1•eIF3 translation initiation complex. *Cell* **158:** 1123–1135.

Fernández IS, Bai XC, Hussain T, Kelley AC, Lorsch JR, Ramakrishnan V, Scheres SH. 2013. Molecular architecture of a eukaryotic translational initiation complex. *Science* **342:** 1240585.

Fernández IS, Bai XC, Murshudov G, Scheres SH, Ramakrishnan V. 2014. Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell* **157:** 823–831.

Fernández-Pevida A, Rodríguez-Galán O, Díaz-Quintana A, Kressler D, de la Cruz J. 2012. Yeast ribosomal protein L40 assembles late into precursor 60S ribosomes and is

required for their cytoplasmic maturation. *J Biol Chem* **287:** 38390–383407.

Fischer N, Konevega AL, Wintermeyer W, Rodnina MV, Stark H. 2010. Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. *Nature* **466:** 329–333.

Fischer N, Neumann P, Konevega AL, Bock LV, Ficner R, Rodnina MV, Stark H. 2015. Structure of the *E. coli* ribosome–EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. *Nature* **520:** 567–570.

Frank J. 2017. The translation elongation cycle—Capturing multiple states by cryo-electron microscopy. *Philos Trans R Soc Lond B Biol Sci* **372:** 20160180.

Frank J, Gao H, Sengupta J, Gao N, Taylor DJ. 2007. The process of mRNA-tRNA translocation. *Proc Natl Acad Sci* **104:** 19671–19678.

Fraser CS, Berry KE, Hershey JW, Doudna JA. 2007. eIF3j is located in the decoding center of the human 40S ribosomal subunit. *Mol Cell* **26:** 811–819.

Gao H, Ayub MJ, Levin MJ, Frank J. 2005. The structure of the 80S ribosome from *Trypanosoma cruzi* reveals unique rRNA components. *Proc Natl Acad Sci* **102:** 10206–10211.

Glaeser RM. 2016. How good can cryo-EM become? *Nat Methods* **13:** 28–32.

Hashem Y, des Georges A, Fu J, Buss SN, Jossinet F, Jobe A, Zhang Q, Liao HY, Grassucci RA, Bajaj C, et al. 2013a. High-resolution cryo-electron microscopy structure of the *Trypanosoma brucei* ribosome. *Nature* **494:** 385–389.

Hashem Y, des Georges A, Dhote V, Langlois R, Liao HL, Grassucci RA, Hellen CUT, Pestova TV, Frank J. 2013b. Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* **153:** 1108–1119.

Hashem Y, des Georges A, Dhote V, Langlois R, Liao HY, Grassucci RA, Pestova TV, Hellen CU, Frank J. 2013c. Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature* **503:** 539–543.

Hussain T, Llácer JL, Fernández IS, Munoz A, Martin-Marcos P, Savva CG, Lorsch JR, Hinnebusch AG, Ramakrishnan V. 2014. Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell* **159:** 597–607.

Julián P, Koneveg AL, Scheres SHW, Lázaro M, Gil D, Wintermeyer W, Rodnina MV, Valle M. 2008. Structure of ratcheted ribosomes with tRNAs in hybrid states. *Proc Natl Acad Sci* **105:** 16924–16927.

Katunin VI, Savelsbergh A, Rodnina MV, Wintermeyer W. 2002. Coupling of GTP hydrolysis by elongation factor G to translocation and factor recycling on the ribosome. *Biochemistry* **41:** 12806–12812.

Khatter H, Myasnikov AG, Natchiar SK. 2015. Structure of the human 80S ribosome. *Nature* **520:** 640–645.

Kim HK, Tinoco I Jr. 2017. EF-G catalyzed translocation dynamics in the presence of ribosomal frameshifting stimulatory signals. *Nucl Acids Res* **45:** 2865–2874.

Klinge S, Voigts-Hoffmann F, Leibundgut M, Arpagaus S, Ban N. 2011. Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. *Science* **334:** 941–948.

Klinge S, Voigts-Hoffmann M, Leibundgut M, Ban N. 2012. Atomic structures of the eukaryotic ribosome. *Trends Biochem Sci* **37:** 189–198.

Koh CS, Brilot AF, Grigorieff N, Korostelev AA. 2014. Taura syndrome virus IRES initiates translation by binding its tRNA-mRNA-like structural element in the ribosomal decoding center. *Proc Natl Acad Sci* **111:** 9139–9144.

Kumar V, Chen Y, Ero R, Ahmed T, Tan J, Li Z, Wong ASW, Bhushan S, Gao YG. 2015. Structure of BipA in GTP form bound to the ratcheted ribosome. *Proc Natl Acad Sci* **112:** 10944–10949.

* Kwan T, Thompson SR. 2018. Noncanonical translation initiation in eukaryotes. *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a032672.

Li Y, Breaker RR. 1999. Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2′-hydroxyl group. *J Am Chem Soc* **121:** 5364–5372.

Li W, Liu Z, Koripella RK, Langlois R, Sanyal S, Frank J. 2015. Activation of GTP hydrolysis in mRNA-tRNA translocation by elongation factor G. *Sci Adv* **1:** e1500169.

Lin J, Gagnon MG, Bulkley D, Steitz TA. 2015. Conformational changes of elongation factor G on the ribosome during tRNA translocation. *Cell* **160:** 219–227.

Ling C, Ermolenko DN. 2016. Structural insights into ribosome translocation. *WIREs RNA* **7:** 620–636.

Liu G, Song G, Zhang D, Zhang D, Li Z, Lyu Z, Dong J, Achenbach J, Gong W, Zhao XS, et al. 2014. EF-G catalyzes tRNA translocation by disrupting interactions between decoding center and codon–anticodon duplex. *Nat Struct Mol Biol* **21:** 817–824.

Liu Z, Gutierrez-Vargas C, Wei J, Grassucci RA, Ramesh M, Espina N, Sun M, Tutuncuoglu B, Madison-Antenucci S, Woolford JL, et al. 2016. Structure and assembly model for the *Trypanosoma cruzi* 60S ribosomal subunit. *Proc Natl Acad Sci* **113:** 12174–12179.

Liu Z, Gutierrez-Vargas C, Wei J, Grassucci RA, Sun M, Espina N, Madison-Antenucci S, Tong L, Frank J. 2017. Determination of the ribosome structure to a resolution of 2.5 Å by single-particle cryo-EM. *Prot Sci* **26:** 82–92.

Llácer JL, Hussain T, Marler L, Aitken CE, Thakur A, Lorsch JR, Hinnebusch AG, Ramakrishnan V. 2015. Conformational differences between open and closed states of the eukaryotic translation initiation complex. *Mol Cell* **59:** 399–412.

Loveland AB, Bah E, Madireddy R, Zhang Y, Brilot AF, Grigorieff N, Korostelev AA. 2016. Ribosome•RelA structures reveal the mechanism of stringent response activation. *eLife* doi: 10.7554/eLife.17029.

Loveland AB, Demo G, Grigorieff N, Korostelev AA. 2017. Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nature* **546:** 113–117.

Maag D, Lorsch JR. 2003. Communication between eukaryotic translation initiation factors 1 and 1A on the yeast small ribosomal subunit. *J Mol Biol* **330:** 917–924.

Mancera-Martínez E, Querido J, Valasek LS, Simonetti A, Hashem Y. 2017. ABCE1: A special factor that orchestrates translation at the crossroad between recycling and initiation. *RNA Biol* doi: 10.1080/15476286.2016.1269993.

McDowall AW, Chang JJ, Freeman R, Lepault J, Walter CA, Dubochet J. 1983. Electron microscopy of frozen hydrated

sections of vitreous ice and vitrified biological samples. *J Microsc* **131:** 1–9.

McMullan G, Chen S, Henderson R, Faruqi A. 2009. Detective quantum efficiency of electron area detectors in electron microscopy. *Ultramicroscopy* **109:** 1126–1143.

Muhs M, Hilal T, Mielke T, Skabkin MA, Sanbonmatsu KY, Pestova TV, Spahn CMT. 2015. Cryo-EM of ribosomal 80S complexes with termination factors reveals the translocated cricket paralysis virus IRES. *Mol Cell* **57:** 422–432.

Munro JB, Sanbonmatsu KY, Spahn CMT, Blanchard SC. 2009. Navigating the ribosome's metastable energy landscape. *Trends Biochem Sci* **34:** 390–400.

Murray J, Savva CG, Shin BS, Dever TE, Ramakrishnan V, Fernández IS. 2016. Structural characterization of ribosome recruitment and translocation by type IV IRES. *eLife* doi: 10.7554/eLife.13567.001.

Nixon PL, Giedroc DP. 2000. Energetics of a strongly pH dependent RNA tertiary structure in a frameshifting pseudoknot. *J Mol Biol* **296:** 659–671.

Ogle JM, Brodersen DE, Clemons WM Jr, Tarry MJ, Carter AP, Ramakrishnan V. 2001. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* **292:** 897–902.

Parsyan A, Shahbaziana D, Martineaua Y, Petroulakisa E, Alaina T, Larssona O, Mathonneta G, Tettweilera G, Hellen CU, Pestova TV, et al. 2009. The helicase protein DHX29 promotes translation initiation, cell proliferation, and tumorigenesis. *Proc Natl Acad Sci* **106:** 22217–22222.

Penczek PA, Frank J, Spahn CM. 2006. A method of focused classification, based on the bootstrap 3D variance analysis, and its application to EF-G-dependent translocation. *J Struct Biol* **154:** 184–194.

Pisareva VP, Pisarev AV, Komar AA, Hellen CU, Pestova TV. 2008. Translation initiation on mammalian mRNAs with structured 5′UTRs requires DExH-box protein DHX29. *Cell* **135:** 1237–1250.

Quade N, Boehringer D, Leibundgut M, van den Heuvel J, Ban N. 2015. Cryo-EM structure of Hepatitis C virus IRES bound to the human ribosome at 3.9 Å resolution. *Nat Commun* **6:** 8646.

Rabl J, Leibundgut M, Ataide SF, Haag A, Ban N. 2011. Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science* **331:** 730–736.

Rodicio R, Heinisch JJ. 2013. Yeast on the milky way: Genetics, physiology and biotechnology of *Kluyveromyces lactis*. *Yeast* **30:** 165–177.

Rodnina MV, Wintermeyer W. 2016. Protein elongation, co-translational folding and targeting. *J Mol Biol* **428:** 2165–2185.

Salsi E, Farah E, Netter Z, Dann J, Ermolenko DN. 2015. Movement of elongation factor G between compact and extended conformations. *J Mol Biol* **427:** 454–467.

Scheres SH. 2016. Processing of structurally heterogeneous cryo-EM data in RELION. *Methods Enzymol* **579:** 125–157.

Schluenzen F, Tocilj A, Zarivach R, Harms J, Gluehmann M, Janell D, Bashan A, Bartels H, Agmon H, Franceschi F, et al. 2000. Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell* **102:** 615–623.

Schuwirth BS, Borovinskaya MA, Hau CW, Zhang W, Vila-Sanjurjo A, Holton JM, Cate JH. 2005. Structures of the bacterial ribosome at 3.5 Å resolution. *Science* **310:** 827–834.

Simonetti A, Brito Querido J, Myasnikov AG, Mancera-Martínez E, Renaud A, Kuhn L, Hashem Y. 2016. eIF3 peripheral subunits rearrangement after mRNA binding and start-codon recognition. *Mol Cell* **63:** 206–217.

Siridechadilok B, Fraser CS, Hall RJ, Doudna JA, Nogales E. 2005. Structural roles for human translation factor eIF3 in initiation of protein synthesis. *Science* **310:** 1513–1515.

Shalev-Benami M, Zhang Y, Matzov D, Halfon Y, Zackay A, Rozenberg H, Zimmerman E, Bashan A, Jaffe CL, Yonath A, et al. 2016. 2.8 Å cryo-EM structure of the large ribosomal subunit from the eukaryotic parasite *Leishmania*. *Cell Rep* **16:** 288–294.

Sharma H, Adio S, Senyushkina T, Belardinelli R, Peske F, Rodnina MV. 2016. Kinetics of spontaneous and EF-G-accelerated rotation of ribosomal subunits. *Cell Rep* **16:** 2187–2196.

Spahn CM, Kieft JS, Grassucci RA, Penczek PA, Zhou K, Doudna JA, Frank J. 2001. Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. *Science* **291:** 1959–1962.

Spahn CM, Gomez-Lorenzo MG, Grassucci RA, Jørgensen R, Andersen GR, Beckmann R, Penczek PA, Ballesta JPG, Frank J. 2004. Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation. *EMBO J* **23:** 1008–1019.

Sun M, Li W, Blomqvist K, Das S, Hashem Y, Dvorin JD, Frank J. 2015. Dynamical features of the *Plasmodium falciparum* ribosome during translation. *Nucl Acids Res* **43:** 10515–10524.

Taylor DJ, Nilsson J, Merrill AR, Andersen GR, Nissen P, Frank J. 2007. Structures of modified eEF2 80S ribosome complexes reveal the role of GTP hydrolysis in translocation. *EMBO J* **26:** 2421–2431.

Tourigny DS, Fernandez IS, Kelley AC, Ramakrishnan V. 2013. Elongation factor G bound to the ribosome in an intermediate state of translocation. *Science* **340:** 1235490.

Valásek L, Phan L, Schoenfeld LW, Valásková V, Hinnebusch AG. 2001. Related eIF3 subunits TIF32 and HCR1 interact with an RNA recognition motif in PRT1 required for eIF3 integrity and ribosome binding. *EMBO J* **20:** 891–904.

Villa E, Sengupta J, Trabuco LG, LeBarron J, Baxter WT, Shaikh TR, Grassucci RA, Nissen P, Ehrenberg M, Schulten K, Frank J. 2009. Ribosome-induced changes in elongation factor Tu conformation control GTP hydrolysis. *Proc Natl Acad Sci* **106:** 1063–1068.

Voorhees RM, Hegde RS. 2016. Structure of the Sec61 channel opened by a signal sequence. *Science* **351:** 88–91.

Voorhees RM, Ramakrishnan V. 2013. Structural basis of the translational elongation cycle. *Annu Rev Biochem* **82:** 203–236.

Voorhees RM, Fernández IS, Scheres SHW, Hegde RS. 2013. Structure of the mammalian ribosome-Sec61 complex to 3.4 Å resolution. *Cell* **157:** 1632–1643.

Voorhees RM, Fernandez IS, Scheres SHW, Hegde RS. 2014. Structure of the mammalian ribosome-Sec61 complex to 3.4 Å resolution. *Cell* **157:** 1632–1643.

Wasserman MR, Alejo, JL, Altman, RB, Blanchard, SC. 2016. Multiperspective smFRET reveals rate-determining late

intermediates of ribosomal translocation. *Nat Struct Mol Biol* **23:** 333–341.

Wilson DN, Doudna Cate JH. 2012. The structure and function of the Eukaryotic ribosome. *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a011536.

Wimberly BT, Brodersen DE, Clemons WM, Morgan-Warren RJ, Carter AP, Vonrhein C, Hartsch T, Ramakrishnan V. 2000. Structure of the 30S ribosomal subunit. *Nature* **407:** 327–339.

Wong W, Bai X, Brown A, Fernandez IS, Hanssen E, Condron M, Tan YH, Baum J, Scheres SHW. 2014. Cryo-EM structure of the *Plasmodium falciparum* 80S ribosome bound to the anti-protozoan drug emetine. *eLife* **3:** e03080.

Wong W, Bai XC, Sleebs BE, Triglia T, Brown A, Thompson JK, Jackson KE, Hanssen E, Marapana DS, Fernandez IS, et al. 2017. Mefloquine targets the *Plasmodium falciparum* 80S ribosome to inhibit protein synthesis. *Nat Microbiol* **2:** 17031.

Yamamoto H, Unbehaun A, Loerke J, Behrmann E, Collier M, Bürger J, Mielke T, Spahn CM. 2014. Structure of the mammalian 80S initiation complex with initiation factor 5B on HCV-IRES RNA. *Nat Struct Mol Biol* **21:** 721–727.

Yamamoto H, Collier M, Loerke J, Ismer J, Schmidt A, Hilal T, Sprink T, Yamamoto K, Mielke T, Bürger J, et al. 2015. Molecular architecture of the ribosome-bound hepatitis C virus internal ribosomal entry site RNA. *EMBO J* **34:** 3042–3058.

Yusupova G, Yusupov M. 2014. High-resolution structure of the eukaryotic 80S ribosome. *Ann Rev Biochem* **83:** 467–486.

Yusupova G, Yusupov M. 2017. Crystal structure of eukaryotic ribosome and its complexes with inhibitors. *Philos Trans R Soc Lond B Biol Sci* **372:** 20160184.

Zaher HS, Green R. 2009. Quality control by the ribosome following peptide bond formation. *Nature* **457:** 161–166.

Zhang W, Dunkle J, Cate JHD. 2009. Structures of the ribosome in intermediate states of ratcheting. *Science* **325:** 1014–1017.

Zhang X, Lai M, Chang W, Yu I, Ding K, Mrazek J, Ng HL, Yang OO, Maslov DA, Zhou ZH. 2016. Structures and stabilization of kinetoplastid-specific split rRNAs revealed by comparing leishmanial and human ribosomes. *Nat Commun* **7:** 13223.

Zhou J, Lancaster L, Donohue JP, Noller HF. 2014. How the ribosome hands the A-site tRNA to the P site during EF-G-catalyzed translocation. *Science* **345:** 1188–1191.

Cold Spring Harbor Perspectives in Biology

New Insights into Ribosome Structure and Function

Amy Jobe, Zheng Liu, Cristina Gutierrez-Vargas and Joachim Frank

Cold Spring Harb Perspect Biol 2019; doi: 10.1101/cshperspect.a032615 originally published online June 14, 2018

References

1. Melnikov, S., et al., *One core, two shells: bacterial and eukaryotic ribosomes.* Nat Struct Mol Biol, 2012. **19**(6): p. 560-7.
2. Ben-Shem, A., et al., *The structure of the eukaryotic ribosome at 3.0 A resolution.* Science, 2011. **334**(6062): p. 1524-9.
3. Stevens, A.R. and P.F. Pachler, *Discontinuity of 26 s rRNA in Acanthamoeba castellani.* J Mol Biol, 1972. **66**(2): p. 225-37.
4. van Keulen, H., et al., *Characterization of a 54-nucleotide gap region in the 28S rRNA gene of Schistosoma mansoni.* Mol Biochem Parasitol, 1991. **45**(2): p. 205-14.
5. Ware, V.C., R. Renkawitz, and S.A. Gerbi, *rRNA processing: removal of only nineteen bases at the gap between 28S alpha and 28S beta rRNAs in Sciara coprophila.* Nucleic Acids Res, 1985. **13**(10): p. 3581-97.
6. Navarro-Rodenas, A., A. Carra, and A. Morte, *Identification of an Alternative rRNA Post-transcriptional Maturation of 26S rRNA in the Kingdom Fungi.* Front Microbiol, 2018. **9**: p. 994.
7. Nomura, T., et al., *Characterization of silk gland ribosomes from a bivoltine caddisfly, Stenopsyche marmorata: translational suppression of a silk protein in cold conditions.* Biochem Biophys Res Commun, 2016. **469**(2): p. 210-5.
8. Pellegrini, M., J. Manning, and N. Davidson, *Sequence arrangement of the rDNA of Drosophila melanogaster.* Cell, 1977. **10**(2): p. 213-4.
9. White, T.C., G. Rudenko, and P. Borst, *Three small RNAs within the 10 kb trypanosome rRNA transcription unit are analogous to domain VII of other eukaryotic 28S rRNAs.* Nucleic Acids Res, 1986. **14**(23): p. 9471-89.
10. Leon, W., D.L. Fouts, and J. Manning, *Sequence arrangement of the 16S and 26S rRNA genes in the pathogenic haemoflagellate Leishmania donovani.* Nucleic Acids Res, 1978. **5**(2): p. 491-504.
11. Burgin, A.B., et al., *The excision of intervening sequences from Salmonella 23S ribosomal RNA.* Cell, 1990. **60**(3): p. 405-14.
12. Eckert, W.A., et al., *Introduction of hidden breaks during rRNA maturation and ageing in Tetrahymena pyriformis.* Eur J Biochem, 1978. **87**(3): p. 607-16.
13. Melen, G.J., et al., *Novel processing in a mammalian nuclear 28S pre-rRNA: tissue-specific elimination of an 'intron' bearing a hidden break site.* Embo j, 1999. **18**(11): p. 3107-18.
14. Azpurua, J., et al., *Naked mole-rat has increased translational fidelity compared with the mouse, as well as a unique 28S ribosomal RNA cleavage.* Proc Natl Acad Sci U S A, 2013. **110**(43): p. 17350-5.
15. Evguenieva-Hackenberg, E., *Bacterial ribosomal RNA in pieces.* Mol Microbiol, 2005. **57**(2): p. 318-25.
16. Liu, Z., et al., *Structure and assembly model for the Trypanosoma cruzi 60S ribosomal subunit.* Proc Natl Acad Sci U S A, 2016. **113**(43): p. 12174-12179.
17. Liu, Z., et al., *Determination of the ribosome structure to a resolution of 2.5 A by single-particle cryo-EM.* Protein Sci, 2017. **26**(1): p. 82-92.

18. Jobe, A., et al., *New Insights into Ribosome Structure and Function.* Cold Spring Harb Perspect Biol, 2019. **11**(1).
19. Schur, F.K., *Toward high-resolution in situ structural biology with cryo-electron tomography and subtomogram averaging.* Curr Opin Struct Biol, 2019. **58**: p. 1-9.
20. Sgro, G.G. and T.R.D. Costa, *Cryo-EM Grid Preparation of Membrane Protein Samples for Single Particle Analysis.* Front Mol Biosci, 2018. **5**: p. 74.
21. Russo, C.J. and L.A. Passmore, *Electron microscopy: Ultrastable gold substrates for electron cryomicroscopy.* Science, 2014. **346**(6215): p. 1377-80.
22. Pantelic, R.S., et al., *Graphene: Substrate preparation and introduction.* J Struct Biol, 2011. **174**(1): p. 234-8.
23. Russo, C.J. and L.A. Passmore, *Controlling protein adsorption on graphene for cryo-EM using low-energy hydrogen plasmas.* Nat Methods, 2014. **11**(6): p. 649-52.
24. Russo, C.J. and L.A. Passmore, *Ultrastable gold substrates: Properties of a support for high-resolution electron cryomicroscopy of biological specimens.* J Struct Biol, 2016. **193**(1): p. 33-44.
25. Drulyte, I., et al., *Approaches to altering particle distributions in cryo-electron microscopy sample preparation.* Acta Crystallogr D Struct Biol, 2018. **74**(Pt 6): p. 560-571.
26. Brilot, A.F., et al., *Beam-induced motion of vitrified specimen on holey carbon film.* J Struct Biol, 2012. **177**(3): p. 630-7.
27. Frank, J., *Three-dimensional electron microscopy of macromolecular assemblies : visualization of biological molecules in their native state*. 2006, New York: Oxford University Press.
28. Thompson, R.F., et al., *An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology.* Methods, 2016. **100**: p. 3-15.
29. Danev, R., D. Tegunov, and W. Baumeister, *Using the Volta phase plate with defocus for cryo-EM single particle analysis.* Elife, 2017. **6**.
30. Zhang, X. and Z.H. Zhou, *Limiting factors in atomic resolution cryo electron microscopy: no simple tricks.* J Struct Biol, 2011. **175**(3): p. 253-63.
31. Wang, L., Y. Shkolnisky, and A. Singer *A Fourier-based Approach for Iterative 3D Reconstruction from Cryo-EM Images*. arXiv e-prints, 2013.
32. Afshari, A., C. Uhde-Stone, and B. Lu, *A cooled CCD camera-based protocol provides an effective solution for in vitro monitoring of luciferase.* Biochem Biophys Res Commun, 2015. **458**(3): p. 543-548.
33. Li, X., et al., *Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM.* Nat Methods, 2013. **10**(6): p. 584-90.
34. Zheng, S.Q., et al., *MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy.* Nat Methods, 2017. **14**(4): p. 331-332.
35. Grant, T. and N. Grigorieff, *Measuring the optimal exposure for single particle cryo-EM using a 2.6 A reconstruction of rotavirus VP6.* Elife, 2015. **4**: p. e06980.
36. Scheres, S.H., *Beam-induced motion correction for sub-megadalton cryo-EM particles.* Elife, 2014. **3**: p. e03665.
37. Rubinstein, J.L. and M.A. Brubaker, *Alignment of cryo-EM movies of individual particles by optimization of image translations.* J Struct Biol, 2015. **192**(2): p. 188-95.

38. Rohou, A. and N. Grigorieff, *CTFFIND4: Fast and accurate defocus estimation from electron micrographs.* J Struct Biol, 2015. **192**(2): p. 216-21.
39. Zhang, K., *Gctf: Real-time CTF determination and correction.* J Struct Biol, 2016. **193**(1): p. 1-12.
40. Ludtke, S.J., P.R. Baldwin, and W. Chiu, *EMAN: semiautomated software for high-resolution single-particle reconstructions.* J Struct Biol, 1999. **128**(1): p. 82-97.
41. Scheres, S.H., *Semi-automated selection of cryo-EM particles in RELION-1.3.* J Struct Biol, 2015. **189**(2): p. 114-22.
42. Singer, A. *Mathematics for cryo-electron microscopy*. arXiv e-prints, 2018.
43. Van Heel, M., *Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction.* Ultramicroscopy, 1987. **21**(2): p. 111-23.
44. Radermacher, M., et al., *Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of Escherichia coli.* J Microsc, 1987. **146**(Pt 2): p. 113-36.
45. Sigworth, F.J., *A maximum-likelihood approach to single-particle image refinement.* J Struct Biol, 1998. **122**(3): p. 328-39.
46. Scheres, S.H., *RELION: implementation of a Bayesian approach to cryo-EM structure determination.* J Struct Biol, 2012. **180**(3): p. 519-30.
47. van Heel, M. and J. Frank, *Use of multivariate statistics in analysing the images of biological macromolecules.* Ultramicroscopy, 1981. **6**(2): p. 187-94.
48. Frank, J., et al., *Reconstruction of glutamine synthetase using computer averaging.* Ultramicroscopy, 1978. **3**(3): p. 283-90.
49. Chen, S., et al., *High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy.* Ultramicroscopy, 2013. **135**: p. 24-35.
50. Liao, H.Y. and J. Frank, *Definition and estimation of resolution in single-particle reconstructions.* Structure, 2010. **18**(7): p. 768-75.
51. Shannon, C.E., *A Mathematical Theory of Communication.* Bell System Technical Journal, 1948. **27**(3): p. 379-423.
52. Neumann, P., A. Dickmanns, and R. Ficner, *Validating Resolution Revolution.* Structure, 2018. **26**(5): p. 785-795 e4.
53. Barad, B.A., et al., *EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy.* Nat Methods, 2015. **12**(10): p. 943-6.
54. Kovalevskiy, O., et al., *Overview of refinement procedures within REFMAC5: utilizing data from different sources.* Acta Crystallogr D Struct Biol, 2018. **74**(Pt 3): p. 215-227.
55. Afonine, P.V., et al., *Real-space refinement in PHENIX for cryo-EM and crystallography.* Acta Crystallogr D Struct Biol, 2018. **74**(Pt 6): p. 531-544.
56. Voorhees, R.M. and V. Ramakrishnan, *Structural basis of the translational elongation cycle.* Annu Rev Biochem, 2013. **82**: p. 203-36.
57. Spahn, C.M., et al., *Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation.* EMBO J, 2004. **23**(5): p. 1008-19.
58. Taylor, D.J., et al., *Structures of modified eEF2 80S ribosome complexes reveal the role of GTP hydrolysis in translocation.* EMBO J, 2007. **26**(9): p. 2421-31.

59. Schuller, A.P. and R. Green, *Roadblocks and resolutions in eukaryotic translation.* Nat Rev Mol Cell Biol, 2018. **19**(8): p. 526-541.
60. Ratje, A.H., et al., *Head swivel on the ribosome facilitates translocation by means of intra-subunit tRNA hybrid sites.* Nature, 2010. **468**(7324): p. 713-6.
61. Dever, T.E. and R. Green, *The elongation, termination, and recycling phases of translation in eukaryotes.* Cold Spring Harb Perspect Biol, 2012. **4**(7): p. a013706.
62. Palade, G.E., *Studies on the endoplasmic reticulum. II. Simple dispositions in cells in situ.* J Biophys Biochem Cytol, 1955. **1**(6): p. 567-82.
63. Lake, J.A., *Ribosome structure determined by electron microscopy of Escherichia coli small subunits, large subunits and monomeric ribosomes.* J Mol Biol, 1976. **105**(1): p. 131-9.
64. Frank, J., et al., *Three-dimensional reconstruction of the 70S Escherichia coli ribosome in ice: the distribution of ribosomal RNA.* J Cell Biol, 1991. **115**(3): p. 597-605.
65. Frank, J., et al., *A model of the translational apparatus based on a three-dimensional reconstruction of the Escherichia coli ribosome.* Biochem Cell Biol, 1995. **73**(11-12): p. 757-65.
66. Wimberly, B.T., et al., *Structure of the 30S ribosomal subunit.* Nature, 2000. **407**(6802): p. 327-39.
67. Ban, N., et al., *The complete atomic structure of the large ribosomal subunit at 2.4 A resolution.* Science, 2000. **289**(5481): p. 905-20.
68. Yusupov, M.M., et al., *Crystal structure of the ribosome at 5.5 A resolution.* Science, 2001. **292**(5518): p. 883-96.
69. Spahn, C.M., et al., *Structure of the 80S ribosome from Saccharomyces cerevisiae--tRNA-ribosome and subunit-subunit interactions.* Cell, 2001. **107**(3): p. 373-86.
70. Taylor, D.J., et al., *Comprehensive molecular structure of the eukaryotic ribosome.* Structure, 2009. **17**(12): p. 1591-1604.
71. Armache, J.P., et al., *Localization of eukaryote-specific ribosomal proteins in a 5.5-A cryo-EM map of the 80S eukaryotic ribosome.* Proc Natl Acad Sci U S A, 2010. **107**(46): p. 19754-9.
72. Armache, J.P., et al., *Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-A resolution.* Proc Natl Acad Sci U S A, 2010. **107**(46): p. 19748-53.
73. Klinge, S., et al., *Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6.* Science, 2011. **334**(6058): p. 941-8.
74. Rabl, J., et al., *Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1.* Science, 2011. **331**(6018): p. 730-6.
75. Brown, A. and S. Shao, *Ribosomes and cryo-EM: a duet.* Curr Opin Struct Biol, 2018. **52**: p. 1-7.
76. Ke, Z., et al., *Translation fidelity coevolves with longevity.* Aging Cell, 2017. **16**(5): p. 988-993.
77. Shao, S., et al., *Decoding Mammalian Ribosome-mRNA States by Translational GTPase Complexes.* Cell, 2016. **167**(5): p. 1229-1240 e15.
78. Fang, X., et al., *Adaptations to a subterranean environment and longevity revealed by the analysis of mole rat genomes.* Cell Rep, 2014. **8**(5): p. 1354-64.

79. Hashem, Y., et al., *High-resolution cryo-electron microscopy structure of the Trypanosoma brucei ribosome.* Nature, 2013. **494**(7437): p. 385-9.
80. Anger, A.M., et al., *Structures of the human and Drosophila 80S ribosome.* Nature, 2013. **497**(7447): p. 80-5.
81. Natchiar, S.K., et al., *Visualization of chemical modifications in the human 80S ribosome structure.* Nature, 2017. **551**(7681): p. 472-477.
82. Emsley, P., et al., *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
83. Liebschner, D., et al., *Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix.* Acta Crystallogr D Struct Biol, 2019. **75**(Pt 10): p. 861-877.
84. Bretscher, M.S., *Translocation in protein synthesis: a hybrid structure model.* Nature, 1968. **218**(5142): p. 675-7.
85. Spirin, A.S., *A model of the functioning ribosome: locking and unlocking of the ribosome subparticles.* Cold Spring Harbor symposia on quantitative biology, 1969. **34**: p. 197-207.
86. Moazed, D. and H.F. Noller, *Intermediate states in the movement of transfer RNA in the ribosome.* Nature, 1989. **342**(6246): p. 142-8.
87. Odom, O.W., W.D. Picking, and B. Hardesty, *Movement of tRNA but not the nascent peptide during peptide bond formation on ribosomes.* Biochemistry, 1990. **29**(48): p. 10734-10744.
88. Frank, J. and R.K. Agrawal, *A ratchet-like inter-subunit reorganization of the ribosome during translocation.* Nature, 2000. **406**(6793): p. 318-322.
89. Valle, M., et al., *Locking and unlocking of ribosomal motions.* Cell, 2003. **114**(1): p. 123-134.
90. Horan, L.H. and H.F. Noller, *Intersubunit movement is required for ribosomal translocation.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(12): p. 4881-4885.
91. Connell, S.R., et al., *Structural basis for interaction of the ribosome with the switch regions of GTP-bound elongation factors.* Molecular cell, 2007. **25**(5): p. 751-764.
92. Schuwirth, B.S., et al., *Structures of the bacterial ribosome at 3.5 A resolution.* Science (New York, N.Y.), 2005. **310**(5749): p. 827-834.
93. Budkevich, T.V., et al., *Regulation of the mammalian elongation cycle by subunit rolling: a eukaryotic-specific ribosome rearrangement.* Cell, 2014. **158**(1): p. 121-131.
94. Zhang, J., et al., *Mechanisms of ribosome stalling by SecM at multiple elongation steps.* eLife, 2015. **4**: p. e09684.
95. Trabuco, L.G., et al., *The role of L1 stalk-tRNA interaction in the ribosome elongation cycle.* Journal of molecular biology, 2010. **402**(4): p. 741-760.
96. Berk, V., et al., *Structural basis for mRNA and tRNA positioning on the ribosome.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(43): p. 15830-15834.
97. Yusupov, M.M., et al., *Crystal structure of the ribosome at 5.5 A resolution.* Science (New York, N.Y.), 2001. **292**(5518): p. 883-896.
98. Korostelev, A., et al., *Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements.* Cell, 2006. **126**(6): p. 1065-77.

99. Selmer, M., et al., *Structure of the 70S ribosome complexed with mRNA and tRNA.* Science, 2006. **313**(5795): p. 1935-42.
100. Fei, J., et al., *Coupling of ribosomal L1 stalk and tRNA dynamics during translation elongation.* Mol Cell, 2008. **30**(3): p. 348-59.
101. Cornish, P.V., et al., *Following movement of the L1 stalk between three functional states in single ribosomes.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(8): p. 2571-2576.
102. Yusupova, G., et al., *Structural basis for messenger RNA movement on the ribosome.* Nature, 2006. **444**(7117): p. 391-394.
103. Valle, M., et al., *Cryo-EM reveals an active role for aminoacyl-tRNA in the accommodation process.* The EMBO journal, 2002. **21**(13): p. 3557-3567.
104. Gao, H., et al., *RF3 induces ribosomal conformational changes responsible for dissociation of class I release factors.* Cell, 2007. **129**(5): p. 929-941.
105. Nakane, T., et al., *Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION.* Elife, 2018. **7**.
106. Alexander, N., N. Woetzel, and J. Meiler, *bcl::Cluster : A method for clustering biological molecules coupled with visualization in the Pymol Molecular Graphics System.* IEEE Int Conf Comput Adv Bio Med Sci, 2011. **2011**: p. 13-18.
107. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis.* J Comput Chem, 2004. **25**(13): p. 1605-12.
108. Ortiz-Meoz, R.F. and R. Green, *Helix 69 is key for uniformity during substrate selection on the ribosome.* J Biol Chem, 2011. **286**(29): p. 25604-10.
109. Liiv, A. and M. O'Connor, *Mutations in the intersubunit bridge regions of 23 S rRNA.* The Journal of biological chemistry, 2006. **281**(40): p. 29850-29862.
110. Piekna-Przybylska, D., et al., *Ribosome performance is enhanced by a rich cluster of pseudouridines in the A-site finger region of the large subunit.* The Journal of biological chemistry, 2008. **283**(38): p. 26026-26036.
111. Yassin, A. and A.S. Mankin, *Potential new antibiotic sites in the ribosome revealed by deleterious mutations in RNA of the large ribosomal subunit.* J Biol Chem, 2007. **282**(33): p. 24329-42.
112. Komoda, T., et al., *The A-site finger in 23 S rRNA acts as a functional attenuator for translocation.* The Journal of biological chemistry, 2006. **281**(43): p. 32303-32309.
113. Rakauskaite, R. and J.D. Dinman, *An arc of unpaired "hinge bases" facilitates information exchange among functional centers of the ribosome.* Molecular and cellular biology, 2006. **26**(23): p. 8992-9002.
114. Sergiev, P.V., et al., *The conserved A-site finger of the 23S rRNA: just one of the intersubunit bridges or a part of the allosteric communication pathway?* Journal of molecular biology, 2005. **353**(1): p. 116-123.
115. Liu, Q. and K. Fredrick, *Contribution of intersubunit bridges to the energy barrier of ribosomal translocation.* Nucleic acids research, 2013. **41**(1): p. 565-574.
116. Fernandez-Leiro, R. and S.H.W. Scheres, *A pipeline approach to single-particle processing in RELION.* Acta Crystallogr D Struct Biol, 2017. **73**(Pt 6): p. 496-502.
117. Lin, J., et al., *Conformational changes of elongation factor G on the ribosome during tRNA translocation.* Cell, 2015. **160**(1-2): p. 219-27.

118. Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.