Manuscript submitted to **Biophysical** *Journal*
Article

Advanced coarse-grained model for fast simulation of nascent polypeptide chain dynamics within the ribosome

Shiqi Yu[1], Artem Kushner[2], Ella Teasell[2], Wenjun Zhao[2], Simcha Srebnik[1,*], and Khanh Dao Duc[2,*]

[1]Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, BC V6T 1Z3, Canada
[2]Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z4, Canada
[*]Correspondence: ssrebnik@mail.ubc.ca, kdd@math.ubc.ca

ABSTRACT The nascent polypeptide exit tunnel (NPET) is a sub-compartment of the ribosome that constrains the dynamics of nascent polypeptide chains during protein translation. Simulating these dynamics has been limited due by the spatial scale of the ribosome and the temporal scale of elongation. Here, we present an automated pipeline to extract the geometry of the NPET and the ribosome surface at high resolution from any ribosome structure. We further convert this into a coarse-grained (CG) bead model that can be used in molecular simulations. This CG model more accurately captures NPET geometry than previous representations and allows for the simulation of co- and post-translational processes that are computationally prohibitive with all-atom approaches. In particular, we illustrate how the CG model may be used to simulate the elongation dynamics of the nascent polypeptide and its escape post-translation, as well as evaluate free energy landscapes and examine the influence of electrostatics on the nascent polypeptide escape.

SIGNIFICANCE The translation of nascent polypeptide chains is mediated by the ribosome, with interactions between the protein and the nascent polypeptide exit tunnel (NPET) impacting the process. However, modeling and simulating protein elongation and its escape from the NPET remain challenging due to computational limitations in spatial and temporal resolution. Here, we develop a computational pipeline for generating coarse-grained (CG) models of the NPET and ribosome surface from any ribosome structure that allow for both effective and accurate computer simulations. We demonstrate how the model can be implemented for various simulations, including the elongation dynamics of the nascent polypeptide and escape of the chain post-translation, as well as in estimating free energy landscapes and examining the impact of the charged environment on the escape time.

INTRODUCTION

The ribosome is a large RNA-protein complex that mediates protein synthesis from messenger RNA (mRNA) during translation. This process requires the formation of a peptide bond at the peptidyl-transferase center (PTC), which is deeply embedded within the large ribosomal subunit (1). As peptides get added, the nascent polypeptide chain passes through the nascent polypeptide exit tunnel (NPET). The NPET is a narrow conduit, approximately 100 Å length, that connects the PTC to the ribosome's solvent-exposed surface (2). Over the past decades, experimental efforts have focused on quantifying interactions between the nascent chain and ribosome and elucidating their impact on protein folding and translation arrest (3–6) to understand how NPET dynamically influences protein synthesis (7, 8). Given the tunnel's dimensions and the complexity of these interactions, molecular dynamics simulations have emerged as both a complement and alternative to experimental methods, providing atomic-level insights into nascent chain interactions at the NPET (9–12).

Simulations involving internal and external features of ribosomes are challenging due to the sheer size a complexity of the ribosome. Bacterial and eukaryotic ribosomes measuring about 20 nm and 30 nm, respectively (13, 14). As such, running even a 10-nanosecond ribosome simulation on a single GPU may take a full day, and microsecond-scale simulations may extend

over several months (9). Whole-ribosome simulations therefore remain exceedingly rare (9) and most molecular dynamics (MD) studies focus exclusively on the tunnel region (15–20). Yet, even these reduced models comprise tens of thousands of atoms (20) and create a substantial computational burden. Several other studies employ cylindrical approximations with varying radii to model the exit tunnel and examine protein escape dynamics (21–23). However, such studies risk oversimplifying some potentially important geometric features of the NPET (24). Moreover, most models fail to consider the ribosome's solvent-exposed surface, which is known to modulate folding and translation dynamics of the nascent polypeptides as they emerge from the NPET (4, 25, 26). Notably, intermolecular interactions between the ribosome and specific amino acid (aa) residues at the N-terminus are known to perturb co-translation folding (27) and regulate translation fidelity (28) of nascent chains, and hence should not be neglected in models aiming to capture the ribosome's role in modulating translation dynamics.

All-atom MD simulations have historically been limited to treating one or few species due to their computational expense. However, the recent surge in high-resolution ribosome structures obtained through cryo-electron microscopy (29) has revealed unprecedented heterogeneity across species and domains of life (30). Comparative studies have demonstrated how some important structural variations in the NPET of bacteria and eukaryotes significantly impact nascent chain dynamics (19, 24). This structural diversity amplifies the need of alleviating computational costs and facilitating efficient modeling of the NPET environment. This expanding structural dataset can be leveraged to comprehensively explore the functional role of the NPET across evolutionary lineages.

To address these computational limitations, we developed a pipeline to construct coarse-grained (CG) models of NPET and ribosome surface from any given ribosome structure that simplifies and reduces the complexity of simulating nascent polypeptide chain dynamics. In particular, when utilized with coarse-grained models of proteins (31), the model allows for simulating dynamics at the time scale of protein translation. Our approach converts empty space voxels from the ribosome 3D structure into a 3D mesh of the NPET, that is subsequently used to create an efficient CG model for studying translation-related processes. We benchmark our model against established tunnel representations, evaluating its geometric accuracy and computational efficiency in simulating nascent polypeptide dynamics. We then demonstrate the versatility of the model through various applications, including simulating co-translational and post-translational processes, as well as comparison of energy profiles, and evaluation of escape kinetics of charged polypeptide chains, showing its potential as a tool for future investigations of nascent polypeptide chain interactions and dynamics across diverse ribosome structures.

METHODS

Modeling the NPET interior and ribosome exterior surfaces

To build a CG model of the NPET and ribosome, we first extract the surface of the ribosome and the interior surface of the NPET as regular meshes. For any arbitrary atomic model of the ribosome (obtained from the PDB or other database (29) as a `.cif` file), we run the following protocols, illustrated in Figure S1:

Exterior ribosome surface

Briefly, we obtain the atomic model of the ribosome deposited in the PDB and parse its (32) `.cif` file to extract 3D coordinates of atoms. This constitutes the initial pointcloud. A pointset is extracted from that cloud that represents only the surface atoms via Delaunay triangulation (33). Further, orientation vectors known as normals are estimated (34) on this surface pointset to establish what constitutes *the interior* and *the exterior* in space relative to the surface pointset. Poisson surface reconstruction (35) is applied to the normal-estimated pointset which yields a smooth mesh of the ribosome surface. Examples of corresponding structures generated during the stages of the modeling of the exterior ribosome surface are shown in Figure S1a.

Interior NPET surface

To model the NPET, we run the following six steps (with visualization shown in Figure S1b.

1. The locations of the PTC and the *Constriction Site* are retrieved from the API of the `riboxyz` project: `api.ribosome.xyz` (for a precise definition of how their coordinates are calculated as well as, see SI file)

2. The pointcloud of the given ribosome structure is obtained from the corresponding PDB file as described in the previous section (*Exterior ribosome surface*). Any fragments that might be occupying the tunnel (*e.g.* nascent peptide chains and ligands) are removed as needed.

3. Using the PTC and the constriction site coordinates as the main axis, we construct a cylindrical mask with radius of 35 Å and height of 120 Å with the PTC at the base. These dimensions are heuristic but conservative enough to contain

any ribosome tunnel across the tree of life. This cylindrical mask is mounted on the axis defined by the PTC and the *Constriction Site* and applied to the ribosome pointcloud.

4. Atoms contained in the mask are projected onto a binary voxel grid of the same dimensions as the cylindrical mask. The positions of atoms which belong to the mask are used to activate (set value to 1) voxels whose centers fall within a sphere of radius 2 Å from the given atom positions. This 2Å sphere is a simplistic representation of the van der Waals radius of an idealized atom. The voxel grid is clipped with the surface of the whole ribosome (produced above in *Exterior ribosome surface*) to discard the inactive voxels (representing empty space past the tunnel exit port).

5. At this point the cylindrical voxel grid contains only values 0 and 1, where inactive (0-value) voxels correspond to a mixture of the NPET interior space and inter-polymeric space pockets and active (1-value) voxels represent space occupied by atoms. We use DBSCAN (36) density-based clustering algorithm to distinguish the central and biggest cluster corresponding to the NPET interior from the other pockets. We run DBSCAN two times, first on the entire grid (with default parameters $\varepsilon = 5.5$ Å and $min_samples = 600$) to detect major clusters, and second on the central cluster ($\varepsilon = 3$ Å and $min_samples = 120$) to sharpen the cluster revealing atomic-scale features. For a structural interpretation of DBSCAN parameters, see SI File.

6. As in the previous section, we estimate the normals on the surface of the NPET interior voxel cloud. Poisson Surface reconstruction (35) is likewise applied to this final voxelcloud to produce a smooth watertight mesh.

Coarse-grained representation of the ribosome tunnel and surface

The aforementioned mesh structure serves as the initial scaffold for the CG model, with CG beads placed at the centroid of the triangular mesh. The beads representing the NPET are placed along the surface normal at a distance corresponding to the radius of the bead to properly capture the cavity. For the ribosome surface, the CG beads are placed directly on the mesh. If necessary, additional CG beads can be added or removed on the mesh vertices or edges. The centroids and normal vectors of the triangle meshes are computed using the PyMeshLab library in Python (37). LAMMPS "delete_atoms" command (38) is used to delete atoms to ensure an open NPET exit port and reduce the computational costs by deleting overlapping CG beads.

Mapping effective charges onto the CG beads

The effective potential at the center of each CG bead at the NPET, j, is calculated using the screened potential, ϕ_j, summed over all charged residues (lysine, arginine, glutamate, and aspartate) and nucleic acids within a certain distance (*e.g.*, 50 Å) of the tunnel centerline, as

$$\phi_j = \sum_{i=1}^{n} \frac{q_i}{4\pi\,\varepsilon_0\,\varepsilon_r} \frac{e^{-\kappa\, r_{ij}}}{r_{ij}} \tag{1}$$

where r_{ij} is the distance between the charged atom, i, and the CG bead, j. q_i is the atom charge, ε_0 is the vacuum permittivity, ε_r is the dielectric constant, and κ is the inverse Debye length. The coordinates of the charged residues are extracted using PyMOL and, for simplicity, assigned to their respective C_α atoms (39). The positions of negatively charged rRNA atoms are placed on the backbone phosphorus atoms. The value of κ is taken to be 0.14 Å^{-1}. The dielectric constant, ε_r, can be set to unity for explicit solvent simulations, can be used to best fit the experimental data (3), or can be estimated in the range of 4 to 80 (40). The effective charge of each CG bead of NPET is then evaluated from $q_{j,eff} \propto \phi_j R$, where R is the bead radius, set as 1.8 Å (41). The proportionality constant is chosen to align the induced potential along the tunnel centerline of the CG model with that from the atomistic model using least squares optimization.

While the effective charge of each CG bead of the ribosome surface can be similarly calculated as above, we alternatively propose to calculate the effective electrostatic potential for implicit solvent simulations using Adaptive Poisson-Boltzmann Solver (APBS) in PyMOL (42, 43). This approach accounts for the distinct solvent environment outside the ribosome compared to the tunnel interior. The calculations of the linearized Poisson–Boltzmann equation (PBE) (43) are performed at 0.15 M ionic strength in monovalent salt, 310K, solute dielectric constant of 4 and a solvent dielectric constant of 80 (44). The grid spacing is set to 1 Å. All other parameters of the APBS input file are kept at the default settings. The effective potentials at the centers of the CG beads are then obtained through three-dimensional cubic interpolation of the grid-based data from APBS.

Dynamic simulations of nascent chains

A simplified CG protein model is used in this study to simulate nascent chain dynamics to illustrate the CG tunnel and ribosome model. Each residue is modeled by a triplet of beads representing the backbone units, where the central bead of each triplet

is designated as either neutral, positively charged, or negatively charged. Lennard-Jones (LJ) potential is used to model the short-range intermolecular interactions between non-bonded residues (45). Intramolecular interactions are modeled with harmonic bond and angle potentials, and the dihedral force between neighboring residues. The screened Coulomb potential is used for electrostatic interactions.

To simulate post-translational dynamics within the NPET, CG beads of the nascent chain are initially aligned along the tunnel centerline with the C-terminus positioned at the PTC. The attached polypeptide chain is first equilibrated as the average internal energy remains constant for 50,000 timesteps. The C-terminus is then released. Co-translational dynamics are modeled by simulating the intermittent addition of beads at the PTC at a constant rate. The chain is advanced by a distance equal to half the bond length towards the exit to introduce new beads. During each interval between bead additions, the nascent chain is attached at the PTC. This process is repeated iteratively until the desired chain length is reached. For details on the MD protocol and force field parameters, see SI file.

Datasets

The structures used to construct the mesh of the NPET and ribosome surface in this paper are high-resolution atomic models of the ribosome from *H.sapiens* and *E.coli*, deposited in the Protein Data Bank (PDB) with codenames 4UG0 and 6WD4 respectively.

Software and implementation

Our pipeline runs in Python, using a combination of packages including `numpy` (46), `pyvista` (33), `scipy` (47), with the code and illustrative notebooks publicly accessible on GitHub. MD simulations were conducted using LAMMPS (released version: 29 Aug 2024) (48) on a single CPU compute node with a consumer-line PC with 1.4 GHz Quad-Core Intel Core i5 and 8GB RAM memory. The coordinate extractions of the tunnel centerline, charged residues and atoms were performed in PyMOL (49). Visualizations of the coarse-grained models and nascent chain dynamics are carried out with VMD (50), and OVITO (51). Cross-sectional views of the ribosome models are generated using ChimeraX (52). All downstream analyses were conducted using custom Python scripts.

RESULTS

Coarse grained modeling of the ribosome surface and NPET

We developed a computational pipeline to generate coarse-grained models of the NPET and ribosome surface from any given structure, as shown in Figure 1a and detailed in the Methods section for a more detailed visualization, see also Supplementary Figure S1. Beginning with an atomistic model of the ribosome, the exit tunnel is first identified based on landmarks that define the PTC and constriction site. Atoms within a cylindrical volume surrounding the tunnel are used to extract the tunnel surface as a mesh through DBSCAN clustering and Poisson surface reconstruction (Figure 1b). CG beads are then placed on the mesh to form a CG representation of the tunnel surface. The ribosome surface is similarly captured (Figure 1c). The CG assembly, integrating both the tunnel and ribosome surface can be seen in 3D in Supplementary Video S1.

Modeling effective charges

As an RNA-protein complex enriched in charged residues, the ribosome will have significant electrostatic interactions with charged polypeptide chains (16, 53). To include these interactions in our CG model, we mapped the effective charges for both the ribosome tunnel and surface onto the CG beads. For the beads at the NPET, we used the ribosomal charged residues and nucleic acids within a certain distance of the NPET, and calculated the associated screened potential at each bead (see Eq. (1)), as described in the Methods section. For the ribosome surface, we used the APBS solver in Pymol (see Methods section) to account for the solvent in our calculation of the effective potential. Figure 2a illustrates the charge distribution of the CG model for the 4UG0 ribosome and NPET (see Figure S2a for 6WD4), with both showing predominantly negative charge. Scattered positively charged regions on the ribosome surface are primarily associated with ribosomal proteins (43) as has been observed in cryo-EM ribosome structures.

To evaluate the accuracy of our mapping, we compared the electrostatic potential profiles along the tunnel centerline generated by our CG model with experimental measurement measured at specific positions of the tunnel for an archaeal ribosome (3) (Figure 2b), as well as the electrostatic profile obtained from atomistic simulation of the bacterial ribosome (4V5D) (44) (Figure S2b). The results indicate that the profile from the CG model aligns closely with empirical measurements (see Figure 2b), and shows the same trend as for the atom-based model. We also noted some more significant deviation between

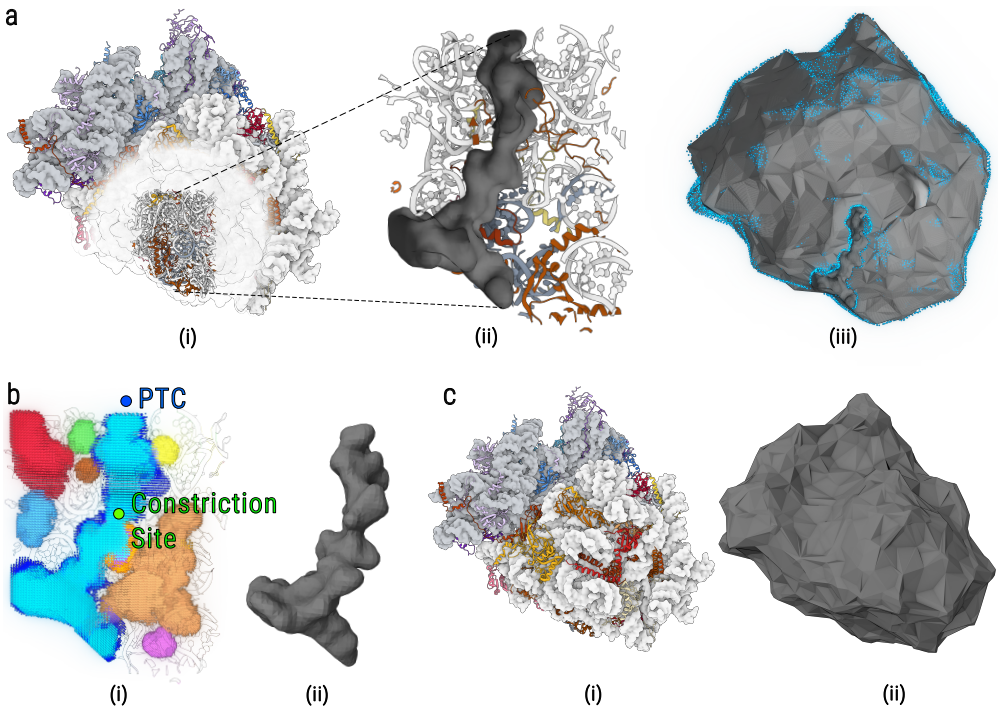

Figure 1: Construction of CG simulation environment illustrated with ribosome structure 4UG0. **(a)** General illustration showing the *(i)* ribosome structure, *(ii)* NPET and *(iii)* ribosome surface. For a better visualization of the tunnel position and geometry, only selected representative beads of the CG model are shown as blue spheres in *(iii)*. **(b)** After *(i)* clustering empty voxels, we *(ii)* build the tunnel mesh. **(c)** Mesh produced from the *(i)* ribosome structure using *(ii)* Poisson surface reconstruction.

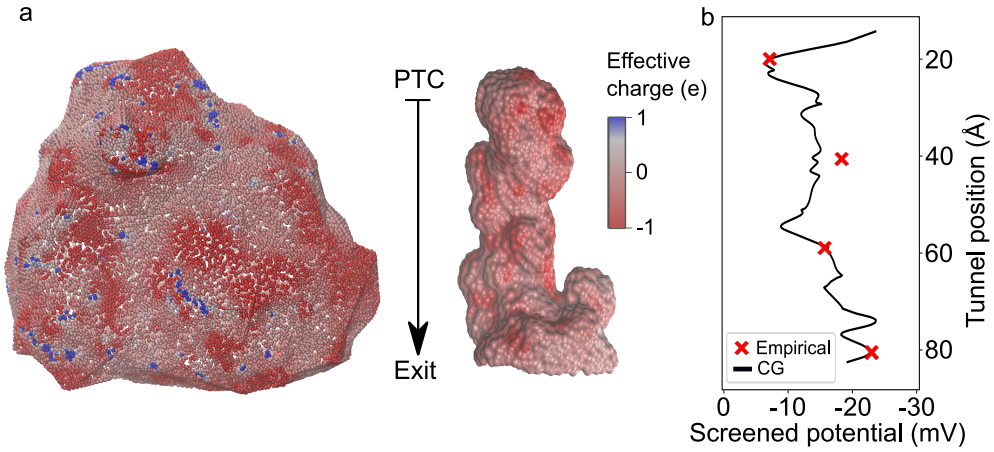

Figure 2: Coarse-grained models of the ribosome surface and tunnel with mapping of effective charges. **(a)** Visualization of effective charges mapped onto the ribosome surface and the exit tunnel of 4UG0. Blue and red beads represent positively and negatively charged beads, respectively. The color scale range is adjusted to improve visualization clarity. **(b)** Electrostatic potential along the tunnel centerline in the archaeal ribosome 1JJ2 derived from the CG model (black curve) compared with experimental data (3) (red crosses).

the CG and atom-based profiles at the exit port (Figure S2b), that can be attributed to the inclusion of magnesium ions and TIP3P waters in the atom-based model (44).

Comparison with other models

The extent to which the CG model more accurately captures the NPET geometry was evaluated by comparing it to the model obtained by running MOLE, a cavity search algorithm (54, 55) that was used in previous recent geometric studies of the NPET (24, 56). As this algorithm yields a set of spheres lining along a centerline from the PTC to the exit port, it likely fails to capture irregularities along cross sections of the NPET. Taking ribosomes 4UG0 (*H.sapiens*) and 6WD4 (*E.coli*) as examples from different life domains, we used MOLE2.5 (54, 55) to extract the NPET model and compared the volume produced with our method. To do so, we also divided the tunnel into three regions shown in Figure 3a, namely the upper, lower and exit regions. These regions are separated by the constriction site (see Methods), and conserved residues K68 in uL23 (bacteria), or R28 in eL39 (eukaryotes) (24, 57). As shown in Figure 3a for 4UG0, the CG model encompasses a volume that is significantly larger, with 3.2×10^4 Å^3 (2.9×10^4 Å^3 for 6WD4, see Figure S3a), compared to 7.3×10^3 Å^3 (9.5×10^3 Å^3 for 6WD4) for the spherical model. While the upper and lower regions in the CG model are 53% and 82% larger (45% and 71% for 6WD4), respectively, we observed that most of the volume discrepancy concentrates in the exit region, where the difference yields an order of magnitude difference, with 1.6×10^4 Å^3 for the CG model compared to 1.5×10^3 Å^3 for the spherical (1.2×10^4 Å^3 and 2.7×10^3 Å^3 for 6WD4). The smaller exit volume observed in the spherical model results from the probe-based algorithm used for extracting cavity terminating earlier, as it could otherwise generate artificial paths (54). Overall, the CG model provides a more comprehensive representation of the NPET's internal space, particularly at the exit port and interface with the external surface. A visualization comparing the volumes of the two tunnel models is provided in the Supplementary Video S2.

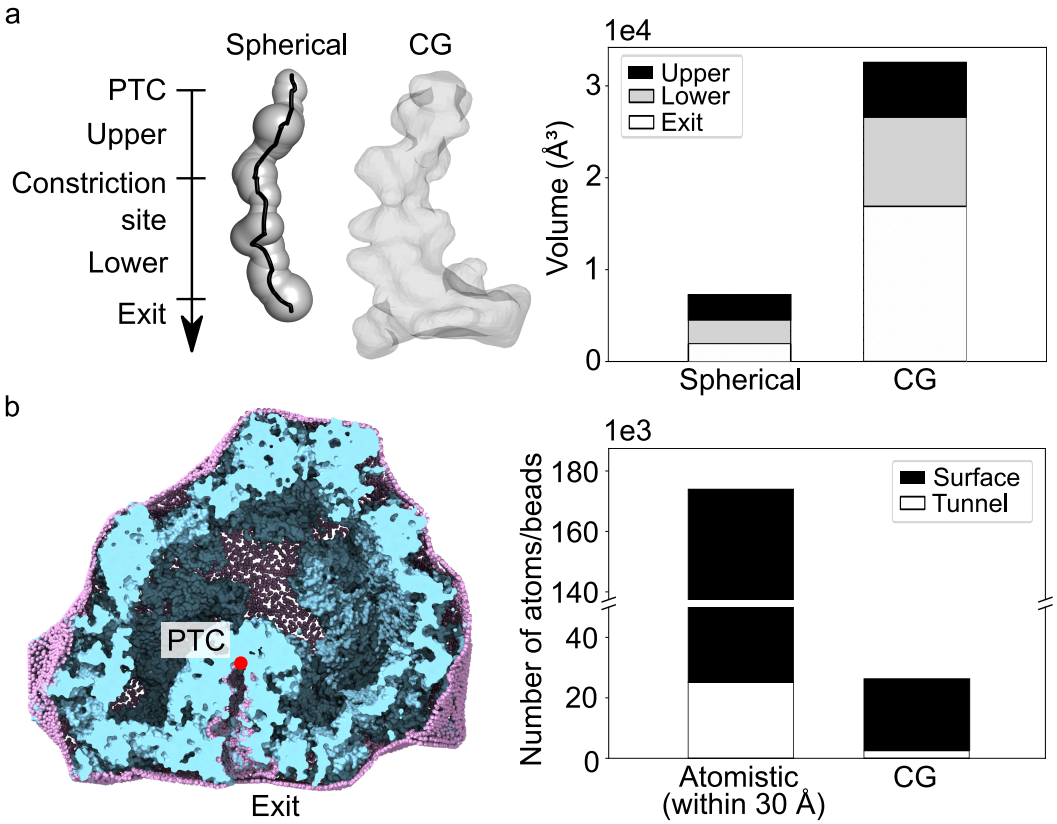

Figure 3: Comparisons of the CG model with the spherical and all-atom model for 4UG0. **(a)** The left panel shows the MOLE-generated spherical tunnel model (left) and the CG tunnel model in mesh (right). The tunnel is divided into three regions. The upper, lower and exit region are separated by the constriction site and conserved residues in uL23 or eL39. The right panel compares the volume of the spherical and CG tunnel models in the upper, lower and exit regions. **(b)** Perspective view of a cross-section of the ribosome, showing the atomistic model including all atoms within 30 Å of the ribosome surface (cyan), and the CG ribosome model (purple). Lighter shades of blue indicate atoms closer to the viewer, while darker shades represent atoms farther away. The PTC is highlighted as a red dot. The right panel shows the comparison of the number of beads or atoms in the tunnel (white) and ribosome surface (black) regions for both models. In the atomistic model, overlapping atoms between the two regions are excluded in the atom count for the surface region.

To evaluate the potential benefit of the CG model in MD simulations, we further quantified the number of atoms that is typically required to simulate the ribosome environment in an atomistic representation. For 4UG0 we extracted all atoms located within 30 Å of the ribosome surface and the tunnel centerline (see Figure 3b). In the NPET region, the number of atoms required for 4UG0 is approximately 25,000 atoms (24,000 atoms for 6WD4). While our CG tunnel model contains 2,500 beads, that were set with a radius of 1.8 Å to be watertight (2,300 beads for 6WD4). When considering an implicit solvent CG protein model like that used in our following simulations, this number can be further reduced to ~ 1,600. For the ribosome surface region, our CG model includes approximately 24,000 beads – about 80.5% reduction compared to the atomistic representation, which contains ~ 123,000 atoms (22,000 and 95,000 atoms for 6WD4, around 76.8% reduction, see Figure S3b). This substantial decrease in system size significantly enhances computational efficiency and makes the model well suited for studying long-timescale processes, such as nascent chain elongation.

Simulations of nascent chain dynamics

In this section, we demonstrate two applications of our CG model for MD simulations using a simplified nascent chain model. The nascent chain is represented using a bead-spring model where each residue comprises a triplet of beads representing backbone units. To isolate geometric and entropic effects, simulations were performed in a neutral CG ribosome environment. Detailed parameters are provided in Methods and SI file. Dynamic trajectory visualizations are provided in Supplementary Videos S3 and S4.

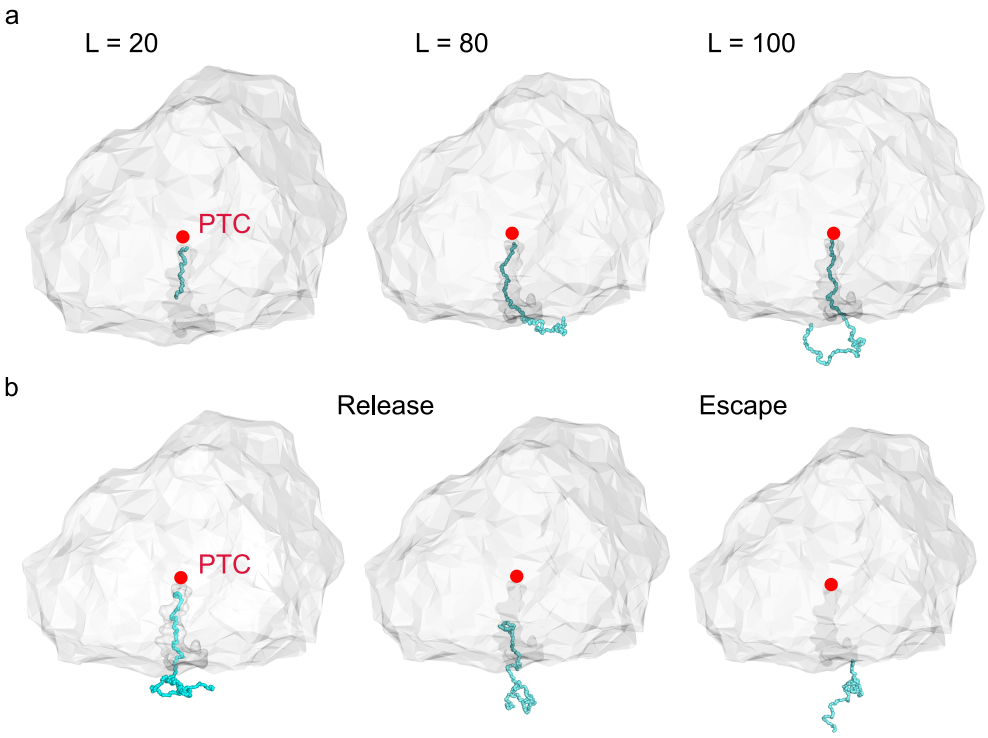

Figure 4: Dynamic simulations of polypeptides within the coarse-grained ribosome model of 4UG0. **(a)** Co-translational process of polypeptide elongating to 20, 80, and 100 amino acid residues. The ribosome surface is modeled to simulate interactions with the protein segment extending beyond the tunnel. Length L represents the number of amino acid residues elongated. **(b)** Post-translational dynamics of a 100-mer polypeptide, illustrating its release from the PTC to its full exit from the tunnel (left to right). The PTC is highlighted as a red dot.

We first simulated the co-translational dynamics (elongation) of the polypeptide in the CG model of 4UG0 encompassing both the NPET and ribosome surface. To do so, we added each residue to the growing chain at the PTC (see Methods section) at 1×10^4 time step intervals. Considering a typical elongation rate of ~ 10 codons per second measured in yeast (58), this corresponds to a time step of 10 μs – a timescale characteristic of molecular CG simulations (31). Figure 4a shows snapshots of the elongation simulations. The 20-mer chain is is seen to be contained within the NPET, which is consistent with previous findings that ribosomes can accommodate up to 50 residues internally (59, 60). As elongation progressed to 80 and 100 residues,

we see the N-terminal portion interacting with the ribosomal surface before extending into the solvent environment where folding might initiate.

Next, we simulated the post-translational escape process of polypeptides from the CG ribosome model of 4UG0, and Figure 4b illustrates the escape snapshots of a 100-mer polypeptide. The escape time of nascent chains can greatly vary within a 1000-fold range (16), with short polypeptides that fully accommodated within the tunnel before release exhibiting particularly extended times compared to other chain lengths (see next section for a more detailed explanation). For example, the simulation of the 20-mer polypeptide completed in approximately 25 minutes, from the equilibrated conformation of the peptide chain with its C-terminus attached to the PTC, until complete escape. While for the majority of the simulation, the chain remains confined within the upper tunnel region (see Figure S4 and next section), suggesting that the translocation of short polypeptide within the tunnel can be strongly hindered by the constriction sites (21). Table 1 summarizes the simulation time steps and corresponding CPU runtimes for both co- and post- translational processes. The simulation of a 500-mer chain elongation completed in approximately 4 hours, demonstrating that our CG model enables long-timescale protein translation simulations to complete within hours, even on a single CPU core.

Table 1: Simulation and CPU times for co- and post-translational processes for different lengths of polypeptides in the coarse-grained ribosome model of 4UG0.

Simulations	**Time step [A.U.]**	**CPU time** [s]
Elongation (50-mer)	1.6×10^6	1,206
Elongation (100-mer)	3.2×10^6	2,282
Elongation (500-mer)	1.6×10^7	15, 040
Post-translation (tunnel escape)	10, 000	2.71

Comparison of energy profiles

Focusing on the post-translation escape process illustrated in our second example, we study the nascent chain dynamics and free energy landscape the protein experiences in the NPET. Previously, we employed a simplified cylindrical model (21) to analyze nascent chain free energy profiles within the tunnel. While this cylindrical model captures the average geometric properties of the NPET across 20 eukaryotic species, significant variations can be observed across species (24), potentially leading to differences in energy profile. Like in our previous study, we considered here 12-mer and 20-mer polypeptides in the CG tunnel model of 4UG0, and calculated their free energy profiles along the tunnel axis. Free energy (f) landscapes for each polypeptide length were derived from simulations, by summing over the canonical partition function along the tunnel, with $f(r, N, T) = -kT/N \ln \sum e^{-E(r,N)/kT}$ (see (21)), where the sum is over all configuration with C-terminus at position r along the tunnel, k is the Boltzmann constant, N is the number of beads in the polypeptide, and E is the energy of each microstate. Figure 5a compares the energy profiles between the previous simplified cylindrical geometry and the current CG model. While both approaches reveal two distinct low-energy states, the energy wells in the CG model are notably shallower with lower energy barriers compared to the cylindrical representation. This difference highlights how a more accurate CG model provides more nuanced insights into the energetic landscape governing nascent chain transit through the ribosomal tunnel, with differences across species that could not be captured using a simplified geometry.

For the 20-mer polypeptide (Figure 5b), we observed more significant qualitative differences between the two models. The cylindrical model only displayed a single energy well at the constriction site, whereas the CG model revealed a two-potential well profile again, with the deeper well located in the upper tunnel region. We attribute this discrepancy to volume underestimation in the cylindrical model's upper tunnel region (as evidenced in Figure 3a), making it insufficient to accommodate the polymer chain. In contrast, the CG model demonstrates that the upper tunnel region can effectively trap both 12-mer and 20-mer polypeptides—a finding consistent with previous observations that short proteins of these lengths can initiate folding near the PTC (19, 61). While both models support the general conclusion that tunnel geometry impedes 20-mer polypeptide escape, it is interesting to note that their biophysical interpretations differ for 4UG0, highlighting again the importance of using a more

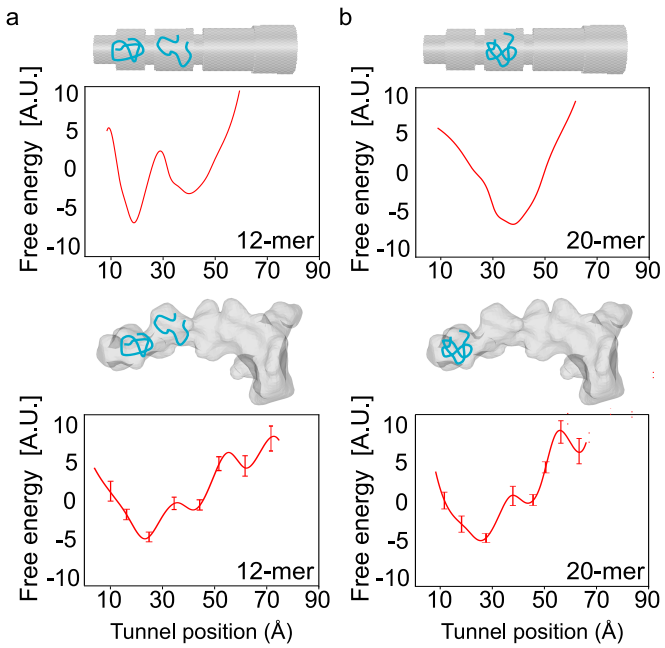

Figure 5: Comparison of free energy landscapes of **(a)** 12-mer and **(b)** 20-mer polypeptide in the cylindrical (upper) (21) and coarse-grained model (lower) of the neutral eukaryotic tunnel 4UG0. The energy is calculated from the sum over the partition function at each position along the tunnel using the C-terminus position. The error bars are calculated based on one hundred independent runs. A Savitzky-Golay filter is applied to reduce the noise.

detailed representation of the NPET.

Escape of charged polypeptide

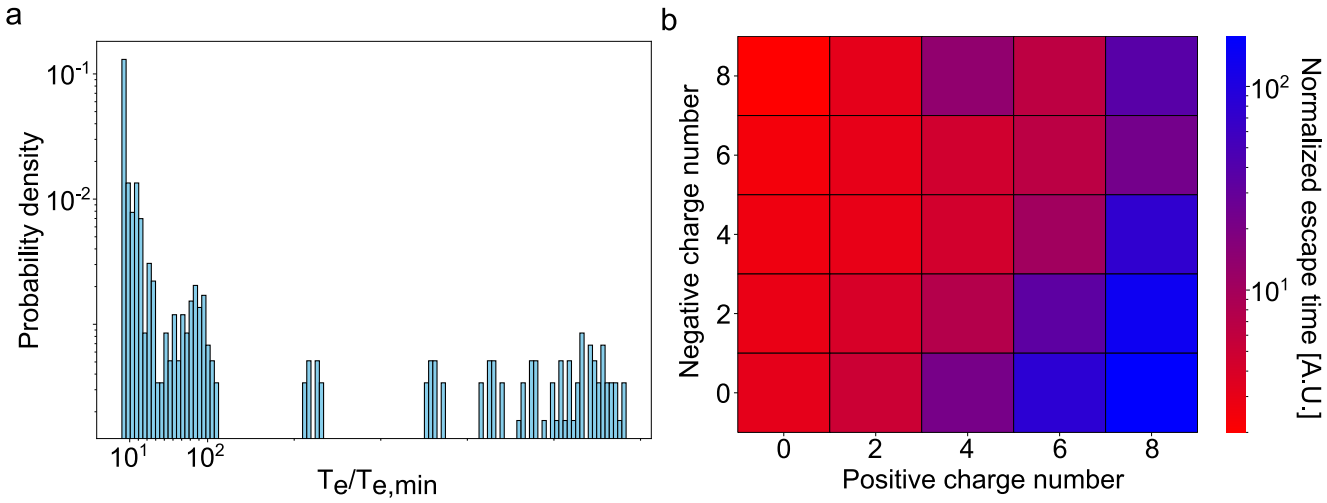

Figure 6: Variations in escape times influenced by charged residues in the C-termini of polypeptides. **(a)** Distribution of normalized escape times for 121 80-mer polypeptide chains simulated using the charged CG ribosome model of 6WD4. Each bar represents the probability density of $T_e/T_{e,\min}$, where T_e is the escape time for each simulation run, and $T_{e,\min}$ is the shortest escape time among all simulations. **(b)** Normalized mean escape times for 25 charge patterns mapped in a heat map. Each cell, except for the neutral protein, represents the average escape time of five charged sequences with randomly distributed charges within the C-terminal 30 residues, normalized by $T_{e,\min}$. Faster and slower escapes are indicated in red and blue, respectively.

We finally illustrate how the CG ribosome model can be used to examine the effects of electrostatics on nascent chain escape dynamics, as previously done in *E.coli* (16). To do so, we constructed 80-mer polypeptides with 25 distinct charge patterns, each consisting of a different number of positively and negatively charged residues. For each non-neutral pattern, five sequences were generated by randomly distributing the charges within the C-terminal 30 residues, resulting in 120 unique charged sequences, along with one neutral sequence. Each sequence was simulated in 10 independent runs. Figure 6a shows the distribution of escape times, T_e, from all simulations, normalized by the shortest escape time, $T_{e,min}$. The distribution spans a wide range of escape times with a peak around 10, similar to the findings of Nissley *et al.* (16). We also note a long tail associated with slow escape events. In Figure 6b, we plot the normalized mean escape times for each charge pattern, revealing a clear charge-dependent bias, with the slow escape times being associated with more positive charges. Polypeptides enriched in positive charges exhibited slower escape, consistent with attractive electrostatic interactions with the negatively charged tunnel surface. Conversely, chains with more negative charges experienced stronger electrostatic repulsion, leading to faster escape. The neutral polypeptide showed an average escape time about 50% slower than the fastest escape and 99% faster than the slowest, highlighting the critical role of electrostatic interactions in modulating escape dynamics (16). Additionally, ribosome profiling analyses revealed that proteins with slower escape exhibited higher ribosome density at the stop codon compared to faster ejectors (16). While our results show consistency with both all-atom simulations and experimental measurements, the potential for slow escape times demonstrates the importance of enhancing computationally efficient simulations to explore rare events.

DISCUSSION

In this paper, we present a pipeline for constructing coarse-grained models of the NPET and ribosome surface for MD simulations of nascent polypeptide chains across various contexts and extended timescales. While traditional methods model the NPET interior using probe-based cavity extraction (55, 62–64), our approach quantizes the volume space into a grid to identify empty voxels. This space-partitioning strategy has recently proven effective for generating highly detailed volumetric mesh representations of biomolecules (65) with a highly parallelizable algorithm that leverages octree data structure. Herein, we leverage KDtree structures due to their available Python implementations for voxel querying.

We also note that while our simulations were conducted on unoptimized CPU environments, substantial computational efficiency improvements can be expected through GPU or accelerator packages (38). We employed a simplified CG protein model with Lennard-Jones interactions to illustrate the nascent chain dynamics in our CG environment. While this model is not designed to capture realistic folding, more sophisticated models using established force fields (66) or all-atom representations (67) can explicitly account for side-chain interactions and dynamic behaviors, thereby providing more detailed and accurate representations of translation and folding dynamics.

One significant limitation of our CG model is its inability to capture potentially flexible elements of the NPET and ribosome (such as the lateral stalk (68)) or conformational changes occurring during different steps of translation (69, 70). Interestingly, Kurkcuoglu *et al.* employed elastic network modeling to investigate ribosomal flexibility (71), identifying three distinct tunnel regions with characteristic harmonic motions: the upper region primarily moving parallel to the tunnel axis, the exit region displaying pronounced perpendicular mobility (potentially facilitating nascent peptide chain dispersion), and the constriction site exhibiting rotational motions that may contribute to polypeptide gating mechanisms. Future work could integrate these insights through a hybrid modeling approach where functionally critical flexible regions—particularly at the constriction site—are represented at atomic resolution. This would maintain computational efficiency while incorporating essential ribosomal dynamics, providing a more complete representation of the translation environment and its impact on nascent chain.

AUTHOR CONTRIBUTIONS

KDD and SS designed the research. AK, SY, ET and WZ implemented the methods. SY carried out all simulations and analyzed the data. SY, AK, KDD and SS wrote the article.

ACKNOWLEDGMENTS

This research is supported by NSERC Discovery Grants RGPIN-2020-05348 and RGPIN-2024-04666.

REFERENCES

1. Rodnina, M. V., M. Beringer, and W. Wintermeyer, 2007. How ribosomes make peptide bonds. *Trends in biochemical sciences* 32:20–26.

2. Thommen, M., W. Holtkamp, and M. V. Rodnina, 2017. Co-translational protein folding: progress and methods. *Current opinion in structural biology* 42:83–89.

3. Lu, J., W. R. Kobertz, and C. Deutsch, 2007. Mapping the electrostatic potential within the ribosomal exit tunnel. *Journal of molecular biology* 371:1378–1391.

4. Masse, M. M., V. Guzman-Luna, A. E. Varela, U. Mahfuza Shapla, R. B. Hutchinson, A. Srivastava, W. Wei, A. M. Fuchs, and S. Cavagnero, 2024. Nascent chains derived from a foldable protein sequence interact with specific ribosomal surface sites near the exit tunnel. *Scientific Reports* 14:12324.

5. Wruck, F., P. Tian, R. Kudva, R. B. Best, G. von Heijne, S. J. Tans, and A. Katranidis, 2021. The ribosome modulates folding inside the ribosomal exit tunnel. *Communications biology* 4:523.

6. Bustamante, C., L. Alexander, K. Maciuba, and C. M. Kaiser, 2020. Single-molecule studies of protein folding with optical tweezers. *Annual review of biochemistry* 89:443–470.

7. Wilson, D. N., and R. Beckmann, 2011. The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. *Current opinion in structural biology* 21:274–282.

8. Berisio, R., F. Schluenzen, J. Harms, A. Bashan, T. Auerbach, D. Baram, and A. Yonath, 2003. Structural insight into the role of the ribosomal tunnel in cellular regulation. *Nature Structural & Molecular Biology* 10:366–370.

9. Bock, L. V., S. Gabrielli, M. H. Kolář, and H. Grubmüller, 2023. Simulation of complex biomolecular systems: the ribosome challenge. *Annual Review of Biophysics* 52:361–390.

10. Huter, P., S. Arenz, L. V. Bock, M. Graf, J. O. Frister, A. Heuer, L. Peil, A. L. Starosta, I. Wohlgemuth, F. Peske, et al., 2017. Structural basis for polyproline-mediated ribosome stalling and rescue by the translation elongation factor EF-P. *Molecular cell* 68:515–527.

11. Zimmer, M. H., M. J. Niesen, and T. F. Miller, 2021. Force transduction creates long-ranged coupling in ribosomes stalled by arrest peptides. *Biophysical Journal* 120:2425–2435.

12. Lind, C., and J. Åqvist, 2016. Principles of start codon recognition in eukaryotic translation initiation. *Nucleic Acids Research* 44:8425–8432.

13. Makarov, G., T. Makarova, N. Sumbatyan, and A. Bogdanov, 2016. Investigation of ribosomes using molecular dynamics simulation methods. *Biochemistry (Moscow)* 81:1579–1588.

14. Bock, L. V., M. H. Kolář, and H. Grubmüller, 2018. Molecular simulations of the ribosome and associated translation factors. *Current opinion in structural biology* 49:27–35.

15. Bui, P. T., and T. X. Hoang, 2020. Protein escape at the ribosomal exit tunnel: Effect of the tunnel shape. *The Journal of Chemical Physics* 153.

16. Nissley, D. A., Q. V. Vu, F. Trovato, N. Ahmed, Y. Jiang, M. S. Li, and E. P. O'Brien, 2020. Electrostatic interactions govern extreme nascent protein ejection times from ribosomes and can delay ribosome recycling. *Journal of the American Chemical Society* 142:6103–6110.

17. Contreras Martínez, L. M., F. J. Martínez-Veracoechea, P. Pohkarel, A. D. Stroock, F. A. Escobedo, and M. P. DeLisa, 2006. Protein translocation through a tunnel induces changes in folding kinetics: a lattice model study. *Biotechnology and bioengineering* 94:105–117.

18. Kirmizialtin, S., V. Ganesan, and D. E. Makarov, 2004. Translocation of a β-hairpin-forming peptide through a cylindrical tunnel. *The Journal of chemical physics* 121:10268–10277.

19. Chwastyk, M., and M. Cieplak, 2021. Nascent folding of proteins across the three domains of life. *Frontiers in Molecular Biosciences* 8:692230.

20. Petrone, P. M., C. D. Snow, D. Lucent, and V. S. Pande, 2008. Side-chain recognition and gating in the ribosome exit tunnel. *Proceedings of the National Academy of Sciences* 105:16549–16554.

21. Yu, S., S. Srebnik, and K. D. Duc, 2023. Geometric differences in the ribosome exit tunnel impact the escape of small nascent proteins. *Biophysical Journal* 122:20–29.

22. Bui, P. T., and T. X. Hoang, 2018. Protein escape at the ribosomal exit tunnel: Effects of native interactions, tunnel length, and macromolecular crowding. *The Journal of Chemical Physics* 149.

23. Bui, P. T., and T. X. Hoang, 2016. Folding and escape of nascent proteins at ribosomal exit tunnel. *The Journal of chemical physics* 144.

24. Dao Duc, K., S. S. Batra, N. Bhattacharya, J. H. Cate, and Y. S. Song, 2019. Differences in the path to exit the ribosome across the three domains of life. *Nucleic acids research* 47:4198–4210.

25. Knight, A. M., P. H. Culviner, N. Kurt-Yilmaz, T. Zou, S. B. Ozkan, and S. Cavagnero, 2013. Electrostatic effect of the ribosomal surface on nascent polypeptide dynamics. *ACS chemical biology* 8:1195–1204.

26. Waudby, C. A., C. Burridge, L. D. Cabrita, and J. Christodoulou, 2022. Thermodynamics of co-translational folding and ribosome-nascent chain interactions. *Current Opinion in Structural Biology* 74:102357.

27. Cassaignau, A. M., T. Włodarski, S. H. Chan, L. F. Woodburn, I. V. Bukvin, J. O. Streit, L. D. Cabrita, C. A. Waudby, and J. Christodoulou, 2021. Interactions between nascent proteins and the ribosome surface inhibit co-translational folding. *Nature Chemistry* 13:1214–1220.

28. Fujii, K., T. T. Susanto, S. Saurabh, and M. Barna, 2018. Decoding the function of expansion segments in ribosomes. *Molecular cell* 72:1013–1020.

29. Kushner, A., A. S. Petrov, and K. Dao Duc, 2023. RiboXYZ: a comprehensive database for visualizing and analyzing ribosome structures. *Nucleic Acids Research* 51:D509–D516.

30. Poitevin, F., A. Kushner, X. Li, and K. Dao Duc, 2020. Structural heterogeneities of the ribosome: new frontiers and opportunities for cryo-EM. *Molecules* 25:4262.

31. Kmiecik, S., D. Gront, M. Kolinski, L. Wieteska, A. E. Dawid, and A. Kolinski, 2016. Coarse-grained protein models and their applications. *Chemical reviews* 116:7898–7936.

32. Cock, P. J., T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski, et al., 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25:1422.

33. Sullivan, C., and A. Kaszynski, 2019. PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK). *Journal of Open Source Software* 4:1450.

34. Zhou, Q.-Y., J. Park, and V. Koltun, 2018. Open3D: A modern library for 3D data processing. *arXiv preprint arXiv:1801.09847* .

35. Kazhdan, M., M. Bolitho, and H. Hoppe, 2006. Poisson surface reconstruction. *In* Proceedings of the fourth Eurographics symposium on Geometry processing. volume 7.

36. Deng, D., 2020. DBSCAN clustering algorithm based on density. *In* 2020 7th international forum on electrical engineering and automation (IFEEA). IEEE, 949–953.

37. Muntoni, A., and P. Cignoni, 2021. PyMeshLab.

38. Thompson, A. P., H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. In't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, et al., 2022. LAMMPS-a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications* 271:108171.

39. Bui, P. T., and T. X. Hoang, 2021. Hydrophobic and electrostatic interactions modulate protein escape at the ribosomal exit tunnel. *Biophysical Journal* 120:4798–4808.

40. Joiret, M., F. Kerff, F. Rapino, P. Close, and L. Geris, 2022. Ribosome exit tunnel electrostatics. *Physical Review E* 105:014409.

41. Jorgensen, W. L., and J. Tirado-Rives, 1988. The OPLS force field for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc* 110:1657–1666.

42. Dolinsky, T. J., J. E. Nielsen, J. A. McCammon, and N. A. Baker, 2004. PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic acids research* 32:W665–W667.

43. Baker, N. A., D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon, 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences* 98:10037–10041.

44. Wang, J., C. Karki, Y. Xiao, and L. Li, 2020. Electrostatics of prokaryotic ribosome and its biological implication. *Biophysical Journal* 118:1205–1212.

45. Rizuan, A., N. Jovic, T. M. Phan, Y. C. Kim, and J. Mittal, 2022. Developing bonded potentials for a coarse-grained model of intrinsically disordered proteins. *Journal of chemical information and modeling* 62:4474–4485.

46. Harris, C. R., K. J. Millman, S. J. Van Der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, et al., 2020. Array programming with NumPy. *Nature* 585:357–362.

47. Virtanen, P., R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, et al., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods* 17:261–272.

48. Plimpton, S., 1995. Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics* 117:1–19.

49. DeLano, W. L., et al., 2002. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr* 40:82–92.

50. Humphrey, W., A. Dalke, and K. Schulten, 1996. VMD: visual molecular dynamics. *Journal of molecular graphics* 14:33–38.

51. Stukowski, A., 2009. Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool. *Modelling and simulation in materials science and engineering* 18:015012.

52. Meng, E. C., T. D. Goddard, E. F. Pettersen, G. S. Couch, Z. J. Pearson, J. H. Morris, and T. E. Ferrin, 2023. UCSF ChimeraX: Tools for structure building and analysis. *Protein Science* 32:e4792.

53. Dao Duc, K., and Y. S. Song, 2018. The impact of ribosomal interference, codon usage, and exit tunnel interactions on translation elongation rate variation. *PLoS genetics* 14:e1007166.

54. Berka, K., O. Hanák, D. Sehnal, P. Banáš, V. Navratilova, D. Jaiswal, C.-M. Ionescu, R. Svobodová Vařeková, J. Koča, and M. Otyepka, 2012. MOLE online 2.0: Interactive web-based analysis of biomacromolecular channels. *Nucleic acids research* 40:W222–W227.

55. Sehnal, D., R. Svobodová Vařeková, K. Berka, L. Pravda, V. Navrátilová, P. Banáš, C.-M. Ionescu, M. Otyepka, and J. Koča, 2013. MOLE 2.0: advanced approach for analysis of biomacromolecular channels. *Journal of cheminformatics* 5:1–13.

56. Yasuda, T., R. Morita, Y. Shigeta, and R. Harada, 2024. Ribosome Tunnel Environment Drives the Formation of α-Helix during Cotranslational Folding. *Journal of Chemical Information and Modeling* 64:6610–6622.

57. Kolář, M. H., H. McGrath, F. C. Nepomuceno, and M. Černeková, 2024. Three stages of nascent protein translocation through the ribosome exit tunnel. *Wiley Interdisciplinary Reviews: RNA* 15:e1873.

58. Ingolia, N. T., L. F. Lareau, and J. S. Weissman, 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 147:789–802.

59. Nilsson, O. B., R. Hedman, J. Marino, S. Wickles, L. Bischoff, M. Johansson, A. Müller-Lucks, F. Trovato, J. D. Puglisi, E. P. O'Brien, et al., 2015. Cotranslational protein folding inside the ribosome exit tunnel. *Cell reports* 12:1533–1540.

60. Kolář, M. H., G. Nagy, J. Kunkel, S. M. Vaiana, L. V. Bock, and H. Grubmüller, 2022. Folding of VemP into translation-arresting secondary structure is driven by the ribosome exit tunnel. *Nucleic acids research* 50:2258–2269.

61. Neidigh, J. W., R. M. Fesinmeyer, and N. H. Andersen, 2002. Designing a 20-residue protein. *Nature structural biology* 9:425–430.

62. Chovancova, E., A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, et al., 2012. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures .

63. Ho, B. K., and F. Gruswitz, 2008. HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures. *BMC structural biology* 8:1–6.

64. Maglic, J. B., and R. Lavendomme, 2022. MoloVol: an easy-to-use program for analyzing cavities, volumes and surface areas of chemical structures. *Applied Crystallography* 55:1033–1044.

65. López, J., S. Botello, R. Herrera, and M. Carrillo-Tripp, 2019. Parallel high-performance computing algorithm to generate FEM-compliant volumetric mesh representations of biomolecules at atomic scale. *In* Supercomputing: 10th International Conference on Supercomputing in Mexico, ISUM 2019, Monterrey, Mexico, March 25–29, 2019, Revised Selected Papers 10. Springer, 318–333.

66. Monticelli, L., S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S.-J. Marrink, 2008. The MARTINI coarse-grained force field: extension to proteins. *Journal of chemical theory and computation* 4:819–834.

67. Fischer, A.-L. M., A. Tichy, J. Kokot, V. J. Hoerschinger, R. F. Wild, J. R. Riccabona, J. R. Loeffler, F. Waibl, P. K. Quoika, P. Gschwandtner, et al., 2024. The role of force fields and water models in protein folding and unfolding dynamics. *Journal of Chemical Theory and Computation* 20:2321–2333.

68. Gonzalo, P., and J.-P. Reboud, 2003. The puzzling lateral flexible stalk of the ribosome. *Biology of the Cell* 95:179–193.

69. Korostelev, A., D. N. Ermolenko, and H. F. Noller, 2008. Structural dynamics of the ribosome. *Current opinion in chemical biology* 12:674–683.

70. Chowdhury, A. R., D. Sapkota, and D. Girodat, 2024. Conformational changes of ribosomes during translation elongation resolved by molecular dynamics simulations. *Current Opinion in Structural Biology* 86:102804.

71. Kurkcuoglu, O., Z. Kurkcuoglu, P. Doruker, and R. L. Jernigan, 2009. Collective dynamics of the ribosomal tunnel revealed by elastic network modeling. *Proteins: Structure, Function, and Bioinformatics* 75:837–845.

SUPPLEMENTARY MATERIAL

Our SI file contains Supplementary methods and figures, and description of the supplementary videos.