

**Ribosome exit tunnel electrostatics**Marc Joiret<sup>1,\*</sup>, Frederic Kerff<sup>2</sup>, Francesca Rapino<sup>3</sup>, Pierre Close<sup>3</sup>, and Liesbet Geris<sup>1,4,5,†</sup><sup>1</sup>*Biomechanics Research Unit, GIGA In Silico Medicine, Liège University, CHU-B34(+5) 1 Avenue de l'Hôpital, 4000 Liège, Belgium*<sup>2</sup>*UR InBios, Centre d'Ingénierie des Protéines, Bât B6a, Allée du 6 Août, 19, B-4000 Liège, Belgium*<sup>3</sup>*Cancer Signaling, GIGA Stem Cells, CHU-B34(+2) 1 Avenue de l'Hôpital, B-4000 Liège, Belgium*<sup>4</sup>*Skeletal Biology & Engineering Research Center, KU Leuven, ON I Herestraat 49 - box 813, 3000 Leuven, Belgium*<sup>5</sup>*Biomechanics Section, KU Leuven, Celestijnenlaan 300C box 2419, B-3001 Heverlee, Belgium* (Received 22 October 2020; revised 8 September 2021; accepted 8 December 2021; published 13 January 2022)

The impact of ribosome exit tunnel electrostatics on the protein elongation rate or on forces acting upon the nascent polypeptide chain are currently not fully elucidated. In the past, researchers have measured the electrostatic potential inside the ribosome polypeptide exit tunnel at a limited number of spatial points, at least in rabbit reticulocytes. Here we present a basic electrostatic model of the exit tunnel of the ribosome, providing a quantitative physical description of the tunnel interaction with the nascent proteins at all centro-axial points inside the tunnel. We show that a strong electrostatic screening is due to water molecules (not mobile ions) attracted to the ribosomal nucleic acid phosphate moieties buried in the immediate vicinity of the tunnel wall. We also show how the tunnel wall components and local ribosomal protein protrusions impact on the electrostatic potential profile and impede charged amino acid residues from progressing through the tunnel, affecting the elongation rate in a range of  $-40%$  to  $+85%$  when compared to the average elongation rate. The time spent by the ribosome to decode the genetic encrypted message is constrained accordingly. We quantitatively derive, at single-residue resolution, the axial forces acting on the nascent peptide from its particular sequence embedded in the tunnel. The model sheds light on how the experimental data point measurements of the potential are linked to the local structural chemistry of the inner wall, shape, and size of the tunnel. The model consistently connects experimental observations coming from different fields in molecular biology, x-ray crystallography, physical chemistry, biomechanics, and synthetic and multiomics biology. Our model should be a valuable tool to gain insight into protein synthesis dynamics, translational control, and the role of the ribosome's mechanochemistry in the cotranslational protein folding.

DOI: [10.1103/PhysRevE.105.014409](https://doi.org/10.1103/PhysRevE.105.014409)**I. INTRODUCTION**

Ribosomes are the cells' manufacturing tools for building up proteins. They decode the 61 sense codons from a primary message encrypted in a messenger RNA (mRNA) single molecule, and translate it with the help of a set of fewer than 61 transfer RNAs (tRNAs) into 20 amino acids to be sequentially polymerized in a nascent polypeptide that will eventually fold into its final structure. At each elongation cycle, the ribosome incorporates a new amino acid into the nascent protein and translocates to the next codon—shifting along the single-stranded mRNA by three nucleotides (triplet). Ribosomes have three binding sites for tRNAs: the aminoacyl (A), the peptidyl (P), and exit (E) sites, each located between the small and the large subunit of the ribosome. The elongation cycle starts with recognition, accommodation by induced fit, and proofreading of an aminoacylated tRNA on the A site of the ribosome if the cognate anticodon pairs the codon being read on the mRNA [1,2]. Elongation proceeds with the binding of the carboxyl terminal end of the peptide

acylated to the previous tRNA at the P site to the amino moiety of the amino acid acylated on the tRNA at the A site. The formation of the new peptide bond between the nascent chain and the new amino acid is catalyzed at the peptidyl transferase center (PTC) (Fig. 1) by a ribozyme belonging to the large subunit of the ribosome [3]. Two energy-rich guanosine triphosphate molecules (GTP) are used, and two elongation factors with GTPase activity assist the ribosome during each elongation cycle. For more than five decades, attempts to model protein synthesis and mRNA translation from first principles have been pursued extensively [4–7]. Although the average codon translation rate is rather constant transcriptome wide, estimated at 5.6 amino acid residues per second in eukaryotes, codon translation rates have been shown to vary up to 100-fold across a single transcript [8,9]. Many factors influence translation speeds across a single transcript (mRNA), including differences in cognate, near-cognate, and noncognate tRNA relative abundance, nascent-chain charged residues inside the ribosome exit tunnel, mRNA secondary structure, proline residues at either A or P site of the ribosome, steric hindrance between contiguous ribosomes translating the same mRNA molecule, and the finite resource of the ribosome pool available in the cell [10–22]. The individual contributions of each of the previous factors to the rate of the translation are

\*marc.joiret@uliege.be

†liesbet.geris@uliege.be

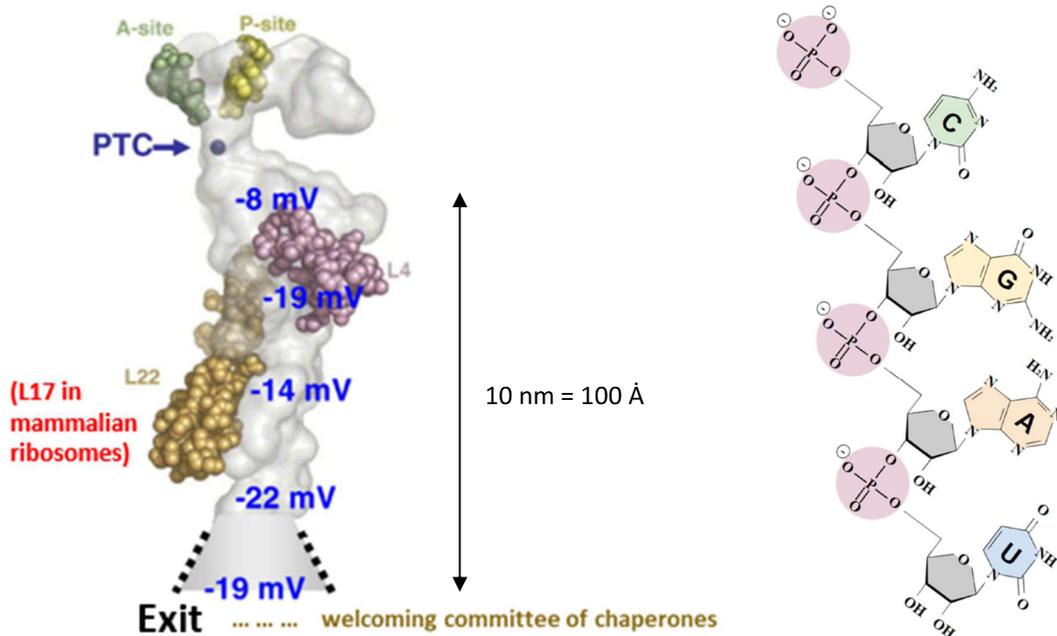


FIG. 1. Left panel: Ribosomal exit tunnel structure. The light gray shape is made up of rRNAs. The peptidyl transfer center (PTC) is where a new amino acid residue is bound to the nascent peptide. The figure is taken with permission from Lu *et al.* [29]. Right panel: RNA molecular structure showing the ribose-phosphate alternating units. The minus signs represent formal negative charges.

difficult to assess quantitatively and separately. Sometimes, depending on the local sequence in the mRNA encrypted message, all these factors interfere and may either antagonize each other or, on the other hand, add up to increase or decrease the rate of translation significantly [12–14,20,21,23]. This hampers our understanding of the dynamics of protein synthesis and specifically of the elongation rate.

Although this has been disputed and it remains a debated question [24–26], some studies have argued that the charged residues are the major determinants of ribosomal velocity [27]. The nascent protein gets out of the ribosome through a tight tunnel approximately 8.5–10 nm long and 1–2 nm wide [28]. The inner wall of the ribosomal exit tunnel is lined with fixed negative charges causing a local negative electrostatic potential inside the tunnel as shown in Fig. 1 [29]. Among the 20 amino acids, two of them are positively charged in physiological conditions, namely, arginine and lysine [30]. A third one, histidine, is only weakly positively charged. When the ribosome incorporates a local increased number of such positively charged amino acid residues in the nascent protein, a local variation in the elongation rate is often reported. This is also true for the negatively charged amino acid residues, i.e., glutamate and aspartate. Ribosome profiling experiments (Ribo-Seq) are difficult to interpret and to reconcile with RNA sequencing (RNA-Seq) profiles and proteome expression results in any given biological condition [20]. *In vitro* laser optical tweezers experiments [31–34], with high-resolution dual traps, involving ribosome specifically, are now being conducted to probe the forces acting upon the mRNA at each translocation step or upon the nascent protein emerging from the exit tunnel [32].

The research community would benefit from highly predictive and quantitatively accurate computational models of

translation dynamics and specifically of elongation rates. A fully realistic model of the electrostatics inside the ribosome exit tunnel is lacking, despite experimental point measurements of the electrostatic scalar potential in the ribosomal exit tunnel being available, at least in one eukaryote species cell type, rabbit reticulocytes [29]. Stochastic models for protein synthesis have been developed for more than 50 years [4–7,14,15]. The extended totally asymmetric simple exclusion process (TASEP) is a widely used stochastic model family dedicated to dynamically simulate the translation rate of a set of transcripts in various conditions [7,35–37]. In the parametrization of TASEP models, most researchers impose an empirical penalty factor to account for the influence of the electrostatic molecular interaction of the ribosome exit tunnel with newly incorporated charged amino residues at the peptidyl transferase center (PTC). For example, a fixed 20% decrease in the translation rate is imposed for those codons that are within five positions downstream of a codon encoding positively charged amino acid residues (lysine, arginine, or histidine) [14]. The negatively charged residues (aspartate and glutamate) are ignored in most studies. This approach is considered inconsistent or too naive if more accurate predictions are expected from TASEP models and to be compared to specific ribosome profiling experimental data [17] or to real-time specific single RNA molecule translation dynamics experiments *in vivo* [9] or *in vitro* [31,32].

In this study, we focus on one of these specific factors which affects the local speed of elongation during protein synthesis, namely, the electrostatic interaction between the charged amino acid residues embedded in the nascent polypeptide chain and the ribosome exit tunnel. To model this electrostatic interaction, we developed a full analytical expression of the electrostatic potential inside the tunnel, start-

ing from two very basic and idealized theoretical geometries for the tunnel. The assumptions on the phenomenological parameters of the model are confronted to the atomic structure of the large ribosomal subunit determined by x-ray crystallography. Together, the model and the experimental structural constraints allow us to investigate the origin of the electrostatic screening that prevails in the very confined environment of the ribosome exit tunnel. The model is used to explore the physical consequences of a possible dynamically variable geometry of the tunnel from a theoretical perspective. The model is used to quantitatively estimate the profile of the axial forces and requires knowledge of the primary sequence of a significant length of the nascent polypeptide chain or its encrypted mRNA to compute the local axial forces acting at the PTC center during elongation. An algorithm is proposed to compute the axial forces acting locally at the PTC and due to a spatially extended electrostatic interaction inside the tunnel. The model is used to conduct comparative analyses of the axial force profiles for different synthetic or real protein sequences. Knowing the axial forces quantitatively allows us to estimate the mechanical work and the biochemical energy required at each elongation step to overcome the electrostatic potential barrier inside the ribosome exit tunnel. These estimations are compared to the energy sources and uptakes involved in the mechanochemistry of the ribosome at each elongation cycle. The ribosome exit tunnel electrostatic model we describe can stand as a building block for computational tools that should be beneficial for the analysis of different experimental techniques like the probing of force by laser optical tweezers, the study of conformational changes with fluorescence resonance energy transfer at the ribosome subunits, the measure of the longitudinal electric field along the ribosome exit tunnel axis by vibrational Stark spectroscopy [38] and for bioinformatic processing of multiomics data.

## II. GEOMETRICALLY IDEALIZED ELECTROSTATIC MODELS OF THE RIBOSOME EXIT TUNNEL

The local negative electrostatic potential inside the ribosome exit tunnel, shown in Fig. 1 (left panel), from which the nascent proteins emerge, originates from the ribosome composition. Ribosomes are composed of two subunits: 50S and 30S in prokaryotes, 60S and 40S in eukaryotes, identified by their sedimentation coefficients, measured in the Svedberg unit  $S$ ; the whole prokaryotic and eukaryotic ribosomes are 70S and 80S, respectively. The ribosome exit tunnel is found in the larger (50S or 60S) of the two subunits. Each of the subunits entails proteins and ribosomal ribonucleic acids (rRNAs). The essential feature of interest of rRNAs, shown in Fig. 1 (right panel), is that, like all RNAs, they are single stranded polymerized molecules with a backbone made up of alternating ribose sugars and phosphate groups all esterified alternatively together. In this long strand, each phosphate group harbors a formal negative charge. The inner wall of the ribosome exit tunnel is mainly lined up with rRNAs (more than 80% w/w in eukaryotic ribosome exit tunnels), though in some locations specific proteins are also present.

### A. General modeling approach and assumptions

Starting with idealized shapes, the ribosome exit tunnel is successively considered as one of the exact three-dimensional surfaces: hollow straight cylinder Fig. 2(a), cone frustum [Fig. 2(b)], and a cone frustum concatenated to a cylinder [Fig. 2(c)] (left panel in each case). The wall material is not of the conductor type with mobile free charges but is rather a dielectric material harboring fixed charges—the fixed phosphate moieties lining the inner wall. As a first reasonable assumption, the fixed charges are supposed to be uniformly distributed on the surface of the inner wall. The size of the hollow cylinder closest to the shape of the ribosome exit tunnel documented in the literature would be 85–100 Å (8.5–10 nm) in length and 10–20 Å (1–2 nm) in diameter [28,39]. The precise length for the ribosomal exit tunnel as measured by cryoelectron microscopy is 9.2 nm on average in prokaryotes and 8.3 nm on average in eukaryotes [39]. The *in vivo* lengths are believed to be a bit larger due to thermal dilatation at the higher temperatures prevailing in living organisms as compared to the cryogenic conditions. For a given uniformly distributed charge density  $\sigma$  on the inner surface wall of the tunnel, the determination of the electrostatic scalar potential  $\Phi(\vec{r})$  and of the electric field  $\vec{E}(\vec{r})$ , at any spatial point close to or far away from the surface, are well stated problems in classical electromagnetism [40]. For the sake of simplicity, we restrict ourselves here on spatial points located on the axis of the tunnel, lying anywhere inside or outside of the tunnel. In this schematic pictorial description, a new amino acid is incorporated into the nascent protein which gets into the tunnel from one side, conventionally from the right of Fig. 2 left panels. The nascent oligopeptide is then pushed by the multitasking ribosomal enzymatic functions inside the tunnel and out of the tunnel at the other side (Fig. 2 left panels). The movement is strictly asymmetric as the nascent protein always enters the tunnel from the same side with the amino terminal end of the protein getting in first and the carboxyl terminal end of the protein getting in last.

The electrical scalar potential  $\Phi(\vec{r})$  at the observed position  $\vec{r}$  is expressed by

$$\Phi(\vec{r}) = \frac{1}{4\pi\epsilon} \iint_S \frac{\sigma(\vec{r}') da}{|\vec{r} - \vec{r}'|}, \quad (1)$$

where  $\sigma(\vec{r}')$  is the surface-charge density (measured in coulombs per square meter) at position  $\vec{r}'$  of the source,  $da$  is the two-dimensional surface element at  $\vec{r}'$  and  $\epsilon$  is the permittivity of the dielectric medium (formula 1.23 in Jackson [40]) with  $\epsilon = \epsilon_r \epsilon_0$ , where  $\epsilon_r$  is the relative permittivity of the medium and  $\epsilon_0$  is the permittivity of free space. We can take advantage of the axial symmetry and restrict to the spatial points on the  $z$ -axis, i.e., for  $\vec{r} = (0, 0, z)$ . The surface integration is conducted on the support of the source charges. The complete detailed derivations of the electrostatic potential and the axial electric field on the tunnel axis are given in the Appendix for three idealized geometries: the hollow straight cylinder, the cone frustum, and the cone frustum concatenated to a hollow cylinder.

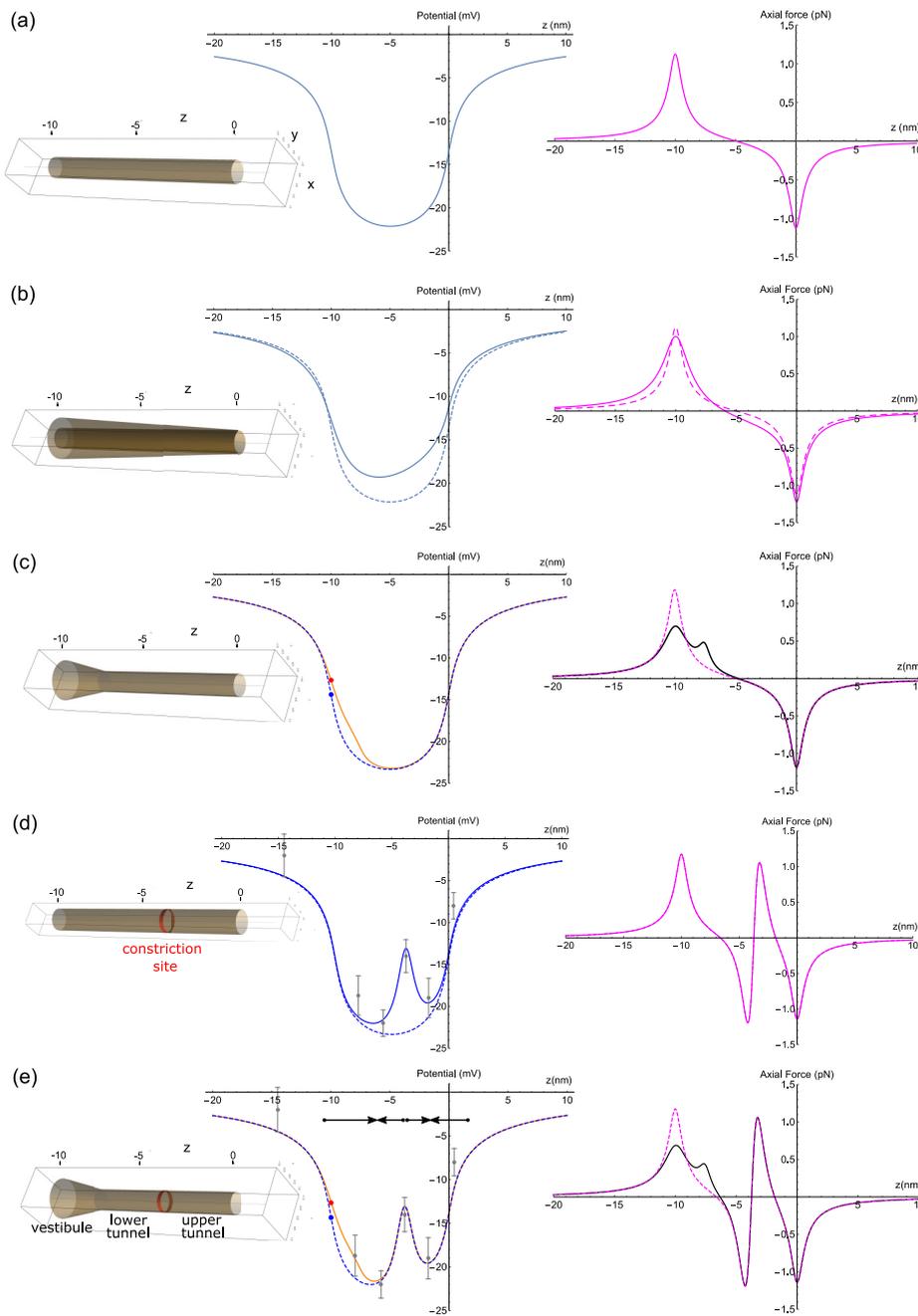


FIG. 2. Idealized and realistic ribosome exit tunnel models. (a) Left panel: Hollow cylinder of length  $L = 10$  nm and  $R = 0.5$  nm with uniformly charged inner wall; (b) Left panel: Normally truncated cone of length  $L = 10$  nm,  $R_{in} = 0.5$  nm and  $R_{out} = 1.0$  nm with uniformly charged inner wall. The entry point is at  $z = 0$  and exit point at  $z = -L$ ; (c) Left panel: Hollow cylinder of length  $L_1 = 8$  nm and  $R = 0.5$  nm concatenated to a truncated cone of length  $L_2 = 2$  nm with a uniformly charged inner wall. Transition between cylinder and truncated cone at 80% of the tunnel total length:  $\lambda = 0.80$ . Electrostatic scalar potential (center panel) and axial force (right panel) profiles for a hollow cylinder (a), a cone (b), and a cylinder concatenated to a truncated cone (c); dashed lines: cylinder only; full lines: cylinder concatenated to cone. Red and blue points: the potential at the tunnel exit point is higher than the potential at the entry point. The axial forces in the tunnel exit region are smoother and more dispersed in the combined geometry (c), (e). (d), (e) Best fitted models for the electrostatic scalar potential of the ribosomal exit tunnel. Center panel: A Lorentzian peak was added locally to the idealized models potential which was fitted to the experimental data points obtained by Lu *et al.* [29]. The protein protrusion’s position is indicated by the red ring (left panel). Dashed blue line: potential resulting from the idealized uniformly charged hollow cylinder. Orange line: potential resulting from the superposition of a cylinder and a truncated cone. The transition between the cylinder and the truncated cone is at  $\lambda = 0.75$  (e). 95% confidence interval error bars computed from the experimental data. Right panel: The axial forces for a positively unit charged test amino acid residue on the tunnel axis as a function of axial position in the tunnel for the best fitted models. Black line: axial forces profile resulting from the superposition of a cylinder and a truncated cone. Right panel: Dashed magenta idealized uniformly charged hollow cylinder. Black arrows (e) indicate the direction of the axial forces upon a positively charged test residue.

### B. Hollow straight cylinder

In a first simplified approach, the ribosome exit tunnel is considered a hollow straight cylinder [Fig. 2(a) left panel]. The electrostatic potential is given by the expression

$$\Phi(z) = \frac{\sigma R}{2\epsilon} \log \frac{\left| \frac{z+L}{R} + \sqrt{\left(\frac{z+L}{R}\right)^2 + 1} \right|}{\left| \frac{z}{R} + \sqrt{\left(\frac{z}{R}\right)^2 + 1} \right|}, \quad (2)$$

and the axial electric field can be written as

$$E_z(z) = -\frac{\sigma R}{2\epsilon} \left( \frac{1}{\sqrt{R^2 + (z+L)^2}} - \frac{1}{\sqrt{R^2 + z^2}} \right). \quad (3)$$

The axial force applied on a test particle is the product of the axial electric field with the charge of the test particle:

$$F_z = qE_z. \quad (4)$$

The plots of electrostatic scalar potential  $\Phi(z)$  and of the axial force  $F_z$  acting on a unit test charge located on the tunnel axis at any point of coordinate  $z$  are displayed in Fig. 2(a), with the medium permittivity prevailing inside the ribosome exit tunnel (see below). A negative force means that the test particle is forced to move towards negative  $z$  values whereas a positive force means that the test particle is forced to move towards

positive  $z$  values. In these plots, the ratio  $\sigma/\epsilon$  is adjusted so that the potential fits the range of the experimentally measured values given, for instance, in Lu *et al.* [29].

### C. Normally truncated straight cone model

An alternative approach would depict the tunnel as a hollow cone normally truncated at both ends [Fig. 2(b) left panel]. The section radius at the entry point is equal to  $R = 0.5$  nm but with a section radius twice that value at the tunnel exit point, and equal to  $R = 1$  nm. With the total axial length kept at  $L = 10$  nm, the half opening angle along the axis is  $\alpha \sim 0.05$  radian (2.86 arc degrees) and exactly such that  $\tan \alpha = R/L$  complying with the observation that the diameter at the exit point is around twice the diameter at the entry point of the tunnel. To alleviate the notations, the two following substitutions are adopted:

$$f_1(z) = R \cos \alpha - z \sin \alpha, \quad (5)$$

$$f_2(z) = R \sin \alpha + z \cos \alpha. \quad (6)$$

$f_1(z)$  is always positive for  $z \leq 0$  (and even for  $z < R/\tan \alpha$ , i.e., the virtual  $z$ -position of the cone summit), which is the domain we are interested in. The  $z$ -position values are negative in the tunnel and beyond its exit point.

The electrostatic potential can be written as

$$\begin{aligned} \Phi_{\text{cone}}(z) = \frac{\sigma}{2\epsilon} \left\{ f_1(z) \cos \alpha \log \left[ \frac{\left| \frac{L}{\cos \alpha} + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + \frac{L^2}{\cos^2 \alpha}} \right|}{\left| f_2(z) + \sqrt{R^2 + z^2} \right|} \right] \right. \\ \left. + \sin \alpha [\sqrt{(z+L)^2 + (R+L \tan \alpha)^2} - \sqrt{R^2 + z^2}] \right\}. \quad (7) \end{aligned}$$

Equation (7), which is valid for any conical geometry with entry section of radius  $R$  and any cone angle  $\alpha$ , replaces Eq. (2) of the cylindrical geometry. Note that the electrostatic potential vanishes at  $z = \pm\infty$  as physically expected. It is also worth noticing that Eq. (7) for the truncated cone restores, as a special case, Eq. (2) for the cylinder when  $\alpha = 0$ , as expected as well.

The axial electric field is given by the expression

$$\begin{aligned} E_{z\text{cone}}(z) = \frac{\sigma}{2\epsilon} \left\{ \sin \alpha \cos \alpha \log \frac{L/\cos \alpha + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + L^2/\cos^2 \alpha}}{f_2(z) + \sqrt{R^2 + z^2}} + \frac{f_1(z) \cos \alpha (\cos \alpha + \frac{z}{\sqrt{R^2 + z^2}})}{f_2(z) + \sqrt{R^2 + z^2}} \right. \\ - f_1(z) \cos \alpha \frac{\cos \alpha + \frac{z+L}{\sqrt{z^2 + 2L(z+R) + R^2 + L^2/\cos^2 \alpha}}}{L/\cos \alpha + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + L^2/\cos^2 \alpha}} \\ \left. - \sin \alpha \left[ \frac{z+L}{(z+L)^2 + (R+L \tan \alpha)^2} - \frac{z}{\sqrt{R^2 + z^2}} \right] \right\}. \quad (8) \end{aligned}$$

Multiplying Eq. (8) by a positive unit test charge yields the axial forces acting on a positive unit test charge. The plot of the axial forces as a function of the position in the tunnel is displayed in Fig. 2(b) right panel for the truncated cone geometry and compared to the cylinder case. Experimental measurements made on ribosome exit tunnels show that the tunnel exit section radius is around 1 nm, i.e., twice the radius of the innermost part of the tunnel. If the ribosome tunnel were of the cone type, the cone opening angle would be around

$\alpha \sim 0.05$  radian (2.86 arc degrees). The consequence on the electrostatic potential profile is of importance because, with this conical geometry, and if the total charges are kept the same for the two surfaces, the electrostatic potential inside the tunnel will necessarily be algebraically higher than the potential profile in the case of the cylinder as displayed in Fig. 2(b) (central panel) where the analytical equation for the electrostatic potential for the truncated cone was plotted and compared to the cylinder case. A simple geometrical calcu-

lation shows that if the two surfaces support the same total charges  $Q_1 = Q_2$ , then  $\sigma_2 = S_{\text{cylinder}}/S_{\text{cone}} \times \sigma_1 = \frac{2}{3} \times \sigma_1$ , for a geometry where both tunnels have the same radius at the entry point, the same total lengths  $L$ , but where the cone exit section has a radius twice as large as the cylindrical radius. The surface charge density  $\sigma_2$  on the lateral truncated cone inner surface would be two-thirds of the surface charge density  $\sigma_1$  prevailing on the lateral inner surface of the cylinder. Moreover, the potential profile in the conical geometry is skewed to the left as compared to the potential profile for the cylindrical geometry. An asymmetry in the potential profile appears due to the change in radius along the  $z$ -axis of the cone. The minimal value of the potential is shifted to the left. The slope of the cylindrical potential profile is steeper than the conical potential at the tunnel exit point, meaning that the electric field intensity will be a bit weaker in that region for the conical geometry as can be seen in Fig. 2(b) of the axial forces curves. The axial forces vary more smoothly and are more dispersed in the conical geometry than in the cylindrical geometry.

$$\Phi(z) = \frac{1}{4\pi} \iint_{K=\{(u,v) \in [0,2\pi] \times [-1,0]\}} \left[ \chi_{[-1, -\lambda]}(v) \frac{\sigma_{\text{cone}}}{\epsilon_{\text{cone}}} \text{Integrand}_{\text{cone}}(u, v, z + \Delta z_{\text{shift}}) + \chi_{[-\lambda, 0]}(v) \frac{\sigma_{\text{cylinder}}}{\epsilon_{\text{cylinder}}} \text{Integrand}_{\text{cylinder}}(u, v, z) \right] du dv, \quad (9)$$

where the characteristic function  $\chi_{[a,b]}(v)$ , used here for the correct setting of the charged sources distribution, is defined by

$$\chi_{[a,b]}(v) = \begin{cases} 1 & \text{if } v \in [a, b] \\ 0 & \text{if } v \notin [a, b] \end{cases}. \quad (10)$$

A  $z$ -shift was also incorporated in Eq. (9) to account for the shift in axial position of the truncated cone, as can be noticed by comparing the left panels in Figs. 2(b) and 2(c). The  $z$ -shift must match the  $\lambda$  value adopted as the interval limit in both characteristic functions in Eq. (9), delineating the limit between the start of the truncated cone and the end of the cylinder when moving axially to the left from  $z = 0$ . For Fig. 2(c),  $z_{\text{shift}} = \lambda L = 0.8L$  ( $L$  is the tunnel total axial length). The  $\alpha$  angle for the cone may be a free parameter to be determined. Different surface charge densities can be incorporated as well, using two different values for  $\sigma$  in the two integrands, providing an extra degree of freedom to fit the model to the experimental observations. Different permittivities can be incorporated as well, considering some heterogeneity in the tunnel medium dielectric response (see next section). In summary, four parameters can be fitted to the experimental data:  $\sigma_{\text{cylinder}}$ ,  $\sigma_{\text{cone}}$  (if the permittivity is hypothesized to be uniform along the tunnel medium length) or  $\sigma_{\text{cylinder}}/\epsilon_{\text{cylinder}}$ ,  $\sigma_{\text{cone}}/\epsilon_{\text{cone}}$  (if a piecewise discontinuity is incorporated for the ratio of the surface charge density over permittivity along the tunnel medium length),  $\lambda$ , and  $\alpha$  which are the surface charge density on the cylindrical surface, the surface charge density on the cone frustum, or their quotients over permittivities, the fraction of the ribosome length occupied by the cylinder, and the cone frustum half opening angle,

#### D. Normally truncated cone concatenated to a cylinder

The question to know whether or not the tunnel is geometrically exactly more like a cylinder or like a truncated cone is less important than the consequence on the electrostatic potential profile. The salient feature of the real ribosome exit tunnel is that there is indeed a widening in the tunnel section at the exit. Therefore, a better simple geometrical model that fits most of the experimental observations to date is a model combining the cylinder and the normally truncated cone as shown in Fig. 2(c). The transition from a cylindrical shape to a conical shape results in an electrostatic potential rise along the tunnel axis when moving from the entry point to the exit point. This is a fundamental difference between the two geometries (truncated cone combined with cylinder versus cylinder alone) that has both energetic and biological consequences. The electrostatic potential resulting from such a configuration results from the linear combination of the straight cylinder model part and the normally truncated cone model part [linear superposition of integrands in Eqs. (A3) and (A24) in the Appendix]:

respectively. Each of these phenomenological parameters can influence the electrostatic potential profile of the tunnel.

The axial electrical field resulting from the combination of the cylindrical geometry for 75% of the tunnel length ( $0.75L$ ) from its entry point and of the truncated cone geometry for the remaining 25% ( $0.25L$ ) of the length in the ribosome exit tunnel, with or without an added Lorentzian peak (see Sec. II E), is the superposition of Eqs. (3) and (8). The parameter settings have to be consistent with the chosen geometry, with the surface charge densities ( $\sigma_1$  and  $\sigma_2$ ) and with the dielectric responses ( $\epsilon_1$  and  $\epsilon_2$ ). More specifically, for a given  $\lambda$ , one would have  $L_1 = \lambda L$  and  $L_2 = (1 - \lambda)L$ . The surface charge density  $\sigma_2$  of the truncated cone that we adopted was such that  $\sigma_2/\sigma_1 = S_{1 \text{ lateral}}/S_{2 \text{ lateral}}$  because it best fits the observational data that were gained from the rabbit reticulocyte ribosomes, given a fixed uniform dielectric response  $\epsilon$  throughout the tunnel. This condition is consistent with the possibility that the conical part of the tunnel end could result from an elastic deformation of an initially cylindrical-shaped tunnel with a uniform surface charge density (conservation of total initial charge before and after this hypothetical elastic deformation of the inner surface turning the cylinder into a truncated cone at the exit side of the tunnel). Alternatively, if the dielectric response along the tunnel medium entails piecewise heterogeneity when going from the cylinder to the cone frustum, the previous assumptions of charge conservation and elastic deformation can be relaxed. In the latter case, the fitted parameters are the ratios  $(\sigma/\epsilon)_{\text{cylinder}}$  and  $(\sigma/\epsilon)_{\text{cone}}$  (or  $\sigma_1/\epsilon_1$  and  $\sigma_2/\epsilon_2$ ). Under the dielectric response piecewise heterogeneity assumption, the individual numerical values of  $\sigma$  and  $\epsilon$  cannot be determined separately; only their ratio  $\sigma/\epsilon$  is numerically

accessible upon fitting to the measured electrostatic potential data points.

The area under the curve of the axial forces profile yields the mechanical energy required for a unit charge to move between two axial points inside the tunnel. Equivalently, the required mechanical energy can easily be computed by multiplying the unit charge with the electrostatic potential difference between the tunnel exit point and the tunnel entry point,

$$W = q(\Phi_{z=-L} - \Phi_{z=0}), \quad (11)$$

where  $W$  is the mechanical work required for a unit test charge  $q$  to move across an electrostatic potential difference (in Volts) from the tunnel entry point ( $z = 0$ ) to the tunnel exit point ( $z = -L$ ). The result is expressed in J/mol by multiplying Eq. (11) with the Avogadro number.

The geometrical asymmetry induced by the widening in the tunnel radius at the exit of the tunnel is important because it introduces a permanent difference in the electrostatic potential between the exit and the entry points of the tunnel as shown in Figs. 2(b) and 2(c) center panel.

This is unfavorable to the positively charged amino acid residues that will have to traverse the tunnel and will require more mechanical energy to overcome this electrostatic potential difference than their negatively charged amino acid counterparts, when moving from the entry point  $z = 0$  to the exit point  $z = -L$ . In an adopted geometry that best fits the experimental observations (see Sec. II E), with  $\lambda = 0.75$ ,  $\alpha = 0.198$  and with  $\sigma_2 = 2/3 \sigma_1$ , the potential difference is  $-12.65 - (-14.35) = 1.70$  mV and the required mechanical work for traversing the tunnel is  $0.164$  kJ/mol  $\sim 0.039$  kcal/mol for a positive unit charge embedded in an otherwise neutral nascent peptide stretch. This provides a rough estimate of the energy required for a single positively charged amino acid residue to traverse the ribosomal exit tunnel if this positive residue is embedded in a completely neutral peptide sequence. The mechanical energy requirement for real sequences depends on the particular distribution of the charged amino acid residues along the primary sequence. There might be particular sequence contexts for which the local mechanical energy requirements could be much higher than the estimated values given above. For a straight cylinder, the mechanical energy is equal to zero (symmetry in the potential between exit and entry points), whereas for the truncated cone concatenated to the cylinder (asymmetry in the potential between the exit and entry points), the mechanical energy uptake when moving a single positive unit charge from the entry point ( $z = 0$ ) to the exit point ( $z = -L$ ) is estimated to be around  $164$  Nm/mol for a stretch of 40 residues in the tunnel, according to our electrostatic interaction model. The estimated mechanical energy uptake per residue incorporation would be around  $4.1$  Nm/mol per residue incorporation ( $\sim 0.001$  kcal/mol per residue incorporation). This is due to the fact that the axial forces profile is not symmetrical in the cone concatenated to the cylinder geometry. However, in this truncated cone geometry (asymmetrical potential), the axial forces amplitudes are reduced and are more spatially dispersed than in the cylindrical geometry as displayed in Fig. 2(c) right panel.

### E. Empirically improved electrostatic potential profile of cylindrical shape best fitted to experimental point measurements

The real electrostatic potential profile inside the ribosome exit tunnel was experimentally measured with an ingenious biochemical technique of molecular tape at least in rabbit reticulocyte ribosomes by Deutsch *et al.* [29]. The measurements were mapped on the x-ray solved spatial structure of the ribosomal tunnel of the archeon *Haloarcula marismortui*. To the authors' knowledge, the potential profiles in other eukaryotic or prokaryotic ribosomes exit tunnels have not been measured experimentally ever since. The real electrostatic potential profile is actually not symmetric. We further need to build an improved and more realistic potential profile by adding to the previous idealized models a small Lorentzian peak function. The motivation for this comes from the experimental data showing that the electrostatic potential locally increases at a distance one-third of the length of the tunnel away from the PTC center (approximately at least 15–17 amino acid residues in the nascent protein upstream from the amino acid residue incorporation site). This local increase in the potential is located near the position of the ribosomal constriction, where specific ribosomal conserved constitutive proteins protrude inward the tunnel, i.e., L4 in both bacteria and eukaryotes, L22 in bacteria and L17 in eukaryotes; see Fig. 1 (left). Dao Duc *et al.* [39] confirmed, with multiple sequence alignments of uL22 and L4 proteins across 20 species in the three domain of life, the presence of a highly conserved sequence enriched in arginine (R) and or lysine (K). In uL22 there are up to seven R or Ks conserved between position 154 and 176. In L4 there are five Rs (or Ks) conserved between position 71 and 92 across eukaryotic species and up to six Ks or Rs conserved between position 69 and 82 across prokaryotic species. Similar conservation has been shown for uL23 (bacteria) and eL39 (eukaryotes) [39]. These positively charged residues protrude near the tunnel constriction and explain the local rise of the potential. The Lorentzian local peak potential as expressed in Eq. (12) that we added was fitted to the experimental data obtained by Lu *et al.* [29]:

$$\Phi_{\text{Lorentzian}} = \frac{f_{\text{scale}}}{\Gamma/2} \frac{\Gamma/2}{(z - z_0)^2 + (\Gamma/2)^2}. \quad (12)$$

The fitted parameters values are  $f_{\text{scale}}/(\Gamma/2) = 9 \times 10^{-9}$  V m for the scale factor divided by  $\Gamma/2$ ,  $\Gamma = 18 \times 10^{-10}$  m for the Lorentzian peak full width at half maximum and  $z_0 = -3.75 \times 10^{-9}$  m for the peak center location, i.e.,  $37.5$  Å measured from the entry side point towards the protein tunnel exit. The experimental data points taken from Lu *et al.* [29] and the fitted adapted function for the ribosome exit tunnel electrostatic potential are displayed in the center panel of Fig. 2(d) when the simple straight cylinder geometry is adopted. The extended expression for the total electrostatic potential in this improved model version is

$$\begin{aligned} \Phi_{\text{total}} &= \Phi_{\text{cylinder}} + \Phi_{\text{Lorentzian}} \\ &= \frac{\sigma R}{2\epsilon} \log \frac{\left| \frac{z+L}{R} + \sqrt{\left(\frac{z+L}{R}\right)^2 + 1} \right|}{\left| \frac{z}{R} + \sqrt{\left(\frac{z}{R}\right)^2 + 1} \right|} \\ &\quad + \frac{f_{\text{scale}}}{\Gamma/2} \frac{\Gamma/2}{(z - z_0)^2 + (\Gamma/2)^2}. \end{aligned} \quad (13)$$

An important characteristic of the Lorentzian function  $\Phi_{\text{Lorentzian}}$  that is shared with the potential  $\Phi_{\text{cylinder}}$  is its vanishing at infinity in both directions. At this point, it is important to underline the physical interpretation of the  $f_{\text{scale}}$  parameter in formulas (12) and (13). Indeed,  $f_{\text{scale}}$  has units of  $\text{V m}^2$  and can be related to a local  $\sigma/\epsilon$  physical value in the constriction site inner volume. A dimensional analysis shows that the  $f_{\text{scale}}$  units are equivalent to the units of  $\sigma/\epsilon$  ( $\text{V/m}$ ) times the units of a volume ( $\text{m}^3$ ). Conventionally naming  $\sigma_3/\epsilon_3$ , the local ratio of charge density over the dielectric response at the constriction site, we have the following dimensional relation:

$$f_{\text{scale}} \equiv \frac{\sigma_3}{\epsilon_3} \pi R^2 \Gamma, \quad (14)$$

where  $R$  is the radius of the cylinder and  $\Gamma$ , the full width at half maximum of the Lorentzian peak, is an approximation of the length of the constriction (electrostatic functional length at the constriction site). From this dimensional relation (14), we can infer the ratio  $\frac{\sigma_3}{\epsilon_3}$  at the constriction site from the above fitted values of  $f_{\text{scale}}$  and  $\Gamma$ . With  $R = 5 \text{ \AA}$ ,  $\sigma_3/\epsilon_3 = 5.73 \times 10^9 \text{ V/m}$  (or  $\sim 57.3 \text{ MV/cm}$ ).

The total axial electric field (and the axial force) is obtained by

$$E_{z,\text{total}} = E_{z,\text{cylinder}} + E_{z,\text{Lorentzian}} \quad (15)$$

$$\begin{aligned} \Phi_{\text{total}} &= \Phi_{\text{cone}}(z + \lambda L, L_2, \sigma_2) + \Phi_{\text{cylinder}}(z, L_1, \sigma_1) + \Phi_{\text{Lorentzian}}(z) \\ &= \frac{\sigma_2}{2\epsilon_2} \left\{ f_1(z + L_1) \cos \alpha \log \left[ \frac{|L_2/\cos \alpha + f_2(z + L_1) + \sqrt{(z + L_1)^2 + 2L_2((z + L_1) + R) + R^2 + \frac{L_2^2}{\cos^2 \alpha}}|}{|f_2(z + L_1) + \sqrt{R^2 + (z + L_1)^2}|} \right] \right. \\ &\quad \left. + \sin \alpha [\sqrt{((z + L_1) + L_2)^2 + (R + L_2 \tan \alpha)^2} - \sqrt{R^2 + (z + L_1)^2}] \right\} \\ &\quad + \frac{\sigma_1 R}{2\epsilon_1} \log \frac{|\frac{z+L_1}{R} + \sqrt{(\frac{z+L_1}{R})^2 + 1}|}{|\frac{z}{R} + \sqrt{(\frac{z}{R})^2 + 1}|} + \frac{f_{\text{scale}}}{\Gamma/2} \frac{\Gamma/2}{(z - z_0)^2 + (\Gamma/2)^2}, \end{aligned} \quad (17)$$

where  $L_1 = \lambda L$ ,  $L_2 = (1 - \lambda)L$ ,  $L = L_1 + L_2$ . Similarly for the axial electric field (and the axial force) along the  $z$ -axis:

$$E_{z,\text{total}} = E_{z,\text{cone}}(z + \lambda L, \sigma_2, L_2) + E_{z,\text{cylinder}}(z, \sigma_1, L_1) + E_{z,\text{Lorentzian}}(z) \quad (19)$$

$$= E_{z,\text{cone}}(z + L_1, \sigma_2, L_2) - \frac{\sigma_1 R}{2\epsilon_1} \left( \frac{1}{\sqrt{R^2 + (z + L_1)^2}} - \frac{1}{\sqrt{R^2 + z^2}} \right) + \frac{f_{\text{scale}}}{\Gamma/2} \frac{2(\Gamma/2)(z - z_0)}{((z - z_0)^2 + (\Gamma/2)^2)^2}, \quad (20)$$

where the detailed expression of the first term in the last right-hand side is easily obtained from Eq. (8) by substituting  $z + \lambda L$  to  $L$ ,  $\sigma_2$  to  $\sigma$  and  $L_2$  to  $L$ . The plots of the electrostatic potential and of the total axial force profiles in the ribosome exit tunnel under this last improvement of the model are displayed in Fig. 2(e) central and right panels. The central panel shows the goodness of the fit with the experimental data of Lu *et al.* [29] The improvement of the fit due to the cone geometry concatenated to the hollow cylinder is worth noticing [orange line in the central panel of Fig. 2(e)]. This

$$\begin{aligned} &= -\frac{\sigma R}{2\epsilon} \left( \frac{1}{\sqrt{R^2 + (z + L)^2}} - \frac{1}{\sqrt{R^2 + z^2}} \right) \\ &\quad + \frac{f_{\text{scale}}}{\Gamma/2} \frac{2(\Gamma/2)(z - z_0)}{((z - z_0)^2 + (\Gamma/2)^2)^2}, \end{aligned} \quad (16)$$

from which the axial force results immediately by  $F_z = qE_z$  and is displayed in Fig. 2(d) right panel in the case of a single positively unit charged amino acid residue.

The charge surface density  $\sigma$  can be made dependent on the  $z$  variable in a stepwise manner to account for local heterogeneity on the ribosome wall and to account for the experimentally observed potential profile. The dielectric constant  $\epsilon$  or the ratio  $\sigma/\epsilon$  can also be made dependent on the  $z$  variable in a stepwise manner.

#### F. Empirically improved electrostatic potential profile of truncated cone combined with cylindrical shape best fitted to experimental point measurements

A still better fit of the experimental data of Deutsch *et al.* [29] is obtained with the truncated cone concatenated to the cylinder geometry. Keeping the same Lorentzian peak, the best extended expression for the total electrostatic potential in this last improved version of the model is

last version of the model perfectly fits the four experimental points located inside the tunnel. The model is a very good fit of all the six experimental data points if the tunnel length is taken in a range from  $L = 8.5 \text{ nm}$  to  $L = 9.5 \text{ nm}$ , keeping all the other parameters constant. Figure 3 shows the plots for the potential curves for these two tunnel lengths boundaries and shows that the six experimental measurements are correctly captured within the 95% confidence intervals of the potential measurements between these two length boundaries. In their study, Lu *et al.* mapped their six experimental points on the

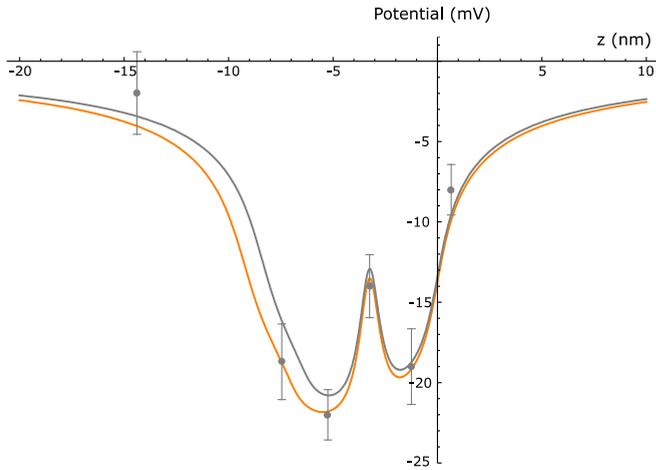


FIG. 3. Electrostatic scalar potential curves for two tunnel length boundaries. Tunnel length = 9.5 nm: orange curve. Tunnel length = 8.5 nm: gray curve. Both curves capture the six experimental points within the 95% confidence interval of the potential measurements that were mapped on the ribosomal crystal structure of *H. marismortui* by Lu *et al.* [29].

ribosomal crystal structure of *H. marismortui* (archae) for which the x-ray or cryoelectron microscopy resolved ribosome structure gives a tunnel length of  $\sim 9.5$  nm [39]. It is recognized that there might be some deviations in mapping distance and with respect to the actual length of the ribosome exit tunnel of the biological material that they used. Also, the actual *in vivo* lengths might slightly differ from the lengths determined in the cryogenic conditions prevailing in cryoelectron microscopy and the functional length might also slightly differ to the geometrical length. The x-ray and cryoelectron microscopy ribosome structure resolution conducted on 23 species across the three domain of life are supporting the fact that the ribosome exit tunnel in bacteria is a bit longer than the one in eukaryotes while archae have intermediate lengths between bacteria and eukaryotes [39]. Hence, a tunnel length of  $L \approx 8.5$  nm, should be adopted if we aim at building a model for the eukaryotic or mammalian ribosome exit tunnel, that could be eventually used for computational biology and bioinformatics purposes on eukaryotic or mammalian omics data.

### III. MODEL ASSUMPTIONS

The main simplifying assumptions related to the physical properties of the tunnel lumen and inner wall on which the models are based are the following:

(1) Uniform distribution of the charges with possible different values for the surface charge densities ( $\sigma$ ) in the different tunnel inner wall parts (upper tunnel, constriction site, lower tunnel and tunnel vestibule). This is the piecewise uniform surface charge density of the tunnel inner wall assumption.

(2) Piecewise continuous constant dielectric responses ( $\epsilon$ ) in the medium prevailing inside the tunnel (upper tunnel, constriction site, lower tunnel and tunnel vestibule). This assumption combined with the previous one is equivalent to the piecewise constant assumption for the ratio  $\sigma/\epsilon$ . A space

dependence of  $\sigma$  is possible if it is compensated for by a similar space dependence of  $\epsilon$  provided the combined ratio  $\sigma/\epsilon$  is constant in a region of interest. This piecewise constant ratio is the strictly necessary assumption for the mathematical surface integration calculations of our models to be analytically tractable.

(3) Polarization surface charge density effects due to media discontinuities which are implicitly incorporated as a net apparent surface charge density  $\sigma$ .

(4) Existence of a strong water screening effect inside the tunnel wall (and not due to mobile ions) which results in the substitution of a *renormalized apparent* surface charge density ( $\sigma$ ) to the *bare* surface charge density ( $\sigma^*$ ).

These assumptions are studied in this section. In order to show how reasonable or realistic these assumptions are, x-ray crystallographic data of a real ribosomal large subunit were analysed and a 2D map at atomic resolution of the tunnel inner wall was built. The model assumptions were confronted to the crystallographic observations. We first present hereafter the tunnel mapping of the x-ray crystallographic solved atomic structure of the large ribosomal subunit.

#### A. Mapping the negatively charged phosphate groups close to the tunnel inner wall and counting the number of charged groups contributing to the formal surface charge density of the ribosome tunnel inner wall from a x-ray solved structure of the large ribosomal subunit

We analyzed the publicly available structure of the large ribosomal subunit of the archeon *H. marismortui* [41] obtained from x-ray crystallography at 2.4 Å [42]. Other ribosomal large subunit structures are available across the three domains of life [39], but we chose 4V9F because it was on *H. marismortui* ribosomal structure that Deutsch *et al.* mapped on their data points measurement of the potential and also because this PDB entry is one of the best resolution of the ribosomal large subunit to date. The structure has an atom count of 103 831 (non-hydrogen atoms) and includes 2808 nucleotides of the 23S RNA, all of the 122 nucleotides in the 5S RNA, and 4311 amino acid residues from 32 ribosomal proteins, as well as 261 ions (138  $\text{Mg}^{2+}$ , 30  $\text{Cl}^{-1}$ , 85  $\text{Na}^{+}$ , three  $\text{K}^{+}$ , and five  $\text{Cd}^{2+}$ ) and 7745 water molecules. The structure has been refined to an *R* factor of 0.166 and an *R* free of 0.206. The refinement statistics indicate root mean square deviations of bond lengths of 0.006 Å and bond angles of 1.08°. It is estimated, that with this 2.4 Å resolution and a *R* factor smaller than 0.2, the upper limit precision on the atoms position is around 0.2 Å for 4V9F. We can therefore consider that atom positions are precise to within one-twelfth of the stated resolution [43]. The atom positions are precise to within one-eighth (12.5%) of the length of a phosphorus-oxygen bond in the phosphodiester bond between a phosphorus atom and the bridging oxygen of the O3'-ribose or O5'-ribose in the nucleotides.

To find the ribosome exit tunnel and extract the tunnel centerline coordinates, we used a tunnel search algorithm developed by Sehnal *et al.* [44], implemented in MOLE 2.0 and the web-based MOLEonline 2.0 tool publicly available online [45,46]. We used PyMOL (PyMOL Molecular Graphics System, Version 2.3.2) and exported the relevant selected atom

positions cartesian coordinates to output files. These files were further processed with mapping algorithms coded in Python to build 2D positional maps of the charged chemical groups on or near the inner surface of the ribosome exit tunnel, as viewed from the tunnel centerline. The protocol to build the 2D maps is detailed in the Appendix. We algorithmically set out the 3D equations of the cylinder and the cone frustum in a reference frame aligned with the centerline of the real x-ray solved structure, to calculate the closest distance of the selected atoms to the surface of the tunnel model. The  $\Delta$  values shown in blue in Fig. 4 were obtained as the closest (orthonormal) distance of the PyMOL selected atoms either from the model cylinder surface or from the model cone surface, depending on where the atoms are located. The 2D maps of the selected atoms within 12 Å of the tunnel centerline, projected from the 3D space surfaces of the tunnel inner wall model as shown in Fig. 4 allow to determine the *formal bare* charge surface densities  $\sigma^*$ . The total phosphorus atom counts is 94 on or near the tunnel model surface, 56 of which are close to the cylinder surface and 38 close to the vestibule surface. Among the 56 phosphorus atoms in the immediate vicinity of the model cylinder surface, 23 are in the upper tunnel, 10 near the constriction site, and 23 in the lower tunnel. We tested the null hypothesis of a uniform distribution of the phosphorus atoms across the cylinder surface. We calculated the  $p$ -value by 100 000 Monte Carlo (or bootstrap) simulations of sets of 56 atoms that were uniformly distributed across the cylinder surface and counted how many of these sets entailed atom counts in the three regions that were equal or more extreme than the actual observed combination of 23, 10, 23. The  $p$ -value to observe the mapped distribution or a more extreme one across the three cylindrical developed areas, i.e., upper tunnel, constriction site, and lower tunnel, is larger than 0.083. The assumption of a uniform 2D-joint distribution of the negatively charged phosphate moieties in the immediate vicinity of the tunnel cylinder wall is reasonable and can not be ruled out. We also tested the null hypothesis of uniformity of the 1D-marginal distributions along the  $z$ -axis or along the  $\phi$ -arc angle with the chi-squared and the Kolmogorov-Smirnov tests (1D marginal along  $\phi$ :  $\chi^2$ -test  $p$ -value = 0.549; 1D marginal along  $z$ :  $\chi^2$ -test  $p$ -value = 0.487, KS-test  $p$ -value = 0.41). Therefore, we concluded that the formal bare charge distribution contributed by the phosphate moieties on the cylinder is  $\sigma_1^* = -2.1 |e|/\text{nm}^2$  and that this value can be retained for the upper cylinder, the constriction site and the lower tunnel altogether. Figure 4(b) shows the 56 phosphorus atoms with an orthogonal distance to the cylinder surface of less than 12 Å. The average distance of a phosphorus atom to the model tunnel inner wall is 6.68 Å. The surface charge density on the 7.45 nm<sup>2</sup> of the model tunnel surface of the vestibule is  $\sigma_2^* = -5.1 |e|/\text{nm}^2$ . The vestibule surface appears to be more enriched in phosphate moieties near the end edges of the tunnel. The 2D-joint distribution of the phosphorus atoms across the cone surface cannot be considered uniform,  $p$ -value  $\leq 0.009$ . This  $p$ -value was calculated using two methods, i.e., the rejection-sampling method and the method of root squared random variable to take into account the increase of the surface elements on a cone surface when moving from the small to the large radius of the cone. We also tested the null hypothesis of uniformity of the 1D-marginal distributions

along the  $z$ -axis or along the  $\phi$ -arc angle for the cone with the chi-squared and the Kolmogorov-Smirnov tests (1D marginal along  $\phi$ :  $\chi^2$ -test  $p$ -value = 0.0847; 1D marginal along  $z$ :  $\chi^2$ -test  $p$ -value = 0.00012, KS-test  $p$ -value = 0.008), where we also included the nonlinearity of the surface metric element of the cone. Altogether, these results show that in the truncated cone, the surface charge density  $\sigma_{\text{cone}} = \sigma_{\text{cone}}(z)$  depends on  $z$ . Around 50%–60% of the phosphate moieties of the cone surface are at the edge of the exit port, in the last 3 Å interval  $z = [-101.80 \text{ \AA}, -98.80 \text{ \AA}]$  of the cone, showing a highly significant enrichment of the phosphates at the exit port where the tunnel radius is the largest.

The potential created by the observed structural distribution of charges as mapped in Fig. 4 along the tunnel centerline is computed in the Appendix to see how it compares with the inverse bell shape results from a uniformly charged tunnel. The constraints of the observed potential data point measurements on the heterogeneity of the phenomenological parameters (dielectric response and electrostatic screening lengths) are also studied and compared in the Appendix A.5.

Figures 4(c) and 4(d) show the maps of the positions of the charged atoms carried by the charged amino acid residues R, K in red and D, E in blue within a distance of 8 Å from the model wall surface. The markers' shapes determine which ribosomal protein these amino acids belong to: squared markers for ribosomal protein uL22, starred markers for uL4, and filled circled markers for eL39 and L24. The threshold distance of 8 Å was adopted for the amino acid residue selection because it is the largest atomic Euclidean length between the side chain distal charged nitrogen (called NH2 in the atomic model PDB format) and the proximal carboxylic carbon in the peptide bond for the arginine amino acid residue. Extending the threshold distance to 12 Å for the amino acid does not affect the maps significantly (not shown here). Figures 7(a) and 7(b) highlight the relative positions of these seven positively charged residues with respect to the tunnel surface. There are no charged amino acid closer than 8 Å to the tunnel surface in the upper tunnel. The salient feature of the constriction site is that there are exactly seven positively charged residues belonging to uL4 (Arg65, Lys72, Arg76, Arg78) and uL22 (Arg125, Arg128, Arg132), with their center of charges at  $z$ -position =  $-40.53 \text{ \AA}$  which is only 3 Å away of  $z_0 = -37.5 \text{ \AA}$  which was the independently fitted location of the Lorentzian peak for the electrostatic potential (see Sec. II E). There is only a single Arg35 (eL39) in the lower tunnel. The average closest distance of the charged atoms from the cylinder tunnel wall is 4.56 Å.

All the five positively charged amino acid residues within close distance of the vestibule tunnel wall are located near the end edge of the cone, with two of them in an enriched hydrophilic region, Arg31 (eL39) and Lys81 (L24).

### B. Dielectric response of the medium prevailing in the tunnel and mobile ions

The medium inside the ribosome wall is of course a dielectric and not a conductor and not free space (vacuum). In formula (2), we see that the potential on the cylinder axis depends on the geometry and surface charge density. The  $\epsilon$  parameter is the medium dielectric permittivity

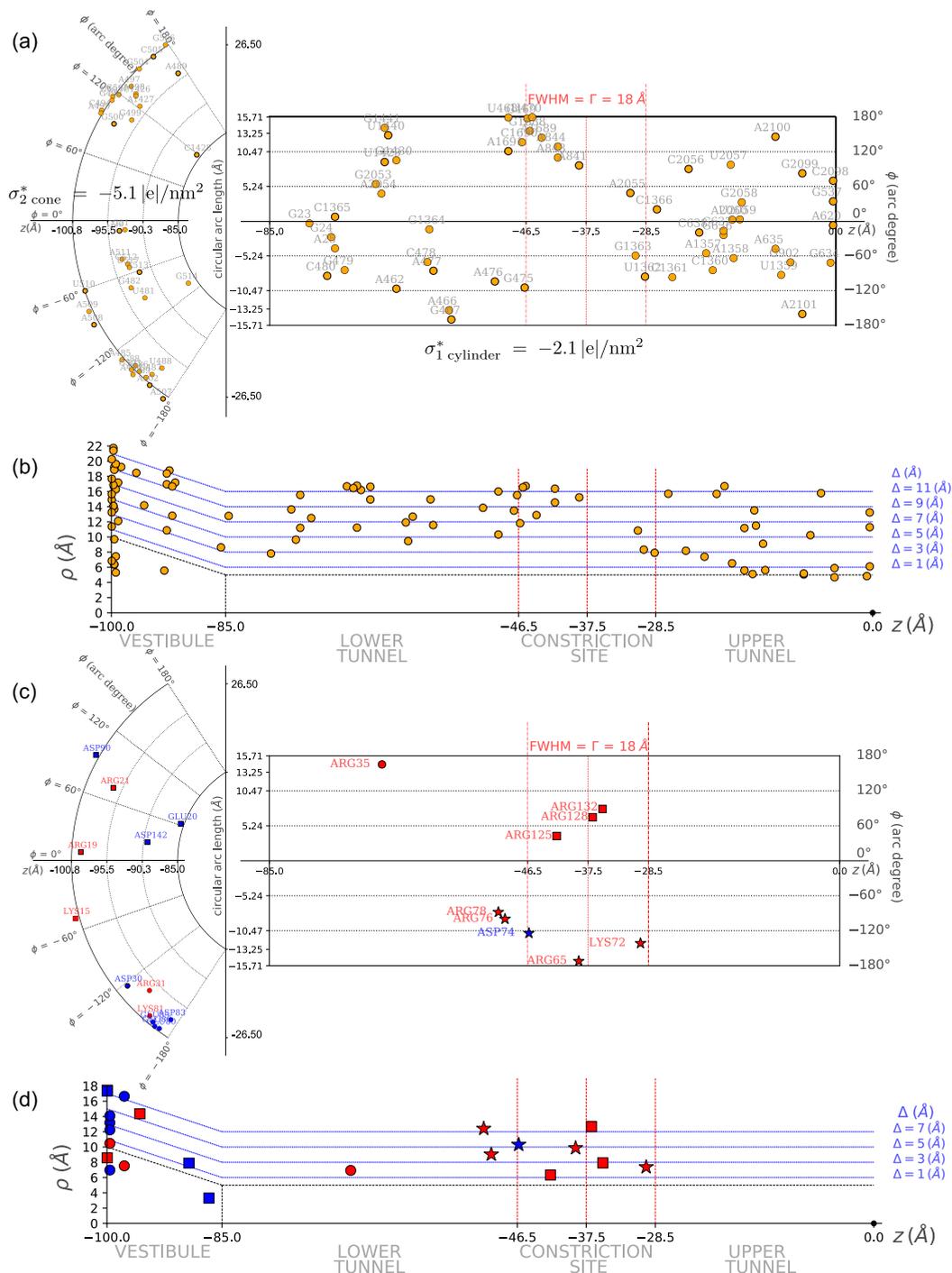


FIG. 4. 2D Mapping of phosphorus atoms and charged amino acid residues close to the ribosome exit tunnel wall surface. Calculated wall surfaces:  $8.95 \text{ nm}^2$  (upper tunnel),  $5.65 \text{ nm}^2$  (constriction site),  $12.09 \text{ nm}^2$  (lower tunnel), and  $7.45 \text{ nm}^2$  (vestibule). Position of the constriction site with full width at half maximum of the Lorentzian peak (dashed red vertical lines). (a) 2D mapping of phosphorus atoms in the phosphate moieties (orange disks) on the tunnel cylindrical surface (right) and tunnel vestibule cone frustum surface (left). Bare surface charge density on cylindrical surface  $\sigma_1^* = -2.1 |e|/\text{nm}^2$ . Bare surface charge density on cone frustum surface  $\sigma_{2(\text{cone})}^* = -5.1 |e|/\text{nm}^2$ . Horizontal axis:  $z$  distance along tunnel centerline, vertical axis, and arc axis: angle to atom line of sight as viewed from centerline (see text for orientation reference frame). (b) Mapping of the phosphorus atoms (orange disks) closest to the tunnel centerline. Horizontal axis:  $z$  distance along tunnel centerline; vertical axis: radial distance from tunnel centerline. (c) 2D mapping of positively charged (red squares) and negatively charged (blue stars) amino acid residues on the tunnel cylindrical surface (right) and tunnel vestibule cone frustum surface (left). (d) Mapping of the positively charged (red squares) and negatively charged (blue stars) amino acid residues closest to the tunnel centerline. Note the enrichment of seven positively charged amino acids at the constriction site.

with  $\epsilon = \epsilon_r \epsilon_0$ . The  $\epsilon_0$  parameter is the vacuum's permittivity:  $8.854 \times 10^{-12}$  Farad/m. The formula applies to the real dielectric medium prevailing inside the ribosome exit tunnel. The ionic strength inside the tunnel and the electric polarizability of all the molecules inside the tunnel are also expected to play a role that is further discussed below.

The volume of a cylinder of length 85 Å and radius 5 Å is  $V_{\text{cyl}} = 6676 \text{ \AA}^3$ . The volume of the cone frustum (tunnel vestibule) is  $V_{\text{cone}} = 3312 \text{ \AA}^3$ . The total volume inside the tunnel is  $V_{\text{tot}} = 9988 \text{ \AA}^3$ . The volume of an individual water molecule is estimated to be  $v_W = 29.9 \text{ \AA}^3$  [47]. Hence, the estimated maximum number of water molecules that could be accommodated in an otherwise empty ribosome exit tunnel is around 333 (223 water molecules in the cylinder and 110 in the cone frustum). In their paper, Dao Duc *et al.* [39] estimated the volume of the exit tunnel in eukarya to be on average  $27.8 \times 10^4 \text{ \AA}^3$ , which is 2.8 times as much as our calculation. This would amount to a maximum number of water molecules of  $\sim 930$  in an otherwise empty exit tunnel (including the large cavity around the PTC). In their paper, Dao Duc *et al.* included the volume of the large cavity between the PTC and the tunnel entry point and this might be the origin of the discrepancy with our estimation. Besides, the tunnel searching algorithm being used may or may not remove solvent water molecules and does not discriminate between the mobile water molecules and the wall constitutive water molecules, increasing the tunnel apparent measured volume. Otherwise the tunnel length and radius averaged values in the Kao Duc *et al.* paper are very similar as the ones used in our calculation. When a nascent polypeptide is progressing through the exit tunnel, the volume left for mobile water molecules is significantly reduced. So the real number of water molecules present in the tunnel during polypeptide elongation is believed to be of the order of  $\sim 100$  and most of them probably present in the vestibule region (tunnel exit end). Water as a bulk solvent has a relative electric permittivity of 78 (25 °C) or 74 (37 °C), while experimental and theoretical evidence suggest that proteins (or the nascent polypeptide) have an average dielectric response that can be approximated with a dielectric constant of about three to four; see [48] and references therein. The dielectric constant of nucleic acids like DNA in bulk solution has been measured to be around 8 [49]. Thus, depending on the abundance of water molecules in the tunnel volume accommodating the nascent chain, the tunnel microenvironment cannot be viewed as uniform. At least two dielectric constants  $\epsilon$  should be used in a range from  $\epsilon = 3\text{--}4$  (polypeptide) to  $\epsilon = 78$  (water). It follows that the dielectric medium permittivity is not uniform along the tunnel axis nor across the tunnel wall and is not necessarily equal to the one of water in bulk solution:  $\epsilon$  could be higher in the vestibule region (more water molecules) than in the narrower section of the tunnel (less water molecules). In our model settings, two different values can be used for  $\epsilon$ : one in the cylinder section (upper and lower tunnel) and another one in the vestibule region. A third permittivity value can also be implicitly used as well in the tunnel constriction through the scale factor  $f_{\text{scale}}$  (see Sec. II E).

Ionic strength and mobile ions effects are very important to study the behavior of biological macromolecules in aqueous bulk solution at equilibrium [48]. The inside of the ribosome

exit tunnel is however too narrow to be considered as a plain aqueous bulk solution. The inside of the ribosome exit tunnel is also far from equilibrium during the dynamical process of protein elongation. The hydrated radius of magnesium  $\text{Mg}^{2+}$  [naked ionic radius of 0.65 Å but with octahedral coordination geometry in  $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ ], calcium  $\text{Ca}^{2+}$ , potassium  $\text{K}^+$ , and sodium  $\text{Na}^+$  ions in bulk aqueous solution at equilibrium are 0.7–0.8 nm, 0.6 nm, 0.3 nm, 0.4–0.45 nm, respectively [50]. The ribosome exit tunnel radius is  $\sim 0.5$  nm in the cylinder, but the tunnel vestibule exit opening is 1 nm in radius. Only monovalent ions and possibly magnesium that has a great affinity for nonbridging phosphate oxygen atoms could be accommodated in the tunnel cylinder region and probably only in the absence of a nascent polypeptide growing inside the tunnel. If there were mobile ions, they would be expected only in the vestibule region of the tunnel. A number of publications have addressed the issue of the presence of metal ions inside the ribosome and particularly underscored the importance and roles of magnesium and potassium as phosphates counter ions stabilizing secondary and tertiary structures of both rRNA and tRNAs, respectively [51–54]. Klein *et al.* [54] numbered 116  $\text{Mg}^{2+}$  ions in total inside the 50S large ribosomal subunit of *H. marismortui* and 88 monovalent ions in total. Most of these metal ions usually bind to the nonbridging phosphate oxygen atoms of the RNA phosphodiester backbone or to specific parts of RNA secondary structure motifs or to highly idiosyncratic sites in 23S RNA. This number of metal ions is at least one order of magnitude lower than the number of phosphate moieties in the rRNAs encountered in the 50 or 60S ribosomal subunit. We could not find supporting published evidence of the presence of (mobile or free) magnesium or potassium ions in abundance inside the upper or lower tunnel during protein synthesis *in vivo*. Furthermore, the very notion of ionic concentration or ionic strength is less clear in the tunnel confined environment where the total number of water molecules is limited and of magnitude order of 100 molecules.

Moreover, as we analyzed above in Sec. III A with PyMOL from the PDB entry id 4V9F of the x-ray solved structure of the 50S subunit of *H. marismortui* at 2.4 Å [42], the total number of water molecules was 7745 in the 50S structure. The number of water molecules found inside the tunnel was 44. The number of Mg atoms in the 50S subunit was 138, the number of Mg atoms inside the tunnel was three at a distance of 12 Å from the tunnel centerline. The total number of Na atoms was 85, three of which were found at a distance of 12 Å from the tunnel centerline. The total number of potassium atoms was three, none of which in or near the tunnel wall. Altogether, the crystallographic data do not support evidence of the presence of free mobile ions in number inside the ribosome exit tunnel.

During the peptide bond formation in the catalytic center of the large ribosomal subunit at the peptidyltransferase center (PTC), the precise positioning [55,56] of the local rRNA, peptidyl-tRNA at the P-site, aminoacyl-tRNA at the A-site and of a possible single solvation water molecule involved in the proton shuttle occurring during the catalytic deacylation and transpeptidation, the introduction of water molecules and or (large) ions such as  $\text{Mg}^{2+}$  or  $\text{K}^+$  or even  $\text{Na}^+$  is sterically hindered (if not impossible). Dynamically, during the pep-

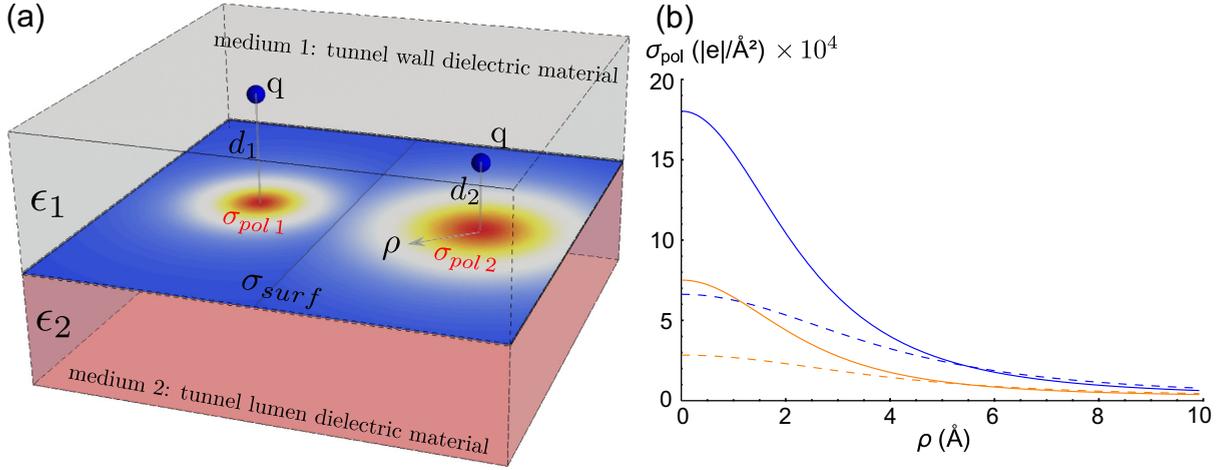


FIG. 5. Polarization surface charge densities on tunnel wall due to discontinuity between dielectric media. (a) Two charges (blue spheres) buried in dielectric medium 1 and their polarization effects at the boundary with dielectric medium 2. (b) Surface polarization charge density  $\sigma_{\text{pol}}$  as a function of the charge burial depth  $d$  in medium 1 and the tangential distance  $\rho$  at the interface. Significant dielectric discontinuity:  $\epsilon_1 = 8$  (rRNA) and  $\epsilon_2 = 78$  (water) for  $d_2 = 3 \text{ \AA}$  (blue line),  $d_1 = 5 \text{ \AA}$  (dashed blue). Weak dielectric discontinuity:  $\epsilon_1 = 8$  (rRNA) and  $\epsilon_2 = 4$  (polypeptide),  $d_2 = 3 \text{ \AA}$  (orange line),  $d_1 = 5 \text{ \AA}$  (dashed orange).

tide elongation process and the progress of the nascent chain into the ribosome exit tunnel from the PTC, it is expected that only aminoacid residues can enter the tunnel and there is hardly space left for third-party water molecules or large ions. It is believed that only small  $\text{H}^+$  or  $\text{OH}^-$  (having the highest electrochemical mobility) could be the counter ions accompanying charged amino acid residues. In any case, the elongation process itself prevents the environment inside the ribosome exit tunnel to be in equilibrium with external aqueous ions. Due to the dynamics of elongation, the ribosome exit tunnel interior is a medium far from equilibrium. For all these reasons, our model neglected such effects as mobile ions and ionic strength. It is possible, however, that mobile ions and ionic strength effects could be present in the vestibule (exit end of the tunnel). The Poisson-Boltzmann theory of charged macromolecules in bulk solution with strong electrolytes or its Debye-Hückel linearized version are not fully appropriate in the context of the ribosomal exit tunnel as their basic assumptions are not met in the very confined microenvironment of the tunnel.

However, a strong screening effect due to the permanent dipole moments of the water molecules that are buried inside the tunnel wall medium and close to the phosphate moieties has to be taken into consideration as will be studied under the Gouy-Chapman approach and shown in Fig. 6(d) below.

### C. Polarization surface charge densities on the tunnel wall due to discontinuity between dielectric media:

#### Image charge effect with dielectrics

Polarization charge surface densities  $\sigma_{\text{pol}}$  appear at media discontinuities between two material with different dielectric constants. This is a well-known classical effect in electrostatics that can be treated in the framework of boundary-value problems with multiple dielectrics; see Jackson [40], chap. 4.4. The polarization charge density appearing on the surface boundary separating the two dielectric media is expressed in

formula (4.47) in Ref. [40]:

$$\sigma_{\text{pol}} = -\frac{q}{2\pi} \frac{\epsilon_0(\epsilon_2 - \epsilon_1)}{\epsilon_1(\epsilon_2 + \epsilon_1)} \frac{d}{(\rho^2 + d^2)^{3/2}}. \quad (21)$$

We consider here a radial discontinuity at the tunnel's inner surface, with dielectric constant  $\epsilon_2$  inside the lumen of the tunnel and dielectric constant  $\epsilon_1$  for the tunnel wall material as shown in Fig. 5(a). When applying formula (21), a fixed phosphate charge moiety is supposed to be in medium 1 at distance  $d$  from the boundary surface separating the media as shown in Fig. 5(a). If the fixed charged phosphates moieties are deeply buried in the tunnel inner wall, two situations should be discussed depending on the dielectric response differences between the media and the importance of the discontinuity.

(1) Case 1:  $\epsilon_2 = \epsilon_{\text{water}} = 78$  and  $\epsilon_1 = \epsilon_{\text{rRNA}} \approx 8$ . The discontinuity is significant.  $\epsilon_2 \gg \epsilon_1$ , the dielectric  $\epsilon_2$  (tunnel lumen) behaves much like a conductor in that the electric field inside it becomes very small and the surface charge density approaches the value appropriate to a conducting surface, apart from a factor of  $\epsilon_0/\epsilon_1$ . The polarization surface charge density has a sign opposed to the buried source fixed charge. The polarization charge partially screens the buried fixed charge:

$$\sigma_{\text{pol}}(d, \rho = 0) = -\frac{q}{2\pi} \frac{\epsilon_0}{\epsilon_1} \left(\frac{70}{86}\right) \frac{1}{d^2}. \quad (22)$$

(2) Case 2:  $\epsilon_2 = \epsilon_{\text{peptide nascent chain}} \approx 4$  and  $\epsilon_1 = \epsilon_{\text{rRNA}} \approx 8$ . The discontinuity is smoother than in the previous case. The polarization surface charge density that appears at the boundary is 2.44 times smaller in absolute value than in the previous case:

$$\sigma_{\text{pol}}(d, \rho = 0) = -\frac{q}{2\pi} \frac{\epsilon_0}{\epsilon_1} \left(\frac{-4}{12}\right) \frac{1}{d^2}. \quad (23)$$

Finally, if the charged phosphate moieties are localized exactly on the surface of the inner wall ( $d = 0$ ), no polarization

appears and  $\sigma_{\text{pol}} = 0$ . Our model relies on the assumption that the negative charges of the phosphate moieties are indeed exposed in numbers at the immediate vicinity of the surface of the wall and are carried by the nonbridging oxygen atoms in the phosphodiester bonds between riboses. A fraction of the inner negative charges buried deeper inside the rRNA material are neutralized by constitutive metal ions like magnesium and sodium as discussed previously or by charged residues of ribosomal proteins. The rest of the inner negative charges buried inside the rRNA that are possibly fixed and not screened directly by counter cations (positively charged amino acid residues from ribosomal proteins or metal cations) will generate local polarization charge densities on the boundary surface between the two dielectric media as calculated above. This polarization due to dielectric discontinuities contributes to the screening of the phosphate moieties.

Media discontinuity surface polarization effects do occur in the ribosome exit tunnel. They are believed to be local and would not affect, on a large scale, the profile of the electrostatic potential along the tunnel length. Note that when moving tangentially along the boundary surface, the polarization surface charge density decreases according to the third power of the tangential distance  $\rho$ . Polarization due to dielectric discontinuities is local and fades away rapidly:  $\sigma_{\text{pol}}$  is 7.34 times smaller at  $\rho = 5 \text{ \AA}$  than at  $\rho = 0 \text{ \AA}$ ; see Fig. 5(b).

We did not explicitly incorporate in our simplified electrostatic model such medium discontinuity surface polarization effects. We took the approach of estimating an *apparent* or *renormalized* fixed negative charge surface density on the tunnel wall (see the Gouy-Chapman screening length approach in next subsection). It is not established whether this net surface charge density  $\sigma$  results from the direct contribution of fixed negative charges directly lying on the surface, i.e.,  $\sigma_{\text{surf}}$ , or from surface polarization effects due to dielectric media discontinuities and due to the charges buried deeper inside the wall material, i.e.,  $\sigma_{\text{pol}}$ . The charged surface density  $\sigma$  that is explicitly used in all our formulas should be considered as a resulting net *apparent* surface charge density  $\sigma_{\text{app}}$  where

$$\sigma = \sigma_{\text{app}} = \sigma_{\text{surf}} + \sigma_{\text{pol}}. \quad (24)$$

Interestingly, a decrease in the ratio  $\sigma/\epsilon$  in the tunnel vestibule (cone frustum) as compared to the ratio  $\sigma/\epsilon$  in the lower or upper tunnel (cylinder) could result from a stronger discontinuity in the dielectric constants between the tunnel inner wall material and the tunnel lumen material. Indeed, if the dielectric constant  $\epsilon$  in the lumen of the vestibule is higher due to the presence of more water molecules there than in the upper tunnel cylinder, the polarization surface charge density on the vestibule wall would be higher and would better screen out the buried negative source charges. The apparent surface charge density would be smaller. Besides, a larger  $\epsilon$  value in a more aqueous medium would also contribute to decrease the ratio  $\sigma/\epsilon$  as compared to a medium containing a nascent polypeptide without water molecules.

#### D. Water screening inside the tunnel wall medium

Electrostatic interactions in ribonucleic structures are potentially quite strong, but these interactions are mitigated by the screening effects of water or nearby protein atoms

[57], even in the absence of mobile ions. The screening of electrostatic interactions results primarily from electronic polarization, reorientation of dipolar groups in the vicinity of charges and dipoles. These effects are well understood and can be accurately determined for interactions in isotropic, homogeneous media. However, in complex inhomogeneous environments such as those near the surface of ribonucleoproteins, dielectric screening is difficult to predict. In the case of the ribosome exit tunnel, the confined geometry and composition of the inner wall close to the tunnel surface and whether the interactions involve direct charges or dipoles are expected to be especially important. The x-ray solved atomic space positions in the immediate 8 to 12  $\text{\AA}$  vicinity of the tunnel wall show that water molecules are indeed good candidates to explain the screening of the *formal bare* charges carried by the nonbridging oxygen atoms bound to the phosphorus atoms as shown in Figs. 6(a) and 6(b). We investigated with PyMOL scripts the number of water molecules (only oxygen atoms of solvent are seen in x-ray crystallography) that are within less than 4, 5, and 6  $\text{\AA}$ , respectively of the 94 phosphorus atoms around the tunnel wall and obtained the following water molecule counts: 110, 208, and 304. It should be noted here that the total number of water molecules sufficiently well organized to be observed by x-ray crystallography inside the LSU is 7745 and that the total number of water molecules observed inside the tunnel is 44 as shown on Fig. 6(c). Figure 6(b) shows the abundance of water molecules in the immediate vicinity of the phosphorus atoms in 23S rRNA near the ribosome tunnel wall. Almost no mobile ions were observed in this region in the x-ray solved atomic structure: only three  $\text{Mg}^{2+}$  were found inside the tunnel, no  $\text{K}^+$ , and only three  $\text{Na}^+$  within a 12  $\text{\AA}$  distance of the 94 selected phosphorus atoms. These results support the assumption that the dominant electrostatic screening is due to constitutive water molecules that are buried in the inner wall medium of the tunnel; each phosphate moiety being surrounded by one to three water molecules in the immediate vicinity of the tunnel wall surface.

#### 1. Gouy-Chapman screening length and charge renormalization incorporating electrostatic screening of buried phosphate moieties due to solvent water molecules inside the tunnel wall medium

Due to the dissociation (charge regularization) of surface groups, the rRNA phosphate moieties support surface acquires a net surface charge density that we call  $\sigma^*$  as shown in Fig. 6(d). These *bare* charges do not stay unbalanced due to a screening effect involving water solvent. The water molecules dipole moments reorient so that a layer of positively charged hydrogens oppose the negatively charged phosphate moieties. The solvent water oxygen atoms also form the next layer and so on. The Poisson-Boltzmann-like derivation in the case of a uniformly charged single planar surface provides a far field solution for the screened electrostatic potential which is equivalent to a Yukawa-Debye-Hückel potential:

$$\Phi_{\text{Yuk}}(\vec{\mathbf{r}}) = \iint_S \frac{\sigma^*(\vec{\mathbf{r}}') da}{4\pi \epsilon \epsilon_0} \cdot \frac{e^{-\frac{|\vec{\mathbf{r}}-\vec{\mathbf{r}}'|}{\xi}}}{|\vec{\mathbf{r}}-\vec{\mathbf{r}}'|}. \quad (25)$$

This is equivalent to a marked exponential damping of the Coulomb interaction where  $\xi$  is a characteristic distance of

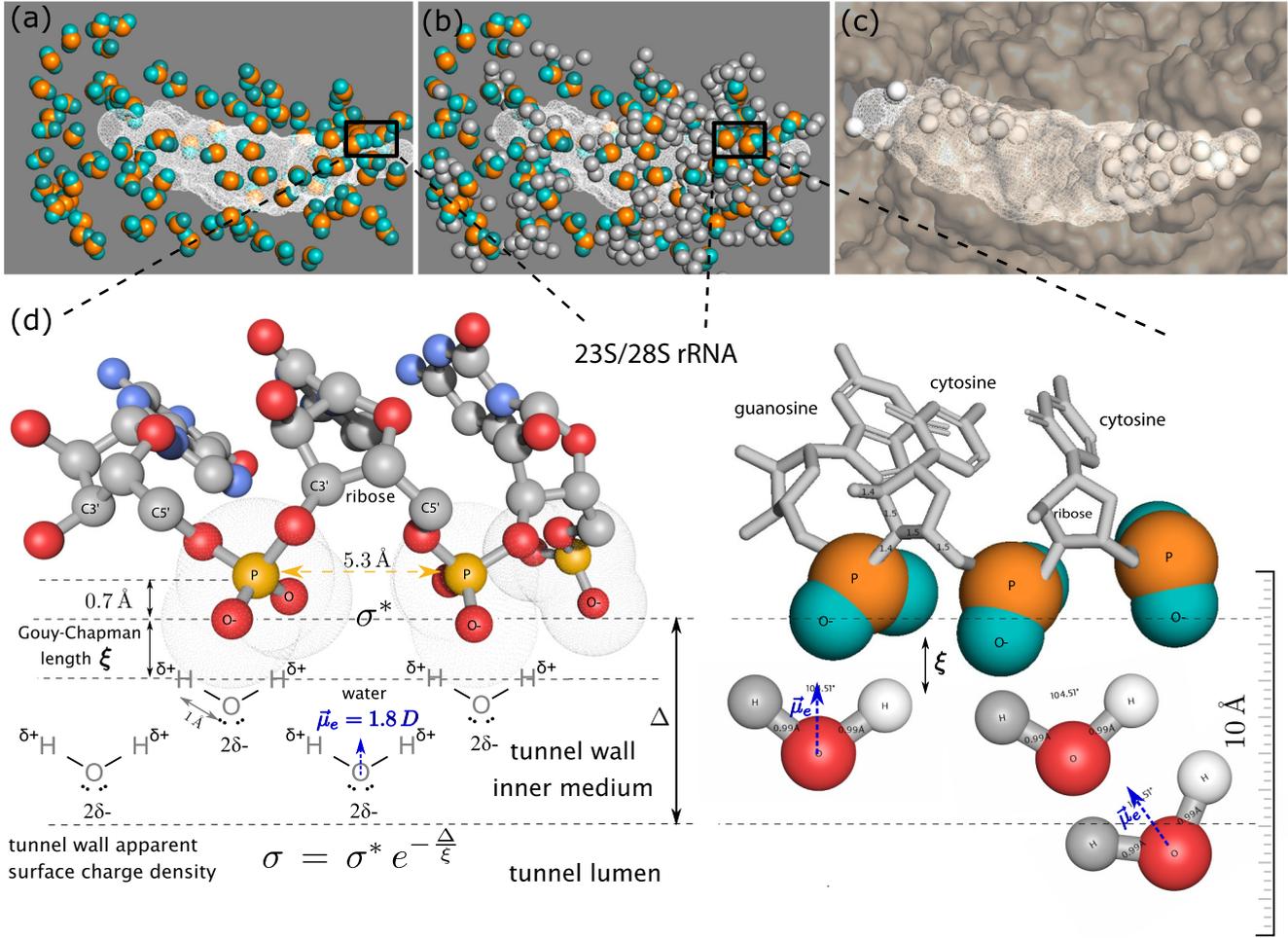


FIG. 6. (a) Representation of the 94 phosphorus atoms (orange spheres) within  $12 \text{ \AA}$  of the tunnel surface (white mesh) with their nonbridging oxygen atoms (teal spheres); (b) 304 oxygen atoms (gray spheres) of water molecules within  $6 \text{ \AA}$  of the 94 phosphorus atoms; (c) 44 oxygen atoms (white spheres) of water molecules within  $5 \text{ \AA}$  of the tunnel centerline; (d) electrostatic screening effect due to water molecules. Phosphate moieties carry bare negative charges reorienting the dipole moment of water molecules  $\vec{\mu}_e$ , shown in blue. The *bare formal* charge surface density  $\sigma^*$  is screened beyond the Gouy-Chapman length  $\xi$  resulting in the *apparent* or *renormalized* charge surface density  $\sigma$  on the tunnel wall. The *bare formal* charge surface density  $\sigma^*$  is screened beyond the Gouy-Chapman length  $\xi$  resulting in the *apparent* or *renormalized* charge surface density  $\sigma$  on the tunnel wall. The cartoon was produced with PyMOL extracted data from PDB entry code 4V9F [41] and ChemDoodle 3D [58].

the exponential screening. If we assume the screening effect to be uniform along the charged surface, we can recast the far-field potential as the result of an *effective* or *renormalized* or *apparent* surface charge density  $\sigma$  that is different from the actual *bare formal* charge  $\sigma^*$ . The “mapping” between the *bare* and the *renormalized apparent* charge density is simply

$$\sigma = \sigma^* e^{-\Delta/\xi}, \quad (26)$$

where  $\Delta$  is the distance between the original bare charged wall and the support wall of the electrostatic surface as obtained within nonlinear screening theory; see Fig. 6(d). The actual *bare formal* surface charge  $\sigma^*$ , when screened by water molecules, results in an *apparent* or *renormalized* surface charge density  $\sigma$  [59]. Unless otherwise stated in this paper,  $\sigma$  always refers to the *apparent* charge surface density obtained within nonlinear screening of water molecules present in the inner wall of the tunnel. Selecting the most appropriate screening theory reduces to knowing which length scale parameter  $\xi$  to use. Three length scales, i.e., the Bjerrum length

( $\lambda_B$ ), the Debye length ( $\lambda_D = \kappa^{-1}$ ), and the Gouy-Chapman length ( $\xi_{GC}$ ) deserve specific attention as highlighted by Van Roij [59]. The Gouy-Chapman length results from the fact that the Bjerrum length multiplied by the number of charges per surface unit in a uniformly charged wall has the dimension of an inverse length and comes from a boundary condition in the Poisson-Boltzmann equation applied to a uniformly charged planar wall complying with the charge neutrality condition [59]. By convention, this so-called Gouy-Chapman length is given by

$$\xi = \frac{1}{2\pi \lambda_B \sigma^*} \quad (27)$$

and is a property of the uniformly charged wall, as  $\xi$  denotes the length scale over which the potential due to the charged wall, in the absence of screening ions but in the presence of the solvent (characterized by  $\lambda_B$ ) equals the agitation thermal energy  $k_B T$ . The atomic x-ray solved structure of the region around the ribosome exit tunnel showed that the wall buried

TABLE I. Screening lengths for three screening theories for different dielectric response values at standard temperature  $T = 298.15$  K.

		Bjerrum length $\lambda_B$	Debye length <sup>a</sup> $\kappa^{-1} = \lambda_D$	Gouy-Chapman length per unit charge <sup>b</sup> $\xi$
Medium		$\lambda_B = \frac{e^2}{4\pi \epsilon \epsilon_0 k_B T}$	$\kappa^{-1} = \left(\frac{2e^2 I N_A}{\epsilon \epsilon_0 k_B T}\right)^{-1/2}$	$\xi = \frac{1}{2\pi \lambda_B \sigma^*}$
Vacuum	$\epsilon = 1(\epsilon_0)$	56 nm	—	—
Protein	$\epsilon = 4$	14 nm	0.18 nm	—
rRNA	$\epsilon = 8$	7 nm	0.25 nm	—
Water	$\epsilon = 78$	0.72 nm	0.78 nm	0.105 nm

<sup>a</sup>Debye length calculated for an ionic strength of  $I = 0.15$  M (= 150 mole/m<sup>3</sup>).  $N_A$  is the Avogadro number.

<sup>b</sup>Gouy-Chapman length calculated for a *bare formal* surface charge density of  $\sigma^* = 2.1|e|/\text{nm}^2$ .

phosphate moieties are surrounded by one to three water molecules and no mobile ions as shown in Fig. 6(b). These experimental facts support the Gouy-Chapman approach for the electrostatic screening in the particular confined microenvironment of the ribosome exit tunnel. The Gouy-Chapman length as shown in Fig. 6(d) can physically be interpreted as the separation between two charged layers. The first wall layer is made of the nonbridging oxygen atoms that carry the negative *formal bare* charges in the phosphodiester backbone of the 23S rRNA close to the tunnel wall as shown in Fig. 6(d). The second layer is made of the hydrogens belonging to the oriented water molecules having their dipole moment pointing towards the first layer. Table I shows how the screening length scales compare in the three screening theories.

The atomic positions mapped on the tunnel surfaces, built from high-precision x-ray solved structures of real ribosomes set experimental constraints on the phenomenological constants  $\sigma$ ,  $\epsilon$ , or their ratio  $\sigma/\epsilon$ , that are complementary to the experimental electrostatic measurements in the ribosome exit tunnel. Altogether, these experimental constraints allow to assess the model assumptions on the piecewise heterogeneity for  $\sigma$ ,  $\epsilon$  along the tunnel length as well as on the particular origin of electrostatic screening effects inside the ribosome exit tunnel. Indeed, taking as screening length scale, the Gouy-Chapman screening length value, i.e.,  $\xi = 1.05$  Å, and taking for  $\Delta = 5.97 = 6.67 - 0.7$  Å, the difference of the average distance between the closest phosphorus atoms to the model cylinder wall surface (6.67 Å), and the distance (0.7 Å) between the phosphorus atoms and the nonbridging oxygen atoms in the phosphodiester bonds with the riboses, the *apparent normalized* surface charge density is

$$\begin{aligned} \sigma_1 &= \sigma_1^* e^{-\frac{\Delta}{\xi_{GC}}} = -2.1 e^{-\frac{5.97}{1.05}} = -2.1 \cdot 0.00339 \\ &= -0.00712 |e|/\text{nm}^2 = -1.141 \text{ mC/m}^2. \end{aligned} \quad (28)$$

Substituting this numerical value in the formula (2) of the potential for the cylinder, with  $z = -50$  Å,  $R = 5$  Å and taking the two extreme assumptions on the dielectric response in the tunnel lumen  $\epsilon = 4$  (nascent protein occupying the tunnel and no water molecule) or  $\epsilon = 78$  (water molecules only occupying the tunnel), we can get an estimate for the range of the electrostatic potential  $\Phi(z)$  expected values at the centerline of the tunnel at position  $z = -50$  Å where the potential reaches its maximal negative value.

With  $\epsilon = 4$  (nascent chain only occupies the tunnel):

$$\begin{aligned} \Phi(z) &= \frac{\sigma_1 R}{2 \epsilon_1 \epsilon_0} \log \left| \frac{7 + \sqrt{7^2 + 1}}{-10 + \sqrt{101}} \right| \\ &= \frac{\sigma_1 R}{2 \epsilon_1 \epsilon_0} \log 282.12 \\ &= \frac{-1.141 \times 10^{-3} \times 5 \times 10^{-10}}{2 \times 4 \times 8.85 \times 10^{-12}} \times 5.6423 \\ &= -0.0455 \text{ V}. \end{aligned} \quad (29)$$

The expected value of the electrostatic potential would be  $-45.5$  mV at the tunnel center if the dielectric response of the medium in the lumen is at the lower limit  $\epsilon_1 = 4$  (nascent protein).

With  $\epsilon = 78$  (water molecules only in the tunnel):

$$\begin{aligned} \Phi(z) &= \frac{-1.141 \times 10^{-3} \times 5 \times 10^{-10}}{2 \times 78 \times 8.85 \times 10^{-12}} \times 5.6423 \\ &= -0.0023 \text{ V}. \end{aligned} \quad (30)$$

The expected value of the electrostatic potential would be  $-2.3$  mV at the tunnel center if the dielectric response of the medium in the lumen is at the upper limit  $\epsilon_1 = 78$  (fully filled with water).

The observed value at the tunnel center is  $-22$  mV. Keeping the apparent charge density on the cylinder wall at  $\sigma_1 = -1.141 \times 10^{-3} \text{ C/m}^2$  (corresponding to a bare charge density  $\sigma_1^* = 2.1 |e|/\text{nm}^2$  for the cylinder as counted on the x-ray solved structure), the mean dielectric response of the medium in the cylinder lumen of the tunnel would be  $\epsilon_1 = 8.3$ . This would be consistent with a medium composition of 94% of nascent protein and 6% of water in the tunnel lumen.

To support the dimensional relation (14) and the stepwise values for the ratio  $\sigma/\epsilon$  in the different tunnel parts, we can estimate the local bare surface charge density  $\sigma_3^*$  from the ratio  $\frac{\sigma_3}{\epsilon_3}$  at the constriction site. In Sec. II E we inferred the ratio  $\frac{\sigma_3}{\epsilon_3}$  at the constriction site from the values of  $f_{\text{scale}}$  and  $\Gamma$ , experimentally fitted to the measured electrostatic potential. With  $R = 5$  Å,  $\sigma_3/\epsilon_3 = 5.73 \times 10^9 \text{ V/m}$ . For a dielectric constant prevailing inside the constriction site, i.e., the dielectric constant of the polypeptide nascent chain ( $\epsilon_r \sim 4$ ) and in the *absence of screening* water molecules in the confined region of the constriction site, we can estimate  $\sigma_3 = \sigma_3^*$ :

$$\sigma_3^* = \epsilon_3 r \epsilon_0 5.7 \times 10^9, \quad (31)$$

$$= 4 \times 8.85 \times 10^{-12} \times 5.73 \times 10^9, \quad (32)$$

$$= 0.2029 \text{ C/m}^2. \quad (33)$$

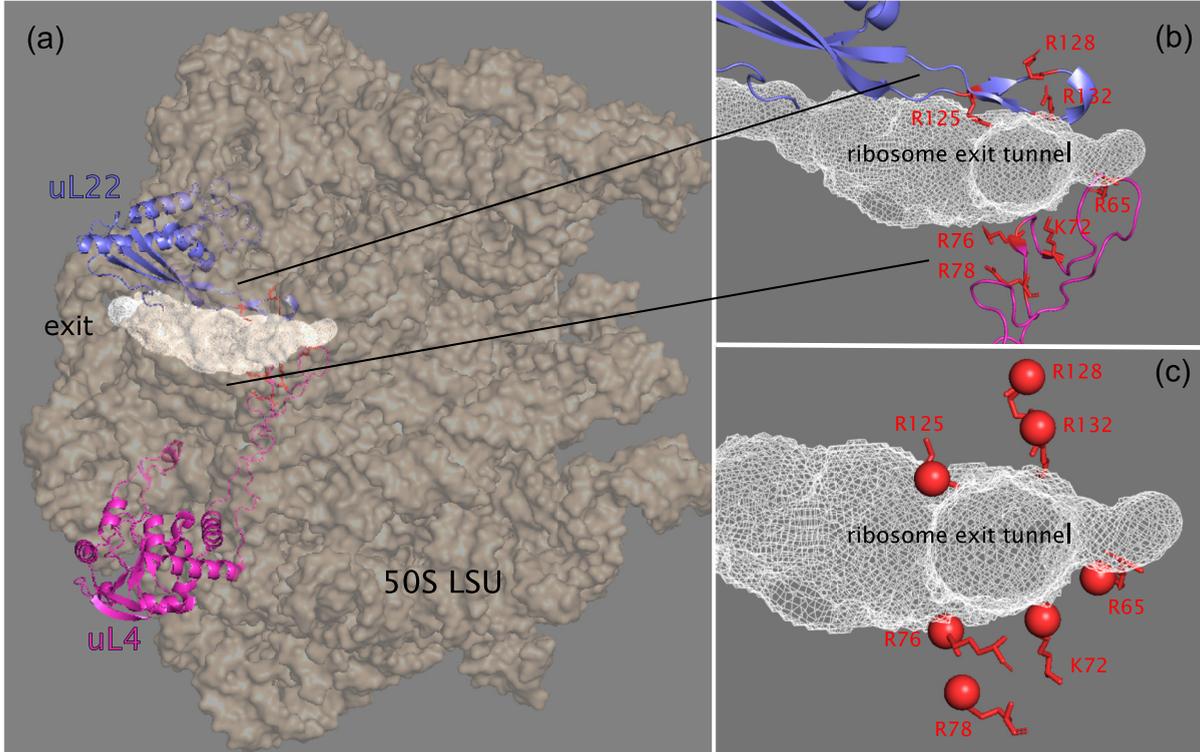


FIG. 7. (a) Ribosome exit tunnel in the 50S large subunit of *H. marismortui* showing two ribosomal proteins uL22 and uL4 and (b) their local protrusion with seven positively charged amino acid residues close to the constriction site; (c) the red spheres represent the nitrogen charged atoms bonded to the  $\epsilon$ -carbon of arginine (NH1 or NH2) or of lysine (NZ).

The lateral surface being  $S = 2\pi R\Gamma = 5.655 \text{ nm}^2$ , the bare formal charge in this local surface is estimated to be  $q_3 \sim 0.2029 \times 5.655 \times 10^{-18} = 1.14 \times 10^{-18} \text{ Coulomb}$  or approximately seven elementary charge units ( $|e| = 1.602 \times 10^{-19} \text{ C}$ ). The fitted values and the dimensional interpretation of section II E indicate that the number of apparent positive charges above the phosphate moieties baseline would be around seven, which is consistent with the number of the positively charged residues at the uL22 and uL4 protrusion (R or K residues in uL22 and uL4 near the tunnel wall) lining in the vicinity of the constriction site surface and within a distance of  $8 \text{ \AA}$  of the tunnel surface as shown in Figs. 4(b) and 4(c) and in Fig. 7.

#### IV. APPLICATION OF THE RIBOSOME EXIT TUNNEL MODEL

##### A. Computing the electrostatic interaction variations of the ribosome exit tunnel for different amino acid sequences in nascent polypeptides

Our established models expressed by Eq. (2), (13), or (18) can be used to quantitatively compare how difficult it is for the ribosome to push a nascent polypeptide chain inside and eventually out of the exit tunnel, depending on the amino acid primary sequence. If a peptide sequence is locally enriched in positively charged amino acids residues inside the tunnel and in negatively charged amino acid residues close to tunnel entry point, the axial forces required to push the nascent protein through the tunnel will be higher than for a peptide

composed of neutral amino acid residues or carrying only a single cluster of positively charged amino acid residues. The electrostatic potential well, locally trapping charged amino acid enriched peptides, needs to be overcome by other forces. These compensation mechanisms are exerted either by the ribosome itself or by third-party proteins with motor domains from specialized chaperone proteins exerting tugging forces outside of the ribosome. The elongation speed also has to be compatible with the decoding speed of the mRNA encrypted message which depends on the codon usage and on codon position autocorrelation, i.e., codon ordering allowing tRNA recycling (reusage of the same tRNA at successive encodings of the same amino acid can speed up translation or favor fidelity [60,61]). The elongation speed may also independently be impeded by downstream mRNA secondary structures [21,22,33,34].

The eukaryote ribosome exit tunnel can accommodate at least 40 amino acid residues and up to more than 70. It is known that the nascent polypeptide can start folding, i.e., finding its final secondary structure, inside the tunnel (and eventually tertiary 3D structure outside). Alpha helices have been shown to be present inside the tunnel close to its exit point. So a variable number of amino acids larger than 40 can actually be hosted inside the tunnel. Again, for the sake of simplicity, here we consider that the number of amino acid residues hosted inside the tunnel is exactly 40 and that the maximum number of amino acid residues that are under the electrostatic influence of the tunnel is exactly 50: five between the PTC and the tunnel entry, 40 in the tunnel, and five out of the tunnel.

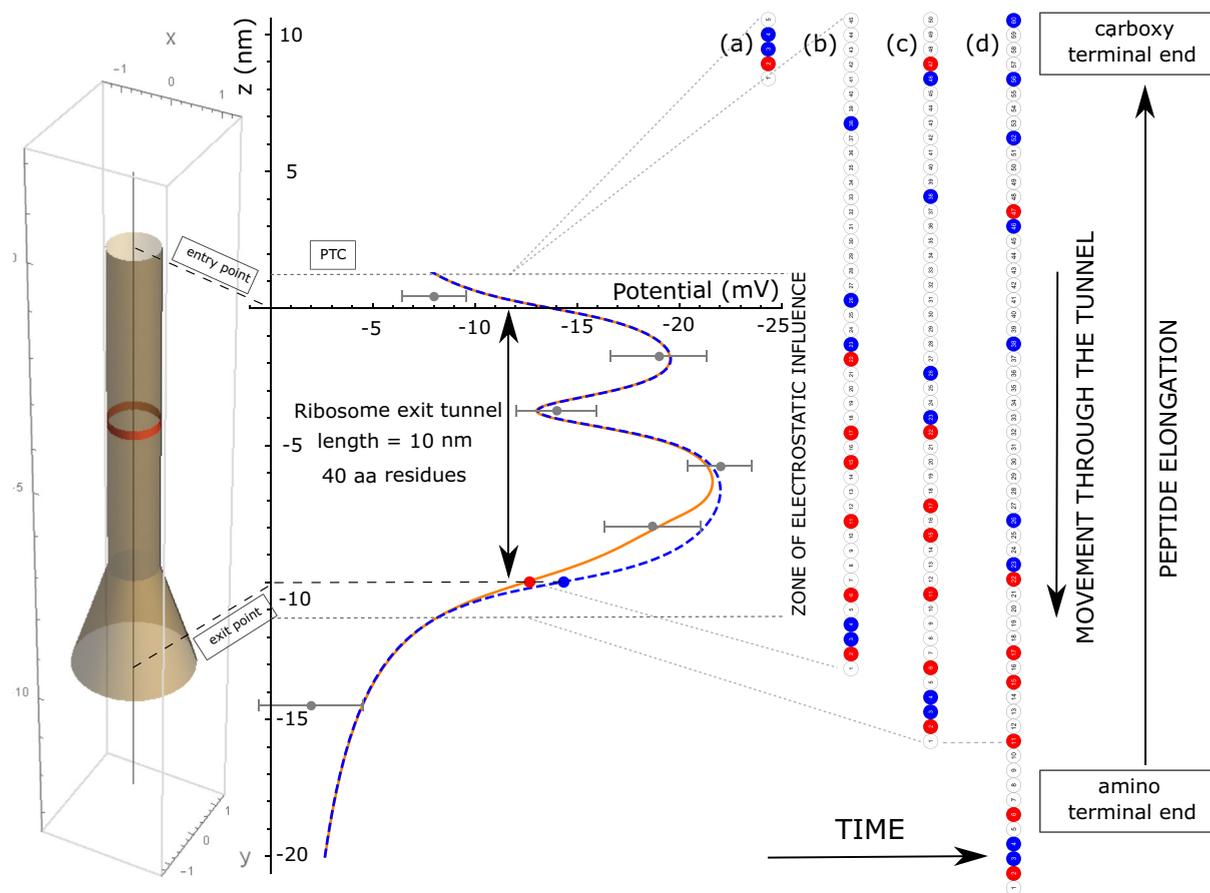


FIG. 8. Algorithm for computing the axial forces acting on a nascent peptide. Only the last 50 residues, at most, from the PTC are under the electrostatic influence of the ribosome exit tunnel. The nascent polypeptide is moving from the top to the bottom through the ribosome exit tunnel during elongation. (a) Five amino acid residues peptide with residue 5 at the PTC and residue 1 at the tunnel entry point. (b) 45 amino acid residues with residue 45 at the PTC and residue 1 at the tunnel exit point. (c) 50 amino acid residues with residue 50 at the PTC and residue 1 emerging out of the electrostatic influence zone. (d) 60 amino acid residues with residue 60 at the PTC and residues 1 to 10 out of the electrostatic influence zone.

The incorporation of a single amino acid to the nascent polypeptide chain takes place at the peptidyl transfer center (PTC), at the so-called P site of a translating ribosome. This PTC center is located around five amino acid residues away from the ribosome tunnel entry point. Stated otherwise, this means that the currently decoded codon, for which the cognate or semicognate aminoacylated tRNA, is five codons downstream the codon for which the amino acid is currently in the entry point of the tunnel. There are five amino acids bound in the oligopeptide part ready to enter the tunnel. Let us also assume that there are five bound amino acids out of the tunnel at the exit side that can feel the electrostatic influence of the tunnel. So, from the start codon (AUG coding for methionine), a nascent peptide starts with a five amino acid residues stretch elongating to the tunnel entry point, building up progressively to a 45 amino acids sequence fully accommodating the whole length of the tunnel, eventually extending to 50 amino acid residues being under a direct influence of the tunnel; see Fig. 8. For this 50-mer stretch to be out of the tunnel influence, another extra 50 amino acids have to be added to the carboxy terminal end of the nascent polypeptide.

Our aim is to compute the force profile and the mechanical power to be applied continuously on peptide stretches to overcome the electrostatic trapping interaction in the tunnel and to exit the ribosome tunnel. The easiest case scenario for computing would be when the downstream sequence is completely neutral. This is, of course, not always the case and the occurrence of charges in the downstream sequence plays a role that should not be neglected.

Multiplying the axial force acting on the stretches with the stretch axial displacement, i.e., the elongation distance towards the ribosome exit tunnel, yields the mechanical work that was delivered. Multiplying the axial forces acting on the stretches with the protein elongation rate, i.e., the speed of the ribosome along the transcript (mRNA) being translated, yields the required instantaneous net mechanical power.

The electrostatic axial force on the tunnel axis felt by an amino acid residue is the product of its net charge by the axial electric field [see Eq. (4)], the latter being the gradient of the electrostatic scalar potential, i.e., the first derivative of the potential with respect to the axial coordinate.

An important simplifying assumption is that all amino acid residues building up the nascent polypeptide are all rigidly

bound together and that the resulting nascent protein can be considered a single linear solid rigid body. This peptide, at least in the tunnel, is considered nondeformable. With this strong assumption, the axial forces individually computed for each charged amino acids act jointly and apply additively on the resulting rigid peptide body.

The local pH along the tunnel is unlikely to be out of the range 6–8 [62]. In this pH range, among the 20 amino acid residues, only three are positively charged and two are negatively charged. Arginine (R), lysine (K), and histidine (H) carry a partial positive charge on the amino moiety in the side chain. The intrinsic pK value, referred to as  $pK_{\text{int}}$ , is the pK value of an ionizable side chain when it is present in pentapeptides [30]. Only arginine,  $pK_{\text{int}} = 12.3$ , and lysine,  $pK_{\text{int}} = 10.4$ , are truly positive in physiological conditions at neutral pH, whereas histidine,  $pK_{\text{int}} = 6.5$ , would be very weakly positive at a pH in the range 6–6.5. For this reason +1, +1, +0.05 net formal charges are arbitrarily adopted for arginine, lysine and histidine respectively. For glutamate (E),  $pK_{\text{int}} = 4.3$  and aspartate (D),  $pK_{\text{int}} = 3.9$ , both carrying a carboxylic moiety on the side chain, the arbitrarily adopted net formal charges are both  $-1$  in physiological conditions. All other amino acid residues are considered neutral. The positively charged residues are represented in red whereas negatively charged residues are represented in blue on the test sequences to be analyzed under our model as displayed in Fig. 8. The neutral residues are insensitive to the electrostatic potential or the axial electric field.

### **1. Algorithm and program pseudocode for computing the axial force on the nascent peptide due to the ribosome exit tunnel interaction**

The algorithm for computing the axial force on a given nascent peptide due to the ribosome exit tunnel electrostatic interaction as a function of the amino acid sequence is schematically depicted in Fig. 8.

#### **2. Reading in the given input peptide sequence**

(Step a) Read in the peptide sequence from the amino terminal end to the carboxy terminal end.

(Step b) Determine the length of the peptide (number of amino acid residues in the given peptide).

(Step c) Convert the sequence of amino acid residues into an ordered list of formal charges using the following charge coding rule:  $K \rightarrow +1$ ,  $R \rightarrow +1$ ,  $H \rightarrow +0.05$ ,  $E \rightarrow -1$ ,  $D \rightarrow -1$ . All other residues are converted to a neutral charge  $X \rightarrow 0$ .

### **3. Computing the axial position of each amino acid in the sequence, compute the axial force acting on the residue at that position, and sum the contributions of all charged residues**

(Step a) Start with the first five residues from the amino terminal end of the peptide (the first five elements in the ordered list) to build the stretch currently computed.

(Step b) Map the axial positions of the residues in the stretch, each separated by a distance  $0.25 \times 10^{-9}$  m. Position  $z = 0$  corresponds to the residue located at the ribosome exit tunnel entry point, position  $z = 5 \times 0.25 \times 10^{-9}$  corresponds to the residue located at the PTC, five residues downstream in

the sequence. All algebraic negative  $z$  positions correspond to residues that have entered the tunnel.

(Step c) Compute the ordered list of axial electric fields for each of the previous axial positions using formula (3) for the idealized cylindrical model, formula (16) or formula (20) for the realistic model, incorporating the Lorentzian peak and the truncated cone geometry at the end side of the tunnel, respectively.

(Step d) Multiply element by element, the ordered list of the axial electric fields by the ordered list of formal charges, to obtain the list of the contributing axial forces acting on the peptide stretch currently computed.

(Step e) Sum all the contributing axial forces in the peptide stretch currently computed and store the result in an ordered list of the total axial forces acting on the stretch upon the carboxy-terminal end of the nascent chain at the PTC site.

(Step f) Repeat *Step b* to *Step e* for all iterated stretches by one residue towards the carboxy terminal end, conditionally on a length of 50 residues, and while the last residue has not reached the end of the given input peptide. The 50 residues condition ensures there are at most 40 residues inside the tunnel, five residues between the PTC site and the tunnel entry point and at most five outside the ribosome exit tunnel, still under the electrostatic influence of the tunnel.

### **4. Plot the total axial force acting on the nascent peptide as a function of the last amino acid residue occupying the ribosomal PTC position**

Positive axial forces are believed to slow down the elongation rate while negative axial forces are believed to speed up the elongation rate of the ribosome.

### **B. Comparing the electrostatic interaction profiles when passing through the ribosome exit tunnel for different amino acid sequences**

#### **1. Simulated synthetic oligopeptide sequences**

It should be emphasized that due to the symmetry of the potential barrier in the idealized cylindrical model and its finite length, a clustered local enrichment in positive (negative) charge in a polypeptide sequence will first be attracted (repelled) when entering into the tunnel and will then be pulled inside (pushed outside) the tunnel when emerging at the tunnel exit point. Hence an inversion in the sign of the force profile should always be observed for locally clustered net charges that are followed by a neutral tail sequence. This inversion spreads over a distance covering the ribosome exit tunnel length which is 40 amino acid residue in length in the adopted simplified model and with equal areas under the curve; see the upper panel of Fig. 9(a).

The situation is more complicated when the tail sequence also includes local charges distribution within a range of 20–40 amino acid residues in the tail sequence (Fig. 9 lower panel) or if the electrostatic potential well barrier is not symmetric as with the truncated cone concatenated to the cylinder geometry; see the upper panel of Fig. 9(b).

To highlight the differences between a symmetric potential (idealized cylindrical model) and an asymmetric potential

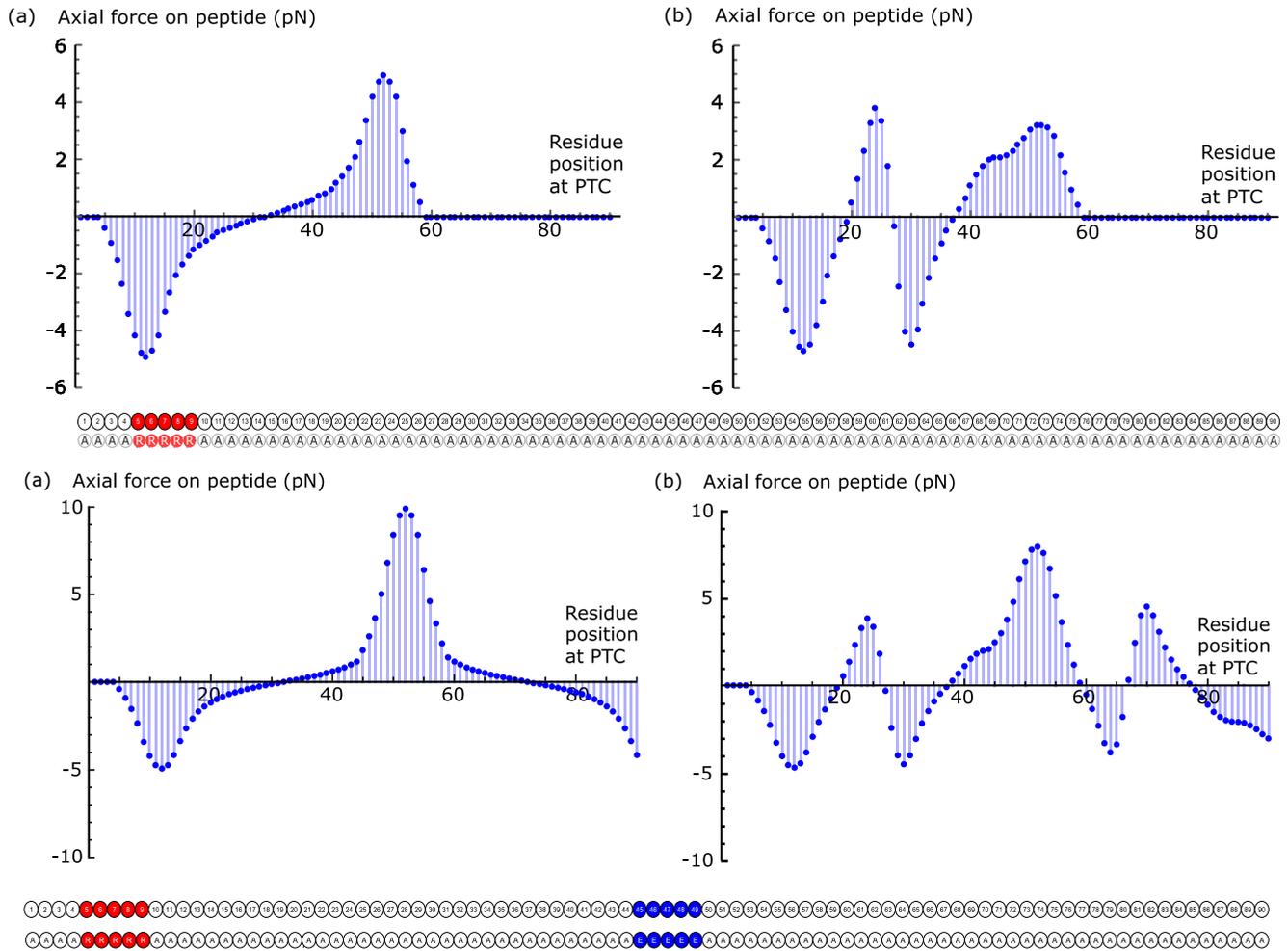


FIG. 9. Axial forces (pN) ordered values acting on the nascent chain at each residue position incorporated in the primary sequence at the PTC center. Upper panels: Symmetric electrostatic potential idealized model (a) with five contiguous arginine. Asymmetric electrostatic potential realistic model (b) with five contiguous arginine, centered at position 7. Positively charged arginine residues are colored in red in the displayed primary sequence (upper inset). Lower panels: Symmetric electrostatic potential idealized model (a) with five contiguous arginine and five glutamate residues clustered as displayed in the figure lower inset. (b) Asymmetric electrostatic potential realistic model. Opposite charges (+ residue positions 5–9: red; – residue positions 45–49: blue).

(cylinder plus truncated cone with Lorentzian peak realistic model), we compared the axial force profiles applied for the same synthetic sequences in both cases with typical clustered net charge distributions.

In Fig. 9 the axial forces acting after each amino acid incorporation at the PTC are displayed for a peptide of 90 residues in length. The upper panel of Fig. 9(a) shows the symmetric potential (idealized cylindrical model) effect on five contiguous positively charged residues between positions 5 and 9 (net positive charge centered at position 7). The nascent peptide is attracted into the tunnel until amino acid residue number 32 ( $= 7 + 25$ ) is incorporated at the PTC. From position 32 to  $59 = 9 + 50$  (position 59 corresponds to the moment when the last positively charged residue is out of the influence zone), the axial forces acting on the peptide tend to pull it back into the tunnel and these forces tend to prevent the peptide from traversing the tunnel easily.

Equivalently, it is hypothesized that the elongation proceeds at a faster rate when residues 5 to 32 are incorporated

at the PTC, and is slower when residues 33 to 59 are incorporated at the PTC. The impact on the elongation speed will be quantitatively assessed with the use of a Maxwell-Boltzmann factor. This Maxwell-Boltzmann factor provides a quantitative modulation of the average elongation speed (see the Appendix).

It is also hypothesized that the ribosome requires more mechanical power to push the nascent chain out of the tunnel when it is repelled due to the electrostatic interactions, when residues 33 to 59, in our example, are incorporated at the PTC. How the extra mechanical power is mobilized is currently unknown. An increased turnover in the biochemical reactions providing Gibbs free energy to the ribosome would probably help. Equivalently, this would require an increased rate in amino acid incorporation because more Gibbs free energy would then be available as there are two energy-rich GTPs hydrolyzed per residue incorporation. The upper panel of Fig. 9(b) shows the asymmetric potential (realistic model) effect on five contiguous positively charged residues between

positions 5 and 9 (net positive charge centered at position 7). The nascent peptide is attracted into the tunnel until amino acid residue number 19 ( $= 7 + 12$ ) is incorporated at the PTC. From position 20 to  $26 = 9 + 17$ , the axial forces acting on the peptide tend to pull it back into the tunnel and these forces tend to prevent the peptide from moving out of the tunnel. Then again, from position 27 to 37, the axial forces acting on the peptide tend to move it out of the tunnel. Finally, from position 38 to 58, the axial forces acting on the peptide tend to pull it back into the tunnel and these forces tend to prevent the peptide from traversing the tunnel easily. Compared with the upper panel of Fig. 9(a), there are two fast moves separated by a short slower move, before residue 38, instead of one single fast move in the symmetric potential case. Equivalently, it is hypothesized that the elongation proceeds at a faster rate when residues 5 to 19, then 27 to 37 are incorporated at the PTC and is slower when residues 20 to 26, then 38 to 58 are incorporated at the PTC. Note the amplitude of the axial forces are smaller but more dispersed in the positive region, for the asymmetric potential (realistic model) [Fig. 9(b)], than for the symmetric potential [Fig. 9(a)]. The lower panel of Fig. 9(a) shows the symmetric potential (idealized model) effect on a peptide with five contiguous positively charged residues between positions 5 and 9 (net positive charge centered at position 7) and five contiguous negatively charged residues between position 45 and 49 (net negative charge centered at position 47) exactly 40 residues away from the first charge cluster. The lower panel of Fig. 9(b) shows the asymmetric potential (realistic model) effect on the same peptide sequence. When the first plus cluster is emerging at the tunnel exit point, the second minus cluster is at the tunnel entry point. This situation results in high axial forces making difficult for the peptide to get out of the tunnel. In the upper panel of Fig. 9(b), there are two fast moves separated by a short slower move, before residue 38, instead of one single fast move in the symmetric potential case. The estimated maximal axial force is more than 8 pN and is reached when residue 52 is at the PTC. The axial forces tend to prevent the peptide to get out of the tunnel when residues 38 to 59 are at the PTC.

## 2. Global and local electrostatic work and energy balance

In the two symmetric electrostatic models (cylinder and cylinder with Lorentzian peak), it is important to note that there is no difference in electrostatic potential between the entry point and the exit point of the ribosome exit tunnel, whatever the form of the potential inside of the tunnel. Because electrostatic interaction is conservative, the total work spent or harnessed by a charged residue when moved from the entry point to the exit point of the tunnel will always be equal to zero. Hence, the global net mechanical work for a full sequence to be moved completely through the ribosome exit tunnel should always be equal to zero. In the two asymmetric electrostatic models (cylinder plus truncated cone with or without the Lorentzian peak), there is a net difference in electrostatic potential between the entry point and the exit point of the ribosome exit tunnel. The total work spent or harnessed by a charged residue when moved from the entry point to the exit point of the tunnel will not be equal to zero

in general. With a potential difference of 1.7 mV between the tunnel exit and entry points, the required mechanical energy is  $-0.164$  kJ/mol (0.039 kcal/mol), or  $\sim 0.3$  pN · nm on a single molecule, to traverse a single positively charged amino acid residue through the tunnel. Moreover, in any case, transiently or locally, the work to overcome positive axial electrostatic forces or the work harnessed in case of negative axial electrostatic forces acting upon any unit charged test residue may not be equal to zero. To illustrate this, the local mechanical work is computed in the case of the simulated synthetic peptide of the lower panel of Fig. 9(b) with five contiguous arginines (+) and five contiguous glutamates (−), separated by 40 neutral residues. When the oligopeptide stretch ranging from residue 5 to 19 is incorporated, the sign of the work is positive (work = +0.67 kcal/mol), according to our adopted conventions in Fig. 2, meaning that the stretch is freely benefiting electrostatic energy to traverse the tunnel during the incorporation of those amino acid residues. On the contrary, when amino acid residues 38 to 59 are incorporated in the nascent chain, the sign for the work (work =  $-1.42$  kcal/mol) is negative, meaning that mechanical energy has to be provided in some way to the nascent chain to help the stretch progressing through the tunnel. It is interesting to compare the computed values for the aforementioned mechanical work that are transiently either harnessed (0.67 kcal for the first stretch of 12 residues) or to be delivered ( $-1.42$  kcal = 9.9 pN nm for the second stretch of 21 residues), to the Gibbs free energy released from biochemical reactions at each residue incorporation, i.e.,  $\Delta G^\circ \sim -18.3$  kcal/mol (per amino acid incorporation) as detailed in the Appendix. If the chemical energy to mechanical work conversion yield is of the order of  $\sim 50\%$ , an estimate of the local required chemical energy to push the nascent chain in the case of the second stretch would be around  $-1.42/0.5 = -2.84$  kcal/mol. This amount of biochemical energy is about  $\sim 15\%$  of the Gibbs free energy released from the biochemical reactions by a single new residue incorporation associated to the ribosome elongation cycle. These simple rough comparisons show that, energetically, the ribosome has enough energy resources to overcome the local electrostatic barrier easily. However, situations may occur for which a nascent peptide will pose difficulties to the ribosome, considering that as much as  $\sim 15\%$ , or possibly more than  $\sim 30\%$  of the Gibbs free energy normally available to the ribosome per elongation cycle could be required to push the nascent chain out of the tunnel, depending on the charged amino acid distribution content of the nascent chain, and depending on the section widening in the region close to the exit point of the ribosomal tunnel.

## 3. Real protein sequences

The purpose of the ribosome exit tunnel electrostatic realistic model is to apply it to real protein sequences, to compare them and to quantitatively determine where are the critical spots for the ribosome elongation process, or what are the axial force profiles acting on proteins during the cotranslational folding process. To illustrate the application of our model to compute the forces acting on real protein sequences, we use it here in the context of neurodegenerative

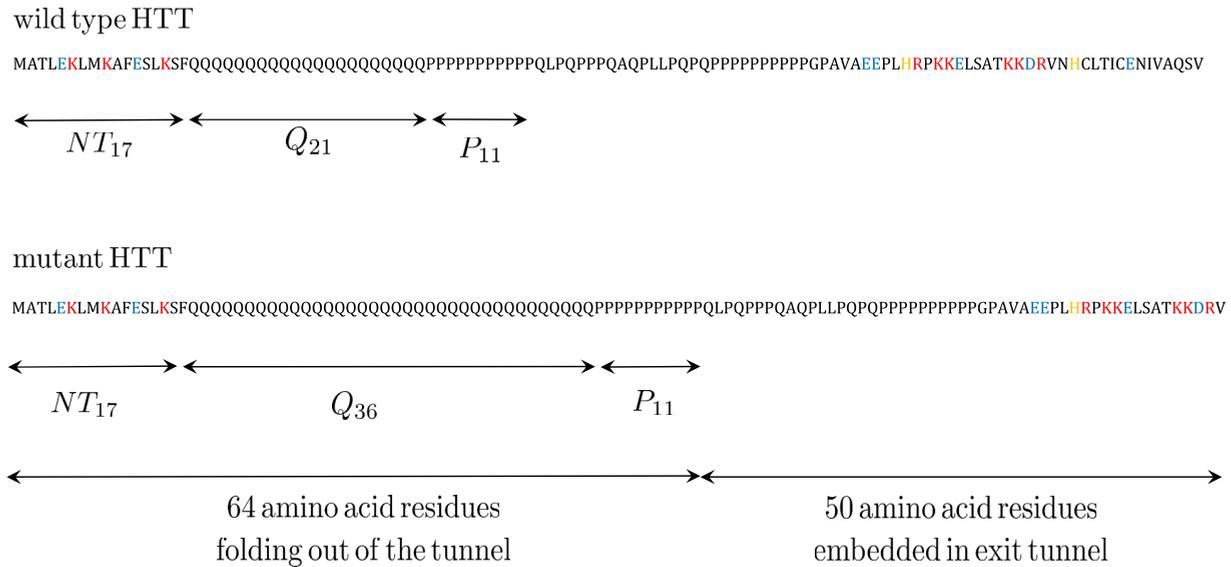


FIG. 10. Wild-type HTT and mutant mHTT huntingtin protein N-terminus starting sequence, showing the lengths of their poly-Q sequences. Positive residues: red; negative residues: blue; histidine residues: orange; neutral residues: black.

diseases like Huntington's, Creutzfeldt-Jakob, or Alzheimer's diseases. These diseases share a common pathology in the deposition of misfolded and aggregated conformations of a particular protein in the central nervous system at sites of neuronal degeneration [63]. The mechanisms of misfolding, aggregation, and their functional consequences are not yet fully elucidated. Huntington's disease is caused by mutations that expand the number of glutamine codons within an existing poly-glutamine (poly-Q) repeat sequence of the gene coding for the huntingtin protein [63–66]. The N-terminus end of a normal huntingtin protein is composed of a N-terminus sequence of 17 residues (NT<sub>17</sub>), a poly-Q sequence with a number of contiguous glutamines anywhere between 6 and 34 (e.g., Q<sub>21</sub>), and a polyproline sequence of around 11 proline residues (e.g., P<sub>11</sub>); see Fig. 10. A mutant allele coding for a number of glutamine repeats exceeding 36 (e.g., Q<sub>36</sub>) will inevitably lead to Huntington's disease if the person carrying this allele lives long enough. Huntingtin has a very long sequence with a total length of 3144 residues in the normal wild-type sequence but the mutated huntingtin is expanded only in the very beginning of the sequence. Here, we do not pretend to solve the mechanism or the detailed molecular steps causing the misfolding of the huntingtin mutant protein but provide an analysis of a possible role of the forces acting on the huntingtin growing sequence while it is biosynthesized by the ribosome and investigate a possible cotranslational misfolding situation. We compare, in Figs. 11(a) and 11(b), the axial forces profiles for the human wild-type huntingtin HTT and a mutant huntingtin mHTT for the first 150 N-terminus residues when their respective transcripts are being translated. The folding conditions and environments are different as the axial forces acted by the exit tunnel of the ribosome on these two growing nascent huntingtins are different. The two sequences embedded in the tunnel are different and cause the two very different net resulting axial forces. The mutant huntingtin has a length of the N-terminus sequence equal to 64 (= NT<sub>17</sub> + Q<sub>36</sub> + P<sub>11</sub>).

The exit tunnel exerts axial forces of electrostatic origin that are either pulling forces or pushing forces. These forces oppose (or not) to the forces generated by the spontaneous folding of the unstructured segments of the nascent polypeptide chain upon lengthening of the chain out of the tunnel. In their computational simulations, Fritch *et al.* estimated that the force difference experienced at the P-site residue upon doubling the length of the chain out of the tunnel was of the order of piconewtons [67]. In Figs. 11(c)–11(f), these folding forces are called tugging forces (black arrows). In the extreme case of SecM mediated ribosomal arrest known in bacteria, Goldman *et al.* provided evidence that the minimal tugging force must be  $\sim 10$  pN to relieve the stalled ribosome [68]. These results show that tugging forces in the range 1–10 pN can possibly be generated by emerging folding nascent chains to reach their native conformation.

As seen from the comparison of the axial forces profiles, the wild-type nascent protein is experiencing a pulling force ( $\sim +2$  pN) of electrostatic origin from the ribosome, when incorporating residue 81 to 90 at the peptidyl transferase center PTC, Fig. 11(a) first arrow and Fig. 11(c), while the length of the nascent chain out of the tunnel is 31 to 40, i.e., when the critical poly-Q segment is fully emerging from the ribosome. In the mutant mHTT, there are no pulling forces from the ribosome at this moment; the axial forces are null at this moment, Fig. 11(b) first arrow and Fig. 11(d).

When the first 64 amino-terminal residues have just emerged out of the exit tunnel of the ribosome and are exploring the folding space, the downstream 50 residues are under electrostatic interaction with the ribosome exit tunnel. The total axial resulting force acting upon residue 114 (= 64 + 50), and indirectly on the whole growing chain outside the tunnel, is different for the mutant mHTT than for the wild-type HTT. For the wild-type HTT, the nascent chain out of the tunnel, 64 residues in length, folds while the ribosome is pulling the chain toward the interior of the tunnel ( $+2$  pN), Fig. 11(a) second arrow and Fig. 11(e), whereas for

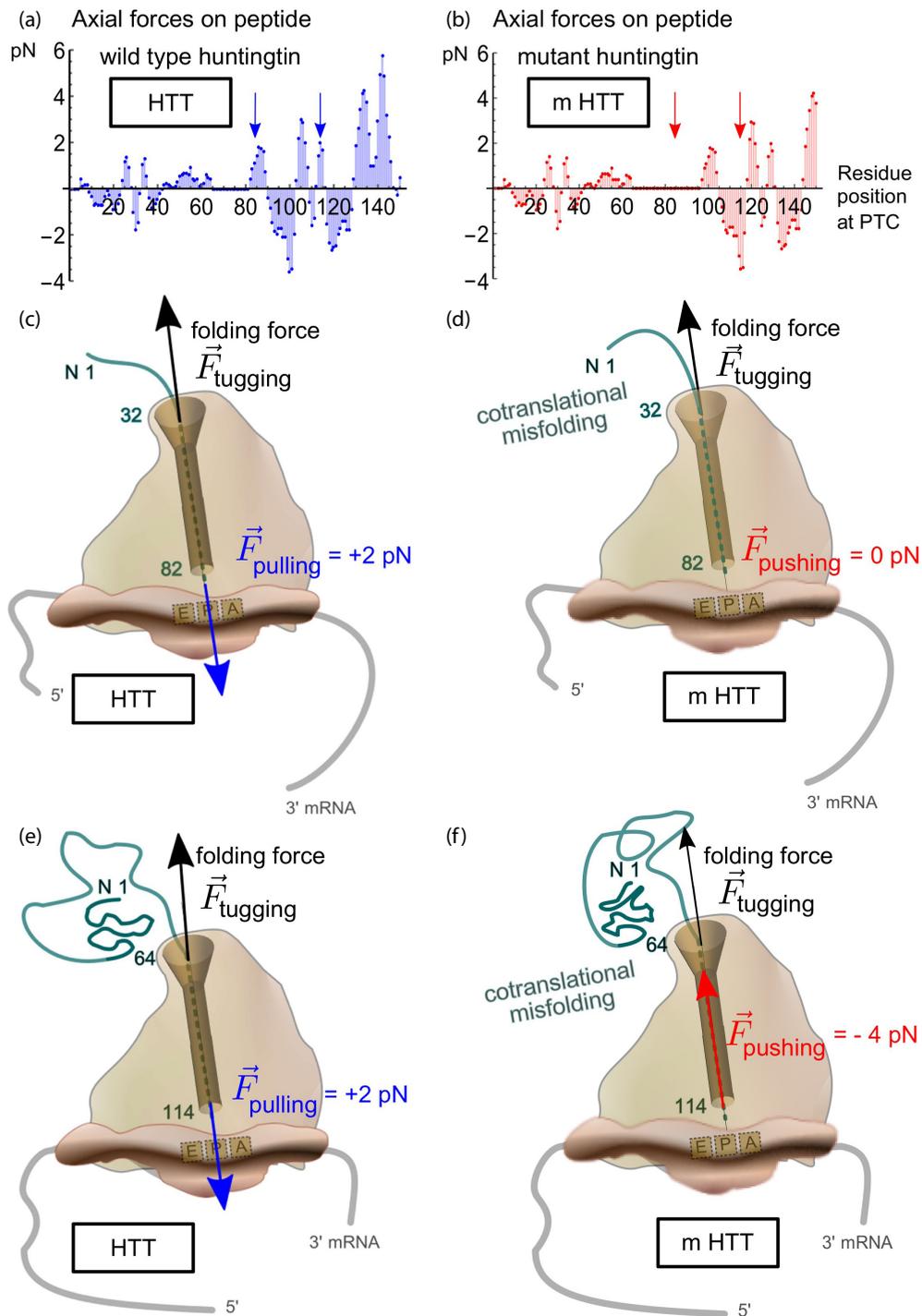


FIG. 11. Comparison of vectorial cotranslational folding for huntingtin wild-type (HTT) and mutant protein (mHTT). (a), (b) Axial forces profiles for wild-type and mutant protein. Blue and red arrows show the values of forces acting on the nascent chain at the PTC when residue 82 and 114, respectively, are incorporated in the peptide at the PTC. (c) Wild-type protein. A 2 pN pulling force due to the tunnel interaction with residues 32–82 opposed to the spontaneous folding force when residues 1–31 are out of the tunnel. (d) Mutant protein. No force is opposed to the spontaneous folding force when residues 1–31 are out of the tunnel. (e) Wild-type protein. A 2 pN pulling force due to the tunnel interaction with residues 64–114 opposed to the spontaneous folding force when residues 1–63 are out of the tunnel. (f) Mutant protein. A pushing force of 4 pN adds to the spontaneous folding force.

the mutant protein mHTT, the nascent chain out of the tunnel, 64 residues in length, folds without opposing force from the ribosome. On the contrary, for the latter, there is a pushing force on the nascent chain (−4 pN) from the ribosome exit

tunnel, Fig. 11(b) second arrow and Fig. 11(f). This analysis would suggest that proper folding of the poly-Q containing segment of huntingtin protein would require a pulling force from the ribosome. If there is no pulling force, as for the

mutant huntingtin, the poly-Q segment would be much more prone to cotranslational misfolding. Interestingly, the effect of expanding the length within the poly-Q segment would just cause a shift in the axial forces profile that reverse the forces acted by the ribosome, upon the folding segment, between the wild-type and the mutant protein.

Overall, these results suggest that a change in the local distribution of charged residues or an insertion or a replacement by neutral residues has impact on the axial forces profile over a spatially extended region of the nascent protein which is in a range corresponding to the length of the ribosome exit tunnel. The proper cotranslational folding of a nascent polypeptide chain out of the ribosome calls for specific patterns in the charged amino acid distribution in the sequence downstream, embedded in the exit tunnel, down to the peptidyl transferase center PTC. The encrypted sequence indirectly dictates, in a spatially extended way, the electrostatic interaction of the charged amino acid residues in the exit tunnel to generate axial forces profiles acted by the ribosome upon the growing chain. These forces would play a key role in the correct cotranslational folding process. The cotranslational folding is vectorial, that is, it involves elements that emerge successively from the N-terminus to the C-terminus [69]. The landscape of cotranslational folding may differ depending on the charged residue distribution which is embedded in the tunnel. Our model sheds light on how the ribosome could affect the folding trajectory.

## V. DISCUSSIONS AND FUTURE PERSPECTIVES

### A. Summary of results

#### 1. A closed form analytical expression of the electrostatic potential with piecewise phenomenological constant parameters

In this study we set out to model the electrostatics of the ribosome exit tunnel to explore quantitatively the impact of the distribution of the charged amino acid residues embedded in the tunnel on the forces acting on the nascent peptide chain during translation. Our approach was to develop a full analytical expression of the electrostatic potential inside the tunnel, starting from two idealized theoretical geometries for the tunnel, i.e., a cylinder and a cone. We eventually concatenated the cylindrical geometry with the conical geometry, and finally added an empirical Lorentzian function motivated by the known experimental observations of local and highly conserved ribosomal protein protrusions inside the tunnel. The precise geometry of the tunnel is important for quantifying the resulting electrostatic potential profile along its centerline. It shows what part of the electrostatic profile is contributed by the shapes and by the sizes of the tunnel and what part is inherent to the physicochemical properties such as the surface charge density contributed by the large number of phosphates moieties lining the tunnel inner wall, or the dielectric responses of the tunnel lumen and inner wall. In our simplified approach, we incorporated partial piecewise heterogeneity for the phenomenological charge densities and dielectric response properties and used three different ratios  $\sigma/\epsilon$ : one in the upper-lower tunnel ( $\sigma_1/\epsilon_1$ ), one in the tunnel vestibule ( $\sigma_2/\epsilon_2$ ) and one in the tunnel constriction site ( $\sigma_3/\epsilon_3$ ).

#### 2. Goodness of fit of the model to experimental electrostatic potential observation

The main results derived from the theoretical analysis of the electrostatics of the ribosome exit tunnel, as displayed in Figs. 2(e) and 3, is the goodness of the fit of the physical model with the measured data points for the electrostatic potential in the ribosome exit tunnel earlier published by Deutsch *et al.* [29]. The geometry, crystallographic data, and physicochemical properties of the tunnel inner walls of the exit tunnel consistently explain the experimentally measured values for the potential from simple first physical principles. The model draws the attention on the main geometrical and physical features as determinants of the electrostatic potential profile and the derived electric field projected along the tunnel axis. Specifically, the geometrical variation induced by widening the tunnel radius at the exit of the tunnel (cone frustum) introduces a permanent difference in the electrostatic potential between the exit and the entry points of the tunnel. The rise in the potential when moving towards the vestibule can also be contributed by an increase in the dielectric response  $\epsilon$  in the vestibule medium. This is energetically unfavorable to the positively charged amino acid residues as compared to their negatively charged amino acid counterparts. This provides a simple bioenergetic explanation to the observation that, proteome wide and across species, the protein sequences are slightly but significantly more enriched in negatively charged amino acid as compared to the positively charged amino acid residues [70]. This observation would be the consequence of a selection pressure in favor of the negatively charged residues as compared to the positively charged residues; the latter requiring more mechanical energy to traverse the ribosome exit tunnel.

The derived axial forces acting upon the nascent polypeptide stretch, within 50 residues upstream of the amino acid residue that is incorporated at the PTC, stand as a valuable quantitative model. The variation in the axial forces due to electrostatic interaction of the charged nascent chain with the ribosome exit tunnel has been estimated in a range from  $-10$  pN to  $+10$  pN in order of magnitude. More importantly, different profiles for these axial forces have been quantitatively related to synthetic polypeptides with arbitrarily charged residue distribution. Arbitrary synthetically engineered transcripts could, in principle, be used in high-resolution optical tweezers multiple traps experiments to test experimentally the theoretical profiles of the axial forces acting upon such nascent polypeptides. The electrostatic model best fitted to the experimental data of Lu *et al.* [29] is the one combining a cone frustum section concatenated to a cylindrical section with a Lorentzian peak roughly located one-third of the tunnel length away from the tunnel entry point. This particular model is used to derive, more accurately, the axial forces acting upon any nascent chain in the tunnel.

#### 3. Model applications to study the contribution of the electrostatic interaction on elongation rate and for ribosomal mechanochemistry

The ordered list of axial forces at single-residue resolution also allows to calculate the mechanical work required to overcome the electrostatic potential real profile in

the exit tunnel at each residue elongation. From this, a Maxwell-Boltzmann correcting factor can be defined following similar developments as the ones exposed in [71,72] and introduced in a seminal article by Bell in the context of cell to cell adhesion [73]. These factors can correct, at single amino acid residue resolution, and in a sequence specific way, the elongation rate in TASEP-like modeling tools. For a given transcript, the specific contribution of the tunnel electrostatic interaction locally modulates the elongation rate in a range from minus 40% to plus 85% when compared to the average elongation rate; see the Appendix. An interesting advantage of these Maxwell-Boltzmann correcting factors lies with the way they are calculated. The exact local memory of the distribution of the charged amino acid residues is conserved for a sliding window of 50 residues that are upstream the site of incorporation of a new residue at the peptidyl transferase center PTC. This extended stretch of 50 residues is expected to be under the influence of the electrostatic interaction caused by the inner wall of the ribosome exit tunnel, along its whole axial length. All charged residues, positive and negative, embedded in the tunnel, additively contribute to the pace of the elongation process. The route of force transmission to the P site residue is through the nascent polypeptide's backbone as it is also the case for the tugging force generated by the spontaneous folding of the lengthening nascent chain out of the ribosome exit tunnel [67]. Mechanical forces can alter the activation energy barriers that reactants have to overcome in the course of a chemical reaction to be converted into products. Intermediate transition states may be more easily attainable from the reactants when the system is experiencing an external force [71]. An effect of the external applied force is to provide mechanical work that will linearly decrease the activation energy even without changing the reactants' configurations or the transition state configuration [72,73]. When the axial forces upon the nascent chain buried in the tunnel are exerted toward the tunnel exit, the Gibbs free energy barrier at the PTC is presumed to be decreased, the rate of the peptidyl-tRNA deacylation step at the P site and the global rate of the peptide bond formation are both expected to be increased. To our knowledge, the model presented here is the first one to take into account the whole size and shape of the ribosome exit tunnel and updates, at single-residue resolution, the mobile 50-mer polypeptide window which is embedded in the tunnel. The position-dependent precise value of the Maxwell-Boltzmann factor is determined by this spatially extended stretch of 50 residues with a specific charge distribution that is encrypted in the transcript being decoded. These elongation rate correcting factors are at codon resolution and keep the memory of the spatially extended stretch of amino acid residues embedded in the tunnel. This is a clear improvement over the current state of the art in terms of realism and consistency of the elongation speed calculation. This can be contrasted, for instance, with studies where only positively charged residues within a limited number of residues upstream the incorporation site are considered and where arbitrarily fixed valued correcting factors are used to adjust for the electrostatic interaction in the tunnel [14].

#### 4. Model applications to study cotranslational folding

The comparison of the axial force profiles of wild-type protein sequences with mutant sequences as illustrated in

the case of huntingtin (Figs. 10 and 11) may help to study the dynamical folding of a nascent protein that is still in contact with the ribosome. The tugging forces generated by the spontaneous folding of the unstructured segments during the peptide lengthening out of the tunnel were estimated by computer simulations at piconewtons order of magnitude by Fritch *et al.* [67]. These spontaneous cotranslational folding tugging forces acting on the nascent chain can be compensated for (or not) by pulling forces from the ribosome due to the electrostatic interaction in the ribosome exit tunnel. The landscape of cotranslational folding of the wild-type and mutant huntingtin nascent proteins may differ by a sheer difference in the distribution of the charged amino acid residues that are embedded in the full length of the tunnel. This would shed light on how the ribosome takes part in configuring folding intermediates [69]. The specific pattern of the axial forces acting on the residues that are incorporated successively from the N-terminus to the C-terminus could prevent the emerging nascent chain from falling in kinetic traps, or in stable misfolded conformations, eventually resulting in protein aggregation. Our model allows a quantitative analysis of these axial forces profiles and a comparison of such profiles between correctly folded and misfolded protein conformations.

### B. Reliability of model assumptions and model limitations

#### 1. X-ray solved spatial structure data provide insights in phenomenological constants of the model

Publicly available x-ray crystallographic data of the large ribosome subunit were used in the study to assess the values of our model main phenomenological parameters for the charge surface densities in the different regions of the tunnel, namely, the upper tunnel, the constriction site, the lower tunnel, and the vestibule. The 2D maps that we produced of the phosphate moieties of the 23S rRNA and of the charged amino acid residues belonging to ribosomal proteins in the immediate vicinity of the tunnel wall were used to estimate the *bare formal* charge densities  $\sigma_1^*$  in the cylinder and  $\sigma_2^*$  in the cone frustum (vestibule). The uniform 2D joint distribution assumption of the phosphate moieties on the cylinder surface of the tunnel wall was not rejected upon Monte Carlo simulations conducted under the null hypothesis of uniformity ( $p$ -value  $> 0.083$ ). Chi-squared and Kolmogorov-Smirnov tests conducted on the 1D marginal distributions did not reject the uniformity hypothesis for the cylinder. The tunnel inner wall bare surface charge density on the cylinder surface is  $\sigma_1^* = -2.1 |e|/\text{nm}^2$ . The surface charge density on the tunnel vestibule is significantly higher at the exit port of the cone than at the entry port of the cone (KS test  $p$ -value  $= 0.008$ ). The Kolmogorov-Smirnov test for the 1D-marginal distribution along  $z$  shows that more than 50% of the phosphate groups on the tunnel vestibule are located at the edge of the exit port of the cone. Structural data, observed potential data point measurements, and coarse-grained calculations of the electrostatic potential in the tunnel also provide different lines of evidence that the dielectric response (permittivity) also increases along the tunnel centerline when moving from the lower tunnel exit point towards the cone exit port, keeping the ratio  $\sigma_2/\epsilon_2 = \sigma_{\text{cone}}/\epsilon_{\text{cone}}$  at least approximately constant.

The average surface charge density on the truncated cone surface is  $\sigma_2^* = -5.1 |e|/\text{nm}^2$  and is higher than the surface charge density on the cylinder. The x-ray crystallographic data showed that water molecules buried in the wall material in the immediate vicinity of the phosphate moieties are good candidates to explain the strong screening of the electrostatic potential. The Gouy-Chapman screening length  $\xi$  links the *formal bare* charge density directly contributed by the non-bridging oxygen atoms in the phosphate groups of the 23S rRNA to the *apparent or renormalized* charge density on the tunnel inner wall electrostatic surface within the nonlinear screening theory. The electrostatic screening exponential factor was estimated under the Gouy-Chapman approach to be  $e^{-\Delta/\xi} = e^{-5.97/1.05} = 0.0034$ . The electrostatic potential in the centerline of the tunnel is reduced by a factor  $\sim 300$  due to this water screening. The enrichment in seven positively charged amino acid residues of ribosomal proteins at the constriction site is quantitatively consistent with the local rise of the electrostatic potential in this region. The parameters of the model fitted on the potential measurements provide an independent estimate of the ratio  $\sigma_3/\epsilon_3 = 57.3 \text{ MV/cm}$  in the constriction site. Using the minimal dielectric response  $\epsilon_r = 4$  for a nascent protein occupying the tunnel medium, and using the constriction site estimated surface, the number of apparent positive charges is calculated to be  $\sim 7$ .

## 2. An alternative approach to the Poisson-Boltzmann theory for a confined microenvironment that is not in equilibrium with strong electrolytes in bulk solution

The Poisson-Boltzmann theory or its linearized Debye-Huckel version are not appropriate in the particular microenvironment inside of the ribosome exit tunnel. Essential assumptions are not met in the context of the ribosome exit tunnel. First, as already noted, it is not a bulk solution with strong electrolytes present. Water molecules are not even abundant and the diameter of the tunnel too small to accommodate monovalent or bivalent ions in numbers. Second, the concept of ionic strength is hard to define inside the tunnel, given both the ill-defined concentration and homogeneity in this particularly confined environment. Third, the exit tunnel is not in equilibrium with any bulk solution during the progress of nascent protein elongation. For all these reasons, we relied directly on the simpler Poisson model only (or equivalently on the Coulomb law) which is believed to be more relevant in the (nonequilibrium) ribosome exit tunnel than a Poisson-Boltzmann model for which both the equilibrium assumption and the presence of abundant strong mobile electrolytes are not fully supported either experimentally or theoretically. The Poisson-Boltzmann theory does apply though on the outer surface of the whole ribosome which is in equilibrium with the electrolytes in the bulk solution of the cytoplasm [74]. The detailed chemical composition of the large ribosomal subunit and of the direct microenvironment of the ribosomal exit tunnel is known. Overall, the phosphate-ribose repeats building up the homopolymeric backbone of the 23S- and 28S-rRNA are the dominant molecular constituents of the tunnel inner wall materials. The fixed (negatively charged) phosphates moieties outnumber by at least one order of magnitudes any other charged chemical

group (be it from the locally protruding ribosomal proteins origin or from the nascent polypeptide itself embedded in the tunnel): the single 23S- and 28S rRNA molecule entails more than 3000 or 5000 phosphate moieties depending on the domain of life. This is primarily the origin of the observation of a net negative electrostatic potential inside the ribosome exit tunnel and even around the external surface of the complete ribosome. There may be polarization and induced dipole effects affecting locally the apparent negative charges harbored by the phosphate moieties inside the tunnel but, from the perspective of the full length of the nascent chain backbone, the curvilinear axis centered in the ribosome exit tunnel experiences the electrostatic effect of these numerous phosphate moieties lining up the tunnel inner wall. The spatial extension of this electrostatic interaction spans throughout the length of the ribosome exit tunnel over a functional distance of around 10 nm. This chemical environment is typical and specific of the ribosome in general and of the ribosomal exit tunnel in particular. With dielectric response of proteins (nascent chain) being  $\epsilon \approx 4$  and of nucleic acid (tunnel wall main component) being  $\epsilon \approx 8$ , both much smaller than the dielectric response of aqueous medium in bulk solution ( $\epsilon = 78$  for water), polarization effects in the exit tunnel due to dielectric discontinuities, in general, are not expected to be as important as the ones that would occur in bulk solution at equilibrium.

## 3. Scope, assumptions, and model limitations

Our model of the electrostatic interaction of the ribosome exit tunnel with the nascent chain polypeptide relies on a number of critical assumptions which prevent to consider the model as a completely realistic representation. As advocated by Lucent *et al.* [75], the understanding of the complexity of molecular behavior in the ribosome exit tunnel should require an atomistic molecular dynamical description including the solvent confined to the tunnel as the medium inside the tunnel does not behave as a continuous isotropic dielectric medium. In our simplified approach, we incorporated partial stepwise heterogeneity for the phenomenological charge densities and dielectric response properties and used three different ratios  $\sigma/\epsilon$ : one in the upper-lower tunnel, one in the tunnel vestibule, and one in the tunnel constriction site. The exact size of the exit tunnel and the number of accommodated residues inside the tunnel are species (life domain) dependent and polypeptides dependent. The Euclidean distance between two consecutive  $\alpha$ -carbons is 3.5 Å in a peptide bond in trans configuration (fully stretched polypeptide). In the most compacted form of a polypeptide, i.e., the  $\alpha$ -helix, the distance between two consecutive  $\alpha$ -carbons, projected along the helix axis, is 1.5 Å. The median distance between these two upper and lower limit cases is 2.5 Å. The number of residues that could be accommodated in a 100 Å long tunnel would be 28, 40, and 66 depending on the polypeptide configuration. We adopted the median length between two  $\alpha$ -carbons (0.25 nm) and hypothesized as a general rule that the number of accommodated residues is 40 for a tunnel length of 10 nm. In electrostatics, the electric polarization is defined macroscopically by the volume density of the sum of all microscopic electric dipole moments. The individual dipole moments are either induced dipole moments (elec-

tronic cloud polarizability) or permanent dipole moments (of intrinsically polar molecules, e.g., water,  $\alpha$ -helix) reoriented by the imposed electric field. The electric polarization is a function of the total electrical field. These physical effects are of fundamental importance in the context of the ribosomal exit tunnel. Unstructured nascent polypeptide sequences do not entail privileged permanent dipoles. On the other hand, it is known that  $\alpha$ -helices secondary structures entail permanent dipoles with values increasing with their length. The permanent electric dipole moment of an alpha helix secondary structure is oriented from the C-terminal to the N-terminal end of the sequence and is of the order  $\bar{\mu}_{e,\alpha\text{ helix}} = 3.5 N D$  where  $N$  is the number of residues in the helix and  $D$  the Debye unit of the electric dipole moment. Such a  $\alpha$ -helix dipole moment is larger than the one of a single water molecule. The upper and lower tunnel cannot in general accommodate secondary structures other than  $\alpha$ -helices. When  $\alpha$ -helices secondary structures appear in the ribosome exit tunnel, they tend to align with the electric field along the  $z$ -axis. In the upper tunnel or in the lower tunnel just after the tunnel constriction,  $\alpha$ -helices orient in a parallel direction to the tunnel axis with the  $\alpha$ -helix dipole moment pointing toward the tunnel exit. This is favorable as it corresponds to the direction of the natural vectorial elongation, resulting in the global movement of the nascent peptide directed toward the tunnel exit, namely from the C-terminal end to the N-terminal end of the peptide embedded in the tunnel. On the contrary, in the end of the lower tunnel or vestibule end, the dipole moment of an  $\alpha$ -helix would tend to point in the opposite direction, namely toward the PTC. In both cases, this would reinforce locally the electric displacement. These events would locally modulate the effects of our simplified electrostatic potential profile on the electric fields and axial forces experienced by a charged residue. Although  $\alpha$ -helix secondary structures might occasionally form in the tunnel during the nascent chain elongation, for the sake of simplicity of our model, and because such events are transients and context dependent, we did not incorporate these events into our electrostatic model at this stage. Our simplified model applies better to elongation situations of unstructured (random coil or extended uncompact) polypeptides. It should be noted however, that the profile of the electrostatic potential and the axial electric field derived from our model provide insights on the orientation trends of  $\alpha$ -helix nascent secondary structures that are consistent with observations of early compaction and early cotranslational folding events occurring within the lower, upper tunnel and vestibule. Indeed, reversion of  $\alpha$ -helices in the lower tunnel end or in the tunnel vestibule have been observed, for instance, in Refs. [76,77]. Reversion of an  $\alpha$ -helix in the tunnel is the observed fact that the N-terminal end of the  $\alpha$ -helix peptide gets back inside the tunnel toward the PTC instead of progressing toward the exit end. These kinds of events are the starting point of tertiary structure sometimes initiating at the lower end of tunnel and or the tunnel vestibule. The reversion propensity of  $\alpha$ -helices in the end of the tunnel might be favored by the orientation of the axial electric field pointing toward the PTC in this zone. Reversion of  $\alpha$ -helices would not be favorable in the very beginning of the upper tunnel near the PTC. The possible secondary structures that can start to form in the tunnel and the impact of their dipole moments on

the local electric field in the tunnel have not been explicitly incorporated into our simplified model. It is also not known with full certainty whether or not the shape and geometry of the ribosome exit tunnel in the large subunit LSU of the ribosome stay the same during the translation process *in vivo* or if reversible continuous elastic deformations occur *in vivo*. We showed that the mechanical energy required to push the growing nascent chain through the LSU exit tunnel, even in difficult scenarios, would be smaller than the Gibbs free energy released from the transpeptidation and the hydrolysis of a single GTP. Overall, the widening of the radius along its central axis toward the exit of the tunnel is however known and contributes to the asymmetric electrostatic potential profile that we estimated. Alternative explanations of the rise in electrostatic potential could also be the rise in the dielectric response when moving from the lower tunnel to the vestibule. This could be due to an increase in abundance of free water molecules. The estimated electrostatic potential profile fits the available observed data rather well at least for the rabbit reticulocytes ribosomes. We must recognize, however, that we relied on only a small sample size of four to six point measurements. Complementary wider experimental studies on both prokaryotes and eukaryotes ribosomes would be beneficial.

To quantitatively estimate the axial forces applied on the nascent chain, we made a rigid body assumption or assumed the nondeformability of the nascent chain inside the tunnel. This assumption is most certainly not valid for all polypeptide chains and most probably not valid locally. However, this assumption could be legitimate on average and proteome wide. Indeed, the ribosome exit tunnel is universal, meaning that all the polypeptides that are naturally occurring in the biosphere did traverse the tunnel at the time of their biosynthesis. All the amino acids have progressed through the entire length of the tunnel after they were incorporated in the nascent chain at the peptidyl transferase site. On average, as a first approximation, we can consider that these amino acids followed a centro-axial trajectory in the tunnel and experienced the effect of the electrostatic interaction upon the charged residues with which they are directly or indirectly bound. Fritch *et al.* recently showed that the spontaneous folding force was transmitted directly from the outside of the tunnel to the PTC center through the backbone of the nascent chain [67]. This direct transmission route supports the rigid body assumption we made for the peptide buried in the tunnel.

### C. Comparison to state of the art and literature in the field

The importance of electrostatic interactions should not be overstated and has to be quantitatively incorporated according to its weight to all other key determinants of the ribosome elongation rate. At least five determinants in the protein elongation rate are known and have been investigated for decades by the scientific community: tRNAs relative abundance and adaptation to mRNA codon usage (local or global tRNA adaptation index); exit tunnel electrostatic interaction; proline residues at the P and/or A site; downstream mRNA secondary structures hampering the movement of the ribosome toward the mRNA 3' terminus; and ribosome interference (traffic jam) and ribosomes' pool-limited resources. Although Charneski

*et al.* [27] argued that the electrostatic interaction was the major determinant, the electrostatic interaction is indeed only one of these five determinants. To dissect the variance in the elongation rate, all factors must be taken into account altogether, especially if these factors are meant to be used as predictors for ribosome profiling (Ribo-Seq) data or are aimed at improving TASEP models. Our study focused only on the electrostatic interaction. Current models of the ribosome exit tunnel electrostatic interaction are incomplete and a correct model is lacking. Our study is aimed to contribute to the improvement of the predictive power of such models. The electrostatic interaction is not the only force acting on the nascent chain during elongation, the entropy driving force and the folding forces acting on the chain also contribute, especially as soon as the N-terminal end of the nascent chain has made its pass through the tunnel (after the ribosome translated the first  $\sim 45$ – $50$  codons). Our study did not address the calculation of the entropy driven forces. Position specific biases in the distribution of the five determinants contributing to the elongation rate are antagonizing each other and blur our understanding of the elongation rate. Two references [20,78] show instances of such position specific biases. Tuller *et al.* [78] initially inferred from Ribo-Seq data that the first 30–50 codons at the 5'-end in yeast transcripts were low-efficiency codons (rare or poorly matched codons to the tRNA pool) and termed this part the “low-efficiency ramp” or the 5'-ramp. This initial 5'-ramp (due to slow accommodation and proof-reading of tRNAs at the A site at the beginning of mRNA sequences) was considered to be the effect of a selection pressure enabling to limit the occurrence of later ribosome traffic jams downstream in the course of mRNA translation. This 5'-ramp explanation was later argued to be insufficient [20]. In contrast, the results of Dao Duc *et al.* [20] suggested that while the N-terminus of the nascent polypeptide has not exited from the tunnel, positively charged amino acids in specific parts of the polypeptide actually facilitate the elongation speed, while the opposite is true for negatively charged amino acids. The statistical analysis by Dao Duc [20] showed that the presence of positive and negative charges in the upper tunnel may respectively facilitate and inhibit elongation as the nascent chain makes its initial pass through the tunnel. These authors found that the number of positively charged residues in the window [1:11] and the number of negatively charged residues in the window [6:14] upstream the A-site are important features with opposite effects; the former facilitates elongation, while the latter slows down elongation. Overall these results are fully consistent with the shape of the electrostatic potential profile of our model: the longitudinal (axial) direction of the force that a positively charged particle would experience along the tunnel points from the P site toward the tunnel exit end at least when the first 15 residues of the nascent chain enter the tunnel. This is the opposite for negatively charged particles. The electrostatic potential profile also shows that the longitudinal (axial) direction of the force that a positively charged particle would experience along the tunnel when the last 15–20 C-terminal residues are still in the tunnel exit points from the exit toward the P-site (upper tunnel). This is also consistent with the results of Nissley *et al.* [79]. Dao Duc *et al.* [20] made the case that once the N-terminus has exited the tunnel, the hydrophathy of the part of the nascent

polypeptide within the ribosome plays a major role in governing the elongation rate variation. These authors concluded that the movement of the polypeptide inside the tunnel is driven by two distinct biophysical mechanisms. First, when the peptide chain has not yet exited the tunnel, electrostatic interactions in the tunnel play a major role in regulating the movement of the chain down the exit tunnel. Second, when the peptide chain has reached a certain length and its N-terminus has exited the tunnel, it is the structure of the chain itself (captured through the hydrophathy) that determines its movement through the tunnel. It should be added that the entropy driving forces upon protein folding outside the tunnel or forces exerted from chaperone proteins also contribute and should be quantified. The electrostatic potential mathematical model that we proposed provides insights into the real measurements that were made in the pioneering experimental studies and that could still be made in the future. It should be emphasized that electrostatic potential measurements should always be conducted in association with precise measurements of size and shape of the tunnel and accurate positional mapping along the tunnel axis.

#### D. Future perspectives

As future perspectives in x-ray crystallographic data mining, we suggest comparing the 2D maps of charged moieties in the immediate vicinity of the tunnel wall for ribosomes from different species across the three domains of life in order to investigate the common shared patterns.

We expect the *in silico* research community to assign itself the task of using our suggested electrostatic model, and the ordered list of Maxwell-Boltzmann factors derived from it, to modulate the elongation rate for a better quantitative account of the effect of the tunnel on the charged amino acid residues. On average, it is believed that the comparison of the electrostatic interaction with the exit tunnel of any two different nascent chains, by applying our model, can provide quantitative insights on the effects of the difference in charged amino acid distribution across their primary sequences. This paves the way to a variety of bioinformatic studies on transcriptomic and proteomic data to shed light on translational control. Immediate perspectives and objectives will address (a) accurate predictions of ribosome footprints in Ribo-Seq profiling ensemble experiments; (b) precise dynamical predictions of the speed of elongation in single mRNA molecules experiments; (c) quantitative predictions of the measured tugging force profiles on nascent polypeptide chain emerging from the ribosome exit tunnel in high-resolution multiple trap optical tweezer experiments to be conducted on tethered ribosomes *in vitro*; (d) experimental measurement of the strength of the electric fields through vibrational Stark spectroscopy; and (e) comparison of axial forces profiles associated to correctly folded or misfolded proteins for the study of cotranslational folding and protein aggregation mechanisms.

The model presented in this study consistently connects different results and experimental observations coming from different fields in molecular biology, x-ray crystallography, structural and physical chemistry, synthetic and multiomics biology and provides a clear picture of the electrostatic in-

teractions in the ribosome exit tunnel and their effects on the protein elongation rate.

### ACKNOWLEDGMENTS

We gratefully acknowledge Prof. John Martin for fruitful discussions on the electrostatic idealized model of the ribosome exit tunnel. We also acknowledge the two anonymous referees for their valuable contribution to the discussions section of the revised paper, for their contribution on the dielectric medium heterogeneity and discontinuity analysis, and for their recommendation of using x-ray solved spatial structures of the ribosome to assess the model phenomenological constants, assumptions, scope, and limits. This work was supported by the FNRS-FWO EOS Grant No. 30480119 (Join-t-against-Osteoarthritis), the FNRS-WELBIO (THERAtRAME) Grant No. WELBIO-CR-2017S-02 in Belgium, and the European Research Council under the European Union's Horizon 2020 Framework Program (H2020/2014-2020)/ERC Grant No. 772418 (INSITE). F.K. is a research associate of the FRS-FNRS, Belgium.

### APPENDIX

#### 1. Hollow straight cylinder model

In a first simplified approach, the ribosome exit tunnel is considered a hollow straight cylinder [Fig. 2(a) left panel]. The wall material is not of the conductor type with mobile free charges but is rather a dielectric material harboring fixed partial charges—the fixed phosphate moieties lining the inner wall. As a first reasonable assumption, the fixed charges are supposed to be uniformly distributed on the surface of the inner wall. The size of the hollow cylinder closest to the shape of the ribosome exit tunnel documented in the literature would be 85–100 Å (8.5–10 nm) in length and 10–20 Å (1–2 nm) in diameter [28,39]. The precise length for the ribosomal exit tunnel as measured by cryo-electron microscopy is 9.2 nm on average in prokaryotes and 8.3 nm on average in eukaryotes [39]. The *in vivo* lengths are believed to be a bit larger due to thermal dilatation at the higher temperatures prevailing in living organisms as compared to the cryogenic conditions.

For a given uniformly distributed charge density  $\sigma$  on the inner surface wall of the cylinder, the determination of the electrostatic scalar potential  $\Phi(\vec{\mathbf{r}})$  and of the electric field  $\vec{\mathbf{E}}(\vec{\mathbf{r}})$ , at any spatial point close to or far away from the cylindrical surface, are well stated problems in classical electromagnetism [40]. For the sake of simplicity, we restrict ourselves here on spatial points located on the axis of the hollow cylinder, lying anywhere inside or outside of the tunnel. In this schematic pictorial description, a new amino acid is incorporated into the nascent protein which gets into the tunnel from one side (conventionally from the right of Fig. 2, left panel). The nascent oligopeptide is then pushed by the multitasking ribosomal enzymatic functions inside the tunnel and out of the tunnel at the other side (left side of Fig. 2) of the hollow cylinder. The movement is strictly asymmetric as the nascent protein always enters the tunnel from the same side with the amino terminal end of the protein getting in first and the carboxy terminal end of the protein getting in last. Under

this idealized model, the hollow cylinder itself is symmetric and has a uniform charge distribution.

The electrical scalar potential  $\Phi(\vec{\mathbf{r}})$  at the observed position  $\vec{\mathbf{r}}$  is expressed by

$$\Phi(\vec{\mathbf{r}}) = \frac{1}{4\pi\epsilon} \iint_S \frac{\sigma(\vec{\mathbf{r}}') da}{|\vec{\mathbf{r}} - \vec{\mathbf{r}}'|}, \quad (\text{A1})$$

where  $\sigma(\vec{\mathbf{r}}')$  is the surface-charge density (measured in coulombs per square meter) at position  $\vec{\mathbf{r}}'$  of the source,  $da$  is the two-dimensional surface element at  $\vec{\mathbf{r}}'$  and  $\epsilon$  is the permittivity of the dielectric medium [formula (1.23) in Jackson [40]] with  $\epsilon = \epsilon_r \epsilon_0$ , where  $\epsilon_r$  is the relative permittivity of the medium and  $\epsilon_0$  is the permittivity of free space. We can take advantage of the axial symmetry and restrict to the spatial points on the  $z$ -axis, i.e., for  $\vec{\mathbf{r}} = (0, 0, z)$ . The surface integration is conducted on the support of the source charges. The cylinder's thin wall is geometrically generated by the  $\gamma(u)$  curve moving axially along the  $z$ -axis from  $z = -L$  to  $z = 0$  as drawn in Fig. 2(a) (left panel) and where  $L$  and  $R$  are the length and radius of the hollow cylinder, respectively:

$$\gamma(u) = (R \cos u, R \sin u, L), \quad u \in [0, 2\pi]. \quad (\text{A2})$$

The cylinder's surface is written as  $S = \phi(K)$  where  $K = \{(u, v) \in [0, 2\pi] \times [-1, 0]\}$  and where  $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}^3: \phi(u, v) = (R \cos u, R \sin u, vL)$ .  $D_u \phi$  is the first partial derivative of the parametric equation of the surface  $\phi(u, v)$  with respect to  $u$ . In the general formula (A1), the surface-charge density  $\sigma(\vec{\mathbf{r}}')$  is dependent of the position  $\vec{\mathbf{r}}'$  on the support of the source charges. Here we will take the simple approximation that  $\sigma$  can be considered a constant parameter over a surface of a given shape, e.g., over a cylinder or over a cone. This is the surface charge uniform distribution assumption for a given shape.

The electrostatic scalar potential results from the surface integral calculation:

$$\Phi(z) = \frac{\sigma}{4\pi\epsilon} \iint_{K=\{(u,v) \in [0,2\pi] \times [-1,0]\}} \frac{1}{\sqrt{(z-vL)^2 + R^2}} \times |D_u \phi \wedge D_v \phi| du dv, \quad (\text{A3})$$

$$D_u \phi = (-R \sin u, R \cos u, 0), \quad (\text{A4})$$

$$D_v \phi = (0, 0, L), \quad (\text{A5})$$

$$|D_u \phi \wedge D_v \phi| = \left| \det \begin{pmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ -R \sin u & R \cos u & 0 \\ 0 & 0 & L \end{pmatrix} \right| \quad (\text{A6})$$

$$= |(RL \cos u, RL \sin u, 0)| \quad (\text{A7})$$

$$= \sqrt{R^2 L^2 (\cos^2 u + \sin^2 u)} \quad (\text{A8})$$

$$= RL, \quad (\text{A9})$$

$$\Phi(z) = \frac{\sigma RL}{4\pi\epsilon} \int_0^{2\pi} du \int_{-1}^0 \frac{dv}{\sqrt{(z-Lv)^2 + R^2}} \quad (\text{A10})$$

$$= \frac{\sigma RL}{2\epsilon} \int_{-1}^0 \frac{dv}{\sqrt{(z-Lv)^2 + R^2}} \quad (\text{A11})$$

$$= \frac{\sigma RL}{2\epsilon} \int_{-1}^0 \frac{dv}{R\sqrt{\left(\frac{z-Lv}{R}\right)^2 + 1}}. \quad (\text{A12})$$

The substitution  $w = \frac{z-Lv}{R}$  yields  $dw = -\frac{L}{R}dv$  and

$$\Phi(z) = -\frac{\sigma RL R}{2\epsilon L} \int_{*}^{*} \frac{dw}{R\sqrt{w^2 + 1}} \quad (\text{A13})$$

$$= -\frac{\sigma R}{2\epsilon} \left. \arg \sinh(w) \right|_{*}^{*} \quad (\text{A14})$$

$$= -\frac{\sigma R}{2\epsilon} \left[ \arg \sinh\left(\frac{z-Lv}{R}\right) \right]_{-1}^0 \quad (\text{A15})$$

$$= \frac{\sigma R}{2\epsilon} \left[ \arg \sinh\left(\frac{z-Lv}{R}\right)_{v=-1} - \arg \sinh\left(\frac{z-Lv}{R}\right)_{v=0} \right] \quad (\text{A16})$$

$$= \frac{\sigma R}{2\epsilon} \left[ \arg \sinh\left(\frac{z+L}{R}\right) - \arg \sinh\left(\frac{z}{R}\right) \right]. \quad (\text{A17})$$

As the  $\arg \sinh$  may be expressed as a logarithm (to prove this, recall that if  $x = \sinh y$ ,  $y = \arg \sinh x$  and so  $\cosh y = \sqrt{1+x^2}$  whence  $\sinh y + \cosh y = e^y$  and we conclude that  $y = \log|x + \sqrt{1+x^2}|$ ), the electrostatic scalar potential finally writes

$$\Phi(z) = \frac{\sigma R}{2\epsilon} \log \frac{\left| \frac{z+L}{R} + \sqrt{\left(\frac{z+L}{R}\right)^2 + 1} \right|}{\left| \frac{z}{R} + \sqrt{\left(\frac{z}{R}\right)^2 + 1} \right|} \quad (\text{A18})$$

$$= \frac{\sigma R}{2\epsilon} \log \frac{z+L + \sqrt{(z+L)^2 + R^2}}{z + \sqrt{z^2 + R^2}}. \quad (\text{A19})$$

The electric field projected along the cylinder axis can be computed as the opposite of the scalar potential gradient, i.e., by taking the first derivative with respect to  $z$  directly from formula (A17):

$$E_z = -\vec{\nabla}\Phi(z) \cdot \vec{e}_z \quad (\text{A20})$$

The cone's surface is written as  $S = \phi(K)$  where  $K = \{(u, v) \in [0, 2\pi] \times [-1, 0]\}$  and where  $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}^3: \phi(u, v) = [R \cdot (1 - \frac{vL \tan \alpha}{R}) \cos u, R(1 - \frac{vL \tan \alpha}{R}) \sin u, vL]$ . The electrostatic scalar potential results from the surface integral calculation:

$$\Phi(z) = \frac{\sigma}{4\pi\epsilon} \iint_{K=\{(u,v) \in [0,2\pi] \times [-1,0]\}} \frac{1}{\sqrt{(z-vL)^2 + R^2(1 - \frac{vL \tan \alpha}{R})^2}} |D_u \phi \wedge D_v \phi| du dv, \quad (\text{A24})$$

$$D_u \phi = \left[ -R \left(1 - \frac{vL \tan \alpha}{R}\right) \sin u, R \left(1 - \frac{vL \tan \alpha}{R}\right) \cos u, 0 \right], \quad (\text{A25})$$

$$D_v \phi = (-L \tan \alpha \cos u, -L \tan \alpha \sin u, L), \quad (\text{A26})$$

$$|D_u \phi \wedge D_v \phi| = \left| \det \begin{pmatrix} -R \left(1 - \frac{vL \tan \alpha}{R}\right) \sin u & R \left(1 - \frac{vL \tan \alpha}{R}\right) \cos u & \vec{e}_z \\ -L \tan \alpha \cos u & -L \tan \alpha \sin u & L \end{pmatrix} \right| \quad (\text{A27})$$

$$= \left| \left( RL \left(1 - \frac{vL \tan \alpha}{R}\right) \cos u, RL \left(1 - \frac{vL \tan \alpha}{R}\right) \sin u, RL \left(1 - \frac{vL \tan \alpha}{R}\right) \tan \alpha \right) \right| \quad (\text{A28})$$

$$= \sqrt{R^2 L^2 \left(1 - \frac{vL \tan \alpha}{R}\right)^2 + R^2 L^2 \left(1 - \frac{vL \tan \alpha}{R}\right)^2 \tan^2 \alpha}$$

$$= -\frac{\partial \Phi(z)}{\partial z} \quad (\text{A21})$$

$$= -\frac{\sigma R}{2\epsilon} \left( \frac{1}{\sqrt{R^2 + (z+L)^2}} - \frac{1}{\sqrt{R^2 + z^2}} \right). \quad (\text{A22})$$

Of course, the axial force applied on a test particle is the product of the axial electric field with the charge of the test particle:

$$F_z = qE_z. \quad (\text{A23})$$

The plots of electrostatic scalar potential  $\Phi(z)$  and of the axial force  $F_z$  acting on a unit test charge located on the tunnel axis at any point of coordinate  $z$  are displayed in Fig. 12, with the medium permittivity prevailing inside the ribosome exit tunnel. A negative force means that the test particle is forced to move towards negative  $z$  values whereas a positive force means that the test particle is forced to move towards positive  $z$  values. In these plots,  $\sigma$  is adjusted so that the potential fits the range of the experimentally measured values given for instance in Lu *et al.* [29].

## 2. Normally truncated straight cone model

An alternative approach would depict the tunnel as a hollow cone normally truncated at both ends [Fig. 2(b)]. The section radius at the entry point is still equal to  $R = 0.5$  nm but with a section radius twice that value at the tunnel exit point and equal to  $R = 1$  nm. With the total axial length kept at  $L = 10$  nm, the half opening angle along the axis is  $\alpha \sim 0.05$  radian (2.86 arc degrees) and exactly such that  $\tan \alpha = R/L$  complying with the observation that the diameter at the exit point is around twice the diameter at the entry point of the tunnel. This better reflects the actual geometry of the real ribosomal exit tunnel as reported in the literature [28].

To analytically derive the correct equation for the potential and axial electrical field in such a conical tunnel, the procedure is the same as the one previously conducted for the cylinder, but this time with the support of the uniformly distributed charges defined by a cone surface normally truncated at both ends.

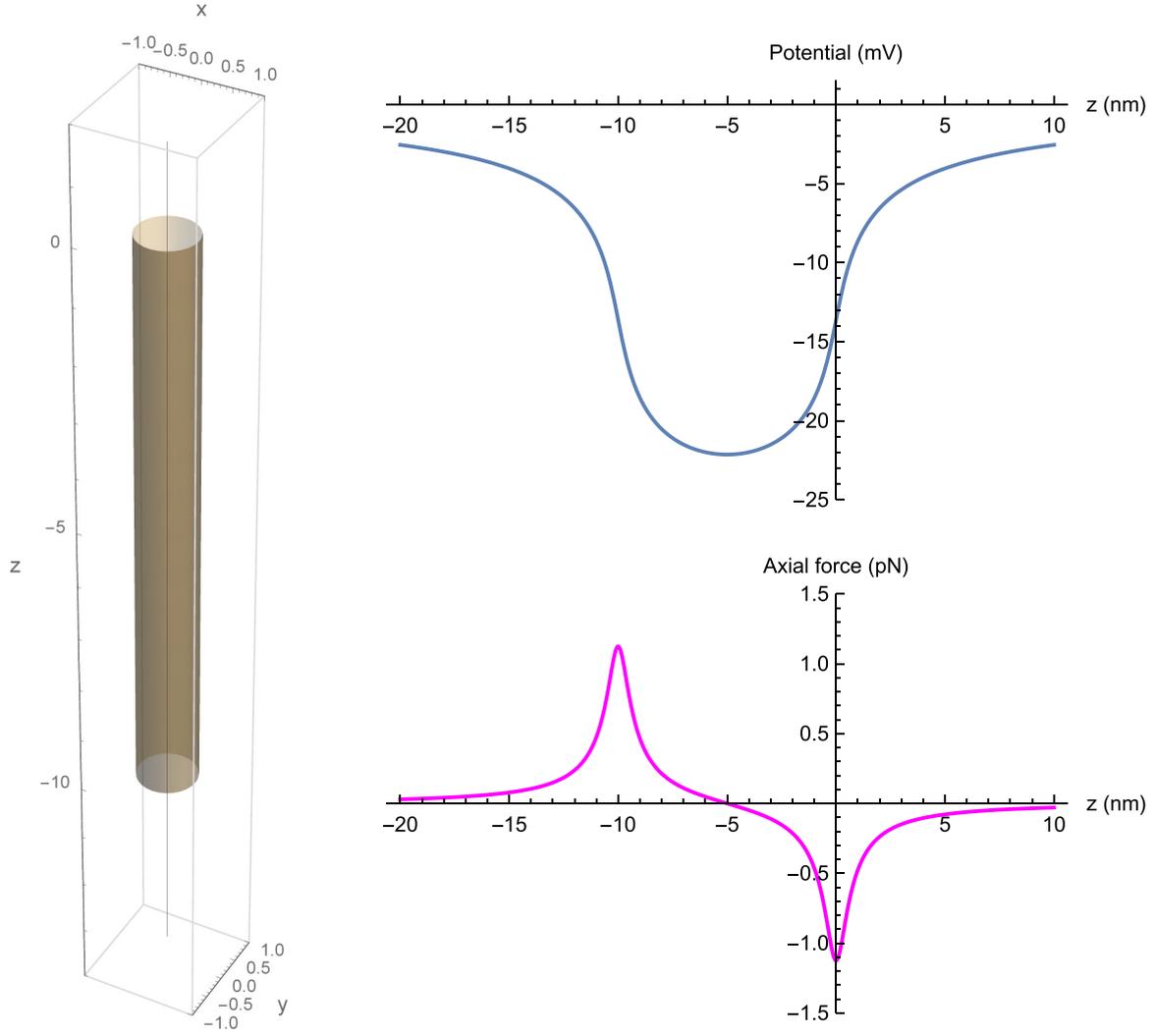


FIG. 12. Electrostatic scalar potential on the axis of the ribosomal exit tunnel (upper panel) and axial force (lower panel) as a function of axial  $z$  position for a positively unit charged test amino acid residue on the tunnel axis. Tunnel idealized as a cylinder.

$$= \sqrt{R^2 L^2 \left(1 - \frac{vL \tan \alpha}{R}\right)^2 (1 + \tan^2 \alpha)} \quad (\text{A29})$$

$$= RL \left(1 - \frac{vL \tan \alpha}{R}\right) \frac{1}{\cos \alpha}, \quad (\text{A30})$$

$$\Phi(z) = \frac{\sigma RL}{4\pi\epsilon \cos \alpha} \int_0^{2\pi} du \int_{-1}^0 \frac{\left(1 - \frac{vL \tan \alpha}{R}\right) dv}{\sqrt{(z - Lv)^2 + R^2 \left(1 - \frac{vL \tan \alpha}{R}\right)^2}} \quad (\text{A31})$$

$$= \frac{\sigma RL}{2\epsilon \cos \alpha} \int_{-1}^0 \frac{\left(1 - \frac{vL \tan \alpha}{R}\right) dv}{\sqrt{(z - Lv)^2 + R^2 \left(1 - \frac{vL \tan \alpha}{R}\right)^2}} \quad (\text{A32})$$

$$= \frac{\sigma L}{2\epsilon \cos \alpha} \int_{-1}^0 \frac{(R - L \tan \alpha v) dv}{\sqrt{(z - Lv)^2 + (R - L \tan \alpha v)^2}} \quad (\text{A33})$$

$$= \frac{\sigma L}{2\epsilon \cos \alpha} \int_0^1 \frac{(R + L \tan \alpha v) dv}{\sqrt{(z + Lv)^2 + (R + L \tan \alpha v)^2}}, \quad (\text{A34})$$

where, in the last line, a dummy integration variable was changed with  $v' = -v \rightarrow dv' = -dv$  and the change of sign was canceled by the integration limits permutation. The complete derivation is given in the following section. To alleviate the

notations, the two following substitutions are adopted:

$$f_1(z) = R \cos \alpha - z \sin \alpha, \quad (\text{A35})$$

$$f_2(z) = R \sin \alpha + z \cos \alpha. \quad (\text{A36})$$

$f_1(z)$  is always positive for  $z \leq 0$  (and even for  $z < R/\tan \alpha$ , i.e., the virtual  $z$  position of the cone summit), which is the domain we are interested in. The  $z$ -position values are negative in the tunnel and beyond its exit point:

$$\begin{aligned} \Phi_{\text{cone}}(z) &= \frac{\sigma}{2\epsilon} \left\{ f_1(z) \cos \alpha \log \left[ \frac{|L/\cos \alpha + f_2(z) + \sqrt{f_1^2(z) + (L/\cos \alpha + f_2(z))^2}|}{|f_2(z) + \sqrt{R^2 + z^2}|} \right] \right. \\ &\quad \left. + \sin \alpha [\sqrt{f_1^2(z) + (L/\cos \alpha + f_2(z))^2} - \sqrt{R^2 + z^2}] \right\} \end{aligned} \quad (\text{A37})$$

$$\begin{aligned} &= \frac{\sigma}{2\epsilon} \left\{ f_1(z) \cos \alpha \log \left[ \frac{|L/\cos \alpha + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + \frac{L^2}{\cos^2 \alpha}}|}{|f_2(z) + \sqrt{R^2 + z^2}|} \right] \right. \\ &\quad \left. + \sin \alpha [\sqrt{(z+L)^2 + (R+L \tan \alpha)^2} - \sqrt{R^2 + z^2}] \right\}. \end{aligned} \quad (\text{A38})$$

Equation (A38), which is valid for any conical geometry with entry section of radius  $R$  and any cone angle  $\alpha$ , replaces Eq. (A19) of the cylindrical geometry. Note that the electrostatic potential vanishes at  $z = \pm\infty$  as physically expected.

It is also worth noticing that Eq. (A38) for the truncated cone restores, as a special case, Eq. (A19) for the cylinder when  $\alpha = 0$ , as expected as well.

The electric field projected along the truncated cone axis can be computed as the opposite of the scalar potential gradient, i.e., by taking the first derivative with respect to  $z$  of Eq. (A38). The full derivation is provided in the next section and the final result is

$$E_{z \text{ cone}}(z) = -\vec{\nabla} \Phi_{\text{cone}}(z) \cdot \vec{e}_z \quad (\text{A39})$$

$$= -\frac{\partial \Phi_{\text{cone}}(z)}{\partial z} \quad (\text{A40})$$

$$\begin{aligned} &= \frac{\sigma}{2\epsilon} \left\{ \sin \alpha \cos \alpha \log \frac{L/\cos \alpha + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + L^2/\cos^2 \alpha}}{f_2(z) + \sqrt{R^2 + z^2}} \right. \\ &\quad + \frac{f_1(z) \cos \alpha (\cos \alpha + \frac{z}{\sqrt{R^2 + z^2}})}{f_2(z) + \sqrt{R^2 + z^2}} - f_1(z) \cos \alpha \frac{\cos \alpha + \frac{z+L}{\sqrt{z^2 + 2L(z+R) + R^2 + L^2/\cos^2 \alpha}}}{L/\cos \alpha + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + L^2/\cos^2 \alpha}} \\ &\quad \left. - \sin \alpha \left( \frac{z+L}{(z+L)^2 + (R+L \tan \alpha)^2} - \frac{z}{\sqrt{R^2 + z^2}} \right) \right\}. \end{aligned} \quad (\text{A41})$$

Multiplying Eq. (A41) by a positive unit test charge yields the axial forces acting on a positive unit test charge. The plot of the axial forces as a function of the position in the tunnel is displayed in Fig. 13 (lower panel) for the truncated cone geometry and compared to the cylinder case.

Experimental measurements made on ribosome exit tunnels show that the tunnel exit section radius is around 1 nm, i.e., twice the radius of the innermost part of the tunnel. If the ribosome tunnel were of the cone type, the cone opening angle would be around  $\alpha \sim 0.05$  radian (2.86 arc degrees).

The consequence on the electrostatic potential profile is of importance because, with this conical geometry, and if the total charges are kept the same for the two surfaces, the electrostatic potential inside the tunnel will necessarily be algebraically higher than the potential profile in the case of the cylinder as displayed in Fig. 13 (upper panel) where

the analytical equation for the electrostatic potential for the truncated cone was plotted and compared to the cylinder case.

A simple geometrical calculation shows that if the two surfaces support the same total charges  $Q_1 = Q_2$ , then  $\sigma_2 = S_{\text{cylinder}}/S_{\text{cone}} \times \sigma_1 = \frac{2}{3} \times \sigma_1$ , for a geometry where both tunnels have the same radius at the entry point, the same total lengths  $L$ , but where the cone exit section has a radius twice as large as the cylindrical radius. The surface charge density  $\sigma_2$  on the lateral truncated cone inner surface would be two third of the surface charge density  $\sigma_1$  prevailing on the lateral inner surface of the cylinder.

Moreover, the potential profile in the conical geometry is skewed to the left as compared to the potential profile for the cylindrical geometry. An asymmetry in the potential profile appears due to the change in radius along the  $z$ -axis of the cone. The minimal value of the potential is shifted to the left. The slope of the cylindrical potential profile is steeper than

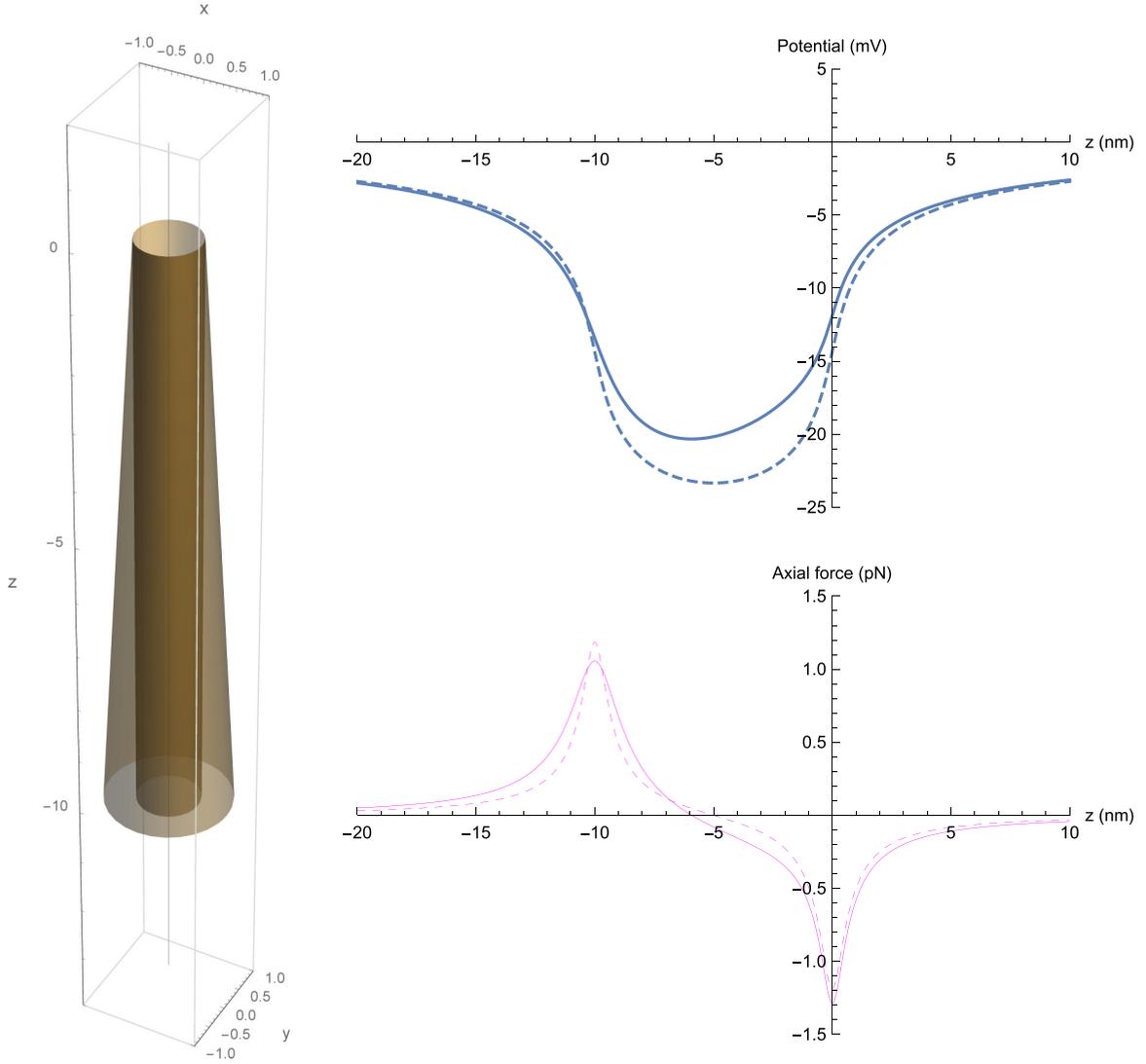


FIG. 13. Electrostatic scalar potential and axial force profiles for a positively unit charged test amino acid residue on the axis of the ribosomal exit tunnel. Comparison of the truncated cone (line) and cylinder (dashed line) geometry with exit section radius of the cone twice as large as the cylinder radius. The lateral surface of the cone is  $3/2$  the lateral surface of the cylinder.  $\sigma_2 = \sigma_1 \times 2/3$  to keep the same total charges on both surfaces.

the conical potential at the tunnel exit point, meaning that the electric field intensity will be a bit weaker in that region for the conical geometry as can be seen in Fig. 13 (lower panel) of the axial forces curves. The axial forces vary more smoothly and are more dispersed in the conical geometry than in the cylindrical geometry.

### 3. Normally truncated straight cone model complete derivation

#### a. Scalar potential

The scalar potential for the normally truncated straight cone model was expressed by Eq. (A34) as

$$\begin{aligned} \Phi_{\text{cone}}(z) &= \frac{\sigma L}{2\epsilon \cos \alpha} \int_0^1 \frac{(R + L \tan \alpha v) dv}{\sqrt{(z + Lv)^2 + (R + L \tan \alpha v)^2}}. \end{aligned} \quad (\text{A42})$$

The expression inside the square root in the denominator of the integrand can be written

$$\begin{aligned} &(z + Lv)^2 + (R + L \tan \alpha v)^2 \\ &= z^2 + L^2 v^2 + 2zLv + R^2 + L^2 \tan^2 \alpha v^2 + 2RL \tan \alpha v \\ &= L^2(1 + \tan^2 \alpha)v^2 + 2L(z + R \tan \alpha)v + z^2 + R^2 \\ &= \left[ \frac{L}{\cos \alpha} v + \cos \alpha (z + R \tan \alpha) \right]^2 + z^2 \\ &\quad + R^2 - \cos^2 \alpha (z + R \tan \alpha)^2 \\ &= \left[ \right]^2 + z^2 + R^2 - \cos^2 \alpha z^2 - 2zR \sin \alpha \cos \alpha \\ &\quad - R^2 \sin^2 \alpha \\ &= \left[ \right]^2 + z^2 \sin^2 \alpha + R^2 \cos^2 \alpha - 2zR \sin \alpha \cos \alpha \\ &= \left[ \right]^2 + (z \sin \alpha - R \cos \alpha)^2, \end{aligned} \quad (\text{A43})$$

and so the square root in the above denominator can be rewritten

$$\begin{aligned} & \sqrt{(z + Lv)^2 + (R + L \tan \alpha v)^2} \\ &= (z \sin \alpha - R \cos \alpha) \sqrt{\left[ \frac{Lv / \cos \alpha + (z \cos \alpha + R \sin \alpha)}{z \sin \alpha - R \cos \alpha} \right]^2 + 1}. \end{aligned}$$

To alleviate the notations, we pose as in (A35) and (A36)

$$f_1(z) = R \cos \alpha - z \sin \alpha, \quad (\text{A44})$$

$$f_2(z) = R \sin \alpha + z \cos \alpha, \quad (\text{A45})$$

and we pose

$$\begin{aligned} w &= \frac{Lv / \cos \alpha + (z \cos \alpha + R \sin \alpha)}{R \cos \alpha - z \sin \alpha} \\ &= \frac{Lv / \cos \alpha + f_2(z)}{f_1(z)}. \end{aligned} \quad (\text{A46})$$

Hence,

$$dw = \frac{L}{\cos \alpha} \frac{1}{f_1(z)} dv, \quad (\text{A47})$$

$$dv = \cos \alpha f_1(z) \frac{1}{L} dw. \quad (\text{A48})$$

The numerator in the integrand of (A38) now writes

$$R + L \tan \alpha v = R + (f_1(z)w - f_2(z)) \sin \alpha, \quad (\text{A49})$$

and Eq. (A34) turns into

$$\begin{aligned} \Phi_{\text{cone}}(z) &= \frac{\sigma L}{2\epsilon \cos \alpha} \int_*^* \frac{\frac{1}{L} \{R + [f_1(z)w - f_2(z)] \sin \alpha\} \cos \alpha f_1(z) dw}{f_1(z) \sqrt{w^2 + 1}} = \frac{\sigma}{2\epsilon} \int_*^* \frac{\{R + [f_1(z)w - f_2(z)] \sin \alpha\} dw}{\sqrt{w^2 + 1}} \\ &= \underbrace{\frac{\sigma}{2\epsilon} \int_*^* \frac{[R - f_2(z) \sin \alpha] dw}{\sqrt{w^2 + 1}}}_I + \underbrace{\frac{\sigma}{2\epsilon} \int_*^* \frac{f_1(z)w \sin \alpha dw}{\sqrt{w^2 + 1}}}_{II}. \end{aligned}$$

These two terms easily integrates. The first one (I) is still simplified further through

$$R - f_2(z) \sin \alpha = R - R \sin^2 \alpha - z \cos \alpha \sin \alpha = R \cos^2 \alpha - z \cos \alpha \sin \alpha = \cos \alpha (R \cos \alpha - z \sin \alpha) = \cos \alpha f_1(z), \quad (\text{A50})$$

and so

$$I = \frac{\sigma}{2\epsilon} \int_*^* \frac{\cos \alpha f_1(z) dw}{\sqrt{w^2 + 1}}. \quad (\text{A51})$$

Substituting  $w = \sinh y$ ,  $w^2 + 1 = \cosh^2 y$ , and  $dw = \cosh y dy$  yields

$$I = \frac{\sigma}{2\epsilon} \int_*^* \frac{\cos \alpha f_1(z) \cosh y dy}{\sqrt{\cosh^2 y}} = \frac{\sigma}{2\epsilon} \int_*^* \cos \alpha f_1(z) dy = \frac{\sigma}{2\epsilon} \cos \alpha f_1(z) y|_*^* \quad (\text{A52})$$

$$= \frac{\sigma}{2\epsilon} \cos \alpha f_1(z) \arg \sinh w|_*^*, \quad (\text{A53})$$

but  $\cosh y = \frac{e^y + e^{-y}}{2}$ ,  $\sinh y = \frac{e^y - e^{-y}}{2}$ , and  $\cosh y + \sinh y = e^y$ , so  $y = \arg \sinh w = \log[\cosh y + \sinh y] = \log(w + \sqrt{1 + w^2})$ . Hence,

$$\begin{aligned} I &= \frac{\sigma}{2\epsilon} \cos \alpha f_1(z) \log \left[ \frac{\frac{Lv}{\cos \alpha} + f_2(z)}{f_1(z)} + \sqrt{1 + \frac{[\frac{Lv}{\cos \alpha} + f_2(z)]^2}{f_1^2(z)}} \right]_{v=0}^{v=1} \\ &= \frac{\sigma}{2\epsilon} \cos \alpha f_1(z) \log \left[ \frac{\frac{L/\cos \alpha + f_2(z)}{f_1(z)} + \sqrt{1 + \frac{[L/\cos \alpha + f_2(z)]^2}{f_1^2(z)}}}{\frac{f_2(z)}{f_1(z)} + \sqrt{\frac{f_1^2(z) + f_2^2(z)}{f_1^2(z)}}} \right] \\ &= \frac{\sigma}{2\epsilon} \cos \alpha f_1(z) \log \left[ \frac{|\frac{L}{\cos \alpha} + f_2(z) + \sqrt{f_1^2(z) + [\frac{L}{\cos \alpha} + f_2(z)]^2}|}{|f_2(z) + \sqrt{R^2 + z^2}|} \right]. \end{aligned}$$

Noticing that

$$f_1^2(z) + f_2^2(z) = R^2 \cos^2 \alpha - 2Rz \cos \alpha \sin \alpha + z^2 \sin^2 \alpha + R^2 \sin^2 \alpha + 2Rz \sin \alpha \cos \alpha + z^2 \cos^2 \alpha = R^2 + z^2 \quad (\text{A54})$$

and that

$$f_1^2(z) + \frac{L^2}{\cos^2 \alpha} + 2\frac{L}{\cos \alpha} f_2(z) + f_2^2(z) = z^2 + R^2 + \frac{L^2}{\cos^2 \alpha} + \frac{2L}{\cos \alpha} (R \sin \alpha + z \cos \alpha) = z^2 + R^2 + \frac{L^2}{\cos^2 \alpha} + 2L(z + R \tan \alpha) \quad (\text{A55})$$

results in

$$I = \frac{\sigma}{2\epsilon} \cos \alpha f_1(z) \log \left[ \frac{|\frac{L}{\cos \alpha} + f_2(z) + \sqrt{z^2 + 2L(z + R) + R^2 + (\frac{L}{\cos \alpha})^2}|}{|f_2(z) + \sqrt{R^2 + z^2}|} \right]. \quad (\text{A56})$$

For the second term (II), substituting  $w = \sinh y$ ,  $\cosh^2 y = \sinh^2 y + 1 = w^2 + 1$ , and  $dw = \cosh y dy$ , we have

$$\begin{aligned} II &= \frac{\sigma}{2\epsilon} \int_*^* \frac{f_1(z) w \sin \alpha dw}{\sqrt{w^2 + 1}} = \frac{\sigma \sin \alpha}{2\epsilon} \int_*^* \frac{f_1(z) \sinh y \cosh y dy}{\sqrt{\cosh^2 y}} = \frac{\sigma \sin \alpha}{2\epsilon} \int_*^* f_1(z) \sinh y dy \\ &= \frac{\sigma \sin \alpha}{2\epsilon} f_1(z) [\cosh y]_*^* = \frac{\sigma \sin \alpha}{2\epsilon} f_1(z) [\sqrt{1 + w^2}]_*^* \\ &= \frac{\sigma \sin \alpha}{2\epsilon} f_1(z) \left\{ \sqrt{1 + \left( \frac{Lv}{\cos \alpha} + f_2(z) \right)_{v=1}^2} - \sqrt{1 + \left( \frac{Lv}{\cos \alpha} + f_2(z) \right)_{v=0}^2} \right\} \\ &= \frac{\sigma \sin \alpha}{2\epsilon} f_1(z) \left\{ \sqrt{1 + \left( \frac{L}{\cos \alpha} + f_2(z) \right)^2} - \sqrt{1 + \left( \frac{f_2(z)}{f_1(z)} \right)^2} \right\} \\ &= \frac{\sigma \sin \alpha}{2\epsilon} \left\{ \sqrt{f_1^2(z) + \left( \frac{L}{\cos \alpha} + f_2(z) \right)^2} - \sqrt{f_1^2(z) + f_2^2(z)} \right\} \\ &= \frac{\sigma \sin \alpha}{2\epsilon} \left\{ \sqrt{f_1^2(z) + \left( \frac{L}{\cos \alpha} + f_2(z) \right)^2} - \sqrt{R^2 + z^2} \right\} = \frac{\sigma \sin \alpha}{2\epsilon} \left\{ \sqrt{(z + L)^2 + (R + L \tan \alpha)^2} - \sqrt{R^2 + z^2} \right\}. \quad (\text{A57}) \end{aligned}$$

Summing the two terms I and II results in Eq. (A38).

### b. Electric field

The complete derivation of the electric field projected along the tunnel axis follows from

$$E_{z \text{ cone}}(z) = -\vec{\nabla} \Phi_{\text{cone}}(z) \cdot \vec{e}_z = -\frac{\partial \Phi_{\text{cone}}(z)}{\partial z} = -\frac{\partial I}{\partial z} - \frac{\partial II}{\partial z}. \quad (\text{A58})$$

We start with  $-\frac{\partial I}{\partial z}$ :

$$\begin{aligned} -\frac{\partial I}{\partial z} &= \frac{\sigma}{2\epsilon} \left\{ \sin \alpha \cos \alpha \log \left[ \frac{\bullet}{\bullet} \right] + f_1(z) \cos \alpha \frac{f_2(z) + \sqrt{R^2 + z^2}}{\frac{L}{\cos \alpha} + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + \frac{L^2}{\cos^2 \alpha}}} \right. \\ &\quad \left. \cdot \frac{\partial}{\partial z} \left[ \frac{\frac{L}{\cos \alpha} + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + \left(\frac{L}{\cos \alpha}\right)^2}}{f_2(z) + \sqrt{R^2 + z^2}} \right] \right\} \\ &= \frac{\sigma}{2\epsilon} \left( \sin \alpha \cos \alpha \log \left[ \frac{\bullet}{\bullet} \right] + f_1(z) \cos \alpha \frac{f_2(z) + \sqrt{R^2 + z^2}}{\frac{L}{\cos \alpha} + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + \frac{L^2}{\cos^2 \alpha}}} \right. \\ &\quad \times \left\{ \left( \frac{L}{\cos \alpha} + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + \left(\frac{L}{\cos \alpha}\right)^2} \right) [f_2(z) + \sqrt{R^2 + z^2}]^{-2} \left[ \cos \alpha + \frac{1}{2}(R^2 + z^2)^{-1/2} 2z \right] \right. \\ &\quad \left. \left. - [f_2(z) + \sqrt{R^2 + z^2}]^{-1} \left[ \cos \alpha + \frac{1}{2} \left( z^2 + 2L(z+R) + R^2 + \frac{L^2}{\cos^2 \alpha} \right)^{-1/2} (2z + 2L) \right] \right\} \right) \\ &= \frac{\sigma}{2\epsilon} \left[ \sin \alpha \cos \alpha \log \frac{L/\cos \alpha + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + L^2/\cos^2 \alpha}}{f_2(z) + \sqrt{R^2 + z^2}} \right. \\ &\quad \left. + \frac{f_1(z) \cos \alpha \left( \cos \alpha + \frac{z}{\sqrt{R^2 + z^2}} \right)}{f_2(z) + \sqrt{R^2 + z^2}} - f_1(z) \cos \alpha \frac{\cos \alpha + \frac{z+L}{\sqrt{z^2 + 2L(z+R) + R^2 + L^2/\cos^2 \alpha}}}{L/\cos \alpha + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + L^2/\cos^2 \alpha}} \right]. \quad (\text{A59}) \end{aligned}$$

We go on with  $-\frac{\partial II}{\partial z}$ :

$$\begin{aligned} -\frac{\partial II}{\partial z} &= -\frac{\sigma}{2\epsilon} \sin \alpha \frac{\partial [\sqrt{\bullet} - \sqrt{\bullet}]}{\partial z} = -\frac{\sigma}{2\epsilon} \sin \alpha \left[ \frac{1}{2} \left( (z+L)^2 + (R+L \tan \alpha)^2 \right)^{-1/2} 2(z+L) - \frac{1}{2} (R^2 + z^2)^{-1/2} 2z \right] \\ &= -\frac{\sigma}{2\epsilon} \sin \alpha \left[ \frac{z+L}{\sqrt{(z+L)^2 + (R+L \tan \alpha)^2}} - \frac{z}{\sqrt{R^2 + z^2}} \right]. \quad (\text{A60}) \end{aligned}$$

Summing the two terms yields the final result for  $E_{z \text{ cone}}(z)$  as in Eq. (A41):

$$\begin{aligned} E_{z \text{ cone}}(z) &= \frac{\sigma}{2\epsilon} \left\{ \sin \alpha \cos \alpha \log \frac{L/\cos \alpha + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + L^2/\cos^2 \alpha}}{f_2(z) + \sqrt{R^2 + z^2}} \right. \\ &\quad \left. + \frac{f_1(z) \cos \alpha \left( \cos \alpha + \frac{z}{\sqrt{R^2 + z^2}} \right)}{f_2(z) + \sqrt{R^2 + z^2}} - f_1(z) \cos \alpha \frac{\cos \alpha + \frac{z+L}{\sqrt{z^2 + 2L(z+R) + R^2 + L^2/\cos^2 \alpha}}}{L/\cos \alpha + f_2(z) + \sqrt{z^2 + 2L(z+R) + R^2 + L^2/\cos^2 \alpha}} \right. \\ &\quad \left. - \sin \alpha \left[ \frac{z+L}{(z+L)^2 + (R+L \tan \alpha)^2} - \frac{z}{\sqrt{R^2 + z^2}} \right] \right\}. \quad (\text{A61}) \end{aligned}$$

## 4. Material and methods for mapping the charged chemical groups on the tunnel inner wall from x-ray solved structures of the ribosomal large subunit

We analyzed the publicly available structure of the large ribosomal subunit of the archeon *H. marismortui* [41] obtained from x-ray crystallography at 2.4 Å [42]. To find the ribosome exit tunnel and extract the tunnel centerline coordinates, we used a tunnel search algorithm developed by Sehnal *et al.* [44], implemented in MOLE 2.0 and the web-based MOLEonline

2.0 tool publicly available online [45,46]. We used PyMOL (PyMOL Molecular Graphics System, Version 2.3.2) and exported the relevant selected atom positions cartesian coordinates to output files. These files were further processed with mapping algorithms coded in Python to build 2D positional maps of the charged chemical groups on or near the inner surface of the ribosome exit tunnel, as viewed from the tunnel centerline. We used as input starting point to search for the tunnel cavity, the coordinates of one of the two nonbridging

oxygen atoms bound to phosphorus in the nucleotide G2099 of the 23SrRNA. This point is known to be close to the PTC. Because of the enlargement in the vestibule area and to avoid artifacts in the search algorithm leading to tunnels with a common entry and multiple exits we used as end points of the tunnel cavity, the geometric center of six possible exit points located in the vestibule area, which were the nonbridging oxygen atoms bound to phosphorus of nucleotides C1426, C492, and A1442, the NH2 atom of R19 in L22, and the NZ atoms of K81 in L24 and K28 in eL39. In case the search algorithm yielded tunnels with multiple exit points sharing a common entry point, we used the midline obtained at the geometric center of the yielded tunnel centerlines and inspected manually the exit region in PyMOL. Once a correct tunnel centerline coordinates set was available, we used it as selection reference in PyMOL. We further used Python and Mathematica scripts to calculate different geometrical features of interest, surface charge density estimates and to produce 2D positional maps of the charged moieties as viewed from the tunnel centerline (Fig. 4). Specifically we translated the crystallographic data model space so that the tunnel entry point would be at the origin and we aligned the direction from the entry point to the exit point along the negative  $z$ -axis. We isolated 250 points along the tunnel centerline for which we calculated the curvilinear distance along the tunnel and determined 250 Frenet-Serret frames, i.e., the tangent, principal normal and binormal unit vectors forming a right-handed trihedron. We calculated the curvatures and torsions of the three dimensional space curve of the tunnel centerline (not shown here). We used the Frenet-Serret frames to determine 250 normal planes to the centerline three-dimensional space curve. We algorithmically scanned all the PyMOL selected atoms of interest to calculate the normal distance of the selected atoms to these 250 planes, found the closest intersection points, calculated the radial distance from the centerline and the elevation angle for the virtual line of sight of the selected atoms as viewed from the right-handed trihedron at the centerline points. The elevation angles of the selected atoms were calculated after two compound rotations of the Frenet-Serret frames so that the unit tangent was first aligned to the  $(0, 0, -1)$  direction, pointing to the negative  $z$ -axis; and the principal normal unit vector was then aligned to the  $(-1, 0, 0)$  direction, pointing toward the negative  $x$ -axis. This procedure is equivalent to a parallel transport of a right-handed trihedron reference frame along the three dimensional space curve of the tunnel centerline when moving from the tunnel entry to the exit points. This parallel transported reference frame is the one used for the angle mapping in Fig. 4. With this convention, 0 degree points toward the negative  $x$ -axis, +90 degrees points toward the positive  $y$ -axis, +180 degrees points toward the positive  $x$ -axis; whereas -90 degrees points toward the negative  $y$ -axis and -180 degrees points toward the positive  $x$ -axis. We algorithmically set out the 3D equations of the cylinder and the cone frustum in this reference frame, to calculate the closest distance of the selected atoms to the surface of the model tunnel. The  $\Delta$  values shown in blue in Fig. 4 were obtained as the closest (orthonormal) distance of the PyMOL selected atoms either from the model cylinder surface or from the model cone surface, depending on where the atoms are located.

### 5. Comparison of coarse-grained electrostatic potential calculated from the observed structural data with the geometrically idealized model

The Coulomb or Yukawa-Debye-Hückel electrostatic potential can be calculated from the x-ray solved exact distribution of the source charges (phosphate moieties and charged amino acids) for which the positional map was shown in Fig. 4. The data set includes the exact 3D coordinate positions of a total of 94 charged atoms that are closest to the tunnel centerline.

In the literature, the method to compute the electrostatic potential based on the real observed atom positions belongs to the coarse-grained modeling methods family. The Yukawa-Debye-Hückel potential is generally used (see for instance [80]) and the exact positions  $\vec{r}_i'$  of the sources and their charges  $q_i$  are summed over all source charges; see Eq. (A62) below.

In this formula, two phenomenological parameters are required which are  $\epsilon_r$ , the relative permittivity of the medium and  $l_D$ , the Debye screening length or  $\xi$  in our model. The Coulomb potential is a particular case of the Yukawa potential when the screening length goes to infinity.

In coarse-grained modeling, the assumption is made that the two phenomenological parameters are constant (homogeneous) in the media where the potential is computed. The standard or defaulted homogeneous values of these parameters are  $\epsilon_r = 78$  (water) and  $l_D = 10 \text{ \AA}$ . The formula also neglects surface charge polarization effects at dielectric media discontinuities. We showed in the paper that the Debye-Hückel theory is not quite appropriate in the confined environment of the tunnel. In particular, the Debye screening length to be used should be larger considering that, in the vicinity of the tunnel walls, the ions contribute weaker to the screening than they do in the inner core of the ribosome. Much weaker ionic strengths and larger Debye screening lengths should be used in the formula to be able to fit the experimentally observed values of the potential.

The electrostatic screening in the tunnel lumen is due to the permanent electric dipoles of constitutive water molecules or to the induced polarization in dielectric neutral media around the charged sources. In strongly confined environment we should resort to Gouy-Chapman screening lengths when there are much less water molecules or when the permittivity of the medium reaches the minimal value ( $\epsilon_{\text{protein}} = 4$ ).

For these reasons, and in a similar way as what was done with the full analytical idealized solutions of the electrostatic potential, we relaxed the homogeneity assumptions of the two phenomenological parameters and also implemented a piecewise continuous assumption for these parameters in the three tunnel regions of interest: cylinder part, cone part, and constriction site:

$$\Phi(z) = \sum_{k \in \text{regions}} \sum_{i \in \text{charged sources}} \frac{q_{i,k}}{4\pi\epsilon_0\epsilon_r(k)} \frac{e^{-\frac{|\vec{r}_i' - (0,0,z)|}{\xi_k}}}{|\vec{r}_i' - (0,0,z)|}. \quad (\text{A62})$$

In the above Coulomb-Yukawa electrostatic potential formula, different values of  $\epsilon_r(k)$  and different  $\xi_k$  screening lengths can be used in the different  $k$ -indexed regions (or

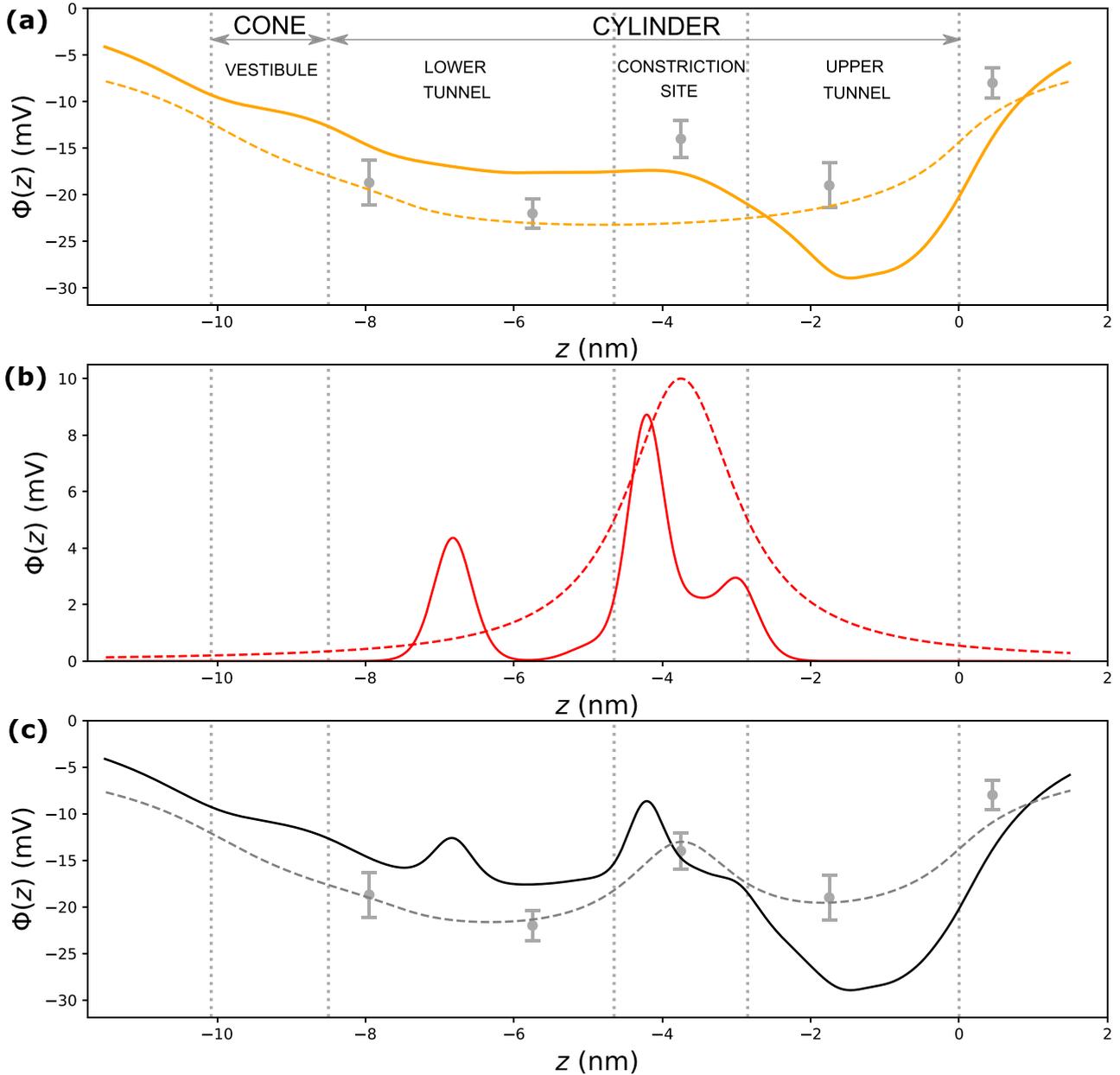


FIG. 14. Electrostatic potential calculated from exact positions of 94 charged atoms close to the tunnel centerline compared to our idealized model. (a) Contribution of phosphate moieties only. Orange solid line: potential profile calculated from mapped atomic positions; orange dashed line: potential profile calculated from idealized model; gray error bars: measured potential data points in [29]. (b) Contribution of the charged amino acid residues around constriction site. Red solid line: potential profile calculated from mapped atomic positions; red dashed line: Lorentzian potential profile from idealized model. (c) sum of (a) and (b) for the 94 atoms as mapped in Fig. 4 (black solid line); potential profile from idealized model (gray dashed line); and measured potential data points in [29] (gray error bars).

media). The elementary unit charge value of  $+|e|$  or  $-|e| = -1.602 \times 10^{-19}$  C is used for each of the charges  $q_i$  associated to the positively or negatively charged atoms at their given  $\vec{r}_i'$  positions. In the model of our paper, the sum over the  $q_i$  results from a surface averaged ensemble and is computed by an integral of the surface charge density over the surface element of interest. Formulas (1) and (25) are the surface integral equivalents of the above discrete sum formula (A62).

Upon implementing this formula (A62) in Python and using the exact positions of Fig. 4 mapped 94 charged atoms, we

obtained the electrostatic potential along the tunnel centerline shown in Fig. 14.

The set of parameters in Table II was used. We used the same set of dielectric responses as the ones estimated from the idealized model and the data fitting constraints of the observed potential data points. Explicitly, we used  $\epsilon_r = 8.3$  in the cylinder,  $\epsilon_r = 20.15$  in the cone, and  $\epsilon_r = 4$  in the constriction site. The screening lengths were adjusted to fit the experimentally observed electrostatic potential data points as well as possible.

It is difficult (or impossible) to get an electrostatic potential profile that fits well the experimentally observed potential

TABLE II. Phenomenological parameter values at standard temperature  $T = 298.15$  K in the tunnel regions of interest.

Region	$k$	Medium permittivity	Screening length		
			$\xi_k$	Ionic strength	$\xi_k$ retained
Tunnel cylinder	1	$\epsilon_k = 8.9$	$\xi_{\text{Debye}} = 2.48$ nm	$I = 15$ mM/L	$\xi_k = 2.48$ nm
Tunnel cone	2	$\epsilon_k = 20.15$	$\xi_{\text{Debye}} = 2.04$ nm	$I = 22$ mM/L	$\xi_k = 2.04$ nm
Constriction site (aa)	3	$\epsilon_k = 4$	$\xi_{\text{Gouy}} = 0.105$ nm		$\xi_k = 0.105$ nm

data points in [29] without resorting to different piecewise constant values of the phenomenological parameters. This shows the scope shortcomings or scope limitations of the current models, whether or not based on exact atomic positions, coarse-grained or ensemble approaches.

Although an inverse bell shape is observed, the results as plotted in Fig. 14 show some discrepancies between the potential profile calculated from the structural data and the data points in [29] or between the structural data potential and the potential profile as obtained from the idealized model (Fig. 2).

The comparison of the potential profile as obtained from the exact atomic positions (structural data) with the potential profile as calculated from our geometrically idealized model using the piecewise  $\sigma/\epsilon$  constant ratio values in the three different regions of concern, calls for the following comments:

(1) Except for the upper tunnel, the contribution of the phosphate moieties to the electrostatic potential profile calculated from the structural data (solid orange line) is parallel to the idealized profile where uniform surface charge distribution is assumed (dashed orange line). The local deeper well of the potential in the upper tunnel is due to the presence of phosphate groups that are closer to the tunnel centerline as compared to other regions; see Figs. 4(a) and 4(b).

(2) The net contribution of the positively charged amino acid residues (minus the negatively charged amino acid residues) calculated from the structural data show that these charged groups cause the peaks increasing the potential locally around the constriction site; see Fig. 4(b). The profiles of the peaks are very similar to the profiles of Lorentzian function as expected. This confirms the Lorentzian function can be used to approximate the contribution of the local enrichment in positively charged amino acid residues at the constriction site. The contribution of the dominant peak in the structural data comes from the close position of R125 (Arg125 of uL22) from the tunnel centerline (R125 is  $\sim 6$  Å from the centerline). The presence of the six or seven positively charged residues in general helps explaining the local rise in the potential as initially experimentally measured in [5].

(3) The transition from the lower tunnel (cylinder end port) to the vestibule (cone entry port) and further to the exit port of the cone is smooth and the potential calculated from the structural data (solid black line) is parallel to the potential profile as calculated from our model (dashed gray line) in Fig. 4(c). The explanation that both potential values diminish in magnitude irrespective of the increase in surface charge

density towards the cone end port comes from the fact that (1) most of the phosphate groups tend to get further away from the tunnel centerline; this is also true for the amino acid charged groups: see Figs. 4(b) and 4(d); (2) the dielectric response of the medium goes from an estimated average value of  $\epsilon = 8.3$  to  $\epsilon = 20.15$  when moving from the lower tunnel to the cone exit port; and (3) the Debye screening length decreases (and the Yukawa exponential factor decreases the potential) due to an expected increase in the ionic strength from the very weak 15 millimole per liter value (lower cylinder estimate) to the still weak 22 millimole per liter value (cone value estimate). The last two facts are both due to the expected increased proportion of water content in the tunnel lumen media and the increased polarizability along the  $z$ -axis centerline towards the cone exit port (from a water content of  $\sim 6\%$  in the cylinder to  $\sim 24\%$  in the cone).

## 6. Specific effect of the tunnel electrostatic interaction on the elongation rate

The electromechanical force due to the tunnel electrostatics acts on the peptide nascent chain and is transmitted inside the ribosomal tunnel up to the peptidyl transfer center (PTC) responsible for the peptide bond formation. The force is transmitted to the PTC through the whole length of the polypeptide chain backbone embedded in the tunnel [67]. At the PTC, the first event that must occur before the peptide bond is built between the peptidyl-tRNA at the P site and the aminoacylated tRNA at the A site is the breaking of the ester covalent bond between the oxygen atom attached on the tRNA 3'-end (3'-carbon at the CCA terminal ribose) and the carbonyl group of the carboxyl terminal end of the peptide. We presume that a force acting on the peptidyl-tRNA peptide directed from the P site toward the N-terminal end of the peptide would help breaking this ester bond. The chemical reaction rate of this ester bond breaking would be increased in the presence of such a force directed toward the exit tunnel. The ribosome elongation average rate can be quantitatively modulated by applying a Maxwell-Boltzmann factor, i.e.,  $\exp \frac{\int \vec{F}_z \cdot dz}{k_B T}$ , resulting from the theoretical treatment of the effect of force on the thermodynamics and kinetics of chemical reactions [71] or as initially introduced by Bell in a cell to cell adhesion context [72,73]. This factor correcting the elongation rate specifically accounts for the electrostatic interaction of the nascent chain in the ribosome exit tunnel. This factor is calculated on the basis of all the 50 residues upstream and is updated at each

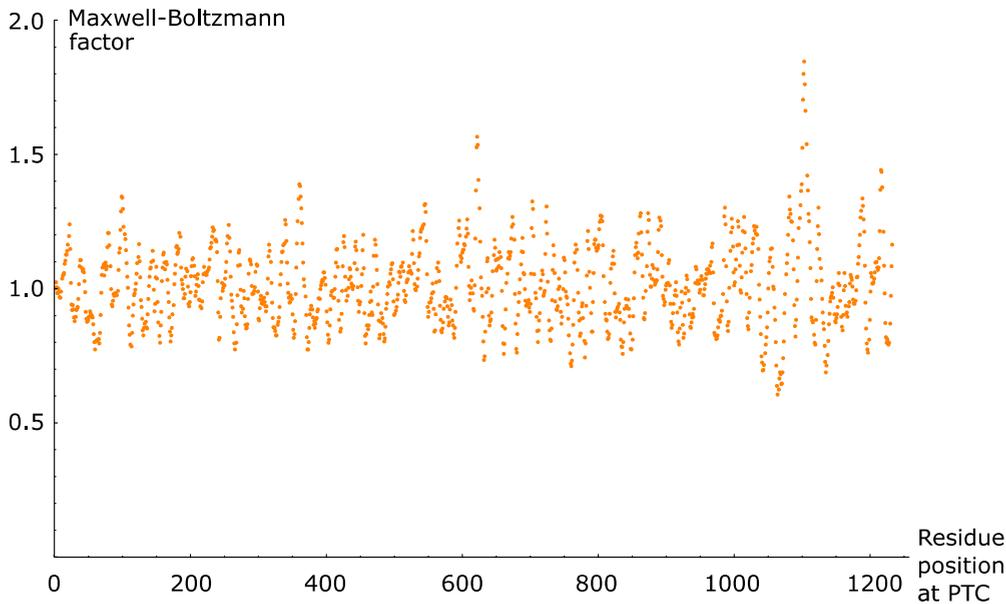


FIG. 15. Maxwell-Boltzmann elongation rate factors weighting for the electrostatic interaction at each incorporation of a new residue at PTC as a function of residue position in the human protein KIF4A. Min = 0.60 and Max = 1.85 occurring at incorporation of residues *E*1064 and *K*1103, respectively. Lower values are associated to smaller elongation rate (slowdown); larger values are associated to higher elongation rates (speeding up).

new incorporation. The numerical value of this factor will be different at each residue incorporation and will always be dependent on the particular amino acid sequence being embedded in the tunnel. For the arbitrarily chosen protein KIF4A (member of the family of human kinesins), all the numerical values for the Maxwell-Boltzmann factors calculated for each residue sequentially incorporated at the PTC are displayed in Fig. 15. The minimal value is 0.60 and the maximal value is 1.85 for the Maxwell-Boltzmann factor in the particular case of KIF4A. 84.5% of the values are in the range [0.80, 1.20]. The mean, 1.01, is very close to 1.0. The minimal value of the Maxwell-Boltzmann factor, 0.60, occurs at incorporation of residue 1064, which is a E (negatively charged glutamate), in KIF4A, when the axial force on the nascent polypeptide stretch in the tunnel is +8.62 pN. The elongation rate is quantitatively slowed down by a factor 0.60. Equivalently, the time spent by the ribosome on codon 1064 is expected to be larger (average time for this type of codon divided by 0.60) at this position because of the most unfavorable electrostatic interaction occurring at the moment of this residue incorporation. The maximal value of the Maxwell-Boltzmann factor, 1.85, occurs at incorporation of residue 1103 which is a K (positively charged lysine), in KIF4A, when the axial force on the nascent polypeptide stretch in the tunnel is  $-10.49$  pN. The elongation rate is quantitatively faster by a factor 1.85. Equivalently, the time spent by the ribosome on codon 1103 is expected to be smaller (average time for this type of codon divided by 1.85) at this position because of the most favorable electrostatic interaction occurring at the moment of this residue incorporation. This illustrates how the Maxwell-Boltzmann factors provide a consistent methodological tool to assess quantitatively the contribution to the elongation rate specifically due to the electrostatic interaction occurring in the tunnel, and in a separate

way from the other factors affecting the mRNA translation rate.

### 7. Energy sources available for the ribosome

The ribosome is a complex macromolecular machine that requires energy to carry out its multiple tasks. During elongation, a ribosome has to translocate the mRNA each time a codon has been paired to its cognate or semicognate tRNA and has to push the nascent protein through the exit tunnel.

The detailed energy balance (energy sources and uptakes) required for elongation has not been fully resolved. Our electrostatic model shows that, in certain situations, a Gibbs free energy fraction in the range 15%–30% of the total biochemical energy available to the ribosome could be required to move the nascent protein through the exit tunnel.

The energy is found in the biochemical reactions taking place in the ribosome with the help of the associated catalytic sites of enzymes like the elongation factors (eEF in eukaryotes) or ribozymes. The elongation factors (EF and EF-G) are GTPases whose activity is controlled by the ribosome. When an aminoacyl group is hydrolyzed from the loaded tRNA, an ester group is broken and energy is released. For each amino acid incorporation cycle, two GTPs molecules are hydrolysed (one with the help of EF in the ternary complex accommodated at the A-site and one with the help of EF-G required for the mechanical translocation). The peptide bond formation itself requires free energy at each chain elongation by one residue. A very rough estimate of the net change in Gibbs free energy for the net balance between peptide bond formation and ester hydrolysis at pH = 7, 25 °C yields  $\Delta G^\circ = -3.7 \pm 1.2$  kcal/mol =  $-15.5 \pm 5.0$  kJ/mol [32,33]. This is known as the transpeptidation Gibbs free energy.

(1) *Peptide bond formation.* the formation of the simplest dipeptide glycylglycine is endergonic and requires 15 kJ/mol (3.6 kcal/mol) per mole of formed peptidic bond:  $\Delta G^\circ = +3.6$  kcal/mol for one residue incorporation (per ribosome cycle).

(2) *Hydrolysis of ester bond in aminoacyl-tRNA.* the hydrolysis of the ester bond in aa-tRNA is exergonic and releases 30.5 kJ/mol (7.3 kcal/mol) per amino acid released from the tRNA:  $\Delta G^\circ = -7.3$  kcal/mol (per ribosome cycle)

(3) *Hydrolysis of 2 GTPs.* the hydrolysis of 2 GTPs is exergonic and releases 30.5 kJ/mol (7.3 kcal/mol) per mole of GTP. Hence, per residue incorporation cycle (two GTPs):  $\Delta G^\circ = -14.6$  kcal/mole (per ribosome cycle)

(4) *Net Gibbs free energy available to the ribosome per aa residue incorporation.*  $\Delta G^\circ = -18.3$  kcal/mol (per residue incorporation)

The net result is that one ester bond to the 3'-hydroxyl of a ribose has been broken (locally in the ribosome) and one peptide bond in the nascent protein has been formed, two GTPs have been hydrolyzed, the ribosome has shifted forward the mRNA by one codon (translocation distance on mRNA,  $\Delta x \sim 1.4$  nm (0.9–1.8), parentheses indicate 95% confidence limits [33]) and the nascent peptide has advanced in the ribosome exit tunnel by one residue (nascent peptide chain distance displacement in the tunnel at each translocation,  $\Delta z \sim 0.25$  nm, which is the estimated distance between two consecutive amino acid  $\alpha$ -carbons as considered in our model). It is not fully elucidated whether (or how) free energy could be stored in the ribosome and used later to catalyze translocation and possibly assist the progression of the nascent protein through the ribosome exit tunnel when needed. Each step in translation involves intrasubunit or intersubunit conformational changes [32–34]. Such conformational changes could store energy that could

be released at a subsequent step, with a thermodynamical yield, providing a conceivable mechanism of harnessing the biochemical energy to use it for mechanical translocation and for moving the nascent peptide through the ribosome exit tunnel when required. The entropy-driven spontaneous or chaperones assisted folding of the protein, generating a tugging force [22] outside of the ribosome exit tunnel, might also help the nascent protein to be pulled out of the tunnel. Optical tweezers assays have opened the way to characterizing the ribosome's full mechanochemical cycle [33,34]. Recently, such *in vitro* assays [33,34] provided an estimate for the maximal mechanical energy required per translocation step (near stalling on the mRNA), 21.2 pN nm = 5.2  $k_B T$ , at 296 K, or  $\sim 3.1$  kcal/mol. As estimated above, the Gibbs free energy available from the transpeptidation step (ester hydrolysis and peptide formation without the help of GTP hydrolysis) is  $\Delta G^\circ = -3.7 \pm 1.2$  kcal/mol. The mechanical work for translocation would be around 80% of the Gibbs free energy available from the transpeptidation. Such a high thermodynamic efficiency for conversion of chemical energy to mechanical motion is higher than occurs in most molecular motor [71]. Instead, efficient translocation would require the hydrolysis of at least one GTP with the help of elongation factor EF-G [33]. EF-G-dependent GTP hydrolysis was shown to precede and greatly accelerate translocation [81]. The mechanical translocation of the ribosome on the mRNA by one codon would take  $\frac{3.1 \text{ kcal/mol}}{7.3 \text{ kcal/mol}}$  or 43% of the Gibbs free energy released by the hydrolysis of one GTP, assisted by elongation factor EF-G. The mechanical energy required to push the nascent peptide chain through the large subunit exit tunnel could be provided by the transpeptidation Gibbs free energy or the hydrolysis of one GTP assisted by elongation factor EF in the ternary complex accommodated in the A site or a combination of both.

- 
- [1] M. Rodnina and W. Wintermeyer, Fidelity of aminoacyl-tRNA selection on the ribosome: Kinetic and structural mechanism, *Annu. Rev. Biochem.* **70**, 415 (2001).
- [2] M. Beringer and M. Rodnina, The ribosomal peptidyl transferase, *Mol. Cell* **26**, 311 (2007).
- [3] P. Nissen, J. Hansen, N. Ban, P. B. Moore, and T. A. Steitz, The structural basis of ribosome activity in peptide bond synthesis, *Science* **289**, 920 (2000).
- [4] C. T. MacDonald, J. H. Gibbs, and A. C. Pipkin, Kinetics of biopolymerization on nucleic acid templates, *Biopolymers* **6**, 1 (1968).
- [5] C. T. MacDonald and J. H. Gibbs, Concerning the kinetics of polypeptide synthesis on polyribosomes, *Biopolymers* **7**, 707 (1969).
- [6] T. von der Haar, Mathematical and computational modelling of ribosomal movement and protein synthesis: An overview, *Comput. Struct. Biotechnol. J.* **1**, e201204002 (2012).
- [7] H. Zur and T. Tuller, Predictive biophysical modeling and understanding of the dynamics of mRNA translation and its evolution, *Nucleic Acids Research* **44**, 9031 (2016).
- [8] N. T. Ingolia, L. F. Lareau, and J. S. Weissman, Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes, *Cell* **147**, 789 (2011).
- [9] T. Morisaki, K. Lyon, K. F. DeLuca, J. G. DeLuca, B. P. English, Z. Zhang, L. D. Lavis, J. B. Grimm, S. Viswanathan, L. L. Looger, T. Lionnet, and T. J. Stasevich, Real-time quantification of single RNA translation dynamics in living cells, *Science* **352**, 1425 (2016).
- [10] A. Fluitt, E. Pienaar, and H. Viljoen, Ribosome kinetics and aa-tRNA competition determine rate and fidelity of peptide synthesis, *Comput. Biol. Chem.* **31**, 335 (2007).
- [11] J. Lu and C. Deutsch, Electrostatics in the ribosomal tunnel modulate chain elongation rates, *J. Mol. Biol.* **384**, 73 (2008).
- [12] T. Tuller, I. Veksler-Lublinsky, N. Gazit, M. Kupiec, E. Ruppim, and M. Ziv-Ukelson, Composite effects of gene determinants on the translation speed and density of ribosomes, *Genome Biol.* **12**, R110 (2011).
- [13] M. V. Rodnina, The ribosome in action: Tuning of translational efficiency and protein folding, *Protein Science* **25**, 1390 (2016).
- [14] A. Sharma, N. Ahmed, and E. O'Brien, Determinants of translation speed are randomly distributed across transcripts resulting in a universal scaling of protein synthesis times, *Phys. Rev. E* **97**, 022409 (2018).

- [15] P. Shah, Y. Ding, M. Niemczyk, G. Kudla, and J. B. Plotkin, Rate-limiting steps in yeast protein translation, *Cell* **153**, 1589 (2013).
- [16] A. Riba, N. Di Nanni, N. Mittal, E. Arhné, A. Schmidt, and M. Zavolan, Protein synthesis rates and ribosome occupancies reveal determinants of translation elongation rates, *Proc. Nat. Acad. Sci. USA* **116**, 15023 (2019).
- [17] A. Dana and T. Tuller, The effect of tRNA levels on decoding times of mRNA codons, *Nucleic Acids Research* **42**, 9171 (2014).
- [18] A. Raveh, M. Margaliot, E. D. Sontag, and T. Tuller, A model for competition for ribosomes in the cell, *J. R. Soc., Interface* **13**, 20151062 (2016).
- [19] M. Y. Pavlov, R. E. Watts, Z. Tan, V. W. Cornish, M. Ehrenberg, and A. C. Forster, Slow peptide bond formation by proline and other *N*-alkylamino acids in translation, *Proc. Nat. Acad. Sci. USA* **106**, 50 (2009).
- [20] K. Dao Duc and Y. Song, The impact of ribosomal interference, codon usage, and exit tunnel interactions on translation elongation rate variation, *PLoS Genet.* **14**, e1007166 (2018).
- [21] J.-R. Yang, X. Chen, and J. Zhang, Codon-by-codon modulation of translational speed and accuracy via mRNA folding, *PLoS Biol.* **12**, e1001910 (2014).
- [22] L. J. Simpson, E. Tzima, and J. S. Reader, Mechanical forces and their effect on the ribosome and protein translation machinery, *Cells* **9**, 650 (2020).
- [23] T. E. Goroehowski, Z. Ignatova, R. A. L. Bovenberg, and J. A. Roubos, Trade-offs between tRNA abundance and mRNA secondary structure support smoothing of translation elongation rate, *Nucleic Acids Res.* **43**, 3022 (2015).
- [24] C. G. Artieri and H. B. Fraser, Accounting for biases in riboprofiling data indicates a major role for proline in stalling translation, *Genome Research* **24**, 2011 (2014).
- [25] R. Sabi and T. Tuller, A comparative genomics study on the effect of individual amino acids on ribosome stalling, *BMC Genomics* **16**, S5 (2015).
- [26] R. D. Requião, H. J. de Souza, S. Rossetto, T. Domitrovic, and F. L. Palhano, Increased ribosome density associated to positively charged residues is evident in ribosome profiling experiments performed in the absence of translation inhibitors, *RNA Biol.* **13**, 561 (2016).
- [27] C. A. Charneski and L. D. Hurst, Positively charged residues are the major determinants of ribosomal velocity, *PLoS Biol.* **11**, e1001508 (2013).
- [28] N. Voss, M. Gerstein, T. Steitz, and P. Moore, The geometry of the ribosomal polypeptide exit tunnel, *J. Mol. Biol.* **360**, 893 (2006).
- [29] J. Lu, W. Kobertz, and C. Deutsch, Mapping the electrostatic potential within the ribosomal exit tunnel, *J. Mol. Biol.* **371**, 1378 (2007).
- [30] C. Pace, G. Grimsley, and J. M. Scholtz, Protein ionizable groups: pK values and their contribution to protein stability and solubility, *J. Biol. Chem.* **284**, 13285 (2009).
- [31] J.-D. Wen, L. Lancaster, H. C. Hodges, A. Zeri, S. Yoshimura, H. Noller, C. Bustamante, and I. Tinoco, Following translation by single ribosomes one codon at a time, *Nature (London)* **452**, 598 (2008).
- [32] C. M. Kaiser and I. Tinoco, Jr., Probing the mechanisms of translation with force, *Chem. Rev.* **114**, 3266 (2014).
- [33] T. Liu, A. Kaplan, L. Alexander, S. Yan, J.-D. Wen, L. Lancaster, C. E. Wickersham, K. Fredrick, H. Noller, I. Tinoco, Jr., and C. J. Bustamante, Direct measurement of the mechanical work during translocation by the ribosome, *eLife* **3**, e03406 (2014).
- [34] V. P. Desai, F. Frank, A. Lee, M. Righini, L. Lancaster, H. F. Noller, I. Tinoco, Jr., and C. Bustamante, Co-temporal force and fluorescence measurements reveal a ribosomal gear shift mechanism of translation regulation by structured mRNAs, *Mol. Cell* **75**, 1007 (2019).
- [35] L. Shaw, R. Zia, and K. Lee, Totally asymmetric exclusion process with extended objects: A model for protein synthesis, *Phys. Rev. E* **68**, 021910 (2003).
- [36] R. K. P. Zia, J. J. Dong, and B. Schmittmann, Modeling translation in protein synthesis with TASEP: A tutorial and recent developments, *J. Stat. Phys.* **144**, 405 (2011).
- [37] P. Greulich, L. Ciandrini, R. Allen, and M. Romano, Mixed population of competing totally asymmetric simple exclusion processes with a shared reservoir of particles, *Phys. Rev. E* **85**, 011142 (2012).
- [38] S. D. Fried and S. G. Boxer, Electric fields and enzyme catalysis, *Annu. Rev. Biochem.* **86**, 387 (2017).
- [39] K. Dao Duc, S. Batra, N. Bhattacharya, J. H. D. Cate, and Y. S. Song, Differences in the path to exit the ribosome across the three domains of life, *Nucleic Acids Res.* **47**, 4198 (2019).
- [40] J. D. Jackson, *Classical Electrodynamics*, 3rd. ed. (John Wiley & Sons, New York, 1998), p. 32.
- [41] PDB code 4V9F, <https://www.rcsb.org>.
- [42] A. Gabdulkhakov, S. Nikonov, and M. Garber, Revisiting the *Haloarcula marismortui* 50S ribosomal subunit model, *Acta Crystallogr. D* **69**, 997 (2013).
- [43] G. Rhodes, *Crystallography Made Crystal Clear*, 3rd. ed. (Academic Press, San Diego, 2006), pp. 183–185.
- [44] D. Sehnal, R. Svobodová Vařeková, K. Berka, L. Pravda, V. Navrátilová, P. Banáš, C.-M. Ionescu, M. Otyepka, and J. Koča, MOLE 2.0: Advanced approach for analysis of biomacromolecular channels, *J. Cheminf.* **5**, 39 (2013).
- [45] K. Berka, O. Hanak, D. Sehnal, P. Banáš, V. Navrátilová, D. Jaiswal, C.-M. Ionescu, R. Svobodová Vařeková, J. Koca, and M. Otyepka, MOLEonline 2.0: Interactive web-based analysis of biomacromolecular channels, *Nucleic Acids Res.* **40**, W222 (2012).
- [46] L. Pravda, D. Sehnal, D. Toušek, V. Navrátilová, V. Bazgier, K. Berka, R. Svobodová Vařeková, J. Koča, and M. Otyepka, MOLEonline: A web-based tool for analyzing channels, tunnels and pores (2018 update), *Nucleic Acids Res.* **46**, W368 (2018).
- [47] M. S. Sansom, G. R. Smith, C. Adcock, and P. C. Biggin, The dielectric properties of water within model transbilayer pores, *Biophys. J.* **73**, 2404 (1997).
- [48] K. Sharp and B. Honig, Electrostatic interactions in macromolecules: Theory and applications, *Annu. Rev. Biophys. Biophys. Chem.* **19**, 301 (1990).
- [49] A. Cuervo, P. D. Dans, J. L. Carrascosa, M. Orozco, G. Gomila, and L. Fumagalli, Direct measurement of the dielectric polarization properties of DNA, *Proc. Nat. Acad. Sci. USA* **111**, E3624 (2014).
- [50] D. C. Harris, *Quantitative Chemical Analysis*, 2nd ed. (W. H. Freeman and Company, New York, 1988), pp. 90–92.
- [51] K. Nierhaus, Mg<sup>2+</sup>, K<sup>+</sup>, and the ribosome, *J. Bacteriol.* **196**, 3817 (2014).

- [52] C. Hsiao and L. Williams, A recurrent magnesium-binding motif provides a framework for the ribosomal peptidyl transferase center, *Nucleic Acids Res.* **37**, 3134 (2009).
- [53] A. Petrov, C. Bernier, C. Hsiao, C. Okafor, E. Tannenbaum, J. Stern, E. Gaucher, D. Schneider, N. Hud, S. Harvey, and L. Williams, RNA-magnesium-protein interactions in large ribosomal subunit, *J. Phys. Chem. B* **116**, 8113 (2012).
- [54] D. Klein, P. Moore, and T. Steitz, The contribution of metal ions to the structural stability of the large ribosomal subunit, *RNA* **10**, 1366 (2004).
- [55] M. Rodnina, M. Beringer, and W. Wintermeyer, Mechanism of peptide bond formation on the ribosome, *Q. Rev. Biophys.* **39**, 203 (2006).
- [56] M. Simonovic and T. Steitz, A structural view on the mechanism of the ribosome-catalyzed peptide bond formation, *Biochim. Biophys. Acta, Gene Regul. Mech.* **1789**, 612 (2009).
- [57] D. J. Lockhart and P. S. Kim, Electrostatic screening of charge and dipole interactions with the helix backbone, *Science* **260**, 198 (1993).
- [58] iChemLabs LLC, ChemDoodle 3D Version 6.2.2, <https://www.chemdoodle.com/3d> (2021).
- [59] R. van Roij, Electrostatics in liquids: Electrolytes, suspension, and emulsions, in lecture notes, Institute of Theoretical Physics, Utrecht University, Utrecht, the Netherlands (unpublished).
- [60] M. T. Friberg, P. Gonnet, Y. Barral, N. N. Schraudolph, and G. H. Gonnet, Measures of codon bias in yeast, the tRNA pairing index and possible DNA repair mechanisms, in *Proc. 6th Workshop on Algorithms in Bioinformatics (WABI)*, edited by P. Bucher and B. Moret, Lecture Notes in Bioinformatics Vol. 4175 (Springer Verlag, Berlin, 2006), pp. 1–11.
- [61] G. Cannarozzi, N. Schraudolph, M. Faty, P. von Rohr, M. Friberg, A. Roth, P. Gonnet, G. Gonnet, and Y. Barral, A role for codon order in translation dynamics, *Cell* **141**, 355 (2010).
- [62] C. Deutsch, Tunnel vision: Insights from biochemical and biophysical studies, in *Regulatory Nascent Polypeptides*, edited by K. Ito (Springer, Tokyo, 2014), pp. 61–86.
- [63] D. M. Hatters, Protein misfolding inside cells: The case of huntingtin and Huntington's disease, *IUBMB Life* **60**, 724 (2008).
- [64] M. Arrasate and S. Finkbeiner, Protein aggregates in Huntington's disease, *Exp. Neurol.* **238**, 1 (2012).
- [65] M. Chen and P. Wolynes, Aggregation landscapes of huntingtin exon 1 protein fragments and the critical repeat length for the onset of Huntington's disease, *Proc. Nat. Acad. Sci. USA* **114**, 4406 (2017).
- [66] S. Bonfanti, M. C. Lionetti, M. R. Fumagalli, V. R. Chirasani, G. Tiana, N. V. Dokholyan, S. Zapperi, and C. A. M. La Porta, Molecular mechanisms of heterogeneous oligomerization of huntingtin proteins, *Sci. Rep.* **9**, 7615 (2019).
- [67] B. Fritch, A. Kosolapov, P. Hudson, D. A. Nissley, H. L. Woodcock, C. Deutsch, and E. P. O'Brien, Origins of the mechanochemical coupling of peptide bond formation to protein synthesis, *J. Am. Chem. Soc.* **140**, 5077 (2018).
- [68] D. Goldman, C. Kaiser, A. Milin, M. Righini, I. Tinoco, and C. Bustamante, Mechanical force releases nascent chain-mediated ribosome arrest in vitro and in vivo, *Science* **348**, 457 (2015).
- [69] M. Thommen, W. Holtkamp, and M. V. Rodnina, Co-translational protein folding: Progress and methods, *Curr. Opin. Struct. Biol.* **42**, 83 (2017).
- [70] R. D. Requião, L. Fernandes, H. J. A. de Souza, S. Rossetto, T. Domitrovic, and F. L. Palhano, Protein charge distribution in proteomes and its impact on translation, *PLoS Comput. Biol.* **13**, e1005549 (2017).
- [71] C. Bustamante, Y. R. Chemla, N. R. Forde, and D. Izhaky, Mechanical processes in biochemistry, *Annu. Rev. Biochem.* **73**, 705 (2004).
- [72] J. Ribas-Arino and D. Marx, Covalent mechanochemistry: Theoretical concepts and computational tools with applications to molecular nanomechanics, *Chem. Rev.* **112**, 5412 (2012).
- [73] G. I. Bell, Models for the specific adhesion of cells to cells, *Science* **200**, 618 (1978).
- [74] N. Baker, D. Sept, S. Joseph, M. Holst, and J. McCammon, Electrostatics of nanosystems: Application to microtubules and the ribosome, *Proc. Nat. Acad. Sci. USA* **98**, 10037 (2001).
- [75] D. Lucent, C. Snow, C. Aitken, and V. Pande, Non-bulk-like solvent behavior in the ribosome exit tunnel, *PLoS Comput. Biol.* **6**, e1000963 (2010).
- [76] E. Mercier and M. Rodnina, Co-translational folding trajectory of the HEMK helical domain, *Biochemistry* **57**, 3460 (2018).
- [77] M. Liutkute, M. Maiti, E. Samatova, J. Enderlein, and M. V. Rodnina, Gradual compaction of the nascent peptide during cotranslational folding on the ribosome, *Elife* **9**, e60895 (2020).
- [78] T. Tuller, A. Carmi, K. Vestsigian, S. Navon, Y. Dorfan, J. Zaborske, T. Pan, O. Dahan, I. Furman, and Y. Pilpel, An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* **141**, 344 (2010).
- [79] D. A. Nissley, Q. V. Vu, F. Trovato, N. Ahmed, Y. Jiang, M. S. Li, and E. P. O'Brien, Electrostatic interactions govern extreme nascent protein ejection times from ribosomes and can delay ribosome recycling, *J. Am. Chem. Soc.* **142**, 6103 (2020).
- [80] B. R. Brooks *et al.*, CHARMM: The biomolecular simulation program, *J. Comput. Chem.* **30**, 1545 (2009).
- [81] M. Rodnina, A. Savelsbergh, V. Katunin, and W. Wintermeyer, Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome, *Nature (London)* **385**, 37 (1997).