Supporting Information

Petrone *et al.* 10.1073/pnas.0801795105

SI Text

Methods. ***Modeling of the ribosome tunnel.*** We used a cut-out of ≈80,000 atoms cropped out from *Haloarcula marismortui* large-subunit (PDB ID code 1S72) (1) crystal structure, capping protein and RNA at the boundaries. The dimensions of the model are $82.5 \times 85 \times 110$ Å^3, and it encompasses the ribosomal exit tunnel and its surrounding areas (Fig. S14) as well as a buffer of exterior water. All atoms at a distance of 10 Å from the inner available volume were frozen. Missing L39E residues (Glu, Val, and Gln) were modeled and energy minimized (see *Simulation details*, below). Parameters for RNA bases pseudouridine and 1-methyl-adeninosine were taken from the AMBER parameter database (http://pharmacy.man.ac.uk/amber/). The remaining protein and RNA parameters come from the AMBER99-ϕ potential set (2). Initial coordinates for crystallographic water, Na^+, Mg^{2+}, K^+, and Cl^- were retained. A Cd atom was replaced with Zn. The system was solvated with TIP3P (3) water. Excess negative charge was neutralized by replacing 338 water molecules with sodium ions.

The x-ray structure provides no information regarding large-scale conformation change of the ribosome in general and specifically those given by the presence of protein inside the tunnel. Cryo-EM data suggests that the RNA helices lining the ribosomal tunnel might not move significantly in the maps of a ribosome stalled with nascent chain compared to the peptide-free ribosome (4). To our knowledge, this study is the only resource available at this time, but not enough to prove that the tunnel does not significantly change shape.

Modeling of the amino acid and ion test probes. To examine physico-chemical properties of the ribosome tunnel, we probed our model with a set of side-chain analogs: tryptophan, alanine, isoleucine, aspartate, and lysine. Even though this set does not attempt to cover the entire amino acid chemical space, it does include aromatic, small, hydrophobic, and charged side-chain species. In addition, we probed the tunnel with two ion species: sodium (Na^+) and chloride (Cl^-).

Because the main focus of our work is the sequence specificity of the ribosome tunnel for different amino acid species, we did not model backbone atoms. To be consistent with the AMBER99-ϕ potential set (2), we modeled amino acid side-chain analogs from standard amino acid AMBER99 libraries as described by Shirts *et al.* in 2004 (5). Starting with a standard amino acid, we replaced backbone atoms (N, C, O, and Cα) with a hydrogen atom bonded to Cβ. All hydrogen atoms bonded to a Cβ are equivalent and bear the same electrostatic charge. Cβ electrostatic charge is modified to keep the molecule neutral (Ile, Trip, Ala), or charged (Asp, Lys).

Probe molecule placement. The solvent-accessible volume of the exit tunnel was divided into a grid of 1-Å^3 cells. Several copies of the chemical probes (amino acid analogs or ion probes) were inserted in the ribosome model (as shown in Fig. S14*B* and tethered to points on the grid (anchors) by a harmonic potential with spring constant $k = 100$ KJ/mol/nm^2 and rest elongation $r_o = 0$ Å.

Multiple probe molecules were run in the same simulation provided that their anchors were separated at least 18 or 24 Å for neutral and charged probes, respectively. This separation distance between mean probe locations is greater than the cut-off for electrostatic and Van der Waals interactions ($r_c = 10$ Å). The probability of probes interacting is low because probes do not generally sample >7 Å from their anchor, as shown in the distance distribution for the tryptophan probe in bulk solvent (Fig. S8).

For efficient coverage of all anchor locations, simulations had different numbers of probes. For this reason, in the case of charged probes (Asp, Lys, Na^+, Cl^-), simulations also had different numbers of Na^+ ions to neutralize the net charge of the system.

Simulation details. We used the all-atom force field of AMBER 99-ϕ modified by Sorin *et al.* (2). A cut-off of 10 Å was used for nonbonded electrostatic interactions. Long-range (>10 Å) electrostatic interactions were approximated with a reaction-field algorithm implemented in GROMACS 3.1 (6), with a dielectric constant of 80. We simulated each system at constant N, $P = 1$ atm, $T = 298$ K, using Berendsen temperature and pressure coupling (7). The initial ribosome model (before insertion of probes) had been equilibrated for 3.5 ns. Once the probes were inserted, the system was subject to energy minimization, followed by 200 ps of molecular dynamics (MD) simulation before production simulations. The integration step was 0.002 ps. Configurations were saved for subsequent analysis at the rate of one snapshot per picosecond. Production varied for different anchors, but an average of 295,000 snapshots were collected for the amino acid probes. Total sampling amounted to 295 ns, divided in trajectories of ≈20 ns long.

Control Simulation. These simulations include only the model cut-out, explicit water, and ions. Starting from the equilibrated conformation, trajectories are divided into 100 runs with 100 clones each, differing in their initial velocities. These runs were parallelized in the Folding@Home distributed computing network. Combining clone simulations, the total production for the system is 12,159.5 ns. However, it should be noted that average trajectory length is ≈25 ns.

Simulation with Side-Chain Analogs. Probe molecules were inserted at 280 grid points as described above. One could, in principle, try to run 280 simulations with one probe in each simulation, or to run one simulation with 280 probes. The former case is computationally inefficient; the latter case forces undesirable direct and indirect interactions between pairs of probes. Instead, we strategically sample multiple probes per simulation, ensuring that the anchor points for all cosimulated probes are at least 18 Å apart (or 24 Å for charged amino acid analogs). As a result, we had to run only 27 different simulation configurations (runs) for noncharged side-chain analogs. Because they were simulated at a larger separation, charged probes were divided among 64 runs. To more rapidly sample the system in parallel, each run was replicated into 40 separate simulations that differed only by their initial velocities (sampled from a Maxwell-Boltzmann distribution at 298 K). The trajectory of each clone extended, on average, 20 ns, adding to a total production of 300-ns sampling per anchor. Cl^- and Na^+ ion probes were treated by following the same procedure as with the charged amino acid probes.

Supplementary Data and Error Analysis. We collected data from probe molecules throughout the accessible tunnel volume. Data were analyzed by using the WHAM (8–12) algorithm, which we extended to three dimensions.

The 3D-WHAM algorithm. WHAM (the extended Ferrenberg-Swendensen algorithm) (9) is a free-energy calculation method designed to obtain an accurate potential of mean force (PMF) by maximizing the amount of information obtained from a

Monte Carlo or MD simulation (12). As in other "umbrella sampling" methods, the system is biased by the addition of an umbrella potential (in our case, springs anchoring our test molecules to places on a 3D grid). This bias drives the system to explore regions of higher free energy that would not otherwise be adequately sampled in a standard simulation (10).

For this work, we are interested in the translational degrees of freedom (x, y, z) of the center of mass of a probe in the ribosome tunnel. The rotational degrees of freedom of the molecule, the relative position of each probes atom, and all external degrees of freedom in the ribosome and solvent will be averaged in the PMF.

We collected the center of mass position for each probes from our MD simulations. We assumed noninteracting probes. Raw data were binned by using a 1-Å grid that covers the entire available volume.

We coded the WHAM equations to process the center of mass positions from each copy of a probe sampling the interior of the ribosome tunnel. The aim is to find

$$\tilde{p}_k^o, \qquad [1]$$

the unbiased probability to find the center of mass of the probe located in bin k. The WHAM equations used follow Karplus *et al.* (10):

$$\tilde{p}_k^o = \frac{\sum_j n_{j,kl}}{\sum_j N_j \tilde{f}_j c_{j,k}} \qquad [2]$$

$$\tilde{f}_j = \frac{1}{\sum_j N_j c_{j,k} \tilde{p}_k^o} \qquad [3]$$

where

$$\tilde{p}_k^o \qquad [4]$$

is the maximum-likelihood estimator for the unbiased probability of finding the probe center in bin k.

$$\tilde{f}_j \qquad [5]$$

is a normalization factor (weight) corresponding to simulation j. The factor

$$c_{j,k} = e^{-\frac{1}{kT} V_{j,k}} \qquad [6]$$

provides information of the potential acting on the center of mass of the probe occupying bin k, during simulation j. Eqs. **2** and **3** are solved iteratively to obtain the unbiased probability of finding the probe

$$\tilde{p}_k^o \qquad [7]$$

in every bin k.

The PMF at bin k is defined as:

$$PMF_k = -k_B T \log(\tilde{p}_k^o) \qquad [8]$$

Error analysis. The estimation of error is crucial when determining a free-energy map because we are trying to isolate signal from noise in statistics of small-molecule binding. To ensure data decorrelation, snapshots from MD simulations are taken at long intervals (1 ps). Neglecting force-field accuracy, we estimate errors that come exclusively from statistics and the histogram analysis. We encounter the following sources of error: (*i*) Discretization error in the biasing potential. The 3D space is discretized to a grid with a spacing of 1 Å. WHAM equations calculated on this grid are subject to error because the biasing potential over the grid is approximated to its value at the center of each bin. (*ii*) Convergence of the WHAM equations: Our criteria for convergence of Eqs. **2** and **3** terminates the iteration when the maximum change in any region of the PMF map (Eq. **8**) is $<0.02\ k_B T$. The number of cycles required for convergence depends on the specific system, but it ranges between 100–200 iterations. (*iii*) Statistical error: This source of error is related to regions in the free-energy map that are not adequately sampled. If a certain bin in the grid is not visited by a probe, it is ambiguous whether the reason is that there is no adequate sampling, or the free energy is very high. In our simulation, we are only going to consider data coming from bins within a radius of 5 Å from any anchor point. The histogram of distances (Fig. S8) shows that bins within 5 Å of the anchor are thoroughly visited, if the potential is given only by the surrounding bulk water. This is the largest source of error, and we estimate it in *Discussion*, below.

One of the advantages (13) of WHAM is that the determination of weights minimizes the statistical error, providing the maximum-likelihood estimate of the unbiased occupancy probability for each bin (10). However, given the raw data, there is uncertainty about the true value of the probability at each bin. This uncertainty not only depends on the number of observed counts for each bin but also depends on the counts for the nearby bins, propagated through the weighting factors f that weight adjacent histograms from different simulations (8). These factors become more efficient if the overlap between histograms is good (40). With 3-Å spacing between probe anchors and a sampled area of radii of 5 Å for each anchor, overlap between histograms is guaranteed.

Test on a toy system. We tested the performance of our algorithm on a toy system where the free energy profile can be analytically calculated: electrostatic potential due to a point charge. The system of choice is a cubic box with a positive charge fixed at a specific location inside its volume. Test probes were other positively charges anchored by springs and distributed throughout the volume and run in parallel in multiple simulations. The grid spacing between the probes, force-field and MD parameters, treatment of electrostatics, and other details of the calculation were the same as the system of study (ribosome tunnel). The main differences between the ribosome tunnel and the toy system resided in the fact that the toy system was smaller, it did not have water molecules, and the probes were monoatomic. The density of counts (raw data) renders a contour map as shown in Fig. S9 *Left*. Processing the raw data using our implementation of the WHAM algorithm, we obtain a free-energy contour map as shown in Fig. S9 *Right*. Fig. S9 agrees with the expected results within the boundaries of the box. However, Fig. S9 *Right* shows that there is high incidence of error in the free energy near the edges. For this reason, we adopt the criteria of restricting our calculation only to points within 5 Å of each anchor. Points within this boundary will be thoroughly sampled and, thus, bear lower uncertainty. Points outside this boundary are omitted from the PMF profiles. Fig. S10 compares the analytical coulombic potential and the free energy (PMF) obtained by using our algorithm, and we can see that the agreement is fairly good.

Discussion. ***Free-energy distribution histograms.*** For all PMF profiles, we obtained the distribution of free energies inside the tunnel relative to the exterior bulk. We observed marked differences between the profiles of the various side-chain analogs, which indicates that the tunnel is selective for different chemical species. Fig. S4 shows that the free-energy histogram for alanine has a prominent peak at $\approx 1\ k_B T$, suggesting that, inside the tunnel, an isolated alanine side chain may approximately experience an environment similar to bulk solvent. The charged probes Asp and Lys exhibit free-energy distributions peaking at

low free energies (<4 k_BT), consistent with the fact that the tunnel is mostly hydrophilic. However, the large width of these histograms reveals the existence of barriers and binding sites for charged probes. Hydrophobic probes (especially Trp) exhibit mean free energies higher than the charged probes; this is reasonable because a hydrophobic environment would otherwise cause aggregation of nascent chain.

Errors of the free-energy distribution. As to estimation of errors, because we have a large amount of sampling, we can afford a block analysis for the estimation of uncertainties, which is less computationally expensive than bootstrapping or Bayesian methods. We randomly divide the data into nine blocks. We use WHAM to obtain a probability profile for each block. Each of these profiles will have more uncertainty than the total raw data. The standard deviation of these nine profiles provides an error in each bin. Assuming a Gaussian distribution of values for the PMF, this error-analysis method easily provides an overestimate for our statistical error. The statistical error for the free-energy maps in Fig. 2 in the main text can be found in Fig. S13. Fig. S11 shows histograms for the uncertainty in the total free-energy calculation for the side chains studied.

Characterizing the RNA–protein interface: gate and latch model. Occupation of A497 in the tunnel exit. To locate the area occupied by the ribosome "latch" A497, we collected the data from our control simulation into a grid that shows the regions most visited by the tip (A497) of the flexible RNA helix 24. The data comprises a total of 5,028 snapshots from 144 different initial structures. Every initial structure has been replicated into 40 clones, each differing in its initial velocities. Snapshots were taken at a rate of 1 every 100 ps. At each snapshot, the position of each of the 30 atoms of A497 has been added to the nearest point on a grid of 1-Å spacing. This results in the specific bin being either empty or occupied. We count how many times a certain bin is occupied. Normalizing by the total amount of snapshots, we obtain the probability of a given grid point to be occupied by A497. The results are shown in Fig. S6.

B-factors. B-factors have been calculated from our trajectory by using the formula: $B = 8\pi^2\langle\bar{x}\rangle^2$. Comparison between crystallographic B-factors and B-factors calculated from our trajectory is shown in Fig. S3. As can be noted in this figure, the gate (tip of helix 24 in rRNA S23) and the latch arginines Arg-39 and Arg-35, have large B-factors ($B > 80Å^2$) in the x-ray structure, consistent with the observed large fluctuations for these residues in our simulation.

Histograms of distances between gate rRNA tetraloop and latch L39. To characterize the interactions between the rRNA tetraloop and L39, we calculated the minimum distance (the shortest distance between two groups of atoms) between A497 and the closest arginines in L39: Arg-20, Arg-31, and Arg-35. Fig. 5*B* in the main text shows a histogram of this distance over 5,028 snapshots taken at a 100-ps time step. It is interesting to note two peaks at 2.8 and 4.8 Å. These lengths correspond well with hydrogen-bond and water-mediated hydrogen-bond distances. The crystallographic minimum distance can be used as reference and is 9.04 Å, as compared with the most probable distance at 7.8 Å. In Fig. 5*B* in the main text, same histograms for G496 and A498 are plotted. These histograms show a less broad distribution, corresponding to less fluctuation, and no close-contact peaks as A497.

To further explore the interactions of A497 to each one of the arginines in L39, we calculated the minimum distance between A497 and each one of the neighboring arginines in L39. For each arginine, we obtain a normal distance distribution (Fig. S12). In Table S1, we list the mode and standard deviation of these distributions.

Even if the most frequent distances reflected by the statistical modes suggest long-range interactions for, at least, Arg-20, Arg-31, Arg-35, and Arg-39, the swaying motion of A497, represented by the large root mean square deviation, seems to include Arg-40, Arg-43, and Arg-44 in its list of visited side chains.

As to short-range interactions, the histograms of distances for Arg-20 and Arg-31 are strikingly similar and exhibit the two close-contact peaks at 2.4 and 4.8 Å mentioned before.

Sequence and structural alignment. We study the sequence and structural conservation of the agents that constitute the electrostatic interface at the exit of the tunnel: ribosomal protein L39 and the tip of rRNA helix 24.

Sequence alignment. Local sequence alignment of the tip of rRNA helix 24 was performed with CLUSTAL W (1.83) Multiple Sequence Alignment Algorithm for the four published x-ray ribosomal structures: *H. marismortui*, *E. coli*, *D. radiodurans*, and *T. thermophilus*. The result is shown in Table S2.

There are several stretches in the tip of helix 24 that are conserved in the four structures. By structural alignment of these conserved rRNA stretches to the *Haloarcula* structure, we obtained the superposed panels *B–D* shown in Fig. 6 of the main text. These alignments show that in the case of *E. coli* and *D. radiodurans*, L23 overlaps with L39 of *Haloarcula*. This is not the case for *T. thermophilus* ' L23, which could mean that the gate–latch mechanism we predict here might not apply to that ribosome.

As to sequence conservation in other species, we used the data in the European Ribosomal RNA database (14) to align the tip of helix 24 from all available species of Archea, Bacteria, and Eukaryotes as long as the name of the species was different. That means, for example, that sequences corresponding to *Bacillus anthracis* and *Bacillus cereus* were both included. In the cases where several sequences have been indexed for the same species, only the first sequence would be included in the multiple alignments. In Archea, A497 is conserved in 84% of a total of 21 sequences. In Bacteria, of 186 sequences, A497 is 32% conserved, 53% mutated to guanine, and never mutated to cytosine. In Eukaryotes, A497 is widely conserved (88% of all 127 sequences).

Secondary structure prediction. We used MFOLD (15) to predict the secondary structure of the tip of helix 24. Fig. S7 shows the predicted secondary structure motifs for this region. In *H. marismortui*, *D. radiodurans*., and *T. thermophilus*, the tip of helix 24 is predicted to be a tetraloop, whereas for *E. coli*, it is predicted as a pentaloop. Looking at the predictions for other species, we found that it can adopt either form or even a hexaloop as in *Homo sapiens* (Fig. S7*F*).

1. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* 289:905–920.
2. Sorin EJ, Pande VS (2005) Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys J* 88:2472–2493.
3. Jorgensen WLJ, Chandrasekhar J, Madura JD, Impey RW, Klein AL (1983) Comparison of simple potential function for simulating liquid water. *J Chem Phys* 79:926–935.
4. Gilbert RJ, *et al.* Three-dimensional structures of translating ribosomes by cryo-EM. *Mol Cell* 14:57–66.
5. Shirts M (2004) in *Chemistry* (Stanford Univ Press, Stanford, CA).
6. Lindahl E, Hess B, van der Spoel D (2001) GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J Mol Model* 7:306–317.
7. Berendsen H, Postma J, Vangunsteren W, Dinola A, Haak J (1984) Molecular-dynamics with coupling to an external bath. *J Chem Phys* 81:3684–3690.
8. Gallicchio E, Andrec M, Felts AK, Levy RM (2005) Temperature weighted histogram analysis method, replica exchange, and transition paths. *J Phys Chem B* 109:6722–6731.
9. Ferrenberg A, Swendsen R (1989) Optimized Monte Carlo data analysis. *Phys Rev Lett* 63:1195–1198.
10. Bartel C, Karplus M (1997) Multidimensional adaptive umbrella sampling: Applications to main chain and side chain peptide conformations. *J Comput Chem* 18:1450–1462.
11. Torrie GM, Valleau JP (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J Comput Phys* 23:187–199.

12. Kumar S, Bouzida D, Swendsen R, Kollman P, Rosenberg J (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 13:1011–1021.
13. Roux B (1995) The calculation of the potential of mean force using computer simulations. *Comput Phys Commun* 91:275–282.
14. European ribosomal RNA database (Antwerp, Belgium). http://bioinformatics.psb.ugent.be/webtools/rRNA/lsu/forms/index.html, accessed January 2008.
15. Zucker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31:3406–3415.

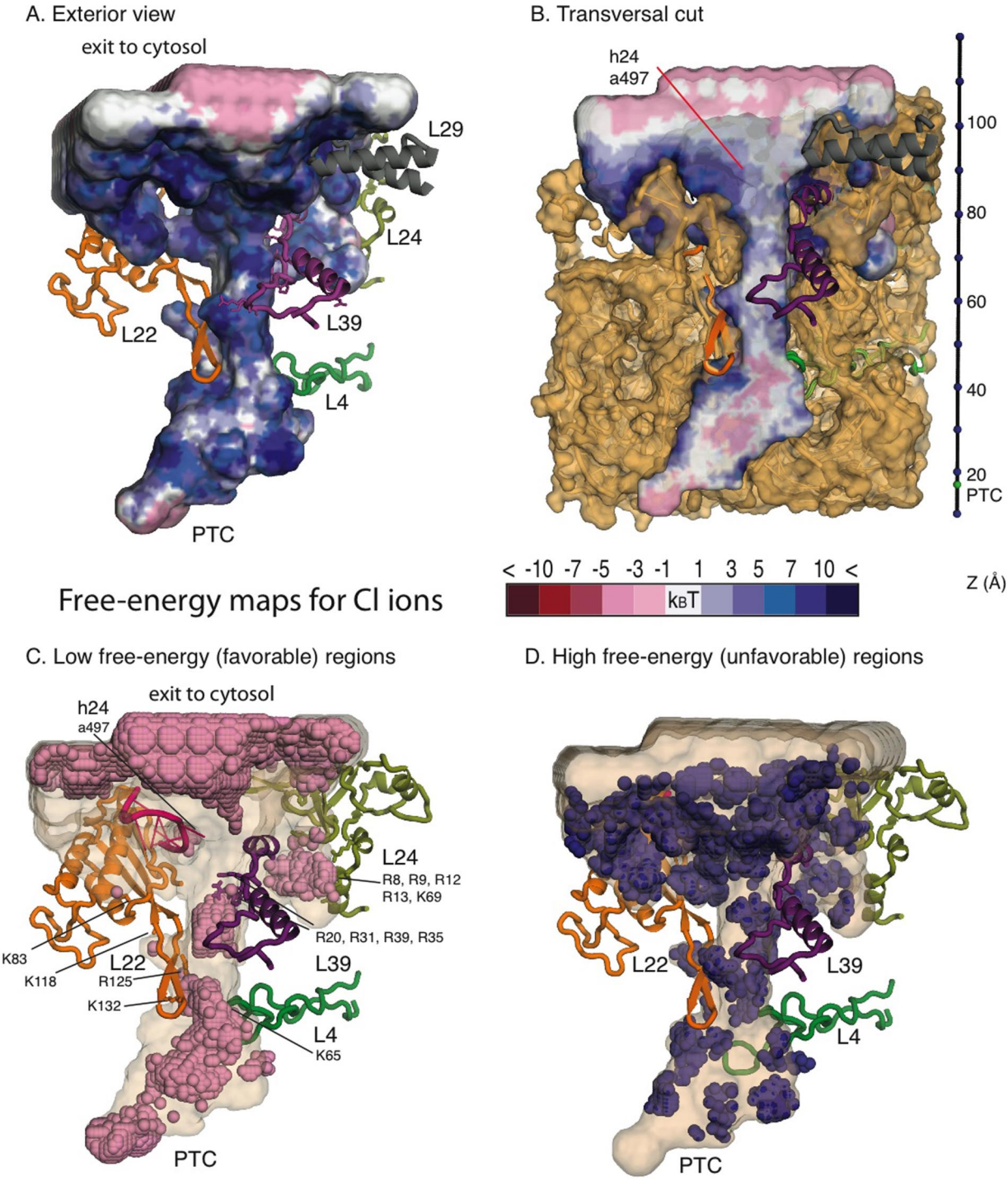


Fig. S1. Free-energy profiles of Cl^- ions in the ribosome tunnel. The color code shows the most-favorable regions in tones of red and the unfavorable regions in darker tones of blue. The reference is solvation energy of the ion given by the free energy in the simulated region farthest away from the ribosome cut-out. (*A*) Exterior view of the tunnel cavity with reference proteins L22 (orange), L23 (cyan), L4 (green), and L39 (magenta). (*B*) Transverse cut 3D free-energy map of Cl^- and the scale of our model showing the distance from the PTC site (green). (*C*) Favorable regions with free energy lower than $-1\ k_BT$. (*D*) Unfavorable regions (free energy $>7\ k_BT$).

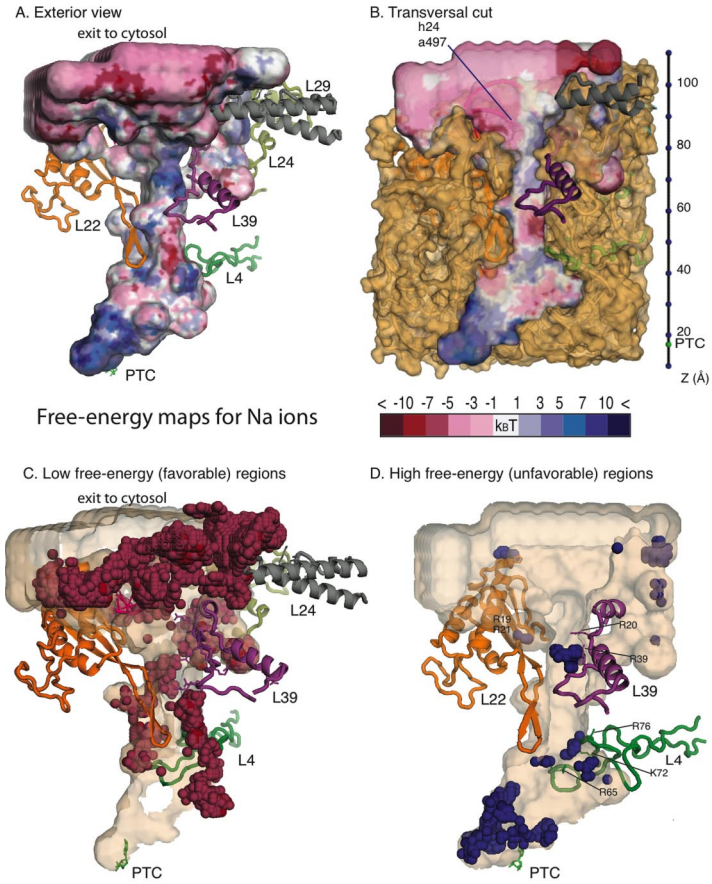


Fig. S2. Free-energy profiles of Na^+ ions in the ribosome tunnel. The color code shows the most-favorable regions in tones of red. The unfavorable regions are colored in darker tones of blue. The reference is solvation energy of the ion given by the free energy in the simulated region farthest away from the ribosome cut-out. (*A*) Exterior view of the tunnel cavity with reference proteins L22 (orange), L23 (cyan), L4 (green), and L39 (magenta). (*B*) Transverse cut 3D free-energy map of Na^+ and the scale of our model showing the distance from the PTC site (green). (*C*) Favorable regions with free energy lower than $-5\ k_BT$. (*D*) Unfavorable regions (free energy $>10\ k_BT$).

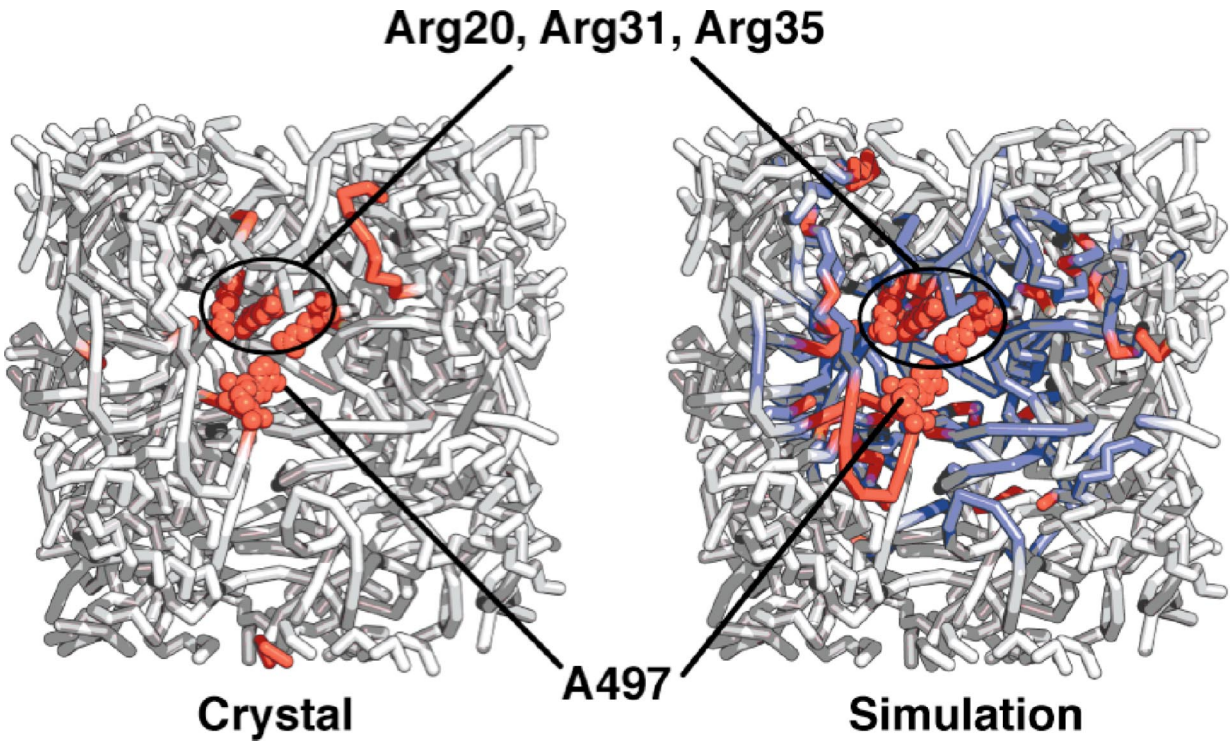


Fig. S3. Comparison of B-factors in the x-ray structure with B-factors from simulation. In red, B-factors $\geq 80 Å^2$. Modeled in spheres, Arg-20, Arg-31, Arg-35, and A497 from S23 rRNA helix 24. (*Left*) B-factors from x-ray structure (PDB ID code 1s72). (*Right*) B-factors from simulation. In blue are residues with rigid constrains in our simulation; in green are flexible residues.

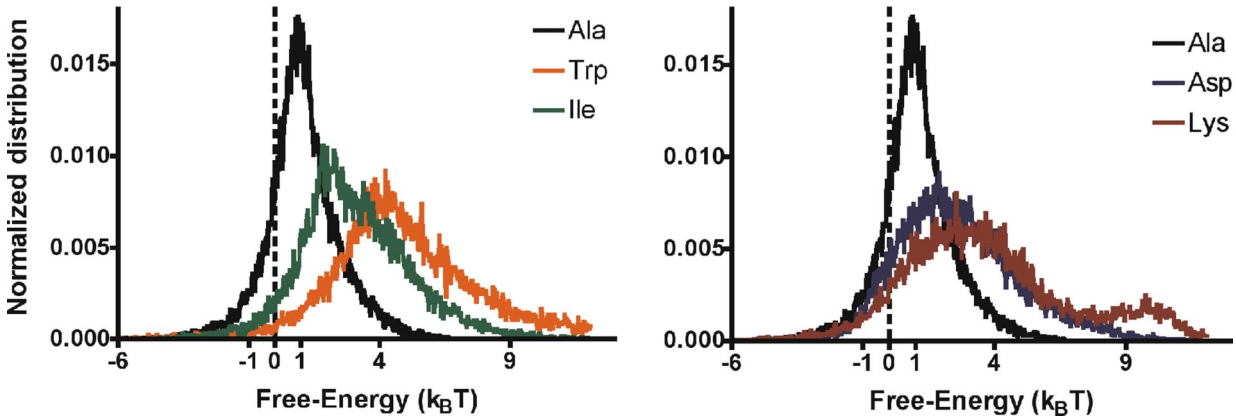


Fig. S4. Histogram of free-energy distributions inside the tunnel for the hydrophobic (*Left*: Ile, Trp) and charged (*Right*: Asp, Lys) probes as compared with Ala. All free energies are calculated with reference to the free energy of that probe in the bulk solvent region.

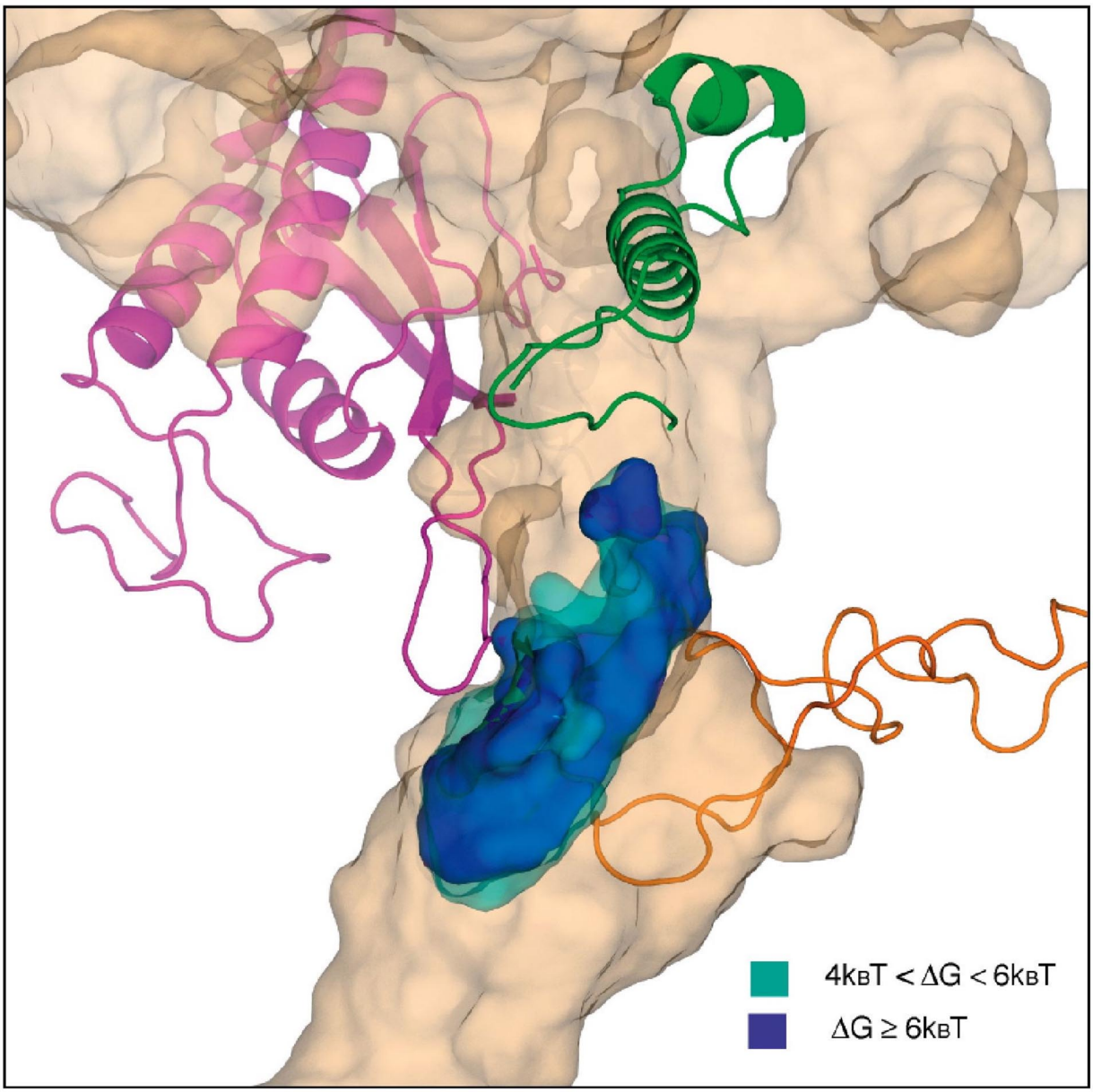


Fig. S5. Constriction site as seen in the solvent-accessible inner volume (sand color). Tryptophan free-energy profile with $\Delta G > 4\ k_BT$ at the constriction site between L22 and L4. Cyan represents free energy $>4\ k_BT$. Dark blue represents volume with free energy $>6\ k_BT$.

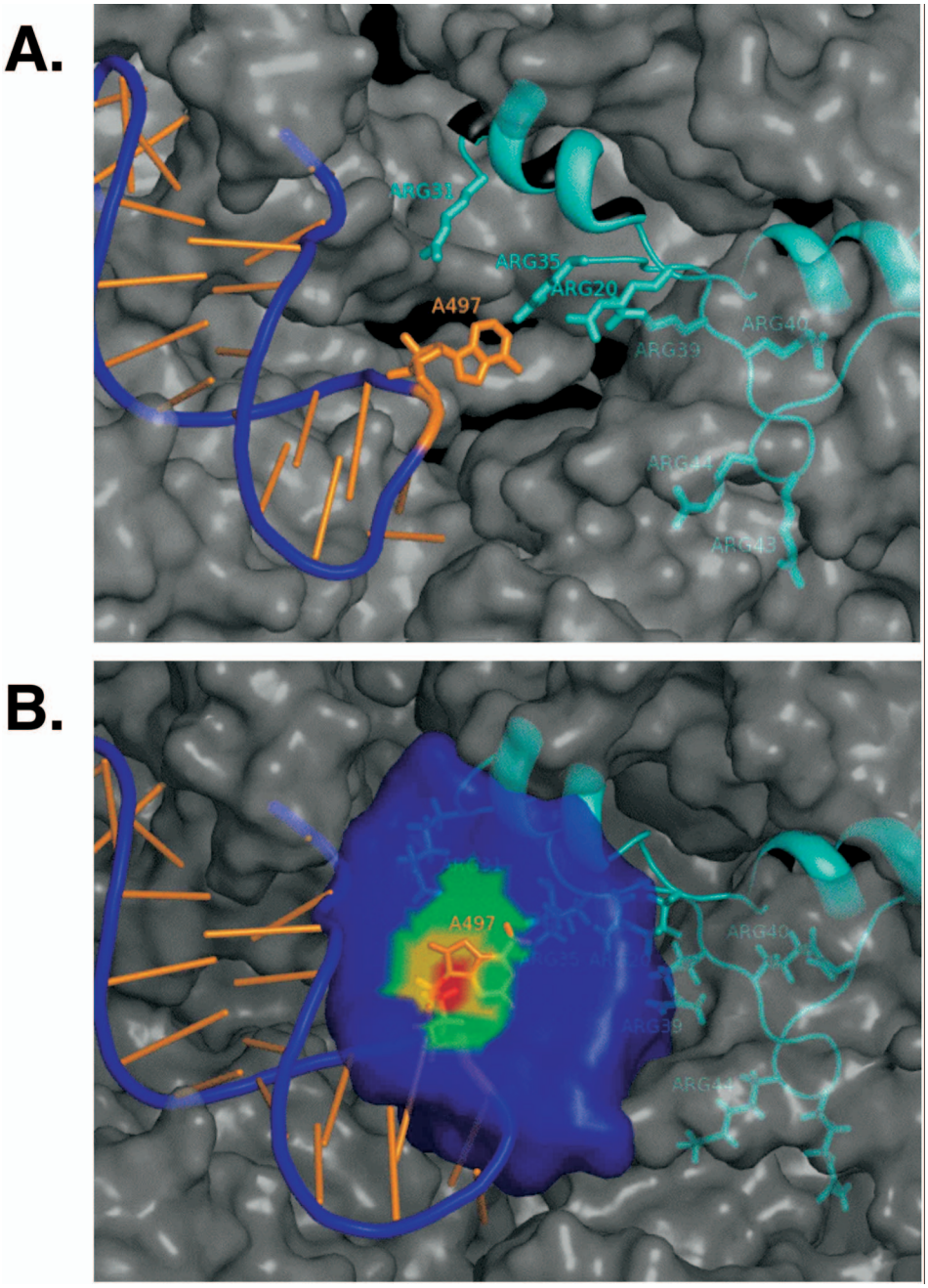


Fig. S6. View into the tunnel. (*A*) X-ray structure PDB I code 1S72. (*B*) Occupation plot for A497. Red, 10–15% occupied; yellow, 5–10%; green, 2–5%; blue, $<2\%$ but $>0\%$.

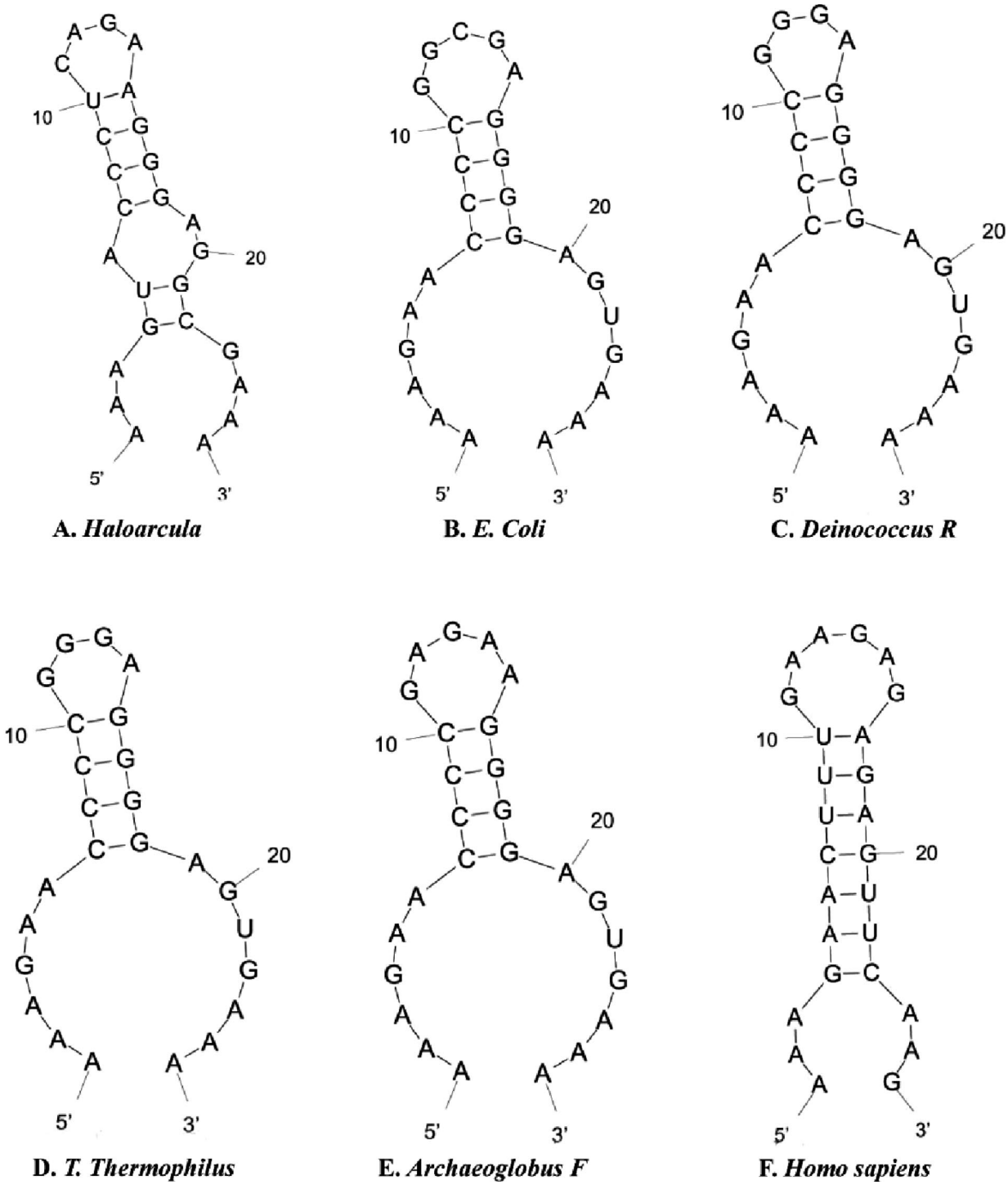


Fig. S7. Local secondary prediction of the tip of helix 24 for 23S rRNA from different species. Predictions and graphs were done with MFOLD (22). The tip of helix 24 can be a tetraloop, a pentaloop, or even a hexaloop.

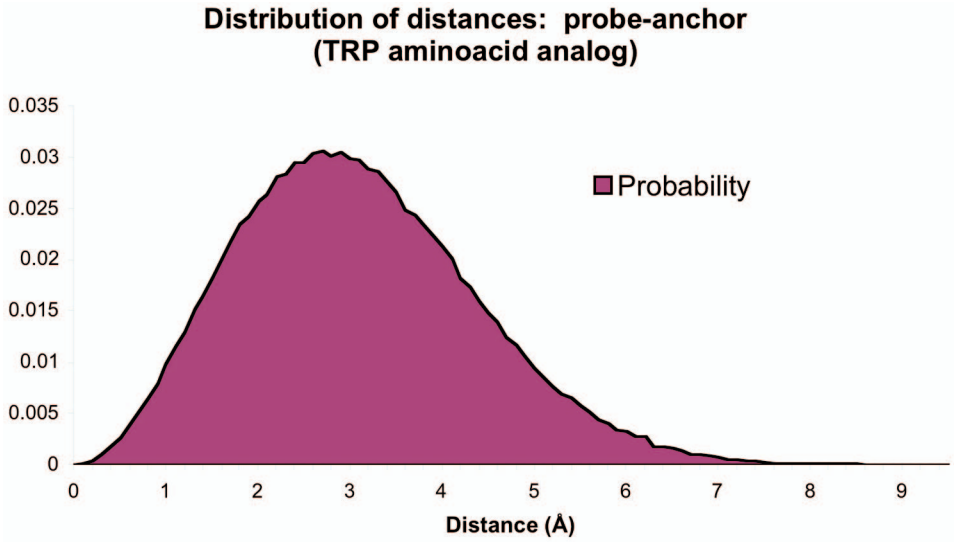


Fig. S8. Normalized distance distribution for the tryptophan probe in bulk solvent (237,388 snapshots).

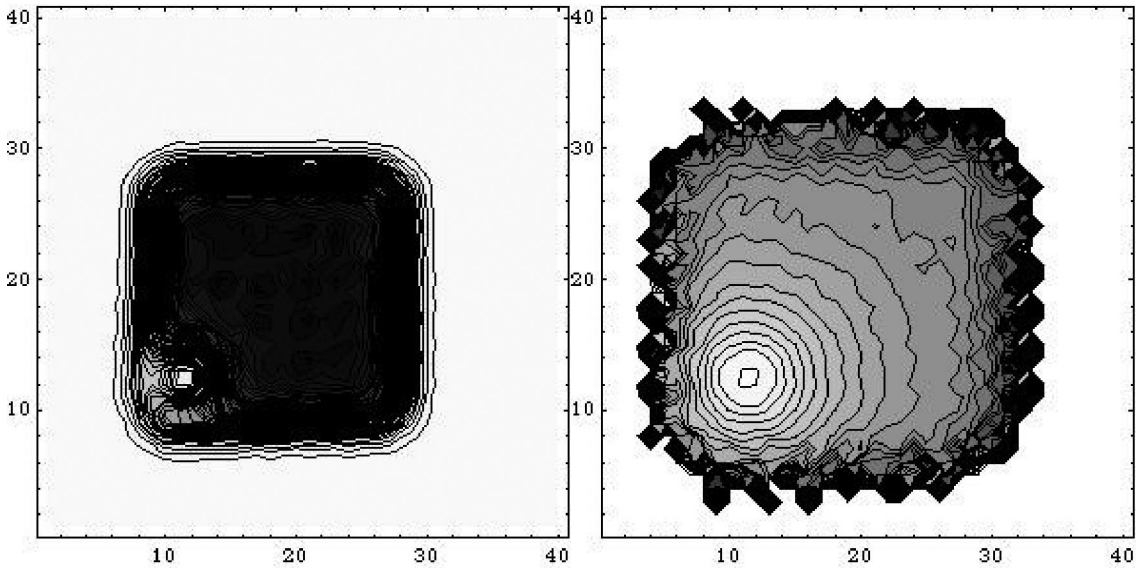


Fig. S9. Testing the 3D-WHAM algorithm. (*Left*) Contour map of raw data as collected from all test-probe simulations of a system with a fixed charge in a corner of a cubic box. The darker color implies a higher number of counts per bin. (*Right*) Free-energy contour for the same plane, as processed with our 3D-WHAM algorithm. The darker color implies lower free energy and, thus, higher probability to find the test particle.

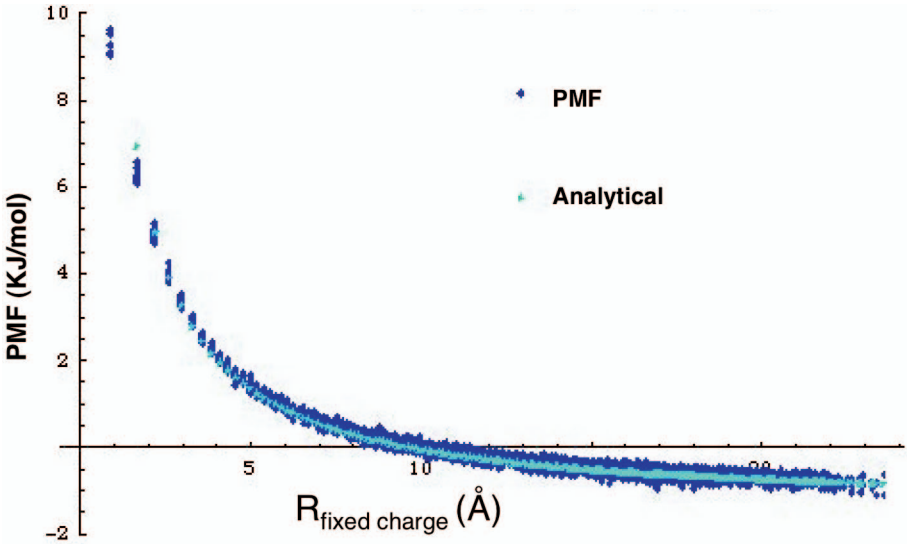


Fig. S10. Comparison between PMF obtained with 3D-WHAM and the analytical value, calculated as the coulomb potential between two repulsive charges.

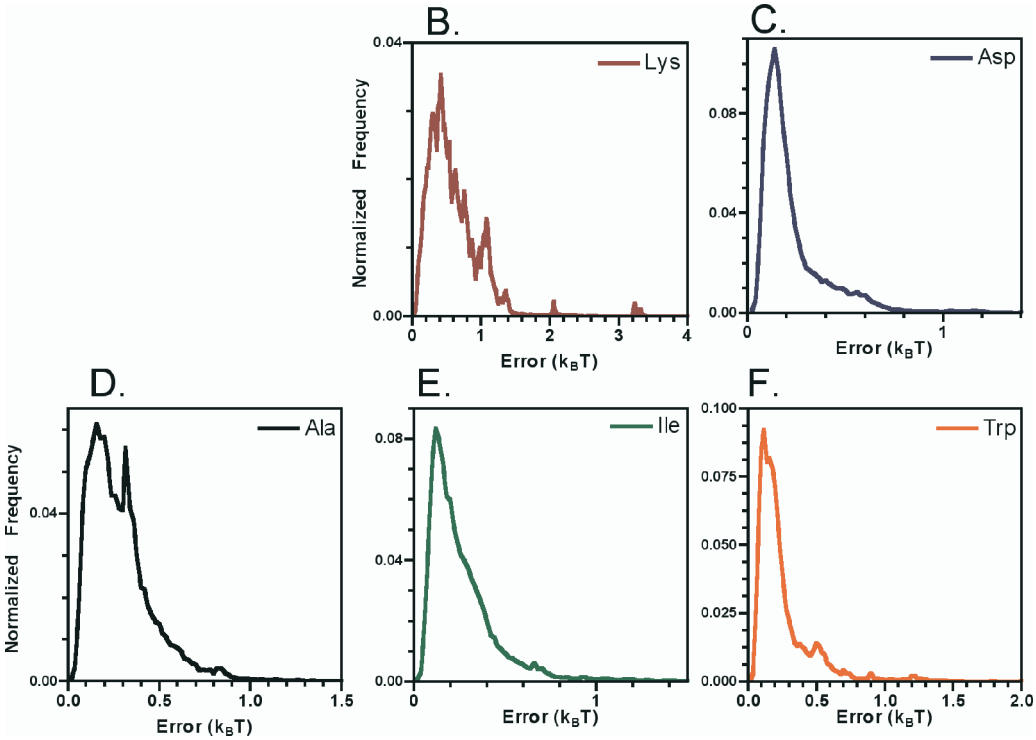


Fig. S11. Histograms for the uncertainty in the free energy corresponding to profiles *B–F* of Fig. 2 in the main text. These histograms show that with the exception of lysine (*B*), most of the uncertainty lies comfortably within 0.5 k_BT.

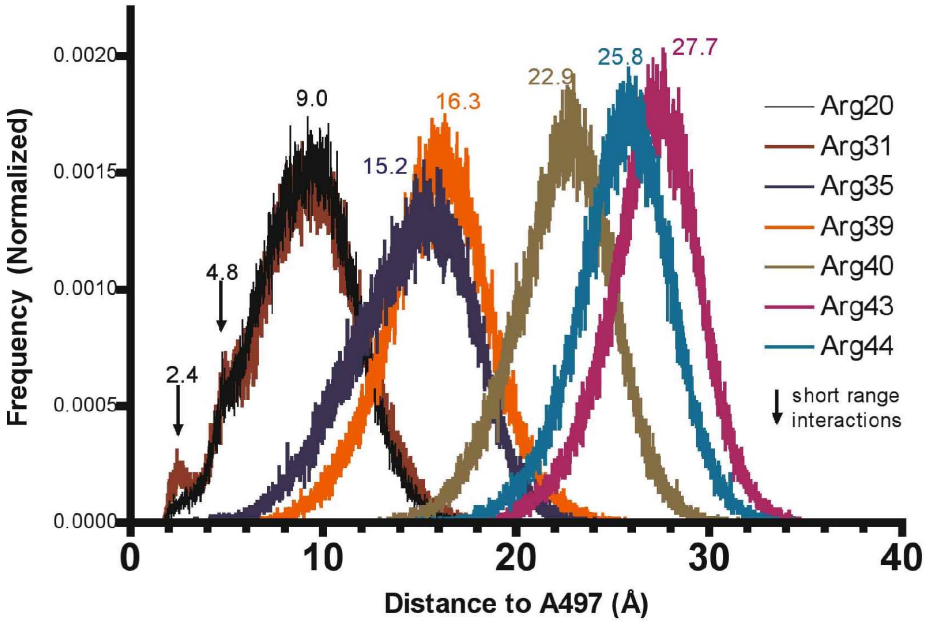


Fig. S12. Minimum-distance histograms between A497 and arginines in L39. Graphs for Arg-20 and Arg-31 are overlapping, indicating that A497 is equidistant to these two arginines. Short-range interactions (<6 Å) are indicated with arrows in the histograms of Arg-20 and Arg-31.

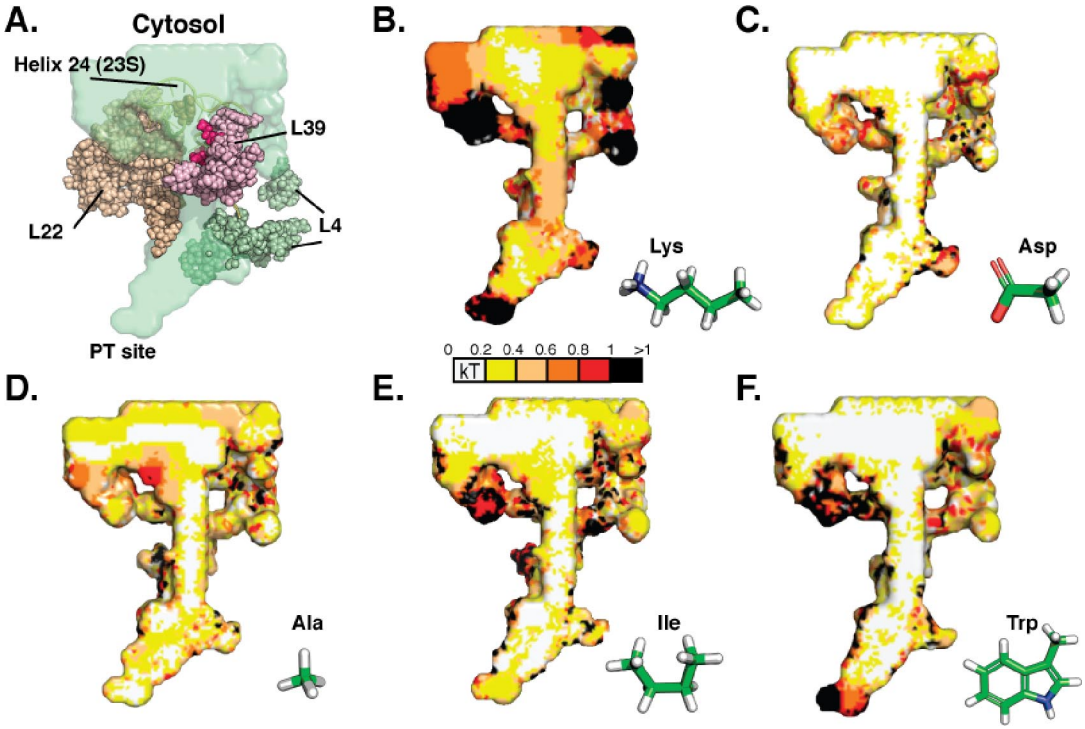


Fig. S13. Uncertainty in the corresponding profiles *B–F* of Fig. 2 in the main text.

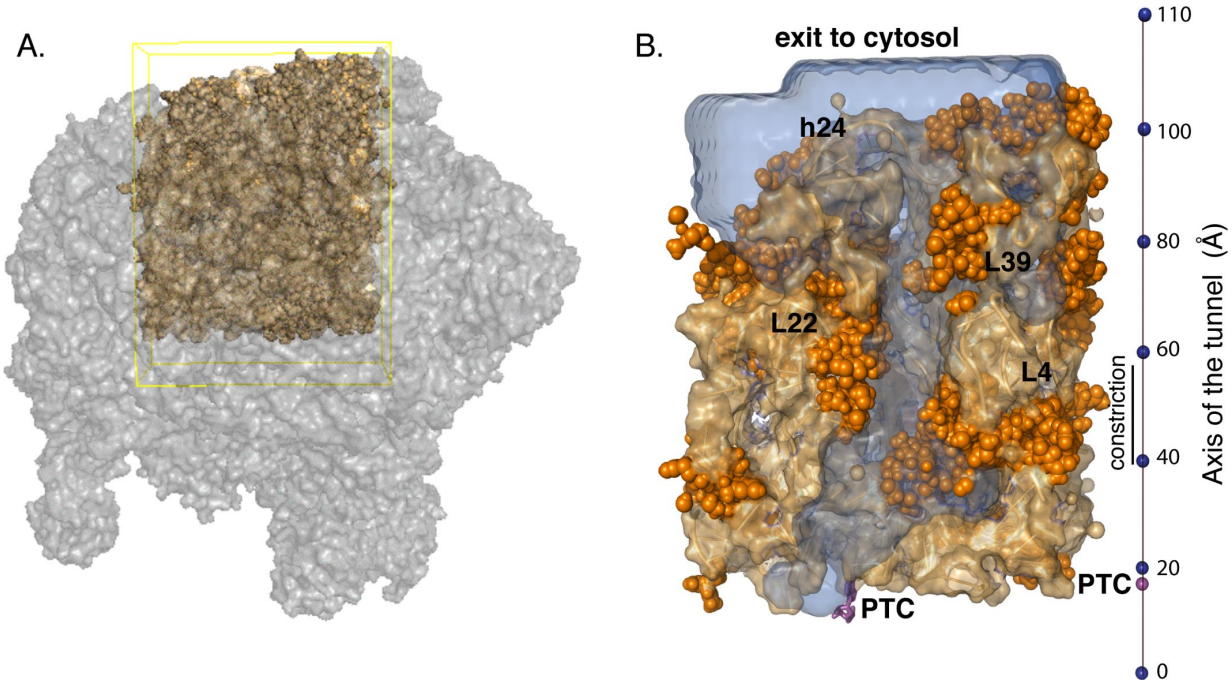


Fig. S14. Model of the ribosome tunnel. (*Left*) In gray, the total ribosome x-ray structure (PDB ID code 1s72) of the large subunit; in brown, our ribosome exit tunnel cut-out model. (*Right*) Ribosome cut-out model in orange. In blue, the ribosome tunnel solvent-accessible volume is shown.

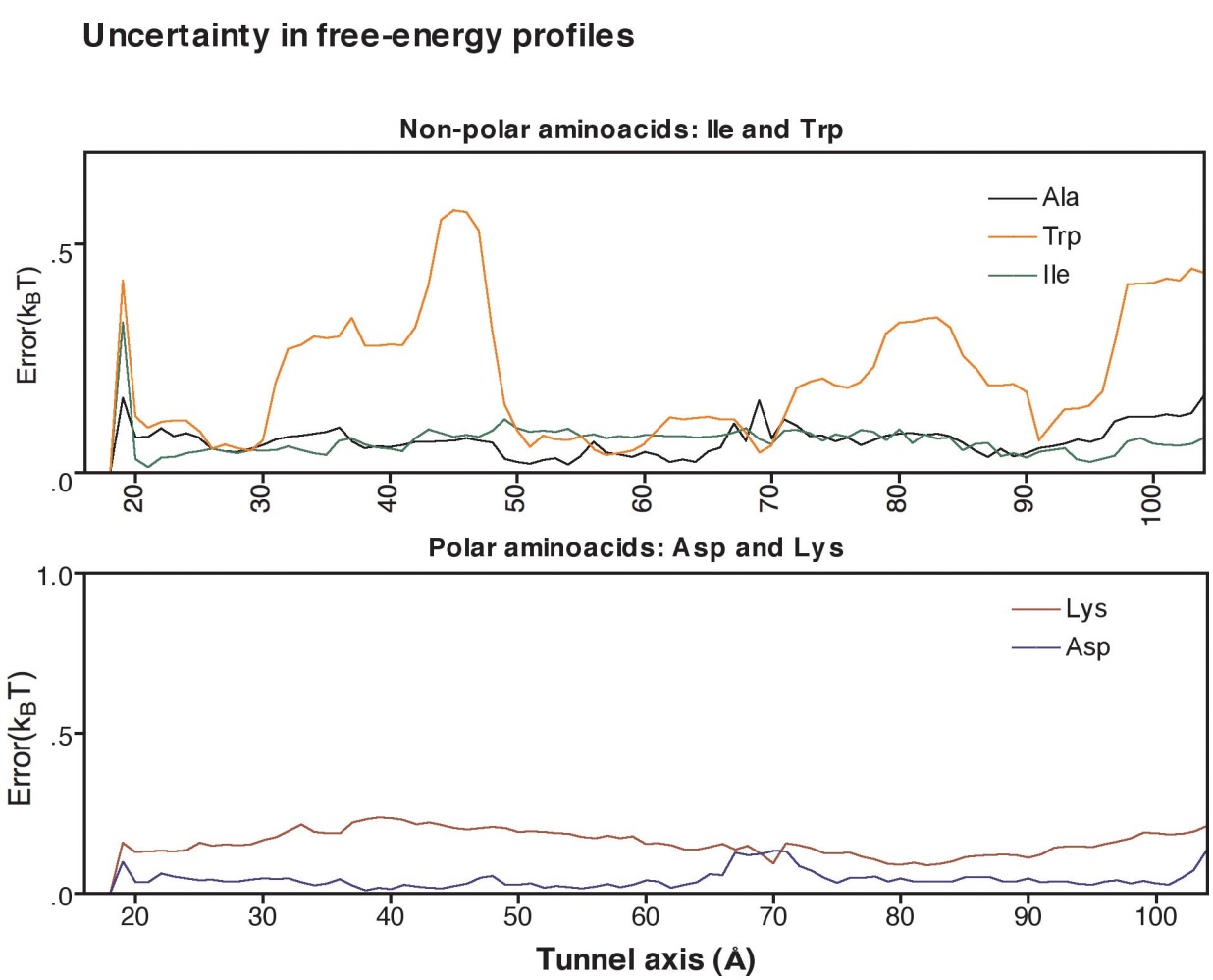
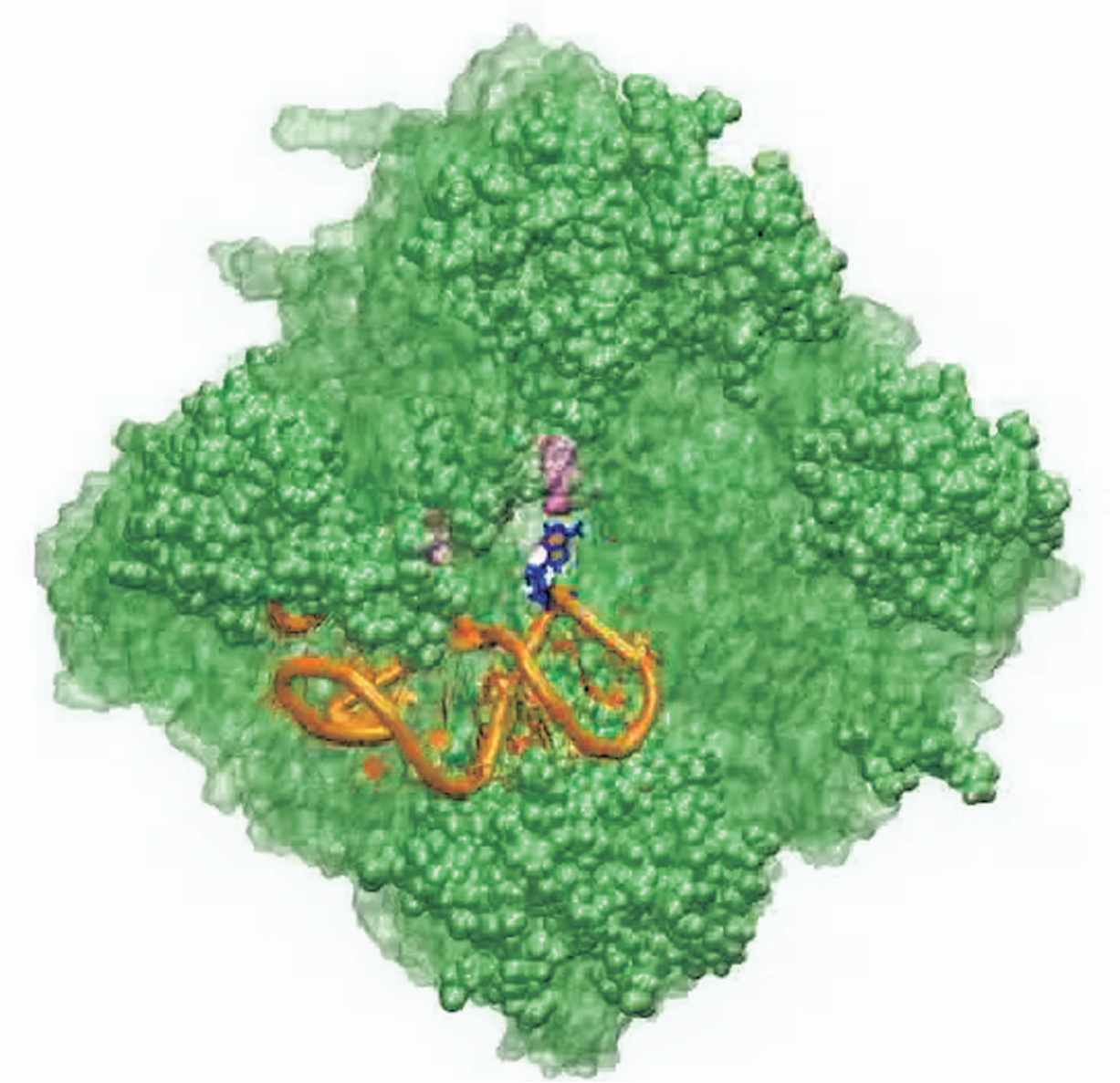
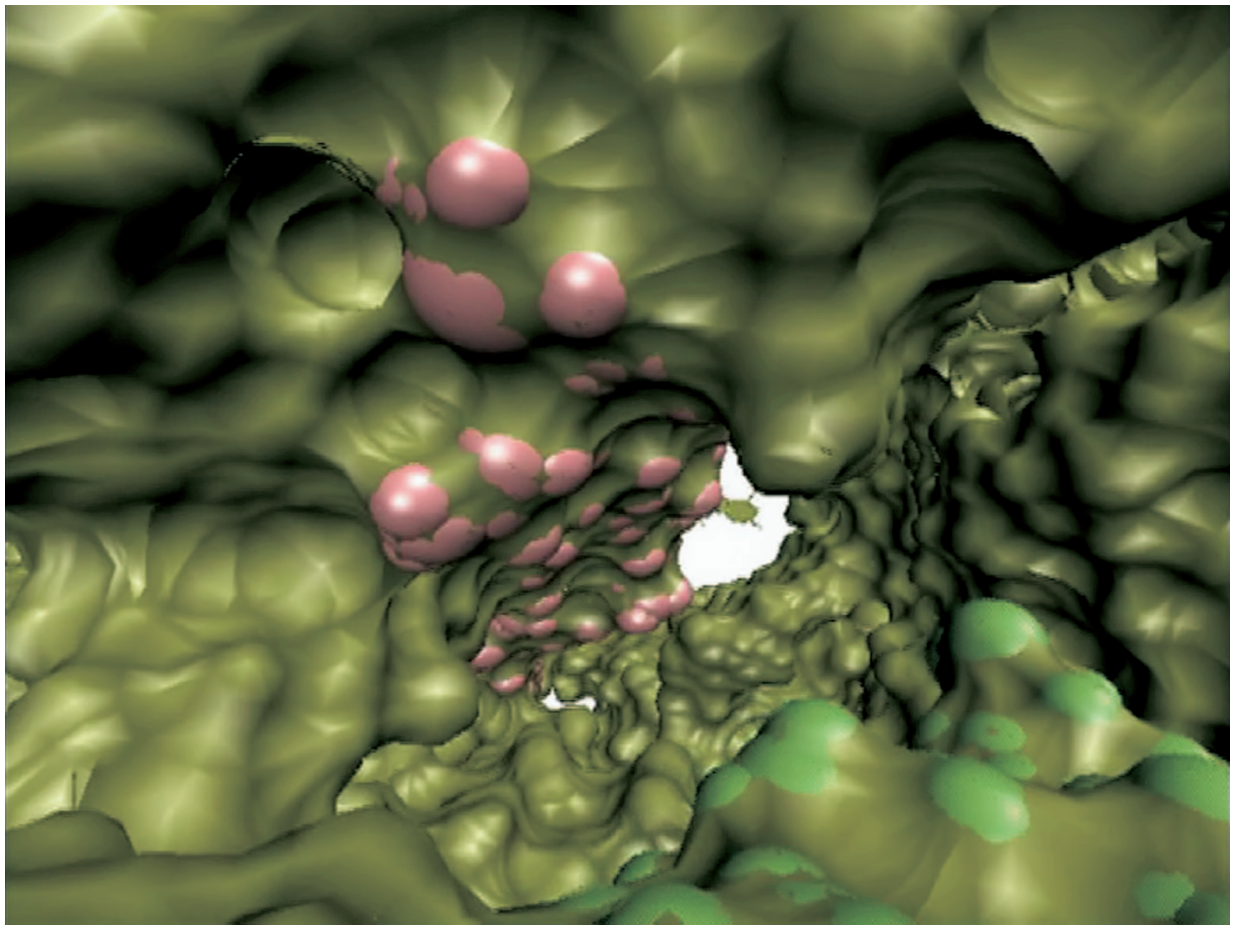


Fig. S15. Uncertainty to the free-energy barriers along the longitudinal axis of the ribosomal tunnel corresponding to Fig. 3 in the main text. The scale corresponds to the structural elements as shown in Fig. S14. Uncertainty calculated as explained in *Supplementary data and error analysis* in *SI Text* above.



Movie S1. Exterior view of the ribosome tunnel and conformational change of the tip of 23S rRNA helix 24 (orange tube). Blue: unpaired base A497 (gate). Purple: arginines in L39 (latch). The MD snapshots are taken at 500-ps intervals, collating trajectories of average length 20 ns.

Movie S1 (MOV)



Movie S2. Tour inside the ribosome tunnel. Proteins L4 (green) and L22 (pink) form the constriction site. L39 is shown in cyan, with its arginines in spheres. Flexible tip of rRNA helix 24 is located at the end of the tunnel. The motion was extracted from the control MD simulation. Tunnel walls are shown rigid for simplicity but are flexible in the original trajectory.

Movie S2 (MOV)

Table S1. Minimum distance between rRNA base A497 and different arginine residues in L39 as given by the distance distributions obtained from MD simulations

	Arg20	Arg31	Arg35	Arg39	Arg40	Arg43	Arg44
Mode, Å	10.4	10.5	15.8	14.8	24.3	26.8	25.6
σ, Å	9.4	9.5	14.9	16.1	22.6	27.2	25.7

Mode and standard deviation of distributions plotted in Fig. S12.

Table S2. Local multiple alignment of the sequence corresponding to the tip of helix 24 for *Haloarcula m.*, *Thermus t.*, *Deinococcus r.*, and *E. coli*

Organism	Sequence
E. coli	GAAAGGCGAAAAGAACCCCGGCGAGGGGAGUGAAAAAGAA–
Deinococcus	GAAAGGUGAAAAGAACCC–CGGGAGGGGAGUGAAAGAGAA–
T. thermophilus	GAAAGGUGAAAAGAACCCCGG–GAGGGGAGUGAAAUAGAGC–
Haloarcula	GAACGCUGCAAAGUACCCUCAGAAGGGAGGCGAAAUAGAGCA

Alignment was performed with CLUSTAL W (1.83) using the implementation at the Mobyle portal (http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.pg) of Institut Pasteur.