The Geometry of the Ribosomal Polypeptide Exit Tunnel

N.R. Voss[1], M. Gerstein[1], T. A. Steitz[1,2,3], and P.B. Moore[2,1]*

[1] Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520-8114.
[2] Department of Chemistry, Yale University, New Haven, CT 06520-8107.
[3] Howard Hughes Medical Institute, New Haven, CT 06520-8114.
* To whom correspondence should be addressed:
Department of Chemistry
Yale University
P.O. Box 208107 (mail) or 350 Edwards St. (shipping)
New Haven, CT 06520-8107

Tel: 203 432-3997
FAX: 203 432-5781
email: peter.moore@yale.edu

ABSTRACT

The geometry of the polypeptide exit tunnel has been determined using the crystal structure of the large ribosomal subunit from *Haloarcula marismortui*. The tunnel is a component of a much larger, interconnected system of channels accessible to solvent that permeates the subunit and is connected to the exterior at many points. Since water and other small molecules can diffuse into and out of the tunnel along many different trajectories. the large subunit cannot be part of the seal that keeps ions from passing through the ribosome-translocon complex. The structure referred to as the tunnel is the only passage in the solvent channel system that is both large enough to accommodate nascent peptides, and that traverses the particle. For objects of that size, it is effectively an unbranched tube connecting the peptidyl transferase center of the large subunit and the site where nascent peptides emerge. At no point is the tunnel big enough to accommodate folded polypeptides larger than α-helices.

INTRODUCTION

The peptidyl transferase center (PTC) is the site where peptide bond formation occurs in the ribosome. It is part of the ribosome's large subunit, and is located in the middle of the face the particle that interacts with the small ribosomal subunit.[1] In 1982, it was reported that nascent proteins first become accessible to antibodies on the side of the large subunit opposite its subunit interface.[2] This surprising observation suggested the large subunit might contain an internal tunnel large enough to accommodate nascent

peptides that connects the PTC to the site where polypeptides emerge, ~100 Å away. The existence of such a tunnel, the *exit tunnel*, was definitively proven by cryo-electron microscopy in 1995.[1] It is an obvious feature in all the high resolution crystal structures of 70S ribosomes and large ribosomal subunits published to date.[3-7]

The physiological properties of the tunnel are not well understood. On the one hand it is clear that nascent proteins having specific sequences can bind so tightly to the wall of the tunnel that protein synthesis is inhibited.[8-11] Furthermore, macrolide antibiotics, which bind to a specific site on the tunnel wall,[12] appear to inhibit protein synthesis by blocking the passage of nascent polypeptides down its lumen. However, it is not known whether interactions between nascent polypeptides and the tunnel wall play a more positive role in protein synthesis, *e.g.* by helping determine which nascent proteins will be secreted.[13] For example, it has been suggested that L22 regulates protein synthesis this way because its globular domain forms part of the rear surface of the subunit where the translocon binds, and its β-hairpin is a component of tunnel wall over almost two-thirds of its length.[14-16] This proposal has been called into question by the finding that cells containing ribosomes from which the loop sequence of L22 has been deleted are viable.[17]

Several proposals have been made about tunnel functions that have geometric implications. In some electron microscopic reconstructions of the ribosome the tunnel appears to branch nears its exit end.[18] This observation has led to the suggestion that nascent peptides might leave the ribosome by two different routes, one used by membrane proteins and the other by cytoplasmic proteins. It has also been suggested that nascent polypeptides fold at the tertiary level while traversing the tunnel,[19] a proposal that could relate to data indicating that 23S rRNA has chaperone activity.[20] Finally, the

ribosome-translocon complex promotes the passage of nascent proteins across membranes that have electrochemical potential gradients across them, but protein translocation is not accompanied by passive ion flow. It has been suggested that the ribosome itself is part of the "seal" that makes the ribosome-translocon complex ion-tight.[21]

Here we present a geometric analysis of the tunnel in the large ribosomal subunit of *Haloarcula marismortui* that was undertaken to illuminate its role in protein synthesis. This analysis does not support any of the hypotheses mentioned above. For objects the size of nascent polypeptides the tunnel is not branched. It is too small to allow nascent proteins to fold, and the ribosome cannot contribute to creating a seal that prevents the passage of ions through the ribosome-translocon complex. The solvent volumes inside the large subunit have been examined using the same methods. The interior of the particle is extremely wet; solvent occupies 39% of its volume. Furthermore, the solvent volume inside the ribosomal particle is a system of connected channels that permeate the entire particle. A water molecule at almost any position inside this system can diffuse to any other position without leaving the ribosome's interior.

RESULTS

Solvent-filled voids are abundant in the large ribosomal subunit. Contrary to the impression conveyed by many of the published images of the large ribosomal subunit, the exit tunnel is not the only volume inside that structure large enough to accommodate solvent molecules. The ribosome is full of such volumes. Interestingly, when the

subunit's interior is viewed slice by slice, as in Figure 1A, virtually every volume seen in every slice that is big enough to contain solvent molecules has at least one ordered solvent molecule in it that is visible crystallographically (Figure 1B). The exit tunnel, which is discernible in the right-hand image of Figure 1A, is distinguished from the rest of the solvent-containing volumes in the subunit mainly by its relative straightness and uniformity of width. It is so straight that you can see all the way through the middle of appropriately oriented ball and stick models of the large ribosomal subunit. The other solvent-containing voids in the subunit are less obvious because they are more irregular in shape and size.

A two-step approach for determining the surface of the tunnel. Many computer programs have been written to assist in the identification and analysis of the voids in the interiors of macromolecules, especially proteins.[22-28] Several of these programs were applied to the large ribosomal subunit but for a variety of reasons were not found useful (see Methods for further discussion). Instead, the rolling ball algorithm of Richards was used for the analysis that follows.[29] This algorithm identifies the solvent-accessible and solvent-contact surfaces of macromolecules using only atomic radii and atomic coordinates as the input information (see Figure 2).

It is not a straightforward task to define the surface of the tunnel using the rolling ball algorithm. The rolling ball algorithm is designed for determining the surfaces of macromolecules, which are fundamentally convex objects. The tunnel is fundamentally concave, being topologically part of the surroundings of the large ribosomal subunit, and thus not a discrete, self-delimited object. Thus the surface of the tunnel had to be worked

out using a two-step process. In the first step, the entire ribosome was surrounded by a limiting surface that made the tunnel finite by capping its ends. In the second step, the rolling ball method was used to define the surface of the cavity that corresponds to the tunnel in that artificially delimited object (see below).

Delimiting the large ribosomal subunit. The tunnel was made discrete by surrounding the entire large ribosomal subunit with a surface we call its *shell*. The essential properties of the shell of the large ribosomal subunit are: (1) that it enclose all the volumes inside the invaginations and other concavities in the surface of the large ribosomal subunit, (2) that the volume of exterior solvent captured within it be otherwise minimal, and (3) that its surface be connected. Every point on a connected surface can be reached from every other point by a trajectory lying entirely on that surface.

Operationally the shell of the large ribosomal subunit was taken to be the contact surface of the particle determined using a spherical probe having a radius of 10 Å. Figure 2 shows what is meant by the term *contact surface* in this context.[29] While it is clear that surfaces surrounding the large subunit can be obtained this way, the use of a 10 Å as the radius of the probe for determining shells requires justification.

The lower bound on the radius of the sphere used to define the shell of the large ribosomal subunit was set by the requirement that the shell be a connected surface. The interiors of space-filling atomic models of macromolecules are riddled with "empty" regions, some big enough to contain solvent molecules, but others so small they cannot. Some of these interior voids may be cavities, which is to say voids big enough to accommodate a sphere that is larger than the largest sphere capable of entering them from

the exterior. A contact surface obtained for a macromolecule using a spherical probe that contains cavities big enough to accommodate the probe will consist of an outer surface that surrounds smaller closed surfaces to which it is not connected. Thus if the rolling sphere contact surface of a molecule is to be a satisfactory shell, *i.e*. to define a connected surface, the radius of the sphere used to define it must be so large that there are no cavities in the macromolecule big enough to accommodate it.

Not surprisingly, the number of volumes inside a macromolecule identified as cavities, *i.e.* volumes big enough to accommodate a sphere of the given radius, depends on probe radius. Also, there is a critical radius beyond which there are no cavities at all. However, as Figure 3 shows, the large ribosomal subunit contains no cavities big enough to enclose a sphere having a radius greater than 10 Å. Therefore, the contact surface defined by a 10 Å sphere is connected, as required for it to be a suitable shell surface.

Granted that 10 Å is the radius of the smallest sphere that can be used to generate a satisfactory shell for the large ribosomal subunit, why would a surface defined using a larger sphere not be just as satisfactory? The reason emerges when the contact surface of the large ribosomal subunit determined using a 10 Å probe is compared to the contact surface of the same object obtained using a 100 Å radius probe (Figure 4). The large ribosomal subunit is a compact, more or less isometric mass from which three large projections protrude. In addition to having small scale (less than 10 Å) surface concavities due to its molecular granularity, its surface has much larger scale (~100 Å) concavities caused by these projections. Most of the small scale surface concavities of the large ribosomal subunit are not represented in the contact surface obtained using a 10 Å radius sphere, but its larger scale concavities remain conspicuous. The large scale

concavities of the particle are less obvious in contact surface obtained using a 100 Å radius sphere because they are partially smoothed, but for the same reason, the surface that emerges encloses large amounts of bulk solvent. Thus to keep the amount of extra solvent encompassed by the shell to a minimum, the shell of a macromolecule should be defined using the smallest spherical probe that yields a connected surface.

The size and shape of the tunnel depends on probe size. Once the large ribosomal subunit had been enclosed in a satisfactory shell, the surface of its tunnel could be defined by placing a sphere of some radius inside its lumen *in silico* and allowing that sphere to roll around inside it. The contact surfaces that emerge are determined in part by the atoms forming the wall of the tunnel, and in part by the surrounding shell.

What should the radius of the sphere be that is used to define the tunnel wall? This is not a trivial question because the size and shape of the surface defined for the tunnel this way is extremely sensitive to probe radius. Figure 5A (green line) shows how the volume of the tunnel accessible to spherical probes depends on probe radius. For comparison, the dependence of the total accessible volume inside the shell of the ribosome on the radius of the probe sphere is shown (Figure 5A, red line). For probes having radii up to 2.5 Å or so the volume accessible from the lumen of the tunnel is almost as large as the entire volume inside the shell accessible to a sphere of that radius. However, as the probe radius increases from 2.5 Å to 3.0 Å, the volume assigned to the tunnel plummets both in absolute terms and as a fraction of the total accessible volume inside the shell. This dramatic fall in volume is associated with an equally dramatic change in tunnel morphology (Figure 5B). Interestingly, when the tunnel surface is determined using a 2.7

Å radius probe, the object so defined is bifurcated at its distal end. However, a slight increase in probe radius eliminates the bifurcation without otherwise affecting the shape of the tunnel surface. When the probe used is 3.0 Å that the tunnel surface that emerges corresponds well to what the casual observer of the structure would consider as constituting the *exit tunnel* (see Figure 1A, right).

It follows that the size and shape of the tunnel cannot be uniquely defined, because its effective size and shape is determined by the shapes and sizes of the molecules that must pass through it. An extended poly-alanine chain has lateral dimensions of roughly 6 Å (see Methods), and nascent polypeptides with bulkier side chains are larger than that. Hence for the purposes of analyzing the passage of polypeptides down the tunnel, it is appropriate to think of the tunnel surface as being the surface defined using a sphere having a radius of 3.0 Å. A stereo view of that tunnel is provided in Figure 6 with some its biochemical landmarks indicated. The volume inside the tunnel surface is about 25,000 $Å^3$, and the distance from the PTC to the distal end of the tunnel is about 80 Å (depending on where one decides its distal end is). Thus if the tunnel were a cylinder of uniform diameter, which it is not, its diameter would be about 20 Å. The wall of the tunnel is RNA-rich; 82% of the atoms that contact its surface are RNA atoms and only 18% of them are protein atoms.

The tunnel is permeable to water. Having defined the tunnel lumen using a 3.0 Å probe (Figure 5B), it is instructive to find out what fraction of the total solvent volume inside the shell of the subunit can be accessed by a water molecule placed in the tunnel lumen. As is evident in Table 1 (and Figure 5A, gray vertical line), the volume accessible

to a water molecule is 93% of the solvent-accessible volume inside the shell of the large ribosomal subunit. Thus effectively, there is no tunnel for water-sized objects. Further, cavities big enough to contain solvent molecules account for only 0.02% of the solvent-containing volume that cannot be accessed from the channel. The rest of the solvent volume that cannot be reached from the tunnel is solvent "trapped" in concavities at the surface of the particle by the shell.

The tunnel so defined comprises a huge network of interconnected channels that permeates the entire structure of the large ribosomal subunit. Since this network reaches the particle's exterior in many places, molecules the size of water molecules should be able to diffuse into and out of the tunnel without difficulty. The tunnel that emerges when a 3.0 Å radius sphere is used, on the other hand, is much smaller. Its volume is about 4.4% of the total solvent volume of the particle (Figure 5A and Table 1).

The interior of the ribosome is very wet. The fraction of the volume inside the shell of the large ribosomal subunit that is solvent-filled is large; just over 39% (see Table 1). To understand what this number means it is useful to compare the ribosome to another macromolecular assembly that has an interior cavity: spinach Rubisco (PDB id, 1RCX).[30] When its structure is analyzed the same way, 17.2% of the volume inside its shell is found to be accessible to solvent, but its interior cavity has no side branches accessible to water-sized probes (Figure 1C, left). This is consistent with the long-standing observation that the interiors of protein domains contain little or no solvent.[31] Further, 56.1% the solvent inside the Rubisco shell is to be found in the thin water layer between its exterior surface and its shell, which is evident upon examination of Figure 1C (right). Thus water

that is truly in the center of the structure accounts for only 7.6% of the volume inside its shell. For the large subunit, only 12.4% of the solvent is found in its outer-most layer and hence interior water still accounts for about 34% of the volume inside the shell. Thus the porous solvent structure that permeates the entire ribosome is not protein-like.

DISCUSSION

Is the exit tunnel branched? In some electron microscopic reconstructions of the ribosome, the tunnel branches at its distal end, and its branches appear to be big enough to accommodate nascent peptides.[18] However, our calculations show that for objects the size of polypeptides the tunnel is not branched. How can this discrepancy be explained? We believe it likely that the branching seen in low resolution electron microscopic images is a Fourier series termination artifact similar to the one that causes α-helices to resemble slender rods in X-ray crystallographic electron density maps calculated at resolutions around 6 Å. The earlier 9 Å resolution crystallographic electron density map published for the large ribosomal subunit looks far more porous than it really is for this same reason.[7] In this connection we note that when low resolution EM images are superimposed on the high resolution X-ray crystal structures, it is found that ribosomal protein occupies the space assigned to the proposed branch (Daniel Klein, personal communication).[1;18]

To what extent do proteins fold in the tunnel? People have speculated about the conformation adopted by nascent proteins as they traverse the exit tunnel ever since the

existence of the tunnel was first suspected. Important insights can be derived by combining what has been learned about the tunnel crystallographically with biochemical data of the sort first reported almost 40 years ago.[32] In 1967, Rich and colleagues demonstrated that the last (*i.e.* C-terminal) 30 residues of nascent peptides are protected from proteolytic attack by their association with the ribosome. Many such experiments have been performed since using ribosomes from different sources and nascent peptides of different kinds (for review see [33]). They have all consistently shown that the length of the peptides protected by the ribosome is 30 to 40 residues. Since the distance from the PTC to the distal end of the tunnel is about 80 Å, the distance of the tunnel traversed by each residue of nascent peptides is, on average, 2.0 to 2.7 Å. This distance is considerably less than the 3.5 Å per residue it would be if nascent proteins were in an extended conformation, but not as short as the 1.5 Å per residue it would be if nascent peptides were entirely α-helical. An increasing body of experimental data supports the view that nascent peptides are indeed α-helical conformation.[34;35] Furthermore, a recent theoretical study by Thirumalai and co-workers has shown that confinement of peptide chains in cylindrical cavities the size of the exit tunnel drives them to adopt α-helical conformations entropically.[36]

All of the above not withstanding, it has been suggested that nascent peptides might become at least partially folded at the tertiary level inside the tunnel.[19] Is this geometrically plausible? Accidentally, we obtained some information about the bore of the tunnel several years ago from a heavy-atom isomorphous replacement experiment done using a cluster compound containing 11 tungsten atoms.[37] This irregularly shaped molecule, which has maximum linear dimensions of 13.6 Å, binds to 3 sites inside the

tunnel as well as to a single site in the tRNA binding cleft (Figure 7C).[38] As it happens, the diameter of the largest sphere that can fit inside the tunnel is only slightly larger, 13.7 Å. The diameter of a poly-alanine α-helix is considerably smaller than this at 9.2 Å, and as expected, α-helices will fit comfortably in the tunnel provided some bends in the helix axis are permitted (Figure 7A). However, small protein domains, such as an IgG domain will not fit (Figure 7B). Thus neither the biochemical data available nor the structure of the large ribosomal subunit itself supports the hypothesis that tertiary folding of nascent peptides occurs in the ribosome.

Those who believe that the tunnel does allow tertiary folding might argue that the conformation of the tunnel is not fixed; that it might be able to accommodate large objects the way a boa constrictor swallows a pig. While it is impossible to rule this hypothesis out, it is implausible. Since the tunnel lies in the middle of a huge macromolecular structure, any conformational change that significantly altered its dimensions would require the disruption of a very large number of tertiary interactions. Furthermore, the tunnel has the same size and shape in the crystal structures now available for large ribosomal subunits from two different species[3;4] and for complete ribosomes from two other species.[5;6]

Finally, in considering the possibility that the tunnel functions as a chaperone, it is useful to compare the interior chamber of a known chaperone, the thermosome,[39] with the interior of the ribosomal tunnel. The two volumes are similar in length, but their diameters differ enormously (Figure 8). Although, a sphere 50 Å in diameter would fit comfortably within the interior cavity of the thermosome, nothing anywhere near this large will fit inside the ribosome.

Is the large ribosomal subunit impermeable to ions? The translocon is a large, membrane protein complex to which the large subunit of the ribosome binds when it is synthesizing proteins destined for insertion into membranes or passage through membranes. The translocon binds to the distal end of the exit tunnel. One of the interesting properties of the translocon-ribosome complex is that protein secretion is not accompanied by ion flows that would discharge the electrochemical gradients that exist across membranes. It is not easy to understand how this is accomplished. Peptide chains are irregular in shape, and one would have anticipated that even a channel that was only just big enough to allow peptides to pass might also accommodate an occasional low molecular weight ion.

Johnson and coworkers have observed that fluorophores inside the peptide exit tunnel of translocon-bound ribosomes cannot be quenched by iodide ions added to the surrounding solution.[21] Consequently they proposed that the ribosome itself is part of the system that makes the ribosome translocon complex ion-tight. This hypothesis appeared to be further supported by a recent report that puromycin treatment of endoplasmic reticulum membranes containing ribosome-translocon complexes with attached nascent polypeptides, which are normally ion-tight, leads to transmembrane calcium ion flow.[40] It is hypothesized that the release of nascent polypeptides from the ribosomes caused by reaction with puromycin opens the PTC end of the exit tunnel allowing ions to flow.

Both Johnson's result and the puromycin-calcium result can be understood geometrically. The Stokes' radii of hydrated cations and anions fall in the range of 3.3 to 4.1 Å (Table 2),[41] which is larger than the radius of the largest sphere (3.0 Å) capable of

leaking out of (or into) the exit tunnel laterally (Figure 5A), assuming that the wall of the tunnel is rigid. One could argue that these ions might enter ribosomal channels having diameters less than their hydrated radii because polar groups in the ribosome can replace waters of hydration. Thus the impermeability of the ribosome to iodide and calcium ions does not mean that the ribosome is impermeable to protons or water, or that it is part of the seal that keeps the ribosome-translocon complex from leaking ions. It should also be noted that recent crystallographic results suggest that the translocon pore itself is likely to be the component of the system that makes it ion-tight.[42]

RNA is "wet". This study shows that the interior of the ribosome includes a large network of interconnected channels; solvent occupies more then one third of the volume inside the shell of the large subunit. Protein assemblies tend to be much drier in their interiors. There is little or no solvent within protein domains, and only a few of them have internal solvent volumes of substantial size, *e.g*. GroEL. It is easy to understand why the ribosome is so wet. Most of the RNA in the ribosome is helical, and no matter how helices ~20 Å in diameter are packed, the gaps between them are bound to be large enough to accommodate solvent. Furthermore, individual RNA helices do not efficiently fill the cylindrical volumes that enclose their structures because their grooves are large. Another reason RNAs pack inefficiently is their packed structures must accommodate the water molecules and counterions required to neutralize their backbone charges. Protein domains, on the other hand, are as densely packed in their interiors as crystals of small organic molecules.[31]

To see whether of not the solvent-rich interior of the ribosome is a generic property of the ribosome, other large RNA molecules (group I intron[43]) and nucleic acid-protein complexes (small ribosomal subunit[48] and nucleosome[49]) were analyzed the same way. They also have large amounts of interior solvent (between 31% and 36% by volume), which indicates that the high solvent content of the interior of the ribosome is a general property of RNA and ribonucleoprotein complexes

Thermal fluctuations. Some of our conclusions about the properties of the polypeptide exit tunnel might be challenged on the grounds that biological structures cannot be fully understood on the basis of their time-averaged structures. For example, the atoms that form the wall of the exit tunnel of the ribosome, like every other atom in the ribosome exhibit thermal fluctuations in position. An estimate of the magnitude of this effect within the populations of macromolecules found in crystals can be obtained from the Debye-Waller factors (B-factors) assigned to the atoms as their structures are refined. The average value of the B-factors for RNA atoms in the *H. marismortui* large ribosomal subunit crystal structure is 30.8 (± 13.9) $Å^2$.[3] This number implies that in any given direction, the root mean square variation in the position of the average RNA atom in the crystals used to obtain the structure in question was 0.63 Å. Since static packing disorder is certain to contribute a significantly to the B-factors of these crystals, 0.63 Å is an upper bound estimate for the magnitude of thermal fluctuations. Random structural fluctuations of this magnitude cannot alter the dimensions of the tunnel enough to create protein accessible side branches, or to allow nascent polypeptides to form tertiary

structure, or to make it significantly more accessible to hydrated ions than its time-averaged structure indicates.

METHODS

Atomic coordinates and VDW radii. All files were taken from the Protein Data Bank (PDB).[44] 1JJ2 was the PDB file used for the *H. marismortui* large ribosomal subunit;[45] nine other large subunit structures gave similar results. For the group I intron the 1U6B[43] was the PDB file used. For the 30S subunit, PDB file 1N32 was used;[46] three other small subunit structures gave the same result. The nucleosome was taken from PDB file 1S32.[47] Spinach rubisco (1RCX)[30] was chosen as a typical protein because it has a central channel, long protein chains and minimal symmetry. The leucine zipper α-helix (Figure 7A) is from PDB file 1YSA[48] and the IgG domain (Figure 7B) is the first domain of NCAM Ig1-2-3 protein, PDB file 1QZ1.[49]

For all PDB files, hydrogens were added using the program REDUCE with default settings.[50] The van der Waals (VDW) radii used for atoms were taken from MSMS program based on stereochemistry of the RNA atoms,[51;52] and merged with coordinate data to create simple XYXR files that specify the coordinates and VDW radius of every atom in a molecule.

Cavity Volumes. Numbers of cavities and cavity volumes were determined using VOIDOO.[22;23] Scripts were written to facilitate the application of VOIDOO to multiple PDB files and probe sizes, as well as to analyze, merge and summarize its output files.

Using VOIDOO version 3.3.2 (040617) with extended array sizes, cavities for the ribosome were calculated using a cubic grid having linear dimensions of 0.4 Å (15 voxels per $Å^3$).

Existing software for analyzing internal solvent volumes. Several programs have been written to find interior voids, cavities and channels, but none proved satisfactory for our purposes.[22-28] For example, VOIDOO,[22;23] which is a program for extracting cavities of large macromolecules, was used to generate Figure 3. Cavities are voids in a structure that can contain a sphere larger than the largest sphere able to enter them from the exterior. Since the ribosomal exit tunnel is an open channel, VOIDOO is not an appropriate tool to use for its characterization.

HOLE[28] is a program that is used to extract ion channels from membrane protein structures by stepping through them from a given start point along a vector. When applied to the ribosome, HOLE terminated without having traversed the entire length of the tunnel because the tunnel axis does is not entirely straight. A new program was written using a stepping algorithm similar to that of HOLE that compensated for changes in the direction of the tunnel axis. While this new program could traverse the entire tunnel, the results were still unsatisfactory because of the large number of channels the structure contains.

SurfNET[24] is designed to extract all the channels in a given structure. It suffers from two shortcomings. First, since the ribosome is filled with channels, it is computationally expensive to use. Second, SurfNET uses the convex hull of the atom centers of a molecule to define the outer surface of molecules. The convex hull of a molecule

corresponds to the shell determined for it using a spherical probe in the limit as the probe radius goes to infinity.[53;54] The volume enclosed by the convex hull surface of the large subunit is much larger than the 100 Å shell surface shown in Figure 4. Hence, the construct analyzed by SurfNET includes much more bulk solvent than the 10 Å shell described above, which is undesirable.

The rolling ball approach to determining volumes of the ribosome. Initially, solvent-contact surface volumes were calculated using MSMS version 2.5.3.[51;52] (Version 2.5.5, which became available subsequently, does not work with structures containing as many atoms as the ribosome.) However, MSMS is designed for speed at the cost of accuracy, and did not always produce consistent results. For this reason we wrote a program of our own for performing these calculations.

This program determines the solvent-accessible volume of macromolecules by reading their XYZR files into memory and assigning to each atom a radius that is its VDW radius plus the radius of the probe (Figure 2). The coordinates of the molecule are then superimposed on a finely spaced, cubic grid. The solvent-accessible volume of the molecule is the set of all grid points that fall within the augmented atomic sphere of any atom. Next, the set of all "empty" grid points that are adjacent to points within the solvent-accessible volume of the molecule is determined. This set defines the solvent-accessible surface. The solvent-contact volume is obtained by removing from the set of points that constitute the solvent-accessible volume all points that are within the one probe radius from any of the points on the solvent-accessible surface (Figure 2). The volume is determined multiplying the total number of occupied grid points by the

individual grid point (voxel) volume. Not surprisingly, this algorithm works best if the grid used is very fine. For all calculations reported above, the linear dimensions of the grid were 0.16 Å to 0.22 Å, *i.e*. 244 to 94 grid points per $Å^3$. Figures showing such surfaces were obtained by writing their grids in EZD electron density file format, converting that file into a CCP4 file using MAPMAN,[22;55] and finally rendering them using the PyMol graphics program.[56] For figures, a 0.3 Å (37 grid points per $Å^3$) grids were calculated, which were then shrunk to 0.6 Å to reduce PyMol rendering time and prevent crashing.

Extraction of Interior Volumes. The first step in the determination of the surface of the polypeptide tunnel was the calculation of the shell volume of the 50S subunit, determined using a 10 Å radius probe, on a fine grid. Second, the accessible surface was determined on the same grid by the rolling ball method (Figure 2) using a smaller probe. Subsequently, the accessible volume grid of the smaller probe was subtracted from the shell volume grid to yield the set of interior grid points that are accessible to the smaller probe. This subtraction yields all points reachable by the probe center within the shell, but not necessarily from the tunnel. The subset of grid points connected to defined tunnel regions was then identified from the larger set. That subset contains only points accessible to the probe center. In order to obtain the contact volume of the tunnel, the accessible volume subset was convoluted with the probe sphere. The set of interior grid points so defined is equivalent to what would be obtained if a probe placed manually in the lumen of the tunnel were allowed to roam freely in the structure without exiting the tunnel. Typical results are shown in Figures 5B. To ensure the continuity of the exit

tunnel with the tRNA binding cleft of the large subunit, the highly dynamic residue U2620 was removed from above the tunnel when generating figures, but not when numerical calculations were being done. The unlabeled exit tunnel surfaces shown in figures were generated using probes between 2.9 and 3.2 Å.

The interior volume of the thermosome, from *Thermus acidophilum* (PDB ID: 1A6D),[39] was computed as described above except that that its shell volume was computed using a 40 Å probe in order to ensure that its interior cavity would be entirely included in its shell. The probe used to obtain its interior volume had a radius of 4 Å; smaller probes can "leak" out of its interior.

Calculation of Linear Dimensions. To determine the width extended poly-alanine chains and α-helices, alanines were modeled in increments of 50 amino acids. The lengths of the models were calculated using Cα distances and volumes using a rolling probe of 5 Å. The cross-section of both alanine chains was then determined from the slopes of a linear regression of the data. Diameters were then calculated assuming the cross-section was circular.

Availability of Software. The volume calculation programs and source code described here are intended to be applicable to similar problems and can be obtained online at *http://geometry.molmovdb.org/3v/*.

ACKNOWLEDGEMENTS

This work was supported by a grant from the National Institute of Health to PBM and TAS (GM022778). NRV thanks Joshua Young for critical reading of the manuscript and Elizabeth Pollock and Gregor Blaha for helpful discussions.

FIGURE LEGENDS

Figure 1: Slices through the interior of the ribosome and selected proteins showing interior solvent volumes. The images of the large subunit shown in (A) and (B) were obtained by cutting the particle into two pieces along a plane perpendicular to the line of view, and exposing its interior by removing the portion nearer to the viewer. In the left pair of images, the particle is oriented in the so-called crown view, and in the right pair of images, the particle was rotated by 90° about the vertical axis prior to being sliced. The cutting plane used to generate the right hand pair of images was chosen to intersect the protein exit tunnel. In all images, RNA is gray, protein is yellow, and solvent is blue. (A) The solvent volume, represented in blue, of a single slice through the large ribosomal subunit. (B) The locations of water molecules (small blue spheres) reported in the crystal structure of the large ribosomal subunit of *H. marismortui* (PDB id: 1JJ2)[45] are displayed that fall within the space visible in part (A) of this figure. (C) Images of spinach rubisco (PDB id: 1RCX) showing a slice through its central cavity (left), and the solvent between its VDW surface and its shell (right).[57] All images are drawn to scale.

Figure 2: Rolling sphere definitions. A sphere of radius R is shown rolling across the surface of a macromolecule defined by atoms 1-12. The contact surface is the surface defined by the exterior of the sphere as it rolls across the array of atoms. The accessible surface is the surface visited by the center of the sphere as it rolls across the molecular surface, and it is called the accessible surface because the center of a sphere of radius R cannot enter the volume inside the accessible surface. [31]

Figure 3: Percent cavity volume as a function of probe radius. The percent of the shell of the large subunit (red) accounted for by cavities is plotted as function probe radius. A cavity is a volume that will accommodate a sphere of some radius, but from which a sphere of that radius cannot escape. Once the probe radius is greater than 9.5 Å, there are no cavities.

Figure 4: The 10 Å contact surface of the *H. marismortui* large subunit compared to its 100 Å contact surface. The structure whose contact surfaces are shown is that described by PDB entry 1JJ2.[45] A stereo view is provided of the contact surface defined for the particle using a sphere of radius 10 Å, *i.e.* its shell, (blue solid) with the contact surface obtained for the same object using a sphere that is 100 Å in radius (yellow mesh) superimposed.

Figure 5: The dependence of volume accessible from the tunnel on probe radius. (A) The total solvent volume (red) and the volume that is accessible to a probe placed in the

tunnel (green) is plotted as a function of probe radius. The gray line highlights the radius of a water-sized probe (1.5 Å). (B) The shapes of the surfaces of the contact volumes that are accessible to probes placed in the tunnel lumen. The location of the active site α-amino is marked in all pictures with a large blue sphere. In the left-most figure the portion of surface that corresponds to the 3 Å shell is shown in yellow.

Figure 6: Stereo diagram of the polypeptide exit tunnel showing the positions of landmarks. The wild-type *H. marismortui* exit tunnel is shown. The figure is oriented with the tRNA binding cleft at the top and the exit at the bottom. L22 is orange, L4 is pink, A-site and P-site CCA tRNA tails are green and yellow, respectively, the active site α-amino is blue, and the contact surface of the exit tunnel is transparent gray.

Figure 7: Stereo projections of space-filling models of molecules inside the peptide exit tunnel. (A) A straight, 41 amino acid α-helix[48] is inserted into the tunnel, oriented to maximize the fraction of the helix inside the tunnel. (B) A model of an IgG domain[49] is placed in the middle of the exit tunnel, again oriented so that the fraction of the domain inside the tunnel lumen will be maximal. (C) Four W11 heavy cluster compounds are displayed lodged in the tunnel and in the cleft immediately above the peptidyl transferase site. Their positions are those occupied by the molecule in question when it is soaked into preformed crystals of the large ribosomal subunit from *H. marismortui*.[38]

Figure 8: The interior volumes of a chaperonin and the polypeptide exit tunnel compared. (A) The contact surface of the internal chamber of the *T. acidophilum*

chaperonin[39] (green). (B) The contact surface of *H. marismortui* large ribosomal subunit exit tunnel (red) determined using a 2.95 Å radius probe. Images are drawn to the same scale.

Table 1: Volumes of the Ribosome. The solvent volume is the shell volume minus solvent-contact volume. The empty volume is the solvent-contact volume minus VDW volume. The fractional solvent volume is the solvent-contact volume divided by the shell volume. The tunnel volume is 93% of the solvent volume at 1.5 Å and 4.4% of the solvent at 3.0 Å.

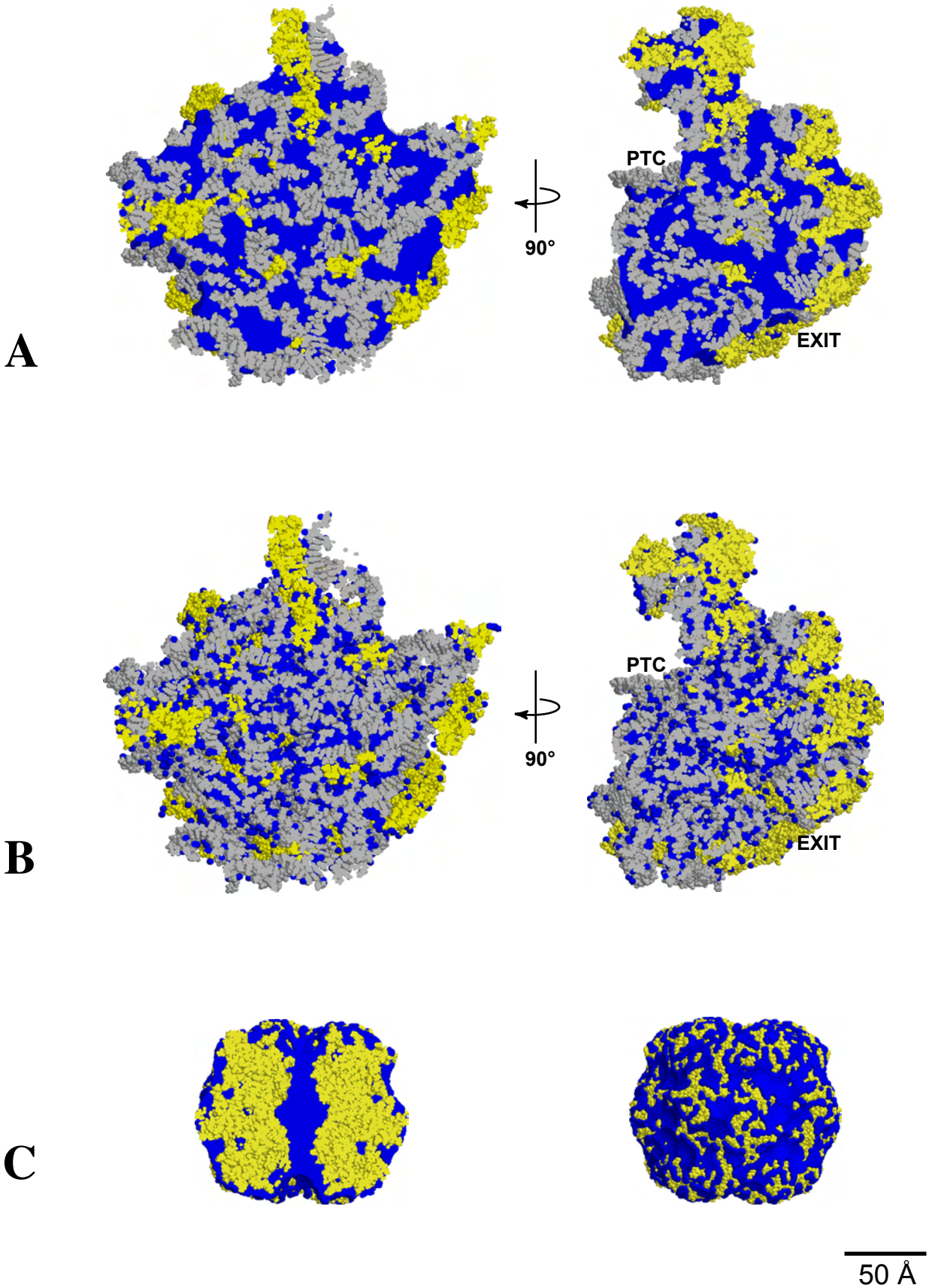
Table 2: Stokes Radii of Ions. The Stokes radii of several ions are shown.[41]

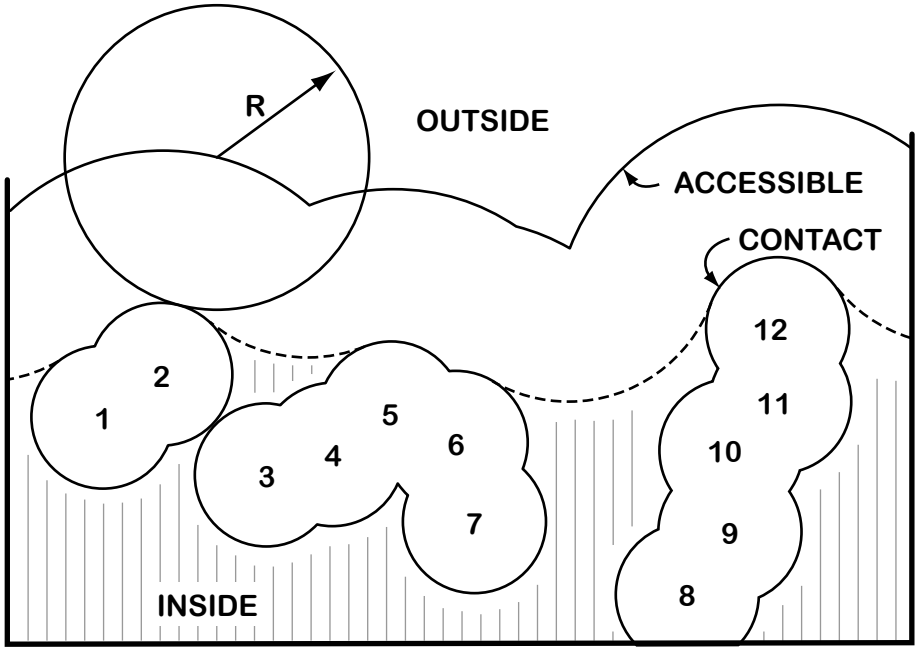
REFERENCES

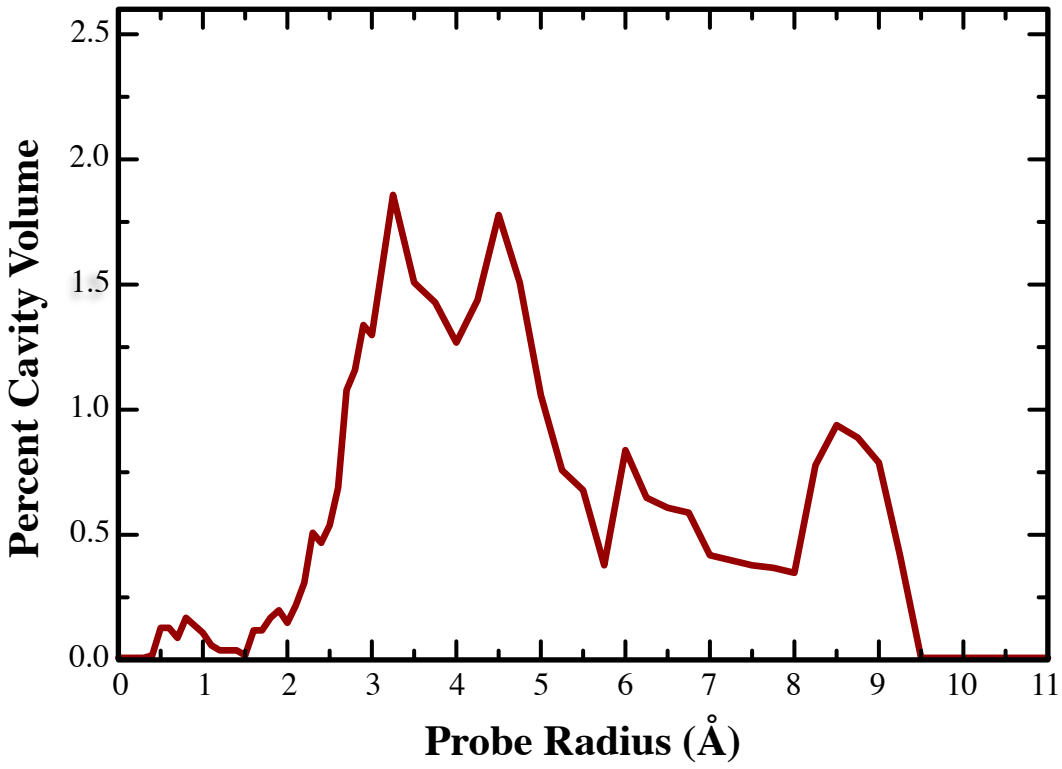
1. Frank, J., Zhu, J., Penczek, P., Li, Y., Srivastava, S., Verschoor, A., Radermacher, M., Grassucci, R., Lata, R. K. & Agrawal, R. K. (1995). A model of protein synthesis based on cryo-electron microscopy of the E. coli ribosome. Nature 376, 441-4.
2. Bernabeu, C. & Lake, J. A. (1982). Nascent polypeptide chains emerge from the exit domain of the large ribosomal subunit: immune mapping of the nascent chain. Proc Natl Acad Sci U S A 79, 3111-5.
3. Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science 289, 905-20.
4. Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F. & Yonath, A. (2001). High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. Cell 107, 679-88.
5. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M. & Cate, J. H. (2005). Structures of the bacterial ribosome at 3.5 A resolution. Science 310, 827-34.
6. Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H. & Noller, H. F. (2001). Crystal structure of the ribosome at 5.5 A resolution. Science 292, 883-96.
7. Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P. B. & Steitz, T. A. (1999). Placement of protein and RNA structures into a 5 A-resolution map of the 50S ribosomal subunit. Nature 400, 841-7.
8. Nakatogawa, H. & Ito, K. (2002). The ribosomal exit tunnel functions as a discriminating gate. Cell 108, 629-36.
9. Nakatogawa, H., Murakami, A. & Ito, K. (2004). Control of SecA and SecM translation by protein secretion. Curr Opin Microbiol 7, 145-50.
10. Tenson, T. & Ehrenberg, M. (2002). Regulatory nascent peptides in the ribosomal tunnel. Cell 108, 591-4.
11. Gong, F. & Yanofsky, C. (2002). Instruction of translating ribosome by nascent peptide. Science 297, 1864-7.

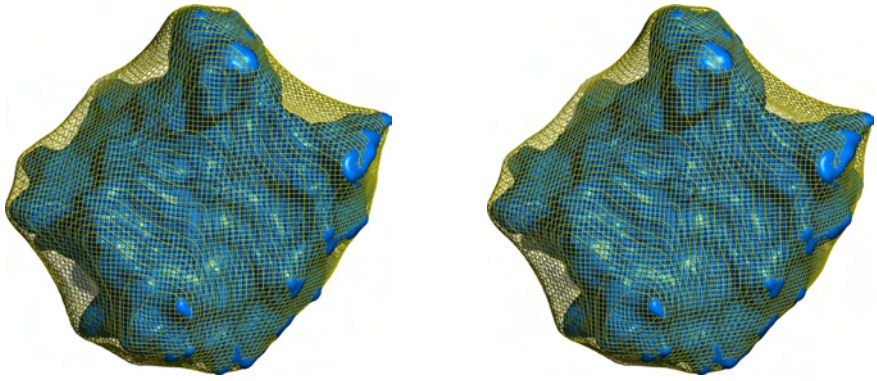
12. Hansen, J. L., Moore, P. B. & Steitz, T. A. (2003). Structures of five antibiotics bound at the peptidyl transferase center of the large ribosomal subunit. J Mol Biol 330, 1061-75.
13. Agmon, I., Auerbach, T., Baram, D., Bartels, H., Bashan, A., Berisio, R., Fucini, P., Hansen, H. A., Harms, J., Kessler, M., Peretz, M., Schluenzen, F., Yonath, A. & Zarivach, R. (2003). On peptide bond formation, translocation, nascent protein progression and the regulatory properties of ribosomes. Derived on 20 October 2002 at the 28th FEBS Meeting in Istanbul. Eur J Biochem 270, 2543-56.
14. Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. (2000). The structural basis of ribosome activity in peptide bond synthesis. Science 289, 920-30.
15. Baram, D., Pyetan, E., Sittner, A., Auerbach-Nevo, T., Bashan, A. & Yonath, A. (2005). Structure of trigger factor binding domain in biologically homologous complex with eubacterial ribosome reveals its chaperone action. Proc Natl Acad Sci U S A 102, 12017-22.
16. Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C. L., 3rd, Ban, N. & Frank, J. (2005). Structure of the E. coli protein-conducting channel bound to a translating ribosome. Nature 438, 318-24.
17. Zengel, J. M., Jerauld, A., Walker, A., Wahl, M. C. & Lindahl, L. (2003). The extended loops of ribosomal proteins L4 and L22 are not required for ribosome assembly or L4-mediated autogenous control. Rna 9, 1188-97.
18. Gabashvili, I. S., Gregory, S. T., Valle, M., Grassucci, R., Worbs, M., Wahl, M. C., Dahlberg, A. E. & Frank, J. (2001). The polypeptide tunnel system in the ribosome and its gating in erythromycin resistance mutants of L4 and L22. Mol Cell 8, 181-8.
19. Gilbert, R. J., Fucini, P., Connell, S., Fuller, S. D., Nierhaus, K. H., Robinson, C. V., Dobson, C. M. & Stuart, D. I. (2004). Three-dimensional structures of translating ribosomes by Cryo-EM. Mol Cell 14, 57-66.
20. Pal, S., Chandra, S., Chowdhury, S., Sarkar, D., Ghosh, A. N. & Gupta, C. D. (1999). Complementary role of two fragments of domain V of 23 S ribosomal RNA in protein folding. J Biol Chem 274, 32771-7.
21. Liao, S., Lin, J., Do, H. & Johnson, A. E. (1997). Both lumenal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration. Cell 90, 31-41.
22. Kleywegt, G. J., Zou, J. Y., Kjeldgaard, M. & Jones, T. A. (2001). Around O. In International Tables for Crystallography, Vol. F. Crystallography of Biological Macromolecules (Rossman, M. G. E. A., ed.), pp. 353-356, 366-367. Kluwer Academic Publishers, The Netherlands.
23. Kleywegt, G. J. & Jones, T. A. (1994). Detection, Delineation, Measurement And Display Of Cavities In Macromolecular Structures. Acta Crystallographica Section D-Biological Crystallography 50, 178-185.
24. Laskowski, R. a. (1995). Surfnet - a Program for Visualizing Molecular-Surfaces, Cavities, and Intermolecular Interactions. Journal of Molecular Graphics 13, 323-&.
25. Bakowies, D. & van Gunsteren, W. F. (2002). Water in protein cavities: A procedure to identify internal water and exchange pathways and application to fatty acid-binding protein. Proteins-Structure Function And Genetics 47, 534-545.
26. Liang, J., Edelsbrunner, H. & Woodward, C. (1998). Anatomy of protein pockets and cavities: Measurement of binding site geometry and implications for ligand design. Protein Science 7, 1884-1897.
27. Laskowski, R. A., Luscombe, N. M., Swindells, M. B. & Thornton, J. M. (1996). Protein clefts in molecular recognition and function. Protein Science 5, 2438-2452.
28. Smart, O. S., Goodfellow, J. M. & Wallace, B. A. (1993). The pore dimensions of gramicidin A. Biophys J 65, 2455-60.
29. Richards, F. M. (1977). Areas, volumes, packing and protein structure. Annu Rev Biophys Bioeng 6, 151-76.
30. Taylor, T. C. & Andersson, I. (1997). The structure of the complex between rubisco and its natural substrate ribulose 1,5-bisphosphate. J Mol Biol 265, 432-44.
31. Richards, F. M. (1974). The interpretation of protein structures: total volume, group volume distributions and packing density. J Mol Biol 82, 1-14.
32. Malkin, L. I. & Rich, A. (1967). Partial resistance of nascent polypeptide chains to proteolytic digestion due to ribosomal shielding. J Mol Biol 26, 329-46.
33. Hardesty, B. & Kramer, G. (2001). Folding of a nascent peptide on the ribosome. Prog Nucleic Acid

Res Mol Biol 66, 41-66.
34. Lu, J. & Deutsch, C. (2005). Folding zones inside the ribosomal exit tunnel. Nat Struct Mol Biol 12, 1123-9.
35. Woolhead, C. A., McCormick, P. J. & Johnson, A. E. (2004). Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. Cell 116, 725-36.
36. Ziv, G., Haran, G. & Thirumalai, D. (2005). Ribosome exit tunnel can entropically stabilize alpha-helices. Proc Natl Acad Sci U S A 102, 18956-61.
37. Wei, Z. Y., Dickman, M. H. & Pope, M. T. (1997). New routes for multiple derivatization of polyoxometalates, bis(acetato)dirhodium-11-tungstophosphate, [(PO4)W11O35{Rh-2(OAc)(2)}](5-). Inorganic Chemistry 36, 130.
38. Ban, N., Freeborn, B., Nissen, P., Penczek, P., Grassucci, R. A., Sweet, R., Frank, J., Moore, P. B. & Steitz, T. A. (1998). A 9 A resolution X-ray crystallographic map of the large ribosomal subunit. Cell 93, 1105-15.
39. Ditzel, L., Lowe, J., Stock, D., Stetter, K. O., Huber, H., Huber, R. & Steinbacher, S. (1998). Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. Cell 93, 125-38.
40. Van Coppenolle, F., Vanden Abeele, F., Slomianny, C., Flourakis, M., Hesketh, J., Dewailly, E. & Prevarskaya, N. (2004). Ribosome-translocon complex mediates calcium leakage from endoplasmic reticulum stores. J Cell Sci 117, 4135-42.
41. Marcus, Y. (1985). Ion solvation, Wiley, Chichester; New York.
42. Menetret, J. F., Hegde, R. S., Heinrich, S. U., Chandramouli, P., Ludtke, S. J., Rapoport, T. A. & Akey, C. W. (2005). Architecture of the ribosome-channel complex derived from native membranes. J Mol Biol 348, 445-57.
43. Adams, P. L., Stahley, M. R., Kosek, A. B., Wang, J. & Strobel, S. A. (2004). Crystal structure of a self-splicing group I intron with both exons. Nature 430, 45-50.
44. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). The Protein Data Bank. Nucleic Acids Res 28, 235-42.
45. Klein, D. J., Schmeing, T. M., Moore, P. B. & Steitz, T. A. (2001). The kink-turn: a new RNA secondary structure motif. Embo J 20, 4214-21.
46. Ogle, J. M., Murphy, F. V., Tarry, M. J. & Ramakrishnan, V. (2002). Selection of tRNA by the ribosome requires a transition from an open to a closed form. Cell 111, 721-32.
47. Edayathumangalam, R. S., Weyermann, P., Gottesfeld, J. M., Dervan, P. B. & Luger, K. (2004). Molecular recognition of the nucleosomal "supergroove". Proc Natl Acad Sci U S A 101, 6864-9.
48. Ellenberger, T. E., Brandl, C. J., Struhl, K. & Harrison, S. C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. Cell 71, 1223-37.
49. Soroka, V., Kolkova, K., Kastrup, J. S., Diederichs, K., Breed, J., Kiselyov, V. V., Poulsen, F. M., Larsen, I. K., Welte, W., Berezin, V., Bock, E. & Kasper, C. (2003). Structure and interactions of NCAM Ig1-2-3 suggest a novel zipper mechanism for homophilic adhesion. Structure 11, 1291-301.
50. Word, J. M., Lovell, S. C., Richardson, J. S. & Richardson, D. C. (1999). Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J Mol Biol 285, 1735-47.
51. Sanner, M. F. (1995). MSMS: Michel Sanner's Molecular Surface 2.5.3 edit. Scripps Research Institute.
52. Sanner, M. F., Olson, A. J. & Spehner, J. C. (1996). Reduced surface: an efficient way to compute molecular surfaces. Biopolymers 38, 305-20.
53. Gerstein, M. & Richards, F. M. (2001). Protein Geometry: Distances, Areas, and Volumes. In International Tables for Crystallography (Rossmann, M. & Arnold, E., eds.), Vol. F: Crystallography of biological macromolecules, pp. 531-539. Springer, Dordrecht.
54. Sanner, M., Olson, A. J. & Spehner, J. C. (1995). Proc. 11th ACM Symp. Comp. Geom.
55. Kleywegt, G. J. & Jones, T. A. (1993). MAPMAN. In RAVE Package 041015 edit. Uppsala Software Factory.
56. DeLano, W. L. (2002). The PyMOL Molecular Graphics System 0.95 edit. DeLano Scientific, San Carlos, CA, USA.
57. Thoden, J. B., Miran, S. G., Phillips, J. C., Howard, A. J., Raushel, F. M. & Holden, H. M. (1998). Carbamoyl phosphate synthetase: caught in the act of glutamine hydrolysis. Biochemistry 37, 8825-31.


A
PTC
90°
EXIT
B
PTC
90°
EXIT
C
50 Å

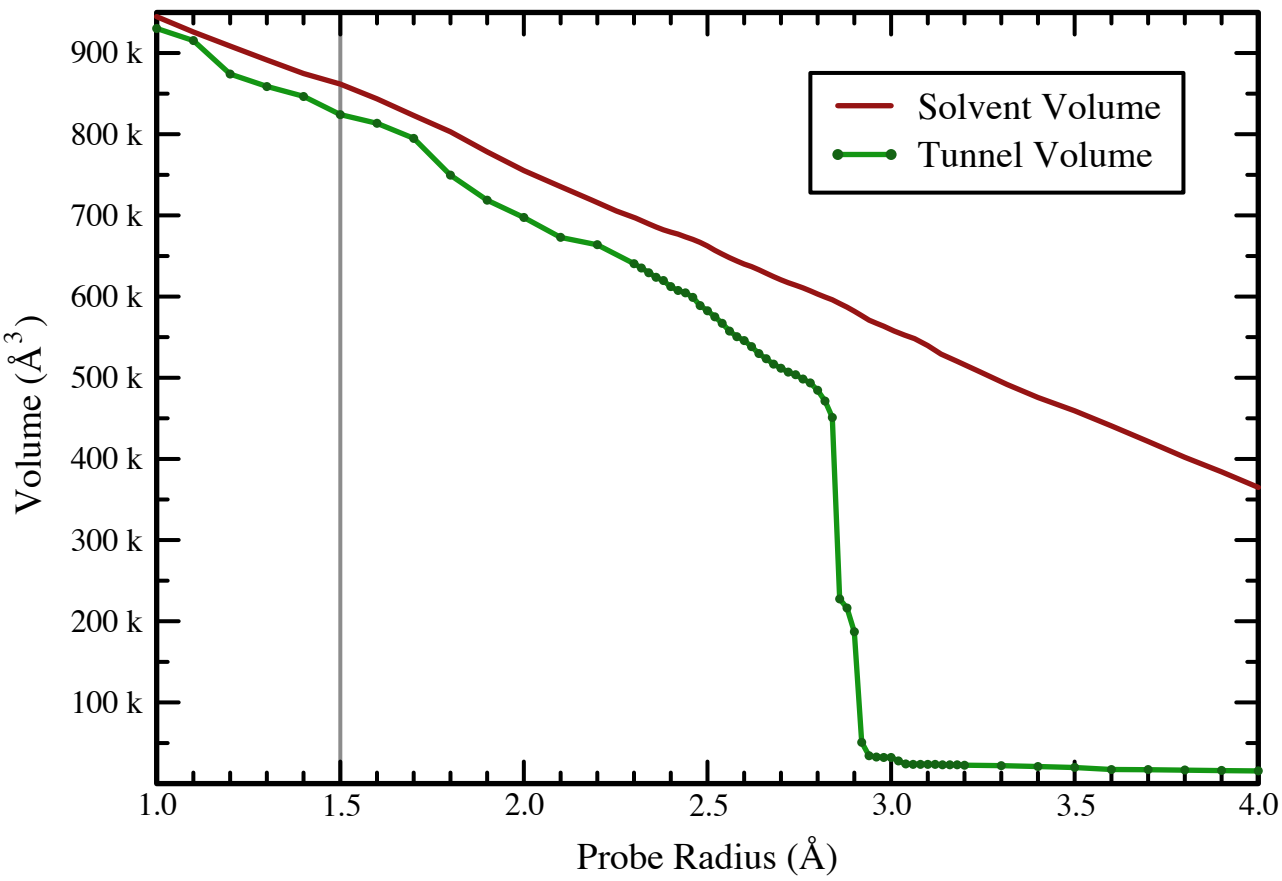

R
OUTSIDE
ACCESSIBLE
CONTACT
1
2
3
4
5
6
7
8
9
10
11
12
INSIDE


Percent Cavity Volume
2.5
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
6
7
8
9
10
11
Probe Radius (Å)

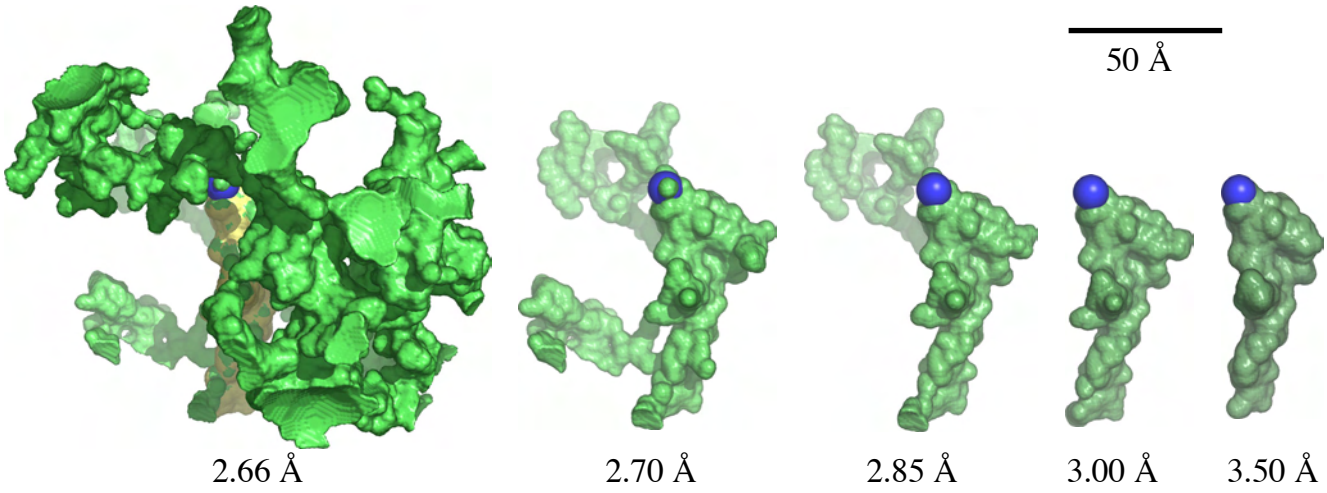


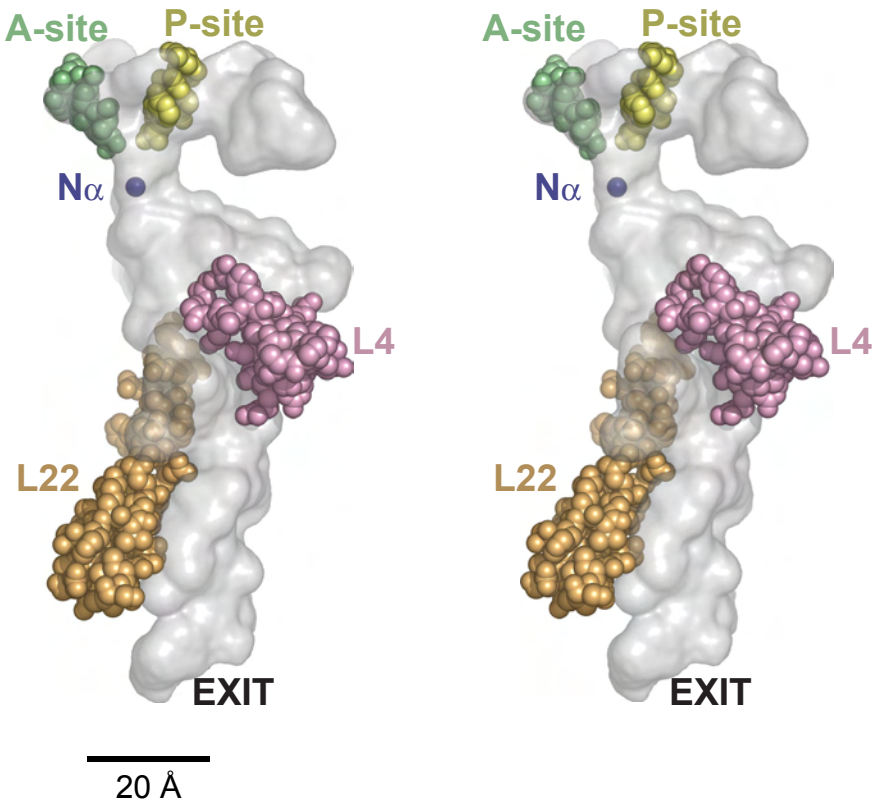
100 Å

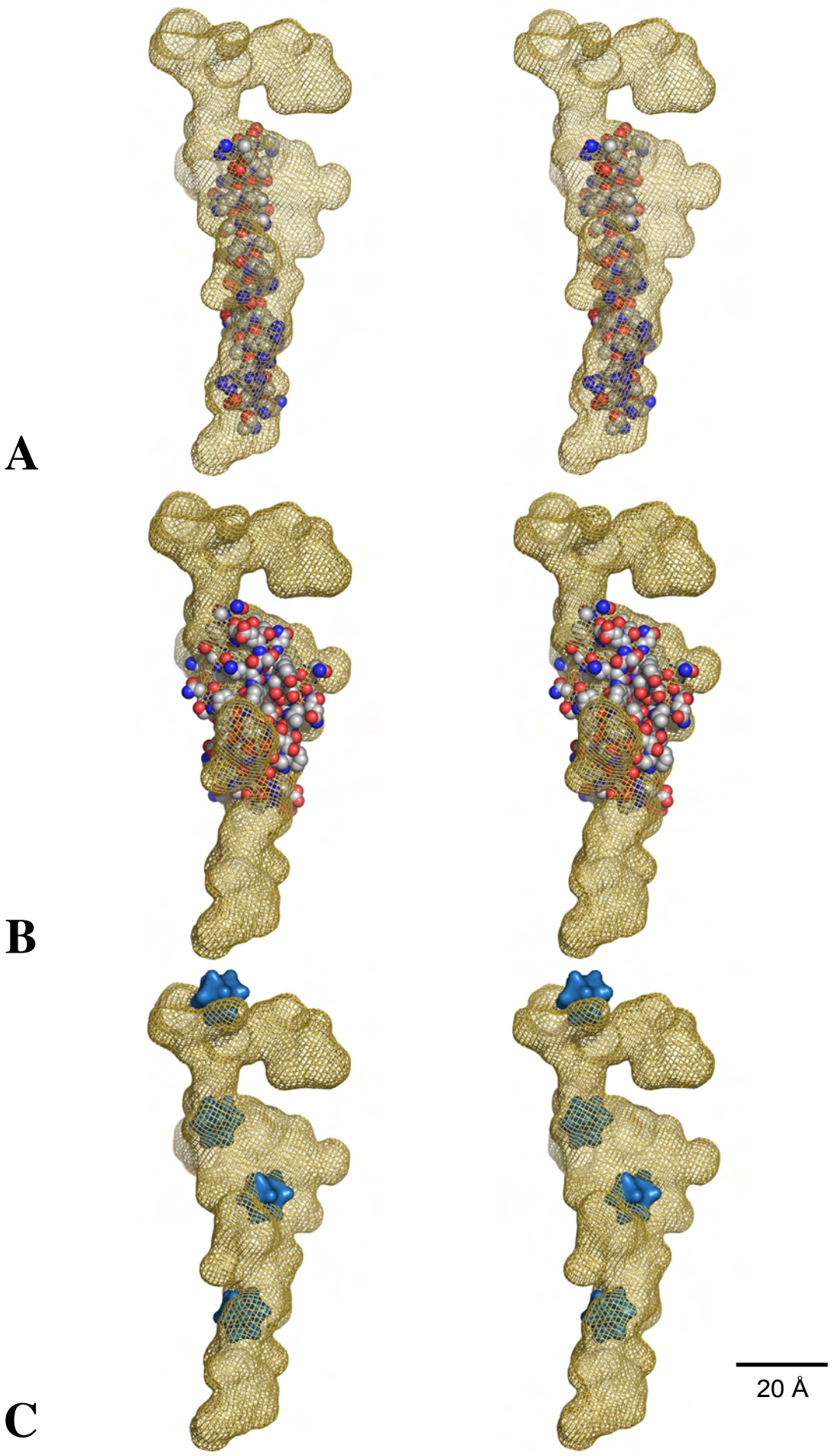
A

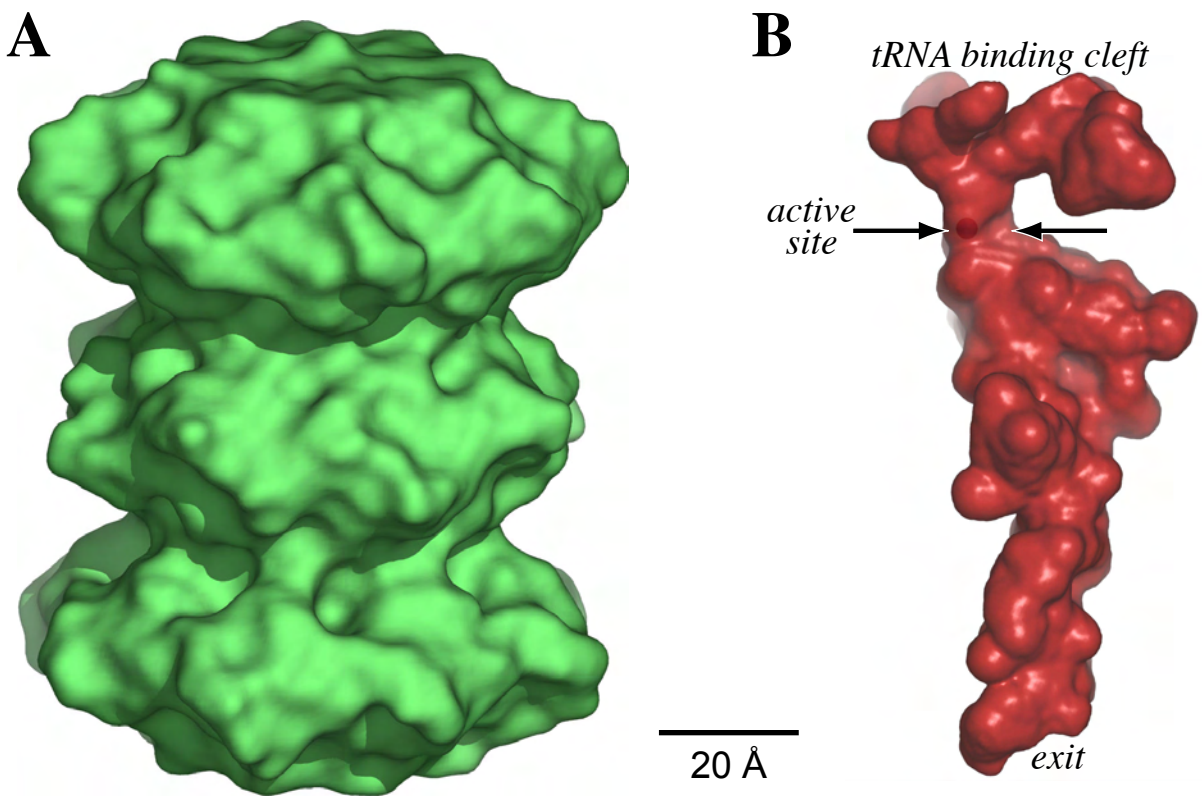

Solvent Volume
Tunnel Volume
Volume (Å^3)
900 k
800 k
700 k
600 k
500 k
400 k
300 k
200 k
100 k
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Probe Radius (Å)

B


50 Å
2.66 Å
2.70 Å
2.85 Å
3.00 Å
3.50 Å


A-site
P-site
Nα
L4
L22
EXIT
A-site
P-site
Nα
L4
L22
EXIT
20 Å


A
B
C
20 Å


A
B
tRNA binding cleft
active
site
20 Å
exit

Exterior Volumes (Å^3)			*Mass*
shell volume	solvent excluded volume	VDW volume	molecular mass (Da)
2,290,980	1,395,480	1,086,421	1,340,664
(1)	(2)	(3)	

Interior Volumes (Å^3)			*Percentage*
solvent volume	empty volume	solvent cavity volume	fractional solvent volume
895,500	309,059	1,578	39.1%
(1)-(2)	(2)-(3)		100%–(2)/(1)

Tunnel Volumes (Å^3)			
1.5 Å tunnel volume	1.5 Å solvent volume	3.0 Å tunnel volume	3.0 Å solvent volume
824,323	889,475	24,028	551,943
	92.68%		4.35%

	Stokes Radius		
ion	***unhydrated***	***hydrated***	***hydration #***
I^-	2.2	3.33	2.8
K^+		3.34	5.1
Cl^-		3.34	3.9
Na^+		3.51	6.5
Ca^{2+}		3.96	10.4
Mg^{2+}		4.12	11.7