Supporting Information for

The History of the Ribosome and the Origin of Translation

Anton S. Petrov[1*], Burak Gulen[1], Ashlyn M. Norris[1], Nicholas A. Kovacs[1], Chad R. Bernier[1], Kathryn A. Lanier[1], George E. Fox[2], Stephen C. Harvey[3], Roger M. Wartell[3], Nicholas V. Hud[1] and Loren Dean Williams[1*]

[1]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, USA

[2]Department of Biology and Biochemistry, University of Houston, Houston, TX, 77204, USA

[3]School of Biology, Georgia Institute of Technology, Atlanta, Georgia 30332, USA

*Corresponding authors:

Name: Anton S. Petrov, Loren Dean Williams

Address: School of Chemistry and Biochemistry, Georgia Institute of Technology,
 Atlanta, GA 30332-0400, USA.

Phone: 404-894-9752, Fax: 404-894-7452

Email: anton.petrov@biology.gatech.edu, loren.williams@chemistry.gatech.edu

Key words: RNA evolution, ribosome, translation, origin of life

Materials and Methods

Secondary structures.

All secondary structures of LSU and SSU rRNAs are taken from our public gallery (http://apollo.chemistry.gatech.edu/RibosomeGallery/). Data are mapped onto secondary structures with RiboVision (1-3).

Three-dimensional structures.

Three *dimensional structures of ribosomal particles were obtained from the PDB database (PDB ID*s: 4V9D (4), 4V51 (5), 4V6W, 4V6X (6). Local and global superimpositions were performed using the built-in cealign functionality of PyMOL (7).

Sequence conservation.

A multiple sequence alignment of SSU rRNA sequences from 133 organisms (DataSet S1), intended to represent a sparse broad sampling of all three domains of life, was generated using the SINA aligner (8) from the SILVA database (9) and manually adjusted to improve the a) 3D superposition of species between the domains of life, and b) relative locations of the eukaryotic expansion segments. The adjustments were done by mapping representative sequences onto the available 3D structures and correcting the positions of nucleotides within the alignment based on their locations in the 3D structures. The sequence conservation of the SSU rRNA is calculated from the alignment and is visualized as Shannon entropy (10-12) mapped onto the secondary structure (Fig. S1).

Types of insertion fingerprints.

Insertion fingerprints can be grouped by morphological relationships between the trunk and branch. Each group is represented within the eukaryotic expansion segments, where the order of expansions and relative ages of RNA elements are well understood. Group descriptions and the number of members of each group within the common core of the SSU are given below.

Y group: A branch helix is inserted into a trunk helix forming a three-way junction. The helical axis of the branch is roughly perpendicular to that of the trunk. Stacking is continuous within the trunk (frequency: 10).

T group: A branch helix caps a trunk helix, forming a three-way junction. Stacking is continuous within the branch. The branch axis is roughly perpendicular to that of the trunk. If a branch inserts into a trunk, the resulting insertion fingerprint is in the T_1 group, if a branch caps a trunk, the fingerprint is in the T_2 group (frequency: 11).

X group: A branch helix is inserted into a trunk helix forming a four-way junction. The branch axis is parallel or slightly skewed relative to that of the trunk (frequency: 5).

I group (or L-group): A branch helix is inserted into a loop so that it extends the trunk. The branch maintains stacking continuity with the original trunk, although the joint may be bent (L-type) (frequency: 10).

The Accretion Model: SSU rRNA

Criteria for Identifying Ancestral Expansion Segments.

Sites of eukaryotic SSU expansions on the common core (Table S1), and ancestral SSU expansions within the common core (Fig. S2, Table S2) are identified by insertion fingerprints (13). An expansion segment is a more recent RNA element (a branch) that projects from a pre-existing RNA helix (a trunk). Insertion fingerprints mark sites where expansions are added to prior RNA, and are observed throughout the LSU and the SSU. Backbone conformation, phosphate proximity, stacking discontinuity, and geometric relationships between helical axes define insertion fingerprints. These characteristics of insertion fingerprints are identical for both subunits.

Helix Elongation.

Because helix elongation does not leave an insertion fingerprint, we assume that our method under-samples the number of ancestral expansion events. In the Accretion Model, AES/aes were added in their present size without elongation. In reality, many AES/aes probably budded and elongated before they were fully buried, subsumed, and frozen with the common core. One helical elongation included in the model is within h44. In this case, distinct elongation steps can be detected by temporal information from intersubunit bridges (Fig. S3) and from a minimal ribosomal model (14). In a few other instances, we could tentatively identify helix elongations either by irregularities in helical structures or by sharp turns of helical axes. In these cases, the elongated portion of an aes was assigned the same number as their founding aes with addition of the suffix "a" and "b".

Deconstruction of the SSU rRNA

In rewinding evolutionary processes, we started the deconstruction of the SSU from the subunit surface. The SSU was 'pruned' according to a structural continuity principle in which existing elements remain connected within a single RNA polymer. Additional information on order of acquisition was inferred from A-minor interactions, stacking continuity, and positioning

of aes. Collectively, the ordering information suggests that aes 1-5, which includes the central pseudoknot, is the oldest component of the SSU rRNA, evolved prior to the association with the LSU. The evolution of aes 1-5 itself requires special consideration to account for the events leading to strand separation and formation of the central pseudoknot (Phases 1-3), and therefore is described separately in detail. Directionality of the A-minor interactions reveals that aes 1 (helix 44 with mRNA and 5' terminus) is the primordial segment from which the SSU evolves (Fig. S5).The peripheral SSU domains incrementally were accreted onto aes 1-5 (Phases 4-6).

Criteria for Evolution Chronology of the SSU rRNA.

Using aes 1-5 as the ultimate primordial ancestors of the SSU rRNA we have recapitulated the timeline of SSU rRNA evolution. The chronology is based on:

a) *Fingerprints*. Trunks are older than branches (a branch is not added until after its trunk has formed).

b) *Dependency within subunits*. The directionality of A-minor interactions is dependent to independent, young to old. A-minor donors are younger than A-minor acceptors [see Steinberg (15)].

c) *Dependency between subunits*. Upon association of the subunits, SSU and LSU evolution is correlated by intersubunit bridges (several bridges contain A-minor interactions).

d) *Continuity*. rRNA evolved continuously from the inside to the outside in three dimensions.

f) *Function*. Evolution of function was gradual and continuous, and included stabilization of tRNAs at the A- and P-sites, SSU association with the LSU, binding of the elongation factors and translocation of mRNA.

The aes are numbered in apparent order of appearance in Table S2. The numbers and the exact temporal order of many peripheral aes are indeterminate, as they were acquired independently of each other. Their ordering cannot be inferred from the parameters listed above,

and relative numbers of these aes are somewhat arbitrary. This ambiguity does not affect conclusions of the study.

Detailed Evolutionary Model of the SSU rRNA.

Using a parsimonious temporal ordering of aes, we partitioned the evolutionary steps into phases. The exact margins between phases are not rigorously defined and are not overly significant. The association of the subunits takes place at the end of Phase 3. Evolution of Phases 1-3 is summarized in Fig. S4.

Phase 1. SSU begins as a primordial stem-loop. The SSU rRNA begins with an A-form helix capped by a tetraloop at base pair 1430-1470 of helix 44. At the other end, helix 44 is extended by a now vanquished helical region composed of the 3' terminus (1431-1440) and its complement modeled by mRNA of *T. thermophilus* [PDB ID: 4V51 (5)].

Phase 2. Dissociation of the termini. aes 2 and 3 are inserted into aes 1 on opposite sides of helix 44, near the termini of aes 1. The resulting five-way junction is unstable and is resolved by rearrangement; aes 3 shifts and rotates and the strand termini dissociate. aes 3 is extended by aes 3a. aes 2 and 3a form tertiary contacts with aes 1 (by a set of A-minor interactions), where universally conserved A1518 and A1519 of aes 2 and A908 and A909 of aes 3a interact with the minor grove of aes 1. A-minor interactions between aes 3a and aes 1 stabilize the newly rearranged conformation.

Phase 3. Formation of the CPK. Accretion of aes 4 provides a docking site for the dissociated 5' terminus, which forms a triple helix with helix 3. This step forms the central pseudoknot (CPK), sequestering the 5' terminus, and preventing re-association of the strand termini. Strand dissociation exposes the 3' terminus, which gains a single-stranded binding functionality. This sequence of events occurs in early SSU evolution, impacting subsequent structure and function of the SSU. aes 4a then aes 5 accrete onto aes 4. aes 5 provides additional stability to the complex, as universally conserved A766, A767, and A768 form A-minor

interactions by donating to aes 2. Finally, aes 3a is extended by aes 3b, which binds to aes 5, as universally conserved A900 and A901 of aes 3b form A-minor interactions by donating to aes 5. Thus, at the end of Phase 3, the SSU represents a stable complex, in which the 5' terminus is engaged in a triple helix within the central pseudoknot, and the unpaired 3' terminus has an ability to bind to a single stranded RNA. Its three-dimensional structure is maintained by interlocking universally conserved A-minor interactions that involve aes 1, aes 2, aes 3a, aes 3b, and aes 5 (Fig. S5). These segments also participate in ancient intersubunit bridges (Fig. S6). Upon association of the two subunits and formation of the initial bridges, the subunit interface is protein free. The association between the LSU and the SSU, mediated by L-shaped tRNA molecules occurs at the end of this phase.

Phase 4.Origin and development of the kernels of the major SSU domains. aes 6-10 are acquired. aes 6, 8, 10 are consequently added to aes 2 to form an initial portion of Domain 3'M. These segments position and stabilize tRNA at the P site of the decoding center. aes 7 and 9 are added to aes 4 to initiate Domain 5'. aes 7 assists in positioning tRNA in the A site. Altogether, aes 6, 7, 8, and 10 embrace two tRNA molecules, suggesting that these segments were added shortly after association of the two subunits to enhance the stability of the associated complex (Figs. S7 and S8). Portions of Domain C (aes 5a) and Domain 3'm (aes 1a) elaborate the SSU interface and form additional bridges with the LSU.

Phase 5. Acquisition of the Major SSU Domains and Development of Translocation. aes 11-16 are acquired. Domain 3'M is extended by aes 11, 14, 16 (Fig. S9). aes 11 cements the A- and P-site tRNAs, while aes 14 and 16 provide an additional integrity to the Head (Domain 3'M). aes 12, 13, 15 extend Domain 5', providing additional stability. aes 7 is decorated by extensions aes 7a and 7b that form a functionel pseudoknot, which is an essential component of the modern energy-driven translational machine (Fig. S10), assisting translocation (16, 17), decoding (18, 19), and proofreading (20).

Phase 6. SSU maturation. Accretion of aes 17-27 completes the rRNA component of the SSU common core (Fig. S11). Domain 3'M is finalized upon accretion of aes 22, 23, 25, 27, and 14a. These ancestral expansion segments form a layer of rRNA distal to the intersubunit interface, providing binding sites for the globular domains of numerous SSU ribosomal proteins (rProteins). Domain 5' is completed by addition of aes 17, 18, 19, and 26. While aes 17 forms additional intersubunit bridges, aes 18, 19, and 26 contribute to the structural integrity and serve as binding sites for SSU proteins. Domain C is completed by addition of aes 20, 21, 24, and 20a. aes 20 embraces nearly half of the SSU, integrating Domains 3'M and 5'. The tip of aes 20, together with its extension aes 20a, form additional intersubunit bridges. Domain 3'm undergoes a final extension (aes 1b) to elaborate the intersubunit interface. All these segments serve as binding sites for SSU proteins. Evolution of Phases 4-6 is summarized in Fig. S12.

Intersubunit Bridges

In the Accretion Model, at the end of Phase 3, intersubunit bridge B3 was established by two A-minor interactions. Universally conserved adenines A1418 and A1483 of aes 1 reach into the minor groove of the AES 15, and interact with paired bases, forming a network of hydrogen bonds (21). The core of the SSU (Phases 1 of the SSU) appears dependent on a more peripheral region of the LSU (Phase 3 of the LSU). This dependence suggests that Phase 1 and 2 of the LSU predate association with elements of the SSU established in Phase 1. Bridge B3 is the most robust link between the subunits, forming a pivot point during translocation in the mature ribosome (21). Additional bridges formed in this phase (B2b and B2c) involve phosphate-Mg^{2+}-phosphate linkages and appear opportunistic (Fig. S6). In general, the bridges formed upon initial association of the two subunits do not involve proteins. The intersubunit bridges are summarized in Fig. S3.

Phase 4. Mg^{2+}-mediated bridge B2b is elaborated by extension aes 5a, interacting with AES 15 of the LSU. As aes 1 is extended by acquisition of aes 1a, Mg^{2+}-mediated bridge B5 is

initially formed between aes 1a and AES 15, and B6 is initially formed between aes1a and AES11.

Phase 5. Accretion of AES 29, 10a, and 37 by the LSU forms bridges B2a, B4, and elaborates B6, by donating adenines and establishing multiple A-minor interactions with aes1 and 1a of the SSU. The adenines in the LSU bridgeheads (A1912, A1913, and A1919 in B2a; A716 in B4, and A1689, A1700, and A1701 in B6) are universally conserved (1, 2). Directionality of the bridgeheads suggests that the accreted LSU segments have been added to the LSU after its association with the SSU and extension of aes 1 with aes 1a to strengthen the intersubutit interactions and to optimize the interface. Additionally, bridges B1a and B1b form, and B5 are elaborated. Bridges B1a, B1b, B4, B5, and B6 are stabilized by interactions with early proteins.

Phase 6. The interface between the two subunits is further strengthened by bridges B7a, B7b, B8, and elaboration of B6. All of these bridges except B7a involve ribosomal proteins.

Defining Trunks and Branches

During each expansion, new branch rRNA is inserted into old trunk rRNA. Stacking is continuous within the trunk and generally within the branch, but not between the trunk and the branch. The characterization of trunk versus branch is in some cases dependent on external indicators of relative age as described above.

Special comments on some noncanonical aes

Most aes are defined by the insertion fingerprints as described above. In general, this method allows easy identification of insertions but not elongations. The Accretion Model treats the addition of each expansion as an instantaneous event, even if it contains multiple helices. For example, aes 20 is composed of helices 21, 22, and 23. However, in a few cases, if a helix bends sharply, it has been divided into two aes: the main segment and an appendage, denoted by suffix "a" or "b" (e.g. aes 14 and 14a). The appendage is assigned the same number as the associated

helix because the criterion of the helix continuity between the main segment and appendage is met. The appendages may appear after the founding aes, if this is supported by structural constraints.

aes 1 (helix 44) is extended by aes 1a and 1b. aes 1, 1a, and 1b form a long linear helix with continuous stacking. However, the elongations of aes 1 by aes 1a and aes 1b are considered independent events because of locations and types of intersubunit bridges and irregularities (bulges) in the helix. This partitioning is also supported by the length of helix 44 in known SSU structures compared to in the model of minimal ribosome, in which h44 is truncated between nucleotides 1419-1481 (14).

aes 4a, which is inserted between aes 4 and 5 in Phase 3, is a short distorted helix, which is inserted into the middle of aes 4 (between helices 19 and 25) and connects it to aes 5 (helices 20 and a portion of helix 24). aes 4a does not merit its own number as it is highly distorted from helical geometry and lacks base-paring.

Helix 24 was partitioned into aes 5 and 5a based on analysis of interactions with the LSU. Helix 24 sharply bends, when it reaches the intersubunit interface, and then extends along it. A portion of helix 24 (within aes 5) is perpendicular to the interface and is engaged in one of the earliest bridges (B2c), formed during the initial association of the subunits. The extension (aes 5a) appears to be a later addition as it transverses along the intersubunit surface and elaborates another bridge (B2b).

Most of helix 18 belongs to aes 7. However, helix 18 contains two small fragments (aes 7a and 7b) that protrude from the main stem. These fragments interact with each other to form a pseudoknot, which is known to play an important role during decoding and translocation of mRNA. Due to their functionality, acquisition of aes 7a and 7b is assumed in Phase 5.

Helix 41 is partitioned by aes 14 and aes 14a as indicated by a sharp bend in aes 14a, which interacts with aes 16. aes 16 was acquired after aes 14, based on structural arguments (as aes 16 branches out from a junction between aes 8 and aes 14). aes 14a forms an A-minor

interaction with aes 16 donating universally conserved A1269. Therefore, aes 14a is younger than aes 16, and was partitioned into Phase 6.

Most of helices 21, 22, and 23 were grouped into a single aes (aes 20), which embraces and integrates the SSU in Phase 6, due to continuity of the backbone and stacking interactions. However, a tip of helix 23 (aes 20a) sharply bends and forms A-minor interactions with aes 5a by donating near-universally conserved A694, A695, and A696.

References

1. Petrov AS, Bernier CR, Gulen B, Waterbury CC, Hershkovitz E, Hsiao C, Harvey SC, Hud NV, Fox GE, Wartell RM, & Williams LD (2014) Secondary Structures of Rrnas from All Three Domains of Life. *PLoS One* 9(2):e88222.
2. Bernier C, Petrov AS, Waterbury C, Jett J, Li F, Freil LE, Xiong b, Wang L, Le A, Milhouse BL, Hershkovitz E, Grover M, Xue Y, Hsiao C, Bowman JC, Harvey SC, Wartel JZ, & Williams LD (2014) Ribovision: Visualization and Analysis of Ribosomes. *Faraday Discuss.* 169(1):195-207.
3. Petrov AS, Bernier CR, Hershkovitz E, Xue Y, Waterbury CC, Grover MA, C. HS, Hud NV, Wartell RM, & Williams LD (2013) Secondary Structure and Domain Architecture of the 23S rRNA. *Nucleic Acids Res.* 41(15):7522-7535.
4. Dunkle JA, Wang LY, Feldman MB, Pulk A, Chen VB, Kapral GJ, Noeske J, Richardson JS, Blanchard SC, & Cate JHD (2011) Structures of the Bacterial Ribosome in Classical and Hybrid States of tRNA Binding. *Science* 332(6032):981-984.
5. Selmer M, Dunham CM, Murphy FV, Weixlbaumer A, Petry S, Kelley AC, Weir JR, & Ramakrishnan V (2006) Structure of the 70S Ribosome Complexed with mRNA and tRNA. *Science* 313(5795):1935-1942.
6. Anger AM, Armache JP, Berninghausen O, Habeck M, Subklewe M, Wilson DN, & Beckmann R (2013) Structures of the Human and Drosophila 80S Ribosome. *Nature* 497(7447):80-85.
7. Schrodinger LLC (2014) The Pymol Molecular Graphics System, Version 1.7.4.
8. Pruesse E, Peplies J, & Glockner FO (2012) Sina: Accurate High-Throughput Multiple Sequence Alignment of Ribosomal RNA Genes. *Bioinformatics* 28(14):1823-1829.
9. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, & Glöckner FO (2013) The Silva Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools. *Nucleic Acids Res.* 41(D1):D590-D596.
10. Shannon CE (1948) A Mathematical Theory of Communication. *At&T Tech J* 27(4):623-656.
11. Sander C & Schneider R (1991) Database of Homology-Derived Protein Structures and the Structural Meaning of Sequence Alignment. *Proteins* 9(1):56-68.
12. Ramazzotti M, Degl'Innocenti D, Manao G, & Ramponi G (2004) Entropy Calculator: Getting the Best from Your Multiple Protein Alignments. *Ital. J. Biochem.* 53(1):16-22.

13. Petrov AS, Bernier CR, Hsiao C, Norris AM, Kovacs NA, Waterbury CC, Stepanov VG, Harvey SC, Fox GE, Wartell RM, Hud NV, & Williams LD (2014) Evolution of the Ribosome at Atomic Resolution. *Proc. Natl. Acad. Sci. U.S.A.* 111(28):10251-10256.

14. Mears JA, Cannone JJ, Stagg SM, Gutell RR, Agrawal RK, & Harvey SC (2002) Modeling a Minimal Ribosome Based on Comparative Sequence Analysis. *J. Mol. Biol.* 321(2):215-234.

15. Bokov K & Steinberg SV (2009) A Hierarchical Model for Evolution of 23S Ribosomal RNA. *Nature* 457(7232):977-980.

16. Rodnina MV, Savelsbergh A, & Wintermeyer W (1999) Dynamics of Translation on the Ribosome: Molecular Mechanics of Translocation. *FEMS Microbiol. Rev.* 23(3):317-333.

17. Noller HF (2004) The Driving Force for Molecular Evolution of Translation. *RNA* 10(12):1833-1837.

18. Ogle JM, Murphy Iv FV, Tarry MJ, & Ramakrishnan V (2002) Selection of tRNA by the Ribosome Requires a Transition from an Open to a Closed Form. *Cell* 111(5):721-732.

19. Ogle JM, Brodersen DE, Clemons WM, Tarry MJ, Carter AP, & Ramakrishnan V (2001) Recognition of Cognate Transfer RNA by the 30S Ribosomal Subunit. *Science* 292(5518):897-902.

20. Jenner L, Demeshkina N, Yusupova G, & Yusupov M (2010) Structural Rearrangements of the Ribosome at the tRNA Proofreading Step. *Nat. Struct. Mol. Biol.* 17(99):1072-1078.

21. Schuwirth BS, Borovinskaya MA, Hau CW, Zhang W, Vila-Sanjurjo A, Holton JM, & Cate JH (2005) Structures of the Bacterial Ribosome at 3.5 Å Resolution. *Science* 310(5749):827-834.

22. Paci M & Fox GE (2015) Major Centers of Motion in the Large Ribosomal RNAs. *Nucleic Acids Res.* 43(9):4640-4649.

23. Pulk A & Cate JH (2013) Control of Ribosomal Subunit Rotation by Elongation Factor G. *Science* 340(6140):1235970.

24. Gerbi SA (1996) Expansion Segments: Regions of Variable Size That Interrupt the Universal Core Secondary Structure of Ribosomal RNA. *Ribosomal RNA—Structure, Evolution, Processing, and Function in Protein Synthesis*, eds Zimmermann RA & Dahlberg AE (CRC Press, Boca Raton, FL), pp 71–87.
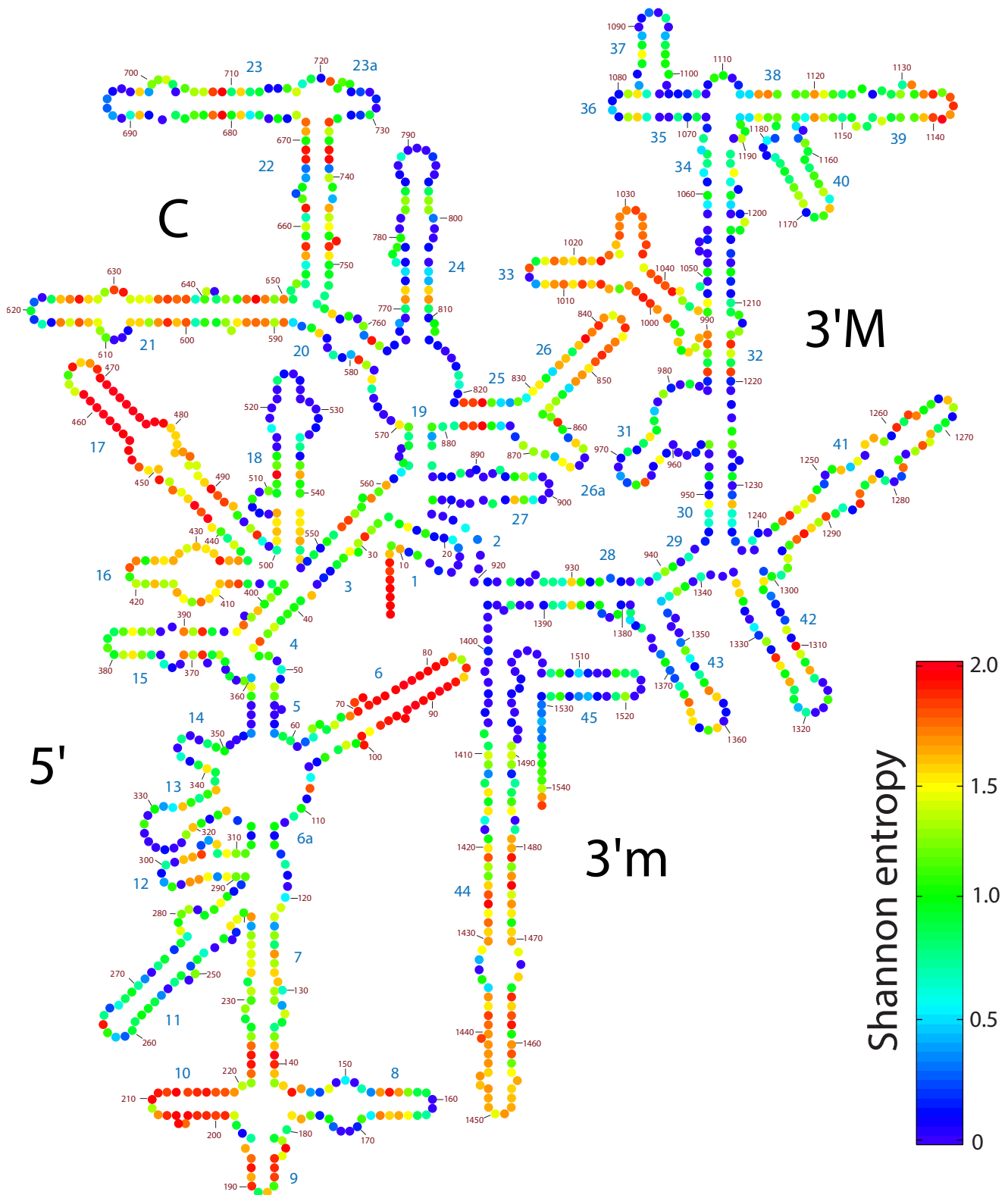
Figure S1. Nucleotide-level Shannon entropies mapped by color onto the secondary structure of the SSU rRNA of *E. coli*. The lowest value (0) indicates universal conservation over the entire tree of life (dark blue). The highest value (2) indicates uniform variation and / or high gap frequency (red). The SSU domains (3'M, 3'm, 5', and C) and helices are labeled. Nucleotides are numbered.

13
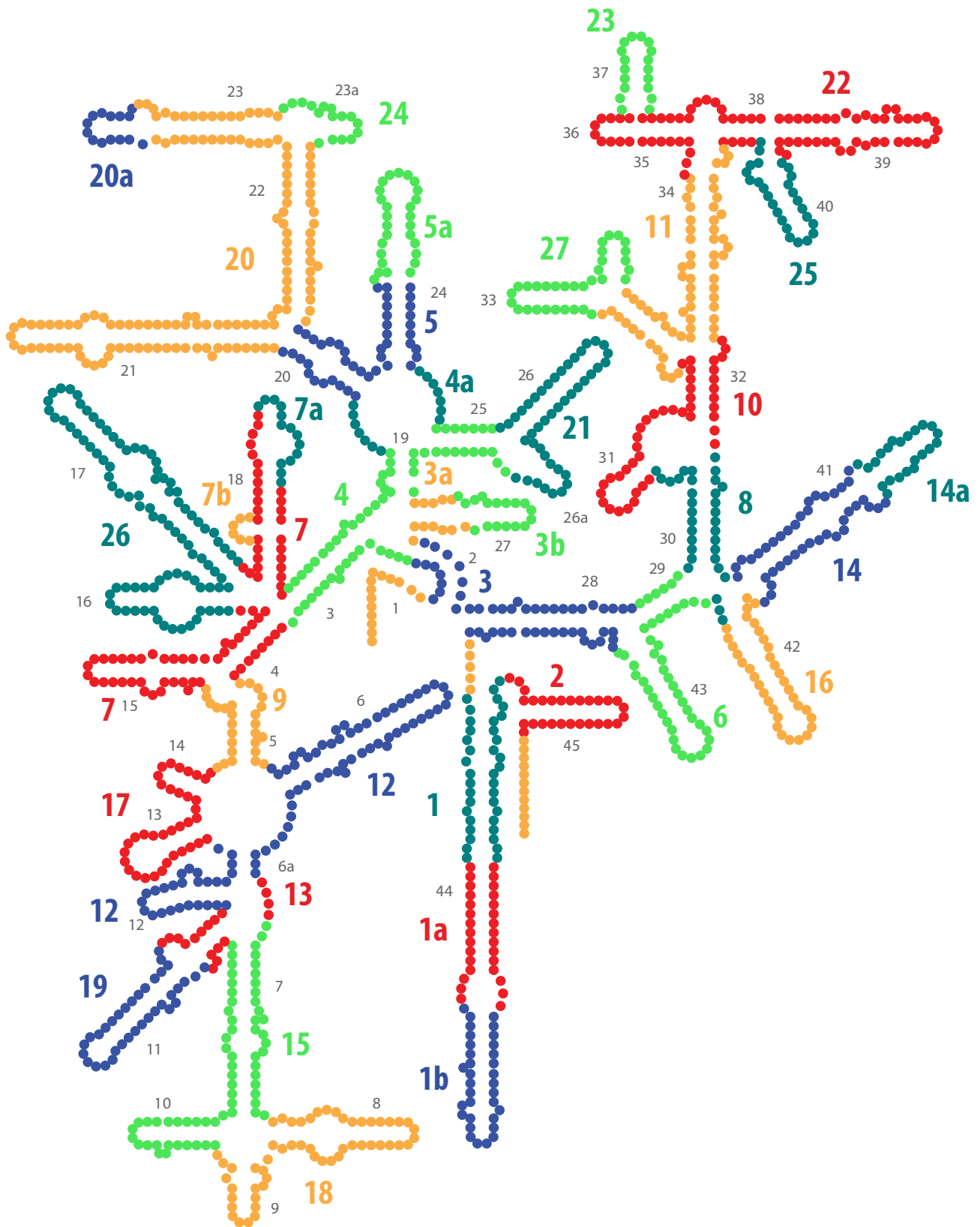
Figure S2. rRNA evolution mapped onto the SSU rRNA secondary structure of *E. coli*. The common core is built by stepwise addition of ancestral expansion segments (aes) at sites marked by insertion fingerprints. Each aes is individually colored and numbered. The aes colors are arbitrary, chosen to distinguish the expansions, such that no aes is the color of its neighbors. Helix numbers are gray.
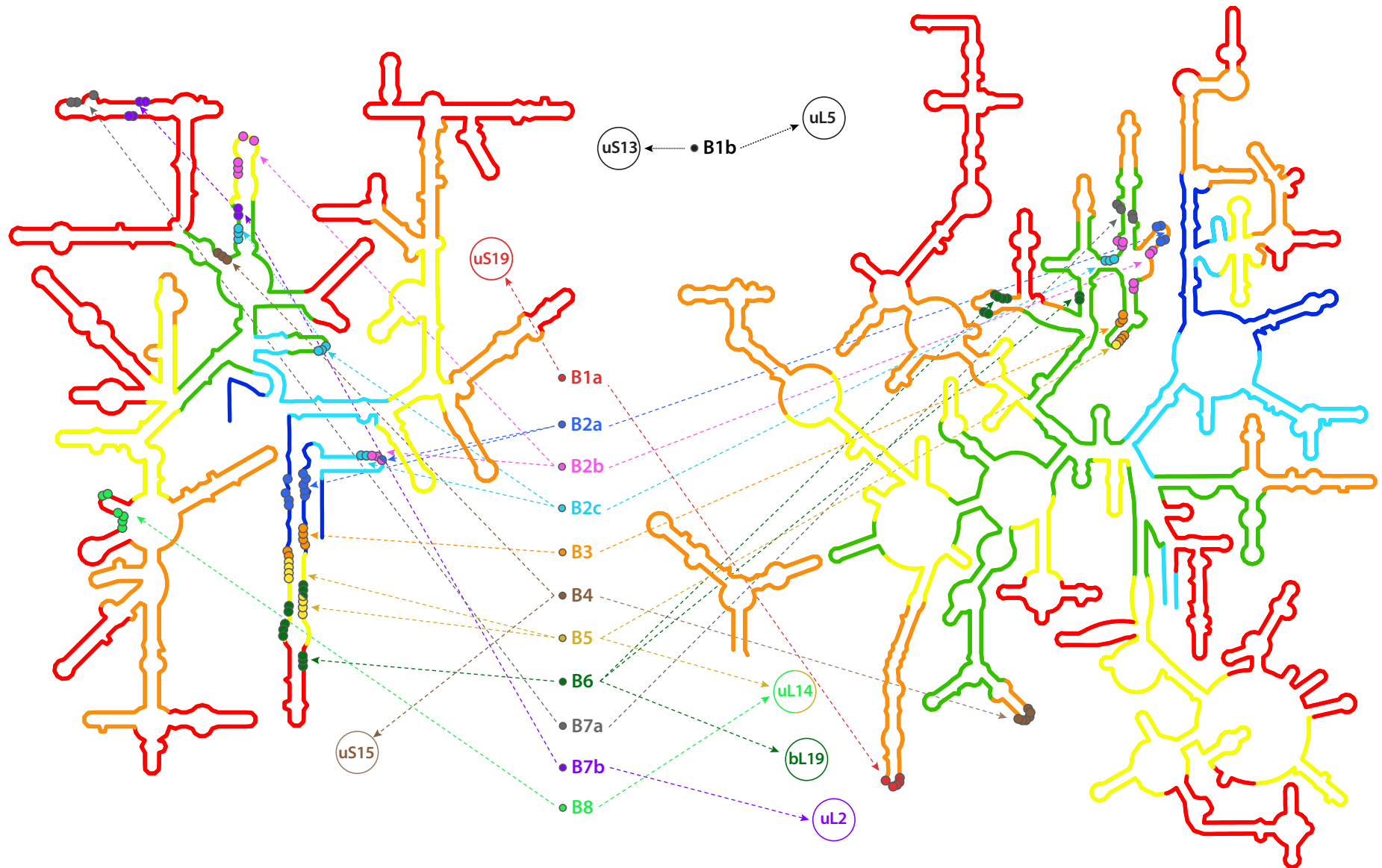
Figure S3. SSU-LSU intersubunit bridges (21) mapped onto the secondary structures of the SSU and LSU rRNAs of *E.coli*. The ribosome is built up in six phases, by stepwise addition of ancestral expansion segments (aes) at sites marked by insertion fingerprints. Phase 1 – dark blue; Phase 2 – light blue; Phase 3 – green; Phase 4 – yellow; Phase 5 – orange; and Phase 6 – red. Bridging nucleotides are indicated by solid circles. Bridging rProteins are represented by open circles.

A) SSU   B) LSU

Phase 1
Phase 2
Phase 3
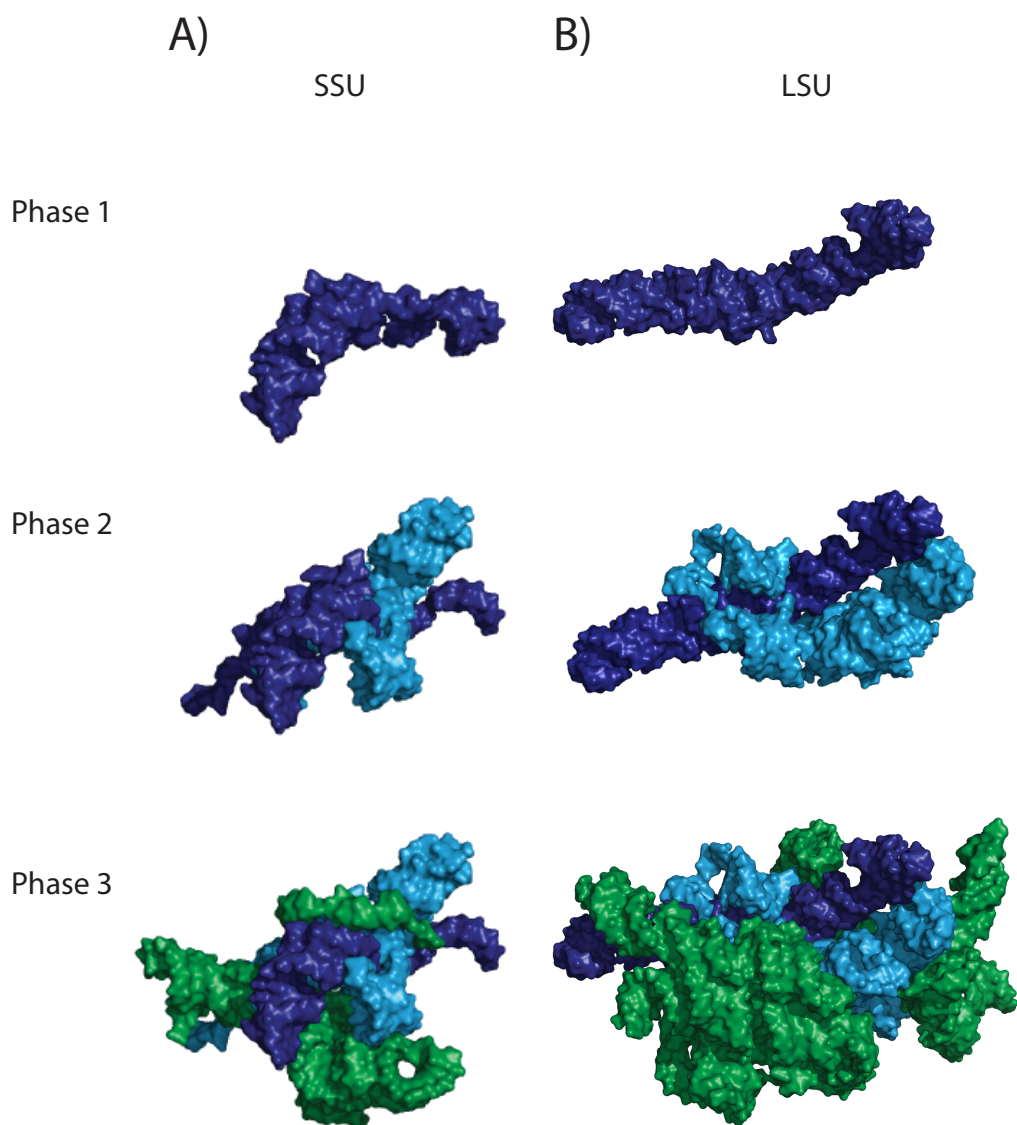
Figure S4. Early evolution of rRNA (Phases 1-3). The structures of ancient rRNAs, before association of the subunits, are inferred from the structure of the bacterial ribosome and the Accretion Model. Phase 1 of the SSU and LSU is dark blue, Phase 2 is light blue, and Phase 3 is green. A) SSU. B) LSU. Temporal relationships between changes in the LSU and SSU are undetermined during Phases 1-3.
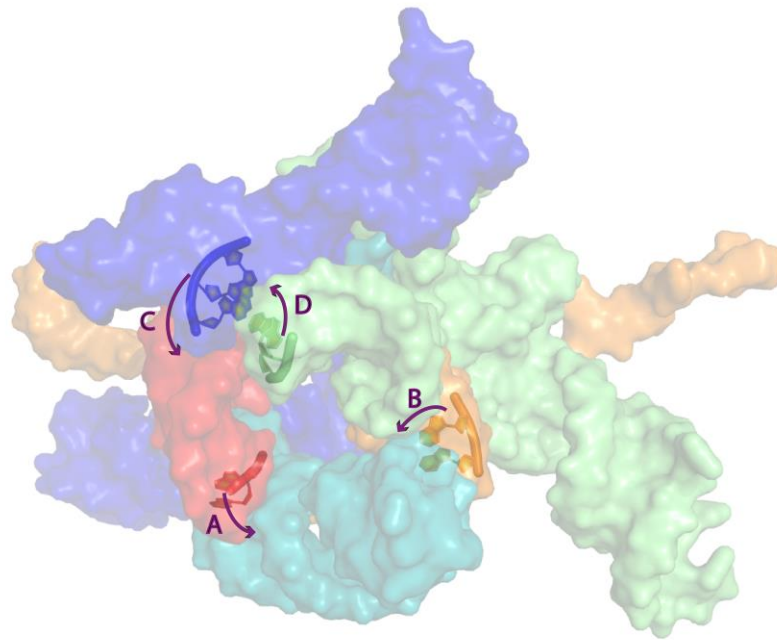
Figure S5. Universally conserved A-minor interactions form an interlocking core within aes 1-5 [viewed from the intersubunit interface]. A) A1518 and A1519 of aes 2 (red) bind to the minor groove of aes 1 (teal). B) A908 and A909 of aes 3a (orange) bind to the minor groove of aes 1 (teal). C) A766, A767, and A768 of aes 5 (blue) bind to the minor groove of aes 2 (red). D) A900 and A901 of aes 3b (light green) bind to the minor groove of aes 5 (blue). The A-minor interactions and their directionalities (from adenine to minor groove) are schematically shown by purple arrows. The interior of aes 1-5 is protein-free. Dirctionalities of the A-minor interactions suggest that aes 1 is the primordial segment from which the SSU evolved.
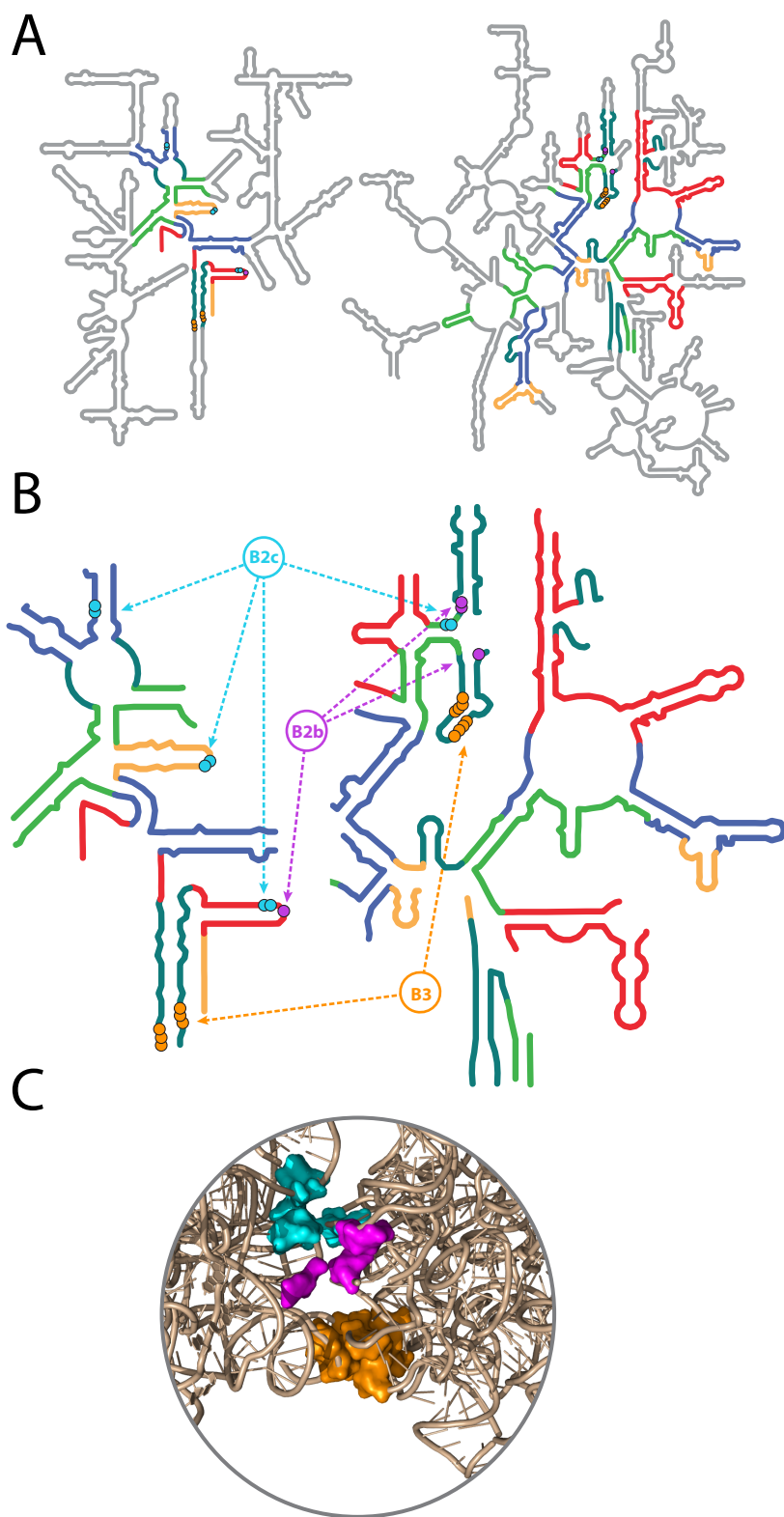
Figure S6. Bridging interactions between the initial assembly formed by the SSU (left) and the LSU (right). A) Full *E. coli* rRNA secondary structures, illustrating development by the subunit association point (Phase 3). Phases 1-3 aes are colored as in Figs. 2A and 2B. Later phase rRNA are in gray. B) Partial secondary structure diagram highlighting the bridging interactions, which are indicated by dashed lines. Bridgeheads are indicted by closed circles. C) Bridge interactions mapped onto the 3D structure of associated proto-SSU and proto-LSU. Bridge B2b is magenta, B2c is cyan and B3 is yellow.
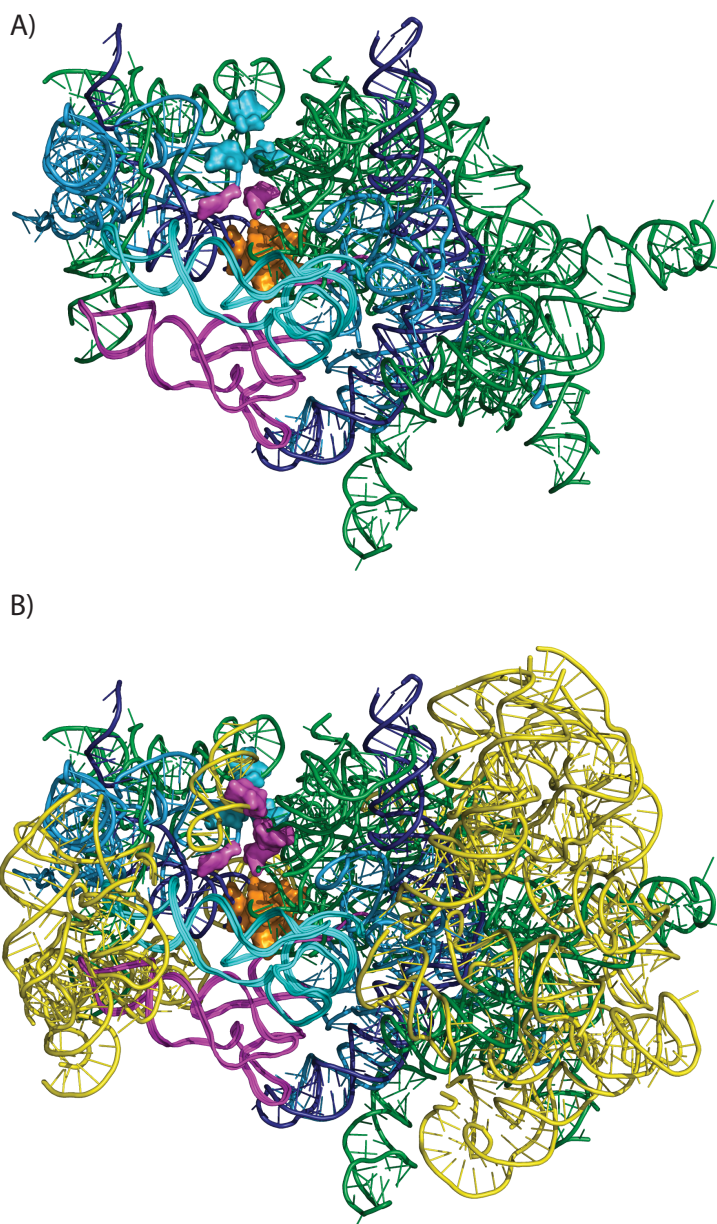
Figure S7. Evolution of the ribosome during Phase 4. The Phase 4 LSU (yellow) expands in the region distal to the SSU interface. The P-site tRNA is in cyan, and the A-site tRNA is in magenta. A) The ribosome at the beginning of Phase 4, containing the LSU rRNA (AES 1-15) and the SSU rRNA (aes 1-5). B) The ribosome at the end of Phase 4, containing the LSU (AES 1-29) and the SSU rRNA (aes 1-11). LSU and SSU rRNAs are colored by phase (as in Fig 2). The initial intersubunit bridges (B2b, B2c, and B3) are in space-fill representation and are colored as in Fig. S3.

Figure S8. Phase 4 SSU rRNA (yellow) contains arms that position the A- and P-site tRNAs (magenta and cyan). The arms emerge after association of the subunits. These arms have no apparent function in the absence of tRNA, so it appears that they arose to stabilize tRNAs within the SSU-mRNA-tRNA-LSU assembly. The SSU rRNA is colored by phase as in Fig. 2.
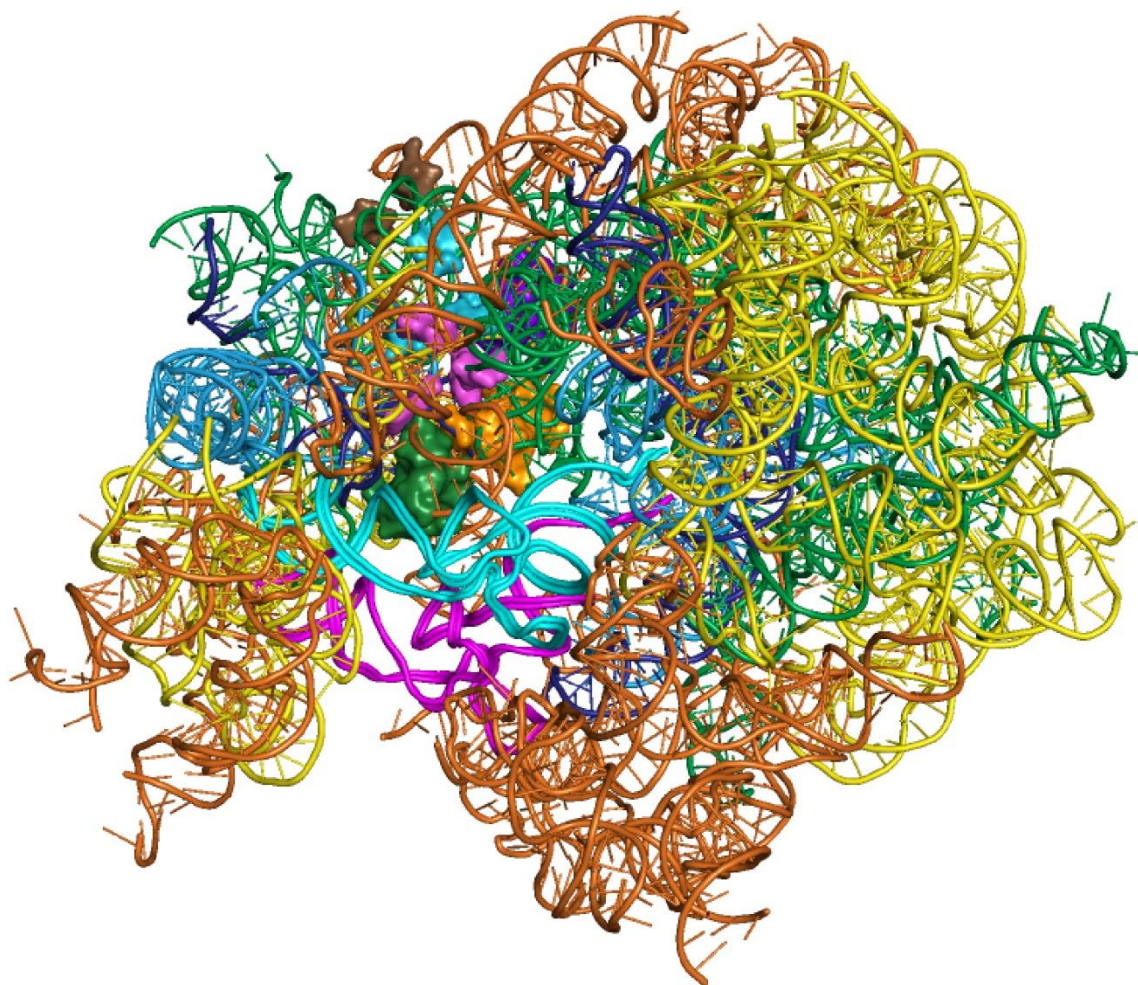
Figure S9. The ribosome at the end of Phase 5: LSU rRNA (AES 1-39 of 23S rRNA and 5S rRNA) and SSU rRNA (aes 1-17) are linked by two tRNAs. The LSU and SSU rRNAs are colored by phase as in Fig. 2. Intersubunit bridges are in surface representation, with colors as in Fig. S3. The A-site tRNA is magenta and the P-site tRNA is cyan.
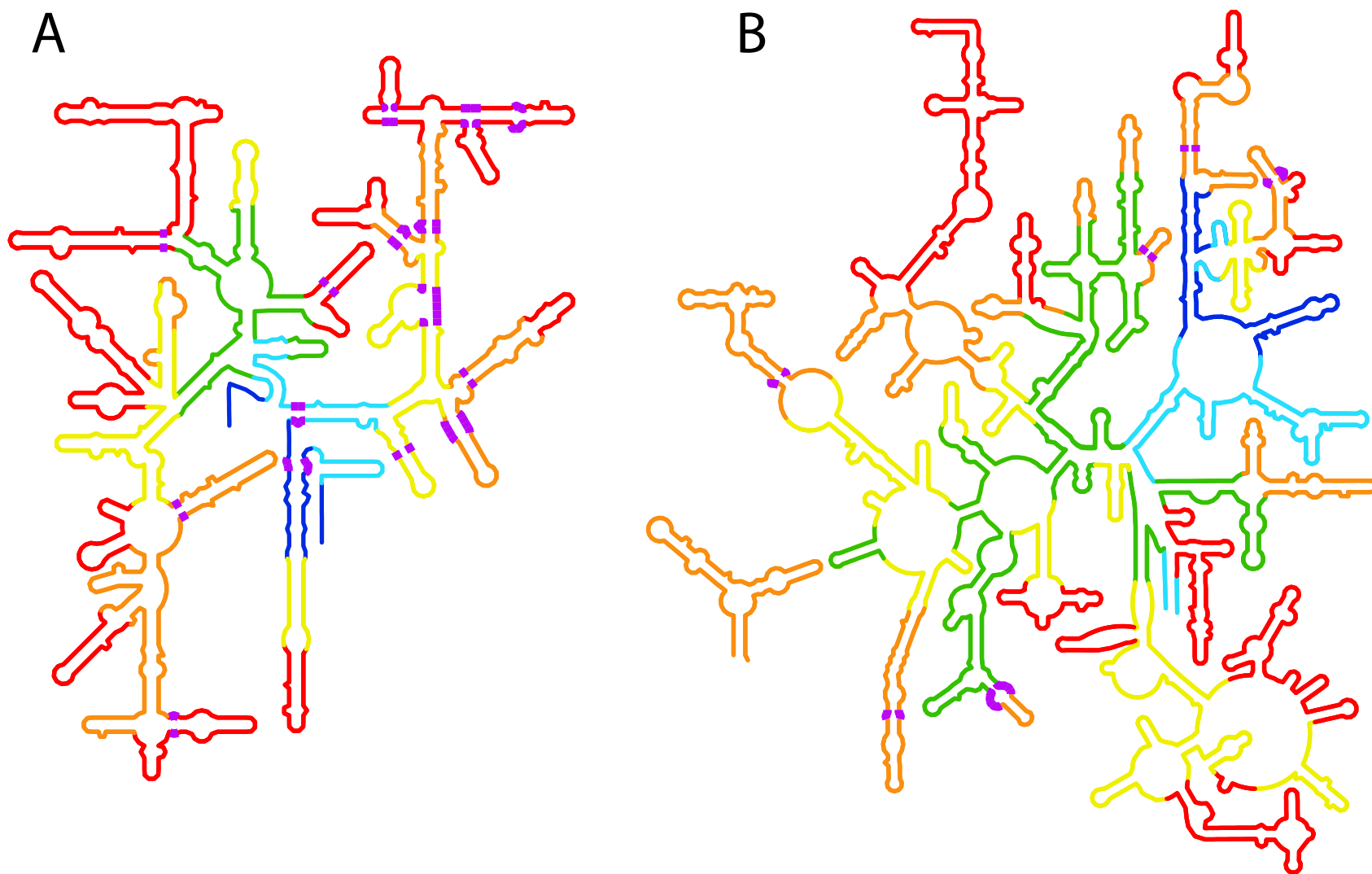
Figure S10. Locations of pivots, identified by Fox (22), in A) SSU, and B) LSU rRNAs of *E. coli* are indicated in purple on secondary structure diagrams. The pivot points were determined by a comparison of crystal structures 4V9D (4) and 4V9O (23). The SSU and LSU rRNAs are colored by phase as in Fig. 2.
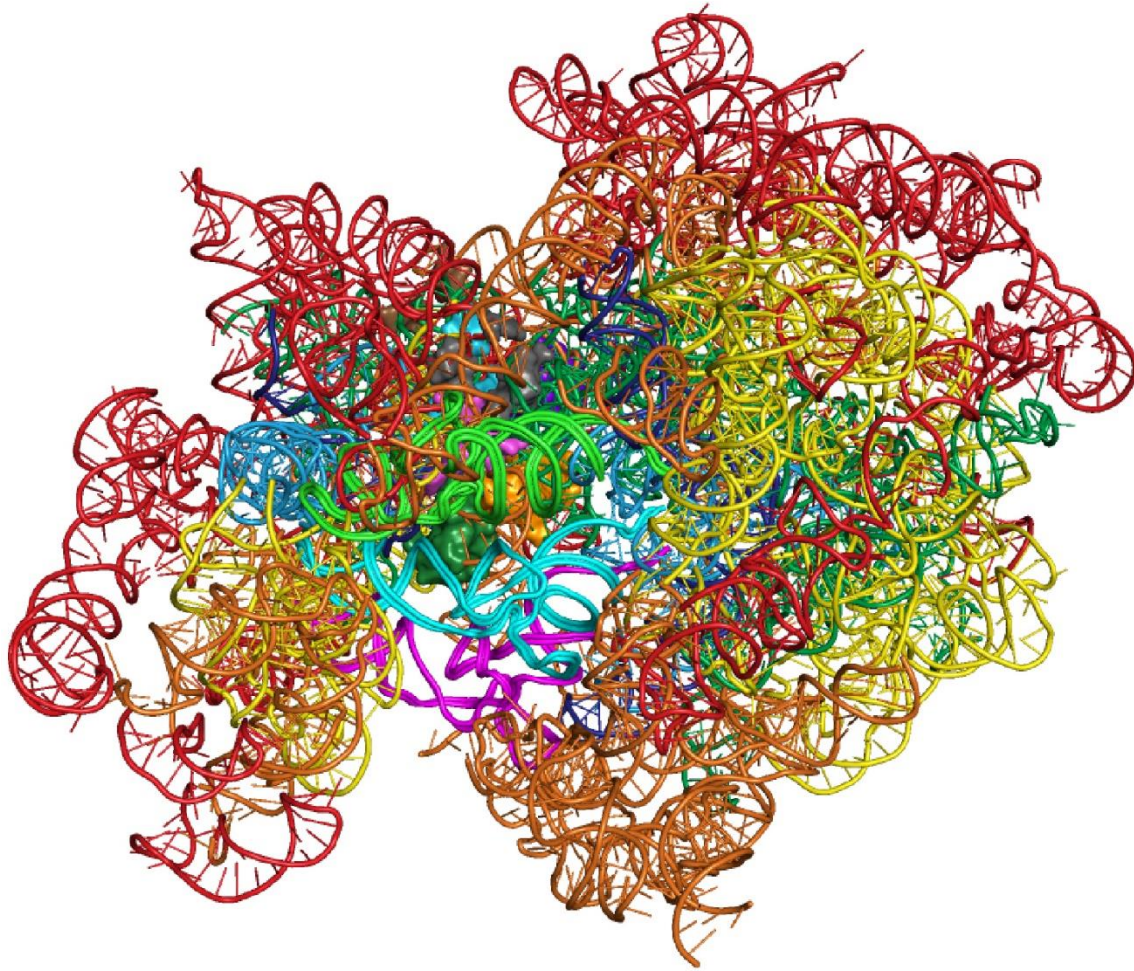
Figure S11. Ribosome particle at the end of Phase 6: LSU rRNA (AES 1-60 of 23S rRNA and 5S rRNA) and SSU rRNA (aes 1-27) are bridged by the P-site tRNA. LSU and SSU rRNAs are color coded according to phase definitions, intersubunit bridges are in surface representation shaded in the colors defined in Fig. S3, and the mRNA is in dark grey. The A-site tRNA is magenta, the P-site tRNA is cyan, and the E-site tRNA is green.
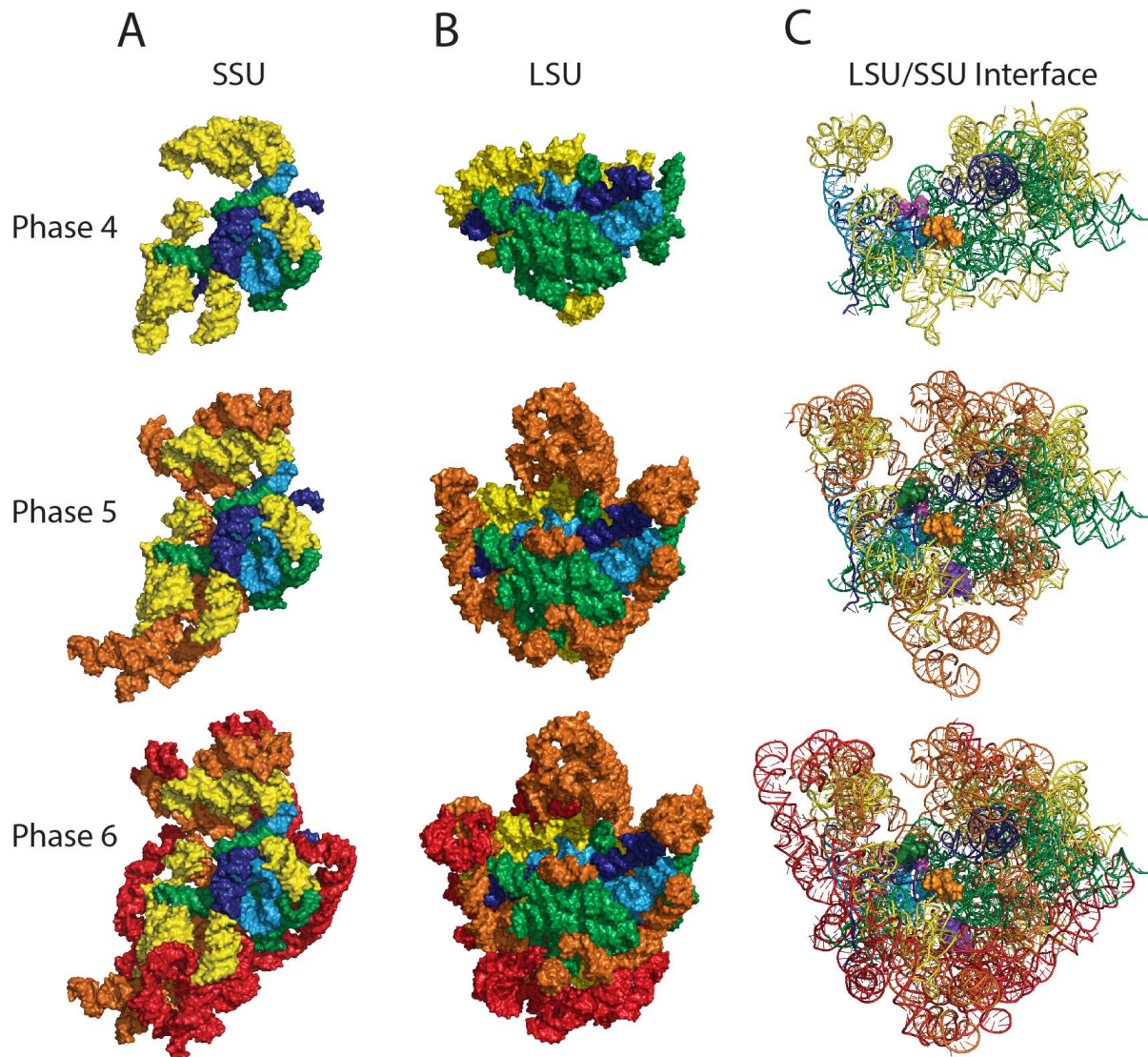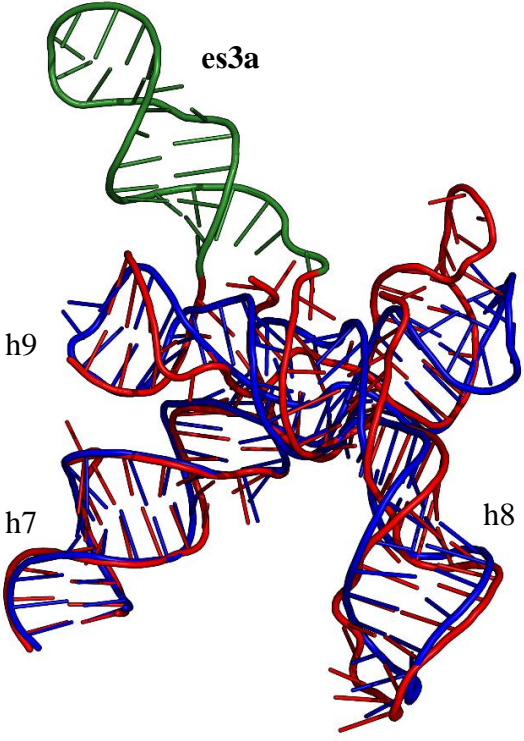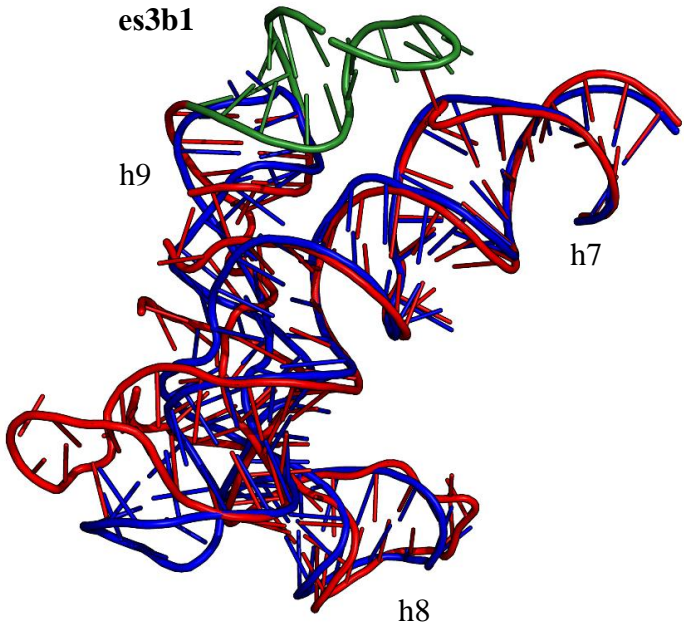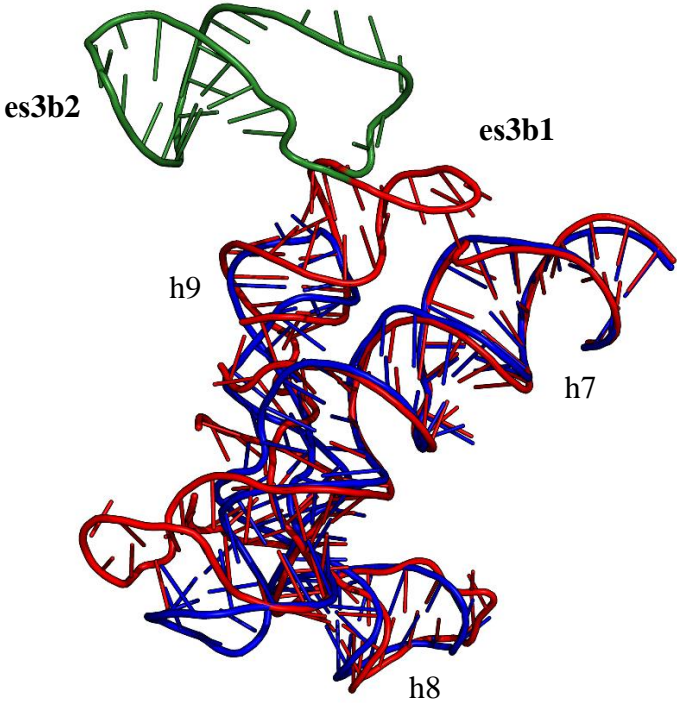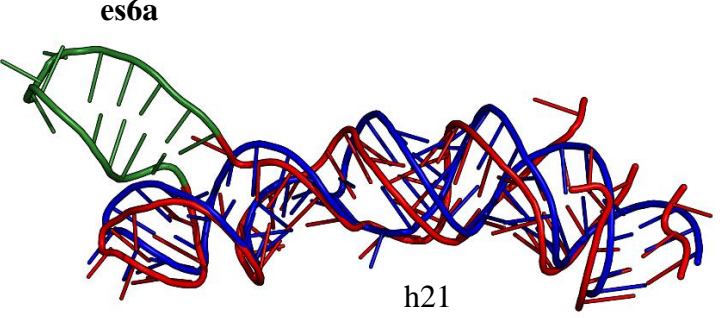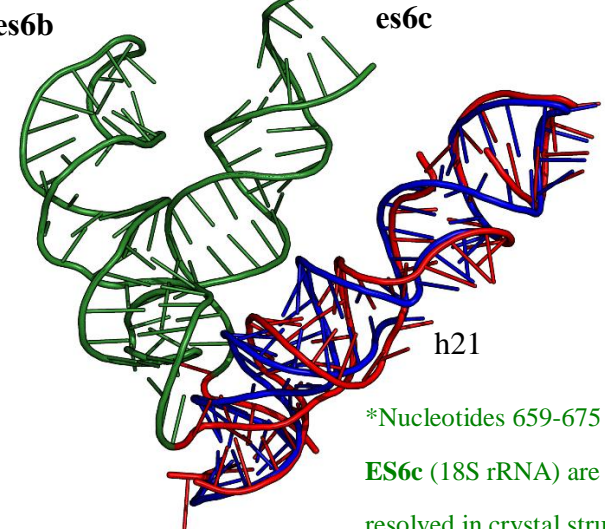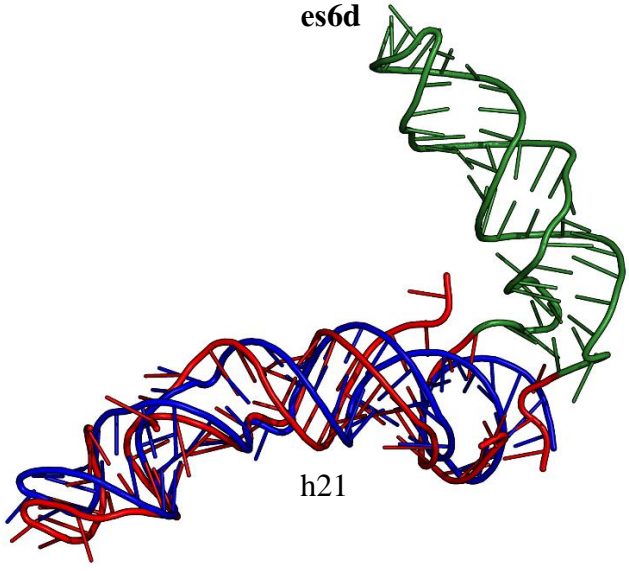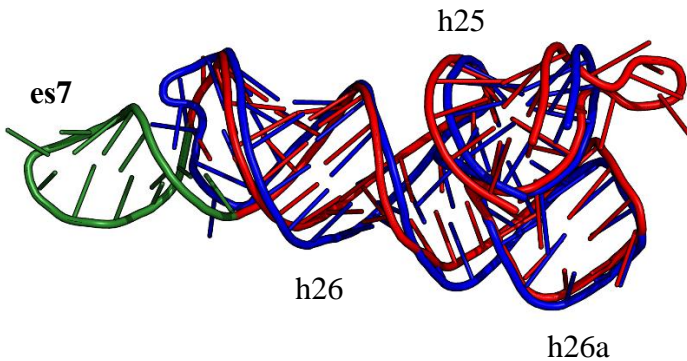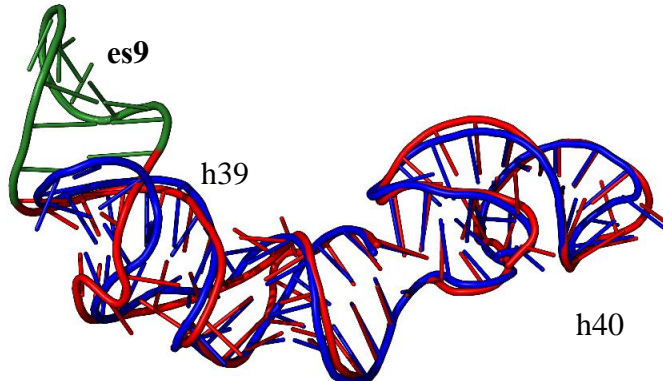
Figure S12. Evolution of the ribosome during Phases 4-6, after association of the SSU and LSU. Phases are colored as in Fig. S3. A) SSU rRNA evolution in surface representation [interface view]. B) LSU rRNA evolution in surface representation [interface view]. C) Cartoon representation of both subunits [side view] with initial intersubunit bridges (B2b, B2c, and B3) in space-fill representation, and colored as in Fig. S3.

Table S1. SSU eukaryotic expansion segments (ES) of *S. cerevisiae* accreted onto the common core (modeled by *E. coli*). *E. coli* trunk helices are blue. *S. cerevisiae* trunk helices are red and the expansion segments (branch helices) are green. Trunk nucleotide numbers of *E. coli* are shown in regular font; those of *S. cerevisiae* are in italics; the expansion segments of *S. cerevisiae* are in bold. The expansion segments are numbered by the scheme of Gerbi (24).

ES	Helix	E. coli (trunk)// *S. cerevisiae (trunk)/* **Expansion Segment**	3-Dimensional View
3a	9	16S:(123-238)// *18S:(107-128); 18S:(140-177); 18S:(204-209); 18S:(256-307)/* **18S:(178-203)**	

3b1	9	16S:(123-238)// *18S:(107-128);* *18S:(140-177);* *18S:(204-209);* *18S:(256-307)/* **18S:(210-214);** **18S:(243-255)**	
3b2	9	16S:(123-238)// *18S:(107-128);* *18S:(140-177);* *18S:(204-214);* *18S:(243-307)/* **18S:(215-242)**	

6a	21	16S:(588-651)// *18S:(635-639);* *18S:(742-773);* *18S:(788-810);* *18S:(859-861)/* **18S:(774-787)**	
6b, 6c	21	16S:(588-651)// *18S:(635-639);* *18S:(742-773);* *18S:(788-810);* *18S:(859-861)/* **18S:(640-741)**	

es6a

h21

es6b   es6c

h21

*Nucleotides 659-675 of **ES6c** (18S rRNA) are not resolved in crystal structure

6d	21	16S:(588-651)// *18S:(635-639);* *18S:(742-773);* *18S:(788-810);* *18S:(859-861)/* **18S:(811-858)**	
7	25, 26, 26a	16S:(821-880)// *18S:(1032-1053);* *18S:(1067-1104)/* **18S:(1054-1066)**	
9	39, 40	16S:(1117-1183)// *18S:(1337-1355);* *18S:(1368-1414)/* **18S:(1356-1367)**	
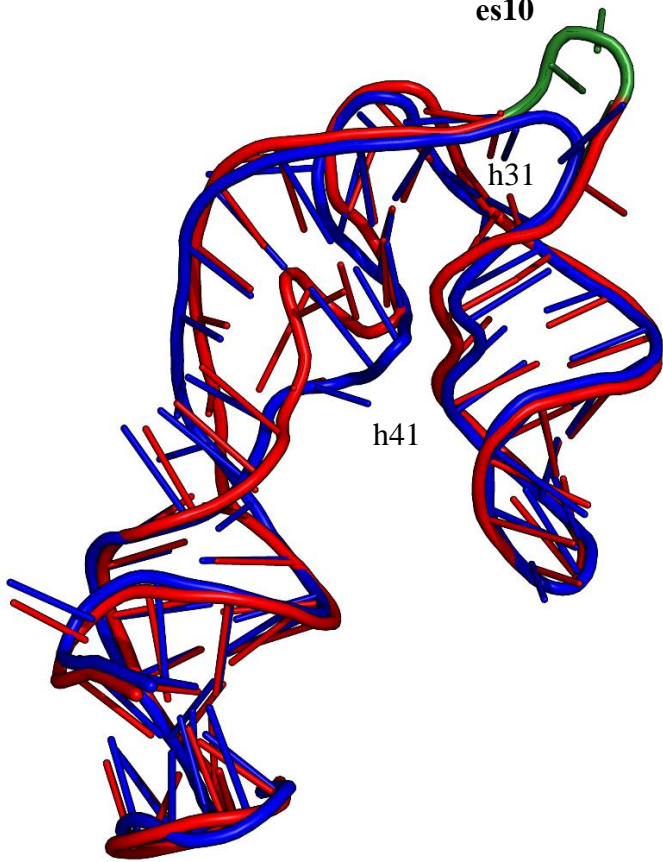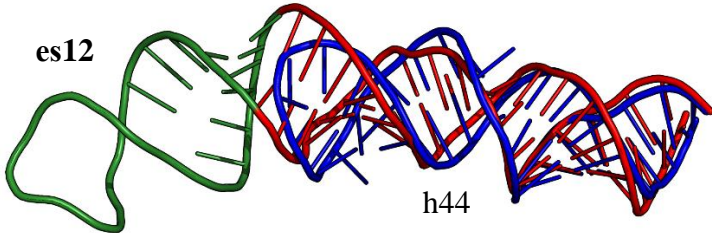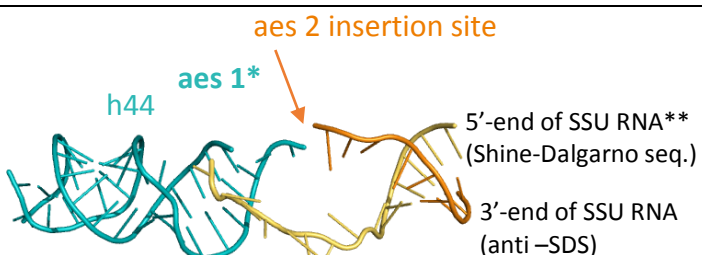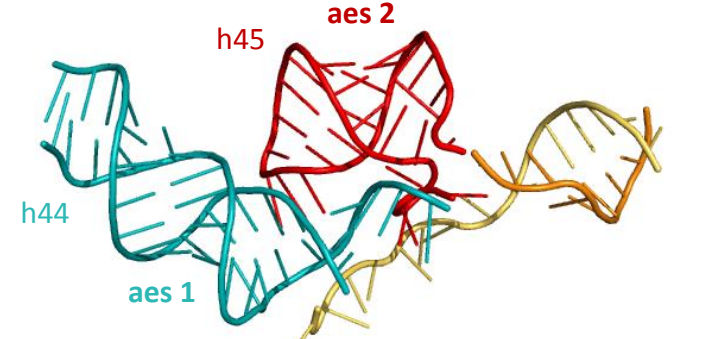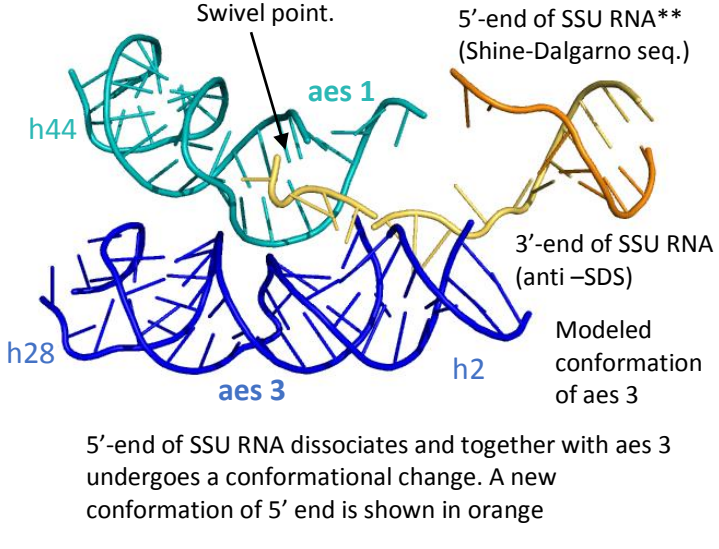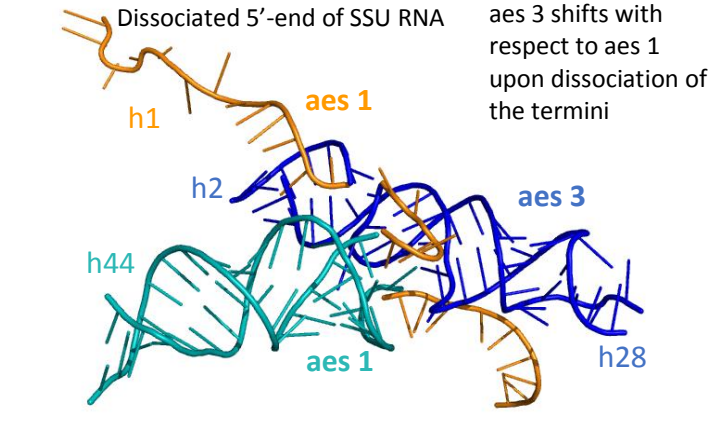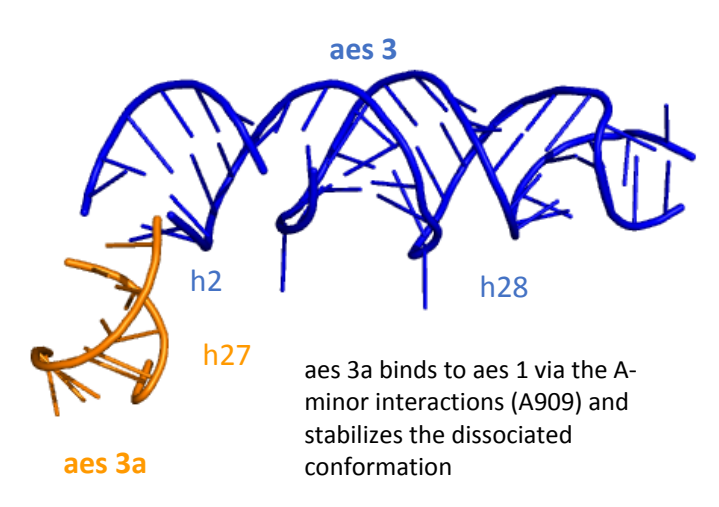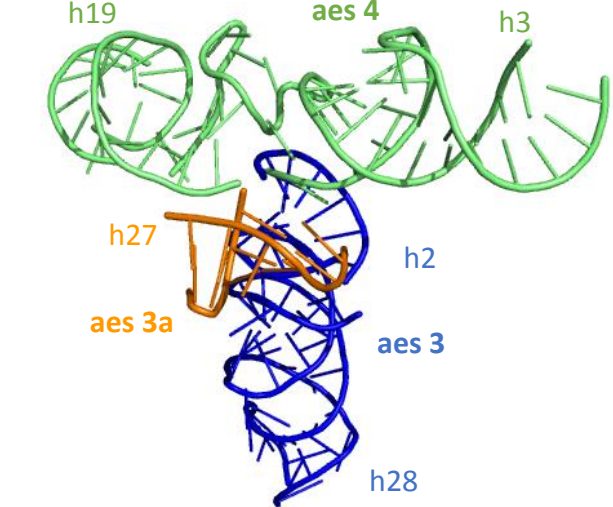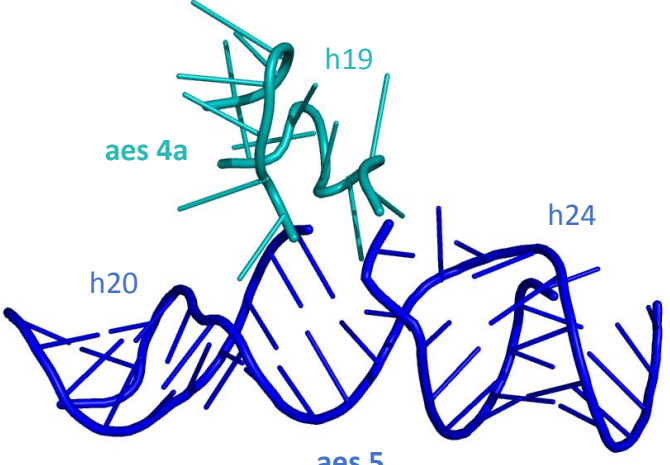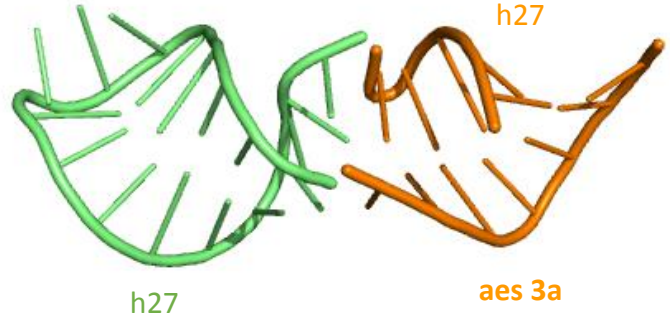
10	41	16S:(1238-1301)// *18S:(1471-1489);* *18S:(1493-1538)/* **18S:(1490-1492)**	
12	44	16S:(1430-1470)// *18S:(1668-1688);* *18S:(1713-1733)/* **18S:(1689-1712)**	
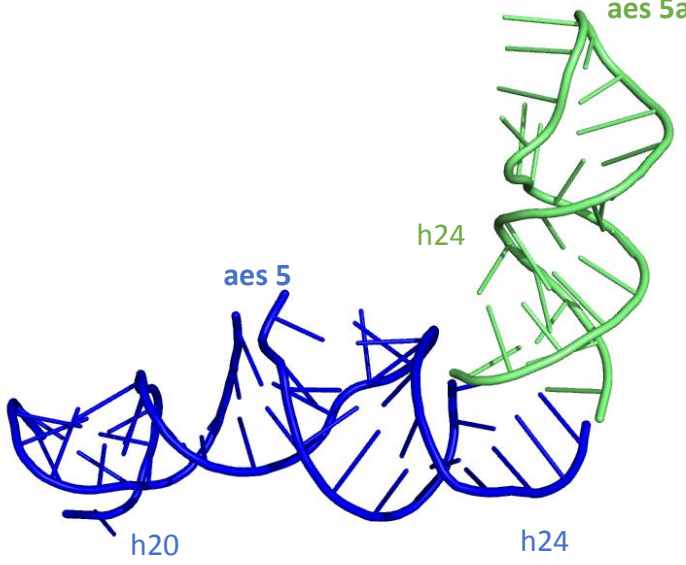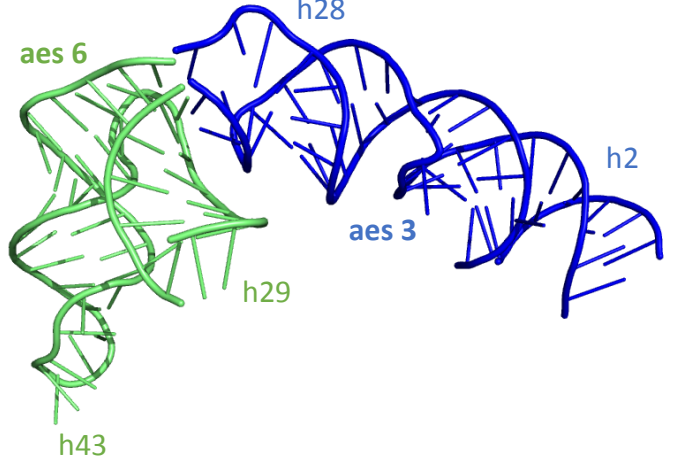
Table S2. Ancestral expansion segments (aes) of a bacterial (*E. coli*) SSU rRNA. aes are colored according to the scheme in Figure S2. Original helix numbers and nucleotide numbers are also provided.
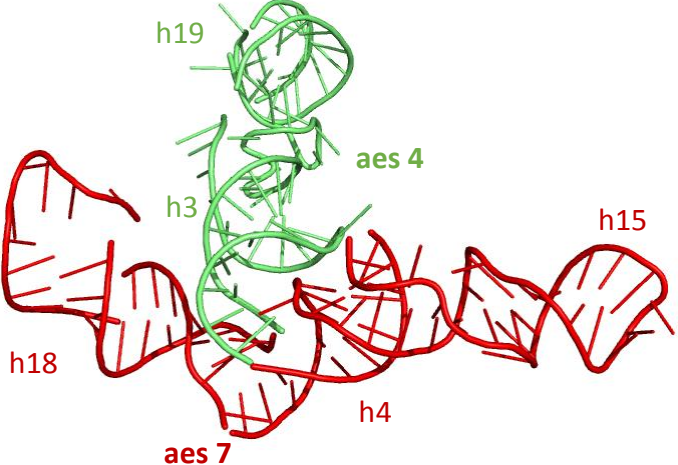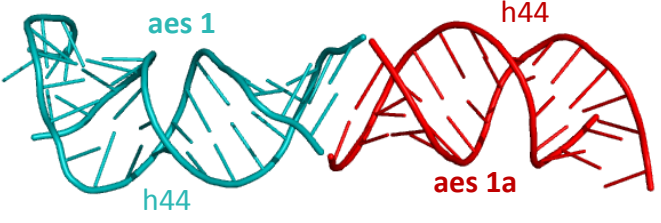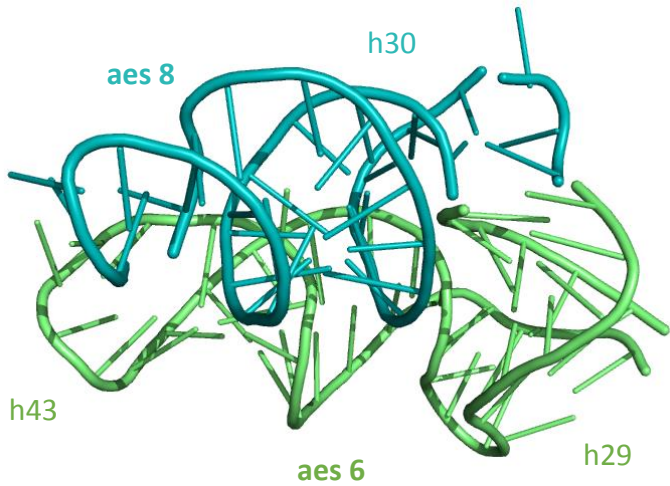
AES trunk/ *AES branch*	Helices trunk/ *Helices branch*	Nucleotides trunk/ *Nucleotides branch (type)*	3-Dimensional View
Phase 1			
1	44+(5'&3')	16S:(1402-1418); 16S:(1482-1502) / *16S:(5-19); 16S:(1531-1540)*	aes 2 insertion site / aes 1* / h44 / 5'-end of SSU RNA** (Shine-Dalgarno seq.) / 3'-end of SSU RNA (anti –SDS) *aes 1 used to be a stem loop composed of h44 and is now a vanquished helix formed by 5' and 3' termini **Modeled by mRNA of *E. coli*
Phase 2			
1 / *2*	44 / *45*	16S:(5-19); 16S:(1402-1418); 16S:(1482-1502); 16S:(1531-1540) / *16S:(1503-1530) (Y-type)*	aes 2 / h45 / h44 / aes 1 aes2 binds to aes1 via the A-minor interactions (A1518 and A1519) and provides an additional stability to the fragment)

30
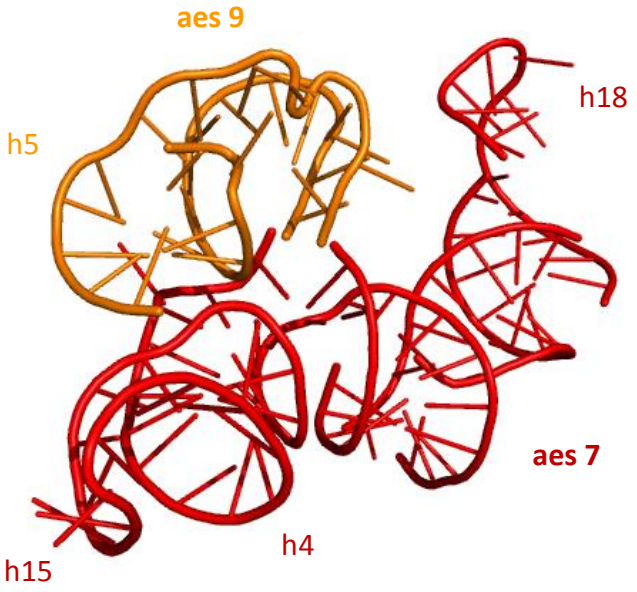
1 / 3	1, 44 / 2, 28	16S:(5-19); 16S:(1402-1418); 16S:(1482-1502); 16S:(1531-1540) / 16S:(15-21); 16S:(914-938); 16S:(1377-1396) (X-type)	Swivel point. 5'-end of SSU RNA** (Shine-Dalgarno seq.) aes 1 h44 3'-end of SSU RNA (anti –SDS) h28 h2 Modeled conformation of aes 3 aes 3 5'-end of SSU RNA dissociates and together with aes 3 undergoes a conformational change. A new conformation of 5' end is shown in orange
1 / 3	1, 44 / 2, 28	Positional shift 16S:(1-14); 16S:(1397-1418); 16S:(1482-1502); 16S:(1531-1540) / 16S:(15-21); 16S:(914-938); 16S:(1377-1396)	Dissociated 5'-end of SSU RNA aes 3 shifts with respect to aes 1 upon dissociation of the termini h1 aes 1 h2 aes 3 h44 aes 1 h28
3 / 3a	2, 28 / 27	16S:(15-21); 16S:(914-938); 16S:(1377-1396) / 16S:(885-889); 16S:(907-913) (L-type)	aes 3 h2 h28 h27 aes 3a binds to aes 1 via the A-minor interactions (A909) and stabilizes the dissociated conformation aes 3a
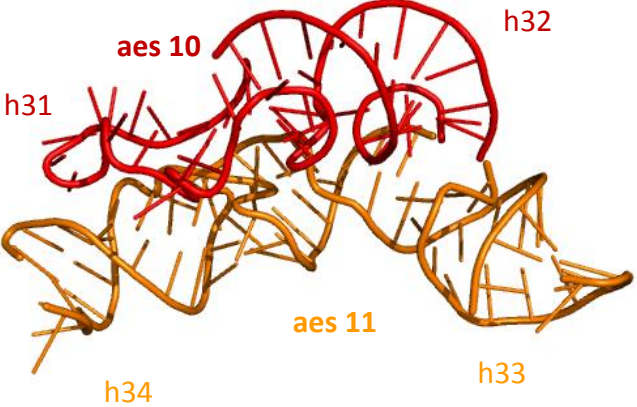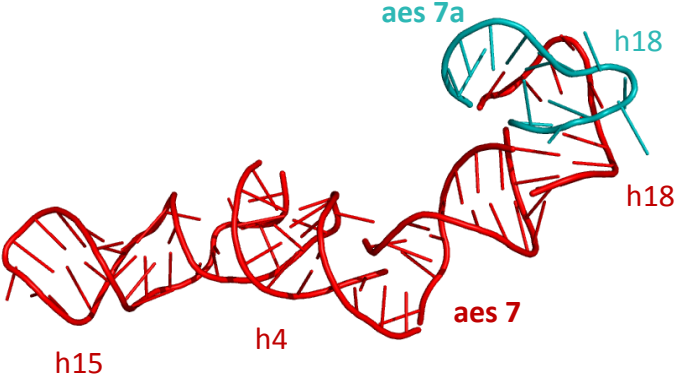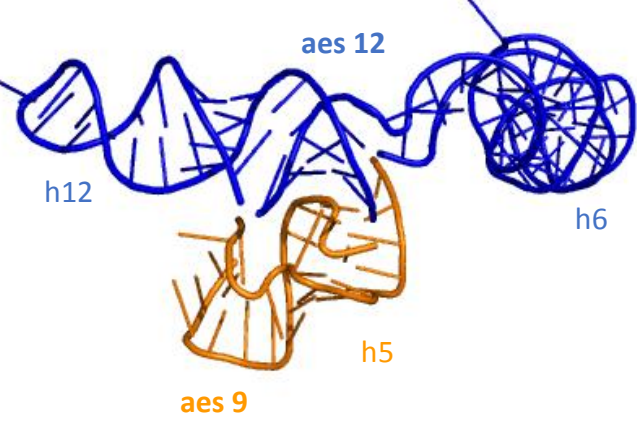
Phase 3			
3, 3a / *4*	**2, 28, 27 /** *3, 19*	**16S:(15-21);** **16S:(885-889);** **16S:(907-938);** **16S:(1377-1396);** / *16S:(22-38);* *16S:(548-569);* *16S:(821-827);* *16S:(871-884)* *(X-type)*	
4 / *4a*	**3, 19** / *19*	**16S:(22-38);** **16S:(548-569);** **16S:(821-827);** **16S:(871-884)** / *16S:(570-576);* *16S:(814-820)* *(Y-type)*	
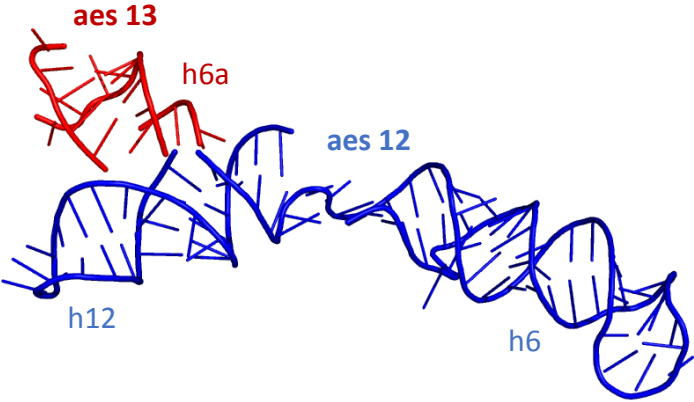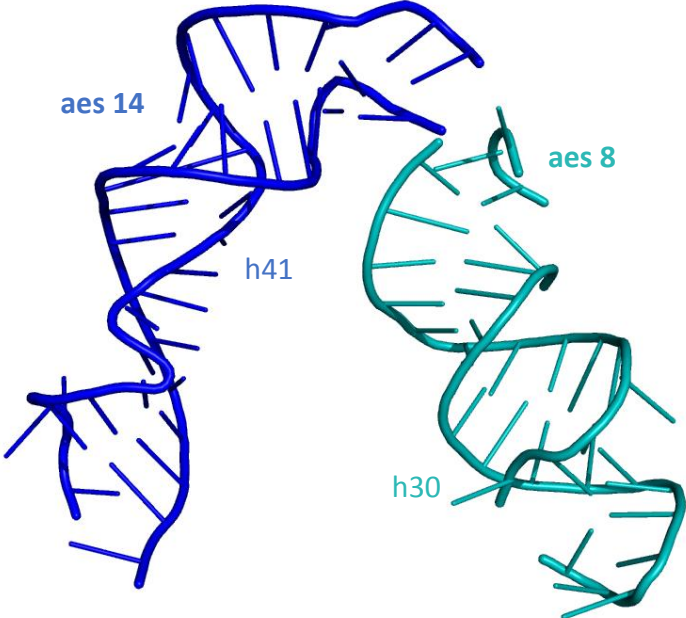
4a / 5	19 / 20, 24	16S:(570-576); 16S:(814-820) / 16S:(577-587); 16S:(754-776); 16S:(804-813) (T-type)	aes 5 binds to aes 2 via the A-minor interactions (A766) and provides an additional stability and integrity)
3a / 3b	27 / 27	16S:(885-889); 16S:(907-913) / 16S: (890-906) (I-type)	aes 3b binds to aes 5 via the A-minor interactions (A900 and A901) and provides an integrity to the entire fragment (aes 1-5) as a majority of aes are interlocked by the A-minor interactions
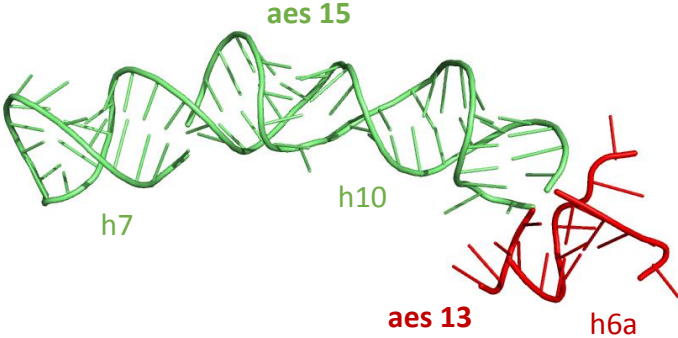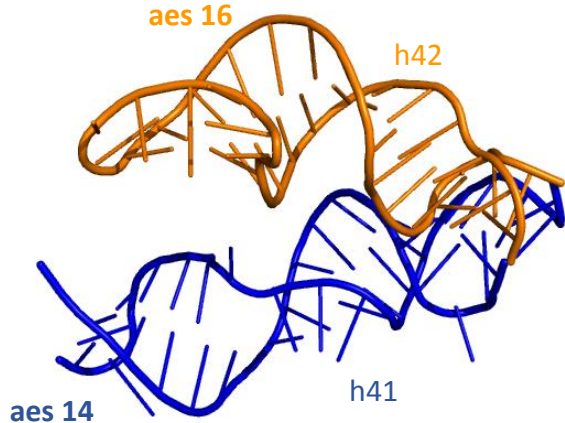
Phase 4			
5 / *5a*	**20, 24** / *24*	**16S:(577-587);** **16S:(754-776);** **16S:(804-813)** / *16S:(777-803)* *(L-type)*	
3 / *6*	**2, 28** / *29, 43*	**16S:(15-21);** **16S:(914-938);** **16S:(1377-1396)** / *16S:(939-944);* *16S:(1338-1376)* *(T-type)*	
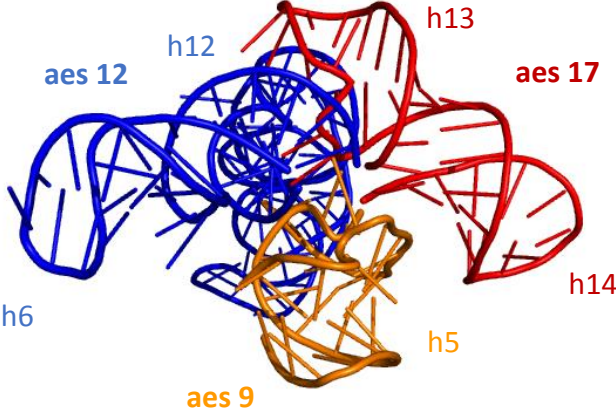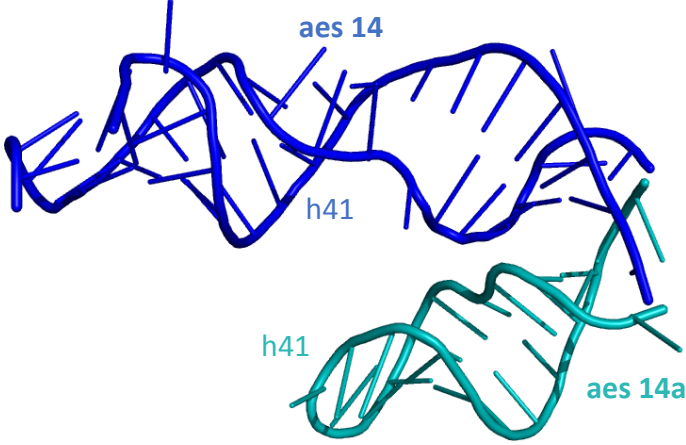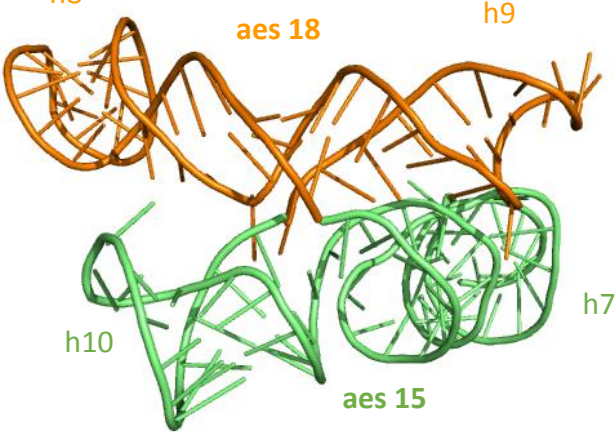
4 / *7*	**3, 19** / *4, 15, 18*	**16S:(22-38);** **16S:(548-569);** **16S:(821-827);** **16S:(871-884)** / *16S:(39-46);* *16S:(366-405);* *16S:(498-504);* *16S:(511-522);* *16S:(537-547)* *(T-type)*	
1 / *1a*	**44** / *44*	**16S:(1402-1418);** **16S:(1482-1502)** / *16S:(1419-1433);* *16S:(1467-1481)* *(I-type)*	
6 / *8*	**29, 43** / *30*	**16S:(939-944);** **16S:(1338-1376)** / *16S:(945-960);* *16S:(1335-1337);* *16S:(1224-1237)* *(T-type)*	
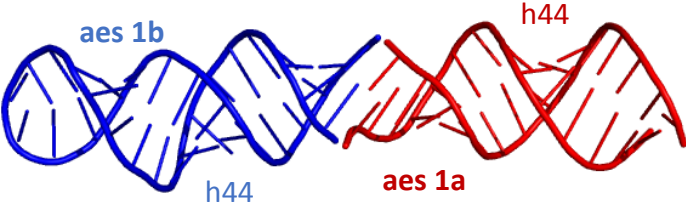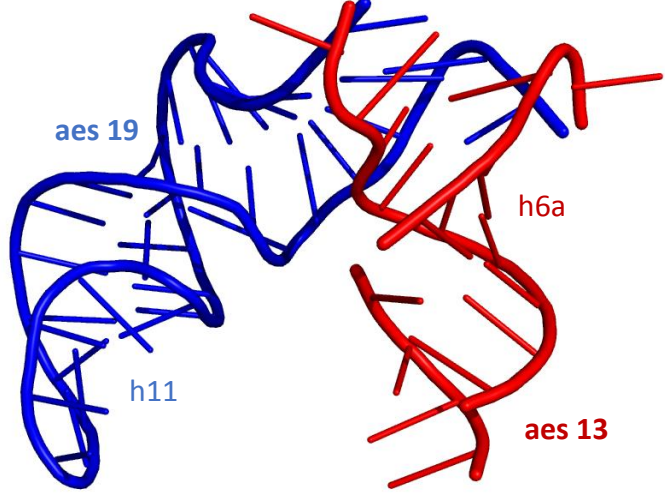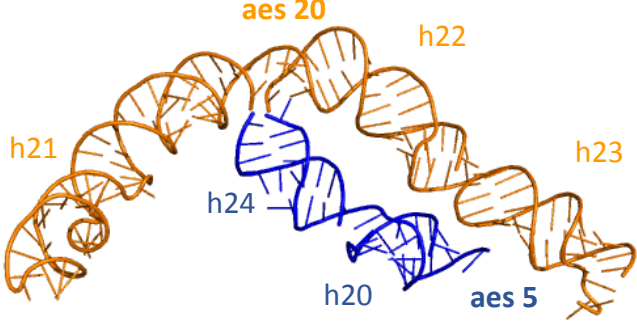
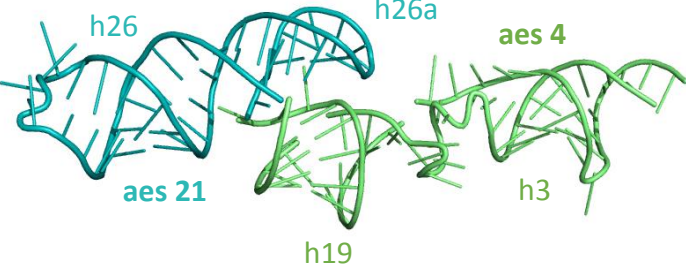| 7 / 9	4, 15, 18 / 5	16S:(39-46); 16S:(366-405); 16S:(498-504); 16S:(511-522); 16S:(537-547) / 16S:(47-59); 16S:(352-365) (Y-type)	
8 / 10	30 / 31, 32	16S:(945-960); 16S:(1335-1337); 16S:(1224-1237) / 16S:(961-991); 16S:(1212-1223) (T-type)	
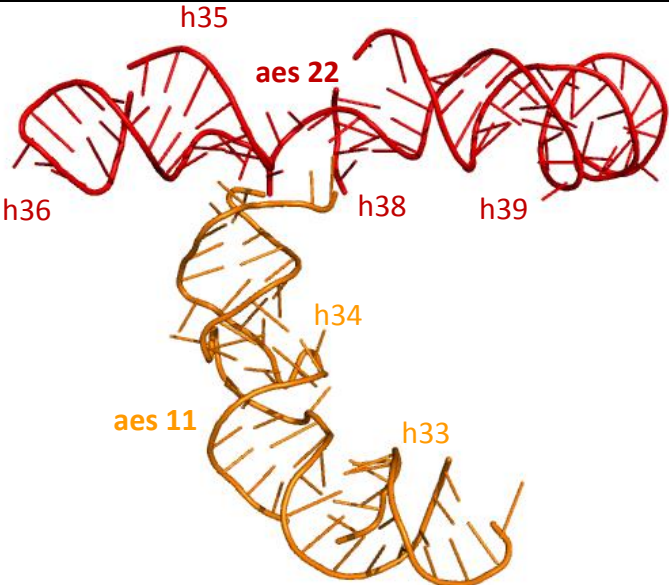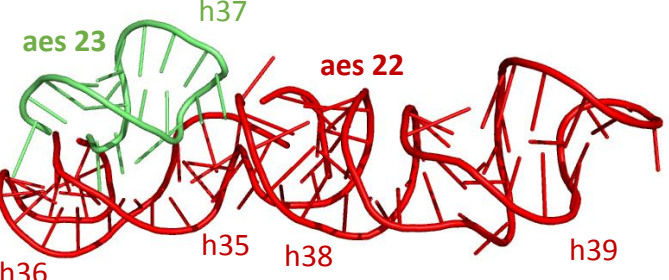
Phase 5			
10 / *11*	**31, 32** / *33, 34*	**16S:(961-991);** **16S:(1212-1223)** / *16S:(992-1004); 16S:(1037-1064); 16S:(1188-1211) (T-type)*	
7 / *7a*	**4, 15, 18** / *18*	**16S:(39-46);** **16S:(366-405);** **16S:(498-504);** **16S:(511-522);** **16S:(537-547);** / *16S:(523-536) (Y-type)*	
9 / *12*	**5** / *6, 12*	**16S:(47-59);** **16S:(352-365)** / *16S:(60-115); 16S:(289-315) (T-type)*	
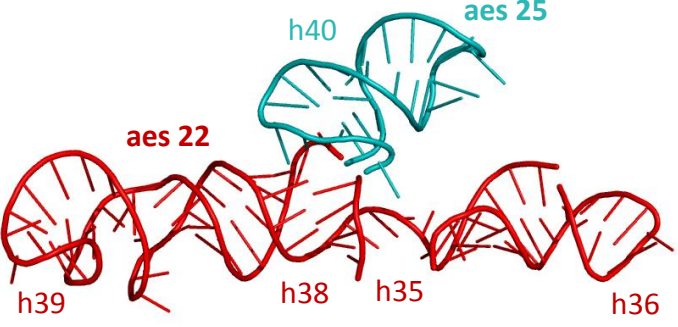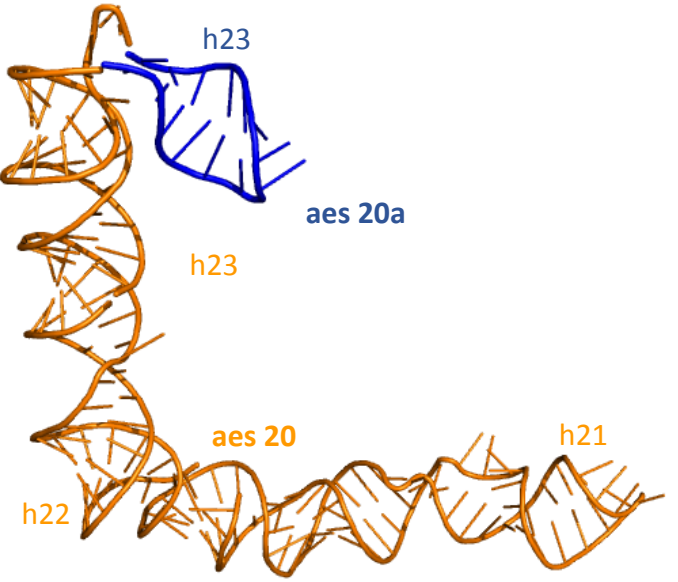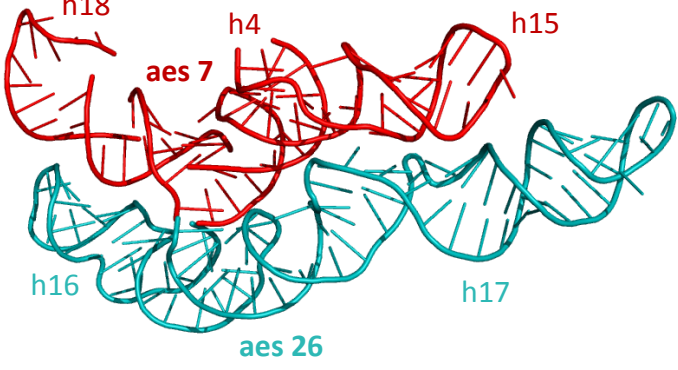
12 / *13*	**6, 12** / *6a*	**16S:(60-115);** **16S:(289-315)** / *16S:(116-119);* *16S:(240-244);* *16S:(280-288)* *(Y-type)*	
8 / *14*	**30** / *41*	**16S:(945-960);** **16S:(1335-1337);** **16S:(1224-1237)** / *16S:(1238-1256);* *16S:(1279-1299)* *(L-type)*	
7 / *7b*	**4, 15, 18** / *18*	**16S:(39-46);** **16S:(366-405);** **16S:(498-504);** **16S:(511-522);** **16S:(537-547)** / *16S:(505-510)* *(Y-type)*	
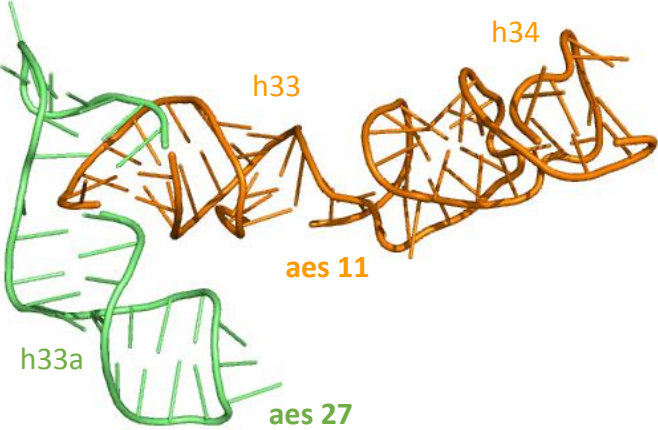
13 / 15	6a / 7, 10	16S:(116-119); 16S:(240-244); 16S:(280-288); / 16S:(120-143); 16S:(198-239) (Y-type)	
14 / 16	41 / 42	16S:(1238-1256); 16S:(1279-1299) / 16S:(1300-1334) (L-type)	

Phase 6			
9, 12 / *17*	**5, 6, 12** / 13, 14	**16S:(47-59);** **16S:(352-365);** **16S:(60-115);** **16S:(289-315)** / *16S:(316-351)* *(X-type)*	
14 / *14a*	41 / 41	**16S:(1238-1256);** **16S:(1279-1299)** / *16S:(1257-1278)* *(L-type)*	
15 / *18*	7, 10 / 8, 9	**16S:(120-143);** **16S:(198-239)** / *16S:(144-197)* *(X-type)*	

1a / 1b	44 / 44	16S:(1419-1433); 16S:(1467-1481) / 16S:(1434-1466) (I-type)	
13 / 19	6a / 11	16S:(116-119); 16S:(240-244); 16S:(280-288) / 16S:(245-279) (L-type)	
5 / 20	20, 24 / 21, 22, 23	16S:(577-587); 16S:(754-776); 16S:(804-813) / 16S:(588-686); 16S:(701-717); 16S:(734-753) (T-type)	
4 / 21	3, 19 / 26, 26a	16S:(22-38); 16S:(548-569); 16S:(821-827); 16S:(871-884) / 16S:(828-870) (T-type)	

11 / *22*	33, 34 / 35, 36, 38, 39	**16S:(992-1004); 16S:(1037-1064); 16S:(1188-1211)** / *16S:(1065-1083); 16S:(1102-1157); 16S:(1184-1187) (T-type)*	
22 / *23*	35, 36, 38, 39 / 37	**16S:(1065-1083); 16S:(1102-1157); 16S:(1184-1187)** / *16S:(1084-1101) (Y-type)*	
20 / *24*	21, 22, 23 / 23a	**16S:(588-686); 16S:(701-717); 16S:(734-753)** / *16S:(718-733) (Y-type)*	

22 / *25*	35, 36, 38, 39 / 40	**16S:(1065-1083); 16S:(1102-1157); 16S:(1184-1187)** / *16S:(1158-1183) (Y-type)*	
20 / *20a*	21, 22, 23 / 23	**16S:(588-686); 16S:(701-717); 16S:(734-753)** / *16S:(687-700) (L-type)*	
7 / *26*	4, 15, 18 / 16, 17	**16S:(39-46); 16S:(366-405); 16S:(498-504); 16S:(511-522); 16S:(537-547)** / *16S:(406-497) (X-type)*	

11 / 27	33, 34 / 33a	16S:(992-1004); 16S:(1037-1064); 16S:(1188-1211) / 16S:(1005-1036) (T-type)	