A structured viral RNA uses molecular mimicry and conformational dynamics to coordinate multiple functions

Steve Bonilla[1], Madeline E. Sherlock[1], Andrea MacFadden[1], and Jeffrey S. Kieft[1,2]

[1]Department of Biochemistry and Molecular Genetics and [2]RNA BioScience Initiative, University of Colorado Denver School of Medicine, Aurora, CO 80045, USA

*To whom correspondence should be addressed:

Jeffrey S. Kieft
Department of Biochemistry and Molecular Genetics
University of Colorado Denver School of Medicine
Mail Stop 8101
Aurora, CO 80045
Telephone: 303-724-3257
Fax: 303-724-3215
Email: Jeffrey.Kieft@cuanschutz.edu

Running Title: Cryo-EM structure of a multifunctional viral tRNA mimic

Keywords: viral RNA, RNA structure, RNA folding, tRNA mimicry, conformational dynamics, cryo-EM

ABSTRACT

Structured RNA elements are essential for biology and are ubiquitously used by viruses to control diverse infection-critical processes. Single-stranded RNA viruses often encode multiple processes into a single RNA element to maximize the functional capacity of their compact genomes. These important and elegant 'multifunctional' RNAs are hypothesized to use programmed structural changes to coordinate several, sometimes opposing, functions. However, detailed molecular mechanisms of such elements are largely mysterious because of the difficulty of solving dynamic RNA structures. We exploited recent advances in cryo-EM to directly visualize the architecture and infer conformational dynamics of the brome mosaic virus tRNA-like structure (BMV TLS), a multifunctional tRNA mimic that participates in replication, translation, and encapsidation of viral RNAs and that has eluded structure determination for decades. We found that although BMV TLS is aminoacylated by cellular tyrosyl-synthetase (TyrRS), its primary conformation is incompatible with TyrRS binding. Rather, the RNA is preorganized for replication, positioning the replicase promoter and the initiation site in close proximity to each other. An alternative 'tyrosylation-ready' conformation requires repositioning of a conformationally dynamic structural domain, and this change must induce additional rearrangements that disrupt the 'replication-ready' configuration. These results demonstrate how programmed RNA dynamics can evolve to coordinate interactions with diverse cellular and viral proteins and thus organize multiple functions on a single RNA platform. Our results support the paradigm that RNA structures are inherently dynamic conformational ensembles, which enables multifunctionality. This work also highlights the emerging power of cryo-EM to dissect the dynamic conformational landscape of small discrete functional RNAs. Furthermore, we anticipate our method of rapidly mapping RNA domains within cryo-EM maps to be broadly useful.

INTRODUCTION

RNA's remarkable functional versatility derives from its unique ability to both store genetic information and form dynamic three-dimensional structures (*1, 2*). This is exploited by RNA viruses, which universally use structured RNA elements to interface with and manipulate the host's cellular machinery and to regulate essential viral processes (*3-5*). Infection-critical structured RNA elements exist in viral clades as diverse as flaviviruses, lentiviruses, coronaviruses, alphaviruses, and picornaviruses (*3, 4, 6-12*). Often, a single viral RNA element regulates several processes using dynamic conformational changes, an important strategy to extend the functional capacity of relatively small RNA viral genomes (*1, 8, 13*). Despite their importance, understanding of such RNA elements is rudimentary, in part because solving the structures of dynamic RNAs is challenging. Crystallization of such elements is difficult due to their inherently dynamic nature, and nuclear magnetic resonance (NMR) is often not tractable for fully functional RNAs due to limitations on the sample molecular weight (*14, 15*). Thus, the structure-function rules that govern the molecular mechanism of multifunctional conformationally dynamic viral RNA elements are not well understood.

Important examples of multifunctional viral RNA elements are the tRNA-like structures (TLSs) at the 3′ ends of certain positive-strand RNA viral genomes (*16, 17*). These elements were thought to be present only in plant-infecting viruses, but recent evidence confirms that they exist in animal-infecting viruses and that other examples likely remain to be discovered (*18*). TLSs were originally identified by their ability to drive aminoacylation of viral genomic 3′ ends by host aminoacyl-tRNA synthetases (aaRSs) and were later found to interact with other tRNA-specific enzymes including CCA-nucleotidyltransferase (CCA-NTase) and translation elongation factor 1A (eEF1A) (*16, 19*). Beyond recruiting tRNA-specific machinery, TLSs have roles in translation, replication, and encapsidation of the viral RNAs (*16, 17, 20*). Some functions are coupled to the aminoacylation state of the TLS and therefore its ability to mimic tRNA, while other functions appear to be independent of this (*16*). In addition, certain TLSs have been proposed to undergo dynamic conformational changes to regulate different functions (*21*). RNA structure-driven coordination of multiple functions is widely utilized across RNA viruses; thus, TLSs are powerful models to reveal fundamental rules applicable to diverse RNA structure-based processes.

The viral TLSs are classified into three types based on their aaRS specificity: valylatable (TLS^{Val}), histidylatable (TLS^{His}), and tyrosylatable (TLS^{Tyr}) (*16, 19*). Each type has a characteristic secondary structure that can differ dramatically from tRNA, highlighting that tRNA mimicry can be achieved in many ways (Figure S1) (*19*). This fact has broad importance as viral tRNA mimicry appears to be widespread, allowing viruses to exploit important cellular machinery such as the ribosome, translation factors, RNase P, and aaRSs. For example, several viral internal ribosome entry sites (IRESs), including that of hepatitis C virus and the *Dicistroviridae*, mimic parts of tRNA to contact tRNA binding sites on the ribosome and facilitate translation of viral proteins (*3*). In addition, the 5′ untranslated region (UTR) of the HIV-1 RNA genome contains a tRNA-like element that binds lysyl-tRNA synthetase and favors the release of bound $tRNA^{Lys3}$, the primer for HIV reverse transcription (*9*), and at least one DNA herpesvirus produces tRNA-like transcripts that may influence pathways linked to virulence (*22*). Thus, tRNA mimicry is a general strategy used by diverse viruses, including some important to human health.

Of the three TLSs classes, TLS^{Tyr} variants differ the most in size and secondary structure from authentic tRNAs and are the most difficult to reconcile with tRNA mimicry (Figure 1A). The prototype TLS^{Tyr} from brome mosaic virus (BMV), found at the 3′ end of each of the three genomic RNAs of the tripartite BMV genome, contains 169 nucleotides and seven helical stems and is a substrate for tyrosyl-tRNA synthetase (TyrRS) (Figure 1A) (*19, 23*). Thus, BMV TLS is substantially larger and more structurally complex than tRNA, presenting an example of highly divergent ‘non-predictable’ tRNA mimicry that promises novel insights into other tRNA mimics. Questions remain about which stems in the TLS serve as the structural analogs of tRNA, with conflicting evidence pointing at different potential anticodon arm analogs (*24-27*). Furthermore, because the complex structure of BMV TLS evolved to facilitate diverse functions, it provides a model of how a single RNA ‘platform’ can coordinate multiple processes. Specifically, in addition to recruiting host biochemical machinery through tRNA mimicry, BMV TLS contains the promoter for the viral replicase, binds the viral coat protein to nucleate virion assembly, serves as a template for genome recombination, and enhances translation of viral proteins (*16, 20, 28-31*). These functions have complex dependences on the structure and tyrosylation state of the TLS that are not fully understood (*16*). Despite the lessons that a structure of BMV TLS could provide, its three-dimensional structure has remained elusive for decades. However, recent advances in cryo-electron microscopy (cryo-EM) offered opportunities to break this impasse.

To uncover rules of diverse tRNA mimicry and gain insights into how multifunctionality can be achieved in a single RNA sequence, we undertook an in-depth structural analysis of the BMV TLS. Using single-particle cryo-EM, we obtained a map of the ~55 kDa BMV TLS to a final overall resolution of 4.2 Å and built a structural model employing a novel strategy to assign secondary structural elements to the density. The structure reveals a distinct type of tRNA mimicry and suggests how the RNA organizes multiple processes using programmed conformational dynamics. Specifically, in the ground state tRNA mimicry is incomplete, with a conformationally flexible anticodon stem analog in a position that differs from the canonical tRNA L-shape. Instead, this conformational state appears organized to facilitate replication by positioning the promoter in immediate proximity to the replication initiation site, which also serves as the tRNA acceptor stem analog. To be aminoacylated, the BMV TLS must undergo a large conformational change that disrupts the promoter position. Thus, structural features that facilitate interactions with either the replication or tyrosylation machinery appear to be mutually exclusive, suggesting a programmed conformational switch within the BMV TLS structure. Overall, our studies provide a structural framework for interpreting decades of functional and biochemical studies of BMV TLS–including reconciling previous conflicting models about the identity of its tRNA mimicking elements–and provide new hypotheses for further experiments. More broadly, these results have implications for our general understanding of RNA folding, dynamics, and function, and highlight the power of cryo-EM for dissecting the dynamic conformational ensembles of small functional RNA elements.

RESULTS AND DISCUSSION

A strategy to reveal the global architecture of BMV TLS by cryo-EM

To generate an RNA sample for structural studies, we *in vitro* transcribed and purified a previously characterized complete BMV TLS RNA sequence (*19, 23*), then tested its ability to be tyrosylated *in vitro* using recombinantly expressed TyrRS from model host *Phaseolus vulgaris* (common bean). The RNA was efficiently tyrosylated relative to a set of controls, indicating that the RNA sample folded to the native functional state (Figure 1B). To visualize the overall architecture of the BMV TLS RNA's functional state, we used conventional defocus-based single particle cryo-EM with a 200 kV electron microscope (Figure S2 & S3). Despite the small size of BMV TLS (55 kDa), a phase plate was not necessary during data collection as the RNA particles displayed high contrast and were readily identifiable in the micrographs (Figure 1C). We

generated an initial 'test' map with an overall resolution of 6.7 Å that readily revealed the global topology of the RNA, with A-form helical features and apparent coaxial stacking between helical stems around a central core (Figure 2A). Although A-form helices were clear, assignment of this density to specific parts of the TLS required additional constraints. To address this, we conceived a strategy to use the well-established concept of RNA structural modularity and the known secondary structure of BMV TLS to design altered RNAs in which specific stems were either extended or shortened by several base pairs (bp). Extension of helices has proven valuable in other applications such as to discern the orientation of helical domains emanating from an RNA catalytic core, infer the conformation of isolated junctions, and to slow overall motions relative to internal motions in NMR studies of small RNAs (*32-34*). Here, we reasoned that localized differences in the cryo-EM maps of such modified RNAs compared to the unmodified RNA would allow identification of the extended/shortened helices and thus assignment of the density.

We applied the helix extension/truncation strategy first with BMV $TLS_{B3ext_Cshort,}$ a modified RNA with stem B3 extended by 5 bp and stem C shortened by 4 bp (Figure 2B). We designed the RNA with both an extension and a truncation to allow unambiguous identification of two helices in a single experiment. A mid-resolution map of this RNA displayed local differences in density relative to that of wild type that were consistent with the extension of one helix and the shortening of another helix, clearly identifying helices B3 and C in the density (Figure 2B). We repeated this analysis with BMV $TLS_{Dext_B2short}$, in which stem D was extended by 4 bp and stem B2 was shortened by 4 bp (Figure 2C). Interestingly, a new tertiary interaction apparently formed between the modified helices; however, this putative interaction did not affect the global architecture of the RNA, and local differences allowed us to assign stems B2 and D (Figure 2C). The constraints imposed by the secondary structure and this strategy allowed us to unambiguously identify all the helices of BMV TLS in the density. Note that because high resolution maps were not necessary to identify the modified helices, each map was obtained in ~ 24 hrs. of microscope and computation time. This method is potentially a relatively rapid and widely applicable strategy for identification of specific regions in cryo-EM maps, especially when robust secondary structures are available, complementing emerging computational modelling tools (*35-37*).

We used our initial 'test' cryo-EM map with the assigned helices to evaluate a decades-old published 3D model of the BMV TLS built using functional data, chemical and enzymatic probing, and phylogenetic

comparisons (Figure S4) (*25*). Remarkably, the model correctly predicted five helical stems emanating from a central core and the formation of a tRNA-like L shape involving helices A, E, and B3. However, the model does not fit well within the experimental map and there are major differences in the relative orientation and position of the helices, motivating higher-resolution studies to understand tRNA mimicry and multifunctionality.

Higher-resolution cryo-EM yields a complete structural model of BMV TLS

To obtain a high-resolution map suitable for building a complete structural model of the BMV TLS RNA, we used a 300 kV electron microscope and used the same sample as above to obtain a map with a final overall resolution of 4.2 Å (Figure 3A; Figure S5). The improvement in resolution was due to a combination of higher microscope power and a larger number of micrographs collected (Figure S2 & S5). As with the initial map, the positions and orientations of the helices were readily discernable (Figure 3A). Furthermore, the minor and major grooves of A-form helices were clearly defined, and the connectivity of the phosphate backbone could be deduced. In regions with relatively high local resolution, stacking and coplanarity of bp was resolved, and “bumps” of phosphate groups were clearly visible along the RNA backbone (Figure S6). As expected, the central core of the map displayed the highest local resolution, but even peripheral region density was well defined (Figure 3A & Figure S6). The exception to this was the density corresponding to the regions containing stems B3 and E and, to a lower extent, the peripheral area of stem C, which were less defined and suggested conformational dynamics (Figure 3A & Figure S6). Interestingly, the density containing stems B3 and E comprises part of the L-shaped region most likely to be mimicking tRNA (Figure 3A).

The secondary structure, helical assignments, and the published computational model were used to build a complete structure of the BMV TLS by docking individual helices into the corresponding density, followed by the application of manual and automated fitting tools and geometry correction using tools based on the Rosetta molecular modelling framework (Figure 3B & 3C) (*38, 39*). The entire structure could be built and refined within the map without substantial steric clashes or breaks in the chain (Figure S6 and Table S1). Structural models of the extended/truncated constructs supported our helical assignments and the accuracy of our BMV TLS model (Figure S7). Overall, we built and refined a three-dimensional model of a complex 55 kDa RNA using a combination of cryo-EM maps (~ 7 to 4 Å resolution), systematic helical extension and

truncation to assign density to specific stems, and tools and resources that are now available to the general scientific community.

Globally, the BMV TLS structure can be described by three domains, each containing a set of coaxially stacked helical elements (Figure 3D). The first comprises helices C, B1, and B2, which form a pseudo-continuous helix that spans the length of the structure, with the apical loops of B2 and C pointing in opposite directions (Figure 3D; center). The second stack comprises helices D and A connected through a central 4-way junction (4wj; Figure 3B-D); helix A contains the 3' CCA and serves as both the acceptor stem for tyrosylation and the replication initiation site (Figure 3D; right)(*16*). The third domain contains stacked helices B3 and E (herein referred to as B3-E domain) and is relatively loosely connected to the rest of the structure (Figure 3D; left). Specifically, it links to the structure through a single strand of RNA connecting B3 to B1 (Figure 3D). As mentioned above, the electron density corresponding to the B3-E domain was weaker and less defined, suggesting that it is dynamic relative to the rest of the structure. While atomic-level details are ambiguous at the overall resolution of the map, the cryo-EM structure of BMV TLS reveals functionally important features and allowed us to infer conformational dynamics with implications to TLS multifunctionality.

BMV TLS positions the replicase promoter in immediate proximity to the replication initiation site

One critical function of BMV TLS is to recruit the replication machinery and present it with the proper initiation site (*20, 28, 40*). Although replication initiation occurs at the 3' end in stem A, the replicase promoter is contained within stem C (Figure 3B) (*28, 41*). Two sequences in stem C are critical for promoting replication: the apical AUA triloop and a UAGA 4-nt bulge formed by nts 114-117 (Figure 3B) (*28, 40, 41*). Our model supports qualitative features revealed by a decades-old NMR model of the AUA apical triloop in the context of an isolated helix C (*41*), although details of the triloop conformation are ambiguous at the overall resolution of the map (Figure S8). Importantly, our structure revealed that stem C is positioned adjacent and nearly parallel to stem A and that, therefore, the essential AUA apical triloop is positioned immediately proximal to the replication initiation site, showing how the BMV TLS facilitates replication initiation by orienting these structural elements in 3D space (Figure 4A). The structure thus explains previous studies that used a ‘minimal promoter’ consisting only of stem C and a single-stranded initiation sequence of varying lengths to show that the position of the initiation site relative to the promoter is essential for efficient replication (*40*). Furthermore,

previous studies suggested that the UAGA 4nt-bulge in stem C may help positioning the promoter (*42*). The structure now suggests that the proximity of the 4-nt bulge to 4wj creates tertiary interactions that may stabilize the orientation and positioning of stem C relative to the initiation site (Figure 4A), consistent with the previous study and with the importance of the 4-nt bulge for replication. It has also been suggested that the 4-nt bulge makes directs interactions with the replicase (*40*). Overall, the structure suggests that, in the ground state, BMV TLS RNA populates a 'replication ready' conformation that positions the promoter in stem C near the initiation site in stem A, reducing the conformational search of the bound replicase for its substrate.

Structure and conservation identify elements that mimic tRNA

Stem A is both the replication initiation site and also the acceptor stem analog for aminoacylation, suggesting that different machineries compete for this element and this is regulated by the overall RNA structure (*25, 28*). The prevailing hypothesis is that the ability of TLS^{Tyr} RNAs to be specifically aminoacylated by TyrRS is dictated by structural rules that are similar to those used by authentic $tRNA^{Tyr}$ (*24, 25*). Because $tRNA^{Tyr}$ recognition by TyrRS occurs via interactions with its acceptor stem and anticodon loop (with the acceptor stem playing a much more important role in tyrosylation than the anticodon), structural analogs of these features are expected to reside in the BMV TLS and a major goal towards understanding BMV TLS tRNA mimicry was to identify them (*16, 25, 43, 44*). Although early studies identified pseudoknotted helix A as the acceptor stem analog, conflicting evidence supported either stem B3 or B2 as the anticodon stem (*24-27*). The structure now resolves this ambiguity. Candidate stem B2 is orientated away from the acceptor stem and is unlikely to directly participate in TyrRS recognition, while the conformationally flexible B3-E domain lies at a right angle to the acceptor stem (Figure 4B). This, combined with previously reported mutagenesis, enzymatic probing, and computational modelling (*24, 25, 27*), makes B3 the best candidate for the anticodon stem analog.

To further examine the putative roles of B3 and E in tRNA mimicry and TyrRS recognition, we generated a consensus sequence and secondary structure model of TLS^{Tyr} using covariation analysis, reasoning that features important for tRNA mimicry would be conserved across diverse TLS^{Tyr} RNAs (Figure 4C). We started with a preliminary alignment of eight TLS^{Tyr} sequences from the Rfam database (*45, 46*), then performed homology searches that resulted in 512 unique sequences that were used to build a consensus

model (Figure 4C), supported by SHAPE chemical probing of three representative variants from different viruses (Figure S9). The consensus shows that the number of bps in B3 and E is highly conserved (Figure 4D), suggesting a strong selective pressure that constrains the geometrical properties of these helices. Although the total number of bp in B3 and E is greater than in the anticodon arm of $tRNA^{Tyr}$ (12 bp in BMV TLS vs. 10 bp in $tRNA^{Tyr}$), this difference is compensated by the smaller sizes of the apical loops of B3 and E relative to those of the anticodon and D loops in $tRNA^{Tyr}$. For BMV TLS, the combined length of E and B3 in the structure is approximately the same as that of the tRNA anticodon+D loop arms (~ 50 Å; Figure 4B), and this is expected to be similar across TLR^{Tyr} RNAs based on the consensus model, consistent with a role in TyrRS recognition. Interestingly, the B3 apical loop does not contain a canonical 'GUA' tyrosine anticodon but contains a conserved 'ACA' that is capable of mimicking 'GUA' based on sequence similarity (*25*), suggesting that the optimal anticodon sequence may be context dependent and/or that there are other selection pressures on the sequence of B3 in addition to TyrRS binding. Of note, the major identity elements of $tRNA^{Tyr}$ lay within its acceptor arm and interactions between TyrRS and the anticodon play a less important role (*43*), suggesting more flexibility and less specificity than is found in tRNA/aaRS pairs with strict anticodon requirements; this would also apply to TLS^{Tyr}.

In comparison to E and B3, stem B2 varies more widely in length, with a few sequences showing as many as 20 bps, and these variations are not compensated by changes in the length of B1 (Figure 4D). In addition, there is no conservation in the apical loop of B2 that would suggest that it serves as an anticodon analog (Figure 4C), arguing against a direct role of B2 in TyrRS recognition. Previously observed aminoacylation defects from structural disruption of B2 can be explained by effects on the overall structure of BMV TLS, as B2 appears integral to stabilizing the global fold (Figure 3C). Thus, we conclude that B3 is the analog for the anticodon stem in the TLS^{Tyr}. However, as stated above, domain B3-E is contained within a region of the cryo-EM map that appears to be relatively flexible. Thus, the tRNA-mimicking characteristics of the BMV TLS appear to be conformationally dynamic, potentially enabling multifunctionality and regulation as described next.

Conformational changes are required to bind TyrRS

Having established the elements within the BMV TLS that mimic tRNA, we now compared them to the structure of an authentic tRNA and to authentic $tRNA^{Tyr}$ bound to TyrRS. Unlike in the classic L-shaped tRNA fold with docked T and D loops, stems A and domain B3-E of the BMV TLS are loosely associated with no interactions analogous to those between the T and D loops (Figure 4B). Thus, the overall dimensions of the tRNA-mimicking portion of the BMV TLS are larger than those of tRNA, reminiscent of an 'extended' tRNA (Figure 4B). We expected these differences to substantially influence potential interactions with the TyrRS. To directly assess this, we used the published crystal structure of a yeast $tRNA^{Tyr}$-TyrRS complex (*44*) to model BMV TLS bound to TyrRS (Figure 5A & B). This approach is valid as the BMV TLS can be efficiently aminoacylated by yeast TyrRS *in vitro* and sequence conservation suggests that TyrRS from yeast and from plant hosts of BMV are highly similar (Figure S10)(*27*). Superposition of the structures of BMV TLS and bound $tRNA^{Tyr}$, based on the alignment of their acceptor stems (Figure S10), suggests that large conformational changes by the anticodon stem of BMV TLS are necessary upon TyrRS binding to prevent steric clashes and to properly align the anticodon sequence with the protein (Figure 5B). Specifically, to properly dock the TyrRS, dynamic domain B3-E must rotate ~180º and translate ~15 Å such that the apical loop of stem E is brought into the core junction, effectively making the TLS structure more compact (Figure 5C & 5D). After the modelled conformational change, the tRNA mimicking elements of BMV TLS remarkably resemble an authentic tRNA (Figure 5E). Interestingly, these changes would induce additional structural changes to avoid new intramolecular steric clashes between stem E and the core of the TLS (Figure 5D); thus, rearrangements in several parts of the molecule appear necessary to reach the tyrosylation-competent state.

Structure suggests conformational changes between 'replication-ready' and 'TyrRS-ready' states

Our data strongly suggests that, to productively interact with the TyrRS, domain B3-E must assume an anticodon stem-like position that would place the apical loop of E into the core of the structure, creating severe steric clashes (Figure 5D). Stem C, which contains the replicase promoter and is positioned directly adjacent to stem A, would require significant reorientation to resolve these steric clashes with the repositioned stem E (Figure 5D). While we cannot confidently predict the position of stem C in the TyrRS bound state, it is possible that it would move away from stem A. Thus, the conformation required for productive TyrRS binding and aminoacylation would disrupt the preorganization of the promoter near the replication initiation site.

Consistent with this, we note that the density for stem C in the cryo-EM map is somewhat less defined than that of other helices in the structure (Figure 3A & Figure S7), suggesting that C may be primed to move. Furthermore, the relative positions of the B3-E domain, stem A, and stem C in the ground state suggest that the B3-E domain is positioned 'out of the way' so that the replication machinery can assemble around the promoter and have direct access to the initiation site. These observations strongly suggest that the dynamic nature of domain B3-E evolved to facilitate accessibility of the replication machinery to the promoter and initiation site while still supporting tRNA mimicry and TyrRS binding by a conformational switch (Figure 5F).

These observations suggest a model for how the BMV TLS can help coordinate replication with other processes during infection. In the ground state, the 'replication-ready' conformation predominates (Figure 5F; 'observed'), allowing the replication machinery to assemble around the promoter and readily access the initiation site. To bind TyrRS and be aminoacylated, BMV TLS switches to an alternative 'TyrRS-ready' conformation (Figure 5F; 'hypothesized') in which the B3-E domain assumes an anticodon arm position and stems C may move away, possibly with other rearrangements across the structure. The hypothesized TyrRS-ready conformation is not significantly populated in the ground state as we do not observe it in our cryo-EM experiments, but is supported by the observed flexibility of the B3-E domain and the model of the BMV bound to TyrRS.

We envision two models for the mechanism of the conformational switch and its relation to TyRS binding. First, in a 'conformational capture' model, the TyrRS-ready conformation is transiently sampled in the ground state as part of the conformational ensemble of BMV TLS and is bound and stabilized by TyrRS. Second, in a 'bound intermediate' model, TyrRS binds first to the acceptor stem of an intermediate state (i.e. a conformation that is not the final bound form) and is followed by reorganization of the BMV structure that includes the re-positioning of the B3-E domain to mimic the anticodon stem. We favor the 'conformational capture' model because it requires fewer assumptions (i.e. no intermediate alternative binding states) and because conformational capture is common in RNA protein interactions and is consistent with the emerging paradigm of functional RNAs existing as dynamic conformational ensembles that populate various functional states (*1, 2*).

Overall, the structure presented here suggests that, in contrast to tRNAs which are largely preorganized for binding their cognate synthetases, the BMV TLS must undergo programmed conformational

changes to bind TyrRS. This also contrasts with the TYMV TLS^{Val}, which nearly perfectly mimics tRNA in the unbound state (*47*). Thus, the BMV TLS reveals a novel dynamic strategy for viral tRNA mimicry that is likely linked to its multifunctionality. As described above, the BMV TLS mimics tRNA to interact with host tRNA-specific proteins and it also interacts with various other host and viral proteins to function in replication, recombination, and encapsidation of the viral RNAs (*28*). Each of these functions may require that BMV TLS assumes a different conformation that allows access to the necessary machinery and that is compatible with the required set of function-specific binding interactions, such as the conformations described here. Future experiments of BMV TLS in complex with protein partners, as well as structural studies of other TLS^{Tyr} variants, will further elucidate these and perhaps other conformational changes.

Cryo-EM is a powerful tool for dissecting RNA's dynamic conformational landscape

The BMV TLS has served as a model system for understanding RNA structure-function relationships for decades but its three-dimensional structure remained elusive. Using cryo-EM we have elucidated the tertiary conformation of BMV TLS and revealed not only structural features, but previously unknown dynamic characteristics that have important regulatory implications. In fact, the dynamic nature of this RNA may explain why decades of previous crystallographic efforts failed. RNAs exist as conformational ensembles, and dynamics have emerged as a critical aspect of RNA function (*1, 14, 48*). Unlike riboswitches, that undergo rearrangements that involve changes in secondary structure (i.e. formation and/or breaking of base pairs), BMV TLS appears to use tertiary changes that spatially rearrange elements in 3D space. This type of dynamics may be widespread but difficult to detect by conventional methods. Recent studies have already explored cryo-EM as a powerful tool for rapidly solving the structure of small RNAs (*35, 49*) and our studies further highlight the great potential of cryo-EM for dissecting not just structures but also the dynamic conformational landscapes of functional RNAs. Here, we presented additional tools that will aid in this task: the use of extensions and/or truncations of modular helical domains in combination with robust secondary structures to rapidly and unambiguously assign specific secondary structural elements to the electron density and provide constraints for structural modelling. This technique, in combination with emerging computational tools, will greatly facilitate the study of diverse dynamic regulatory RNAs such as viral RNAs, riboswitches, ribozymes, and aptamers by cryo-EM.

MATERIALS AND METHODS

In vitro transcription, purification, and folding of RNA

Templates for *in vitro* transcription were generated by PCR amplification of sequence-verified double-stranded DNA fragments (gBlocks; Integrated DNA technologies). Sequences for all constructs are provided in supplementary material (Table S2). To reduce the amount of N+1 products generated by *in vitro* transcription and ensure correct termination at CCA_{OH}, which is necessary for aminoacylation, PCR amplification was performed using a reverse primer with a 2′OMe modification at the first 5′ nucleotide, as described previously (*21*). As one of the negative controls for BMV TLS aminoacylation, we produced BMV TLS (2'–3' cP), which is identical to BMV TLS with the exception of a terminal 2′-3′ cyclic phosphate that prevents efficient aminoacylation. To introduce the terminal 2′-3′ cyclic phosphate, we included a 3′ flanking 69 nt hepatitis delta virus (HDV) self-excising ribozyme that was removed after transcription (Table S2); the product of the self-excision results in the production of the terminal 2′-3′ cyclic phosphate. A 'GG' was added on the 5' end of the sequences of interest to facilitate transcription by T7 RNA polymerase. The transcription reactions contained 8 mM of each NTP (ATP, UTP, CTP, and GTP), 60 mM $MgCl_2$, 30 mM Tris, pH 8.0, 10 mM DTT, 0.1 % spermidine, 0.1 % Triton X-100, 0.24 u/µl RNasin RNase inhibitor (Promega), and ~ 0.14 mg/ml T7 RNA polymerase (in-house prepared). The transcription reactions were incubated for 3 hrs. at 37°C. To cleave the ribozyme in BMV TLS (2'–3' cP), $MgCl_2$ was raised to ~ 120 mM after transcription, followed by incubation at 60 °C for 15 minutes. The transcribed RNA constructs were ethanol precipitated, purified by denaturing gel electrophoresis (8% acrylamide), cleaned using magnetic beads (Ampure), and buffer exchanged into RNase free water using 10 KDa cutoff centrifugal filters (Amicon). To fold the RNAs into their native conformation, the RNAs were resuspended at the desired concentration into a buffer solution containing 50 mM MOPS pH 7.0, 30 mM Na^+, and 10 mM $MgCl_2$, incubated at 80 - 90 °C for 1 minute, and cooled down to 23 °C for 15 minutes.

Preparation of TyrRS

To test the aminoacylation capacity of our BMV TLS construct we used TyrRS from a model host of BMV, *Phaseolus vulgaris* (common bean), as previous studies have shown that this TyrRS variant efficiently tyrosylates BMV RNA *in vitro* (*50*). We obtained the sequence of this TyrRS from the UniProt database (gene:

PHAVU_002G027700g) and optimized it for expression in *E. coli* using the online IDT Codon Optimization Tool. A gBlock (Integrated DNA Technologies) encoding the TyrRS sequence, cleavage sequences for restriction enzymes NdeI and BamHI, and a 6-His-tag was cloned into a pET15b vector using standard molecular biology techniques for ligation with T4 DNA ligase. The protein was expressed in BL21 (DE3) cells in Luria broth (LB) containing ampicillin at 37 °C until OD600 reached 0.6 and then translation was induced with the addition of 0.5 mM IPTG for 4 hours. The cells were harvested by centrifugation at 5488 x g for 12 minutes at 6 °C. The pellets were resuspended in lysis buffer (20 mM Tris-HCl pH 7, 500 mM NaCl, 2 mM β-mercaptoethanol, 10% glycerol, and 1 EDTA-free protease inhibitor tab (Roche)) and sonicated for a total of 2 minutes (20 sec on and 40 sec off intervals) at 75% amplitude. The lysate was centrifuged at 30,000 x g for 30 min at 4 °C before being loaded into a gravity flow column (Bio-Rad) and purified by Ni-NTA resin (UBP Bio) followed by size exclusion using FPLC and a Sepax 300 SEC column. The protein was stored in buffer containing 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 2 mM β-mercaptoethanol, and 10% glycerol.

Aminoacylation assay

Tyrosylation of BMV TLS and control RNAs was assayed using a protocol adapted from a previous study of TYMV TLS^{Val} (*21*). Briefly, the RNAs were folded by heating and cooling in a buffer containing 10 mM $MgCl_2$, 30 mM Na^+, and 50 mM MOPS pH 7.0, as described above. The RNAs were diluted to 10 X the final reaction concentrations (0 – 100 nM; Figure 1B). Prior to the aminoacylation experiment, the TyrRS was exchanged into a buffer containing 50 mM Tris pH 8.0, 5 mM $MgCl_2$, 1 mM TCEP, and 5 % glycerol using a 30 kDa cutoff centrifugal filter (Amicon) and diluted to 2 μM. The reaction buffer consisted of 2 mM ATP, 30 mM KCl, 5 mM $MgCl_2$, 5 mM DTT, and 30 mM HEPES-KOH pH 8.0. The total volume of each aminoacylation reaction was 20 μl. Reaction buffer, TyrRS solution (final concentration of TyrRS was 200 nM), and folded RNA solution (final concentration of RNA was 0 – 100 nM) were mixed and incubated in ice. The aminoacylation reaction was initiated by the addition of 1 μl of ^{3}H-labeled L-tyrosine (40 – 60 Ci/mmol) and followed by incubation at 30 °C for 30 minutes. The reactions were then diluted by adding 200 μl of wash buffer (20 mM Bis-tris pH 6.5, 10 mM NaCl, and 1 mM $MgCl_2$, and trace amounts of bromophenol blue dye) and immediately deposited in a well of a vacuum filter blotting apparatus. The reaction solution was filtered through a layer of Tuffryn membrane (PALL Life Sciences), Hybond positively charged membrane (GE Healthcare), and gel dryer filter paper (BioRad). The negatively charged RNA binds to the Hybond membrane.

Immediately after the diluted reaction solution ran through the filter membranes, 600 µl of wash buffer was added to the well and allowed to flow through the membranes. The Hybond membrane was allowed to dry and the locations marked by the bromophenol blue from the wash buffer were cut out. A scintillation counter (Perkin-Elmer Tri-Carb 2910 TR) was used to measure the radioactivity of the Hybond membrane pieces.

Bioinformatic analysis of TLS^{Tyr} variants

To build a sequence and secondary structure consensus model of TLS^{Tyr}, we used a preliminary alignment of the TLS^{Tyr} class obtained from the Rfam database (family RF01084) as a starting point, which contained eight sequences lacking the 5´ end of the structure corresponding to stems E and B3 stems (*45, 46*). These eight sequences were manually expanded to include the additional base pairing elements based on published secondary structure models for BMV TLS. Homology searches were performed using Infernal (*51*) to query a database of positive sense RNA viral sequences, resulting in 512 unique sequences derived from eleven distinct viruses, all in the *Bromoviridae* family (Supplemental Files 1 & 2). We evaluated the resulting alignment for nucleotide sequence conservation, as well as significant covariation of base pairing elements, using RNA Structural Covariation Above Phylogenetic Expectation (R-scape) (*52*) analysis to generate a consensus sequence and secondary structure model of TLS^{Tyr} RNAs.

SHAPE chemical probing

A published protocol was used to probe the secondary structure of representative TLS^{Tyr} variants using *N*-methyl isatoic anhydride (NMIA) as a chemical modifier (*53*). Briefly, the RNAs were folded to their native conformations using the steps described above and reacted for 30 minutes at 23 ^{o}C with either NMIA or DMSO (control). The reactions contained 3 mg/ml NMIA or DMSO, 50 mM HEPES-KOH pH 8.0, 10 mM $MgCl_2$, 60 mM RNA, and 3 nM fluorescently labeled DNA primer for reverse transcription. The reactions were then quenched by adding NaCl to a final concentration of 500 mM and Na-MES pH 6.0 to final concentration of 50 mM. Magnetic beads with poly-dT strands complementary to a poly-A sequence contained within the DNA primer were used to purify the RNA using a magnetic stand (Invitrogen Poly(A)Purist MAG Kit). After purification, the RNAs were resuspended in water to ~ 240 nM. The RNAs were then reverse transcribed using SuperScript III (Invitrogen) by incubating at 50 ^{o}C for 45 minutes and degraded by the addition of NaOH to a final concentration of 200 mM followed by incubation for 5 minutes at 90 ^{o}C. The reactions were quenched by addition of an acidic solution (final concentrations: 250 mM sodium acetate pH 5.2, 250 mM HCl, 500 mM

NaCl). To produce RNA ladders, separate reverse transcriptions were performed with each of four ddNTPs. The remaining DNA was then purified using the magnetic stand and eluted in a solution containing HiDi formamide (ThermoFisher) and trace amounts of GeneScan 500 ROX Dye Size Standard (ThermoFisher). Capillary electrophoresis (Applied Biosystems 3500 XL) was used to probe the remaining fluorescently labeled DNA products. Data analysis–including background subtraction and signal normalization using the reactivity of flanking hairpins common to all constructs–was performed using the HiTrace RiboKit computational toolkit (https://ribokit.github.io/HiTRACE/), as described previously (*54-56*).

cryo-EM grid preparation and imaging

Cryo-EM grids with RNA samples were prepared using a standard procedure. Briefly, 3µl of folded RNA solution (final concentrations: 10 – 25 µM RNA, 10 mM $MgCl_2$, 30 mM Na^+, 50 mM MOPS pH7) was deposited on C-Flat holey carbon grids (hole size: 1.2 µm; spacing: 1.3 µm; mesh: 400; manufacturer: VWR) and a FEI Vitrobot Mark IV was used to plunge freeze the grids. Our initial studies were performed using a ThermoFisher Talos Arctica transmission electron microscope (TEM) operating at an accelerating voltage of 200 kV and equipped with a Gatan K3 Summit direct electron detector (DED). Leginon software was used to collect the data. For BMV TLS, we collected 2389 movies (50 frames) at a magnification of 45,000 (raw pixel size: 0.899 Å) with defocus ranging from –1 to –2.5 µm. The spherical aberration was 2.7 mm. The total dose was ~60 e/Å^2. For BMV_TLS$_{B3ext_Cshort}$ we collected 2494 movies (60 frames) with a total dose of ~75 e/Å^2 and for BMV_TLS$_{Dext_B2short}$ we collected 505 movies (60 frames) with a total dose of ~75 e/Å^2. The magnification, defocus range, and spherical aberration was the same for all constructs. To obtain a higher resolution map of BMV TLS we used a ThermoFisher Krios TEM equipped with a Falcon 3 DED and a Bioquantum K3 imaging filter at the Pacific Northwest Cryo-EM Center. We collected 4463 movies (28 frames) with a total dose of 30 e/Å^2, raw pixel size of 0.413 Å, and a defocus range of –1 to –2.5 µm.

Cryo-EM data analysis

The imaging data was processed using cryoSPARC (*57*). The general procedure for obtaining 3D maps from the imaging data is summarized in Figures S2 & S5. The data analysis process was essentially the same for all constructs. Briefly, movies were imported into cryoSPARC, and motion correction ('Patch motion correction') and CTF estimation ('Patch CTF estimation') were performed using the built-in applications within the cryoSPARC platform using default parameters. Exposures were curated to eliminate those with obvious

defects such as damaged grid areas or excessive ice contamination. Two different methods of picking particles were used, both of which produced identical results. In the first method, we manually picked ~ 3000 particles and created templates for automated particle picking using 2D classification (Figure S2). These templates were then used to pick particles in all exposures. In the second method, we used the 'Blob Picker' application (the particle diameter range was set to 70 – 150 Å) on a subset of exposures and performed multiple rounds of 2D classification to eliminate wrong picks and generate templates for automated particle picking (Figure S5). We also tested Relion (*58*) for this task, but the quality of the 2D classes was significantly worse than what we obtained with cryoSPARC. The picked particles were then used to generate *ab initio* 3D reconstructions. A fraction of the particles generated 3D volumes that did not display recognizable RNA features and were not interpretable (Figure S2 and S3); further analysis of these particles suggested that they corresponded to high-contrast noise, damaged, and/or misfolded particles (Figure S3). A main 'good' volume displayed clear RNA secondary structural features and was refined further. Initial refinement was performed by reclassifying the particles using the 'heterogenous refinement' function within cryoSPARC. The 'bad' volumes generated previously by the *ab initio* reconstructions were used as 'sinks' for several rounds of particle 3D reclassification. At each round, the particles that were grouped into the 'good' RNA volume were kept and submitted for another round of 3D classification. Although the global features of the RNA volume remain unchanged during each round of classification, the resolution was incrementally better. The final result was the same regardless of the number of *ab initio* reconstructions we began with. The particles remaining from the several rounds of 3D classification and the refined volume were then used in a 'homogenous refinement' to generate maps that ranged between 5 and 7 Å in resolution. For the BMV TLS data obtained with the 300 kV TEM microscope, we performed an additional 'non-uniform refinement' in cryoSPARC and a 'density modification' analysis in Phenix (*59*) that resulted in maps with 4.6 and 4.2 Å resolution, respectively (Figure S6). Stem B3 in modified construct BMV TLS_{B3ext_Cshort}, which was 5 bp longer than the original stem, displayed great flexibility and was difficult to visualize. To better visualize this stem, we performed a '3D Variability Analysis' in cryoSPARC that performs a principal component analysis to divide a map into an ensemble and this analysis better revealed the extended B3.

Structural modeling

To build a structural model of BMV TLS, we began with the helical assignments discussed in the main text and the published computational model (*25*), which agreed for the most part with the known secondary structure of BMV TLS. Each helical domain was extracted from the published model and docked into the corresponding density using UCSF Chimera (*60*). Manual modelling was performed in Chimera to join the helical domains and to obtain an initial model for flexible fitting. We then performed molecular dynamics flexible fitting using the web server Namdinator (*37*). The resulting model was imported into Phenix (*61*) to perform several rounds of real-space refinements, with secondary structural constraints. ERRASER was used to correct the geometry of the model (*38, 39*). The final structure of BMV TLS displayed a MolProbity clashscore of 9.63, 0 bad bonds, 2 bad angles, 29 outlier backbone conformations, and 0 probably wrong sugar puckers. These statistics were obtained using the MolProbity web server (http://molprobity.biochem.duke.edu) (*62*) The outlier conformations were examined and were generally located in complex areas of the structure expected to have non-canonical conformations. Similar methods were used to build structural models for BMV TLS_{B3ext_Cshort} and BMV $TLS_{Dext_B2short}$ (Figure S7). Pymol (Schrödinger, Inc) and Chimera were used to build models for BMV TLS bound to yeast TyrRS and for the conformational changes of domain B3-E.

DATA DEPOSITION

Accession code for coordinates and cryo-EM maps will be supplied prior to publication.

ACKNOWLEDGMENTS

Peter Van Blerkom (Univ. of Colorado Anschutz Medical Campus Cryo-EM Facility) and Theo Humphreys (Pacific Northwest Center for Cryo-EM) assisted with microscope operation. Erik Hartwick provided assistance with the aminoacylation assays. The authors thank current and former Kieft Lab members for thoughtful discussions and technical assistance, and Anna-Lena Steckelberg, Benjamin Akiyama, and Quentin Vicens for critical reading of this manuscript and assistance with structural modelling.

Funding: This work was supported by NIH grants R35GM118070 (JSK) and F32GM139385 (SB). MES. is a Jane Coffin Childs Postdoctoral Fellow. Use of Pacific Northwest Center for Cryo-EM was supported by NIH grant U24GM129547.

Author contributions: SB conceptualized the study, performed cryo-EM and aminoacylation experiments, analyzed data, performed structural modelling, and wrote the manuscript in collaboration with JSK. MES performed bioinformatic analysis and chemical probing and edited the manuscript. AM purified proteins and edited the manuscript. JSK conceptualized the study, performed structural modelling, acquired funding, and wrote the manuscript in collaboration with SB.

Competing interests: The authors declare no competing interests.

Data and materials availability. All data is available in the manuscript and supplementary material. Maps and atomic coordinates will be deposited in public databases before publication.

FIGURES AND LEGENDS

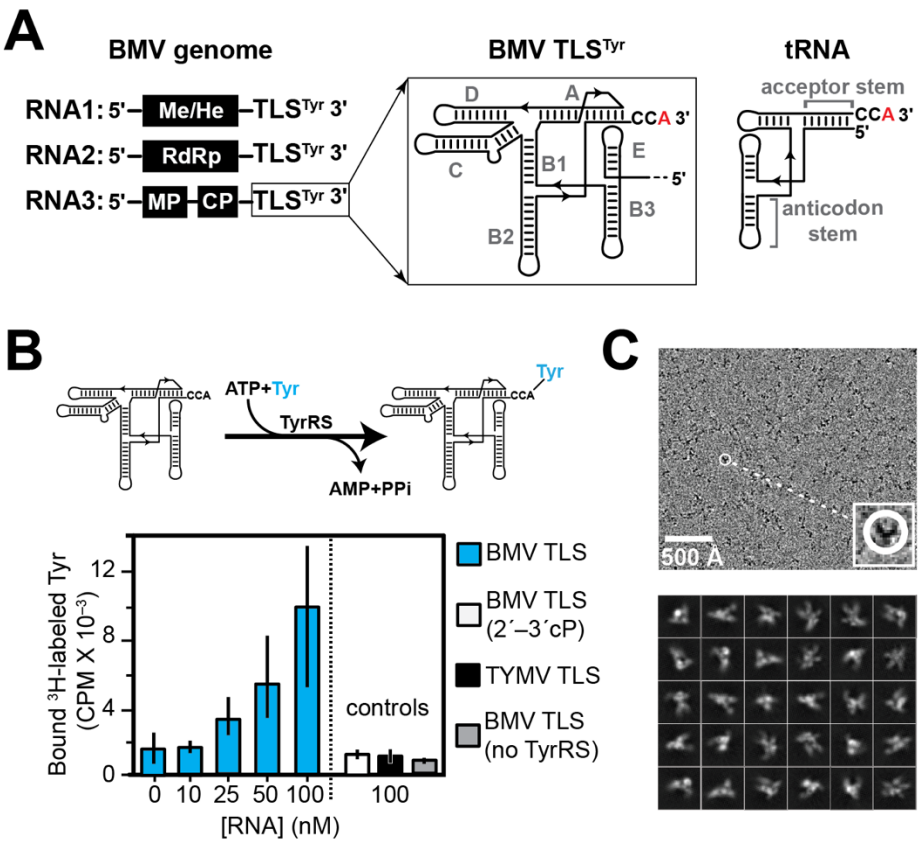

Figure 1. Functional and initial structural characterization of BMV TLS RNA. (A) Organization of the tripartite BMV genome with tyrosylatable TLS at the 3' end of each RNA. Me/He: methyltransferase/helicase; RdRp: RNA-dependent RNA polymerase; MP: movement protein; CP: coat protein. The conserved secondary structure of the BMV TLS^{Tyr} is shown next to that of canonical tRNA. Structural domains and the terminal CCA are shown, with the red A designating the site of aminoacylation. (B) Top: Cartoon representation of chemical reaction catalyzed by TyrRS. Bottom: Tyrosylation of the BMV TLS RNA used for cryo-EM studies. Incorporation of radioactive tyrosine onto RNA was measured using an adapted published protocol (*21*). BMV TLS (2'–3' cP) contains a terminal 2'–3' cyclic phosphate and therefore is not a substrate of TyrRS. (C) Top: Representative micrograph of the BMV TLS^{Tyr} RNA. For clarity, 30% of the field of view is shown. Defocus range used in our experiments was –1 to –2.5 µm. Bottom: Classified projections from particles (n = 139,010) used to build initial 3D map of BMV TLS (Figure S2).

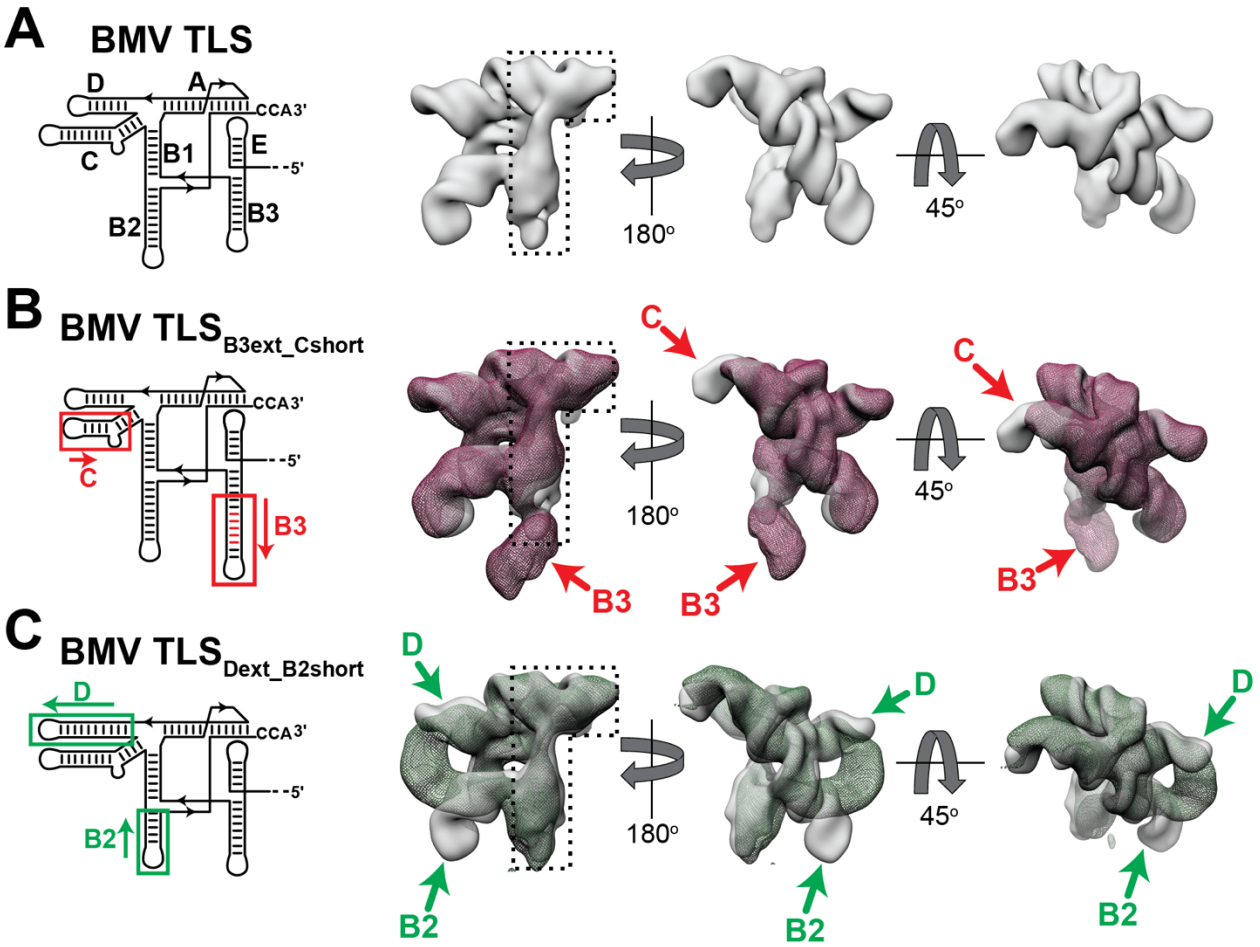

Figure 2. Cryo-EM maps of BMV TLS and helix assignments. Proposed secondary structures and cryo-EM maps of (A) BMV TLS, (B) BMV TLS_{B3ext_Cshort}, and (C) BMV $TLS_{Dext_B2short}$. The maps of modified constructs BMV TLS_{B3ext_Cshort} (red mesh) and BMV $TLS_{Dext_B2short}$ (green mesh) are superimposed on BMV TLS (solid gray). Arrows point to differences in electron density that correspond to altered stems. Box indicates L-shape density resembling tRNA. General method used to obtain maps is presented in Figure S2.

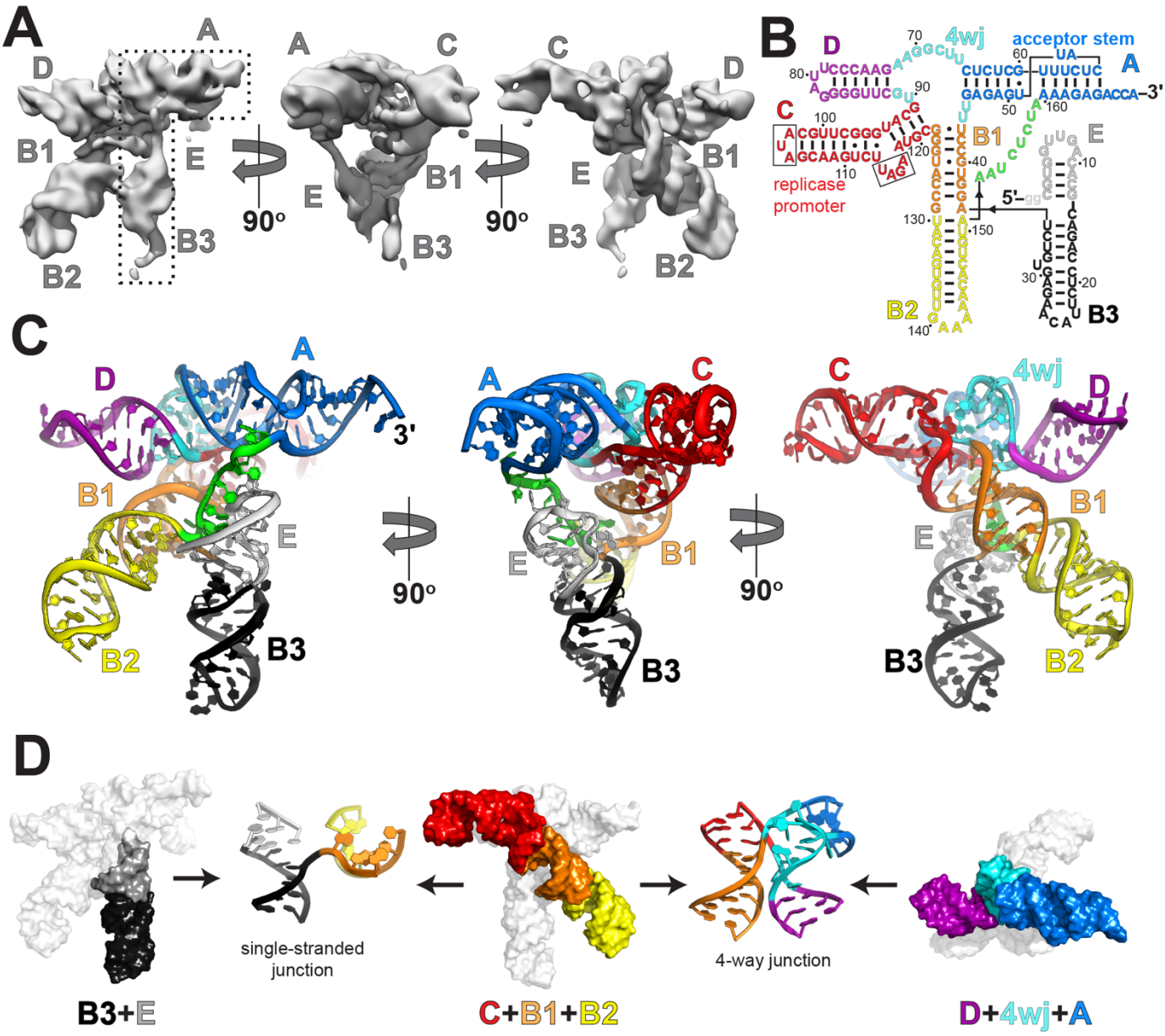

Figure 3. Overall tertiary structure of the BMV TLS^{Tyr}. (A) Three views of the cryo-EM map of BMV TLS^{Tyr} obtained with a 300 kV electron microscope, with a final overall resolution is 4.2 Å (Figure S5). Dashed box shows L-shape density resembling tRNA. (B) Secondary structure of the BMV TLS^{Tyr} construct used for cryo-EM is labeled and colored by domain, with the location of the replication promoter and the acceptor stem labeled. (C) Structural model of BMV TLS viewed from three directions. Colors match those in the secondary structure. (D) The three domains of BMV TLS^{Tyr} formed by helical stacks, with the connecting junctions shown between them. Colors match those in (B) and (C).

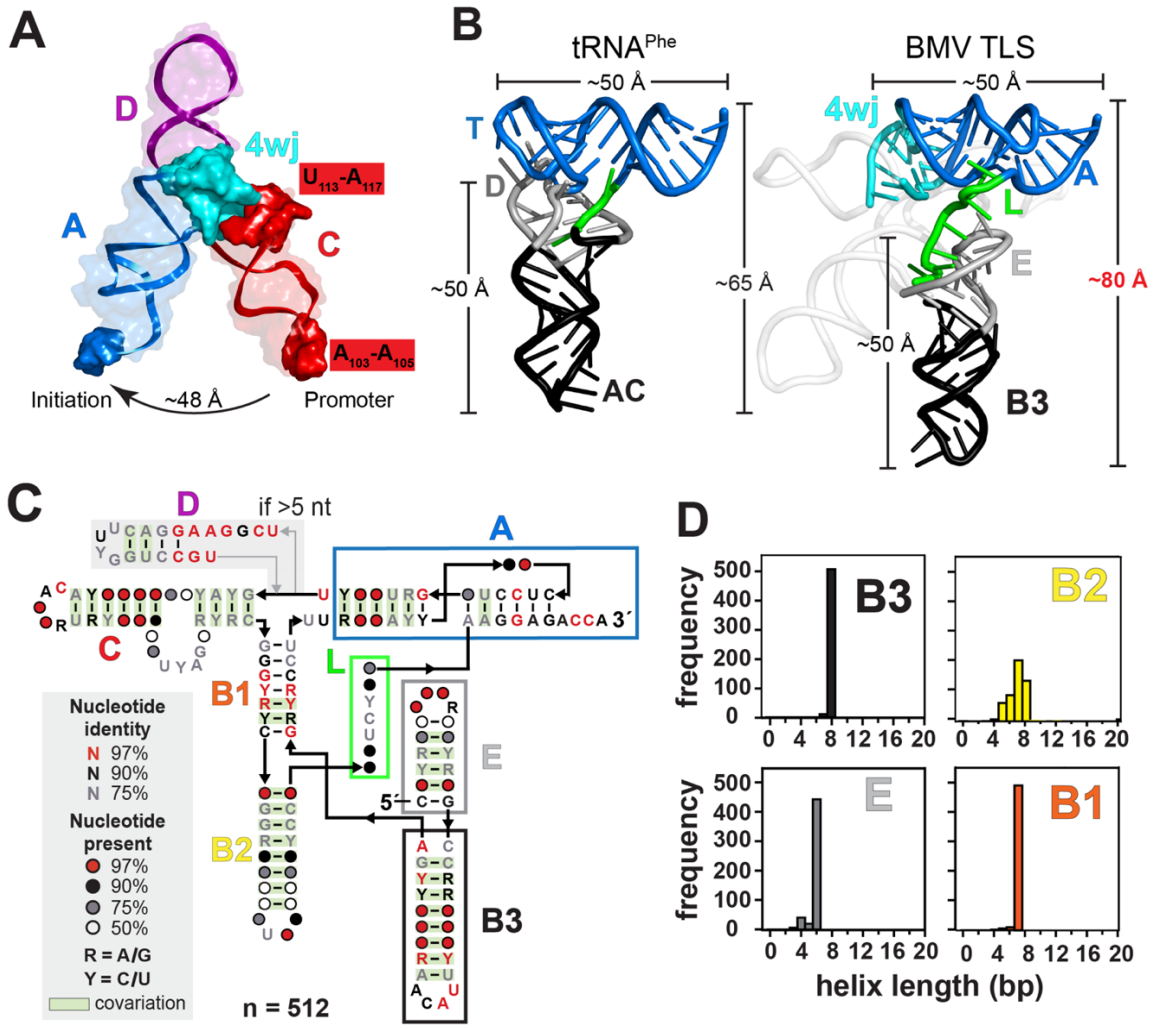

Figure 4. Structural features of BMV TLS^{Tyr} important for replication or tRNA mimicry. (A) The relative positions of stems A, C, D, and 4wj are shown, colored as in Figure 3. The locations of replication promoter signals are indicated as solid red surfaces and labeled, the replication initiation site is in solid blue, and the 4wj is in solid cyan. The distance between the critical AUA and the replication initiation site is shown. (B) Comparison of the structure of unbound tRNA^{Phe} (left) and tRNA mimicking portions of the BMV TLS^{Tyr} (right). The structurally important loops T and D are labeled in tRNA^{Phe}, as is the anticodon (AC) loop. Analogous structural features are colored as per Figure 3 and various molecular dimensions are shown. (C) Consensus covariation and sequence model of the TLS^{Tyr} from 512 sequences. Helices are labeled according to the naming in BMV TLS. The elements shown in panel (B) are boxed in matching colors. (D) Graphs of the distribution of lengths of four helices according to the structural alignment of 512 sequences. The high conservation of the lengths of E and B3 supports the role of these helices as the anticodon stem analog in BMV TLS.

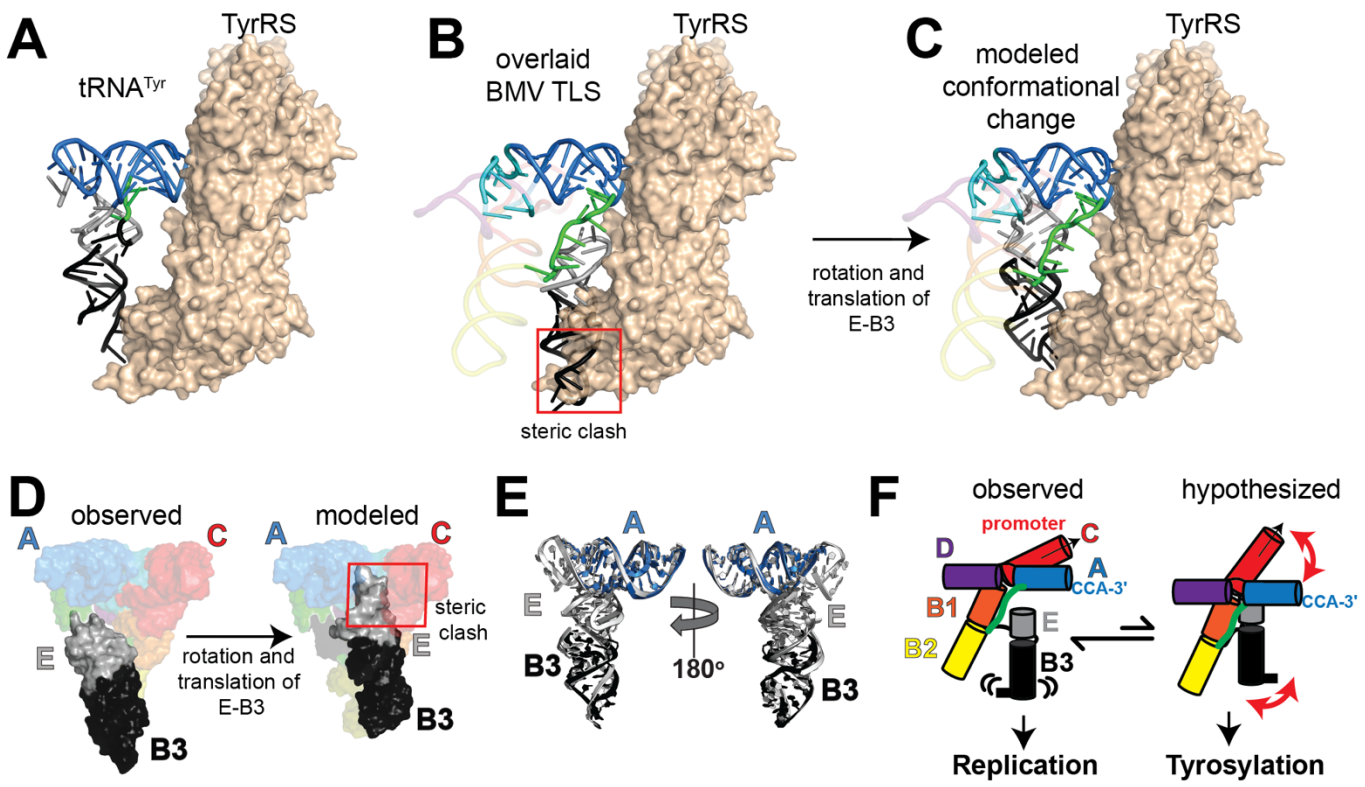

Figure 5. BMV TLS must undergo conformational changes to bind TyrRS. (A) Published crystal structure of yeast tRNA^{Tyr} bound to yeast TyrRS (*44*). (B) The cryo-EM derived BMV TLS structure overlaid on the bound tRNA^{Tyr} by superimposing the acceptor arms of each (Figure S10). The location of a steric clash between B3 (anticodon stem analog) and TyrRS is boxed in red. (C) Model of the BMV TLS^{Tyr} bound to the TyrRS after rotation and translation of the B3-E domain to match the position of the anticodon stem in tRNA (Figure S10). (D) Left: the structure of the BMV TLS with the 'ground state' position of domain B3-E. Right: modeled position of domain B3-E, in the position shown in panel (C). The movement creates a steric clash between the apical loop of E and the core of the molecule. (E) The tRNA mimicking elements of BMV TLS (i.e. stems A, E, and B3) after the conformational change modelled in (C) are superimposed on the structure of tRNA^{Tyr} (light gray). (F) Cartoon of the conformational change needed to transition from the observed ground state ('replication-ready') to the hypothesized conformation that can bind TyrRS ('tyrosylation-ready'). This hypothesized conformational change involves large rearrangement of domain B3-E and reorganization of stem C, along with smaller rearrangements of other regions of the TLS.

SUPPLEMENTAL FIGURES

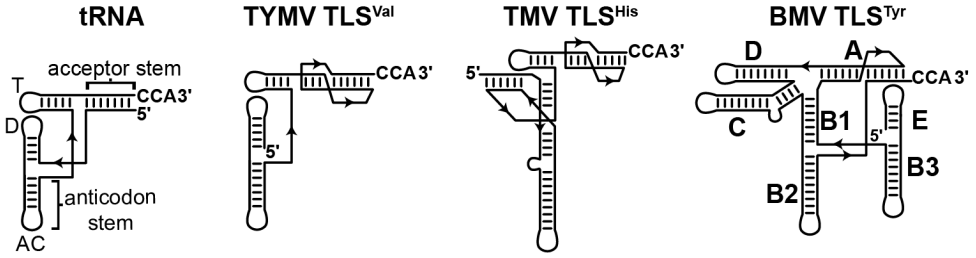

Figure S1. Comparison of the secondary structures of tRNA and TLS from different classes. TLSs are classified based on their aminoacylation specificities. Three classes have been identified in plant viruses: valylatable (TLS^{Val}), histidylatable (TLS^{His}), and tyrosylatable (TLS^{Tyr}) (*16*). Cartoon secondary structures of representative variants from each class are shown. These are turnip yellow mosaic virus (TYMV), tobacco mosaic virus (TMV), and brome mosaic virus (BMV). The canonical structural features of tRNA are labeled: the anticodon and acceptor stems and the T, D and anticodon (AC) loops. The helical stems of BMV TLS are labeled with the names assigned in previous studies (*25*).

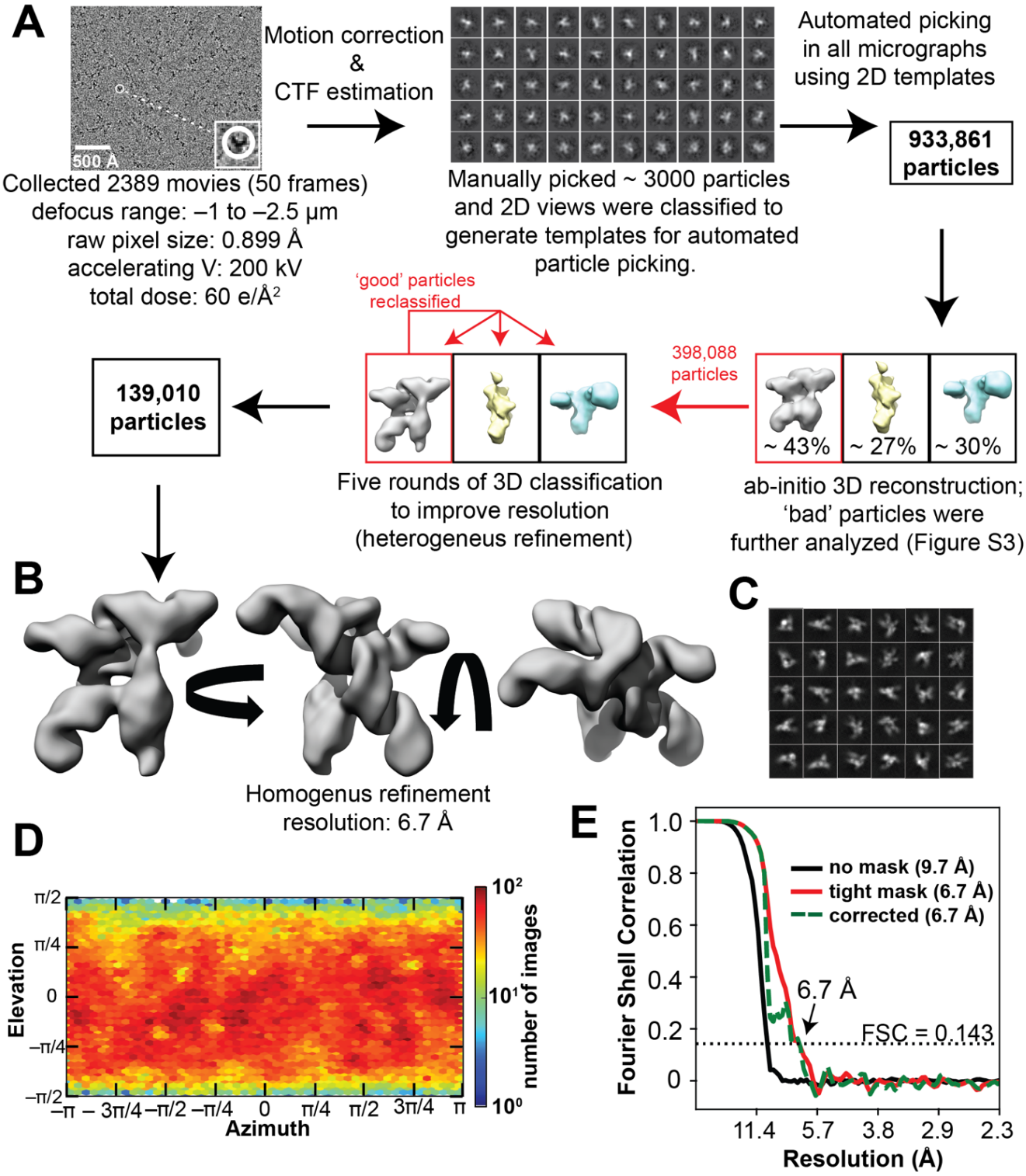

Figure S2. Analysis of data collected with 200 kV electron microscope. (A) Steps followed to generate cryo-EM reconstruction of BMV TLS using cryoSPARC (*57*). After particle picking, several rounds of 3D classification were performed to increase the resolution of the final map. Volumes created with 'bad' particles that were not used in the *ab initio* reconstruction to generate main volume were used as 'sinks' for iterative 3D classification. Further analysis of these bad particles did not result in interpretable maps (Figure S3), suggesting that they may correspond to damage and/or misfolded particles. (B) Map generated by homogeneous refinement in cryoSPARC. (C) Classified 2D projections of particles used for building 3D model in (B). (D) Distribution of viewing angles in 3D reconstruction showed no strong biases in particle orientation. (E) Estimation of overall resolution of cryo-EM map using half maps and gold standard FSC of 0.143.

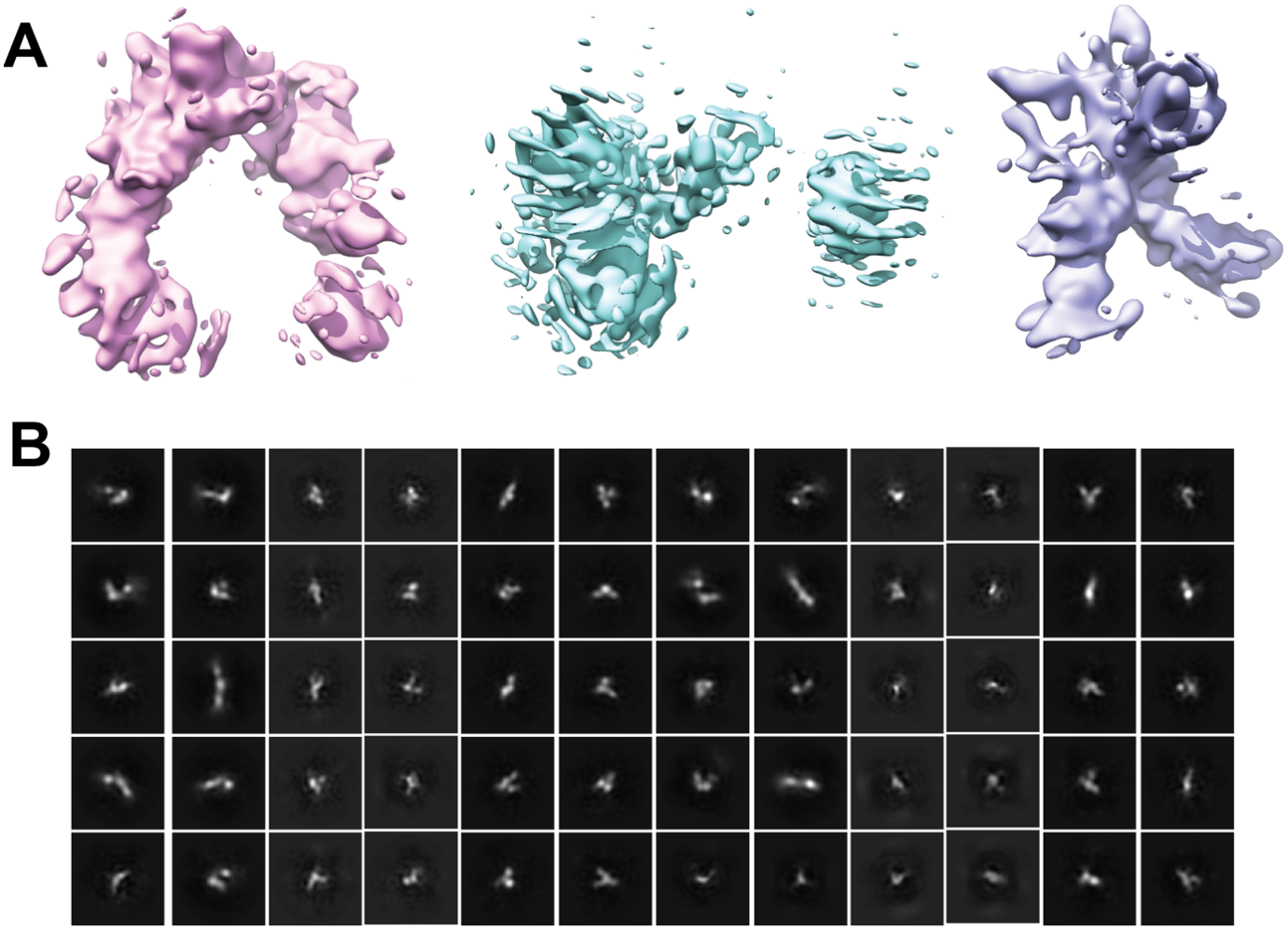

Figure S3. Analysis of 'bad' particles that were excluded from BMV TLS modelling during *ab initio* reconstruction. (A) Refinement of 3D reconstructions built from bad particles. The bad particles were used to generate three *ab initio* reconstructions and we attempted to refine them using the same process (and multiple variations of it) used to refine the final map of BMV TLS. The resolution did not improve, and the resulting maps did not show recognizable RNA features and were inconsistent with folded BMV TLS. (B) Classified 2D projections of bad particles. Many of these particles show features that are inconsistent with genuine BMV TLS, such as discontinuities (more than one particle picked), elongation (possible misfolded or aggregated particles), or small size (possibly high contrast noise). A few particles resemble genuine BMV TLS and may have been misclassified in the *ab initio* reconstruction.

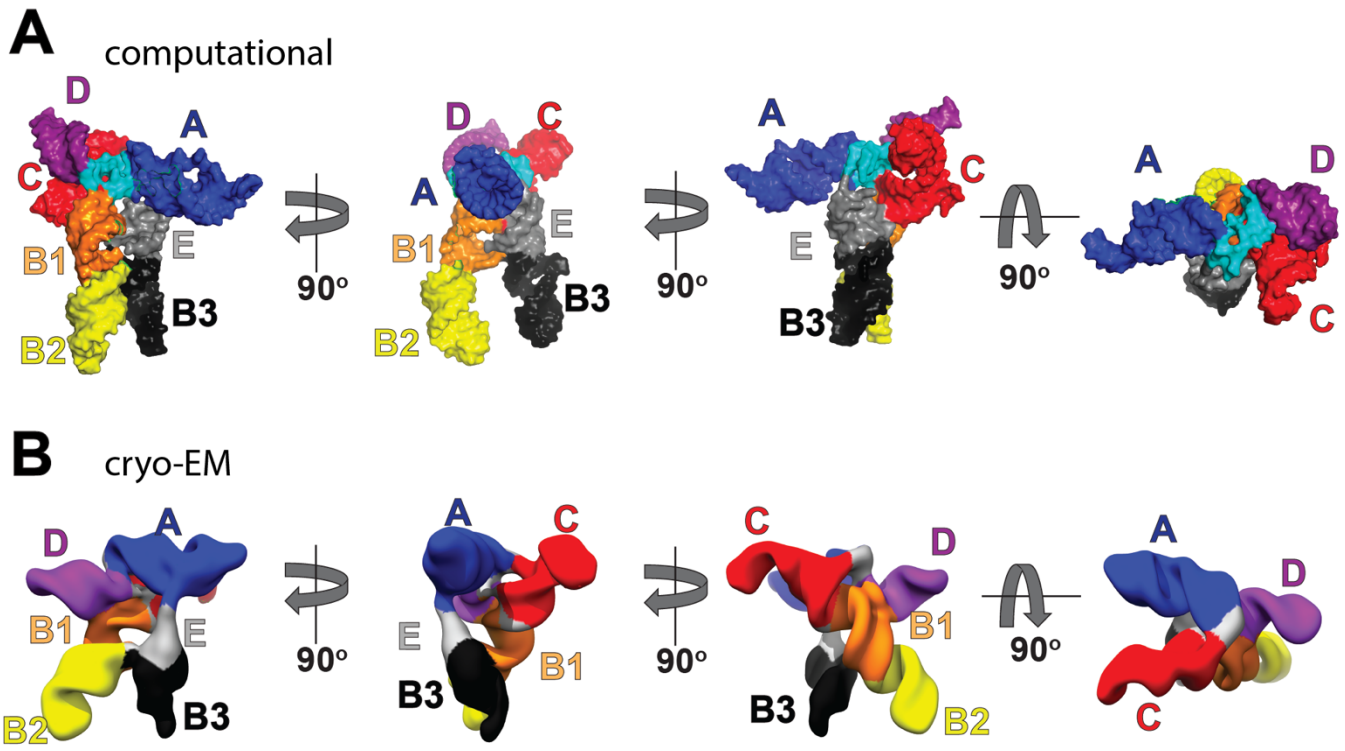

Figure S4 Comparison between cryo-EM map and published computational model. Helices are colored according to secondary structure in Figure 3 of the main text. (A) Published computational model was built using chemical and enzymatic probing data, and phylogenetic secondary structure comparisons (*25*). (B) The map obtained by cryo-EM is oriented similarly to the computational model in (A) for comparison. Although the global architectures of the cryo-EM map and the computational model agree qualitatively, the location and orientation of the helices differ.

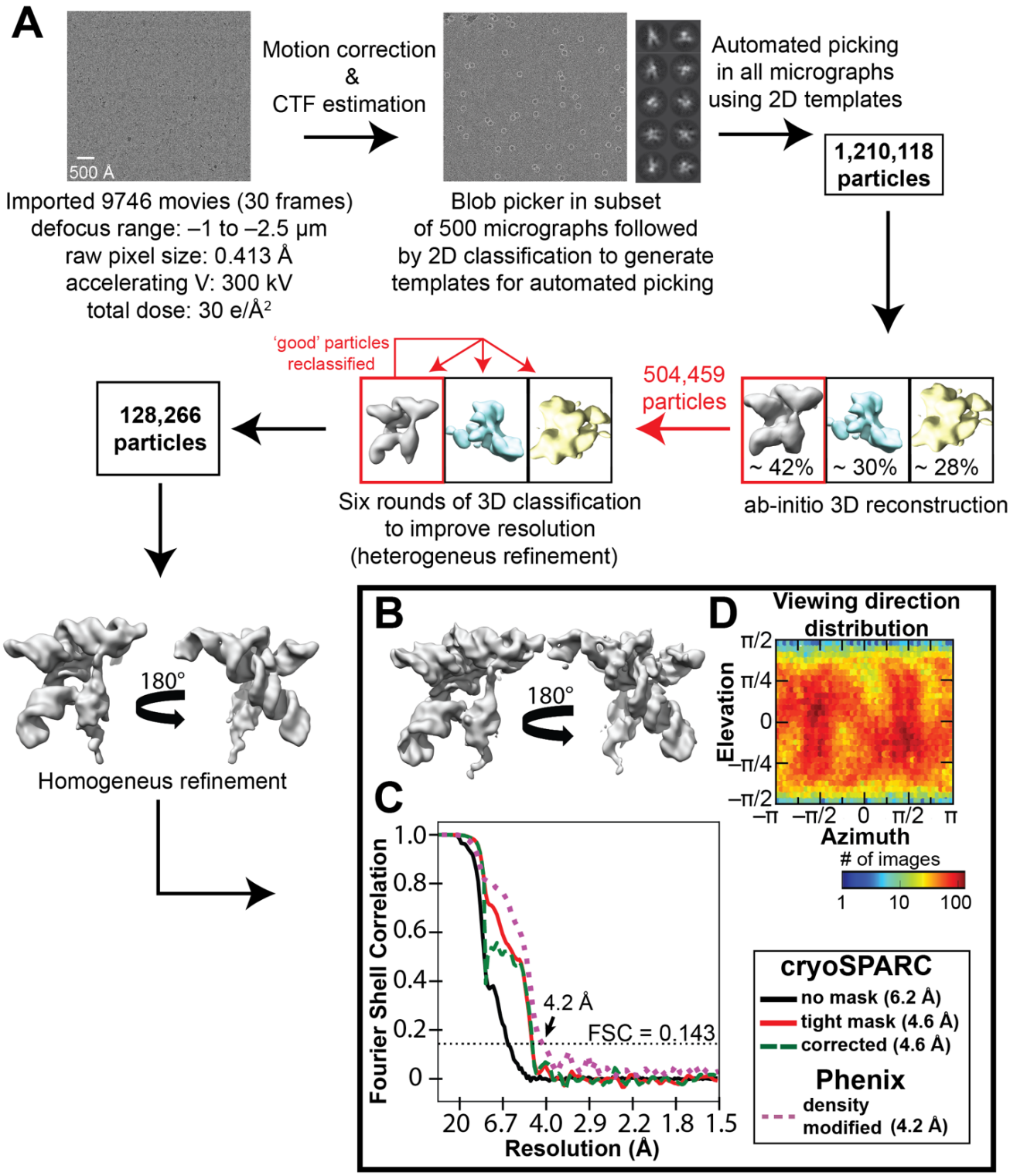

Figure S5. Analysis of data collected with 300 kV electron microscope. (A) Steps followed to generate cryo-EM reconstruction of BMV TLS using cryoSPARC (*57*). After particle picking, several rounds of 3D classification were performed to increase the resolution of the final map. Volumes created with particles that were not used in the *ab initio* reconstruction to generate main volume were used as 'sinks' for iterative 3D classification. (B) Map generated using non-uniform refinement application in cryoSPARC. The overall resolution of this map was 4.6 Å. (C) Estimation of overall resolution of cryo-EM maps using half maps and gold standard FSC of 0.143. Map reconstruction from cryoSPARC was processed using density modification application in Phenix to produce map with final resolution of 4.2 Å (*59*). (D) Distribution of viewing angles.

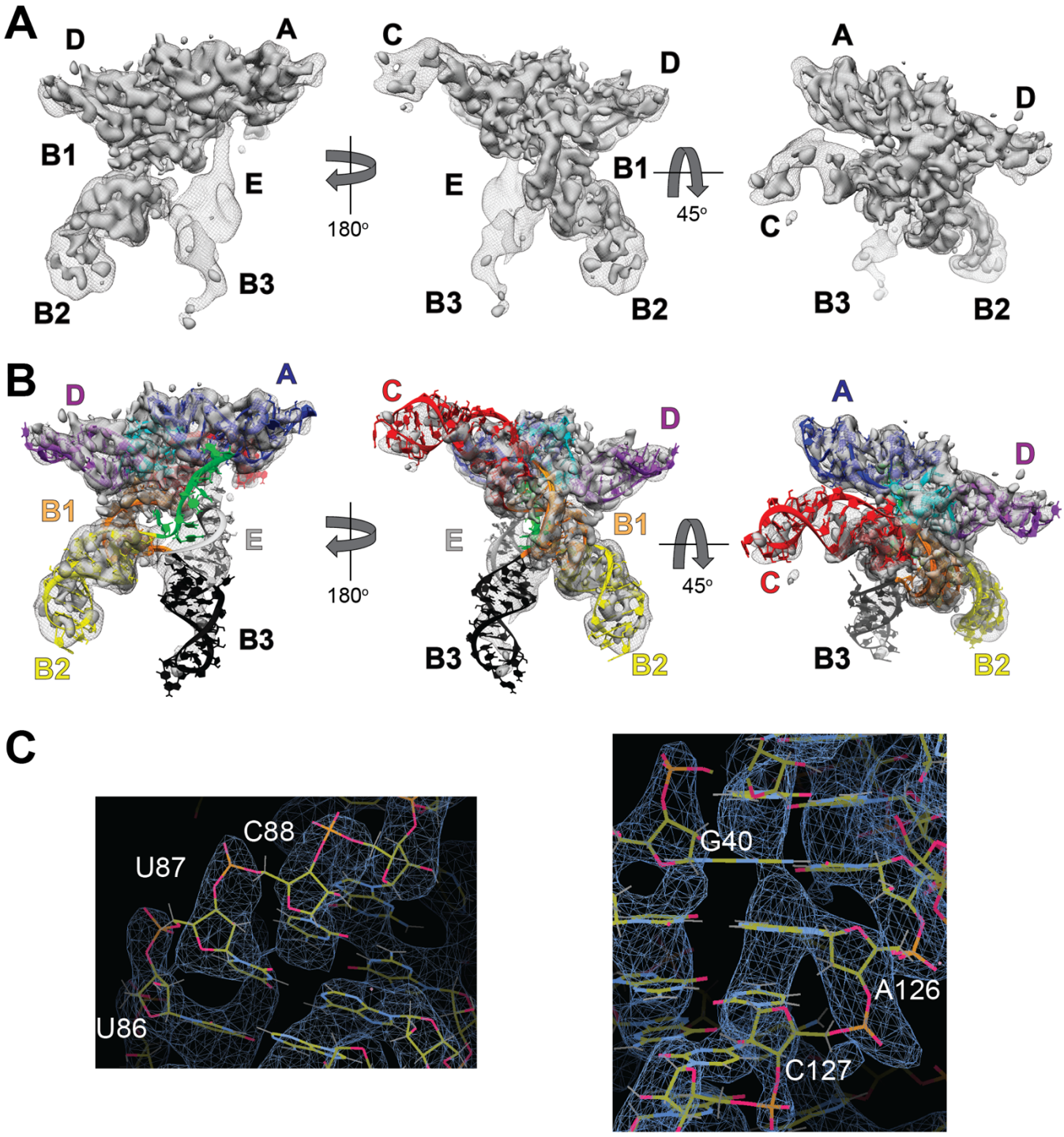

Figure S6. Model fit to density. (A) Two maps are shown: the map obtained from cryoSPARC (4.6 Å resolution) and the final map obtained after performing density modification in Phenix (4.2 Å resolution) (*59*). (B) Structural model of BMV TLS fitted to cryo-EM density. (C) Examples of density displaying phosphate 'bumps' and base stacking. Figures were generated in Coot using the 4.2 Å resolution map (level: 0.5159 V; rmsd: 4.9) (*63*)

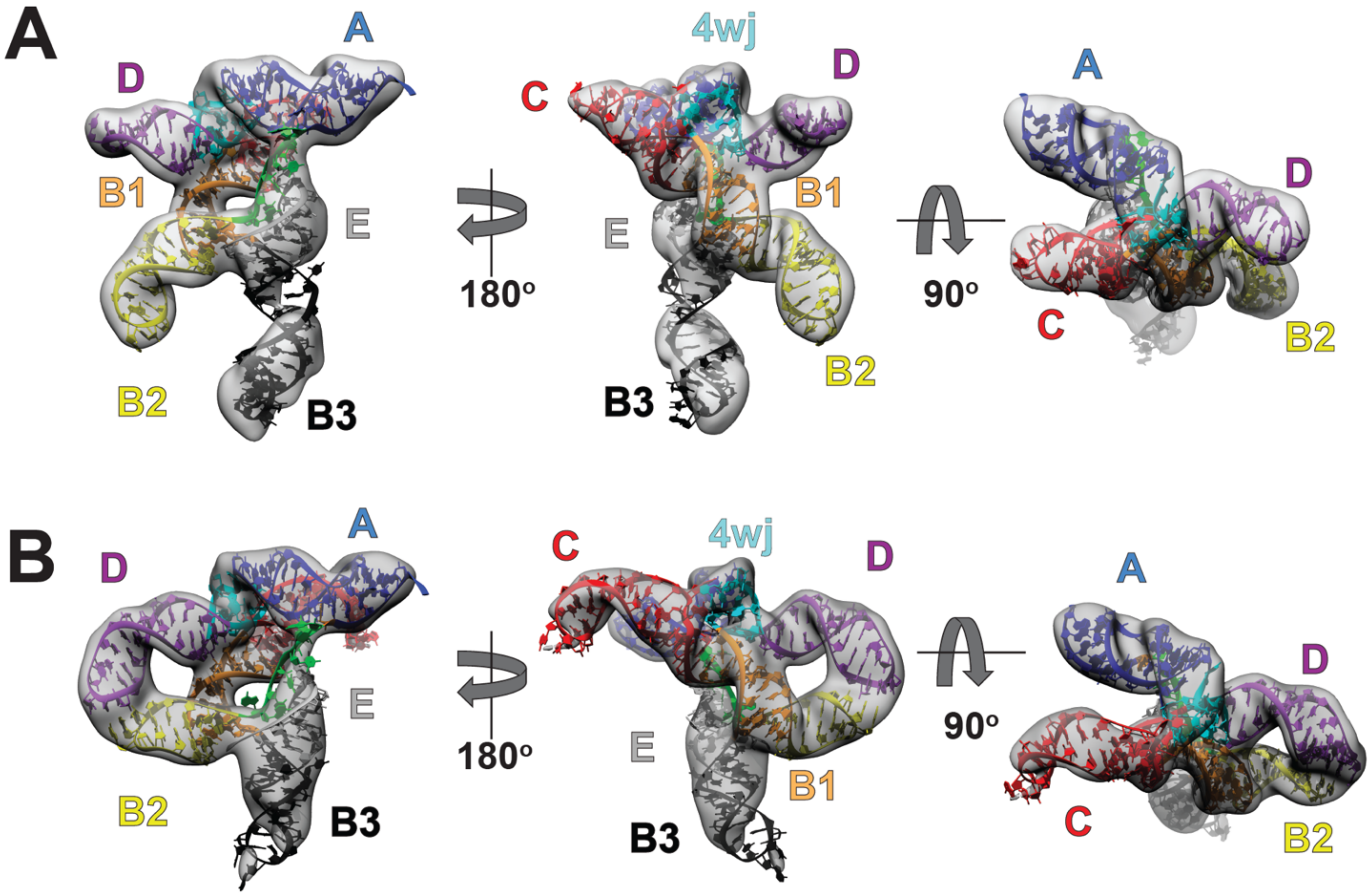

Figure S7. Structural models of modified constructs support architecture of BMV TLS. (A) Density of BMV TLS_{B3ext_Cshort}, with the structural model docked inside. In this model, extra base-pairs have been added to stem B3 and base pairs have been removed from stem C. The density matches a model with these altered helical lengths. (B) Same as panel A, but with BMV $TLS_{Dext_B2short}$. The lengthening of stem D and the shortening of B2 in the model match the density. Unexpected, these changes induced a new tertiary contact between the apical loops of B2 and D, apparently due to a fortuitous alignment of the bases in their apical loops.

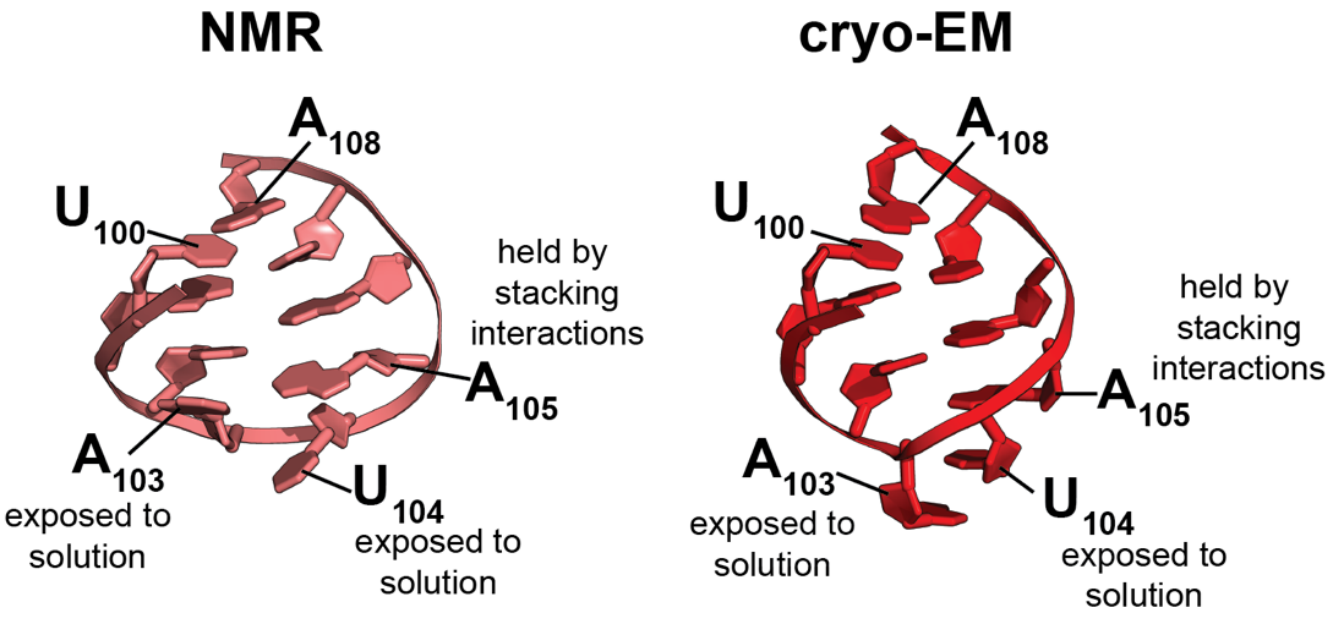

Figure S8. Comparison between NMR and cryo-EM models of the AUA apical loop of stem C. A published NMR structure (*41*) of the UAU triloop within a minimal isolated helix reported qualitative features that are also shown in our cryo-EM model. The 5´ adenine of the triloop is exposed to solution and the 3´ adenine is held by stacking interactions with the closing base pairs. The detailed configuration of the triloop in the cryo-EM structure is ambiguous as the local resolution of the cryo-EM maps at the apical loops was relatively low.

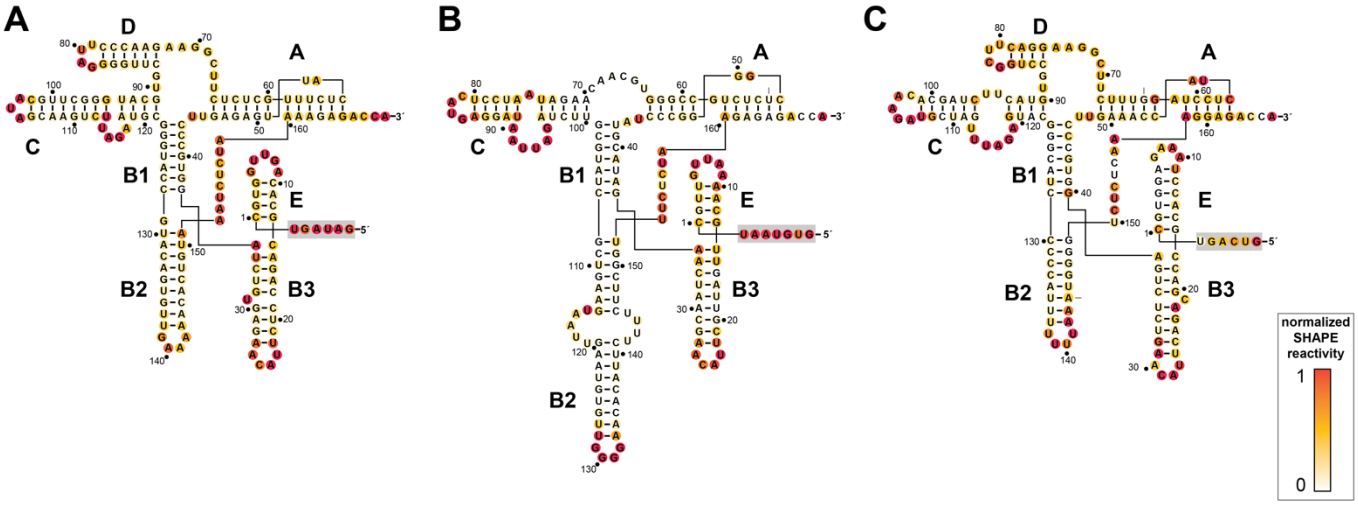

Figure S9. Chemical probing of representative TLS^{Tyr} variants supports secondary structure consensus model. SHAPE probing of TLS^{Tyr} belonging to three different viruses: (A) BMV, (B) pelargonium zonate spot virus, and (C) cucumber mosaic virus. Helices were labeled using the BMV TLS naming. Structures were drawn to match the consensus sequence and secondary structure model in the main text (Figure 4). Variant from pelargonium zonate spot virus lacks a stem D. 5´ sequences in grey box were added for chemical probing and are not part of the model.

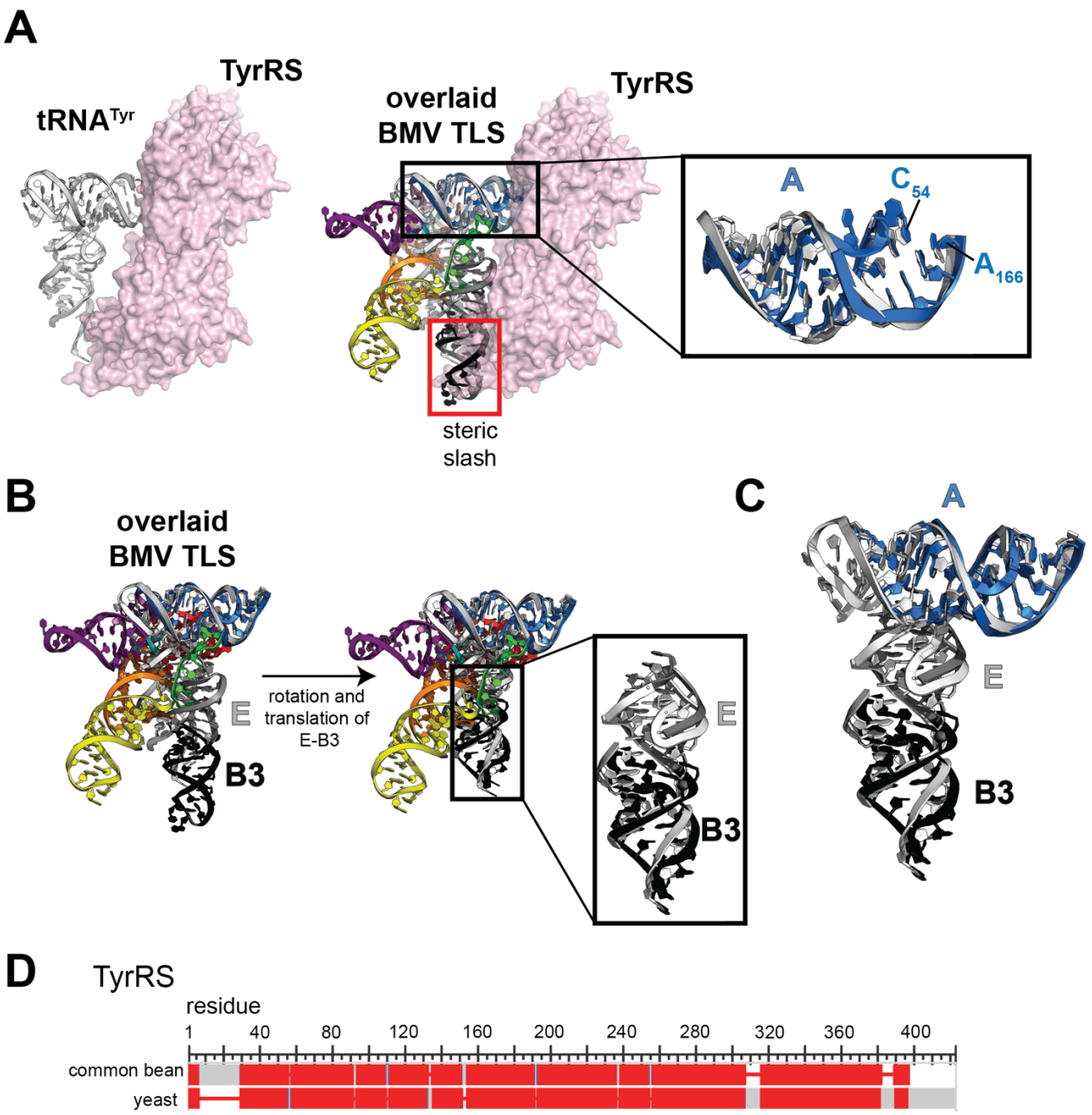

Figure S10. Structural modelling of BMV TLS bound to yeast TyrRS and conformational changes. (A) Left: Published structure of yeast $tRNA^{Tyr}$ bound to its cognate TryRS (*44*). Right: Identical structure to panel (A), with the BMV TLS superimposed on the $tRNA^{Tyr}$ by aligning the structures of the acceptor stems. Boxed: Close-up of the aligned acceptor stems, with tRNA in grey and the BMV TLS stem A in blue. The alignment was performed prioritizing base pairs near the terminal CCA, as this region is known to be critical for tyrosylation (*43, 44*). Selected nucleotides are labeled for reference. (B) Left: Overlaid ground state BMV TLS and $tRNA^{Tyr}$ from (A) with the TyrRS removed. Rotation and translation of domain B3+E leads to the model to the right, in which domain B3+E superimposes remarkably well with the analogous portion of $tRNA^{Tyr}$ (boxed). (C) Superposition of $tRNA^{Tyr}$ with the modeled tRNA-mimicking portions of the BMV TLS after the modeled translation and rotation. The two molecules show remarkable similarity after this proposed change in the position of conformationally dynamic domain B3+E. (D) Sequence alignment of TyrRS from BMV model host *Phaseolus vulgaris* (common bean) and yeast performed using the online NCBI tool COBALT (https://www.ncbi.nlm.nih.gov/tools/cobalt/re_cobalt.cgi). Red is strong conservation, blue is weaker conservation, and grey is gap in conservation.

Table S1. Statistics of BMV TLS structure obtained using the MolProbity web server (http://molprobity.biochem.duke.edu) (*62*)

Clashscore, all atoms	9.63
Bad bonds	0 of 4021
Bad angles	2 of 6263
Outlier backbone conformations	29 of 169
Probably wrong sugar puckers	0 of 169

Table S2. Sequences of constructs used. Sequence in bold is T7 promoter. Lower case 'gg' was added to facilitate *in vitro* transcription. Sequence underlined corresponds to HDV ribozyme. Sequences were ordered from Integrated DNA Technologies and synthesized as gBlocks.

Construct	Sequence
BMV TLS	GCGAGTGAATTC**TAATACGACTCACTATA**ggCGTG GTTG ACAC GCAG ACCT CTTA CAAG AGTG TCTA GGTG CCTT TGAG AGTT ACTC TTTG CTCT CTTC GGAA GAAC CCTT AGGG GTTC GTGC ATGG GCTT GCAT AGCA AGTC TTAG AATG CGGG TACC GTAC AGTG TTGA AAAA CACT GTAA ATCT CTAA AAGA GACC A
BMV TLS_{B3ext_Cshort}	GCGAGTGAATTC**TAATACGACTCACTATA**ggCGTGGTTGA CACGCAGACCTAGTACCTTACAAGGTACTAGTGTCTAGG TGCCTTTGAGAGTTACTCTTTGCTCTCTTCGGAAGAAC CCTT AGGG GTTC GTGC ATGG GCAT AGCC TTAG AATG CGGG TACC GTAC AGTG TTGA AAAA CACT GTAA ATCT CTAA AAGA GACC A
BMV $TLS_{Dext_B2short}$	GCGAGTGAATTC**TAATACGACTCACTATA**ggCGTG GTTG ACAC GCAG ACCT CTTA CAAG AGTG TCTA GGTG CCTT TGAG AGTT ACTC TTTG CTCT CTTC GGAA GAAC AGUUCCTT AGGGAACU GTTC GTGC ATGG GCTT GCAT AGCA AGTC TTAG AATG CGGG TACC GTAC A TGA AAA T GTAA ATCT CTAA AAGA GACC A
BMV TLS (2'–3' cP)	GCGAGTGAATTC**TAATACGACTCACTATA**ggCGTG GTTG ACAC GCAG ACCT CTTA CAAG AGTG TCTA GGTG CCTT TGAG AGTT ACTC TTTG CTCT CTTC GGAA GAAC CCTT AGGG GTTC GTGC ATGG GCTT GCAT AGCA AGTC TTAG AATG CGGG TACC GTAC AGTG TTGA AAAA CACT GTAA ATCT CTAA AAGA GACC A <u>GGGC GGCA TGGT CCCA GCCT CCTC GCTG GCGC CGCC TGGG CAAC ATGC TTCG GCAT GGCG AATG GGACC</u>GGATCCGCGAT

REFERENCES

1. L. R. Ganser, M. L. Kelly, D. Herschlag, H. M. Al-Hashimi, The roles of structural dynamics in the cellular functions of RNAs. *Nat Rev Mol Cell Biol* **20**, 474-489 (2019).
2. D. Herschlag, S. Bonilla, N. Bisaria, The Story of RNA Folding, as Told in Epochs. *Cold Spring Harb Perspect Biol* **10**, (2018).
3. Z. A. Jaafar, J. S. Kieft, Viral RNA structure-based strategies to manipulate translation. *Nat Rev Microbiol* **17**, 110-123 (2019).
4. B. M. Akiyama, D. Eiler, J. S. Kieft, Structured RNAs that evade or confound exonucleases: function follows form. *Curr Opin Struct Biol* **36**, 40-47 (2016).
5. K. Lu, X. Heng, M. F. Summers, Structural determinants and mechanism of HIV-1 genome packaging. *J Mol Biol* **410**, 609-633 (2011).
6. B. M. Akiyama *et al.*, Zika virus produces noncoding RNAs using a multi-pseudoknot structure that confounds a cellular exonuclease. *Science* **354**, 1148-1152 (2016).
7. J. M. Watts *et al.*, Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* **460**, 711-716 (2009).
8. K. Lu *et al.*, NMR detection of structures in the HIV-1 5'-leader RNA that regulate genome packaging. *Science* **334**, 242-245 (2011).
9. C. P. Jones, W. A. Cantara, E. D. Olson, K. Musier-Forsyth, Small-angle X-ray scattering-derived structure of the HIV-1 5' UTR reveals 3D tRNA mimicry. *Proc Natl Acad Sci U S A* **111**, 3395-3400 (2014).
10. D. Yang, J. L. Leibowitz, The structure and functions of coronavirus genomic 3' and 5' ends. *Virus Res* **206**, 120-133 (2015).
11. R. Madhugiri *et al.*, Structural and functional conservation of cis-acting RNA elements in coronavirus 5'-terminal genome regions. *Virology* **517**, 44-55 (2018).
12. O. Fernandez-Miragall, S. Lopez de Quinto, E. Martinez-Salas, Relevance of RNA structure for the activity of picornavirus IRES elements. *Virus Res* **139**, 172-182 (2009).
13. A. E. Simon, L. Gehrke, RNA conformational changes in the life cycles of RNA viruses, viroids, and virus-associated RNAs. *Biochim Biophys Acta* **1789**, 571-583 (2009).
14. X. Shi, S. Bonilla, D. Herschlag, P. Harbury, Quantifying Nucleic Acid Ensembles with X-ray Scattering Interferometry. *Methods Enzymol* **558**, 75-97 (2015).
15. K. Murata, M. Wolf, Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochim Biophys Acta Gen Subj* **1862**, 324-334 (2018).
16. T. W. Dreher, Viral tRNAs and tRNA-like structures. *Wiley Interdiscip Rev RNA* **1**, 402-414 (2010).
17. R. Giege, Interplay of tRNA-like structures from plant viral RNAs with partners of the translation and replication machineries. *Proc Natl Acad Sci U S A* **93**, 12078-12081 (1996).
18. M. E. Sherlock, E. W. Hartwick, A. MacFadden, K. J.S., Structural Diversity and Phylogenetic Distribution of Valyl tRNA-like Structures in Viruses. *bioRxiv*, (2020).
19. J. A. Hammond, R. P. Rambo, M. E. Filbin, J. S. Kieft, Comparison and functional implications of the 3D architectures of viral tRNA-like structures. *RNA* **15**, 294-307 (2009).
20. T. W. Dreher, Role of tRNA-like structures in controlling plant virus replication. *Virus Res* **139**, 217-229 (2009).
21. E. W. Hartwick *et al.*, Ribosome-induced RNA conformational changes in a viral 3'-UTR sense and regulate translation levels. *Nat Commun* **9**, 5074 (2018).
22. R. J. Bowden, J. P. Simas, A. J. Davis, S. Efstathiou, Murine gammaherpesvirus 68 encodes tRNA-like sequences which are expressed during latency. *J Gen Virol* **78 (Pt 7)**, 1675-1687 (1997).
23. M. Vieweger, E. D. Holmstrom, D. J. Nesbitt, Single-Molecule FRET Reveals Three Conformations for the TLS Domain of Brome Mosaic Virus Genome. *Biophys J* **109**, 2625-2636 (2015).

24. V. Perret, C. Florentz, T. Dreher, R. Giege, Structural analogies between the 3' tRNA-like structure of brome mosaic virus RNA and yeast tRNATyr revealed by protection studies with yeast tyrosyl-tRNA synthetase. *Eur J Biochem* **185**, 331-339 (1989).
25. B. Felden, C. Florentz, R. Giege, E. Westhof, Solution structure of the 3'-end of brome mosaic virus genomic RNAs. Conformational mimicry with canonical tRNAs. *J Mol Biol* **235**, 508-531 (1994).
26. P. Fechter, R. Giege, J. Rudinger-Thirion, Specific tyrosylation of the bulky tRNA-like structure of brome mosaic virus RNA relies solely on identity nucleotides present in its amino acid-accepting domain. *J Mol Biol* **309**, 387-399 (2001).
27. T. W. Dreher, T. C. Hall, Mutational analysis of the tRNA mimicry of brome mosaic virus RNA. Sequence and structural requirements for aminoacylation and 3'-adenylation. *J Mol Biol* **201**, 41-55 (1988).
28. A. L. Rao, C. Cheng Kao, The brome mosaic virus 3' untranslated sequence regulates RNA replication, recombination, and virion assembly. *Virus Res* **206**, 46-52 (2015).
29. A. L. Rao, T. W. Dreher, L. E. Marsh, T. C. Hall, Telomeric function of the tRNA-like structure of brome mosaic virus RNA. *Proc Natl Acad Sci U S A* **86**, 5335-5339 (1989).
30. Y. G. Choi, T. W. Dreher, A. L. Rao, tRNA elements mediate the assembly of an icosahedral RNA virus. *Proc Natl Acad Sci U S A* **99**, 655-660 (2002).
31. S. Barends *et al.*, tRNA-like structure regulates translation of Brome mosaic virus RNA. *J Virol* **78**, 4003-4010 (2004).
32. T. M. Nakamura, Y. H. Wang, A. J. Zaug, J. D. Griffith, T. R. Cech, Relative orientation of RNA helices in a group 1 ribozyme determined by helix extension electron microscopy. *EMBO J* **14**, 4849-4859 (1995).
33. D. M. Lilley, Analysis of branched nucleic acid structure using comparative gel electrophoresis. *Q Rev Biophys* **41**, 1-39 (2008).
34. A. T. Frank, A. C. Stelzer, H. M. Al-Hashimi, I. Andricioaei, Constructing RNA dynamical ensembles by combining MD and motionally decoupled NMR RDCs: new insights into RNA dynamics and adaptive ligand recognition. *Nucleic Acids Res* **37**, 3670-3679 (2009).
35. K. Kappel *et al.*, Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. *Nat Methods* **17**, 699-707 (2020).
36. R. McGreevy, I. Teo, A. Singharoy, K. Schulten, Advances in the molecular dynamics flexible fitting method for cryo-EM modeling. *Methods* **100**, 50-60 (2016).
37. R. T. Kidmose *et al.*, Namdinator - automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCrJ* **6**, 526-531 (2019).
38. F. C. Chou, P. Sripakdeevong, S. M. Dibrov, T. Hermann, R. Das, Correcting pervasive errors in RNA crystallography through enumerative structure prediction. *Nat Methods* **10**, 74-76 (2013).
39. F. C. Chou, N. Echols, T. C. Terwilliger, R. Das, RNA Structure Refinement Using the ERRASER-Phenix Pipeline. *Methods Mol Biol* **1320**, 269-282 (2016).
40. M. R. Chapman, C. C. Kao, A minimal RNA promoter for minus-strand RNA synthesis by the brome mosaic virus polymerase complex. *J Mol Biol* **286**, 709-720 (1999).
41. C. H. Kim, C. C. Kao, I. Tinoco, Jr., RNA motifs that determine specificity between a viral replicase and its promoter. *Nat Struct Biol* **7**, 415-423 (2000).
42. T. W. Dreher, T. C. Hall, Mutational analysis of the sequence and structural requirements in brome mosaic virus RNA for minus strand promoter activity. *J Mol Biol* **201**, 31-40 (1988).
43. P. Fechter, J. Rudinger-Thirion, A. Theobald-Dietrich, R. Giege, Identity of tRNA for yeast tyrosyl-tRNA synthetase: tyrosylation is more sensitive to identity nucleotides than to structural features. *Biochemistry* **39**, 1725-1733 (2000).
44. M. Tsunoda *et al.*, Structural basis for recognition of cognate tRNA by tyrosyl-tRNA synthetase from three kingdoms. *Nucleic Acids Res* **35**, 4289-4300 (2007).
45. I. Kalvari *et al.*, Non-Coding RNA Analysis Using the Rfam Database. *Curr Protoc Bioinformatics* **62**, e51 (2018).
46. I. Kalvari *et al.*, Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res* **46**, D335-D342 (2018).

47. T. M. Colussi *et al.*, The structural basis of transfer RNA mimicry and conformational plasticity by a viral RNA. *Nature* **511**, 366-369 (2014).
48. L. Salmon, S. Yang, H. M. Al-Hashimi, Advances in the determination of nucleic acid conformational ensembles. *Annu Rev Phys Chem* **65**, 293-316 (2014).
49. S. L. Kaiming Zhang, Kalli Kappel, Grigore Pintilie, Zhaoming Su, Tung-Chung Mou, Michael F. Schmid, Rhiju Das, Wah Chiu, Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. *Nature Communications* **10**, (2019).
50. R. J. Kohl, T. C. Hall, Aminoacylation of RNA from several viruses: amino acid specificity and differential activity of plant, yeast and bacterial synthetases. *J Gen Virol* **25**, 257-261 (1974).
51. E. P. Nawrocki, S. R. Eddy, Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**, 2933-2935 (2013).
52. E. Rivas, J. Clements, S. R. Eddy, Estimating the power of sequence covariation for detecting conserved RNA structure. *Bioinformatics* **36**, 3072-3076 (2020).
53. P. Cordero, W. Kladwang, C. C. VanLang, R. Das, The mutate-and-map protocol for inferring base pairs in structured RNA. *Methods Mol Biol* **1086**, 53-77 (2014).
54. S. Yoon *et al.*, HiTRACE: high-throughput robust analysis for capillary electrophoresis. *Bioinformatics* **27**, 1798-1805 (2011).
55. W. Kladwang *et al.*, Standardization of RNA chemical mapping experiments. *Biochemistry* **53**, 3063-3065 (2014).
56. H. Kim, P. Cordero, R. Das, S. Yoon, HiTRACE-Web: an online tool for robust analysis of high-throughput capillary electrophoresis. *Nucleic Acids Res* **41**, W492-498 (2013).
57. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).
58. S. H. Scheres, RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-530 (2012).
59. T. C. Terwilliger, S. J. Ludtke, R. J. Read, P. D. Adams, P. V. Afonine, Improvement of cryo-EM maps by density modification. *Nat Methods* **17**, 923-927 (2020).
60. E. F. Pettersen *et al.*, UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004).
61. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877 (2019).
62. C. J. Williams *et al.*, MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293-315 (2018).
63. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).