



Supplementary Materials for  
**Geometric deep learning of RNA structure**

Raphael J. L. Townshend *et al.*

Corresponding authors: Rhiju Das, [rhiju@stanford.edu](mailto:rhiju@stanford.edu); Ron O. Dror, [ron.dror@stanford.edu](mailto:ron.dror@stanford.edu)

*Science* **373**, 1047 (2021)  
DOI: 10.1126/science.abe5650

**The PDF file includes:**

Materials and Methods  
Figs. S1 to S9  
Tables S1 to S5  
References

**Other Supplementary Material for this manuscript includes the following:**

MDAR Reproducibility Checklist

## Materials and Methods

### Equivariance

Fundamentally, the goal of ARES's design is to be able to encode detailed geometric patterns while also automatically being able to recognize and compose them at different positions and orientations. We refer to such patterns as structural motifs: a specific arrangement of atoms in 3D. A key to this ability is a property known as *equivariance*. Informally, a function or neural network layer is equivariant to some transformation (such as a rotation or translation in 3D space) if a transformation of the input results in the same transformation of the output. *Invariance* is a related property in which the function output does not vary with respect to transformations of the input. See the work by Cohen and Welling (37) for more formal definitions and additional background. We require that all network layers be equivariant to rotations and translations. Furthermore, the initial layers accumulate information locally and preserve information about orientation and position of their inputs, while later layers gather information globally. In the following, we focus on rotational equivariance, as translational equivariance for neural network architectures is more common and easier to achieve (e.g., by encoding relative positions of atoms).

It is important to note that when training and testing ARES we never explicitly rotate the network inputs; instead, the inbuilt equivariance property and ability to preserve information about orientations allow the network to learn to recognize a pattern in any orientation while also keeping track of the orientation in which the pattern occurs. ARES can then learn higher-level patterns, which combine lower-level patterns in specific orientations. As an example, we could imagine a single network layer learning to recognize aromatic rings. We could then further compose this layer with a second layer to discover  $\pi$  interaction patterns, by having the second layer reason about the relative orientations and positions of aromatic rings. If the first layer did not preserve information about orientations and positions, the second layer would be unable to tell if the stacking was in a sandwich conformation (where the two rings would be parallel to one another) versus a t-shape conformation (where the two rings would be perpendicular). By keeping the layers and their compositions equivariant and preserving orientation information, we can not only recognize the specific conformation, but also pass on the overall orientation of the stack to further layers.

### Equivariant convolution

One of our primary layers that preserves information about orientations and positions is the equivariant convolution, originally defined in the publication on Tensor Field Networks (16). Our equivariant convolutions take in a set of atoms in three-dimensional (3D) space, represented as points in 3D with associated features for each point, and compute new features for each atom. The function that computes this output is learnable. We describe our implementation below.

For a given atom  $a$  (referred to as the source atom), the equivariant convolution is based on filter functions applied one at a time to each atom  $b$  within its local neighborhood (referred to as the neighbor atoms). We define  $\vec{r}_{ab}$  as the 3D vector between the source atom and a given neighbor atom. The filter functions only take as input this vector  $\vec{r}_{ab}$ , and their output is combined with a given neighbor atom's current features to produce updated features for the source atom. In this way, a neighboring atom's information is shared with the source atom. The design of the filter functions, as well how their outputs are combined with neighbor's features, is

the key to ensuring the layer can capture detailed geometric information while still maintaining equivariance and preserving orientation information.

In this work, these filter functions are composed by combining two classes of sub-functions: radial ( $R$ ) and angular ( $A$ ) functions. We define these two sets of sub-functions next.

### Radial functions

The radial functions encode the distance relationships between atoms, without considering their relative orientations. Our radial functions take the form of a dense neural network (11). The inputs to this network,  $G$ , are computed by applying a filter bank of Gaussians (examples shown in Figure S1B) to the magnitude  $r_{ab} = \|\vec{r}_{ab}\|$ :

$$G(r_{ab}) = [G_0(r_{ab}), G_1(r_{ab}), \dots, G_n(r_{ab})]$$

with:

$$G_j(r_{ab}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(r_{ab}-\mu_j)^2}{2\sigma^2}}$$

where  $\sigma = 1$ ,  $n = 11$ , and  $\mu_j = \frac{12}{11}j$  and  $r_{ab}$  is expressed in units of ångstrom (we omit units for simplicity). The dense network has one hidden layer of dimension 12, with a ReLU activation (38) before the hidden layer, and produces  $E$  scalar radial filter outputs:

$$[R_0(r_{ab}), R_1(r_{ab}), \dots, R_E(r_{ab})] = \text{Dense}(G(r_{ab}))$$

The dense network contains learnable bias parameters as part of the hidden and output layers. As these functions only consider distances between atoms, they are invariant to translations and rotations.

### Angular functions

The angular functions, on the other hand, only consider orientations between atoms, not distances. We use the real spherical harmonics,  $Y$ , as our angular functions. Spherical harmonics are grouped by their angular resolution  $l \in \mathbb{Z}_0^+$ , which we refer to as *angular order*—there are  $2l + 1$  harmonics per order. To index within each order, we use an *angular index*  $m$ , with  $m \in \{-l, -l + 1, \dots, l - 1, l\}$ . These functions are applied to the unit vector  $\hat{r}_{ab} = \vec{r}_{ab}/r_{ab}$ .

We define  $L$  as the maximum order used, thus using  $M = \sum_{l=0}^L (2l + 1)$  scalar-valued angular functions total. We use  $L = 2$ , giving us the zeroth-, first-, and second-order harmonics (Figure S1B). Spherical harmonics form an orthonormal basis for functions on the unit sphere and are equivariant with respect to rotations in 3D Euclidean space. The maximum angular order governs the resolution to which angular dependencies can be captured within a network layer.

### Filter functions

We define filter functions,  $F_{cm}^l$ , as combinations of our scalar-valued radial and angular functions. These form the core of the equivariant convolution:

$$F_{cm}^l(\vec{r}_{ab}) = R_c(r_{ab}) Y_m^l(\hat{r}_{ab})$$

where  $c \in \{0, 1, \dots, E\}$ ,  $l \in \{0, 1, \dots, L\}$ ,  $m \in \{-l, -l + 1, \dots, l\}$ . We refer to  $E$  as the dimension of the equivariant convolution. The three equivariant convolutions have dimensions 24, 12, and 4 (see Figure S1A). As the radial sub-function is invariant to rotations, and the angular sub-function is equivariant to rotations within an angular order, each filter function is equivariant to rotations within an angular order.

### Combining filter functions and features

The features,  $V$ , associated with a point (here point index  $a$ ) have multiple components indexed by angular, radial, and order indices  $m$ ,  $c$ , and  $l$ , respectively. For the input to the first network layer, we only have scalar features (angular order  $l = 0$ ) and a total of  $E = 3$  radial features for the three possible element types that we encode.

The equivariant convolution,  $L$ , uses a tensor product that allows us to combine the point-associated features and filter functions in an equivariant manner to update our features  $V$ :

$$L_{acm_o}^{l_o}(V_{acm_i}^{l_i}) = \sum_{m_i, m_f} C_{(l_f, m_f)(l_i, m_i)}^{(l_o, m_o)} \sum_{b \in \text{neighbors}(a)} F_{cm_f}^{l_f}(\vec{r}_{ab}) V_{bcm_i}^{l_i}$$

where the subscripts  $f$ ,  $i$ , and  $o$  denote filter, input, and output feature, respectively.  $C$  are Clebsch-Gordan coefficients which are non-zero only for  $|l_i - l_f| \leq l_o \leq l_i + l_f$  (39). Note that different combinations of  $l_f$  and  $l_i$  can yield outputs of the same angular order  $l_o$ . Thus, we compute the new value of  $V$  for each of these combinations of  $l_f$  and  $l_i$  and concatenate these outputs together along the  $c$  dimension, which effectively increases the dimensionality of  $V$  along the  $c$  dimension. We restrict the point convolution to the  $K$  nearest neighbors of a given point to account for the fact that the laws of physics that govern inter- and intramolecular interactions are local (17). Note that the only learned parameters in the equivariant convolution are the weights of the dense network that operate on the inter-point distance.

We describe other used layers next; these are more straightforwardly equivariant to rotations as they only operate on individual atoms (atomic embedding, pointwise normalization, pointwise non-linearity, and pointwise self-interaction) or only operate on rotationally invariant features (per-channel mean and subsequent layers). Composing these individually equivariant layers together yields a network that is overall equivariant (16, 17).

### Pointwise normalization

The pointwise normalization operation,  $N$ , acts separately on the features associated with each atom (17):

$$N(V_{acm}^l) = \frac{V_{acm}^l}{\sqrt{\sum_{c,m} (V_{acm}^l)^2}}$$

where  $m$ ,  $c$ , and  $l$  are the same angular, radial, and order indices as defined in previous layers.

### Pointwise non-linearity

The pointwise non-linearity operation,  $P$ , also acts separately on the features associated with each atom. We use a non-linearity adapted from Tensor Field Networks (16):

$$P(V_{acm}^l) = \begin{cases} \eta(V_{acm}^l) & \text{if } l = 0 \\ V_{acm}^l \cdot \eta\left(\sqrt{\sum_m (V_{acm}^l)^2} + b^l\right) & \text{otherwise} \end{cases}$$

where  $b^l$  is a learnable scalar bias term (one per order),  $m$ ,  $c$ , and  $l$  are the same angular, radial, and order indices as defined in previous layers, and  $\eta$  is a shifted soft plus non-linearity, as in SchNet (40):

$$\eta(x) = \ln(0.5e^x + 0.5)$$

### Pointwise self-interaction

We use self-interaction layers,  $S$ , as in SchNet (40) to mix information across radial channels between equivariant convolution layers and update our features  $V$ :

$$S_{adm}^l(V_{acm}^l) = b_d + \sum_c V_{acm}^l W_{cd}$$

where  $W$  is a learnable weight matrix,  $b$  is a learnable bias term,  $m$ ,  $c$ , and  $l$  are the same angular, radial, and order indices as defined in previous layers, and  $d$  is the new radial index. Note that the bias term is only used when operating on angular order 0 (i.e.  $l = 0$ ). Within a given self-interaction layer, the number of output radial indices is the same for each angular order of spherical harmonics; this value is referred to as the dimension of the pointwise self-interaction. The 6 self-interaction layers have dimensions 24, 24, 12, 12, 4, and 4, respectively.

### Atomic embedding

The atomic embedding is used to generate the initial features associated with each atom (which only inhabits angular order 0). We use a one-hot representation which encodes if the atom is a carbon, nitrogen, or oxygen. All atoms of other element types are not represented:

$$V_{a00}^0 = 1 \text{ if atom } a \text{ has element type carbon}$$

$$V_{a10}^0 = 1 \text{ if atom } a \text{ has element type oxygen}$$

$$V_{a20}^0 = 1 \text{ if atom } a \text{ has element type nitrogen}$$

### Per-channel mean

After these initial layers, we drop the positions of the atoms, as well as any features that do not correspond to the zeroth-order harmonic. We then compute the average, across all atoms, of

each of the remaining features. This averaging produces a molecule-wide embedding that is insensitive to the original RNA's size. As we are keeping only the entries corresponding to the zeroth-order harmonic, this causes further layers to be invariant to rotations, as the zeroth-order harmonic is itself invariant to rotations. This results in new features  $E$  that are indexed only by the radial channel  $c$ :

$$E_c = \sum_a V_{ac0}^0$$

### Dense layers

Our dense layers are standard fully connected neural network layers:

$$E_d = b_d + \sum_c W_{cd} E_c$$

where  $W$  and  $b$  are a learnable weight matrix and learnable bias parameters, respectively.

### Network architecture

In total, there are 15 layers with learnable parameters (6 self-interactions, 3 equivariant convolutions, 3 pointwise non-linearities, and 3 fully connected), and 5 layers with fixed parameters (1 atomic embedding, 3 pointwise normalizations, and 1 per-channel mean) (16, 17). The first fully connected layer uses an ELU non-linearity (41) while the other two use no non-linearities. All learnable biases were initialized to 0, and all learnable weight matrices were initialized using Xavier uniform initialization (42). Each equivariant convolution uses the real spherical harmonics of orders 0, 1, and 2, for a total of 9 angular sub-functions. We define the local neighborhood of an atom as the nearest 50 atoms (including the source atom itself). The overall network design, the dimension of the equivariant convolution and pointwise self-interaction layers, and the number of neurons in the dense layers, are shown in Figure S1A.

### Training

The network was trained with the Adam optimizer (43) to minimize the Huber (44) loss, as applied to the difference between the predicted and true root mean square deviation (RMSD) between the atoms of the experimentally determined structure and a candidate structural model after alignment:

$$RMSD = \sqrt{\frac{1}{N} \sum_a |\vec{p}_a - \vec{p}_a'|^2}$$

where  $N$  is the total number of atoms present, and  $\vec{p}_a$  and  $\vec{p}_a'$  are the positions of atom  $a$  in the candidate model and the experimentally determined structure, respectively. We performed all RMSD calculations using Rosetta, excluding hydrogen atoms as well as the rare bases and sugars that make no atomic contacts in the experimentally determined structure. RMSD is a common evaluation metric in the field of RNA structure prediction (45).

We trained on 18 RNAs selected from the FARNA benchmark (18). For each RNA, we randomly selected 1,000 structural models generated by FARFAR2 from the FARFAR2-Classics dataset (19). The experimentally determined structure for each of these RNAs was taken from the FARNA benchmark (18). We randomly split the dataset by RNA, with 14 RNAs used for training the network’s parameters and the other 4 for selecting hyperparameters that specify characteristics of the training process (Figure S2). The RNAs used for hyperparameter selection were those corresponding to PDB codes 1Q9A, 1I9X, 1A4D, and 1KKA, with the rest used for training network parameters.

The network was trained using 4 Titan X Pascal GPUs in parallel with Horovod (46), using synchronous gradient descent with the batch evenly split across the GPUs. We optimized several hyperparameters that specify characteristics of the training process: batch size, number of epochs, learning rate, and number of candidate structural models per RNA to feed in. We considered ~100 sets of hyperparameter values. For each of these sets, we trained the network parameters using structural models for 14 RNAs, and then evaluated the loss of the trained network over structural models for another four RNAs (1Q9A, 1I9X, 1A4D, and 1KKA). The set of hyperparameter values with the lowest loss was a batch size of 16, 1,000 randomly selected structural models for each sequence, a learning rate of 0.01, and a single training epoch. The specific network parameters that yielded this lowest loss were used for all ARES results reported in this manuscript, including the blind structure prediction competition.

### Benchmark 1 dataset

We benchmarked on the first 21 RNAs from the RNA-Puzzles blind RNA structure prediction challenges (45). For each RNA, we used all the structural models from the FARFAR2-Puzzles dataset (19). These RNAs are generally quite large, ranging in size from 41–188 nucleotides, and include riboswitch aptamers, ribozymes, and viral noncoding RNAs, providing a diverse range of targets against which to assess performance. On average, over 20,000 structural models were available per RNA.

Depending on the difficulty of the prediction target, the best models generally range from 2–10 Å RMSD to the experimentally determined structure. To guarantee the generation of near-native structural models (< 2 Å RMSD) for each sequence in FARFAR2-Puzzles, we employed the same FARFAR2 simulation conditions as used for the FARFAR2-Puzzles dataset, but we added in an energetic restraint to the experimentally determined structure and omitted the exclusion of homologous fragments (19). Specifically, we used flat-bottomed harmonic functions  $y$  restraining every heavy atom of the RNA model to the corresponding coordinates from the experimentally determined structure, piecewise such that for each atom  $a$ :

$$y_a = 0 \text{ if } |\vec{p}_a - \vec{p}_a'| \leq 1 \text{ \AA}$$

$$y_a = |\vec{p}_a - \vec{p}_a'| \text{ if } |\vec{p}_a - \vec{p}_a'| > 1 \text{ \AA}$$

For each RNA, the number of restrained structural models was 1% the number of unrestrained models from the FARFAR2-Puzzles dataset.

### Benchmark 2 dataset

To choose RNAs for our *de novo* prediction benchmark in an unbiased fashion, we began by identifying all structures in the PDB (accessed on January 30<sup>th</sup>, 2021) that met several quality criteria. First, structures must be more recent than 2007 (that is, more recent than any of those in the training set). Second, structures must be relatively high-quality: cryo-EM and X-ray structures must have better resolution than 3.0 Å; NMR structures must have a clashscore of zero. Third, each RNA must comprise 25 to 150 nucleotides, in order to eliminate ribosomes, spliceosome states, and other targets too large to sample thoroughly, as well as simple duplexes and stem loops whose successful prediction is not a meaningful indication of scoring function quality.

We filtered these structures to remove any molecules that shared a family-level Rfam classification with cases from the training set, Benchmark 1, or the four blind predictions. Where this reduced set contained more than one structure in the same Rfam family (47), we applied two criteria to select only one structure per family. First, we favored structures in an *apo* state or with their cognate ligand bound over structures in a *holo* state or with a non-cognate ligand bound. Second, we favored more recently deposited structures over older structures. This deduplication and filtering procedure yielded 16 structures for 16 distinct RNAs. These structures are shown in Figure S1. The corresponding PDB codes, nucleotide sequences, and RNA names are listed in Table S4.

For each of these 16 RNAs, we generated 5,000 structural models with the FARFAR2 *rna\_denovo* application, using only sequence and secondary structure information (Table S4). All secondary structure information was predicted from Rfam alignments (47) or from sequence alone (48), or based on predictions published before the corresponding 3D structures (49, 50). We scored the resulting models with seven scoring functions: RASP (21), 3dRNAscore (22), SimRNA (51), ARES, and three versions of the Rosetta scoring function published in 2007 (*rna\_lores.wts*) (18), 2010 (*rna\_hires.wts*) (52), and 2020 (*stepwise/rna/rna\_res\_level\_energy4.wts*) (19). For each RNA, we selected the best-scoring model and the ten best-scoring models according to each scoring function; these were used to compute the results shown in Figure S5, panels A and C.

We also performed a bootstrap analysis to determine the extent to which accuracy of the top-scoring models depends on the precise set of candidate models considered. For each RNA, we selected 20,000 bootstrap samples, each comprising 1,000 structural models from the 5,000 models available for that RNA. For each scoring function, we calculated the RMSD of the best-scoring model in each bootstrap sample and the lowest RMSD among the ten best-scoring models in each sample, and then calculated the median of these values across the 20,000 bootstrap samples. We used the same set of bootstrap samples to calculate 95% confidence intervals. To estimate p-values for difference-of-medians between two scoring functions *a* and *b*, we calculated the fraction of bootstrap samples for which  $\text{median}(a) < \text{median}(b)$ . These bootstrapped median results appear in Figure 2D and Figure S5B.

Scripts for scoring models with each scoring function and parsing the resulting output, as well as the raw data for each model, are deposited online (35, 36).

### Study of complex tertiary interactions (Figure S9)

To test whether ARES correctly identifies various canonical tertiary interactions, we conducted an additional FARFAR2 simulation of an adenine riboswitch (PDB code: 1Y26). We used the same FARFAR2 protocol as in Benchmark 1 but omitted any input templates. The structural model with the best ARES score (Figure S9A, B) exhibits closely packed helices and base triplets of correct geometry. We also illustrate RNAs from Benchmark 1 where the structural model with the best ARES score included an intercalated T-loop (Fig. S9C; RNA Puzzle 15; a synthetic hammerhead ribozyme; PDB code 5DI4) and a loop-into-helix motif (Fig S9D; RNA Puzzle 21; *Thermobifida fusca* guanidinium III riboswitch; PDB code: 5NWQ).

### Helix width experiment (Figure 4A)

An ideal five-base-pair A-form helix was constructed using Rosetta's RNA tools:

```
rna_helix.py -seq gcgcg cgcgc
```

The two strands from the resulting PDB were translated relative to each other using PyRosetta (53): the rigid-body degree-of-freedom between the two strands was placed between the C4' atom of the third and eighth residues of the PDB (i.e., the middle residue of each strand), and Rosetta's RigidBodyTransMover was employed to sample strand translations every 0.1 Å for 10 Å in both directions.

### RNA characteristics experiment (Fig. 4B)

We assembled a set of non-redundant RNA structures based on the Representative Sets of RNA 3D Structures (54). We used version 3.91 at 2.5 Å resolution cutoff, which consisted of 1952 structures total. We removed structures with fewer than 100 non-hydrogen atoms (348 structures, 17.8% of total), because RNAs with fewer than five nucleotides generally do not form well-defined monomeric structures in solution. We also removed structures with more than 20,000 non-hydrogen atoms (29 structures, 1.5% of total) due to memory limitations of the GPUs we used.

To extract ARES's learned features for a given RNA, we applied the final selected ARES network to this RNA and took the penultimate dense layer's activations (a vector of length 256). We repeated this process for every RNA in the nonredundant RNA structures set, and then computed the principal components of the resulting set of activations, comparing them against RNA properties as computed by x3dna (55).

### Blind structure prediction

The format of the RNA Puzzles competition involves submitting 10 predictions for an RNA, whose structure is only released after the competition. The FARFAR2 protocol (19) was used to generate 87170, 52659, 93798, and 150810 candidate structural models for the blind prediction challenge of RNA-Puzzles 24, 26, 27, and 28, respectively. Exact inputs to the FARFAR2 protocol are described in Table S1 and include secondary structures previously predicted in the literature (56–58) as well as threaded local template structures from homologous RNAs. Puzzle 24 is the Adenovirus VA-I RNA (59) (PDB code: 6OL3), Puzzle 26 is the *Geobacillus kaustophilus* T-box discriminator-tRNA<sup>Gly</sup> complex (60) (PDB code: 6PMO),

Puzzle 27 is the *Bacillus subtilis* T-box–tRNA<sup>Gly</sup> complex (60) (PDB code: 6POM), and Puzzle 28 is the *Nocardia farcinic* T-box–tRNA<sup>Ile</sup> complex (61) (PDB code: 6UFM).

These candidates were ranked by ARES, and then either the top 1% of structures (for Puzzle 24) or the best 400 structures (for Puzzles 26, 27, and 28) were clustered using Rosetta's `rna_cluster` application, using a cluster radius of either 5.0 Å (for Puzzle 24) or a function of the simulation 'convergence' as measured by the mean pairwise RMSD among the top 10 models (19, 62) (for Puzzles 26, 27, 28). This resulted in a cluster radius of 5.0 Å for Puzzles 26 and 27. For Puzzle 28, three distinct model sets starting from slightly different template structures and modeling conditions were analyzed in this fashion and their resulting clusters pooled (see Table S1). The radii in this case were 9.0 Å (for the "cst" subset) and 14.0 Å (for the other two subsets). For each cluster, we submitted as a representative the model with the best Rosetta score. The best 10 cluster representatives, again ranked by ARES score, were submitted to each competition. These are available at <https://www.rnapuzzles.org> under the name DasTFN for Puzzle 24 and the name TFN for Puzzles 26, 27, and 28.

For each competition, we also submitted another set using the same procedure applied to the same pool of candidate structural models, but instead using the all-atom Rosetta energy function for initial and cluster rankings, under the name FARFAR2.

The structural models submitted for Puzzles 24, 26, 27, and 28 were locked down upon submission to the RNA Puzzles competition. The dates they were submitted, as well as the organizers of each Puzzle, can be found on [rnapuzzles.org](http://rnapuzzles.org). Note that FARFAR2 does not predict coordinates for phosphate atoms on 5' chain termini; for RNA Puzzles submissions, they are grafted on using a deterministic procedure (the `rna_graft` executable with the flag `-unvirtualize_phosphate`), simply to meet model formatting requirements. For Puzzles 26, 27, and 28, where 5' phosphate atoms had been erroneously omitted in the original submission, the competition organizers accepted updated models with 5' phosphate atoms placed.

After the release of the experimentally determined structure, the exact same procedures (ranking, selecting the top 400 or top 1% of structures, and clustering) were applied to the same pool of candidate structural models, using other scoring functions for the ranking of both individual models and clusters. These scoring functions were RNA3DCNN (23), 3dRNAScore, and RASP. All of our blind and retrospective submissions are listed in Table S5.

#### ARES usability and performance

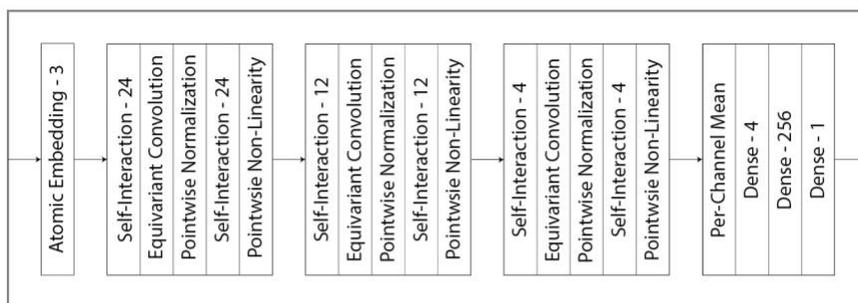
We have created an ARES webserver that accepts candidate structural models and provides the ARES score for each: <http://drorlab.stanford.edu/ares.html>. The candidate structural models described in this manuscript were created with the FARFAR2 software (19). FARFAR2 is available as a webserver at <http://rosie.rosettacommons.org/farfar2>.

ARES itself is computationally inexpensive, and its memory and compute requirements scale linearly with the number of nucleotides in an RNA. As a concrete data point, on a single Titan X Pascal GPU, scoring all 87,000 structural models for blind prediction A (Figure 3) took less than 3.5 hours (i.e., ARES scored 7 models per second). This RNA has 112 nucleotides and 2352 non-hydrogen atoms. While the Titan X Pascal GPU's 12 gigabytes of memory can

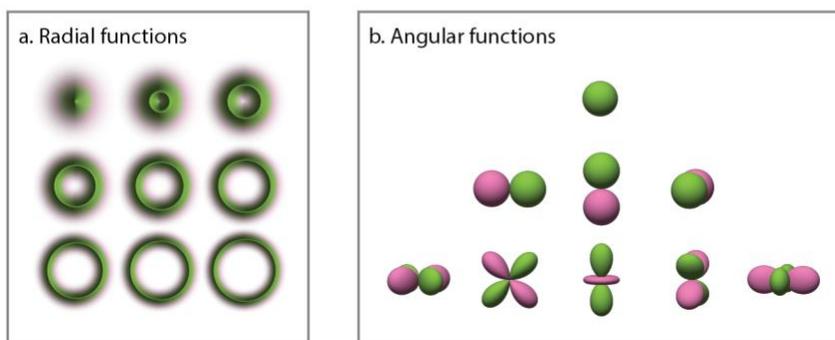
accommodate up to 20,000 non-hydrogen atoms (~1,000 nucleotides), more recent GPUs are much faster and have much more memory, accommodating RNAs of over 6,500 nucleotides. RNAs larger than this are extremely rare but could be accommodated by parallelization across multiple GPUs. We provide a Pytorch (63) implementation of ARES (32, 33). This implementation draws on the E3NN library (64).

Use of ARES, like any other scoring function, requires that one first generate candidate models to be scored. This “sampling” procedure is typically more computationally expensive than the scoring itself. For example, generating the 87,000 structural models mentioned above with FARFAR2 required 906 CPU-hours on a 32-core, 2.5 GHz AMD EPYC 7502 CPU. More powerful CPUs are now available, and this calculation can easily be parallelized both across cores and across CPUs, substantially reducing the compute time. FARFAR2’s memory and compute requirements scale sub-quadratically with the number of nucleotides in an RNA. Using 16 GB of memory, FARFAR2 can generate structural models for RNAs with over 6,500 nucleotides.

A.

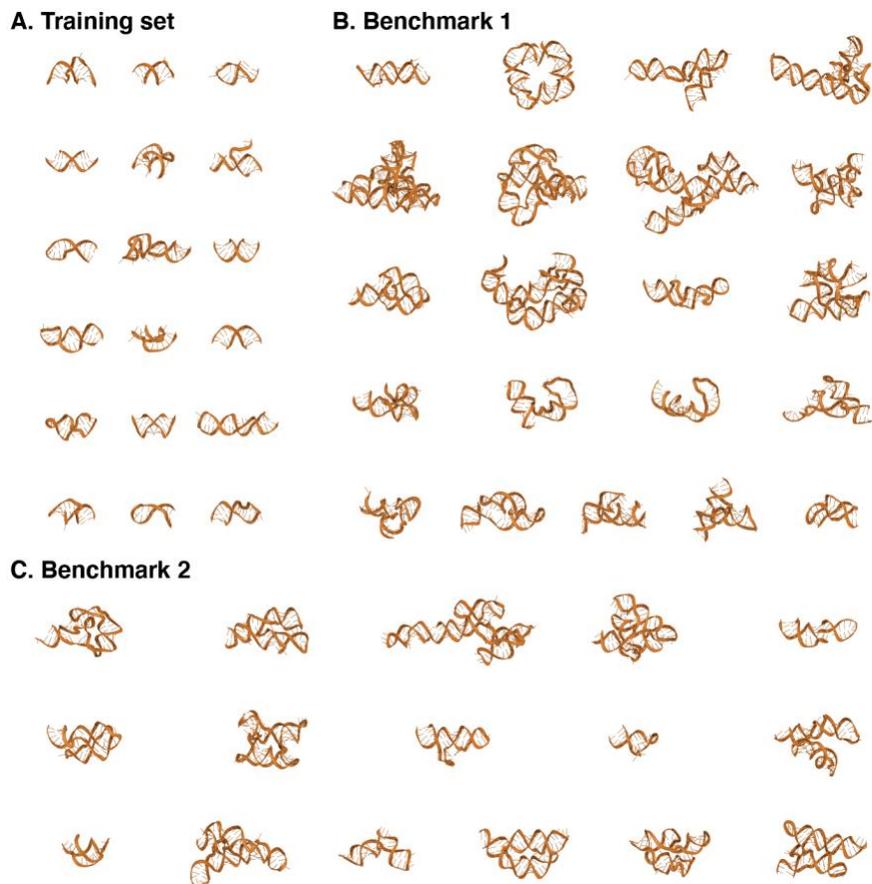


B.



**Fig. S1.**

**ARES architecture.** **A**, Composition of layers in the ARES network. A structural model is provided as input, and mapped through the layers, resulting in a scalar prediction of the RMSD to the experimentally determined structure (12). **B**, Equivariant convolutions combine radial (a) and angular (b) sub-functions, allowing for the efficient capture of detailed geometric information. The radial sub-functions only take into account distance between atoms, whereas the angular sub-functions only take into account their relative orientations. Together, they can represent complex geometric patterns.



**Fig. S2.**

**Experimentally determined structures for RNAs in the training and benchmark.** **A**, The training set is composed of small RNA structures. They are derived from FARFAR2-Classics, which consists of RNAs whose structures were published in 2007 or earlier. The final four RNAs are used for hyperparameter tuning (the three in the last row, and the rightmost one in the penultimate row). **B**, Benchmark 1 is composed of larger RNA structures. These are from FARFAR2-Puzzles, which consists of RNA molecules used in more recent blind structure prediction challenges (structures published in 2011 or later). **C**, Benchmark 2 is composed of RNAs whose structures were published in 2008 or later. Each is in a different Rfam family from any other and from any structure in the training set, Benchmark 1, or the blind prediction set. Each includes one or more of the following hallmarks of structural complexity: multiway junctions, tertiary contacts, and ligand binding sites.

Lowest RMSD among N best-scoring structural models

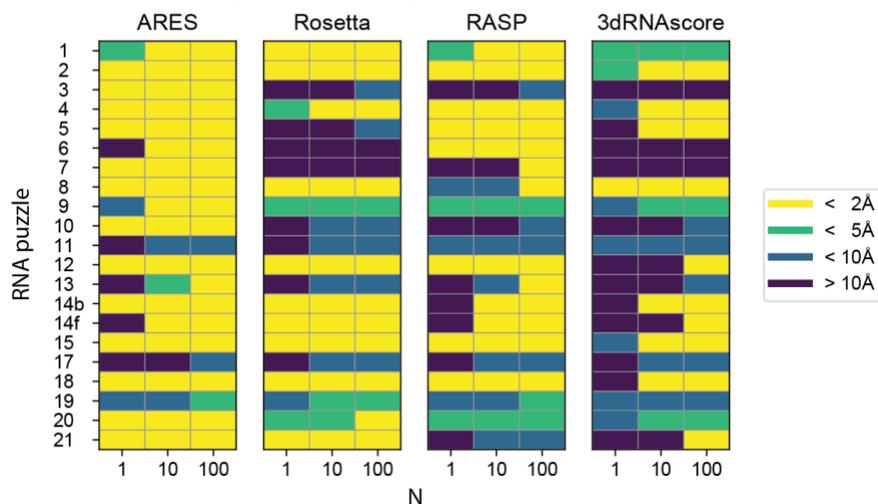


Fig. S3.

**Detailed analysis of near-native ranking task from Figure 2.** For each RNA in Benchmark 1, we rank a large set of structural models using ARES as well as three leading scoring functions. For each scoring function, we then select the  $N$  best-scoring structural models for each RNA, where  $N$  takes on the values 1, 10, and 100. For each scoring function, RNA, and value of  $N$ , we record the lowest RMSD across structural models in this set. We quantize these results by determining if this RMSD is below 2 Å, between 2 Å and 5 Å, between 5 Å and 10 Å, or above 10 Å. For each value of  $N$  and for each RMSD threshold (2 Å, 5 Å, or 10 Å), the number of RNAs with at least one selected model whose RMSD is below the threshold is greater when using ARES than when using any of the other three scoring functions.

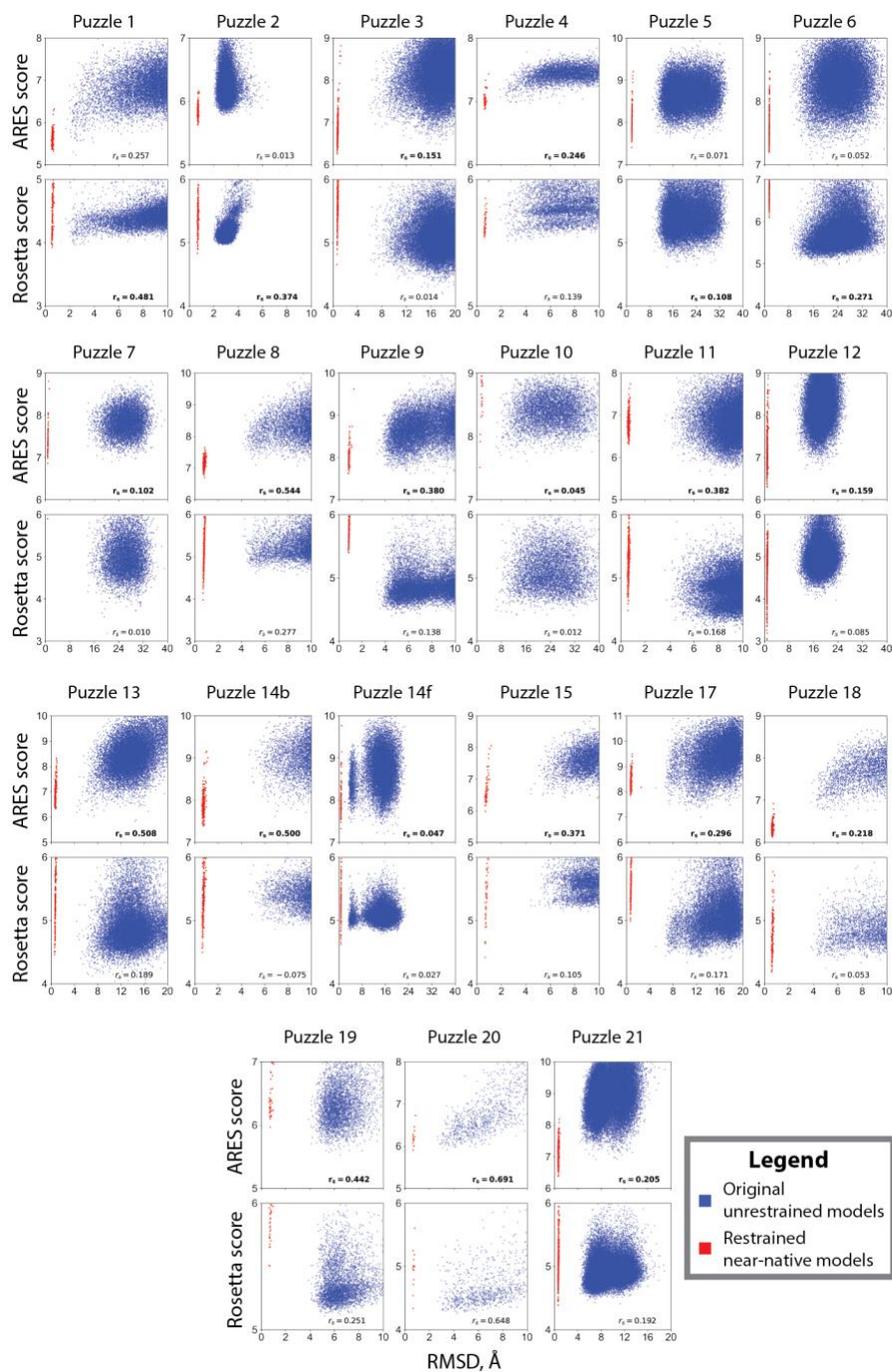
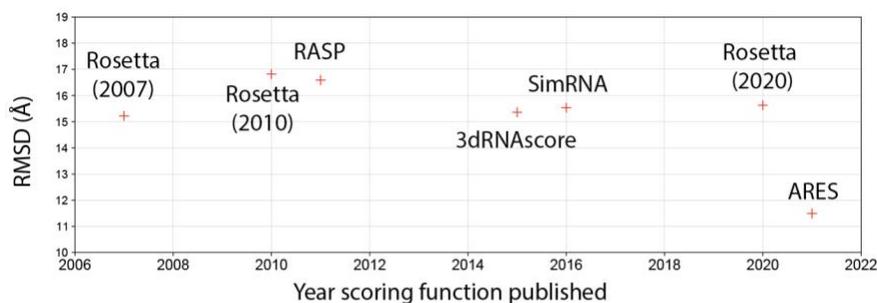


Fig. S4.

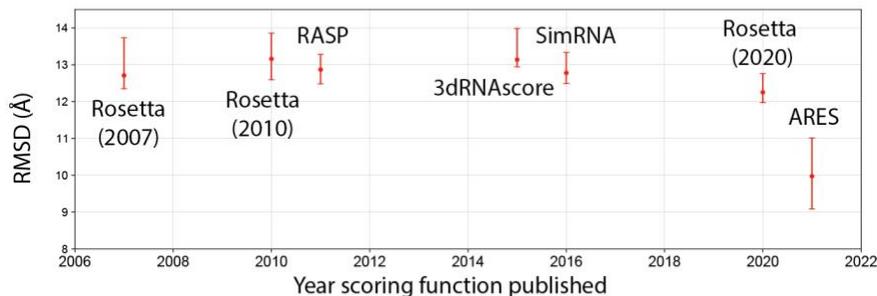
**Scatter plots comparing performance of the Rosetta and ARES scoring functions for Benchmark 1.** We consider again the FARFAR2-Puzzles dataset, which consists of FARFAR2-generated models of RNA molecules included in previous RNA Puzzles blind structure prediction challenges. Each pair of scatter plots compares ARES and Rosetta scores on a different RNA from this set, plotting the score assigned to each structural model versus its RMSD to the experimentally determined structure. To aid with visualization, the “Rosetta score”

plotted here is computed by adding 105 to the original Rosetta score and then taking the natural logarithm of the sum. We first consider the structural models provided within FARFAR2-Puzzles (blue), finding that ARES achieves a better Spearman correlation (bottom right of each plot, denoted as  $r_s$ ) on 17 of the 21 RNAs when ranking these models. To ensure that some models were near-native (i.e., within 2 Å RMSD of the experimentally determined native structure), we then added a smaller number of models (red) that were generated with energetic restraints to the native structure's coordinates (12). When using only the FARFAR2-Puzzles dataset, ARES and Rosetta tend to choose models of similar accuracy, but when the restrained near-natives are added in, ARES generally chooses much more accurate models than Rosetta (see also Figure 2). Note that analyses based on the FARFAR2-Puzzles data set are biased in favor of the Rosetta scoring function, because the FARFAR2 sampling method was tuned using this scoring function—specifically, FARFAR2 was tuned to ensure that the best models it produced for the RNAs in this particular dataset would be ranked highly by the Rosetta scoring function.

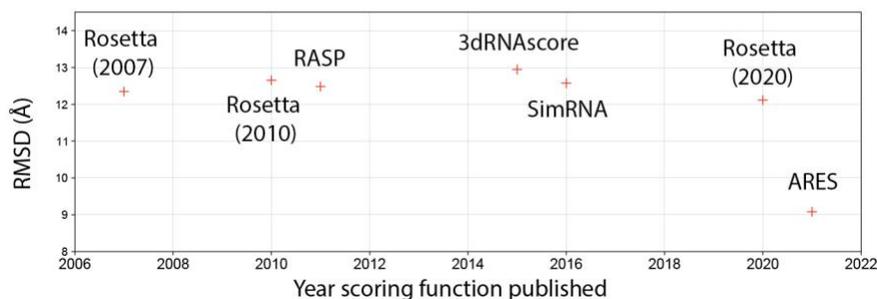
**A. RMSD of best-scoring structural model (full model set, no bootstrapping)**



**B. Lowest RMSD among 10 best-scoring structural models (bootstrapped)**

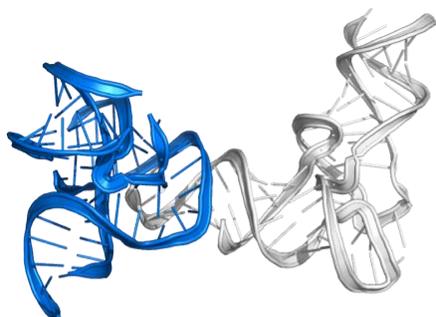


**C. Lowest RMSD among 10 best-scoring structural models (full model set, no bootstrapping)**



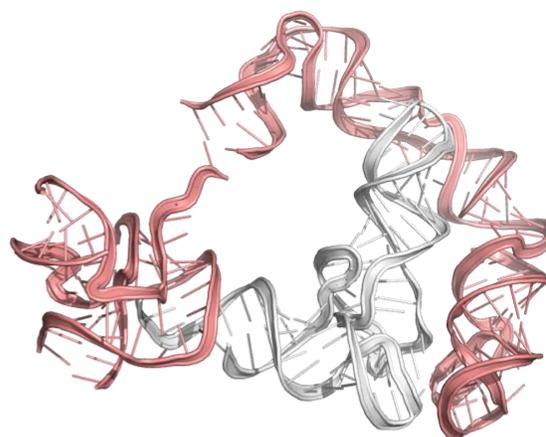
**Fig. S5.**

**Further results for Benchmark 2. A,** As in Figure 2D, we determine the RMSD of the model scored as best by each of seven scoring functions for each RNA, and then plot the median RMSD across RNAs for each scoring function. This plot differs from that of Figure 2D because here, we score all 5,000 structural models for each RNA and then select the best-scoring one, whereas in Figure 2D, we consider 1,000 structural models in each bootstrap sample. **B,** For each RNA and for each of the seven scoring functions, we determine the minimum RMSD across the 10 best-scoring structural models. For each scoring function, we plot the median across RNAs, with a 95% confidence interval determined by bootstrapping (12). ARES significantly outperforms each of the other scoring functions (p values 0.0004–0.0033; (12)). Of the other scoring functions, none significantly outperforms any other (p values 0.065–0.842). **C,** Results of an analysis identical to that of panel B, except that we score all 5,000 structural models for each RNA and then select the 10 best-scoring ones, rather than bootstrapping.



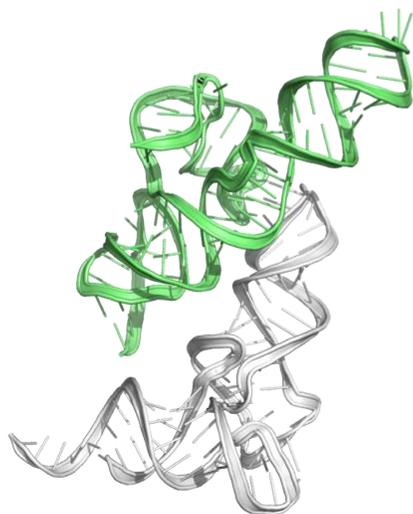
### RNA B

*G. kaustophilus* T-box discriminator/tRNA-Gly



### RNA C

*B. subtilis* T-box/tRNA-Gly



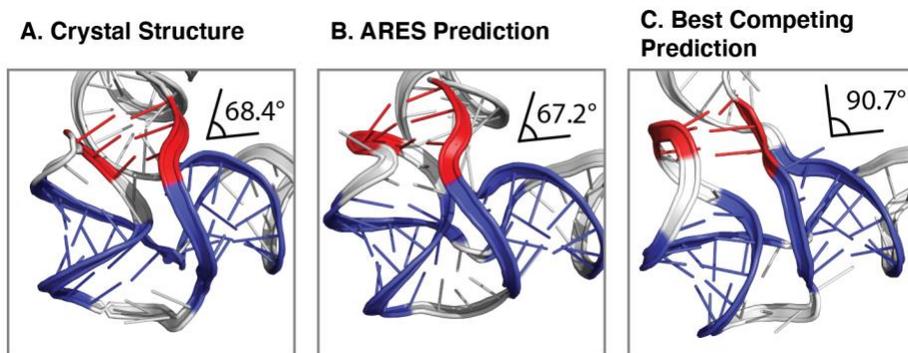
### RNA D

*N. farcinica* T-box/tRNA-Ile

**Fig. S6.**

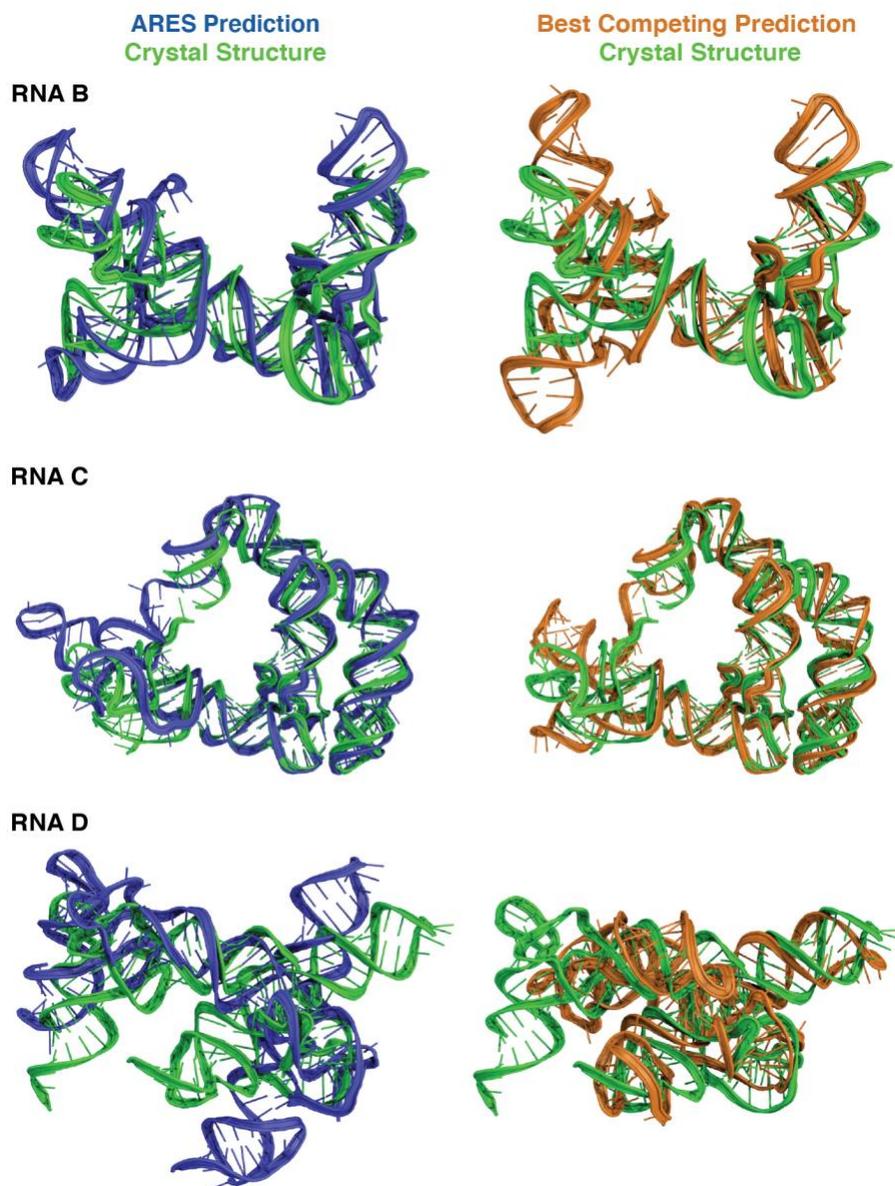
**The blind predictions on T-box riboswitches are significantly distinct challenges.** The three T-box riboswitch complexes (RNA B: the *Geobacillus kaustophilus* T-box discriminator–tRNA<sup>Gly</sup>, PDB code 6PMO; RNA C: the *Bacillus subtilis* T-box–tRNA<sup>Gly</sup>, PDB code 6POM; and RNA D: the *Nocardia farcinica* T-box–tRNA<sup>Ile</sup>, PDB code 6UFM), with their tRNA ligands shown in gray. The T-box nucleotides of RNA-Puzzle 26 (RNA B), 27 (RNA C), and 28 (RNA D) are blue, pink, and lime green respectively. This alignment emphasizes that each of these blind prediction challenges involves a distinct subset of residues from different T-box riboswitches. Even in the one instance of substantial overlap (RNA B consists exclusively of a tRNA bound to a discriminator domain, while RNA C features a larger T-box construct that

includes a discriminator domain), there is a visually apparent difference in the conformation of the overlapping domain.



**Fig. S7.**

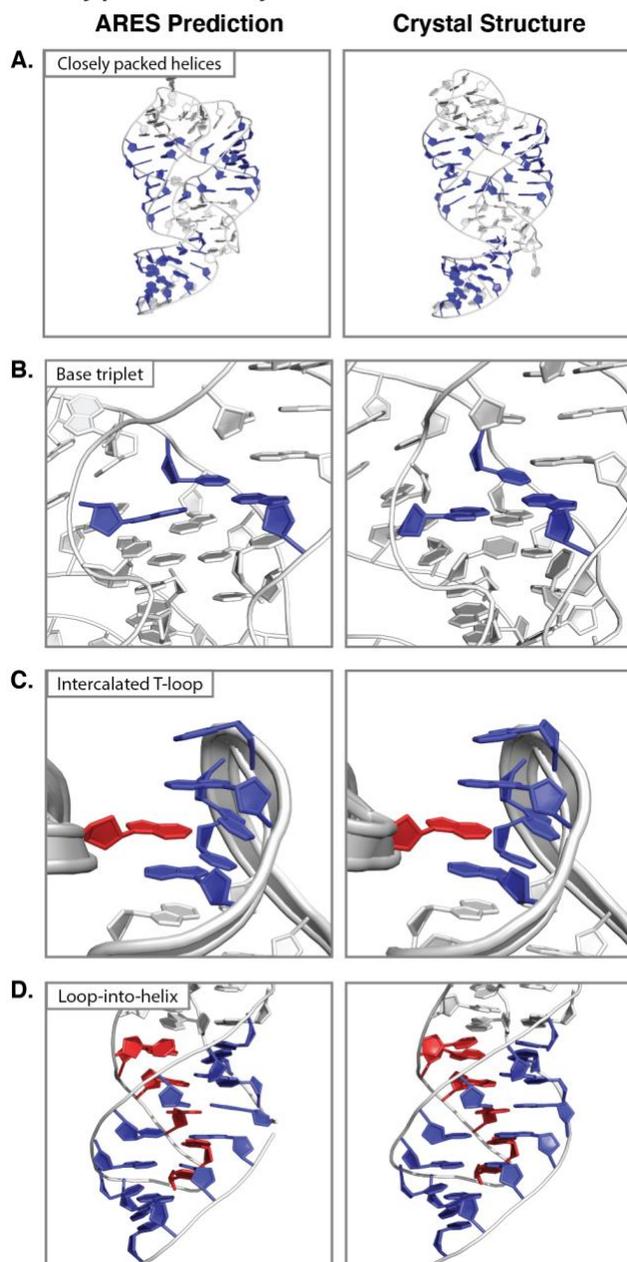
**ARES accurately recovers a structural motif that differs substantially from any in the structures used to train ARES.** **A**, In the hinge region shown above (RNA A, PDB code 6OL3, the adenovirus VA RNA I), a pseudoknot (red) defines the geometry of an unusual three-way junction (helices forming junction in blue). **B**, ARES accurately recovers the 3D geometry of this interaction, and with it the global fold of RNA A (Figure 3). **C**, The most accurate RNA-Puzzles submission for this RNA from any other method( specifically, from Rosetta) recovers this hinge region much less accurately. Inter-stem angles shown in top right of each panel are calculated between bases 59, 94, and 137, with base numbering as in 6OL3. We note that the hinge region bears no resemblance to any interaction found in the FARFAR2-Classics dataset used to train ARES. FARFAR2-Classics contains six duplexes, eight stem-loops, two H-type pseudoknots, and a hook-turn. Nowhere does a multiway junction appear, let alone one whose geometry is constrained by a pseudoknot.



**Fig. S8.**

**Further blind challenge visualizations.** RNAs are as defined in Figure 3 (B: PDB code 6PMO, the *Geobacillus kaustophilus* T-box discriminator-tRNA<sup>Gly</sup>; C: PDB code 6POM, the *Bacillus subtilis* T-box-tRNA<sup>Gly</sup>; D: PDB code 6UFM, the *Nocardia farcinic* T-box-tRNA<sup>Ile</sup>). For RNA B, the best ARES structural model (blue) achieved a 12.5 Å RMSD to the experimentally determined structure (green). The best competing structural model (orange) for RNA B, identified by human experts in the Das lab, achieved an RMSD of 13.3 Å. For RNA C, the best ARES structural model achieved 9.5 Å RMSD. The best competing structural model for RNA C, identified by Rosetta automated selection and also by human experts in the Das lab, achieved an RMSD of 10.0 Å. For RNA D, the best ARES structural model achieved a 14.5 Å RMSD. The best competing structural model for RNA D, submitted by the Adamiak group, achieved an RMSD of 18.2 Å.

Closely packed tertiary structure results



**Fig. S9.**

**ARES correctly identifies intimate tertiary interactions.** We examine the structural model with the best ARES score for each of three RNA molecules. Sugar and base atoms of nucleotides involved in highlighted interactions are colored in blue and/or red. **A**, For the *Vibrio vulnificus* purine riboswitch (PDB code 1Y26), ARES accurately places three closely packed helices. **B**, For the same purine riboswitch, ARES correctly identifies a base triplet motif (in blue). **C**, For a synthetic hammerhead ribozyme (RNA Puzzle 15, PDB code 5DI4), ARES identifies an

intercalated T-loop motif (T-loop in blue, intercalated base in red). **D**, For the *Thermobifida fusca* guanidinium-III riboswitch (RNA Puzzle 21, PDB code 5NWQ), ARES identifies a loop-into-helix motif (helix in blue, loop in red).



METHOD	RNA			
	A	B	C	D
<b>ARES</b>	<b>5.6</b>	<b>15.2</b>	<b>11.9</b>	<b>17.5</b>
Adamiak	12.0	22.1	24.3	23.4
Bujnicki	11.2	17.1	21.9	28.5
Chen	13.2	21.8	14.4	41.7
Ding	30.3	25.0	—	52.0
Das (Human)	15.8	16.1	12.7	34.7
iFoldRNA	12.1	—	108.3	36.2
RNAComposer	12.6	23.7	18.7	24.3
Rosetta	9.0	18.0	12.7	27.0
SimRNA	20.1	21.7	88.2	34.3
Xiao	18.9	40.5	43.9	48.3

**Table S2.**

**Blind RNA structure prediction results quantified with deformation index metric.** We report the deformation index across base pairing and base stacking interactions (DI all) (65) for the blind challenge submissions. Each entry in the table represents the best (lowest) DI across the models submitted for one method, as reported by the RNA-Puzzles organizers. RNAs are lettered as in Figure 3 (RNA A: the Adenovirus VA-I RNA, PDB code 6OL3; RNA B: the *Geobacillus kaustophilus* T-box discriminator-tRNA<sup>Gly</sup>, PDB code 6PMO; RNA C: the *Bacillus subtilis* T-box-tRNA<sup>Gly</sup>, PDB code 6POM; and RNA D: the *Nocardia farcinic* T-box-tRNA<sup>Ile</sup>, PDB code 6UFM). For all four RNAs, ARES produced the most accurate structural model of any method as measured by the DI all metric.

METHOD	RNA			
	A	B	C	D
<b>ARES</b>	<b>4.8</b>	<b>12.5</b>	9.5	<b>14.5</b>
3dRNAscore	9.2	14.9	<b>8.9</b>	24.6
RASP	<b>4.8</b>	14.1	9.8	24.8
RNA3DCNN	7.7	13.9	10.3	25.0

**Table S3.**

**Analysis of scoring method accuracies on blind prediction challenges.** ARES outperforms other scoring methods run on the same model sets used to select its own blind challenge submissions. These other methods were run after the experimentally determined structures became available, using the same ranking and clustering procedure as in the ARES submissions. For each of these methods, the tables shows the RMSD (Å) of the most accurate of the ten structural models that would have been submitted. RNA3DCNN (23) included the benchmark set in its training set and therefore could not be compared in other assessments. RNAs are as defined in Figure 3 (A: PDB code 6OL3, the Adenovirus VA-I RNA; B: PDB code 6PMO, the *Geobacillus kaustophilus* T-box discriminator-tRNA<sup>Gly</sup>; C: PDB code 6POM, the *Bacillus subtilis* T-box-tRNA<sup>Gly</sup>; D: PDB code 6UFM, the *Nocardia farcinic* T-box-tRNA<sup>Ile</sup>).



Exact inputs to FARFAR2 used in producing candidate structural models for each RNA in Benchmark 2. These inputs can consist of primary sequence and secondary structure. No RNA templates were employed.

METHOD	RNA	RMSD	SUBMISSION_ID	TAG	SUBSET	RANK
3dRNAScore	A	10.231		S_000007_228		1
3dRNAScore	A	19.669		S_000007_1345		2
3dRNAScore	A	16.628		S_000011_452		3
3dRNAScore	A	9.176		S_000001_1984		4
3dRNAScore	A	17.985		S_000002_110		5
3dRNAScore	A	13.223		S_000023_1559		6
3dRNAScore	A	17.288		S_000010_753		7
3dRNAScore	A	10.686		S_000001_1325		8
3dRNAScore	A	19.640		S_000037_320		9
3dRNAScore	A	11.573		S_000008_179		10
3dRNAScore	B	23.552		340S_000040		1
3dRNAScore	B	20.168		386S_000046		2
3dRNAScore	B	15.78		23S_000042		3
3dRNAScore	B	17.33		448S_000001		4
3dRNAScore	B	17.487		756S_000047		5
3dRNAScore	B	19.3		788S_000075		6
3dRNAScore	B	18.794		693S_000014		7
3dRNAScore	B	21.831		103S_000014		8
3dRNAScore	B	20.389		692S_000039		9
3dRNAScore	B	14.875		461S_000030		10
3dRNAScore	C	10.318		1311S_000068		1
3dRNAScore	C	15.532		950S_000013		2
3dRNAScore	C	8.879		173S_000008		3
3dRNAScore	C	13.413		1264S_000061		4
3dRNAScore	C	12.383		1288S_000094		5
3dRNAScore	C	16.388		425S_000071		6
3dRNAScore	C	10.341		825S_000076		7
3dRNAScore	C	11.47		627S_000052		8
3dRNAScore	C	10.329		328S_000078		9
3dRNAScore	C	11.928		1126S_000044		10
3dRNAScore	D	29.828		783S_000017	no_cst	1
3dRNAScore	D	24.666		897S_000060	cst	2
3dRNAScore	D	32.915		1159S_000093	cst	3
3dRNAScore	D	32.963		507S_000099	cst	4
3dRNAScore	D	24.550		467S_000027	cst	5
3dRNAScore	D	30.861		256S_000009	no_cst	6
3dRNAScore	D	28.652		651S_000005	no_cst	7
3dRNAScore	D	27.617		1903S_000003	no_cst	8
3dRNAScore	D	35.452		397S_000001	no_cst	9
3dRNAScore	D	24.684		422S_000063	cst	10
ARES	A	10.865	24_DasTFN_1_rpr_0001	S_000001_1684		1
ARES	A	13.228	24_DasTFN_2_rpr_0001	S_000029_113		2
ARES	A	16.643	24_DasTFN_3_rpr_0001	S_000002_1479		3
ARES	A	21.778	24_DasTFN_4_rpr_0001	S_000001_1899		4
ARES	A	19.542	24_DasTFN_5_rpr_0001	S_000027_1292		5
ARES	A	6.53	24_DasTFN_6_rpr_0001	S_000009_1646		6
ARES	A	16.713	24_DasTFN_7_rpr_0001	S_000009_699		7
ARES	A	10.595	24_DasTFN_8_rpr_0001	S_000001_2684		8
ARES	A	8.886	24_DasTFN_9_rpr_0001	S_000007_1799		9
ARES	A	4.828	24_DasTFN_10_rpr_0001	S_000001_2963		10
ARES	B	15.78	PZ26_TFN_1_0001	957S_000037		1
ARES	B	17.66	PZ26_TFN_2_0001	248S_000050		2
ARES	B	12.472	PZ26_TFN_3_0001	700S_000015		3
ARES	B	17.295	PZ26_TFN_4_0001	812S_000012		4
ARES	B	22.753	PZ26_TFN_5_0001	425S_000027		5
ARES	B	17.107	PZ26_TFN_6_0001	170S_000039		6
ARES	B	12.763	PZ26_TFN_7_0001	184S_000018		7
ARES	B	14.756	PZ26_TFN_8_0001	436S_000007		8
ARES	B	18.029	PZ26_TFN_9_0001	902S_000042		9
ARES	B	19.92	PZ26_TFN_10_0001	506S_000011		10

ARES	C	12.949	PZ27_TFN_1_0001	867S_000066		1
ARES	C	9.454	PZ27_TFN_2_0001	407S_000004		2
ARES	C	11.227	PZ27_TFN_3_0001	596S_000003		3
ARES	C	11.589	PZ27_TFN_4_0001	654S_000048		4
ARES	C	13.691	PZ27_TFN_5_0001	944S_000046		5
ARES	C	12.476	PZ27_TFN_6_0001	461S_000063		6
ARES	C	13.303	PZ27_TFN_7_0001	176S_000044		7
ARES	C	15.104	PZ27_TFN_8_0001	216S_000015		8
ARES	C	11.559	PZ27_TFN_9_0001	133S_000010		9
ARES	C	10.866	PZ27_TFN_10_0001	115S_000021		10
ARES	D	25.162	PZ28_TFN_1_0001	590S_000004	no_cst	1
ARES	D	35.976	PZ28_TFN_2_0001	921S_000013	no_cst	2
ARES	D	14.49	PZ28_TFN_3_0001	1394S_000008	no_cst	3
ARES	D	31.924	PZ28_TFN_4_0001	1660S_000011	no_cst	4
ARES	D	30.054	PZ28_TFN_5_0001	1206S_000006	no_cst	5
ARES	D	31.364	PZ28_TFN_6_0001	1194S_000023	cst	6
ARES	D	30.525	PZ28_TFN_7_0001	489S_000006	cst	7
ARES	D	33.036	PZ28_TFN_8_0001	1801S_000007	no_cst	8
ARES	D	29.111	PZ28_TFN_9_0001	346S_000003	no_cst	9
ARES	D	23.035	PZ28_TFN_10_0001	534S_000001	no_cst	10
Das (Human)	A	15.866	24_Das_1_rpr_0001	n/a		1
Das (Human)	A	17.21	24_Das_2_rpr_0001	n/a		2
Das (Human)	A	16.757	24_Das_3_rpr_0001	n/a		3
Das (Human)	A	14.965	24_Das_4_rpr_0001	n/a		4
Das (Human)	A	14.323	24_Das_5_rpr_0001	n/a		5
Das (Human)	A	14.525	24_Das_6_rpr_0001	n/a		6
Das (Human)	A	13.57	24_Das_7_rpr_0001	n/a		7
Das (Human)	A	14.882	24_Das_8_rpr_0001	n/a		8
Das (Human)	A	14.056	24_Das_9_rpr_0001	n/a		9
Das (Human)	A	16.417	24_Das_10_rpr_0001	n/a		10
Das (Human)	B	15.243	PZ26_Das_1_0001	n/a		1
Das (Human)	B	14.302	PZ26_Das_2_0001	n/a		2
Das (Human)	B	15.631	PZ26_Das_3_0001	n/a		3
Das (Human)	B	16.832	PZ26_Das_4_0001	n/a		4
Das (Human)	B	15.26	PZ26_Das_5_0001	n/a		5
Das (Human)	B	17.084	PZ26_Das_6_0001	n/a		6
Das (Human)	B	16.277	PZ26_Das_7_0001	n/a		7
Das (Human)	B	17.57	PZ26_Das_8_0001	n/a		8
Das (Human)	B	13.329	PZ26_Das_9_0001	n/a		9
Das (Human)	B	17.483	PZ26_Das_10_0001	n/a		10
Das (Human)	C	12.219	PZ27_Das_1_0001	n/a		1
Das (Human)	C	10.046	PZ27_Das_2_0001	n/a		2
Das (Human)	C	12.12	PZ27_Das_3_0001	n/a		3
Das (Human)	C	14.725	PZ27_Das_4_0001	n/a		4
Das (Human)	C	10.63	PZ27_Das_5_0001	n/a		5
Das (Human)	C	11.742	PZ27_Das_6_0001	n/a		6
Das (Human)	C	12.042	PZ27_Das_7_0001	n/a		7
Das (Human)	C	10.788	PZ27_Das_8_0001	n/a		8
Das (Human)	C	11.774	PZ27_Das_9_0001	n/a		9
Das (Human)	C	13.149	PZ27_Das_10_0001	n/a		10
Das (Human)	D	30.843	PZ28_Das_1_0001	n/a		1
Das (Human)	D	35.372	PZ28_Das_2_0001	n/a		2
Das (Human)	D	31.461	PZ28_Das_3_0001	n/a		3
Das (Human)	D	31.445	PZ28_Das_4_0001	n/a		4
Das (Human)	D	32.647	PZ28_Das_5_0001	n/a		5
Das (Human)	D	31.702	PZ28_Das_6_0001	n/a		6
Das (Human)	D	31.239	PZ28_Das_7_0001	n/a		7
Das (Human)	D	30.989	PZ28_Das_8_0001	n/a		8
Das (Human)	D	33.151	PZ28_Das_9_0001	n/a		9
Das (Human)	D	28.84	PZ28_Das_10_0001	n/a		10
RASP	A	9.42		S_000014_030		1
RASP	A	13.223		S_000023_1559		2
RASP	A	16.777		S_000001_2849		3

RASP	A	6.083		S_000001_1308		4
RASP	A	8.580		S_000003_1458		5
RASP	A	10.665		S_000009_077		6
RASP	A	15.639		S_000005_1330		7
RASP	A	19.045		S_000001_707		8
RASP	A	17.059		S_000022_035		9
RASP	A	4.828		S_000001_2963		10
RASP	B	17.382		276S_000054		1
RASP	B	18.271		61S_000062		2
RASP	B	15.612		480S_000009		3
RASP	B	21.219		756S_000041		4
RASP	B	23.21		555S_000013		5
RASP	B	16.352		652S_000058		6
RASP	B	21.594		6S_000060		7
RASP	B	14.062		289S_000053		8
RASP	B	14.974		411S_000024		9
RASP	B	16.193		259S_000063		10
RASP	C	11.141		952S_000047		1
RASP	C	12.767		1104S_000030		2
RASP	C	11.426		1099S_000003		3
RASP	C	11.3		274S_000057		4
RASP	C	12.696		1165S_000014		5
RASP	C	13.03		1120S_000075		6
RASP	C	11.25		331S_000054		7
RASP	C	13.084		1069S_000031		8
RASP	C	10.852		471S_000063		9
RASP	C	9.761		196S_000028		10
RASP	D	29.828		783S_000017	no_cst	1
RASP	D	35.976		921S_000013	no_cst	2
RASP	D	27.702		1572S_000007	no_cst	3
RASP	D	24.815		1679S_000004	no_cst	4
RASP	D	29.020		825S_000039	cst	5
RASP	D	32.963		507S_000099	cst	6
RASP	D	31.364		1194S_000023	cst	7
RASP	D	30.231		155S_000014	no_cst	8
RASP	D	26.681		371S_000061	cst	9
RASP	D	27.768		639S_000029	full_naive_28	10
RNA3DCNN	A	16.550		S_000001_2992		1
RNA3DCNN	A	14.045		S_000006_2657		2
RNA3DCNN	A	16.726		S_000009_181		3
RNA3DCNN	A	16.628		S_000011_452		4
RNA3DCNN	A	9.420		S_000014_030		5
RNA3DCNN	A	10.449		S_000002_084		6
RNA3DCNN	A	7.710		S_000001_2510		7
RNA3DCNN	A	12.112		S_000008_1580		8
RNA3DCNN	A	14.886		S_000005_1029		9
RNA3DCNN	A	19.050		S_000011_643		10
RNA3DCNN	B	16.193		259S_000063		1
RNA3DCNN	B	18.059		328S_000085		2
RNA3DCNN	B	17.112		596S_000019		3
RNA3DCNN	B	18.617		350S_000010		4
RNA3DCNN	B	16.099		56S_000006		5
RNA3DCNN	B	13.861		682S_000011		6
RNA3DCNN	B	18.679		545S_000027		7
RNA3DCNN	B	17.815		717S_000056		8
RNA3DCNN	B	14.056		817S_000079		9
RNA3DCNN	B	18.546		23S_000009		10
RNA3DCNN	C	12.789		1431S_000056		1
RNA3DCNN	C	10.944		658S_000041		2
RNA3DCNN	C	10.497		1219S_000032		3
RNA3DCNN	C	12.042		271S_000062		4
RNA3DCNN	C	13.143		1440S_000008		5
RNA3DCNN	C	10.256		914S_000063		6

RNA3DCNN	C	10.75		1035S_000030		7
RNA3DCNN	C	11.123		1128S_000074		8
RNA3DCNN	C	11.938		934S_000068		9
RNA3DCNN	C	12.345		191S_000046		10
RNA3DCNN	D	12.843		509S_000021	cst	1
RNA3DCNN	D	24.957		340S_000009	no_cst	2
RNA3DCNN	D	25.627		825S_000039	cst	3
RNA3DCNN	D	27.635		1130S_000013	no_cst	4
RNA3DCNN	D	33.111		539S_000026	no_cst	5
RNA3DCNN	D	32.581		632S_000074	full_naive_28	6
RNA3DCNN	D	32.789		1074S_000002	no_cst	7
RNA3DCNN	D	27.189		1422S_000011	no_cst	8
RNA3DCNN	D	29.828		783S_000017	no_cst	9
RNA3DCNN	D			1572S_000010	no_cst	10
Rosetta	A	16.643	24_FARFAR2_1_rpr_0001	S_000002_1479		1
Rosetta	A	16.741	24_FARFAR2_2_rpr_0001	S_000008_153		2
Rosetta	A	19.045	24_FARFAR2_3_rpr_0001	S_000001_707		3
Rosetta	A	13.461	24_FARFAR2_4_rpr_0001	S_000007_1249		4
Rosetta	A	18.381	24_FARFAR2_5_rpr_0001	S_000002_2275		5
Rosetta	A	8.886	24_FARFAR2_6_rpr_0001	S_000007_1799		6
Rosetta	A	12.112	24_FARFAR2_7_rpr_0001	S_000008_1580		7
Rosetta	A	12.948	24_FARFAR2_8_rpr_0001	S_000009_1583		8
Rosetta	A	7.636	24_FARFAR2_9_rpr_0001	S_000001_2948		9
Rosetta	A	10.113	24_FARFAR2_10_rpr_0001	S_000001_1304		10
Rosetta	B	14.302	PZ26_FARFAR2_1_0001	623S_000057		1
Rosetta	B	16.832	PZ26_FARFAR2_2_0001	205S_000077		2
Rosetta	B	15.26	PZ26_FARFAR2_3_0001	96S_000070		3
Rosetta	B	17.084	PZ26_FARFAR2_4_0001	111S_000002		4
Rosetta	B	14.888	PZ26_FARFAR2_5_0001	267S_000043		5
Rosetta	B	16.277	PZ26_FARFAR2_6_0001	772S_000058		6
Rosetta	B	17.57	PZ26_FARFAR2_7_0001	264S_000069		7
Rosetta	B	15.015	PZ26_FARFAR2_8_0001	789S_000021		8
Rosetta	B	19.795	PZ26_FARFAR2_9_0001	218S_000064		9
Rosetta	B	15.631	PZ26_FARFAR2_10_0001	634S_000002		10
Rosetta	C	10.788	PZ27_FARFAR2_1_0001	917S_000001		1
Rosetta	C	10.046	PZ27_FARFAR2_2_0001	699S_000081		2
Rosetta	C	12.219	PZ27_FARFAR2_3_0001	262S_000075		3
Rosetta	C	12.12	PZ27_FARFAR2_4_0001	451S_000060		4
Rosetta	C	14.725	PZ27_FARFAR2_5_0001	177S_000089		5
Rosetta	C	10.308	PZ27_FARFAR2_6_0001	588S_000029		6
Rosetta	C	12.042	PZ27_FARFAR2_7_0001	271S_000062		7
Rosetta	C	10.63	PZ27_FARFAR2_8_0001	264S_000055		8
Rosetta	C	11.742	PZ27_FARFAR2_9_0001	302S_000065		9
Rosetta	C	16.525	PZ27_FARFAR2_10_0001	414S_000035		10
Rosetta	D	29.828	PZ28_FARFAR2_1_0001	783S_000017	nocst	1
Rosetta	D	30.241	PZ28_FARFAR2_2_0001	701S_000025	cst	2
Rosetta	D	22.141	PZ28_FARFAR2_3_0001	10S_000005	nocst	3
Rosetta	D	33.932	PZ28_FARFAR2_4_0001	1831S_000008	nocst	4
Rosetta	D	35.164	PZ28_FARFAR2_5_0001	1412S_000001	nocst	5
Rosetta	D	35.372	PZ28_FARFAR2_6_0001	478S_000033	cst	6
Rosetta	D	34.574	PZ28_FARFAR2_7_0001	776S_000057	cst	7
Rosetta	D	28.84	PZ28_FARFAR2_8_0001	650S_000062	full_naive_28	8
Rosetta	D	34.954	PZ28_FARFAR2_9_0001	867S_000075	full_naive_28	9
Rosetta	D	29.97	PZ28_FARFAR2_10_0001	424S_000067	full_naive_28	10

**Table S5.**

Comparison of the models selected by ARES and other scoring functions for the 4 RNAs shown in Figure 3 (RNA A: the Adenovirus VA-I RNA, PDB code 6OL3; RNA B: the *Geobacillus kaustophilus* T-box discriminator-tRNA<sup>Gly</sup>, PDB code 6PMO; RNA C: the *Bacillus subtilis* T-

box-tRNA<sup>Gly</sup>, PDB code 6POM; and RNA D: the *Nocardia farcinic* T-box-tRNA<sup>Ile</sup>, PDB code 6UFM). For each RNA, each method ranks the same set of structural models, all generated by the FARFAR2 sampling procedure. For each scoring function, we list the 10 best-ranked structural models for each RNA, as the allow RNA Puzzles competition allows submission of ten models per method (12). We report the official submissions (ARES, Das (Human), and Rosetta) to the RNA Puzzles competition, as well as the structural models selected by the other scoring functions listed in Table S3. The official submissions have their public submission identifier included with them. For RNA D, three different subsets were analyzed, as documented in Table S1. The human experts in the Das lab also considered additional candidate structural models from simulations that featured manual buildup of RNA modules rather than purely relying on the automated FARFAR2 protocol; these additional models have no associated tag. All submitted models for the blind challenges may be found at <https://www.rnapuzzles.org>.

## References and Notes:

1. T. R. Cech, J. A. Steitz, The noncoding RNA revolution-trashing old rules to forge new ones. *Cell*. **157**, 77–94 (2014).
2. K. D. Warner, C. E. Hajdin, K. M. Weeks, Principles for targeting RNA with drug-like small molecules. *Nat. Rev. Drug Discov.* **17**, 547–558 (2018).
3. A. Churkin, M. D. Retwitzer, V. Reinharz, Y. Ponty, J. Waldispühl, D. Barash, Design of RNAs: Comparing programs for inverse RNA folding. *Brief. Bioinform.* **19**, 350–358 (2018).
4. ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature*. **489**, 57–74 (2012).
5. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
6. S. Jain, D. C. Richardson, J. S. Richardson, Computational methods for RNA structure validation and improvement. *Methods Enzymol.* **558**, 181–212 (2015).
7. D. S. Marks, L. J. Colwell, R. Sheridan, T. A. Hopf, A. Pagnani, R. Zecchina, C. Sander, Protein 3D structure computed from evolutionary sequence variation. *PLOS ONE* **6**, e28766 (2011).
8. A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Žídek, A. W. R. Nelson, A. Bridgland, H. Penedones, S. Petersen, K. Simonyan, S. Crossan, P. Kohli, D. T. Jones, D. Silver, K. Kavukcuoglu, D. Hassabis, Improved protein structure prediction using potentials from deep learning. *Nature* **577**, 706–710 (2020).
9. H. Kamisetty, S. Ovchinnikov, D. Baker, Assessing the utility of coevolution-based

- residue-residue contact predictions in a sequence- and structure-rich era. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15674–15679 (2013).
10. F. Pucci, M. B. Zerihun, E. K. Peter, A. Schug, Evaluating DCA-based method performances for RNA contact prediction by a well-curated data set. *RNA* **26**, 794–802 (2020).
  11. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015).
  12. Materials and methods are available as supplementary materials.
  13. D. E. Worrall, S. J. Garbin, D. Turmukhambetov, G. J. Brostow, "Harmonic networks: deep translation and rotation equivariance" in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2017), pp. 7168–7177.
  14. B. Anderson, T. S. Hy, R. Kondor, "Cormorant: Covariant Molecular Neural Networks" in *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, R. Garnett, Eds. (Curran Associates, 2019), pp. 14537–14546.
  15. M. Weiler, M. Geiger, M. Welling, W. Boomsma, T. S. Cohen, "3D Steerable CNNs: Learning Rotationally Equivariant Features in Volumetric Data" in *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, R. Garnett, Eds. (Curran Associates, 2018), pp. 10381–10392.
  16. N. Thomas, T. Smidt, S. Kearnes, L. Yang, L. Li, K. Kohlhoff, P. Riley, Tensor Field Networks: Rotation- and Translation-Equivariant Neural Networks for 3D Point Clouds. [arXiv:1802.08219](https://arxiv.org/abs/1802.08219) [cs.LG] (2018).

17. S. Eismann, R. J. L. Townshend, N. Thomas, M. Jagota, B. Jing, R. O. Dror, Hierarchical, rotation-equivariant neural networks to select structural models of protein complexes. *Proteins* **89**, 493–501 (2021).
18. R. Das, D. Baker, Automated de novo prediction of native-like RNA tertiary structures. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14664–14669 (2007).
19. A. M. Watkins, R. Rangan, R. Das, FARFAR2: Improved De Novo Rosetta Prediction of Complex Global RNA Folds. *Structure* **28**, 963-976.e6 (2020).
20. Z. Miao, R. W. Adamiak, M. Antczak, M. J. Boniecki, J. Bujnicki, S.-J. Chen, C. Y. Cheng, Y. Cheng, F.-C. Chou, R. Das, N. V. Dokholyan, F. Ding, C. Geniesse, Y. Jiang, A. Joshi, A. Krokhotin, M. Magnus, O. Mailhot, F. Major, T. H. Mann, P. Piątkowski, R. Pluta, M. Popena, J. Sarzynska, L. Sun, M. Szachniuk, S. Tian, J. Wang, J. Wang, A. M. Watkins, J. Wiedemann, Y. Xiao, X. Xu, J. D. Yesselman, D. Zhang, Y. Zhang, Z. Zhang, C. Zhao, P. Zhao, Y. Zhou, T. Zok, A. Żyła, A. Ren, R. T. Batey, B. L. Golden, L. Huang, D. M. Lilley, Y. Liu, D. J. Patel, E. Westhof, RNA-Puzzles Round IV: 3D structure predictions of four ribozymes and two aptamers. *RNA* **26**, 982–995 (2020).
21. E. Capriotti, T. Norambuena, M. A. Marti-Renom, F. Melo, All-atom knowledge-based potential for RNA structure prediction and assessment. *Bioinformatics* **27**, 1086–1093 (2011).
22. J. Wang, Y. Zhao, C. Zhu, Y. Xiao, 3dRNAscore: A distance and torsion angle dependent evaluation function of 3D RNA structures. *Nucleic Acids Res.* **43**, e63 (2015).
23. J. Li, W. Zhu, J. Wang, W. Li, S. Gong, J. Zhang, W. Wang, RNA3DCNN: Local and global quality assessments of RNA 3D structures using 3D deep convolutional neural

- networks. *PLOS Comput. Biol.* **14**, e1006514 (2018).
24. J. Behler, M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
  25. J. Xu, Distance-based protein folding powered by deep learning. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 16856–16865 (2019).
  26. J. S. Smith, O. Isayev, A. E. Roitberg, ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **8**, 3192–3203 (2017).
  27. M. Ragoza, J. Hochuli, E. Idrobo, J. Sunseri, D. R. Koes, Protein–Ligand Scoring with Convolutional Neural Networks. *J. Chem. Inf. Model.* **57**, 942–957 (2017).
  28. K. Xu, Z. Wang, J. Shi, H. Li, Q. C. Zhang, "A<sup>2</sup>-Net: Molecular Structure Estimation from Cryo-EM Density Volumes" in *Proceedings of the AAAI Conference on Artificial Intelligence* (AAAI Press, 2019), pp. 1230–1237.
  29. F. Noé, S. Olsson, J. Köhler, H. Wu, Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. *Science* **365**, eaaw1147 (2019).
  30. M. Alquraishi, End-to-End Differentiable Learning of Protein Structure. *Cell Syst.* **8**, 292–301.e3 (2019).
  31. K. M. Kutchko, A. Laederach, Transcending the prediction paradigm: Novel applications of SHAPE to RNA function and evolution. *WIREs RNA* **8**, e1374 (2017).
  32. R. J.L. Townshend, S. Eismann, A. M. Watkins, R. Rangan, M. Karelina, R. Das, R. O. Dror, Training code for ARES neural network, Version 1.0, Zenodo (2021); [doi:10.5281/zenodo.5088971](https://doi.org/10.5281/zenodo.5088971).

33. R. J.L. Townshend, S. Eismann, A. M. Watkins, R. Rangan, M. Karelina, R. Das, R. O. Dror, ARES-specific adaptation of E3NN, Version 1.0, Zenodo (2021); [doi:10.5281/zenodo.5090151](https://doi.org/10.5281/zenodo.5090151).
34. R. J.L. Townshend, S. Eismann, A. M. Watkins, R. Rangan, M. Karelina, R. Das, R. O. Dror, Auxiliary code related to the publication "Geometric Deep Learning of RNA Structure", Version 1.0, Zenodo (2021); [doi:10.5281/zenodo.5090157](https://doi.org/10.5281/zenodo.5090157).
35. R. J. L. Townshend, A. M. Watkins, S. Eismann, R. Rangan, M. Karelina, R. Das, R. O. Dror, Structural data used to train, test, and characterize a new geometric deep learning RNA scoring function, Stanford Digital Repository (2021); <https://doi.org/10.25740/bn398fc4306>
36. A. M. Watkins, R. Rangan, R. J. L. Townshend, S. Eismann, M. Karelina, R. O. Dror, R. Das, Structural data used to test a new geometric deep learning RNA scoring function emulating fully de novo modeling conditions, Stanford Digital Repository (2021); <https://doi.org/10.25740/sq987cc0358>
37. T. S. Cohen, M. Welling, "Group equivariant convolutional networks" in *Proceedings of the 33rd International Conference on Machine Learning*, M. F. Balcan, K. Q. Weinberger, Eds. (PMLR, 2016), pp. 2990–2999.
38. R. H. R. Hahnloser, R. Sarpeshkar, M. A. Mahowald, R. J. Douglas, H. S. Seung, Digital selection and analogue amplification coexist in a cortex- inspired silicon circuit. *Nature*. **405**, 947–951 (2000).
39. M. Reisert, H. Burkhardt, "Spherical tensor calculus for local adaptive filtering" in *Tensors in Image Processing and Computer Vision*, S. Aja-Fernández, R. de Luis García,

- D. Tao, X. Li, Eds. (Springer, 2009), pp. 153–178.
40. K. T. Schütt, P.-J. Kindermans, H. E. Saucedo, S. Chmiela, A. Tkatchenko, K.-R. Müller, "SchNet: A continuous-filter convolutional neural network for modeling quantum interactions" in *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, U. von Luxburg, I. Guyon, S. Bengio, H. Wallach, R. Fergus, Eds. (Curran Associates, 2017), pp. 992–1002.
41. D. A. Clevert, T. Unterthiner, S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (ELUs)" in *Proceedings of the 4th International Conference on Learning Representations (ICLR 2016)*, Y. Bengio, Y. LeCun, Eds. (2016).
42. X. Glorot, Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks" in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, Y. W. Teh, M. Titterton, Eds. (PMLR, 2010), pp. 249–256.
43. D. P. Kingma, J. L. Ba, "Adam: A method for stochastic optimization" in *Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015)*, Y. Bengio, Y. LeCun, Eds. (2015).
44. P. J. Huber, Robust Estimation of a Location Parameter. *Ann. Math. Stat.* **35**, 73–101 (1964).
45. Z. Miao, E. Westhof, RNA Structure: Advances and Assessment of 3D Structure Prediction. *Annu. Rev. Biophys.* **46**, 483–503 (2017).
46. A. Sergeev, M. Del Balso, Horovod: fast and easy distributed deep learning in TensorFlow. [arXiv:1802.05799](https://arxiv.org/abs/1802.05799) [cs.LG] (2018).
47. I. Kalvari, E. P. Nawrocki, N. Ontiveros-Palacios, J. Argasinska, K. Lamkiewicz, M.

- Marz, S. Griffiths-Jones, C. Toffano-Nioche, D. Gautheret, Z. Weinberg, E. Rivas, S. R. Eddy, R. D. Finn, A. Bateman, A. I. Petrov, Rfam 14: expanded coverage of metagenomic, viral and microRNA families. *Nucleic Acids Res.* **49**, D192–D200 (2021).
48. C. B. Do, D. A. Woods, S. Batzoglou, CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics.* **22**, e90–e98 (2006).
49. M. E. Sherlock, N. Sudarsan, S. Stav, R. R. Breaker, Tandem riboswitches form a natural Boolean logic gate to control purine metabolism in bacteria. *eLife* **7**, e33908 (2018).
50. E. B. Porter, J. G. Marcano-Velázquez, R. T. Batey, The purine riboswitch as a model system for exploring RNA biology and chemistry. *Biochim. Biophys. Acta* **1839**, 919–930 (2014).
51. M. J. Boniecki, G. Lach, W. K. Dawson, K. Tomala, P. Lukasz, T. Soltysinski, K. M. Rother, J. M. Bujnicki, SimRNA: A coarse-grained method for RNA folding simulations and 3D structure prediction. *Nucleic Acids Res* **44**, e63 (2016).
52. R. Das, J. Karanicolas, D. Baker, Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat. Methods* **7**, 291–294 (2010).
53. S. Chaudhury, S. Lyskov, J. J. Gray, PyRosetta: A script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics* **26**, 689–691 (2010).
54. N. B. Leontis, C. L. Zirbel, "Nonredundant 3D structure datasets for RNA knowledge extraction and benchmarking" in *RNA 3D Structure Analysis and Prediction*, N. Leontis, E. Westhof, Eds. (Springer, 2012), pp. 281–298.
55. X. Lu, W. K. Olson, 3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **31**, 5108–

- 5121 (2003).
56. K. Launer-Felty, C. J. Wong, J. L. Cole, Structural analysis of adenovirus VAI RNA defines the mechanism of inhibition of PKR. *Biophys. J.* **108**, 748–757 (2015).
57. J. Zhang, A. R. Ferré-D’Amaré, Direct evaluation of tRNA aminoacylation status by the T-box riboswitch using tRNA-mRNA stacking and steric readout. *Mol. Cell* **55**, 148–155 (2014).
58. A. V. Sherwood, F. J. Grundy, T. M. Henkin, T box riboswitches in Actinobacteria: Translational regulation via novel tRNA interactions. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 1113–1118 (2015).
59. I. V. Hood, J. M. Gordon, C. Bou-Nader, F. E. Henderson, S. Bahmanjah, J. Zhang, Crystal structure of an adenovirus virus-associated RNA. *Nat. Commun.* **10**, 2871 (2019).
60. S. Li, Z. Su, J. Lehmann, V. Stamatopoulou, N. Giarimoglou, F. E. Henderson, L. Fan, G. D. Pintilie, K. Zhang, M. Chen, S. J. Ludtke, Y.-X. Wang, C. Stathopoulos, W. Chiu, J. Zhang, Structural basis of amino acid surveillance by higher-order tRNA-mRNA interactions. *Nat. Struct. Mol. Biol.* **26**, 1094–1105 (2019).
61. K. C. Suddala, J. Zhang, High-affinity recognition of specific tRNAs by an mRNA anticodon-binding groove. *Nat. Struct. Mol. Biol.* **26**, 1114–1122 (2019).
62. K. Kappel, K. Zhang, Z. Su, A. M. Watkins, W. Kladwang, S. Li, G. Pintilie, V. V. Topkar, R. Rangan, I. N. Zheludev, J. D. Yesselman, W. Chiu, R. Das, Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. *Nat. Methods* **17**, 699–707 (2020).
63. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N.

- Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, "PyTorch: An imperative style, high-performance deep learning library" in *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, R. Garnett, Eds. (Curran Associates, 2019), pp. 8024–8035.
64. M. Geiger, T. Smidt, A. K. Miller, W. Boomsma, B. Dice, K. Lapchevskyi, M. Weiler, M. Tyszkiewicz, S. Batzner, J. Frellsen, N. Jung, S. Sanborn, J. Rackers, M. Bailey, E3NN, GitHub (2021); <https://github.com/e3nn/e3nn>.
65. M. Parisien, J. A. Cruz, É. Westhof, F. Major, New metrics for comparing and assessing discrepancies between RNA 3D structures and models. *RNA* **15**, 1875–1885 (2009).