## SUPPLEMENTARY MATERIAL

# All-atom knowledge-based potential for RNA structure prediction and assessment

Emidio Capriotti, Tomas Norambuena, Marc A. Marti-Renom and Francisco Melo

### 1.1 Calculation of knowledge-based RNA potentials

Four different knowledge-based potentials were calculated. The main difference between them was the number and type of atoms used to represent a RNA nucleotide (Sup. Table 1). Pairwise distance-dependent energy score functions were calculated as previously described (Melo and Feytmans, 1997; Melo, et al., 2002), using the following equation:

$$\Delta E_k^{ij}(d) = RT\ln\left[1+M_k^{ij}\cdot\sigma\right] - RT\ln\left[1+M_k^{ij}\cdot\sigma\frac{f_k^{ij}(d)}{f_k^{xx}(d)}\right] \qquad (1)$$

where $R$ is the gas constant, $T$ is the absolute temperature, which value was set to 298 K so that $RT$ is equivalent to 0.582 kcal/mol. $M_k^{ij}$ is the total number of interactions observed between atom types $i$ and $j$ below the maximum distance range threshold (20 Å) at a given value of topological factor or sequence separation ($k$) and it was calculated as follows:

$$M_k^{ij} = \sum_{d=1}^{N} F_k^{ij}(d) \qquad (2)$$

$F_k^{ij}(d)$ is the absolute frequency of observations between atom types $i$ and $j$ at the distance class $d$, and $N$ is the total number of distance classes defined. The topological factor or sequence separation $k$ between nucleotides $n$ and $m$ is defined by $k = |m\text{-}n| - 1$, where $n$ and $m$ correspond to the observed residue indexes in the RNA chain. The potentials were calculated using a maximum distance threshold of 20 Å and distance bins of 1 Å each in the range of 0 to 20 Å. The constant weight factor $\sigma$ given to each pairwise energy score function was set to 0.02, as previously described (Sippl, 1990). $f_k^{ij}(d)$ is the relative frequency of interactions between atom types $i$ and $j$ at the distance class $d$ and sequence separation $k$, and it was defined as follows:

$$f_k^{ij} = \frac{F_k^{ij}(d)}{M_k^{ij}} \qquad (3)$$

$f_k^{xx}(d)$ is the reference system and corresponds to the relative frequency of observations between any two atom types in the distance class $d$ with sequence separation $k$. This quantity was calculated by using the following equation:

$$f_k^{xx} = \frac{\sum_{i=1}^{C}\sum_{j=1}^{C} F_k^{ij}(d)}{\sum_{i=1}^{C}\sum_{j=1}^{C}\sum_{d=1}^{N} F_k^{ij}(d)} \qquad (4)$$

where $C$ is the number of different atom types and $N$ is the number of distance classes.

All potentials were derived asymmetrically, which means that the interaction between atoms $x$ and $y$ is not necessarily equivalent to that between $y$ and $x$. The $x$ atom in this case is defined as the one located topologically closer to the 5' end of the RNA molecule. All potentials calculated here are available as supplementary data at: http://melolab.org/sup-mat.html.

## 1.2 Optimization of knowledge-based RNA potentials

The RASP variants have been calculated with different values of sequence separation (*k*) and also considering as non-local interactions all those with a sequence separation larger or equal than a given threshold (*K*). The optimal value for the separation between local and non-local interactions (*ie.* the sequence separation threshold *K*) has been optimized by calculating the information product (*IP*) at different thresholds spanning from 1 to 19 and selecting as optimal the one that resulted in an increment of *IP* smaller than 5%. The *IP* of a potential was calculated as previously described (Ferrada and Melo, 2009), by using the following equation:

$$IP = \sqrt{\bar{n}} \cdot \Delta\bar{E}^{ij}_{NL} \tag{5}$$

where $\bar{n}$ is the mean number of interactions that will be observed in a typical RNA structure when using the potential and corresponds to:

$$\bar{n} = \frac{1}{N}\sum_{i=1}^{N} n_i \tag{6}$$

where $n_i$ is the number of score events (*ie.* those interactions that will be considered by a potential according to its utilization parameters) in the native RNA structure *i* and *N* is the total number of native RNA structures used to derive the potential. $\Delta\bar{E}^{ij}_{NL}$ is the average energy score value per interaction observed in those native RNA molecules used to derive the potential:

$$\Delta\bar{E}^{ij}_{NL} = \frac{1}{X}\sum_{x=1}^{X} \Delta E^{ij}_{NL}(d) \tag{7}$$

where *X* corresponds to the total number of interactions scored in the native RNA structures when the potential was used to calculate their total score. Therefore, *X* corresponds to:

$$X = N \times \bar{n} = \sum_{i=1}^{N} n_i \tag{8}$$

In the case of the distance-dependent potentials calculated here, $\Delta\bar{E}^{ij}_{NL}$ constitutes the best estimate of mutual information because it naturally takes into account the sensible issue of sparse data in the calculation of informatics quantities and accordingly adjusts the estimate of the energy score (Solis and Rackovsky, 2006; Solis and Rackovsky, 2008).

## 1.3 Performance assessment measures

The structural divergence between native and decoy RNA structures was calculated by the root mean square deviation (RMSD) and by the GDT-TS measures. They were both computed for C3' atoms after the optimal superposition of the decoy and native structures. The RMSD was calculated as follows:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \delta_i^2} \tag{9}$$

where *N* is the total number of atoms and $\delta_i$ is the distance between the atoms *i* after the rotation that minimized the distances. The GDT-TS was the mean value of GDT at 1, 2, 4, and 8 Å distance cutoffs.

$$GDT - TS = \frac{GDT1 + GDT2 + GDT4 + GDT8}{4} \tag{10}$$

We calculated the GDT-TS as the fraction of corresponding C3' atoms that were superimposed under a selected distance threshold *x* after the rotation that minimized the distances. GDT-TS scores ranged from 0 to 1, with decoy structures very similar to native conformations resulting in GDT-TS close to 1. The Pearson correlation coefficient between the energy scores and the structure diversity measures (*ie.* RMSD and GDT-TS) was calculated as follows:

$$r \frac{n\sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{\sqrt{n\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2}\sqrt{n\sum_{i=1}^{N} y_i^2 - \left(\sum_{i=1}^{N} y_i\right)^2}} \qquad (11)$$

where $x_i$ are the values of RMSD or GDT-TS and and $y_i$ values correspond to the normalized scores.

### 1.4 Computer software that uses RASP potentials to assess RNA structures

The computer software developed here was written in C++ computer language and it is freely available to academic and non-academic users at http://melolab.org/sup-mat.html.

## References

Ferrada, E. and Melo, F. (2009) Effective knowledge-based potentials, Protein Sci. 18, 1469-1485.

Melo, F. and Feytmans, E. (1997) Novel knowledge-based mean force potential at atomic level, J. Mol. Biol. 267, 207-222.

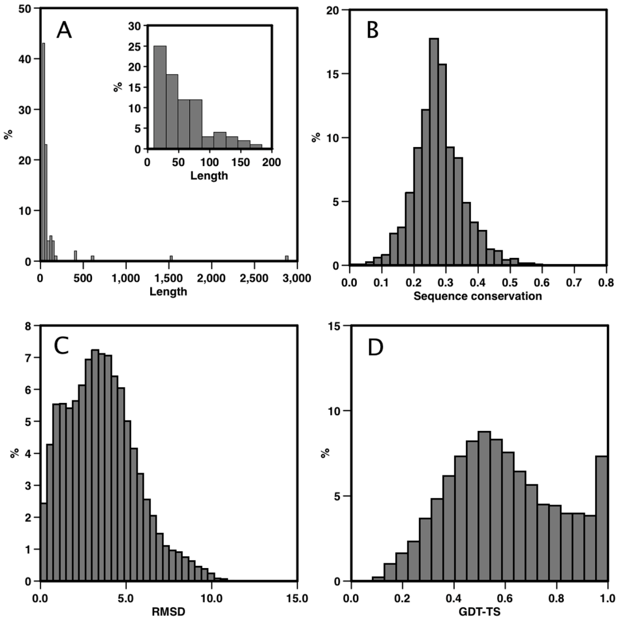
Melo, F., Sanchez, R. and Sali, A. (2002) Statistical potentials for fold assessment, Protein Sci. 11, 430-448.

Sippl, M.J. (1990) Calculation of conformational ensembles from potentials of mean force. An approach to the knowledge-based prediction of local structures in globular proteins. J. Mol. Biol. 213, 859-883.

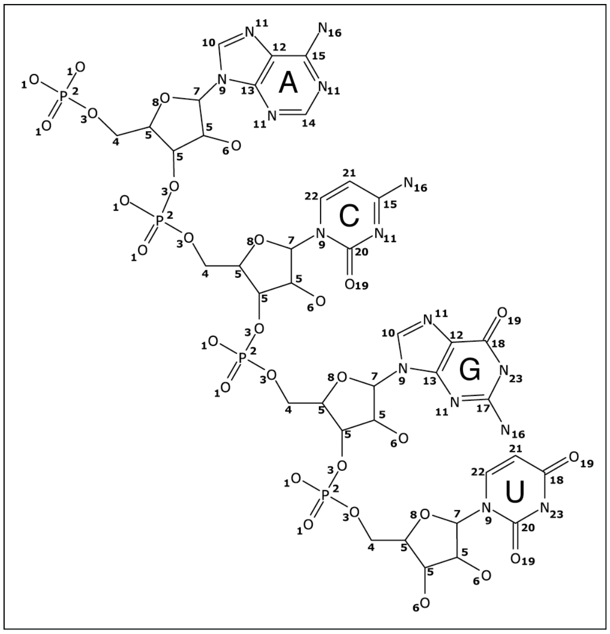
Solis, A.D. and Rackovsky, S. (2006) Improvement of statistical potentials and threading score functions using information maximization. Proteins 62, 892-908.

Solis, A.D. and Rackovsky, S. (2008) Information and discrimination in pairwise contact potentials. Proteins 71, 1071-1087.

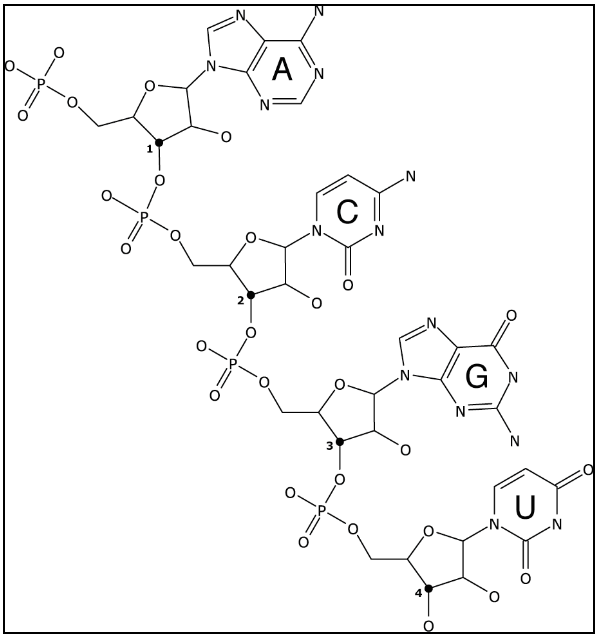
## Supplementary Figures



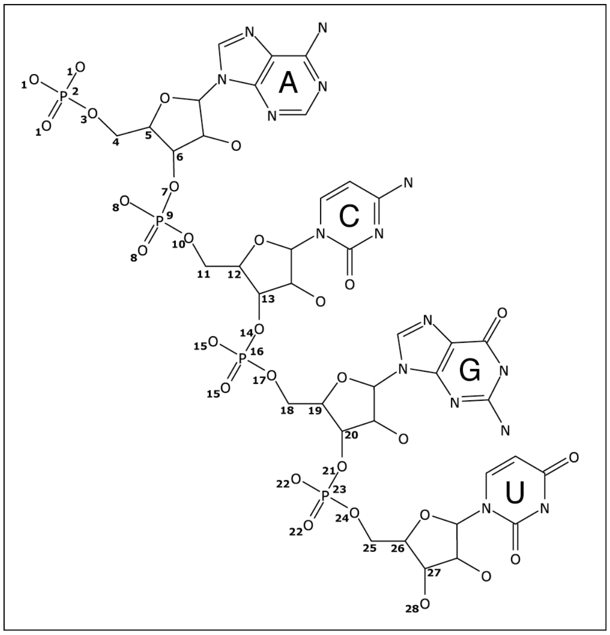
**Supplementary Figure 1.** Features of the benchmark datasets. A) Sequence length distribution for the non-redundant RNA dataset. B) Percentage sequence identity distribution for the randseq decoy set. C) Root mean square deviation (RMSD) distribution of C3’ atoms for the randstr decoy set. D) GTD-TS distribution of C3’ atoms for the randstr decoy set.



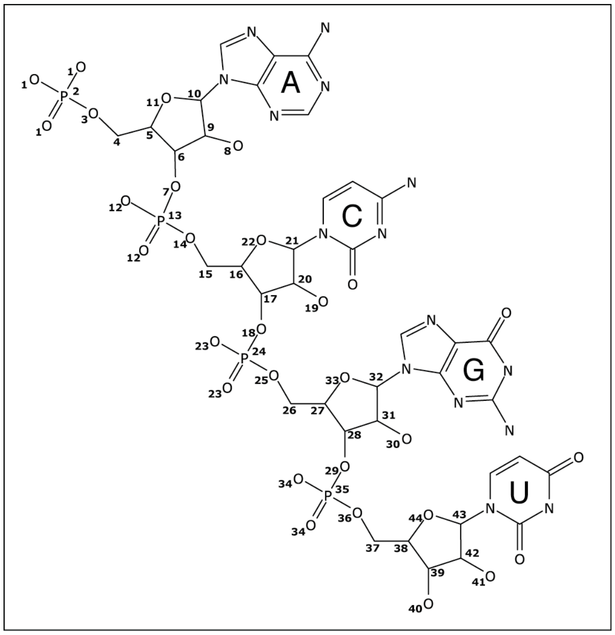
**Supplementary Fig 2.** Atom type definition in RASP-ALL potential. The definition of atom types is based on chemical nature, atom connectivity and chemical group location (backbone, ring and nitrogenated base). Note that this definition of atom types does not differentiate between common backbone and ring atoms belonging to different nucleotide groups.



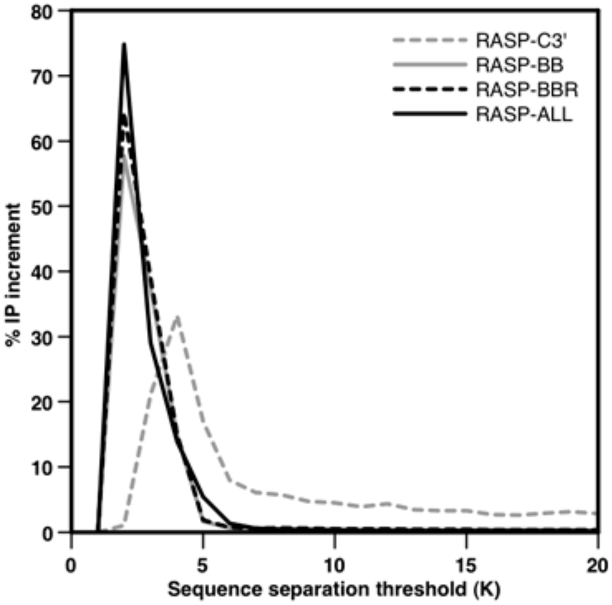
**Supplementary Figure 3**. Atom types of RASP-C3 potential



**Supplementary Figure 4.** Atom types of RASP-BB potential



**Supplementary Figure 5.** Atom types of RASP-BBR potential



**Supplementary Figure 6.** Information product (IP) variation as a function of the topological factor threshold (K) for the four statistical potentials

## Supplementary Tables

**Supplementary Table 1.** Features of RASP potentials

| Potential | Atoms† | N | *K* |
|---|---|---|---|
| RASP-C3 | C3' atoms | 4* | 8 |
| RASP-BB | Backbone atoms | 28* | 4 |
| RASP-BBR | Backbone and ribose atoms | 44* | 4 |
| RASP-ALL | All atoms | 23** | 5 |

K is the optimal topological factor threshold. N is the number of atom types. † Only non-hydrogen atoms are considered. *The detailed description of the atom type definition is provided as Sup. Tables 2 and 3. **See Sup. Figs. 2, 3, 4, 5 and 6.

**Supplementary Table 2.** Atom type definition in RASP-ALL potential

| Type | Atom Names and Chemical Groups+ | Atom Connectivity Features |
|---|---|---|
| 1 | OP1, OP2, OP3 (all nucleotides) | O-sp$^3$ or O-sp$^2$ bonded to P |
| 2 | P (all nucleotides) | P-sp$^3$ bonded to O |
| 3 | O5' (all nucleotides) | O-sp$^3$ bonded to P-sp$^3$ and C-sp$^3$ |
| 4 | C5' (all nucleotides) | C-sp$^3$ bonded to O-sp$^3$ and C-sp$^3$ |
| 5 | C4', C3', C2' (all nucleotides) | C-sp$^3$ bonded to two C-sp$^3$ and to an O-sp$^3$ |
| 6 | O2', O3'-terminal (all nucleotides) | O-sp$^3$ bonded to a C-sp$^3$ |
| 7 | C1' (all nucleotides) | C-sp$^3$ bonded to a C-sp$^3$, an O-sp$^3$ and an N-sp$^2$ |
| 8 | O4' (all nucleotides) | O-sp$^3$ bonded to two C-sp$^3$ |
| 9 | N1 (Pyrimidines); N9 (Purines) | N-sp$^2$ (pyrrolic) |
| 10 | C8 (Purines) | C-sp$^2$ bonded to a N-sp$^2$ and to a pyrrolic N-sp$^2$ |
| 11 | N3, N7 (Purines); N1 (ADE); N3 (CYT) | N-sp$^2$ bonded to two C-sp$^2$ and not to an H |
| 12 | C5 (Purines) | C-sp$^2$ bonded to a C-sp$^2$ and to an N-sp$^2$ |
| 13 | C4 (Purines) | C-sp$^2$ bonded to a C-sp$^2$, to an N-sp$^2$ and to a pyrrolic N-sp$^2$ |
| 14 | C2 (ADE) | C-sp$^2$ bonded to two N-sp$^2$ |
| 15 | C6 (ADE); C4 (CYT) | C-sp$^2$ bonded to an N-sp$^3$, to an N-sp$^2$ and to a C-sp$^2$ |
| 16 | N6 (ADE); N4 (CYT); N2 (GUA) | N-sp$^3$ bonded to C-sp$^2$ |
| 17 | C2 (GUA) | C-sp$^2$ bonded to two N-sp$^2$ and to an N-sp$^3$ |
| 18 | C6 (GUA); C4 (URI) | C-sp$^2$ bonded to C-sp2, N-sp$^2$ and O-sp$^2$ |
| 19 | O2 (Pyrimidines); O6 (GUA); O4 (URI) | O-sp$^2$ bonded to a C-sp$^2$ |
| 20 | C2 (Pyrimidines) | C-sp$^2$ bonded to a pyrrolic N-sp$^2$, to an N-sp$^2$ and to an O-sp$^2$ |
| 21 | C5 (Pyrimidines) | C-sp$^2$ bonded to two C-sp$^2$ |
| 22 | C6 (Pyrimidines) | C-sp$^2$ bonded to a C-sp$^2$ and to a pyrrolic N-sp$^2$ |
| 23 | N1 (GUA); N3 (URI) | N-sp$^2$ bonded to two C-sp$^2$ and to an H |

+Atom names in IUPAC format as found in PDB files were used. Atom types defined are graphically illustrated in Sup. Fig. 2.

**Supplementary Table 3.** Atom type definition in RASP potentials (C3, BB and BBR)

| ATOM TYPES DEFINITION | | | DESCRIPTION OF ATOMS AND NUCLEOTIDES |
|---|---|---|---|
| C3 | BB | BBR | |
| --- | 1 | 1 | OP1, OP2, OP3 (ADE) |
| --- | 2 | 2 | P (ADE) |
| --- | 3 | 3 | O5’ (ADE) |
| --- | 4 | 4 | C5’ (ADE) |
| --- | 5 | 5 | C4’ (ADE) |
| 1 | 6 | 6 | C3’ (ADE) |
| --- | 7 | 7 | O3’ (ADE) |
| --- | --- | 8 | O2’ (ADE) |
| --- | --- | 9 | C2’ (ADE) |
| --- | --- | 10 | C1’ (ADE) |
| --- | --- | 11 | O4’ (ADE) |
| --- | 8 | 12 | OP1, OP2, OP3 (CYT) |
| --- | 9 | 13 | P (CYT) |
| --- | 10 | 14 | O5’ (CYT) |
| --- | 11 | 15 | C5’ (CYT) |
| --- | 12 | 16 | C4’ (CYT) |
| 2 | 13 | 17 | C3’ (CYT) |
| | 14 | 18 | O3’ (CYT) |
| --- | --- | 19 | O2’ (CYT) |
| --- | --- | 20 | C2’ (CYT) |
| --- | --- | 21 | C1’ (CYT) |
| --- | --- | 22 | O4’ (CYT) |
| --- | 15 | 23 | OP1, OP2, OP3 (GUA) |
| --- | 16 | 24 | P (GUA) |
| --- | 17 | 25 | O5’ (GUA) |
| --- | 18 | 26 | C5’ (GUA) |
| --- | 19 | 27 | C4’ (GUA) |
| 3 | 20 | 28 | C3’ (GUA) |
| --- | 21 | 29 | O3’ (GUA) |
| --- | --- | 30 | O2’ (GUA) |
| --- | --- | 31 | C2’ (GUA) |
| --- | --- | 32 | C1’ (GUA) |
| --- | --- | 33 | O4’ (GUA) |
| --- | 22 | 34 | OP1, OP2, OP3 (URI) |
| --- | 23 | 35 | P (URI) |
| --- | 24 | 36 | O5’ (URI) |
| --- | 25 | 37 | C5’ (URI) |
| --- | 26 | 38 | C4’ (URI) |
| 4 | 27 | 39 | C3’ (URI) |
| --- | 28 | 40 | O3’ (URI) |
| --- | --- | 41 | O2’ (URI) |
| --- | --- | 42 | C2’ (URI) |
| --- | --- | 43 | C1’ (URI) |
| --- | --- | 44 | O4’ (URI) |

**Supplementary Table 4.** Ranking test

| | *randseq* | | *randstr* | |
|---|---|---|---|---|
| | Top 1 | Top 10 | Top 1 | Top 10 |
| C3 | 0.52 | 0.72 | 0.08 | 0.32 |
| BB | 0.54 | 0.71 | 0.35 | 0.78 |
| BBR | 0.62 | 0.80 | 0.89 | 0.93 |
| FULL | --- | --- | 0.93 | 0.95 |
| NAST | --- | --- | 0.22 | 0.65 |
| ROSETTA | --- | --- | 0.62 | 0.75 |
| ROSETTAmin | --- | --- | 0.85 | 1.00 |
| AMBER99 | --- | --- | 0.73 | 0.88 |

Fraction of native RNA structures correctly ranked by each scoring function in the randseq and randstr datasets. Top 1 ranking represents the fraction of cases where the native RNA structure had the lowest score. Top 10 ranking represents the fraction of cases where the native structure had a score within the lowest ten scores out of the 500 decoys.

**Supplementary Table 5.** Statistical significance analysis of correlation tests between energy scores and RMSD

| | C3 | BB | BBR | ALL | NAST | R | Rmin | AMB |
|---|---|---|---|---|---|---|---|---|
| C3 | N.A. | 0.694 | 0.753 | 0.859 | 0.329 | 0.287 | 0.301 | 0.153 |
| BB | $<10^{-3}$ | N.A. | 0.141 | 0.412 | 0.729 | 0.800 | 0.679 | 0.729 |
| BBR | $<10^{-3}$ | **0.340** | N.A. | 0.341 | 0.800 | 0.882 | 0.751 | 0.800 |
| FULL | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | N.A. | 0.894 | 0.929 | 0.869 | 0.882 |
| NAST | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | N.A. | 0.120 | 0.505 | 0.235 |
| R | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | **0.550** | N.A. | 0.506 | 0.244 |
| Rmin | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | N.A. | 0.305 |
| AMB | **0.251** | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | 0.015 | 0.011 | 0.001 | N.A. |

The statistical significance of the observed differences between two distributions of Pearson correlation coefficients using RASP (C3, BB, BBR and ALL), NAST, ROSETTA (R), ROSETTAmin (with energy minimization, labeled as Rmin) and AMBER (labeled as AMB) energy scores and the RMSD of all atoms in the randstr dataset was evaluated with the non-parametric and distribution free Kolmogorov-Smirnov (KS) test. The cells below the diagonal of the table show the p-value for the comparison of two potentials using the KS test. The cases where the difference is not statistically significant, at the confidence level of 95%, are shown in bold. The cells above the diagonal of the table show the D statistic for all pairs of potentials. N.A. stands for "Not Applicable".

**Supplementary Table 6.** Statistical significance analysis of correlation tests between energy scores and GDT-TS

| | C3 | BB | BBR | ALL | NAST | R | Rmin | AMB |
|---|---|---|---|---|---|---|---|---|
| C3 | N.A. | 0.788 | 0.835 | 0.894 | 0.400 | 0.252 | 0.363 | 0.153 |
| BB | $<10^{-3}$ | N.A. | 0.235 | 0.518 | 0.777 | 0.835 | 0.750 | 0.788 |
| BBR | $<10^{-3}$ | 0.015 | N.A. | 0.426 | 0.859 | 0.905 | 0.797 | 0.835 |
| FULL | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | N.A. | 0.918 | 0.941 | 0.882 | 0.906 |
| NAST | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | N.A. | 0.183 | 0.622 | 0.294 |
| R | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | **0.105** | N.A. | 0.558 | 0.244 |
| Rmin | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | N.A. | 0.375 |
| AMB | **0.251** | $<10^{-3}$ | $<10^{-3}$ | $<10^{-3}$ | 0.001 | 0.010 | $<10^{-3}$ | N.A. |

The statistical significance of the observed differences between any two distributions of Pearson correlation coefficients using RASP (C3, BB, BBR and ALL), NAST, ROSETTA (R), ROSETTAmin (with energy minimization, labeled as Rmin) and AMBER (labeled as AMB) energy scores and the GDT-TS of all atoms in the randstr dataset was evaluated with the non-parametric and distribution free KS test. See legend of Sup. Table 5 for more details.

**Supplementary Table 7.** Model ranking results in ROSETTA benchmark set

| **Seq Num** | **MOTIF NAME** | **Number of Residues** | **Number of Chains** | **RMSD of Lowest Energy Model†** | | **RMSD of nearest-native Model** |
|---|---|---|---|---|---|---|
| | | | | **ROSETTAmin** | **RASP-ALL** | |
| **1** | G-A base pair | 6 | 2 | 3.106 | **1.190** | **1.190** |
| **2** | Fragment with G/G and G/A pairs, SRP helix VI | 8 | 2 | 2.620 | **1.832** | **1.832** |
| **3** | Helix with A/C base pairs | 12 | 2 | **1.814** | 2.451 | **1.814** |
| **4** | Four-way junction, HCV IRES | 13 | 4 | 11.351 | 11.351 | 10.031 |
| **5** | Loop 8, A-type Ribonuclease P | 7 | 1 | 4.501 | **3.440** | 1.384 |
| **6** | Helix with U/C base pairs | 8 | 2 | 3.006 | **2.095** | **2.095** |
| **7** | Curved helix with G/A and A/A base pairs | 12 | 2 | **1.062** | 1.743 | 0.998 |
| **8** | Pre-catalytic conformation, hammerhead ribozyme | 19 | 3 | 12.287 | **7.659** | **7.659** |
| **9** | Loop E motif, 5S RNA | 18 | 2 | **2.163** | 2.269 | 1.641 |
| **10** | UUCG tetraloop | 6 | 1 | 1.148 | **1.143** | 1.122 |
| **11** | Rev response element high affinity site | 9 | 2 | 4.112 | **4.078** | 3.859 |
| **12** | Fragment with A/C pairs, SRP helix VI | 12 | 2 | 5.456 | **3.270** | **3.270** |
| **13** | Signal recognition particle Domain IV | 12 | 2 | 3.227 | **2.346** | 1.217 |
| **14** | Bulged G motif, sarcin/ricin loop | 13 | 2 | **1.659** | 5.160 | 1.460 |
| **15** | Tertiary interaction, hammerhead ribozyme | 16 | 3 | **7.942** | 9.863 | 7.565 |
| **16** | GAGA tetraloop from sarcin/ricin loop | 6 | 1 | 0.921 | **0.819** | 0.819 |
| **17** | Pentaloop from conserved region of SARS genome | 7 | 1 | 3.110 | **3.109** | 0.999 |
| **18** | L2/L3 tertiary interaction, purine riboswitch | 18 | 2 | **9.461** | 9.590 | 8.081 |
| **19** | L3, thiamine pyrophosphate riboswitch | 7 | 1 | 3.750 | **1.995** | **1.995** |
| **20** | Kink-turn motif from SAM-I riboswitch | 13 | 2 | **1.365** | 8.447 | 1.220 |
| **21** | Active site, hammerhead ribozyme | 17 | 3 | **10.210** | 11.721 | 8.643 |
| **22** | P1/L3, SAM-II riboswitch | 23 | 2 | **9.762** | 12.289 | 7.397 |
| **23** | J4/5 from P4-P6 domain, Tetrahymena ribozyme | 9 | 2 | **2.125** | 2.352 | 1.759 |
| **24** | Stem C internal loop, L1 ligase | 12 | 2 | 2.416 | 2.416 | 2.240 |
| **25** | J5/5a hinge, P4-P6 domain, Tetr. ribozyme | 17 | 2 | **10.897** | 10.973 | 9.941 |
| **26** | Three-way junction, purine riboswitch | 13 | 3 | **6.578** | 7.126 | 6.099 |
| **27** | J4a/4b region, metal-sensing riboswitch | 14 | 2 | 4.479 | **3.523** | 3.428 |
| **28** | Kink-turn motif | 15 | 2 | 10.049 | **9.724** | 8.580 |
| **29** | Tetraloop/helix interaction, L1 ligase crystal | 10 | 3 | 1.214 | **0.857** | **0.857** |
| **30** | Hook-turn motif | 11 | 3 | **4.569** | 5.495 | 1.718 |
| **31** | Tetraloop/receptor, P4-P6 domain, Tetr. ribozyme | 15 | 3 | 8.232 | **6.770** | 2.891 |
| **32** | Pseudoknot, domain III, CPV IRES | 18 | 2 | 5.835 | **3.808** | 3.380 |

Benchmark set of 407 structure models built with ROSETTA and the FARFAR potential for 32 RNA motifs with non-canonical base pairs (Das et al., 2010). † The RMSD of the lowest energy model of ROSETTAmin and RASP-ALL is shown in bold face when it is the lowest value between the two methods. The RMSD value of the nearest-native model in each set is shown in bold type when either ROSETTAmin or RASP-ALL potentials were able to select it with the lowest energy score.