Blind tests of RNA nearest-neighbor energy prediction

Fang-Chieh Chou[a], Wipapat Kladwang[a], Kalli Kappel[b], and Rhiju Das[a,b,c,1]

[a]Department of Biochemistry, Stanford University, Stanford, CA 94305; [b]Biophysics Program, Stanford University, Stanford, CA 94305; and [c]Department of Physics, Stanford University, Stanford, CA 94305

Edited by David Baker, University of Washington, Seattle, WA, and approved May 31, 2016 (received for review November 25, 2015)

The predictive modeling and design of biologically active RNA molecules requires understanding the energetic balance among their basic components. Rapid developments in computer simulation promise increasingly accurate recovery of RNA's nearest-neighbor (NN) free-energy parameters, but these methods have not been tested in predictive trials or on nonstandard nucleotides. Here, we present, to our knowledge, the first such tests through a RECCES–Rosetta (reweighting of energy-function collection with conformational ensemble sampling in Rosetta) framework that rigorously models conformational entropy, predicts previously unmeasured NN parameters, and estimates these values' systematic uncertainties. RECCES–Rosetta recovers the 10 NN parameters for Watson–Crick stacked base pairs and 32 single-nucleotide dangling-end parameters with unprecedented accuracies: rmsd of 0.28 kcal/mol and 0.41 kcal/mol, respectively. For set-aside test sets, RECCES–Rosetta gives rmsd values of 0.32 kcal/mol on eight stacked pairs involving G–U wobble pairs and 0.99 kcal/mol on seven stacked pairs involving nonstandard isocytidine–isoguanosine pairs. To more rigorously assess RECCES–Rosetta, we carried out four blind predictions for stacked pairs involving 2,6-diaminopurine–U pairs, which achieved 0.64 kcal/mol rmsd accuracy when tested by subsequent experiments. Overall, these results establish that computational methods can now blindly predict energetics of basic RNA motifs, including chemically modified variants, with consistently better than 1 kcal/mol accuracy. Systematic tests indicate that resolving the remaining discrepancies will require energy function improvements beyond simply reweighting component terms, and we propose further blind trials to test such efforts.

RNA helix | ensemble prediction | simulated tempering | thermodynamics | blind prediction

RNA plays central roles in biological processes, including translation, splicing, regulation of genetic expression, and catalysis (1, 2), and in bioengineering efforts to control these processes (3–5). These critical RNA functions are defined at their most fundamental level by the energetics of how RNA folds and interacts with other RNAs and molecular partners, and how these processes change upon naturally occurring or artificially introduced chemical modifications. Experimentally, the folding free energies of RNA motifs can be precisely measured by optical melting experiments, and a compendium of these measurements have established the nearest-neighbor (NN) model for the most basic RNA elements, including double helices with the four canonical ribonucleotides (6). In the NN model, the stability of a base pair is assumed to only be affected by its adjacent base pairs, and the folding free energy of a canonical RNA helix can be estimated based on NN parameters for each stacked pair, an initialization term for the entropic cost of creating the first base pair, and corrections for different terminal base pairs. Although next-NN effects and tertiary contacts are not treated in the NN model (7–9), the current NN model gives accurate predictions for the folding free energies of canonical RNA helices (<0.5 kcal/mol for helices with 6–8 base pairs) (10, 11) and can be extended to single-nucleotide dangling ends, chemically modified nucleotides, and more complex motifs, such as noncanonical base pairs, hairpins, and internal loops (11–14). However, it is currently not feasible to experimentally characterize the energetics of all RNA motifs due to the large number of possible motif sequences and the requirement of specialized experiments to address complex motif topologies, such as three-way junctions (15–17). These considerations, and the desire to test physical models of RNA folding, have motivated several groups to pursue automated computational methods to calculate the folding free energies of RNA motifs.

Current computational approaches are beginning to recover NN parameters for the simplest RNA motifs with accuracies within a few-fold of the errors of experimental approaches. For example, the Rosetta package has been developed and extensively tested for structure prediction and design of macromolecules, including RNA. Recent successes at near-atomic resolution have leveraged an all-atom "score function" that includes physics-based terms (for hydrogen bonding, van der Waals packing, and orientation-dependent implicit solvation) and knowledge-based terms (for, e.g., RNA torsional preferences) (18). When interpreting the total score as an effective energy for a conformation, simple Rosetta calculations recover the NN parameters for all canonical stacked pairs with an rmsd of less than 0.5 kcal/mol upon fitting two phenomenological parameters, the Rosetta energy scale and a constant offset parameterizing the conformational entropy loss upon folding each base pair (ref. 18 and see below). In parallel, molecular dynamics studies have demonstrated calculation of folding free energies of short RNA hairpins using umbrella sampling, molecular mechanics–Poisson Boltzmann surface area (MM–PB/SA), free energy perturbation, and other methods (19–22). Although these calculations have not yet accurately recovered folding free energies (errors > 10 kcal/mol) (21, 22), relative differences of NN parameters between different sequences and other aspects of RNA motif energetics have been recovered with accuracies between 0.6–1.8 kcal/mol (22–24). These error ranges are similar or lower than uncertainties of

Significance

Understanding RNA machines and how their behavior can be modulated by chemical modification is increasingly recognized as an important biological and bioengineering problem, with continuing discoveries of riboswitches, mRNA regulons, CRISPR-guided editing complexes, and RNA enzymes. Computational strategies for understanding RNA energetics are being proposed, but have not yet faced rigorous tests. We describe a modeling strategy called RECCES–Rosetta (reweighting of energy-function collection with conformational ensemble sampling in Rosetta) that models the full ensemble of motions of RNA in single-stranded form and in helices, including nonstandard nucleotides, such as 2,6-diaminopurine, a variant of adenosine. When compared with experiments, including blind tests, the energetic accuracies of RECCES–Rosetta calculations are at levels close to experimental error, suggesting that computation can now be used to predict and design basic RNA energetics.

Author contributions: F.-C.C., W.K., K.K., and R.D. designed research; F.-C.C., W.K., and K.K. performed research; F.-C.C. contributed new reagents/analytic tools; F.-C.C. and K.K. analyzed data; and F.-C.C. and R.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: rhiju@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1523335113/-/DCSupplemental.

empirically defined NN parameters for most motifs, which are on the scale of 1 kcal/mol. For example, original NN energy estimates for G–U stacked pairs, single-nucleotide bulges, and tetraloop free energies have been corrected by >1 kcal/mol when revisited in detailed studies (11, 25–27). Overall, computational approaches may be ready for calculations of new energetic parameters, including parameters for these uncertain motifs as well as for motifs involving nonstandard nucleotides that are being found throughout natural coding and noncoding RNAs (28, 29) or used to engineer new RNA systems (30, 31). However, the predictive power of these methods has not been evaluated through tests on previously unmeasured NN parameters. Predictive tests are particularly important because models are increasing in complexity and risk overtraining on previously available data.

Here we report, to our knowledge, the first blind tests of a method to computationally predict NN energetic parameters. The newly measured parameters involve RNA stacked pairs with the nonnatural nucleotide 2,6-diaminopurine (D) paired to uracil (Fig. 1). To ensure a rigorous comparison, calculations were carried out by one author (F.-C.C.) and subsequently tested in independent experiments by another author (W.K.). In preparation for this blind test, we developed a reweighting of energy-function collection with conformational ensemble sampling in Rosetta (RECCES–Rosetta) framework to calculate free energies based on density-of-states estimation and expected errors from statistical precision, inaccuracies in the NN assumption, and uncertainties in the weights of the underlying energy function. Furthermore, to address previous ad hoc assumptions used to fit conformational entropy from data, RECCES calculates the conformational entropy of helix and single-stranded states without fitting of additional parameters. These systematic improvements—and calibration based on previously measured NN parameters—ensured that our blind tests carried sufficient power to rigorously establish the accuracy and limitations of NN energy calculations that seek to make nontrivial predictions.

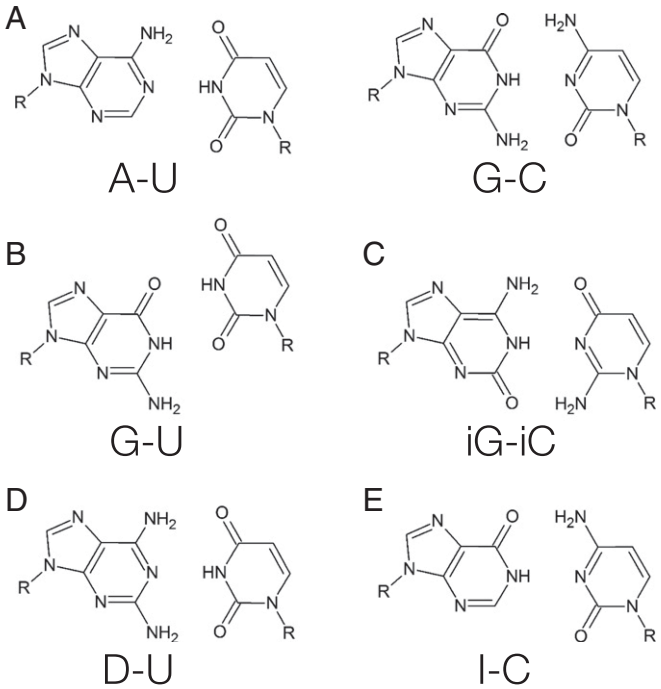


Fig. 1. Base pairs involved in NN parameters considered in this study. (*A*) Canonical pairs adenosine–uracil and guanosine–cytidine, (*B*) guanosine–uracil wobble pair, (*C*) nonnatural isoguanosine–isocytidine, (*D*) nonnatural 2,6-diaminopurine–uracil, and (*E*) inosine–cytidine.

Results

Recovery of Canonical Helix and Dangling-End Parameters. Blind tests of a prediction method are not worthwhile if the expected prediction errors significantly exceed the range of possible experimental values—on the order of several kilocalories per mole for NN parameters. We therefore first sought to determine whether folding free-energy calculations with the Rosetta all-atom energy function, previously developed for RNA structure prediction and design, could recover NN energetics for canonical Watson–Crick stacked pairs and whether these calculations' uncertainties were acceptable for making blind predictions. The Rosetta energy function involves separate component terms for hydrogen bonding, electrostatics, van der Waals interactions, nucleobase stacking, torsional potentials, and an orientation-dependent solvation model. Prior structure prediction and design studies did not strongly constrain the weights of these components (18). Thus, we anticipated that NN parameter prediction would require optimization of the weights and care in uncertainty estimation. To assess whether the errors due to weight uncertainties would allow nontrivial predictions, we sought not just a single weight set but instead a large collection of weight sets consistent with available data.

To discover these weight sets, we developed the RECCES framework for sampling conformational ensembles of the single-stranded and helix conformations relevant to NN energy estimation (Fig. 2 and *SI Appendix*, Table S1). Through the use of a density-of-states formalism, simulated tempering, and weighted histogram analysis method (WHAM) integration, RECCES allowed the estimation of free energies with bootstrapped errors of less than 0.003 kcal/mol, significantly less than systematic errors of 0.3 kcal/mol (estimated below; *SI Appendix*, Tables S2–S4), using two central processing unit (CPU) hours of computation per molecule. These methods are similar to replica exchange methods in common use in molecular dynamics studies, but are simpler in that they do not require running multiple parallel processes (*SI Appendix, Supporting Methods*). Importantly, the overall RECCES framework did not require separate fitting of conformational entropy factors, reducing the likelihood of overfitting. Furthermore, starting from these initial simulations, RECCES enabled evaluation of alternative weight sets with negligible additional computation (<0.1 s) through a rapid reweighting of cached energies. Though noisy at low energies (compare green to blue curves in Fig. 2*C*), we confirmed that this reweighting procedure nevertheless led to an acceptable mean calculation error of 0.28 kcal/mol (*SI Appendix*, Table S4), significantly smaller than the several kilocalories per mole range of experimental NN parameters (*SI Appendix*, Table S1). Further tests of the NN assumption, based on simulations with different helix contexts for each stacked pair, also gave systematic errors of 0.2–0.3 kcal/mol (*SI Appendix*, Table S2). Hereafter, we conservatively describe the systematic errors of the RECCES–Rosetta NN parameter estimates to be the higher value in this range, 0.3 kcal/mol.

To obtain a collection of weight sets, we used RECCES to optimize the weights of all terms in the Rosetta score function over numerous runs with different initial values. These optimization runs minimized the mean square error with respect to the NN parameters of 10 canonical stacked pairs (four base pairs next to four base pairs, removing symmetric cases), 32 single-nucleotide dangling ends (four nucleotides at either the 5′ or 3′ end of four base pairs), and the terminal penalty for A–U vs. G–C. The resulting 9,544 minimized weight sets were highly diverse, even after discarding the weight sets with 5% worst rmsd agreement to training data (*SI Appendix*, Table S5, describes score terms and summarizes mean and SDs of weights; *SI Appendix*, Table S6, gives five example weight sets). Most score terms were recovered with mean weights greater than zero by more than one SD, confirming their importance for explaining RNA structure and energetics. These terms included *stack_elec*, which models the

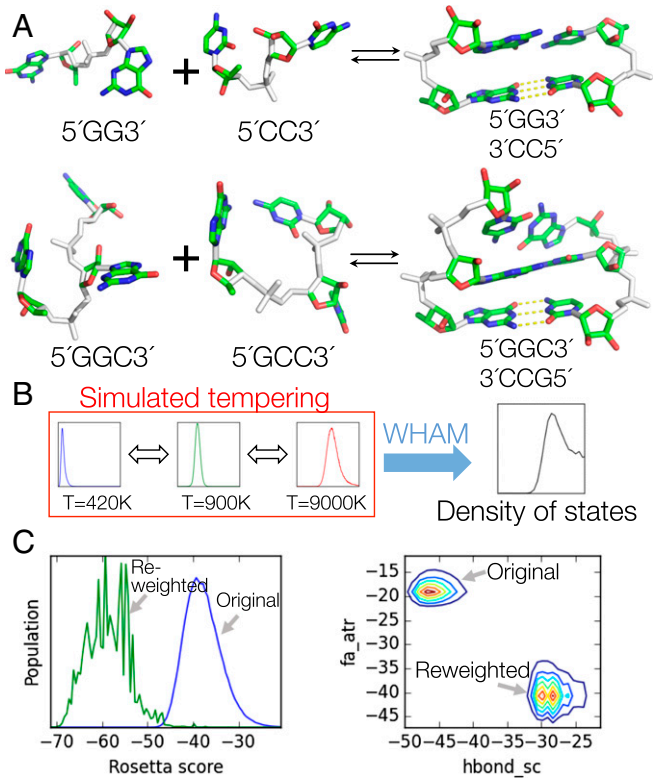


Fig. 2. RECCES thermodynamic framework and reweighting. (*A*) Example systems simulated for this study. Degrees of freedom sampled are colored in white. The relative orientation of first base pair in each helix was fixed (*Right*, yellow dashes). (*Upper* and *Lower*) Folding reactions of two-base-pair and three-base-pair systems, respectively. (*B*) Density of state estimation by simulated tempering and WHAM. (*C*) Reweighting demonstration. (*Left*) State population at room temperature before (blue) and after (green) reweighting. (*Right*) Two-dimensional population histograms of *fa_atr* (Lennard–Jones attraction) vs. *hbond_sc* (hydrogen bonds) energy components, before and after reweighting.

electrostatic interaction between stacked nucleobases, an effect previously posited by several groups to be important for understanding fine-scale RNA energetics (14, 32). Terms with wider variance across weight sets could be explained through their covariance with other terms. For example, some pairs of score terms, such as the nucleobase stacking term *fa_stack* and the van der Waals term *fa_atr*, model similar physical effects, but other pairs model opposing effects in helix association, such as hydrogen bonding *hbond_sc* and the solvation term for burying polar moieties *geom_sol_fast* (*SI Appendix*, Table S7). The weights of these pairs varied significantly across optimized weight sets, but linear combinations of these weight pairs were nearly invariant across the weight set collection (*SI Appendix*, Fig. S1).

Despite the variations and covariations observed across this large collection of weight sets, each weight set gave an rmsd accuracy of better than 0.58 kcal/mol for canonical base pairs and dangling ends, with a mean accuracy of 0.40 kcal/mol across all training data. These accuracies were significantly better than rmsds of 1.51 kcal/mol and 1.23 kcal/mol, respectively, obtained with the original structure prediction weights, supporting the need for reweighting (*SI Appendix*, Table S6). The rmsd over just the canonical stacked base pairs was 0.28 kcal/mol (Fig. 3*A*), comparable in accuracy to the initial experimental estimates of these values (10, 12) and consistent with the estimated systematic errors of our calculation strategies (0.3 kcal/mol) (*SI Appendix*, Tables S2 and S4). For the dangling-end data, RECCES–Rosetta also gave an excellent rmsd of 0.41 kcal/mol (Fig. 3*B*). For these data, the largest deviations from experiment were tagged as having the highest expected error from weight uncertainties by RECCES, supporting this method of error computation (see, e.g., $^{5'CG}_{3'G}$ dangling end in *SI Appendix*, Table S1). For both sets of NN parameters, the rmsd errors were significantly smaller than the range of experimental values (2.5 kcal/mol and 1.5 kcal/mol for canonical stacked pairs and dangling ends, respectively), leading to the visually clear correlations in Fig. 3 *A* and *B*. The terminal penalty for A–U relative to G–C was also recovered with a similar error (0.3 kcal/mol) (*SI Appendix* includes further discussion and computation of other terminal base pair contributions).

Because we directly trained the RECCES score function against the experimental dataset, the accuracies of these results were expected. Nevertheless, we gained further confidence in the use of Rosetta-derived energy functions and RECCES framework by comparing its performance to the results of two simpler models trained on the same data. First, a three-parameter hydrogen-bond counting model, similar to simple phenomenological models that inspired the NN parametrization (10) (*SI Appendix*, *Supporting Methods*), achieved rmsd accuracies of 0.29 kcal/mol and 0.45 kcal/mol on canonical stacked pairs and dangling ends, respectively—slightly worse than the RECCES results (0.28 kcal/mol and 0.41 kcal/mol, respectively), despite including fitted parameters that account for conformational entropy loss of base pairs and dangling ends. Second, a prior single-conformation Rosetta method, which uses the same energy function as RECCES–Rosetta but evaluates the score only for a minimized helix conformation (18) achieved accuracies of 0.30 kcal/mol and 0.44 kcal/mol for canonical stacked pairs and dangling ends, respectively—again worse than the RECCES–Rosetta results despite including separately fitted conformational entropy terms. For all three models, the largest deviation was for the stacked pair $^{5'CG}_{3'GC}$, which is less stable than the other stacked pairs with two G–C pairs by 1 kcal/mol; still, even for this parameter, the RECCES–Rosetta calculations were more accurate than the simpler models. These comparisons supported the utility of

Table 1. Accuracies of nearest-neighbor parameter predictions

		Rmsd accuracy (kcal/mol)			
RNA motif category	No. motifs	Hydrogen-bond counting	Single-conformation Rosetta	RECCES–Rosetta	RECCES–Rosetta refitted*
Canonical†	10	0.29	0.30	0.28	0.41
Dangling†	32	0.45	0.44	0.41	0.43
G–U‡	8	0.59	0.49	0.32	0.32
iG–iC‡	7	0.79	0.85	0.99	1.08
D-U§	4	0.48	0.40	0.63	0.46
All	61	0.50	0.49	0.50	0.53

*The model was trained with all data available, so all entries in the column are training data.
†Data used in training the models.
‡Data set aside for testing.
§Blind test data.

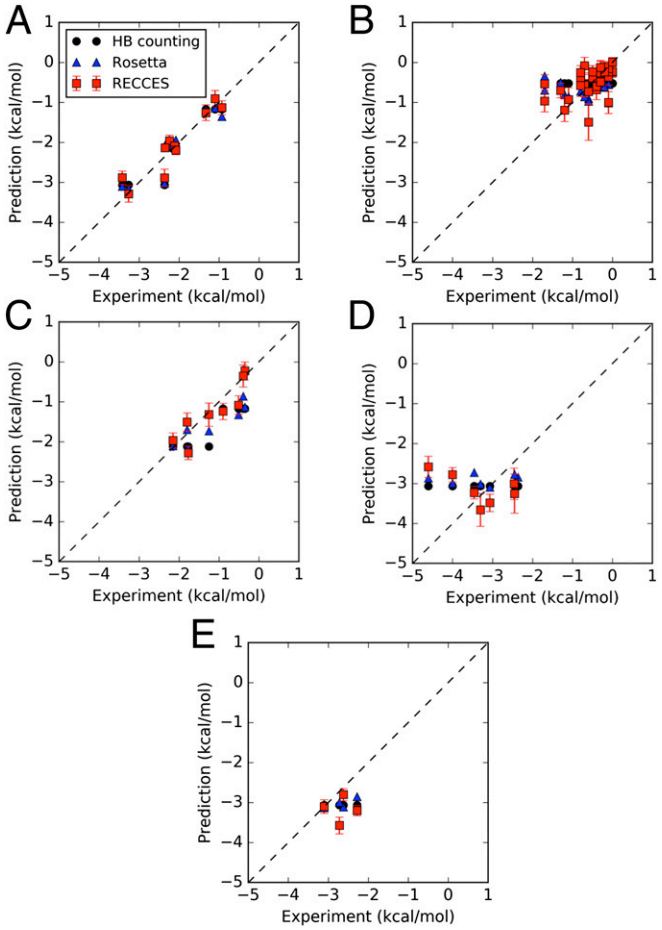


Fig. 3. Calculations vs. experiment for each NN parameter set. (*A*) Canonical stacked pairs; (*B*) single-nucleotide dangling ends; (*C*) stacked pairs including one G–U pair; (*D*) stacked pairs including at least one iG–iC pair; (*E*) stacked pairs including one D–U pair. All panels are drawn with the same axis limits and a line of equality (dashed) to aid cross-panel visual comparison.

the RECCES–Rosetta method compared with less generalizable models. However, the significance in the accuracy improvement was difficult to rigorously evaluate because the models contained different numbers and types of parameters; we therefore turned to independent test sets and blind predictions.

Tests on Independent Nearest-Neighbor Parameter Measurements. Recent comprehensive experimental measurements have updated the NN parameters for stacked pairs involving G–U wobble pairs next to canonical Watson–Crick pairs (11). Because these values were not used in the training of the models herein and because the geometry of G–U wobble pair is distinct from G–C and A–U pairs (Fig. 1*B*), this set of measurements offered strong tests of modeling accuracy. Furthermore, the expected error in the RECCES–Rosetta calculations from weight uncertainties, based on variation across the large collection of weight sets, was 0.22 kcal/mol (*SI Appendix*, Table S1), less than the estimated ∼0.3 kcal/mol systematic error (*SI Appendix*, Tables S2 and S4). Both error contributions were significantly less than the full range of predicted NN parameters (2.1 kcal/mol), supporting the strength of this test. The actual rmsd accuracy across these G–U NN measurements was 0.32 kcal/mol for RECCES–Rosetta (Table 1), nearly as accurate as the recovery of training set stacked pairs (0.28 kcal/mol) and comparable to expected systematic errors. Furthermore, this accuracy over G–U-containing stacked pairs outperformed the rmsd values calculated from hydrogen-bond counting and single-conformation Rosetta scoring methods (0.59 and 0.49 kcal/mol, respectively) by 50–80%, supporting the importance of carrying out detailed physical simulations of the conformational ensemble via RECCES over simpler approaches. Here and below, the predictions and their estimated errors were calculated by computing means and SDs of NN parameters across the full collection of weight sets discovered by RECCES. Compared with this averaging over multiple models, using the single weight set with best fit to the training data gave slightly worse accuracies on the test data (*SI Appendix*, Table S6) (33).

A more difficult test involved seven previously measured NN parameters of a nonnatural base pair, iG–iC (Fig. 1*C*) (34). The rmsd for the iG–iC test case was 0.99 kcal/mol, mainly due to two significant outliers: $^{5'iGiC}_{3'iCiG}$ and $^{5'GiC}_{3'CiG}$ (Fig. 3*D*). The predicted NN parameters for these outliers were larger than experimental values (less stable) by 2.2 and 1.3 kcal/mol, respectively. Nevertheless, over the other five iG–iC NN parameters, the rmsd was 0.51 kcal/mol, and the discrepancies appeared primarily due to a systematic offset in the predictions (Fig. 3*D*). The accuracy was comparable to the maximum errors expected from weight uncertainties (0.4–0.5 kcal/mol) and similar, in terms of relative accuracies, to the canonical and G–U-containing stacked pairs above. Compared with RECCES–Rosetta, the simpler hydrogen-bond counting and single-conformation Rosetta scoring models gave 15–20% better accuracies (0.79 and 0.85 kcal/mol, respectively; 0.47 and 0.44 kcal/mol, excluding outliers); but both simple models gave near-constant NN parameters (range less than 0.3 kcal/mol) over all stacked pairs, providing no explanation for the 2.2 kcal/mol range in experimental measurements or for the outliers (Fig. 3*D*). On one hand, the two outliers suggest that some important physical effect is missing or incorrectly implemented in the current calculation procedure (see *Discussion*). On the other hand, the excellent accuracies over the other iC–iG-containing stacked pairs, along with the performance in the G–U test set, motivated us to continue with blind comparisons.

Blind Tests Involving Diaminopurine–Uracil Base Pairs. As a blind test, we applied RECCES–Rosetta to predict the NN parameters for stacked pairs involving a distinct nonnatural base pair, 2,6-diaminopurine paired with uracil (D–U) (Fig. 1*D*). Predictions of these parameters (*SI Appendix*, Table S1) suggested a wide range of NN values and confirmed that errors from weight uncertainties were smaller or comparable to other systematic sources of error (0.3 kcal/mol). To test these predictions, we measured NN parameters for the four stacked pairs involving D–U next to G–C pairs, which were expected to have a range of 0.8 kcal/mol. *SI Appendix*, Table S8, gives construct sequences and experimental folding free-energy values for these constructs, and Table 1 and *SI Appendix*, Tables S1 and S9, summarize the NN parameter estimation. The rmsd of the RECCES–Rosetta blind predictions was 0.63 kcal/mol (Fig. 3*E*). The hydrogen-bond counting and single-conformation Rosetta scoring models, which fared worse than RECCES–Rosetta in most tests above, gave rmsds of 0.48 and 0.40 kcal/mol, respectively, better than RECCES–Rosetta by 24–37% (Table 1). This result is similar to what we observed in the iG–iC test case; indeed, the two simple models again produced near constant predictions (range < 0.2 kcal/mol) for the D–U stacked pairs that did not account for the 0.8 kcal/mol range of the measured values (Fig. 3*E*). Given the blind nature of the test and our attempts to ensure its power to falsify our calculations, this test unambiguously indicated that some physical term is missing in the current Rosetta all-atom energetic model (as well as simpler models). Nevertheless, the results are encouraging: the blind predictions from each of the three models over each of four NN values separately achieved better than 1 kcal/mol accuracy compared with subsequent experimental measurements.

Post Hoc Fit Across All Data. Though post hoc tests of models on prior collected data are less rigorous than blind trials, they can

help guide future work. As a final test, we wished to understand possible explanations for the worse accuracy of RECCES–Rosetta in iG–iC test cases and blind D–U trials compared with the G–U test cases. One model for this inaccuracy was that overfitting of energy function weights to the training data worsened predictive power over the new data. Another (not necessarily exclusive) model was that the underlying energy function derived from Rosetta score terms was fundamentally incapable of modeling the available NN data under any weight set with the RECCES procedure. We were able to test these models by carrying out a post hoc global fit of energy function weights over all available NN data (Fig. 4 and Table 1). As expected, we observed better fits to the test data, including an improvement in rmsd accuracy for the four D–U stacked pairs from 0.63 kcal/mol to 0.46 kcal/mol; this result suggests a modest overfitting to the training set in the studies above. However, we observed somewhat worse fits to the training data, including a worsening of rmsd accuracy for the 10 canonical stacked pairs from 0.28 to 0.41 kcal/mol, worse than expected systematic errors in our calculations (0.3 kcal/mol) (*SI Appendix*, Tables S2 and S4) supporting the second model of fundamental energy function inaccuracy. Furthermore, this global fit still failed to account for the two striking outliers involving iG–iC base pairs, again giving evidence for the second model: energetic calculations based on the current Rosetta score function are fundamentally incapable of accounting for all of the data within expected error, even with a post hoc optimized weight set.

Discussion

This study reports, to our knowledge, the first blind test of the predictive power of high-resolution, all-atom modeling methods for RNA folding energetics. We developed a RECCES strategy in the Rosetta framework that rigorously models conformational ensembles of single strand and helical states, is computationally efficient (hours with currently available CPUs), and brackets systematic errors based on comprehensive reweighting tests. Compared with simpler phenomenological methods, RECCES–Rosetta achieved excellent rmsd accuracies for the NN parameters of canonical base pairs, dangling ends, and G–U pairs, but somewhat worse results for NN parameters involving nonnatural base pairs iG–iC and D–U. The latter D–U parameters were measured after the predictions as a blind test. The computational accuracies were better than 1 kcal/mol in all cases, based on rmsd values over each separate set of NN parameters (0.28, 0.41, 0.32, 0.99, and 0.63 kcal/mol for canonical, dangling end, G–U, iG–iC, and D–U parameters, respectively) and also individually for each of the four blind predictions. These rmsd values are significantly smaller than the 2–3 kcal/mol ranges measured for these sets of NN values (Fig. 4 and *SI Appendix*, Table S1), are comparable to errors in ad hoc fits used in the current NN model for most motifs (11, 25–27), and are generally smaller than molecular dynamics calculations that remain significantly more expensive (21, 22). The generality of the RECCES–Rosetta framework and this level of success in initial tests support the further development of RECCES–Rosetta for nonnatural nucleotides and for motifs more complex than the helical stacked pairs and dangling ends considered herein.

While achieving consistently sub-kcal/mol accuracies, there is room for improvement in the RECCES–Rosetta approach. For example, the modeling does not account for the 1 kcal/mol stability increase of the $^{5'GC}_{3'CG}$ NN parameter relative to $^{5'CG}_{3'GC}$; the electrostatic term *stack_elec* does favor the former, but is not assigned a strong enough weight in the final fits to account for the stability difference. Also, the rmsd accuracies still remain larger than estimated systematic errors (0.3 kcal/mol), particularly for the nonnatural base pairs in the test data, and the discrepancies remain even if those data are included in a post hoc fit of the energy function weights to all available measurements. Our results help bracket which strategies might improve the accuracy and which might not. On one hand, nonnatural pairs present their atomic moieties in different bonded contexts, which might modulate the strengths of hydrogen bonds or other interactions that they form. For example, a previous analysis suggested that the hydrogen bonds in an iG–iC base pair might be stronger than in a G–C base pair by ∼0.4 kcal/mol (14). Accounting for this effect would be predicted to offset our calculated NN parameters for all iG–iC stacked pairs, without changing their relative ordering, and cannot account for strong outliers. Indeed, if we added an extra fitting term for stabilizing iG–iC pairing, the rmsd accuracy over these data did not significantly improve (0.96 kcal/mol vs. 0.99 kcal/mol without the extra term). On the other hand, several unmodeled factors are sensitive to the ordering of base pairs within stacked pairs and could affect the relative ordering of NN parameters within each set. For example, the current Rosetta all-atom score function models electrostatics through fixed charges with a distance-dependent dielectric and does not explicitly model water or counterions that may differentially stabilize the base pair steps (35, 36). Recent and planned additions of nonlinear Poisson–Boltzmann solvation models, polarizable electrostatic models, and a potential of mean force for water-mediated hydrogen bonding into the Rosetta framework should allow evaluation of whether these physical effects can improve accuracy of NN

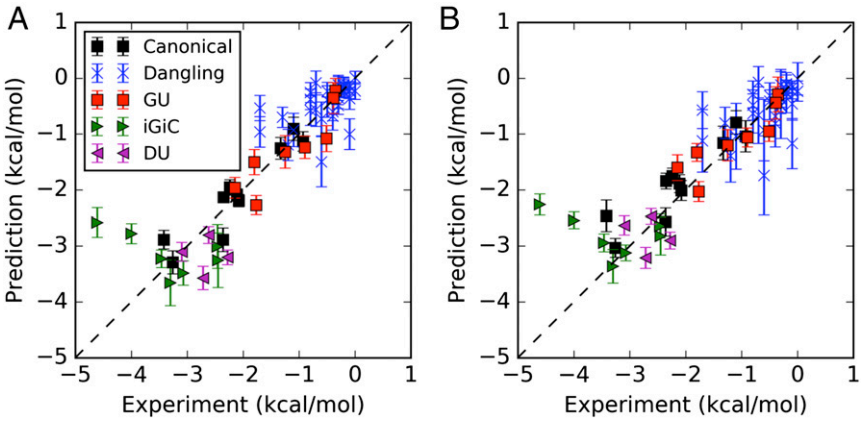


Fig. 4. Calculations vs. experiment across all NN parameters. Comparisons are based on (*A*) RECCES–Rosetta weight sets trained on canonical and dangling-end data (same values as in Fig. 3) and (*B*) "best-case" weight sets fitted post hoc over all available NN parameters, including D–U stacked pairs measured for blind predictions.

parameter calculations to the 0.3 kcal/mol fundamental limit of the RECCES method. If these models can also be expanded to calculate the temperature dependence of solvation, it may also become possible to compare calculated and measured entropies and enthalpies of the NN parameters, which are well measured but may be dominated by solvation effects. In addition, we propose that calculations for recently characterized stacked pairs that give anomalous NN parameters, including some tandem G–U stacked pairs (11) and pseudouridine-A–containing stacked pairs (37, 38), could offer particularly stringent tests.

Continuing work in modeling RNA energetics will benefit from further blind trials, perhaps in a community-wide setting analogous to the ongoing RNA-puzzle structure prediction trials (39, 40). The prediction of two kinds of parameters could serve as future blind tests. First, based on the results herein, nonnatural base pairs offer good test cases and require the same amount of computational power as canonical base pair NN parameter estimation. Alternative approaches based on, e.g., molecular dynamics, should also be applicable to these cases. We have completed RECCES–Rosetta predictions for additional stacked pairs involving iG–iC and D–U pairs, as well as for inosine–cytosine (I–C) base pairs (Fig. 1*E* and *SI Appendix*, Table S1), but are waiting to make experimental measurements until there are comparison values from other groups and approaches. Second, future blind trials might involve predicting energetics of RNA motifs more complex than those considered herein, such as apical loops, internal loops, multihelix junctions, and tertiary interactions. For these cases, an expansion of the RECCES approach in which physically realistic candidate conformations of each motif are first estimated with structure prediction (18, 41) and then subjected to rigorous RECCES-based free-energy calculations may offer predictive power. Such an approach may also allow calculations of next-NN effects and development of rapid approximations to estimate conformational entropy of candidate conformations, which would be useful for structure prediction and design (*SI Appendix*, Fig. S2). A new generation of high-throughput RNA biochemistry platforms (42–44) offers the prospect of both training these next-generation energetic prediction algorithms and carrying out blind tests with many thousands of measurements.

Materials and Methods

Details of NN parameter estimation with RECCES (including basic equations, simulation parameters, and energy function) and with simple single-conformation methods, as well as methods used to experimentally estimate NN parameters for helices with D–U base pairs, are presented in *SI Appendix*.

ACKNOWLEDGMENTS. We thank J. Yesselman for help in generating the partial charges for nonnatural RNA bases; and P. Sripakdeevong, K. Beauchamp, W. Greenleaf, and members of R.D.'s laboratory for useful discussions. Calculations were performed using the Texas Advanced Computing Center Stampede cluster through Extreme Science and Engineering Discovery Environment (XSEDE) Allocation Project MCB120152, and the Stanford BioX^3 cluster. This work is supported by a Howard Hughes Medical Institute International Student Research Fellowship (to F.-C.C.); a Stanford BioX Graduate Student Fellowship (to F.-C.C.); a Burroughs-Wellcome Career Award at Scientific Interface (to R.D.); and NIH Grants NIGMS R21 GM102716 and R01 GM102519 (to R.D.).

1. Bloomfield VA, Crothers DM, Tinoco I (2000) *Nucleic Acids: Structures, Properties, and Functions* (University Science Books, Sausalito, CA).
2. Atkins JF, Gesteland RF, Cech TR (2010) *RNA Worlds: From Life's Origins to Diversity in Gene Regulation* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY).
3. Cong L, et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121):819–823.
4. Mali P, et al. (2013) RNA-guided human genome engineering via Cas9. *Science* 339(6121):823–826.
5. Hannon GJ (2002) RNA interference. *Nature* 418(6894):244–251.
6. Cantor CR, Tinoco I, Jr (1965) Absorption and optical rotatory dispersion of seven trinucleoside diphosphates. *J Mol Biol* 13(1):65–77.
7. Kent JL, et al. (2014) Non-nearest-neighbor dependence of stability for group III RNA single nucleotide bulge loops. *RNA* 20(6):825–834.
8. Lavery R, et al. (2010) A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic Acids Res* 38(1):299–313.
9. Vanegas PL, Horwitz TS, Znosko BM (2012) Effects of non-nearest neighbors on the thermodynamic stability of RNA GNRA hairpin tetraloops. *Biochemistry* 51(11):2192–2198.
10. Xia T, et al. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* 37(42):14719–14735.
11. Chen JL, et al. (2012) Testing the nearest neighbor model for canonical RNA base pairs: revision of GU parameters. *Biochemistry* 51(16):3508–3522.
12. Serra MJ, Turner DH (1995) Predicting thermodynamic properties of RNA. *Methods Enzymol* 259:242–261.
13. Mathews DH, Sabina J, Zuker M, Turner DH (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* 288(5):911–940.
14. Yildirim I, Turner DH (2005) RNA challenges for computational chemists. *Biochemistry* 44(40):13225–13234.
15. Liu B, Diamond JM, Mathews DH, Turner DH (2011) Fluorescence competition and optical melting measurements of RNA three-way multibranch loops provide a revised model for thermodynamic parameters. *Biochemistry* 50(5):640–653.
16. Mathews DH, Turner DH (2002) Experimentally derived nearest-neighbor parameters for the stability of RNA three- and four-way multibranch loops. *Biochemistry* 41(3):869–880.
17. Diamond JM, Turner DH, Mathews DH (2001) Thermodynamics of three-way multibranch loops in RNA. *Biochemistry* 40(23):6971–6981.
18. Das R, Karanicolas J, Baker D (2010) Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat Methods* 7(4):291–294.
19. Sarzynska J, Nilsson L, Kulinski T (2003) Effects of base substitutions in an RNA hairpin from molecular dynamics and free energy simulations. *Biophys J* 85(6):3445–3459.
20. Deng N-J, Cieplak P (2007) Molecular dynamics and free energy study of the conformational equilibria in the UUUU RNA Hairpin. *J Chem Theory Comput* 3(4):1435–1450.
21. Deng N-J, Cieplak P (2010) Free energy profile of RNA hairpins: A molecular dynamics simulation study. *Biophys J* 98(4):627–636.
22. Spasic A, Serafini J, Mathews DH (2012) The Amber ff99 force field predicts relative free energy changes for RNA helix formation. *J Chem Theory Comput* 8(7):2497–2505.
23. Réblová K, et al. (2010) An RNA molecular switch: Intrinsic flexibility of 23S rRNA helices 40 and 68 5′-UAA/5′-GAN internal loops studied by molecular dynamics methods. *J Chem Theory Comput* 2010(6):910–929.
24. Van Nostrand KP, Kennedy SD, Turner DH, Mathews DH (2011) Molecular mechanics investigation of an adenine-adenine non-canonical pair conformational change. *J Chem Theory Comput* 7(11):3779–3792.
25. Znosko BM, Silvestri SB, Volkman H, Boswell B, Serra MJ (2002) Thermodynamic parameters for an expanded nearest-neighbor model for the formation of RNA duplexes with single nucleotide bulges. *Biochemistry* 41(33):10406–10417.
26. Blose JM, et al. (2007) Non-nearest-neighbor dependence of the stability for RNA bulge loops based on the complete set of group I single-nucleotide bulge loops. *Biochemistry* 46(51):15123–15135.
27. Sheehy JP, Davis AR, Znosko BM (2010) Thermodynamic characterization of naturally occurring RNA tetraloops. *RNA* 16(2):417–429.
28. Roost C, et al. (2015) Structure and thermodynamics of N6-methyladenosine in RNA: A spring-loaded base modification. *J Am Chem Soc* 137(5):2107–2115.
29. Zhao BS, He C (2015) Pseudouridine in a new era of RNA modifications. *Cell Res* 25(2):153–154.
30. Ng EW, et al. (2006) Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat Rev Drug Discov* 5(2):123–132.
31. Sipa K, et al. (2007) Effect of base modifications on structure, thermodynamic stability, and gene silencing activity of short interfering RNA. *RNA* 13(8):1301–1316.
32. Proctor DJ, et al. (2004) Folding thermodynamics and kinetics of YNMG RNA hairpins: Specific incorporation of 8-bromoguanosine leads to stabilization by enhancement of the folding rate. *Biochemistry* 43(44):14004–14014.
33. Opitz D, Maclin R (1999) Popular ensemble methods: An empirical study. *J Artif Intell Res* 11:169–198.
34. Chen X, Kierzek R, Turner DH (2001) Stability and structure of RNA duplexes containing isoguanosine and isocytidine. *J Am Chem Soc* 123(7):1267–1274.
35. Auffinger P, Westhof E (2000) Water and ion binding around RNA and DNA (C,G) oligomers. *J Mol Biol* 300(5):1113–1131.
36. Chen Z, Znosko BM (2013) Effect of sodium ions on RNA duplex stability. *Biochemistry* 52(42):7477–7485.
37. Hudson GA, Bloomingdale RJ, Znosko BM (2013) Thermodynamic contribution and nearest-neighbor parameters of pseudouridine-adenosine base pairs in oligoribonucleotides. *RNA* 19(11):1474–1482.
38. Kierzek E, et al. (2014) The contribution of pseudouridine to stabilities and structure of RNAs. *Nucleic Acids Res* 42(5):3492–3501.
39. Cruz JA, et al. (2012) RNA-puzzles: A CASP-like evaluation of RNA three-dimensional structure prediction. *RNA* 18(4):610–625.
40. Miao Z, et al. (2015) RNA-puzzles round II: Assessment of RNA structure prediction programs applied to three large RNA structures. *RNA* 21(6):1066–1084.
41. Sripakdeevong P, Kladwang W, Das R (2011) An enumerative stepwise ansatz enables atomic-accuracy RNA loop modeling. *Proc Natl Acad Sci USA* 108(51):20573–20578.
42. Buenrostro JD, et al. (2014) Quantitative analysis of RNA-protein interactions on a massively parallel array reveals biophysical and evolutionary landscapes. *Nat Biotechnol* 32(6):562–568.
43. Tome JM, et al. (2014) Comprehensive analysis of RNA-protein interactions by high-throughput sequencing-RNA affinity profiling. *Nat Methods* 11(6):683–688.
44. Ozer A, Pagano JM, Lis JT (2014) New technologies provide quantum changes in the scale, speed, and success of SELEX methods and aptamer characterization. *Mol Ther Nucleic Acids* 3:e183.

Supporting Information Appendix for: "Blind tests of RNA nearest-neighbor energy prediction"

Fang-Chieh Chou[1], Wipapat Kladwang[1], Kalli Kappel[2], and Rhiju Das[1,2,3]*

[1] Department of Biochemistry, Stanford University, Stanford, CA 94305, USA
[2] Biophysics Program, Stanford University, Stanford, CA 94305, USA
[3] Department of Physics, Stanford University, Stanford, CA 94305, USA

* Correspondence to: Rhiju Das. Phone: (650) 723-5976. Fax: (650) 723-6783. E-mail: rhiju@stanford.edu.

Supporting Methods

Nearest-neighbor Parameter Estimation with RECCES

In this section, we describe computation of NN parameters under the RECCES framework. First, we review how NN parameters can be determined from either experimental measurements or computational estimates of the folding free energies of individual helix motifs. Hereafter, 'folding free energies' refers to the free energy of association of two separate strands into a folded A-form helix at 1 M standard state. First, we discuss a simple example involving association of two GC segments with each other, in the context of a longer helix. Second, we describe the calculations of NN parameters for the general case where the helix strand segments have different sequences as well as calculations for dangling-ends, again in terms of well-defined folding free energies of specific complexes. The presented relations apply to both experimental and computational studies. Third, for computational approaches, we show how each of these folding free energies can be computed based on the free energies of the complex and of the separate strands (in a 'random coil' state). Last, we describe the RECCES sampling framework, which uses simulated-tempering Monte Carlo to efficiently sample the density of states of and then evaluate the free energy of each single-strand and helix molecule.

Nearest-neighbor Parameters with Symmetric G-C pairs

Here we compute a NN parameter of the stacked pair $\begin{pmatrix} 5'GC \\ 3'CG \end{pmatrix}$ using the folding free energies of two helix motifs with different lengths. The NN model assumes that the folding free energy of an RNA helix can be decomposed into the sum of the contribution of each NN fragment. For example,

$$\Delta G_f \begin{pmatrix} 5'\text{GG} \\ 3'\text{CC} \end{pmatrix} = \Delta G_{NN} \begin{pmatrix} 5'\text{GG} \\ 3'\text{CC} \end{pmatrix} + \Delta G_{init}$$
$$\Delta G_f \begin{pmatrix} 5'\text{GGC} \\ 3'\text{CCG} \end{pmatrix} = \Delta G_{NN} \begin{pmatrix} 5'\text{GG} \\ 3'\text{CC} \end{pmatrix} + \Delta G_{NN} \begin{pmatrix} 5'\text{GC} \\ 3'\text{CG} \end{pmatrix} + \Delta G_{init} \quad (1)$$

where ΔG_f is the folding free energy of the given helix; the terms ΔG_{NN} are the NN parameters; and ΔG_{init} accounts for the entropic penalty of initiating the helix with the first G-C base pair. Hereafter, the free energies are computed at standard states in which each molecule is at 1 M concentration and the temperature is 37 ˚C, conditions at which most values of NN parameters are tabulated in the experimental literature. (Similar expressions for the temperature dependence of energetics lead to relations involving ΔH and ΔS . Calculating these parameters from simulations requires understanding the dependence of solvation and other physical effects on temperature and will not be considered herein.)

In the above example, the folding free energy expression of the second helix contains the $\Delta G_{NN} \begin{pmatrix} 5'\text{GC} \\ 3'\text{CG} \end{pmatrix}$ parameter, which we wish to determine from experimental measurements or computer simulations. Taking the difference of the two folding free energies above, we have:

$$\Delta G_{NN} \begin{pmatrix} 5'\text{GC} \\ 3'\text{CG} \end{pmatrix} = \Delta G_f \begin{pmatrix} 5'\text{GGC} \\ 3'\text{CCG} \end{pmatrix} - \Delta G_f \begin{pmatrix} 5'\text{GG} \\ 3'\text{CC} \end{pmatrix} \quad (2)$$

Note that ΔG_{init} cancels out in the above equation. Calculating ΔG_{init} requires accounting for the translational and rotational entropy lost upon helix association and requires separate computations beyond the scope of the present work.

Nearest-neighbor Parameter Estimation, General Case

The example above illustrates the evaluation of NN parameters involving two helix-associating segments with the same two-nucleotide sequence. The more general case involves taking into account an additional parameter, the terminal penalty, as discussed next.

As an example, we describe the computation of the nearest-neighbor parameter $\Delta G_{NN}\begin{pmatrix}5'\text{GA}3'\\3'\text{CU}5'\end{pmatrix}$.

Following the procedure in eqs. (1)-(2) above, this can be evaluated from folding free energies $\Delta G_f\begin{pmatrix}5'\text{GGA}\\3'\text{CCU}\end{pmatrix}$ and $\Delta G_f\begin{pmatrix}5'\text{GG}\\3'\text{CC}\end{pmatrix}$. The relationships assumed by the NN model are:

$$\begin{aligned}\Delta G_f\begin{pmatrix}5'\text{GGA}\\3'\text{CCU}\end{pmatrix} &= \Delta G_{NN}\begin{pmatrix}5'\text{GG}\\3'\text{CC}\end{pmatrix} + \Delta G_{NN}\begin{pmatrix}5'\text{GA}\\3'\text{CU}\end{pmatrix} + \Delta G_{init} + \Delta G_{Terminal-AU}\\ \Delta G_f\begin{pmatrix}5'\text{GG}\\3'\text{CC}\end{pmatrix} &= \Delta G_{NN}\begin{pmatrix}5'\text{GG}\\3'\text{CC}\end{pmatrix} + \Delta G_{init}\end{aligned} \qquad (3)$$

Here $\Delta G_{Terminal-AU}$ is the terminal contribution for having an A-U pair at the terminal of the helix instead of G-C. Determining this additional term requires additional folding free energies, as follows. The definition of the NN model requires $\Delta G_{NN}\begin{pmatrix}5'\text{GA}3'\\3'\text{CU}5'\end{pmatrix}$ to be the same as $\Delta G_{NN}\begin{pmatrix}5'\text{UC}3'\\3'\text{AG}5'\end{pmatrix}$, flipping the upper and lower sequence. Analogous to eq. (3), we can write the parameter in terms of $\Delta G_f\begin{pmatrix}5'\text{GUC}\\3'\text{CAG}\end{pmatrix}$ and $\Delta G_f\begin{pmatrix}5'\text{GU}\\3'\text{CA}\end{pmatrix}$:

$$\Delta G_f\begin{pmatrix}5'\text{GUC}\\3'\text{CAG}\end{pmatrix}=\Delta G_{NN}\begin{pmatrix}5'\text{GU}\\3'\text{CA}\end{pmatrix}+\Delta G_{NN}\begin{pmatrix}5'\text{UC}\\3'\text{AG}\end{pmatrix}+\Delta G_{init}$$
$$\Delta G_f\begin{pmatrix}5'\text{GU}\\3'\text{CA}\end{pmatrix}=\Delta G_{NN}\begin{pmatrix}5'\text{GU}\\3'\text{CA}\end{pmatrix}+\Delta G_{init}+\Delta G_{Terminal-AU} \quad (4)$$

The NN parameter and the terminal A-U contribution can now be computed in terms of measurable folding free energies ΔG_f of helix association as:

$$\Delta G_{NN}\begin{pmatrix}5'\text{GA}\\3'\text{CU}\end{pmatrix}=\frac{1}{2}\left[\Delta G_f\begin{pmatrix}5'\text{GGA}\\3'\text{CCU}\end{pmatrix}-\Delta G_f\begin{pmatrix}5'\text{GG}\\3'\text{CC}\end{pmatrix}+\Delta G_f\begin{pmatrix}5'\text{GUC}\\3'\text{CAG}\end{pmatrix}-\Delta G_f\begin{pmatrix}5'\text{GU}\\3'\text{CA}\end{pmatrix}\right]$$
$$\Delta G_{Terminal-AU}=\frac{1}{2}\left[\Delta G_f\begin{pmatrix}5'\text{GGA}\\3'\text{CCU}\end{pmatrix}-\Delta G_f\begin{pmatrix}5'\text{GG}\\3'\text{CC}\end{pmatrix}-\Delta G_f\begin{pmatrix}5'\text{GUC}\\3'\text{CAG}\end{pmatrix}+\Delta G_f\begin{pmatrix}5'\text{GU}\\3'\text{CA}\end{pmatrix}\right] \quad (5)$$

Other NN parameters and terminal contributions can be similarly evaluated in terms of folding free energies of two-base-pair and three-base-pair helices. In our calculations, we checked that using helices of different lengths or with a different first base pair led to systematic errors of 0.26 kcal/mol or less in final NN values (Table S2).

The Dangling-End Nearest-neighbor Parameters

The other NN parameters considered herein are for single-nucleotide dangling ends. These dangling-end parameters contribute to the folding free energies of complexes such as $\Delta G_f\begin{pmatrix}5'\text{GAA}\\3'\text{CU}\end{pmatrix}$:

$$\Delta G_f\begin{pmatrix}5'\text{GAA}\\3'\text{CU}\end{pmatrix}=\Delta G_{NN}\begin{pmatrix}5'\text{GA}\\3'\text{CU}\end{pmatrix}+\Delta G_{NN}\begin{pmatrix}5'\text{AA}\\3'\text{U}\end{pmatrix}+\Delta G_{init}+\Delta G_{Terminal-AU}$$
$$\Delta G_f\begin{pmatrix}5'\text{GA}\\3'\text{CU}\end{pmatrix}=\Delta G_{NN}\begin{pmatrix}5'\text{GA}\\3'\text{CU}\end{pmatrix}+\Delta G_{init}+\Delta G_{Terminal-AU} \quad (6)$$

Similar to above [eq. (2) and (5)], the dangling-end NN parameter $\Delta G_{NN}\begin{pmatrix}5'\text{AA}\\3'\text{U}\end{pmatrix}$ can be estimated from the difference of folding free energies for complexes with and without the dangling-ends:

$$\Delta G_{NN}\begin{pmatrix}5'\text{AA}\\3'\text{U}\end{pmatrix} = \Delta G_f\begin{pmatrix}5'\text{GAA}\\3'\text{CU}\end{pmatrix} - \Delta G_f\begin{pmatrix}5'\text{GA}\\3'\text{CU}\end{pmatrix} \quad (7)$$

As a brief note of clarification, the experimental measurements of the dangling-end parameters were presented in a 1995 study, before the A-U terminal contribution was introduced into the NN model (1); eq. (6) is the appropriate update to include the terminal contribution. In either case, however, the simple expression (7) applies to determine the dangling end NN parameters from folding free energies, and the tabulated computational values below use this shared expression.

Free Energy Computation and Score Function Reweighting

In the above sections, we expressed the NN parameters as linear combinations of the folding free energies of strands folding into helices, which can be measured experimentally or estimated through computation. In our computational approach, the folding free energy of a helix is defined as the difference in the free energies of the helix and of the two separated single-strand molecules, i.e. the two states involved in the folding equilibrium (main text Figure 2a):

$$\begin{gathered}
\mathrm{GGC} + \mathrm{GCC} \rightleftharpoons \begin{matrix}\mathrm{5'GGC} \\ \mathrm{3'CCG}\end{matrix} \\
\Delta G_f\begin{pmatrix}\mathrm{5'GGC} \\ \mathrm{3'CCG}\end{pmatrix} = G\begin{pmatrix}\mathrm{5'GGC} \\ \mathrm{3'CCG}\end{pmatrix} - G\left(\mathrm{GGC}\right) - G\left(\mathrm{GCC}\right) \\
\mathrm{GG} + \mathrm{CC} \rightleftharpoons \begin{matrix}\mathrm{5'GG} \\ \mathrm{3'CC}\end{matrix} \\
\Delta G_f\begin{pmatrix}\mathrm{5'GG} \\ \mathrm{3'CC}\end{pmatrix} = G\begin{pmatrix}\mathrm{5'GG} \\ \mathrm{3'CC}\end{pmatrix} - G\left(\mathrm{GG}\right) - G\left(\mathrm{CC}\right)
\end{gathered} \tag{8}$$

Here the terms $G\begin{pmatrix}\mathrm{5'GGC} \\ \mathrm{3'CCG}\end{pmatrix}$, $G\left(\mathrm{GGC}\right)$, etc. are not free energy differences (ΔG) but are instead the 'raw' free energies of the helices and single-strands, each defined by an integral over the system's conformational space:

$$\begin{aligned}
G &= -k_{\mathrm{B}} T \ln Z \\
Z &= \iint \ldots \int d\theta_1 \, d\theta_2 \ldots d\theta_n \exp\left(-\frac{E(\theta_1, \theta_2, \ldots, \theta_n)}{k_{\mathrm{B}} T}\right)
\end{aligned} \tag{9}$$

Here the partition function *Z* is the integration of the Boltzmann factor over all accessible torsion angles θ_i of the molecule, *E* is the internal energy of the molecule at a conformation specified by the torsion angles, k_{B} is the Boltzmann constant, and *T* is the system's temperature. In the folded state, because the energy of a single molecule is independent of translational and rotational degrees of freedom, those degrees of freedom only contribute to a constant factor to the free energy of the system, which is omitted in the above expressions. For bimolecular systems (e.g., the two disassociated strands), the system free energy does depend on the relative positions and orientations of the molecules, but the translational/rotational entropy from these terms cancel out in the determination of the NN parameters [see, e.g., eqs. (2) and (7) above] considered herein.

While eq. (9) can theoretically be used to directly compute the molecule's free energy, it is challenging to integrate the expression over all torsion angles, especially when the system has a large number of degrees of freedom. Instead we compute the partition function by estimating the molecule's density of states (DOS), $g(E)$:

$$
\begin{aligned}
Z &= V_p \int_{-\infty}^{\infty} dE g(E) \exp\left(-\frac{E}{kT}\right) \\
g(x) &= \frac{\int \ldots \int d\theta_1 \ldots d\theta_n \delta(x - E(\theta_1, \cdots, \theta_n))}{V_p} \\
V_p &= \int \ldots \int d\theta_1 \ldots d\theta_n
\end{aligned}
\tag{10}
$$

Here δ is the Dirac delta function; and V_p is the total phase space volume available to the molecule, which can be calculated exactly (see below). Due to normalization by the V_p factor, the DOS $g(E)$ integrates to unity. The DOS describes the probability distribution of the molecule's energy at infinite temperature. Once determined, it allows the calculation of the partition function and hence free energies at any temperature. However, in practice, an infinite temperature simulation gives negligible sampling of the DOS for low energy states. Therefore, to estimate the DOS precisely at all temperatures, we carry out simulations of the conformational ensemble at different temperatures, providing estimates of $g(E)$ within different, overlapping temperature ranges, up to a different scaling factor for each simulation. Overlaying these distributions in overlapping energy regimes defines the unknown scaling factors at each temperature and yields a portrait of g(E) across all temperature ranges, with the final overall normalization set by the property that g(E) integrates to unity.

To carry out simulations of conformational ensembles, we used Metropolis Monte Carlo sampling, with states defined as follows. For a single RNA strand, we allowed all backbone and side-chain torsions to freely sample the entire range of 2π = 360°, except for the sugar puckers, which were only allowed to sample two conformations, the ideal 2'-endo and 3'-endo puckers (main text Figure 2A, left panels). The phase space volume is $V_p = 2^n (2\pi)^n (2\pi)^{5(n-1)}$, where *n* is the number of nucleotides in the strand. Here 2^n represents the two sugar pucker forms,

$(2\pi)^n$ represents the side-chain torsions (χ angle), and $(2\pi)^{5(n-1)}$ represent the backbone torsions connecting the sugars [five torsions (ε, ζ, α, β, γ) for each connection, *n* – 1 connections between *n* nucleotides]. The 5′ and 3′ terminal phosphates were omitted in the simulated constructs. When included, these phosphates did not make stable interactions and instead gave constant entropic contributions that canceled out during the folding free energy evaluation.

For the helix state, we froze the relative position between the two strands by forcing the first base pair to take an ideal geometry (main text Figure 2A, right panels). This constraint eliminates translation and rotational entropy contributions to computed helix free energy, which cancel out during the calculations of NN parameters; see, e.g., eq. (2). This cancellation allows the sampling to focus on estimating energy fluctuations and conformational entropy important for the considered NN parameters. The sugar puckers of all nucleotides were restricted to the 3'-endo form, and the other torsions were allowed to sample values between ±60˚ around ideal A-form torsion angles (α = –64˚, β = 176˚, γ = 53˚, ε = –150˚, ζ = –71˚; χ =79°). The phase space volume over both strands is therefore $V_p = \left(\frac{2\pi}{3}\right)^{2n} \left(\frac{2\pi}{3}\right)^{2\times 5(n-1)}$, where *n* is the length of each strand in the helix. We note that the constraints described above provide our working definition of the helix state. We confirmed in separate calculations that changing the backbone torsion constraint from ±60˚ to ±40˚ and using alternative ideal pucker conformations, led to negligible

changes in computed free energies (0.08 and 0.13 kcal/mol error; Table S2). For both helix and strand states, we allowed the torsion angle for 2′-OH to sample 360°, leading to additional phase space volume contributions that canceled out during evaluation of the folding free energy.

For all Monte Carlo runs, we used a new application ('recces_turner') in Rosetta, with the computed energy based on the Rosetta scoring function. (See Supporting Results for command-lines and version numbers; Rosetta is freely available for academic users at www.rosettacommons.org.) The function is a linear combination of multiple component terms, including a Lennard-Jones potential, hydrogen bonding terms, an orientation-dependent solvation term, and long-range electrostatics (2). As noted above, we performed Monte Carlo simulations at multiple temperatures, ranging from room temperature to infinity. We applied the simulated tempering method (3) to facilitate the conformational search and barrier crossing during the Monte Carlo runs by allowing the temperature to vary during the simulation in a manner satisfying detailed balance (main text Figure 2B; see Supporting Methods for details). For each molecule, a typical simulation took less than an hour on a single CPU to generate up to 10 million Monte Carlo samples. The simulated tempering parameters were determined using short single-temperature simulations, as described by Huang et al. (4) We combined all simulations into one DOS using the Weighted Histogram Analysis Method (WHAM) (5, 6).

The conformational ensemble obtained from the above simulation scheme allowed rapid tests of alternative energy functions based on different weights on the Rosetta score terms. Caching the contribution of each score term for each sampled conformation allowed rapid calculation of the entire simulation ensemble with different score-term weights; see also refs (7, 8). For each molecular ensemble, this reweighting step was a linear matrix operation that took less than 0.1 s on a single CPU.

Main text Figure 2C illustrates this procedure. This rapid reweighting enabled optimization of the weight set, by minimizing the difference between the computed NN parameters and the experimental values in a training set using a standard quasi-Newton minimization algorithm. Because the target cost function is not convex and has a large number of local minima, we repeated the minimization using thousands of randomly initialized starting weights. These separate minimization runs gave a collection of score functions, all compatible with the experimental NN parameters in the training set, and allowed an estimate of the systematic errors for the prediction of new NN parameters.

Details of Simulated Tempering Monte Carlo simulations

The free energy of each sequence (single-strand or helix) was evaluated using simulated tempering Monte Carlo (MC). For each sequence, we performed simulations at 7 temperatures T, with k_BT = 0.8, 1, 1.4, 1.8, 3, 7 and 30 Rosetta units (RU), where k_B is the Boltzmann constant. In later reweighting stages we calibrated the score so that one RU equals one k_BT for T = 310.15 K (37 °C, at which most nearest-neighbor parameters are tabulated). Before the simulated tempering run, we first performed short regular MC simulations (300,000 steps) at each of the seven temperatures, to determine parameters that govern switching between temperatures during the simulated tempering run. The simulated tempering parameters were computed by numerically solving the weight difference for each pair of neighbor temperatures such that the mean probability of moving upward and downward in the temperature ladder was the same (4). The initial values of these parameters, used by the numerical equation solver, were computed using the average energy for each temperature (9). With the computed simulated tempering parameters, we performed a long simulated tempering simulation (9,000,000 steps). In addition to simulated tempering, we also performed a regular MC simulation at infinite temperature, to calibrate the full DOS profile. For all regular MC and simulated tempering simulations,

acceptance rates for the conformational moves and temperature switches were above 10%. The simulations stored the total score and each separate score term component (Tables S3, S6, and S7) for all sampled conformations, so that scores could be recomputed for different weight sets without rerunning the simulations. Example Rosetta command lines (for Rosetta release 3.6 and following releases) are given below:

1. Short pre-runs:

   ```
   recces_turner -score:weights stepwise/rna/turner -n_cycle 300000 -seq1 gaa -temps 0.8 -
   out_prefix prerun
   ```

   The above command performs a short single-temperature pre-run for the determination of the simulated tempering parameters. Here k_BT = 0.8 RU, and the simulated sequence is GAA (single strand).

2. Simulated tempering:

   ```
   recces_turner -score:weights stepwise/rna/turner -seq1 gu -seq2 aac -n_cycle 9000000 -
   temps 0.8 1 1.4 1.8 3 7 30 -st_weights 0 8.04 15.39 17.78 17.76 14.96 11.81 -out_prefix
   ST
   ```

   This command performs simulated tempering simulation on $\begin{matrix}5'\text{GU}\\3'\text{CAA}\end{matrix}$ (dangling end).

3. Infinite temperature:

   ```
   recces_turner -score:weights stepwise/rna/turner -seq1 gu -seq2 aac -n_cycle 300000 -
   temps -1 -out_prefix kT_inf
   ```

   The "-1" temperature stands for infinite temperature, where all MC moves are accepted.

These simulation results were then combined into a single DOS using WHAM. First the simulation data at each temperature were aggregated into histograms with bin size of 0.1 RU. For the infinite temperature

simulation, the simulated conformation could have very high energy (> 1000 RU). Since these high energy conformations are not sampled at room temperature, and are important only for normalizing the density of states, we binned all conformations with scores higher than 800 RU into a single bin. We verified that using different bin sizes and score cutoffs in this procedure gave negligible changes to the results (<0.1 kcal/mol, Table S2). These histograms were then combined into one DOS using WHAM. During the combination calculation, WHAM assigned a weight to each energy bin (so all conformations in the same bin share the same weight). We recorded these weights for all conformations, which were needed for reweighting, as discussed next.

Details of Reweighting and Training Scheme

After collecting the conformational samples for all relevant sequences, we reweighted the Rosetta score function to minimize the prediction error on a limited training set of canonical and dangling end NN parameters. In this section we present the technical details of these reweighting and training steps.

For each sampled conformation, the Rosetta score is the weighted linear combination of all individual score terms:

$$Score(w_1,\cdots,w_m)=\sum_{i=1}^{m} w_i s_i \qquad (11)$$

Table S5 gives a short description of each score term; all were introduced in ref. (10) except for an electrostatic interaction between nucleobases (*stack_elec*), described below. Here w_i is the weight for each score term, and s_i is the value of the score term for the conformation. To reweight the score with a new set of weights, we updated the value of w_i in equation (11) to obtain a new score for each

conformation. The reweighted scores were then combined into DOS, using the WHAM weights computed previously (see section above). This reweighted DOS was then used to compute the free energy under the new weight set.

To optimize weight sets from training data (i.e. canonical and dangling-end parameters), we minimized the training error with respect to the score term weights. The training error was a weighted square error between the experimental and predicted NN parameters:

$$Error = \sum \left(G_{\text{canonical,expt}} - G_{\text{canonical,pred}} \right)^2 + 0.1 \times \sum \left(G_{\text{dangling,expt}} - G_{\text{dangling,pred}} \right)^2 + \left(G_{\text{terminal,expt}} - G_{\text{terminal,pred}} \right)^2 \quad (12)$$

Each dangling end data point was given a weight of 0.1 compared to canonical data points, to avoid overfitting to the dangling end values, which were measured less accurately with fewer experimental measurements. During the training step, we minimized the error function (12) with respect to the score term weights, using the truncated Newton minimizer in Scipy (method *TNC* in scipy.optimize.minimize). To ensure the minimized weights were reasonable, we constrained the score term weights to be within certain ranges during minimization. Here the weights for *fa_atr* and *fa_rep* were constrained to be in [0.1, 20], weights for *hbond_sc* and *rna_torsion* were constrained to be in [0.5, 20], and the weight for *fa_intra_rep* was constrained to be in [0, 0.01]. All other score terms were constrained to be in [0, 20]. Before the minimization, the score term weights were randomly initialized to be between 1/10 to 10 times of the initial weights (Table S4). Minimizing with all the sampled data was computationally slow due to the large number of conformations being reweighted in each stage. Instead, we used only 1% of the data (randomly selected) for each minimization run. The final training errors for the minimized weight sets were computed using the full dataset. Because the target error function is not convex, the

minimization generated different results when initialized with different values. Rather than attempt an aggressive global optimization, we chose to assemble a large collection of locally optimized weight sets to help bracket systematic errors in predicting new energetic parameters. Here we repeated the minimization 9,544 times to obtain a diverse set of minimized weights (see Table S6 for example weights). These minimized weight sets were not all distinct; clustering analysis with a cluster radius of 0.4 (based on the Euclidean distance between the weight vectors) on 95% of the lowest error weight sets led to 1315 unique clusters. All the reweighting and minimization stages were carried out in separate Python scripts, also available in Rosetta (in subdirectory tools/rna_tools/recces).

Electrostatic interactions across stacked bases

Electrostatic interactions are critical components of the energetics of biomolecules. In previous Rosetta RNA modeling work, hydrogen bonding interactions were modeled using a potential derived from database hydrogen bond geometries and tested with quantum mechanical calculations (11). Electrostatic contributions beyond this hydrogen bonding term were modeled in a limited manner (10, 12-14), through a weak carbon hydrogen bond potential (not included here due to lack of constraints from NN data; not shown) and a highly screened electrostatic repulsion between the backbone phosphates (*fa_elec_rna_phos_phos*)(10). Electrostatic interactions between bonded atoms were included implicitly in the RNA torsional potential, which was derived in a knowledge-based fashion using crystal structures in Protein Data Bank (PDB). The rather limited modeling of electrostatics followed the philosophy of Rosetta protein modeling, in which complex physical effects were left out until strong evidence from structure prediction or design tests have supported and helped calibrate the inclusion of these terms (15).

Recent studies of RNA model systems have suggested that the electrostatic interaction between the stacking bases might play an important role in the NN interactions of RNA (16). To test this hypothesis, we implemented a new score term called *stack_elec*. This new term models the electrostatic interaction between atom pairs in different bases through Coulomb's law with a distance-proportional dielectric (the same as the Rosetta *fa_elec* term), but the interaction is suppressed to zero for atom pairs whose inter-atom vector *r* is perpendicular to the base normal of the first or second base (parametrized by angles κ_1 and κ_2): $\frac{q_1 q_2}{r^2}(\cos^2\kappa_1 + \cos^2\kappa_2)$, where q_1 and q_2 are the atom partial charges. As a result of the orientation dependence, the electrostatic interactions between stacked but not co-planar bases are captured in this term, and so *stack_elec* could be optimized separately from the knowledge-based Rosetta hydrogen bond potential. Without the orientation-dependent suppression (i.e., use of the original Rosetta *fa_elec* term), we observed strong repulsion between fixed positive charges of hydrogen atoms in, e.g., A-U pairs that could not be reconciled with the stability of the Watson-Crick arrangement. RECCES modeling of canonical stacked pairs with *fa_elec* instead of *stack_elec* gave worse RMSDs to canonical pair NN parameters (0.48 kcal/mol compared to 0.33 kcal/mol); and in simulations with both terms, optimization of these terms' weights recovered *stack_elec* at positive weight and *fa_elec* at zero weight. The partial charges for non-natural RNA residues were derived using the MATCH method (multipurpose atom-typer for CHARMM) (17). For canonical RNA residues, we used the default atomic partial charges in Rosetta for the *stack_elec* calculations. Calculations for NN parameters for canonical stacked pairs using MATCH charges instead of default Rosetta charges gave similar results, with an RMSD of 0.21 kcal/mol (comparable or less than other sources of systematic error; see Table S2).

Simple single-conformation methods

As baseline comparisons for the RECCES calculations above, we sought NN parameter predictions from phenomenological models that required negligible computation: a hydrogen-bond counting model and a single-conformation Rosetta scoring model. These two models used the same pipeline as RECCES to predict the NN parameters, by first computing the folding free energies for each helix motif, then combining these free energies into NN parameters. However instead of performing comprehensive ensemble sampling, these two models used simple approximations to evaluate the folding free energies.

Hydrogen-bond Counting Model

For the hydrogen-bond counting model, the folding free energy for a helix construct is

$$\begin{aligned}\Delta G_f(S_1,S_2) = N_{HB}(S_1,S_2)\Delta G_{HB} + N_{Dangle}(S_1,S_2)\Delta G_{Dangle} + \\ [L(S_1)+L(S_2)-2]\Delta G_{Stack} + [L(S_1)+L(S_2)]\Delta G_{Base} + \Delta G_{init}\end{aligned} \quad (13)$$

Here S_1 and S_2 are the two strand sequences in a helix construct. $N_{HB}(S_1,S_2)$ and $N_{Dangle}(S_1,S_2)$ are the number of hydrogen-bonds and dangling-ends in a fully associated helix conformation. For example, $\begin{matrix}5'\text{GA}\\3'\text{CU}\end{matrix}$ has 5 hydrogen bonds (3 for the G-C, plus 2 for the A-U) and no dangling-ends; $\begin{matrix}5'\text{GAA}\\3'\text{CU}\end{matrix}$ also has 5 hydrogen bonds, plus 1 dangling-end. $L(S_1)$ and $L(S_2)$ are the lengths of the single-strand S_1 and S_2. ΔG_{HB} is the free energy contribution per hydrogen-bond, ΔG_{Dangle} is the free energy contribution for the conformational entropy of each dangling-end, ΔG_{Stack} is the free energy contribution for each base-base stacking (L - 1 base-base stacking per strand), and ΔG_{Base} is the entropy loss during folding for each base in the single-strands. This model assumes the folding free energy of a helix is a simple linear function of

the number of hydrogen bonds, as hydrogen-bonding is the dominant interaction in helix formation; the contributions of other relevant physical factors, such as base stacking and entropy loss upon helix formation, are approximated to be sequence independent.

We may further rewrite eq. (13) as follows:

$$\Delta G_f(S_1,S_2) = N_{HB}(S_1,S_2)\Delta G_{HB} + N_{Dangle}(S_1,S_2)\Delta G_{Dangle} + [L(S_1)+L(S_2)]\Delta G_S + \Delta G_{const} \quad (14)$$

Here $\Delta G_S = \Delta G_{Stack} + \Delta G_{Base}$ and $\Delta G_{const} = \Delta G_{init} - 2G_{Stack}$. In eq. (14), ΔG_{HB}, ΔG_{Dangle}, and ΔG_S are model parameters; ΔG_{const} is a constant factor that cancels out when we compute the NN parameters.

With the above folding free energy expressions, we can linearly combine them into NN parameters, as we did in the RECCESS framework. For example, in this model, the NN parameters for $\begin{matrix}5'\text{GC}\\3'\text{CG}\end{matrix}$ and $\begin{matrix}5'\text{AA}\\3'\text{U}\end{matrix}$ can be expressed as

$$\Delta G_{NN,HB}\begin{pmatrix}5'\text{GC}\\3'\text{CG}\end{pmatrix} = 3\Delta G_{HB} + 2\Delta G_S$$
$$\Delta G_{NN,HB}\begin{pmatrix}5'\text{AA}\\3'\text{U}\end{pmatrix} = \Delta G_{Dangle} + \Delta G_S \quad (15)$$

The model parameters were determined by least-squares regression against the experimental NN parameters for canonical base pairs and dangling-ends (see Supporting Methods). The optimized parameters were ΔG_{HB} = –1.89 kcal/mol, ΔG_S = +1.31 kcal/mol, and ΔG_{Dangle} = –1.83 kcal/mol .

Single-conformation Rosetta Scoring Method

Instead of simply counting the hydrogen bonds, the single-conformation Rosetta scoring method uses the Rosetta score to approximate the free energy of a construct. Similar to eq. (13), the model is

$$\Delta G_f(S_1,S_2) = k_{Rosetta} Score(S_1,S_2) + N_{Dangle}(S_1,S_2)\Delta G_{Dangle} + [L(S_1)+L(S_2)]\Delta G_S + \Delta G_{const} \quad (16)$$

Here, $Score(S_1,S_2)$ is the Rosetta score of a representative conformation of the target helix. This representative conformation was obtained by minimizing the helix conformation under the Rosetta score function. We used the standard Rosetta score function for RNA structure prediction (*rna_hires*). $k_{Rosetta}$ is a model parameter that sets the scale between the Rosetta score and the folding free energy in kcal/mol; see also ref. (10).

For example, with this model, the NN parameters for $\begin{matrix}5'\text{GC}\\3'\text{CG}\end{matrix}$ and $\begin{matrix}5'\text{AA}\\3'\text{U}\end{matrix}$ can be expressed as

$$\Delta G_{NN}\begin{pmatrix}5'\text{GC}\\3'\text{CG}\end{pmatrix} = k_{Rosetta}\left[Score\begin{pmatrix}5'\text{GGC}\\3'\text{CCG}\end{pmatrix} - Score\begin{pmatrix}5'\text{GG}\\3'\text{CC}\end{pmatrix}\right] + 2\Delta G_S$$
$$\Delta G_{NN}\begin{pmatrix}5'\text{AA}\\3'\text{U}\end{pmatrix} = k_{Rosetta}\left[Score\begin{pmatrix}5'\text{GAA}\\3'\text{CU}\end{pmatrix} - Score\begin{pmatrix}5'\text{GA}\\3'\text{CU}\end{pmatrix}\right] + \Delta G_{Dangle} + \Delta G_S \quad (17)$$

The corresponding model parameters, determined by training against experimental NN parameters with canonical base pairs and dangling-ends, are $k_{Rosetta}$ = 0.7 kcal/mol/Rosetta-unit, ΔG_S = +2.12 kcal/mol, and ΔG_{Dangle} = +1.54 kcal/mol.

Optical Melting Measurements on Helices with D-U Base pairs

To test the prediction accuracy of RECCES, we experimentally measured the folding free energies of RNA helices containing D-U, a non-natural base pair for which the NN parameters have not been previously determined. We performed optical melting experiments on six helices containing D-U using a Shimadzu Spectrophotometer UV-1800, as well as two helices with canonical base pairs, which were confirmed to give thermodynamic parameters that reproduced literature values. The RNA constructs were ordered from Dharmacon with HPLC purification. For each construct, we obtained 12 melting curves at three different concentrations (15, 25 and 35 μM), measured at 260 nm. The buffer system was 1.0 M NaCl, 20 mM sodium cacodylate at pH 7.0, and 0.5 mM Na_2EDTA. For each sample, RNA concentrations were determined *in situ* using the high-temperature absorbances and extinction coefficients of RNA single-strands (18, 19). The extinction coefficient parameter for 2,6-diaminopurine is unknown. The free energies computed herein are insensitive to small changes of the extinction coefficients (20); we assumed that 2,6-diaminopurine has the same parameters as adenine. The melting curves were fitted with a two-state model with linear baselines to obtain the corresponding folding free energies for all constructs (21-23). To obtain the NN parameters from the helix folding free energies, recall that the helix folding free energies can be written as linear combinations of NN parameters (e.g., see eqs. (1), (3) and (5)). The NN parameters can be solved through least-squares regression, with the exception of $\Delta G_{Terminal-DU}$, the terminal contribution of D-U relative to G-C. Determining $\Delta G_{Terminal-DU}$ requires measurements on RNA sequences with D at the 5'-end of strands and with D on the 3'-end of the strands (see, e.g., eq. (4)), but the latter are not commercially available. For results in the main text, we have assumed that this parameter is zero, as it should have the same geometry and same number of hydrogen-bonds (3) as G-C. Assuming different terminal D-U contributions from –0.2 to +0.3 kcal/mol

(compiled in Table S9) gives similar RMSD accuracies of our predictions compared to experimental values.

Supporting Results

Accuracies of terminal base pair penalty calculations

In addition to predicting NN parameters for two-base-pair segments, we were able to calculate terminal contributions with RECCES-Rosetta (see main text, eq. (4)). The terminal contribution is the free energy contribution for having a non-G-C base pair instead of a G-C pair at the end of a helix. For example, because an A-U base pair only has two hydrogen bonds, one less than that of G-C pair, putting it at the end of the helix incurs a penalty in folding free energy. Our method predicted the terminal contribution for A-U pair to be 0.75 ± 0.15 kcal/mol, somewhat higher than the experimental value (0.45 kcal/mol) (21). For G-U pairs, our prediction suggested it to be similar but slightly less stable than A-U pairs at helix terminal (0.87 ± 0.3 kcal/mol), while recent experiments have shown that it is equally stable as G-C pair at terminal (24) (0 terminal contribution). The higher experimental stability of G-U pair at the terminal may be due to the variety of non-A-form alternative conformations that it is known to the sample (24); our current calculations only sampled near-A-form conformations for helices. For the iG-iC pair, our method predicted a terminal contribution of –0.11 ± 0.16 kcal/mol, within error of the experimental value (–0.19 ± 0.07 kcal/mol) (20).

Supporting Tables

Table S1. Experimental NN parameters (20, 21, 25) and RECCES-Rosetta predictions.

NN Sequence	RECCES-Rosetta	Expt.	NN Sequence	RECCES-Rosetta	Expt.
Canonical[a]					
5'AA3' 3'UU5'	-1.13 ± 0.17	-0.93 ± 0.03	5'UC3' 3'AG5'	-2.13 ± 0.09	-2.35 ± 0.06
5'AU3' 3'UA5'	-0.91 ± 0.21	-1.1 ± 0.08	5'UG3' 3'AC5'	-2.09 ± 0.10	-2.11 ± 0.07
5'AC3' 3'UG5'	-1.95 ± 0.14	-2.24 ± 0.06	5'CC3' 3'GG5'	-3.29 ± 0.21	-3.26 ± 0.07
5'AG3' 3'UC5'	-2.19 ± 0.11	-2.08 ± 0.06	5'CG3' 3'GC5'	-2.89 ± 0.21	-2.36 ± 0.09
5'UA3' 3'AU5'	-1.26 ± 0.20	-1.33 ± 0.09	5'GC3' 3'CG5'	-2.88 ± 0.17	-3.42 ± 0.08
Terminal A-U	0.76 ± 0.15	0.45 ± 0.04			
G-U pairs[b]					
5'AG3' 3'UU5'	-0.22 ± 0.22	-0.35± 0.08	5'GG3' 3'CU5'	-1.50 ± 0.23	-1.8 ± 0.09
5'AU3' 3'UG5'	-1.24 ± 0.20	-0.9 ± 0.08	5'GU3' 3'CG5'	-1.96 ± 0.19	-2.15 ± 0.10
5'CG3' 3'GU5'	-1.32 ± 0.29	-1.25 ± 0.09	5'GA3' 3'UU5'	-1.08 ± 0.23	-0.51 ± 0.08
5'CU3' 3'GG5'	-2.27 ± 0.17	-1.77 ± 0.09	5'UG3' 3'AU5'	-0.35 ± 0.28	-0.39 ± 0.09
Terminal G-U	0.87 ± 0.30	0			
Isoguanosine(iG)-Isocytidine(iC)[b]					
5'AiG3' 3'UiC5'	-2.34 ± 0.26	N.A.	5'CiG3' 3'GiC5'	-3.01 ± 0.39	-2.46 ± 0.08
5'AiC3' 3'UiG5'	-1.84 ± 0.27	N.A.	5'CiC3' 3'GiG5'	-3.23 ± 0.16	-3.46 ± 0.11
5'UiG3' 3'AiC5'	-1.93 ± 0.24	N.A.	5'iGiG3' 3'iCiC5'	-3.66 ± 0.41	-3.30 ± 0.17
5'UiC3' 3'AiG5'	-2.14 ± 0.23	N.A.	5'iGiC3' 3'iCiG5'	-2.58 ± 0.26	-4.61 ± 0.17
5'GiG3' 3'CiC5'	-3.48 ± 0.22	-3.07 ± 0.11	5'iCiG3' 3'iGiC5'	-3.25 ± 0.49	-2.45 ± 0.17
5'GiC3' 3'CiG5'	-2.78 ± 0.18	-4.00 ± 0.09	Terminal iG-iC	-0.11 ± 0.16	-0.19 ± 0.07
2,6-Diaminopurine(D)-U[c,d]					
5'AD3' 3'UU5'	-2.32 ± 0.17	N.A.	5'CD3' 3'GU5'	-3.57 ± 0.21	-2.72 ± 0.20
5'AU3' 3'UD5'	-1.85 ± 0.19	N.A.	5'CU3' 3'GD5'	-3.20 ± 0.13	-2.28 ± 0.22
5'UD3' 3'AU5'	-2.61 ± 0.18	N.A.	5'DD3' 3'UU5'	-3.46 ± 0.18	N.A.
5'UU3' 3'AD5'	-2.28 ± 0.16	N.A.	5'DU3' 3'UD5'	-3.01 ± 0.21	N.A.
5'GD3' 3'CU5'	-3.10 ± 0.17	-3.10 ± 0.21	5'UD3' 3'DU5'	-3.64 ± 0.27	N.A.

5'GU3' 3'CD5'	-2.80 ± 0.15	-2.62 ± 0.14	Terminal D-U	-0.02 ± 0.19	N.A.
Inosine(I)-C[c]					
5'AI3' 3'UC5'	-1.09 ± 0.14	N.A.	5'CI3' 3'GC5'	-1.98 ± 0.20	N.A.
5'AC3' 3'UI5'	-0.95 ± 0.21	N.A.	5'CC3' 3'GI5'	-2.24 ± 0.16	N.A.
5'UI3' 3'AC5'	-0.98 ± 0.25	N.A.	5'II3' 3'CC5'	-1.18 ± 0.29	N.A.
5'UC3' 3'AI5'	-1.06 ± 0.11	N.A.	5'IC3' 3'CI5'	-1.01 ± 0.23	N.A.
5'GI3' 3'CC5'	-2.07 ± 0.26	N.A.	5'CI3' 3'IC5'	-0.96 ± 0.31	N.A.
5'GC3' 3'CI5'	-1.96 ± 0.13	N.A.	Terminal I-C	0.88 ± 0.21	N.A.
Dangling ends[a,e]					
5'AA3' 3'U5'	-0.58 ± 0.27	-0.8	5'U3' 3'AA5'	-0.35 ± 0.12	-0.3
5'AU3' 3'U5'	-0.39 ± 0.31	-0.5	5'A3' 3'UA5'	-0.49 ± 0.16	-0.3
5'AC3' 3'U5'	-0.25 ± 0.18	-0.8	5'G3' 3'CA5'	-0.60 ± 0.16	-0.4
5'AG3' 3'U5'	-0.73 ± 0.30	-0.6	5'C3' 3'GA5'	-0.12 ± 0.09	-0.2
5'UA3' 3'A5'	-0.96 ± 0.27	-1.7	5'U3' 3'AU5'	-0.23 ± 0.10	-0.5
5'UU3' 3'A5'	-0.38 ± 0.18	-0.8	5'A3' 3'UU5'	-0.27 ± 0.13	-0.3
5'UC3' 3'A5'	-0.53 ± 0.23	-1.7	5'G3' 3'CU5'	-0.13 ± 0.14	-0.2
5'UG3' 3'A5'	-1.19 ± 0.28	-1.2	5'C3' 3'GU5'	-0.15 ± 0.10	-0.1
5'CA3' 3'G5'	-0.93 ± 0.30	-1.1	5'U3' 3'AC5'	-0.23 ± 0.10	-0.2
5'CU3' 3'G5'	-0.68 ± 0.25	-0.4	5'A3' 3'UC5'	-0.12 ± 0.18	-0.3
5'CC3' 3'G5'	-0.70 ± 0.18	-1.3	5'G3' 3'CC5'	-0.23 ± 0.15	0
5'CG3' 3'G5'	-1.50 ± 0.45	-0.6	5'C3' 3'GC5'	0.01 ± 0.12	0
5'GA3' 3'C5'	-0.56 ± 0.30	-0.7	5'U3' 3'AG5'	-0.26 ± 0.29	-0.3
5'GU3' 3'C5'	-0.36 ± 0.20	-0.1	5'A3' 3'UG5'	-0.23 ± 0.13	-0.1
5'GC3' 3'C5'	-0.08 ± 0.22	-0.7	5'G3' 3'CG5'	-0.16 ± 0.19	-0.2
5'GG3' 3'C5'	-1.00 ± 0.27	-0.1	5'C3' 3'GG5'	-0.17 ± 0.09	-0.2

All values in kcal/mol. N.A., not available.

[a] Used in training the parameters.

[b] Used in model testing.

[c] Blind predictions.

[d] Experimental values determined by a linear regression to helix folding free energies, assuming a zero terminal contribution. See Methods for details.

[e] No error estimates were given in the original paper.

Table S2. Systematic errors due to different system setups, sampling schemes, and parameters used in WHAM analysis.

	Original	No initial BP[1]	Initial BP = A/U[1]	Initial BP = GG/CC[1]	Reduced A-form range[2]	Alternative Sugar[3]	Modified bin-size and cutoff[4]
5'AC3' 3'UG5'	-1.79	-1.91	-1.62	-1.49	-1.70	-1.72	-1.79
5'AG3' 3'UC5'	-2.14	-2.17	-1.83	-1.92	-2.08	-2.31	-2.14
5'UC3' 3'AG5'	-2.10	-2.32	-1.94	-1.82	-2.00	-1.92	-2.10
5'UG3' 3'AC5'	-1.83	-1.46	-1.52	-1.58	-1.89	-1.91	-1.83
Error (kcal/mol)[5]		0.23	0.26	0.27	0.08	0.14	5.86×10^{-5}

[1] The initial base pairs for the simulation (see methods).
[2] Use ±40˚ as the angle constraint for A-form helix instead of ±60˚.
[3] Use alternative conformation for the sugar rings during sampling ("-rna::corrected_geo false" in Rosetta command line).
[4] Use different bin-size and high-score cutoff (0.05 and 1600) instead. The original value is 0.1 and 800.
[5] Error relative to the original.

Table S3. Statistical error on sampling computed based on bootstrapping (resampling of model ensembles with replacement).

Sequence	Simulated molecule	Mean free energy (kcal/mol)	Standard deviation (kcal/mol)
5'GGG3' 3'CCC5'	Complex	-5.54	0.0023
5'GU3' 3'CA5'	Complex	-5.12	0.0010
5'ACC3'	Single strand	10.86	0.0017
5'GU3'	Single strand	5.60	0.0014
5'GC3' 3'CGA5'	Complex, dangling end	-5.91	0.0010

Table S4. Systematic error from RECCES reweighting procedure.
NN predictions from five weight sets were computed by applying a fast reweighting procedure to models simulated with an arbitrarily chosen starting weight set, and compared to predictions from explicit simulation of the conformational ensemble with five arbitrarily chosen new weight sets.

Score term	Starting	Weight 1	Weight 2	Weight 3	Weight 4	Weight 5
fa_atr	0.37	0.58	0.28	0.10	0.33	0.93
fa_rep	0.2	0.1	1.25	0.72	0.32	1.15
fa_intra_rep	0.0035	0.0022	0.0046	0.0018	0.01	0
fa_stack	0.00001	0.0203	6.00	0.40	0.25	0
rna_torsion	9.5	4.90	5.79	5.65	3.29	6.78
hbond_sc	3.8	2.61	0.61	5.39	3.30	5.85
lk_nonpolar	0.51	0	0.61	1.83	0.038	0.92
geom_sol_fast	0.40	0	1.68	1.40	0	2.02
stack_elec	1.7	3.26	2.29	2.41	1.92	1.14
fa_elec_rna_phos_phos	1.6	6.77	1.34	0	4.47	0.72
Reweighting error (kcal/mol)	n.a.	0.48	0.26	0.25	0.22	0.23

Table S5. Mean and standard deviation of the weights of each score term from different optimization runs.

Score term	Description	Mean	Standard deviation
fa_atr	Lennard-Jones attraction	0.52	0.23 (45%)
fa_rep	Lennard-Jones repulsion (inter-residue)	0.27	0.51 (186%)
fa_intra_rep	Lennard-Jones repulsion (intra-residue)	0.067	0.038 (57%)
fa_stack	Extra Lennard-Jones attraction for stacking atoms	0.044	0.080 (181%)
rna_torsion	RNA torsional potential	9.3	4.7 (50%)
hbond_sc	Hydrogen bond	3.8	1.6 (42%)
lk_nonpolar	Lazaridis-Karplus solvation (Ref. (26))	1.1	0.94 (85%)
geom_sol_fast	Geometric solvation (Ref. (10))	0.83	1.2 (139%)
stack_elec	Electrostatics for stacking atoms	1.4	1.0 (74%)
fa_elec_rna_phos_phos	Electrostatics for backbone phosphates	9.6	7.6 (80%)

Table S6. Example component weights in RECCES collection of energy functions.

Score term	Structure prediction[1]	Best[2]	Weight 1	Weight 2	Weight 3	Weight 4
fa_atr	0.23	0.73	0.72	0.94	0.50	0.39
fa_rep	0.12	0.1	0.17	0.39	0.1	0.19
fa_intra_rep	0.0029	0.0071	0.01	0.0029	0.0042	0.0100
fa_stack	0	0	0	0	0	0.051
rna_torsion	0.1	4.26	4.30	3.46	8.45	18.64
hbond_sc	3.4	2.46	2.48	2.36	3.11	5.99
lk_nonpolar	0.32	0.25	1.66	0.82	0.55	3.63
geom_sol_fast	0.62	0	0	0	0.38	2.51
stack_elec	0	1.54	0.77	1.15	0.75	1.49
fa_elec_rna_phos_phos	1.05	4.54	0	0.73	13.6	20
Training error (kcal/mol)	1.23	0.35	0.38	0.41	0.36	0.38
Test error, GU (kcal/mol)	0.93	0.34	0.43	0.38	0.34	0.51
Test error, iGiC (kcal/mol)	1.79	1.06	0.96	1.10	1.00	1.03
Test error, iGiC with outlier exclusion (kcal/mol)	1.26	0.52	0.54	0.69	0.62	0.66

[1] The default score weights used in Rosetta structure prediction. In addition to the listed score terms, it also includes the terms *ch_bond* (0.42), *hbond_sr_bb_sc* (0.62), *hbond_lr_bb_sc* (3.4) (weights in parentheses). Here *ch_bond* is the hydrogen bond interaction between C-H to polar atom (O and N). The other two terms are hydrogen bonds between side-chain and backbone atoms. The terms do not contribute in a sequence-dependent manner to helix NN parameters and were omitted from the simulations herein. The training and test errors presented in the table are based on calculations optimizing a global scaling factor converting Rosetta energy units to k_BT (final value of 1 k_BT /0.218) over training data (canonical stacked pairs and dangling ends).

[2] The best weight with the lowest training error in the reweighting.

Table S7. Highly correlated score terms in the collection of RECCES score functions.
Each score function uses a different, locally optimized weight set fitted to canonical Watson-Crick stacked pairs and dangling ends for natural bases. Terms with absolute correlation > 0.1 are listed.

Score term 1	Score term 2	Correlation coefficient
fa_atr	*fa_rep*	0.31
fa_atr	*fa_stack*	-0.62
fa_atr	*rna_torsion*	-0.35
fa_atr	*stack_elec*	-0.11
fa_atr	*fa_elec_rna_phos_phos*	-0.17
fa_rep	*rna_torsion*	-0.12
fa_rep	*lk_nonpolar*	0.11
fa_rep	*stack_elec*	0.14
fa_rep	*fa_elec_rna_phos_phos*	-0.18
fa_stack	*rna_torsion*	-0.12
fa_stack	*lk_nonpolar*	-0.19
fa_stack	*stack_elec*	0.35
fa_stack	*fa_elec_rna_phos_phos*	-0.10
rna_torsion	*hbond_sc*	0.73
rna_torsion	*lk_nonpolar*	0.47
rna_torsion	*geom_sol_fast*	0.72
rna_torsion	*fa_elec_rna_phos_phos*	0.61
hbond_sc	*lk_nonpolar*	0.29
hbond_sc	*geom_sol_fast*	0.99
hbond_sc	*stack_elec*	0.32
hbond_sc	*fa_elec_rna_phos_phos*	0.63
lk_nonpolar	*geom_sol_fast*	0.32
lk_nonpolar	*fa_elec_rna_phos_phos*	0.26
geom_sol_fast	*stack_elec*	0.29
geom_sol_fast	*fa_elec_rna_phos_phos*	0.64

Table S8. Experimental measurements and free energy predictions of helices containing 2,6-diaminopurine (D)-uracil base pairs.

Sequence	Experiment	RECCES-Rosetta	Hydrogen-bond counting	Rosetta score
GDGCUC	-9.56 ± 0.37	-11.51 ± 0.38	-10.04 ± 0.24	-10.34 ± 0.24
CUGCDG	-9.06 ± 0.38	-12.45 ± 0.42	-11.11 ± 0.23	-10.61 ± 0.23
CDCGUG	-8.34 ± 0.11	-10.58 ± 0.44	-10.04 ± 0.24	-10.04 ± 0.24
DGCGCU	-9.16 ± 0.63	-11.14 ± 0.42	-10.84 ± 0.38	-10.46 ± 0.38
GDCGUC	-9.38 ± 0.35	-9.64 ± 0.4	-11.11 ± 0.23	-10.91 ± 0.23
DCCGGU	-9.67 ± 0.22	-10 ± 0.43	-10.5 ± 0.37	-10.62 ± 0.37
RMSE (kcal/mol)		2.02 (21.9%)	1.52 (16.5%)	1.34 (14.6%)

Predicted values are computed by linear combinations of the predicted NN parameters; errors for the predictions were estimated by error propagation.

Table S9. Nearest-neighbor parameters for 2,6-Diaminopurine/U containing stacked pairs assuming different terminal D-U contributions (in kcal/mol).

Terminal D-U	5'DG/3'UC	5'GD/3'CU	5'DC/3'UG	5'CD/3'GU	RMSE (vs. RECCES)
-0.2	-2.08	-3.30	-2.42	-2.92	0.68
-0.1	-2.18	-3.20	-2.52	-2.82	0.65
0	-2.28	-3.10	-2.62	-2.72	0.63
0.1	-2.38	-3.00	-2.72	-2.62	0.63
0.2	-2.48	-2.90	-2.82	-2.52	0.65
0.3	-2.58	-2.80	-2.92	-2.42	0.67
RECCES-Rosetta	-3.2	-3.1	-2.8	-3.57	n.a.

Supporting Figures

Figure S1. Correlation plots for the component score terms.
Each red cross represents a minimized score weight set from the RECCES-Rosetta energy function collection. The blue lines are linear regressions of the red crosses.

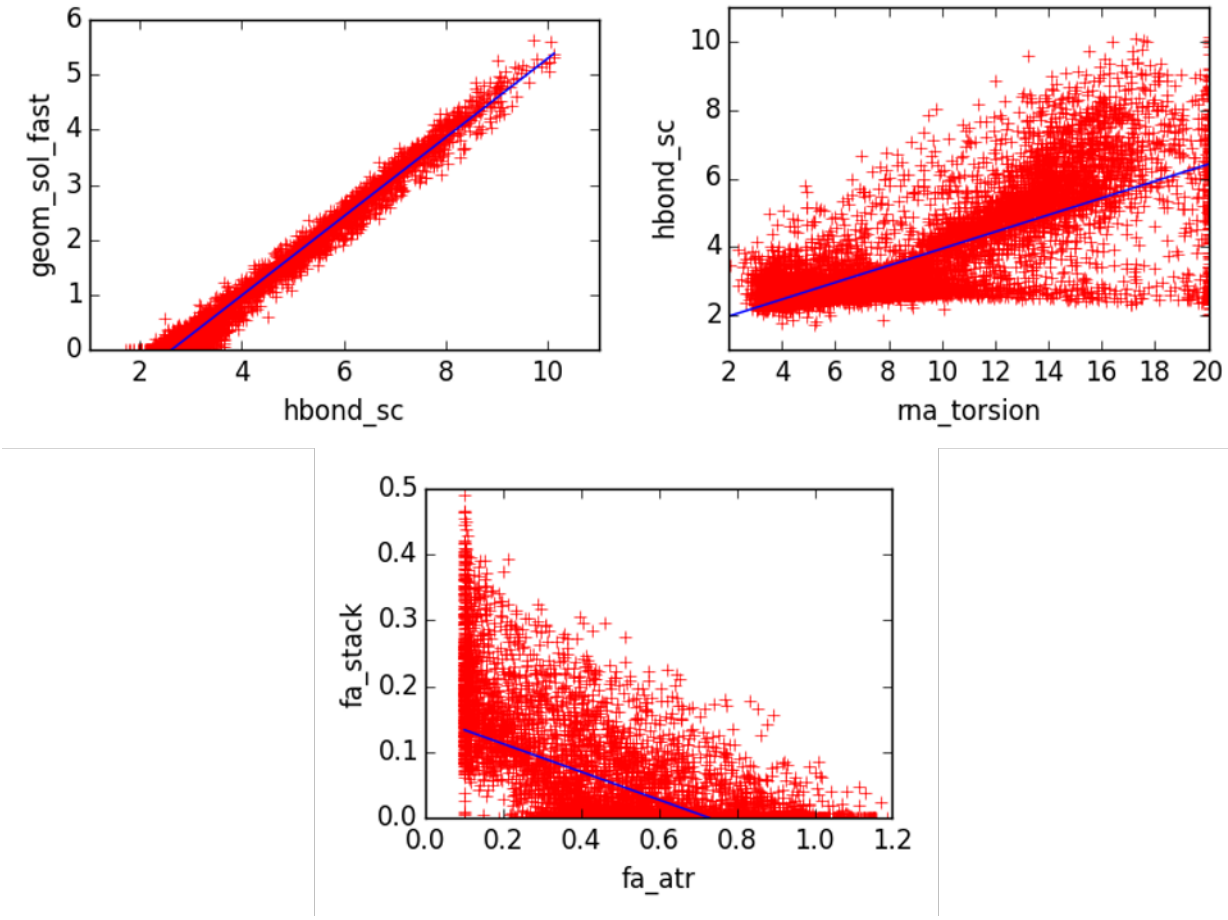
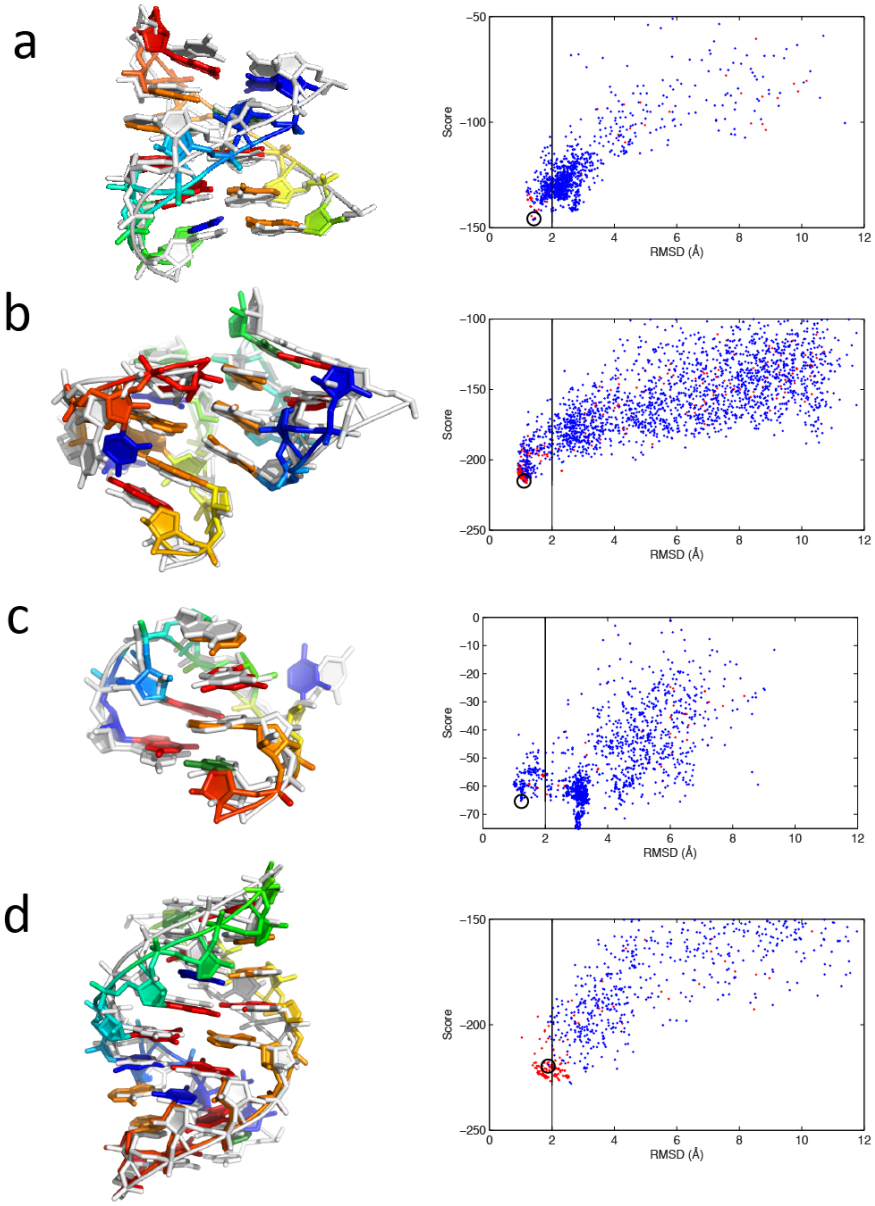


Figure S2. RECCES weights tested in structure prediction.

De novo 3D structure prediction (5000 fragment assembly steps, 2000 models each) repeated for all noncanonical motifs in the FARFAR benchmark (Das, Karanicolas, Baker, Nature Methods, 2010) confirms that the 'best' RECCES weight set (Table S5) recovers 12 of the 16 motifs previously recovered at better than 2 Å resolution. Examples include (a) domain IV from signal recognition particle RNA (1LNT) and (b) the tetraloop/receptor from the P4-P6 domain (2R8S). For the other 4 motifs, models better than 2 Å accuracy are still sampled, are stable to refinement in the RECCES energy function, and give energies similar to the lowest energy models. Examples include (c) GAGUA pentaloop from a SARS conserved domain (1XJR) and (d) the bacterial 5S ribosomal RNA loop E motif. Each panel gives single-model RECCES energies for *de novo* models (blue) and refined native models (red) vs. RMSD to crystallographic structure (right); and one 3D model (colored) overlaid on crystallographic structure (white) (on left). The scores computed in this figure are for single conformations; they might achieve higher accuracy through RECCES ensemble modeling for each conformation, whose computation may be possible with extensions of RECCES and would also allow inclusion of structure prediction during weight training.



References for Supporting Information Appendix

1. Serra MJ & Turner DH (1995) Predicting thermodynamic properties of RNA. *Meth. Enzymol.* 259:242-261.
2. Das R, Karanicolas J, & Baker D (2010) Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat Methods* 7(4):291-294.
3. Marinari E & Parisi G (1992) Simulated Tempering: A New Monte Carlo Scheme. *EPL* 19(6):451.
4. Huang X, Bowman GR, & Pande VS (2008) Convergence of folding free energy landscapes via application of enhanced sampling methods in a distributed computing environment. *The Journal of Chemical Physics* 128(20):205106.
5. Kumar S, Rosenberg J, Bouzida D, Swendsen R, & Kollman P (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* 13(8):1011-1021.
6. Chodera JD, Swope WC, Pitera JW, Seok C, & Dill KA (2006) Use of the Weighted Histogram Analysis Method for the Analysis of Simulated and Parallel Tempering Simulations. *J. Chem. Theory Comput.* 3(1):26-41.
7. Guerois R, Nielsen JE, & Serrano L (2002) Predicting Changes in the Stability of Proteins and Protein Complexes: A Study of More Than 1000 Mutations. *Journal of Molecular Biology* 320(2):369-387.
8. Leaver-Fay A*, et al.* (2013) Chapter Six - Scientific Benchmarks for Guiding Macromolecular Energy Function Improvement. *Meth. Enzymol.,* Methods in Protein Design, ed Keating AE (Academic Press), Vol 523, pp 109-143.
9. Park S & Pande VS (2007) Choosing weights for simulated tempering. *Phys. Rev. E* 76(1).
10. Das R, Karanicolas J, & Baker D (2010) Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat Meth* 7(4):291-294.
11. Morozov AV, Kortemme T, Tsemekhman K, & Baker D (2004) Close agreement between the orientation dependence of hydrogen bonds observed in protein structures and quantum mechanical calculations. *PNAS* 101(18):6946-6951.
12. Sripakdeevong P, Kladwang W, & Das R (2011) An enumerative stepwise ansatz enables atomic-accuracy RNA loop modeling. *Proceedings of the National Academy of Sciences* 108(51):20573-20578.
13. Chou F-C, Sripakdeevong P, Dibrov SM, Hermann T, & Das R (2013) Correcting pervasive errors in RNA crystallography through enumerative structure prediction. *Nat Meth* 10(1):74-76.
14. Sripakdeevong P*, et al.* (2014) Structure determination of noncanonical RNA motifs guided by 1H NMR chemical shifts. *Nat Meth* 11(4):413-416.
15. O'Meara MJ*, et al.* (2015) A Combined Covalent-Electrostatic Model of Hydrogen Bonding Improves Structure Prediction with Rosetta. *J Chem Theory Comput* 11(2):609-622.
16. Yildirim I & Turner DH (2005) RNA Challenges for Computational Chemists†. *Biochemistry* 44(40):13225-13234.
17. Yesselman JD, Price DJ, Knight JL, & Brooks CL (2012) MATCH: An atom-typing toolset for molecular mechanics force fields. *Journal of Computational Chemistry* 33(2):189-202.
18. Cantor CR & Tinoco Jr I (1965) Absorption and Optical Rotatory Dispersion of Seven Trinucleoside Diphosphates. *Journal of Molecular Biology* 13(1):65-77.
19. Murugaiah V (2011) Determination of Extinction Coefficient. *Handbook of Analysis of Oligonucleotides and Related Products*, ed Srivatsa G (CRC Press), pp 351-358.

20. Chen X, Kierzek R, & Turner DH (2001) Stability and Structure of RNA Duplexes Containing Isoguanosine and Isocytidine. *Journal of the American Chemical Society* 123(7):1267-1274.
21. Xia T*, et al.* (1998) Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson−Crick Base Pairs. *Biochemistry* 37(42):14719-14735.
22. Mathews DH, Sabina J, Zuker M, & Turner DH (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of Molecular Biology* 288(5):911-940.
23. Andronescu M, Condon A, Turner DH, & Mathews DH (2014) The Determination of RNA Folding Nearest-neighbor Parameters. *RNA Sequence, Structure, and Function: Computational and Bioinformatic Methods,* Methods in Molecular Biology, eds Gorodkin J & Ruzzo WL (Humana Press), pp 45-70.
24. Chen JL*, et al.* (2012) Testing the Nearest-neighbor Model for Canonical RNA Base Pairs: Revision of GU Parameters. *Biochemistry* 51(16):3508-3522.
25. Chen JL*, et al.* (2012) Testing the nearest-neighbor model for canonical RNA base pairs: revision of GU parameters. *Biochemistry* 51(16):3508-3522.
26. Lazaridis T & Karplus M (1999) Effective energy function for proteins in solution. *Proteins: Structure, Function, and Bioinformatics* 35(2):133-152.