Supplementary Information

for

RNA-PAIRS: RNA Probabilistic Assignment of Imino Resonance Shifts

Arash Bahrami[a], Lawrence J. Clos II[a], John L. Markley[a,b], Samuel E. Butcher[a,b], and Hamid R. Eghbalnia[c,1]

[a]National Magnetic Resonance Facility at Madison, [b]Biochemistry Department, University of Wisconsin-Madison, Madison, WI 53706, [c]Department of Molecular and Cellular Physiology, University of Cincinnati, Cincinnati, OH 45267

Author contributions: H.R.E, S.E.B, L.J.C., A.B., and J.L.M designed the research. L.J.C. and A.B contributed equally to various aspects of the algorithm development. A.B. implemented the software; L.J.C. collected, assigned, and simulated NMR data sets. L.J.C. and A.B. tested and validated PARS results. All authors contributed to the writing of the paper.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed: E-mail: eghbalhd@uc.edu.

Visual Presentation of Assignment and Base Pairing Results. The imino assignment and base pairing results for the RNA molecules with "direct" experimental data is visually coded in Figure S1.

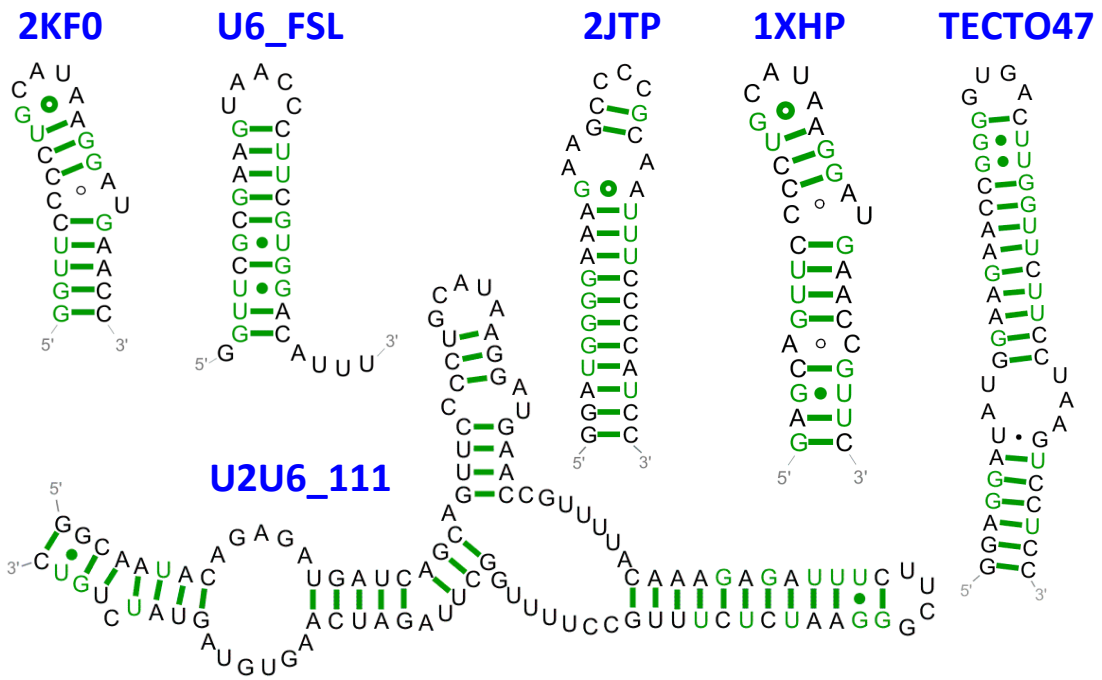

Fig S1. Green residue letters in the figure specify correct imino proton assignment within the best 3 choices. Green base-pair connectors (dashes for Watson-Crick, or circles for mismatch pairs) specify correct best choice base pair prediction.

Formal outline of the Algorithm (Figure 2 main text)

Inputs: NMR Experiments[1], Sequence

Optional Inputs: User-defined base pairing

P: Probability of states of the network signifying assignment/secondary structure

P_{new}/$P_{previous}$ are respectively the current probabilities and previous probabilities

Begin RNA-PAIRS

Begin Process Input

Predict secondary structure ensemble

Align peaks and cluster

Perform NOESY cluster analysis

Load resonance distributions derived from BMRB

End Process Input

$P_{previous}$ = F(initial energy)[2]

P_{new} = 0;

Begin Loop While ($P_{new} \neq P_{previous}$)[3]

Update P_{new}

End Loop

End

1 Currently supported experiments: 1H-1H 2D NOESY, 1H-15N 2D HMQC

2 F(initial energy) is calculated according to formula 2.

3 (P_{new} = $P_{previous}$) if probabilities have not changed more than a small value

4 Update P_{new} is performed according to equation 2 (or 3 as appropriate)

Secondary structure ensemble is based on equation 1 in the main text

Equations 2 and 3 are restated below (from the main text).

$$E_S = \sum_i U_i(\lambda(v_i)) + \sum_{i,j} U_{ij}(\lambda(v_i), \lambda(v_j)) \quad (2)$$

$$E_S = \sum_i U_i(\gamma(\rho_i)) + \sum_{i,j} U_{ij}(\gamma(\rho_i), \gamma(\rho_j)) + \sum_{i,j,k} U_{ijk}(\gamma(\rho_i), \gamma(\rho_j), \gamma(\rho_k)) \quad (3)$$