Supplementary Information:

Ensemble Analysis of Primary miRNA Structure Reveals an Extensive Capacity to Deform near the Drosha Cleavage Site

Kaycee A. Quarles [1], Debashish Sahu [1], Mallory A. Havens [2], Ellen R. Forsyth [1], Christopher Wostenberg [1], Michelle L. Hastings [2], and Scott A. Showalter[1]*

[1]Department of Chemistry and Center for RNA Molecular Biology, The Pennsylvania State University, 104 Chemistry Building, University Park, PA 16802, USA.

[2]Department of Cell Biology and Anatomy, Chicago Medical School, Rosalind Franklin University of Medicine and Science, North Chicago, IL 60064, USA.

Figures S1-S6.

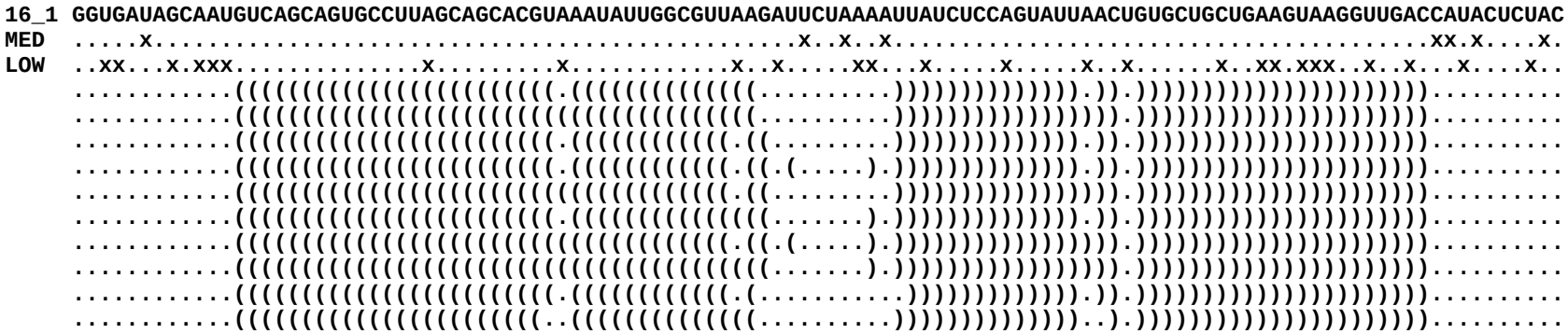

Figure S1. The ten most probable secondary structure assignments predicted by MC-Fold with SHAPE constraints for pri-mir-16-1. The nucleotide sequence of pri-mir-16-1 is displayed above the MC-Fold output. Each secondary structure prediction is displayed in dot-bracket format, with those at the top of the figure being more probable than those closer to the bottom. The constraint mask used in the generation of these secondary structure predictions is provided below the nucleotide sequence.

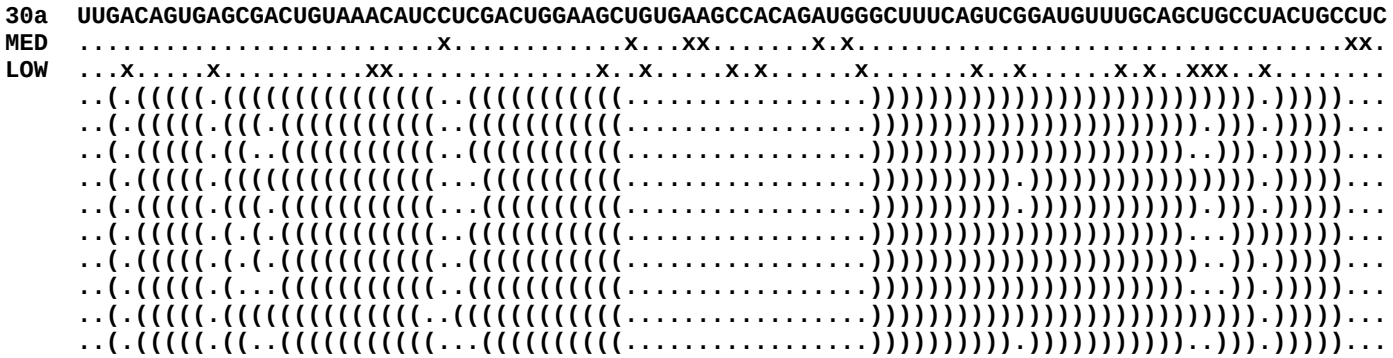

Figure S2. The ten most probable secondary structure assignments predicted by MC-Fold with SHAPE constraints for pri-mir-30a. The nucleotide sequence of pri-mir-30a is displayed above the MC-Fold output. Each secondary structure prediction is displayed in dot-bracket format, with those at the top of the figure being more probable than those closer to the bottom. The constraint mask used in the generation of these secondary structure predictions is provided below the nucleotide sequence.

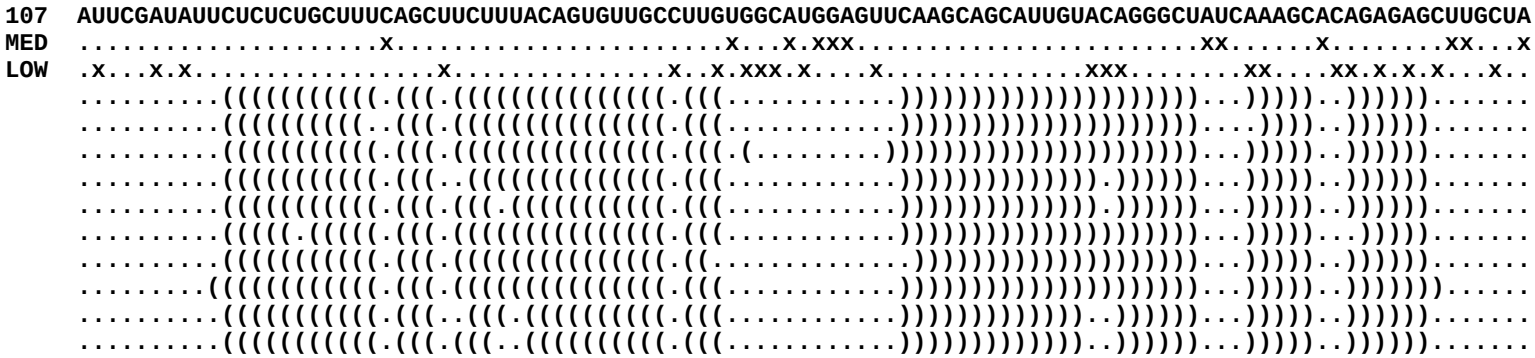

Figure S3. The ten most probable secondary structure assignments predicted by MC-Fold with SHAPE constraints for pri-mir-107. The nucleotide sequence of pri-mir-107 is displayed above the MC-Fold output. Each secondary structure prediction is displayed in dot-bracket format, with those at the top of the figure being more probable than those closer to the bottom. The constraint mask used in the generation of these secondary structure predictions is provided below the nucleotide sequence.

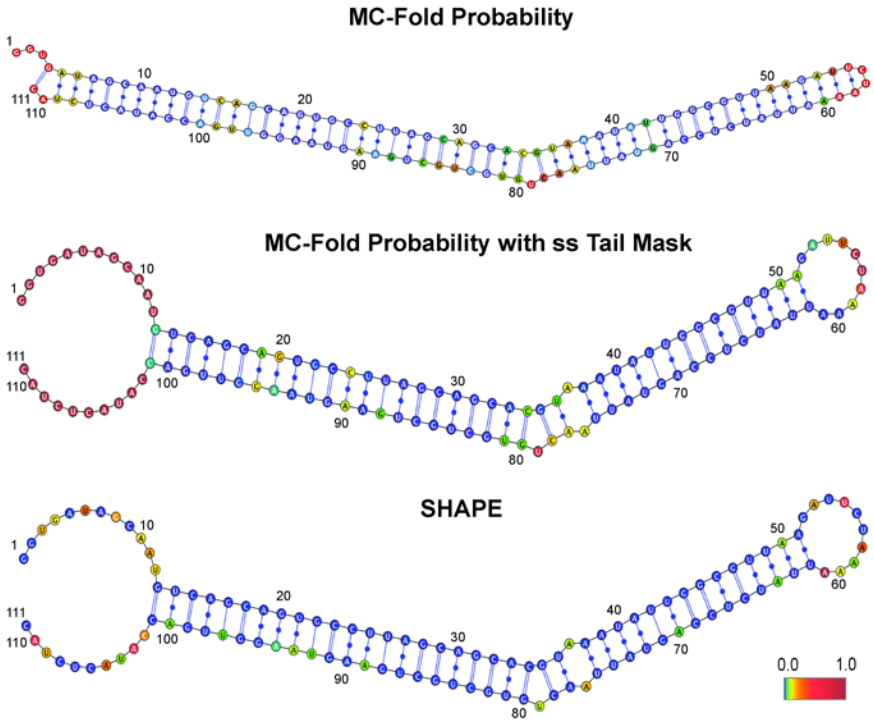

Figure S4. Inclusion of SHAPE data significantly influences the most probable secondary structure reported by MC-Fold for pri-mir-16-1. The top model is the most probable secondary structure of pri-mir-16-1 predicted by MC-Fold using no constraints and is color coded according to the probability of being single-stranded predicted by MC-Fold alone. In the middle and bottom panels, the most probable secondary structure from the MC-Fold simulations conducted with SHAPE data added as a low resolution constraint is used (and is identical to the secondary structure depicted in Figure 1 of the main text). In contrast to Figure 1, the color coding reflects the single-stranded probability calculated by MC-Fold without SHAPE constraints as shown for the top model, but with the tails constrained to be single-stranded (middle); and the normalized SHAPE intensity alone, without re-weighting by the MC-Fold procedure (bottom). Note that the base-pairing register changes from the top model to the bottom two models when the tails are constrained to be single-stranded, as predicted by SHAPE.

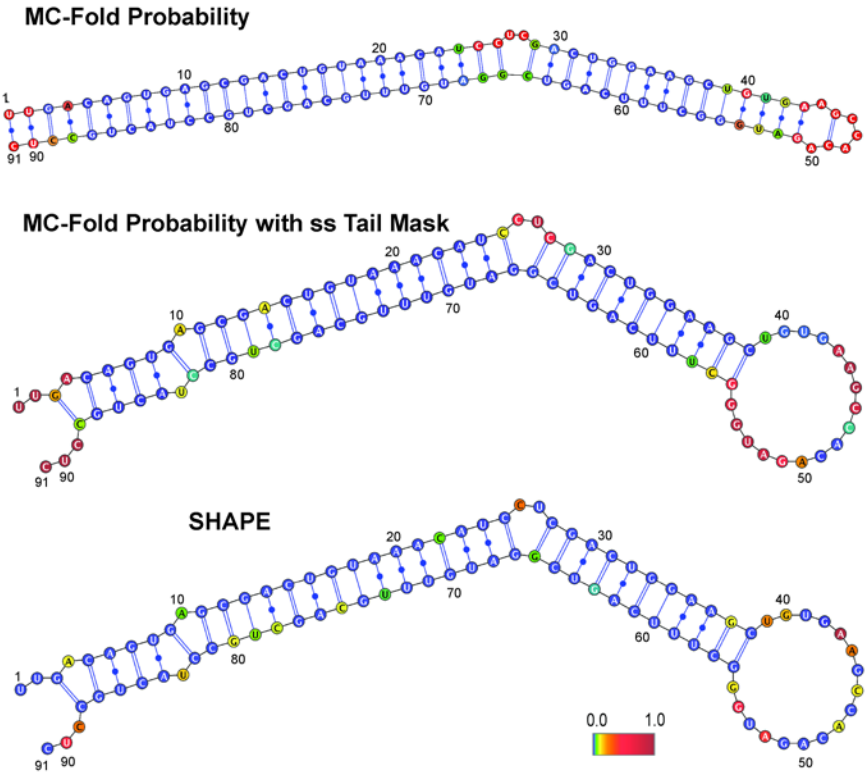

Figure S5. Inclusion of SHAPE data significantly influences the most probable secondary structure reported by MC-Fold for pri-mir-30a. The top model is the most probable secondary structure of pri-mir-30a predicted by MC-Fold using no constraints and is color coded according to the probability of being single-stranded predicted by MC-Fold alone. In the middle and bottom panels, the most probable secondary structure from the MC-Fold simulations conducted with SHAPE data added as a low resolution constraint is used (and is identical to the secondary structure depicted in Figure 1 of the main text). In contrast to Figure 1, the color coding reflects the single-stranded probability calculated by MC-Fold without SHAPE constraints as shown for the top model, but with the tails constrained to be single-stranded (middle); and the normalized SHAPE intensity alone, without re-weighting by the MC-Fold procedure (bottom). Note that the base-pairing register changes from the top model to the bottom two models when the tails are constrained to be single-stranded, as predicted by SHAPE.

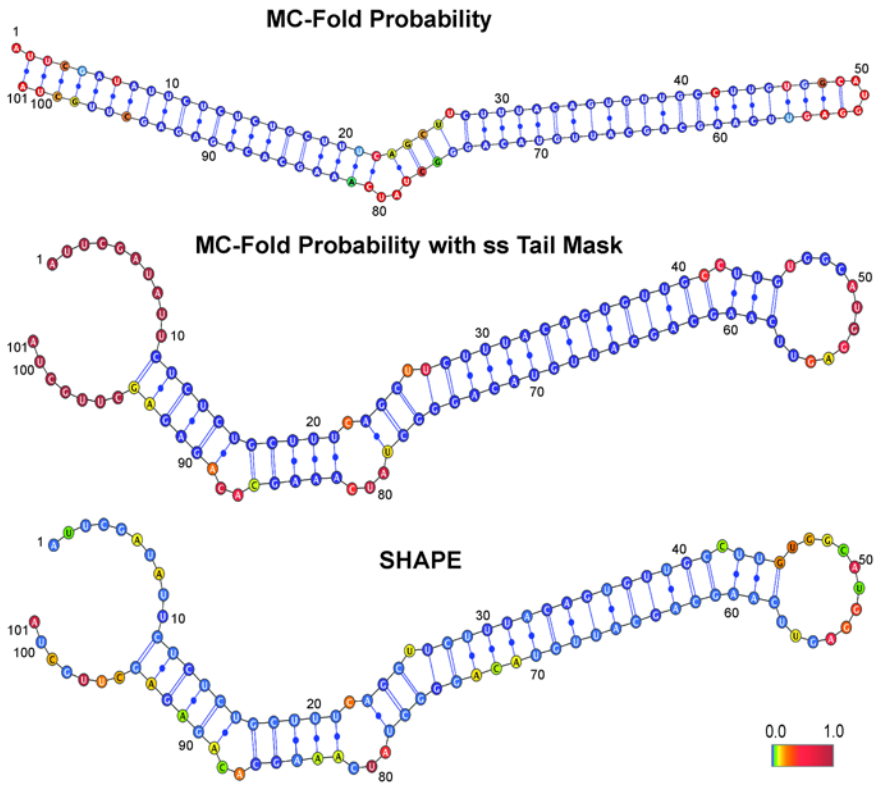

Figure S6. Inclusion of SHAPE data significantly influences the most probable secondary structure reported by MC-Fold for pri-mir-107. The top model is the most probable secondary structure of pri-mir-107 predicted by MC-Fold using no constraints and is color coded according to the probability of being single-stranded predicted by MC-Fold alone. In the middle and bottom panels, the most probable secondary structure from the MC-Fold simulations conducted with SHAPE data added as a low resolution constraint is used (and is identical to the secondary structure depicted in Figure 1 of the main text). In contrast to Figure 1, the color coding reflects the single-stranded probability calculated by MC-Fold without SHAPE constraints as shown for the top model, but with the tails constrained to be single-stranded (middle); and the normalized SHAPE intensity alone, without re-weighting by the MC-Fold procedure (bottom). Note that the base-pairing register changes from the top model to the bottom two models when the tails are constrained to be single-stranded, as predicted by SHAPE.