Figure S1 (see next page for figure legend)

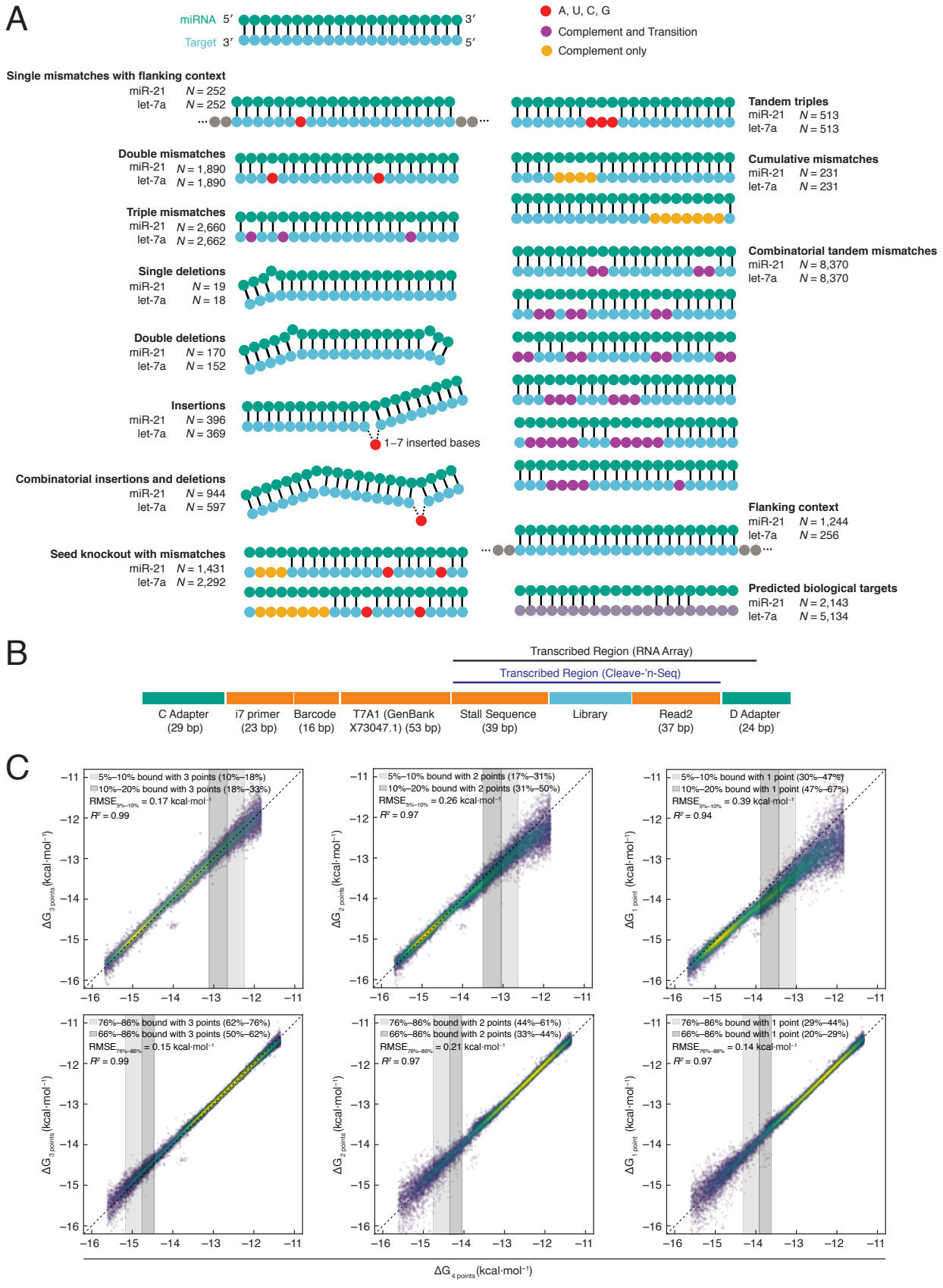

Figure S1. AGO2 Library Design and Construction, Related to Figure 1
(A) Schematic of designed targets included in miR-21 and let-7a libraries.
(B) Construct used in array experiments. Transcribed region is indicated above the schematic.
(C) By constraining the maximum fluorescence values to an empirical distribution defined by high affinity variants, we were able to estimate K_D values between 10 pM and 10 nM. These limits correspond to variants with only 5% bound at our highest concentration (10 nM K_D) and variants with 86% bound at our lowest concentration (10 pM K_D). To test whether K_D estimates near the detection limit were robust, we progressively removed the highest and lowest concentrations from our K_D curves and found that even when K_D values were defined by a single point, they were fit to nearly identical K_D values as were originally determined using all concentrations. Top row: comparisons of ΔG values obtained from 4-point curves (*x*-axis) and ΔG values computed from 63, 125, and 250 pM (left), 63 and 125 pM (middle), and only 63 pM (right). The set of variants for which the binding at the highest concentration is 5–10% or 10–20% in the subsampled data is shaded in gray (i.e., the subsampled K_D for these variants is defined by a single point with a low fraction bound). The corresponding fraction bound for these variants at the 500 pM point in the 4-point curves is shown in parentheses in the legend. The RMSE reported reflects the RMSE for the variants that are 5–10% bound at the highest concentration in the subsampled fit. Bottom row: removing low concentration points. Comparisons of ΔG values obtained from 4-point curves (*x*-axis) and ΔG values computed from 125, 250, and 500 pM (left), 250 and 500 pM (middle), and only 500 pM (right). The set of variants for which the binding at the lowest concentration is 76–86% or 70–76% in the subsampled data is shaded in gray (i.e., the subsampled K_D for these variants is defined by a single point with <86% bound). The corresponding fraction bound for these variants at the 63 pM point in the 4-point curves is shown in parentheses in the legend. The RMSE reported reflects the RMSE for the variants that are 76–86% bound at the lowest concentration in the subsampled fit.

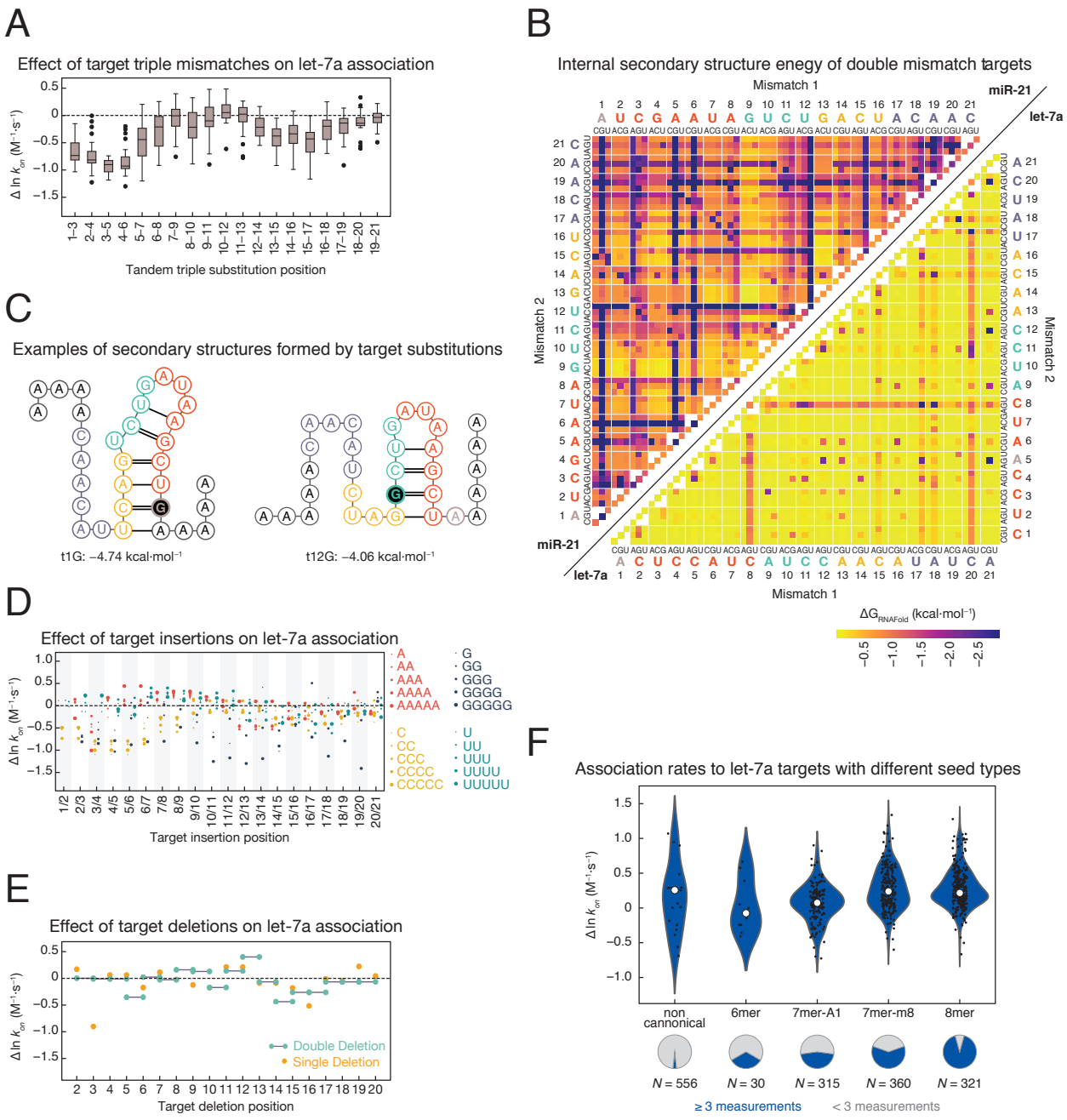

Figure S2. Factors Contributing to AGO2 Association Kinetics, Related to Figures 2

(A) Association rates for let-7a RISC binding to targets with consecutive triple mismatches.

(B) Internal structure for double mismatched target sequences of miR-21 and let-7a as predicted by RNAfold.

(C) Predicted secondary structures for miR-21 target containing t1G and t12G substitutions. The ensemble free energy predicted by RNAfold is shown below each structure.

(D) Association rates for let-7a targets containing 1–5 insertions of each base.

(E) Association kinetics for let-7a RISC to targets containing single and double target deletions.

(F) Association rates measured for let-7a RISC binding to predicted targets grouped by the site type. Predicted targets are filtered to remove those with stable predicted structure ($\Delta G_{RNAfold}$ <−2 kcal·mol^{-1}).

Figure S3 (see next page for figure legend)

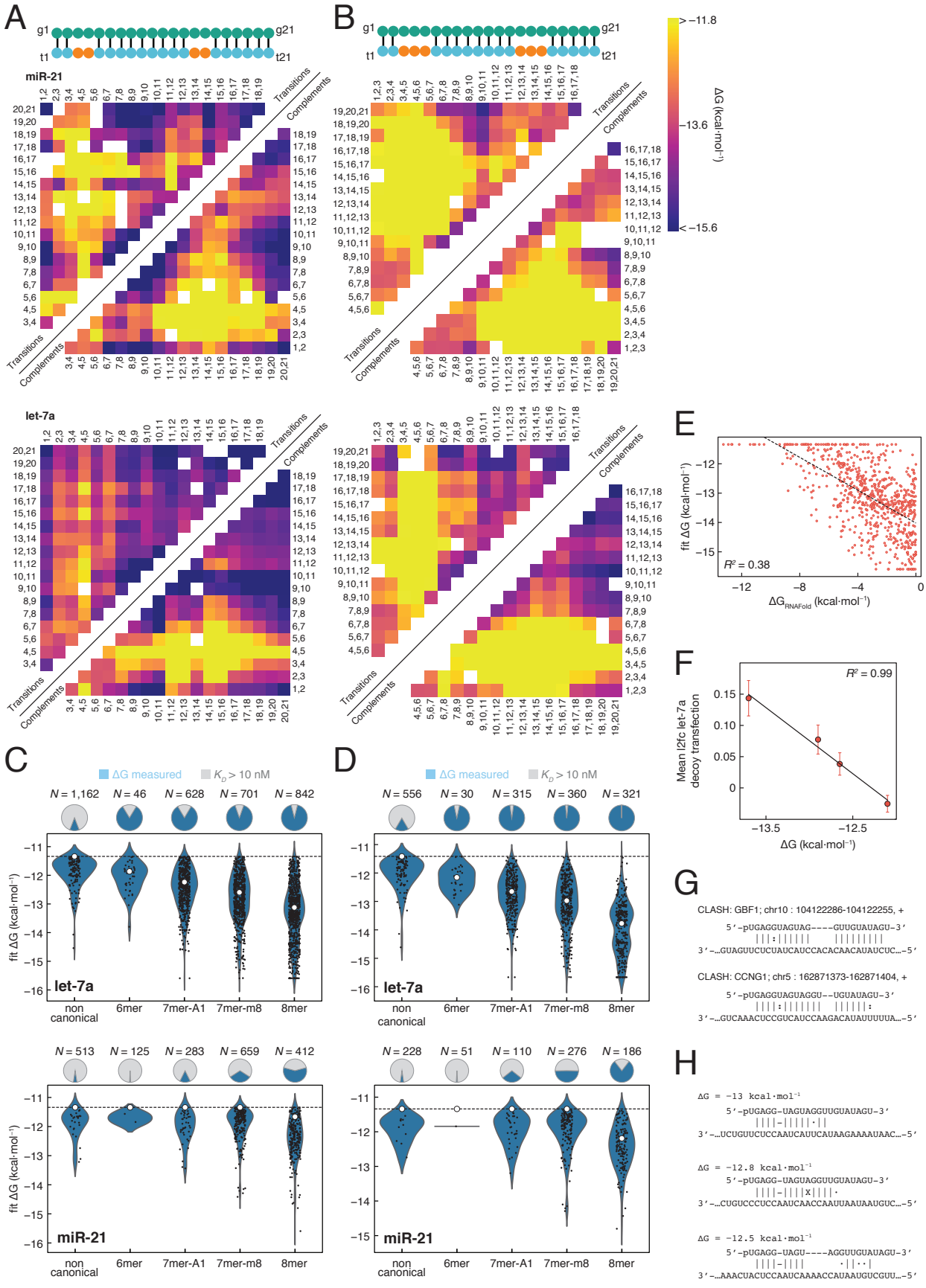

Figure S3. Binding Affinity of AGO2 to Predicted and Designed Targets, Related to Figure 3

(A) RISC binding affinity to miR-21 (top) and let-7a (bottom) targets containing two by two substitutions. A schematic of the targets is shown on top of the panel. Transition (A↔G and C↔U) substitutions are above the diagonal and complement (C↔G and A↔U) substitutions are shown below.

(B) RISC binding affinity to miR-21 (top) and let-7a (bottom) targets containing three by three substitutions. A schematic of the targets is shown on top of the panel.

(C) Affinities measured for let-7a (top) and miR-21 (bottom) loaded RISC binding to all predicted targets grouped by the site type. Dashed lines represent the minimum binding affinity that we could resolve experimentally.

(D) Affinities measured for let-7a (top) and miR-21 (bottom) predicted targets filtered to keep those that are predicted to form less internal structure ($\Delta G_{RNAfold} < -2$ kcal·mol^{-1}).

(E) Relationship between measured target affinity and RNAfold predicted internal structure.

(F) Relationship between binding affinity and mean (± SEM) change in target abundance following transfection of a let-7a decoy for canonical seed types. l2fc: $\log_2$ fold change.

(G) Highest affinity noncanonical targets measured for let-7a RISC, related to panels C and D.

(H) Highest affinity CUCCAAU nucleation bulge targets measured for let-7a RISC.

Figure S4 (see next page for figure legend)

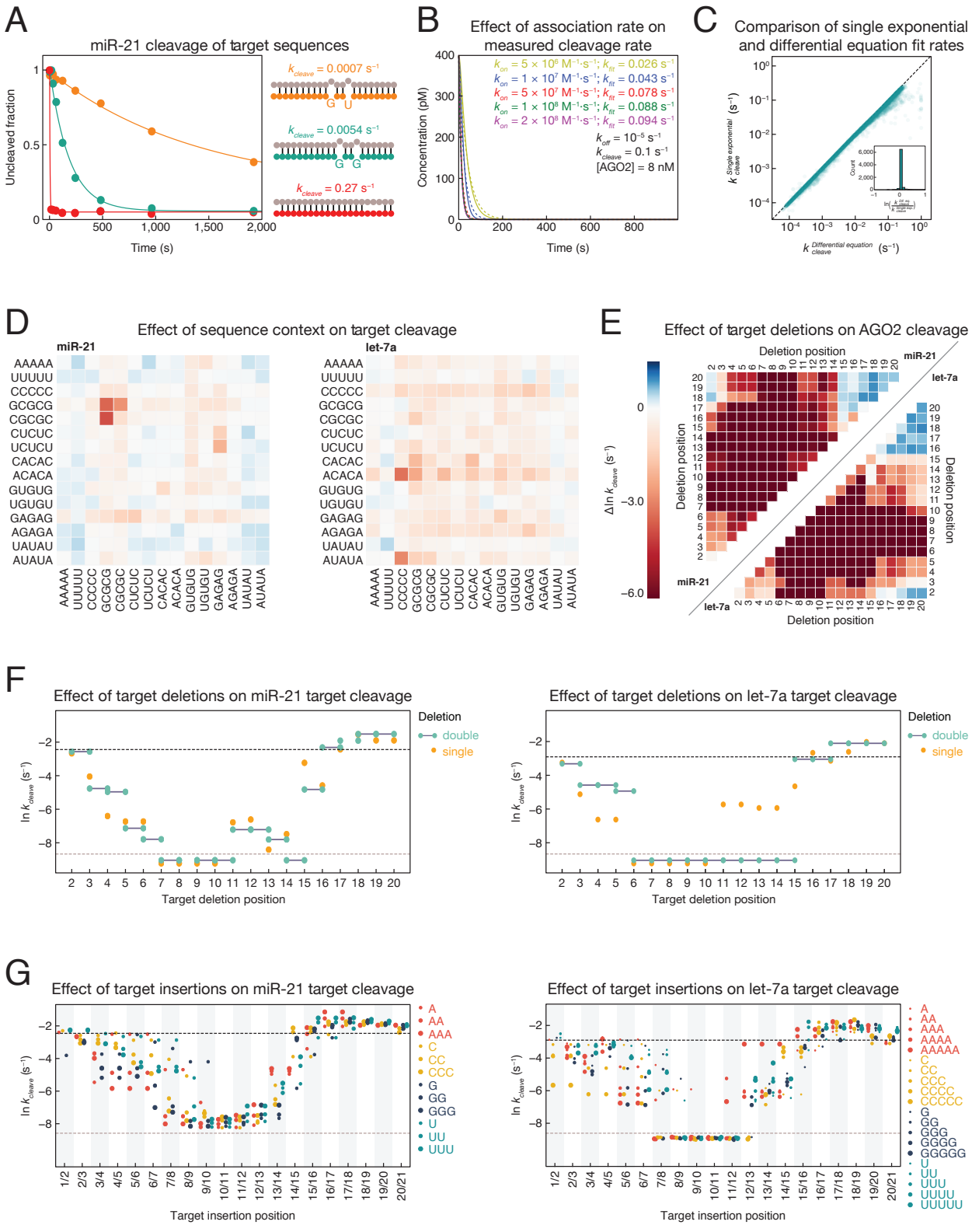

Figure S4. AGO2 Cleavage Kinetics, Related to Figure 4

(A) Fraction of uncleaved RNA for 3 target sequences as a function of time obtained from the RISC-Cleave ‘n Seq experiments. The lines represent the fit to a single exponential.

(B) Simulations of cleavage rates for different association rates. Solid lines represent simulated data and dotted lines represent the single exponential fit to the simulated data.

(C) Simulated relationship between fit cleavage rate and true cleavage rate for different association rates.

(D) Cleavage rates measured for perfectly complementary targets with different 5 nucleotide flanking sequences. Color bar as in (E).

(E) Cleavage rates of target sequences containing double deletions for miR-21 (top left) and let-7a (lower right).

(F) Cleavage rates of miR-21 (left) and let-7a (right) targets containing single and double consecutive deletions.

(G) Cleavage rates of miR-21 (left) and let-7a (right) targets containing multiple nucleotides inserted into the target sequence.

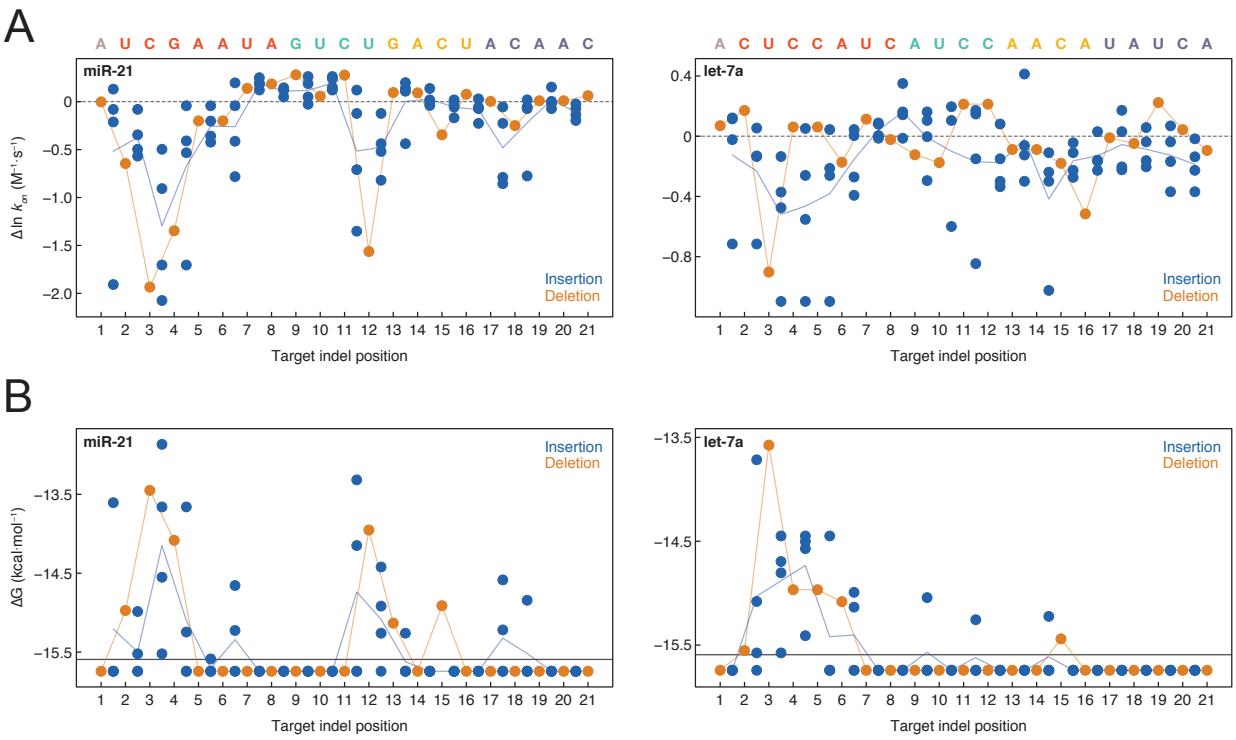

Figure S5. Insertions and Deletions Have Similar Effects on RISC Binding, Related to Figure 5

(A) Association rates for miR-21 (left) and let-7a (right) single insertions (blue dots) and single deletions (orange dots). Insertions and deletions that correspond to multiple target positions, such as insertion of the same base either before or after a base, or deletion of a repeated base, are plotted in all possible target positions. The fully complementary sequence is shown on top of each plot, and its association rate is indicated by the dotted line. Orange line, all single deletions; blue line, mean of the single insertions.

(B) Binding affinity for miR-21 (left) and let-7a (right) single insertions and single deletions. Points below the solid line were too high affinity to be accurately measured (K_D <10 pM).

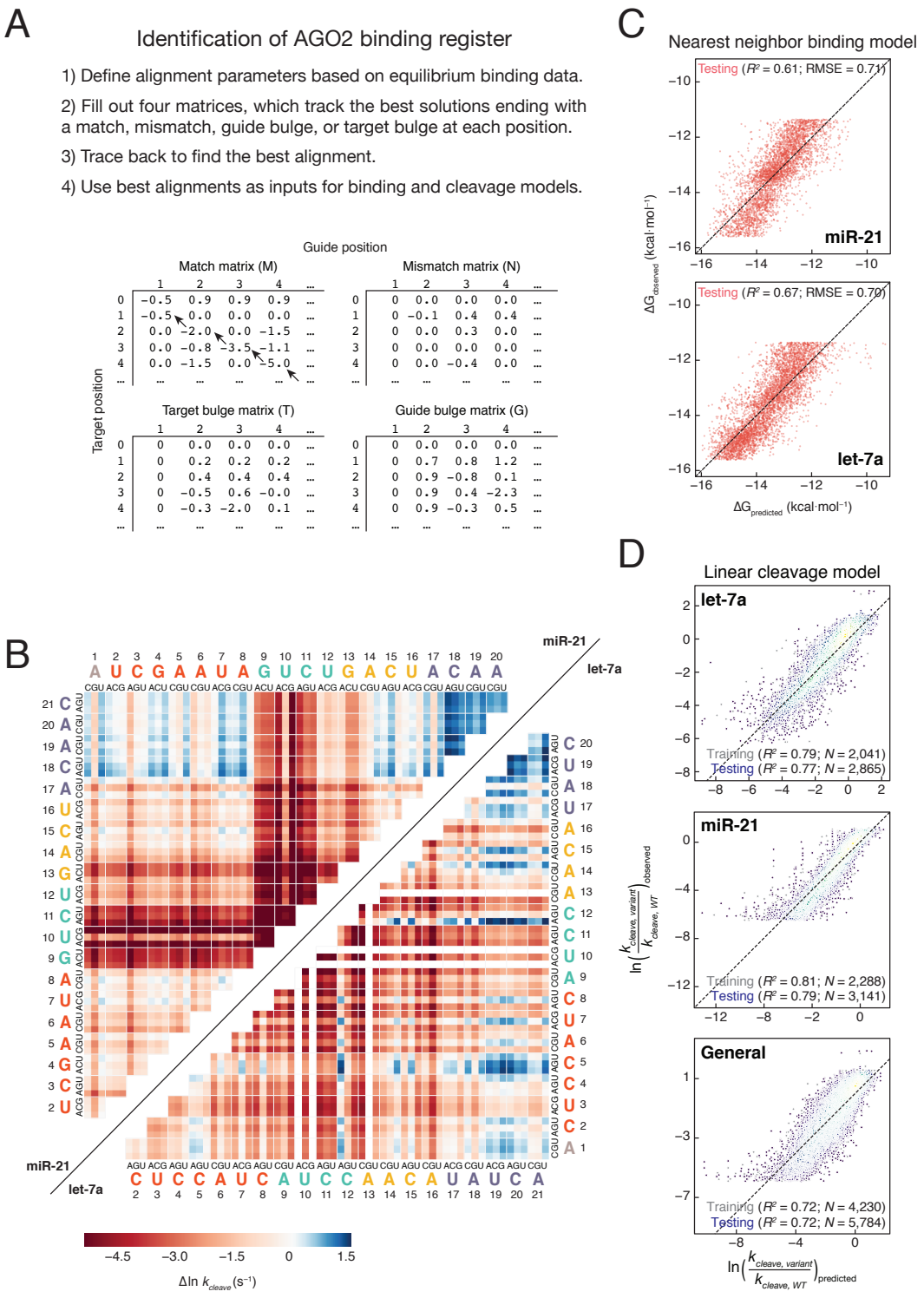

Figure S6. Models of AGO2 Binding Affinity and Cleavage Kinetics, Related to Figure 6
(A) Overview of dynamic programming alignment algorithm used to align sequences.
(B) Predicted double substitution cleavage rates from single substitution cleavage rates.
(C) Comparison of binding affinity predicted by miR-21 (top) and let-7a (bottom) specific nearest-neighbor models to observed binding affinities.
(D) Performance of cleavage model when trained and tested on a random split of the data for let-7a cleavage model (top), miR-21 cleavage model (middle), and general cleavage model (bottom).

Figure S7 (see next page for figure legend)

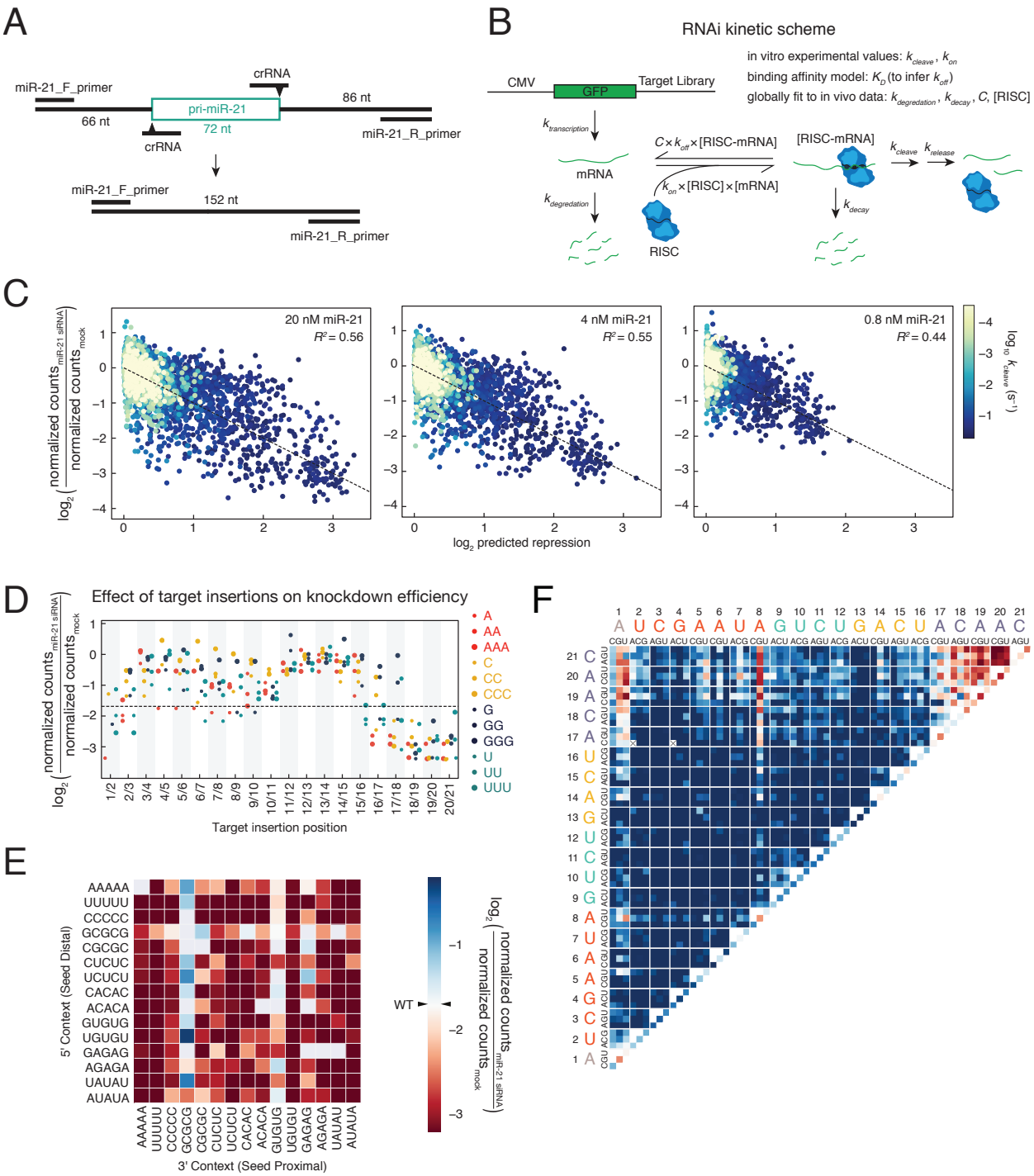

Figure S7. Additional in Cell Target Knockdown Analysis, Related to Figure 7

(A) CRISPR-Cas9 editing design to generate the miR-21 knockout cell line. Guide RNAs were designed to cut both sides of the primary miR-21 hairpin. Successful editing was confirmed by PCR of the edited region and confirming loss of the 72-nt hairpin sequence, as well as by a TaqMan assay specific for mature miR-21.

(B) Kinetic model of in cell RISC activity. *C*, the dissociation rate scaling factor.

(C) Additional demonstrations of kinetic biochemical model of miR-21 knockdown. The miR-21 transfection concentration is indicated in each panel. Individual targets are colored according to their measured cleavage rates. The dotted lines each have a slope of –1 and an intercept of 0.

(D) Knockdown efficiencies for miR-21 targets containing 1–3 insertions of each base.

(E) Knockdown of perfectly complementary targets in varying 5′ and 3′ sequence contexts. Color bar is centered on the knockdown of a perfectly complementary target in the poly(A) sequence context.

(F) Knockdown of all miR-21 double mismatched targets. Color bar as in (E).