Supporting Information

Selective nucleic acid capture with shielded covalent probes

Jeffrey R. Vieregg, Hosea M. Nelson, Brian M. Stoltz, and Niles A. Pierce

Contents

S1 Target and probe sequences

The sequences for all DNA and RNA targets, probes, and helper strands are shown in Tables S1–S4. The family of targets used for the experiments shown in Figures 3–6 was designed to be free of secondary structure using NUPACK.[1] Once the targets were designed, NUPACK was used to design hairpin loop sequences so as to minimize the ensemble defect with respect to the desired hairpin structure.[2] The FP target sequences (Figure 7) were chosen because they have the least secondary structure (as measured by the calculated ensemble free energy for 30mer DNA target sequences, considering dimers and pseudoknots) of the the six regions of the DsRed2 and mCherry sequences that differ by single 2-nt substitutions when aligned using Clustal W.[3] FP probe design is discussed in Section S7.

Table S1. Target sequences. Mismatches in orange using IUB nucleotide codes (N = A, C, G, or T/U; D = A, G or T/U; H = A, C or T/U; V = A, C or G). Crosslink partner in blue. 6-FAM represents 6-carboxyfluorescein.

Material	Name	Sequence (5' to 3')	Length (nt)	Figures
DNA	C	Cy5-CTAAGCCATGTCAACCGTTCCTGCGAGTA	29	3, 6a, S4–S11, S13, S21, S23, S34
DNA	M1	Cy5-CTAAGCCATGTATACCGTTCCTGCGAGTA	29	3, 6a, S10, S13, S14, S22, S24, S34
DNA	M2	Cy5-CTAAGCCATGTCAACCGAGCCTGCGAGTA	29	3, S13, S14
DNA	M3	Cy5-CTAAGCCATGTCAACCGTTCCTGGAAGTA	29	3, S11, S13, S14
RNA	C	Cy5-CUAAGCCAUGUCAACCGUUCCUGCGAGUA	29	3, 4, S11, S13, S15, S16, S29–S31
RNA	M1	Cy5-CUAAGCCAUGUAUACCGUUCCUGCGAGUA	29	3, S13
RNA	M2	Cy5-CUAAGCCAUGUCAACCGAGCCUGCGAGUA	29	3, S13
RNA	M3	Cy5-CUAAGCCAUGUCAACCGUUCCUGGAAGUA	29	3, S11, S13, S29–S31
DNA	C	Cy5-CATGTCAACCGTTCCTGCGAGTA	23	1, 4, 6b, 6c, S15, S25–S28
DNA	S1	Cy5-CATGTAAACCGTTCCTGCGAGTA	23	4, 6b, 6c, S15, S25, S26
DNA	S2	Cy5-CATGTTAACCGTTCCTGCGAGTA	23	4, 6b, 6c, S15, S25, S26
DNA	S3	Cy5-CATGTCTACCGTTCCTGCGAGTA	23	1, 4, 6b, 6c, S15, S25, S26
DNA	S4	Cy5-CATGTCAACCGATCCTGCGAGTA	23	4, 6b, 6c, S15, S25, S26
DNA	S5	Cy5-CATGTCAACCGTACCTGCGAGTA	23	4, 6b, 6c, S15, S25, S26
DNA	S6	Cy5-CATGTCAACCGTTGCTGCGAGTA	23	4, 6b, 6c, S15, S25, S26
RNA	S1	Cy5-CUAAGCCAUGUAAACCGUUCCUGCGAGUA	29	4, S15, S16
RNA	S2	Cy5-CUAAGCCAUGUUAACCGUUCCUGCGAGUA	29	4, S15, S16
RNA	S3	Cy5-CUAAGCCAUGUCAACCGAUCCUGCGAGUA	29	4, S15, S16
RNA	S4	Cy5-CUAAGCCAUGUCAACCGUACCUGCGAGUA	29	4, S15, S16
RNA	S5	Cy5-CUAAGCCAUGUCAGCCGUUCCUGCGAGUA	29	4, S15, S16
RNA	S6	Cy5-CUAAGCCAUGUCAACCGUUCCUGCGGGUA	29	4, S15, S16
RNA	S7	Cy5-CUAAGCCAUGUCAACCCUUCCUGCGAGUA	29	4, S15, S16
DNA	C	6-FAM-CTAAGCCATGTCAACCGTTCCTGCGAGTA	29	5, S17
DNA	P1	Cy5-CTAAGCCAVNTCAACCGTTCCTGCGAGTA	29	5, S17
DNA	P2	Cy5-CTAAGCCATHNCAACCGTTCCTGCGAGTA	29	5, S17
DNA	P3	Cy5-CTAAGCCATGVNAACCGTTCCTGCGAGTA	29	5, S17
DNA	P4	Cy5-CTAAGCCATGTCAACDNTTCCTGCGAGTA	29	5, S17
DNA	P5	Cy5-CTAAGCCATGTCAACCHNTCCTGCGAGTA	29	5, S17
DNA	P6	Cy5-CTAAGCCATGTCAACCGVNCCTGCGAGTA	29	5, S17
DNA	P7	Cy5-CTAAGCCATGTCAACCCTTCCTGCHNGTA	29	5, S17
DNA	P8	Cy5-CTAAGCCATGTCAACCCTNDCTGCGAGTA	29	5, S17
RNA	C	6-FAM-CUAAGCCAUGUCAACCGUUCCUGCGAGUA	29	5, S18
RNA	P1	Cy5-CUAAGCCAVNUCAACCGUUCCUGCGAGUA	29	5, S18
RNA	P2	Cy5-CUAAGCCAUHNCAACCGUUCCUGCGAGUA	29	5, S18
RNA	P3	Cy5-CUAAGCCAUGVNAACCGUUCCUGCGAGUA	29	5, S18
RNA	P4	Cy5-CUAAGCCAUGUCAACDNUUCCUGCGAGUA	29	5, S18
RNA	P5	Cy5-CUAAGCCAUGUCAACCHNUCCUGCGAGUA	29	5, S18
RNA	P6	Cy5-CUAAGCCAUGUCAACCGVNCCUGCGAGUA	29	5, S18
RNA	P7	Cy5-CUAAGCCAUGUCAACCCUUCCUGCHNGUA	29	5, S18
RNA	P8	Cy5-CUAAGCCAUGUCAACCCUNDCUGCGAGUA	29	5, S18
DNA	M4	Cy5-CATGTCAACCAATCCTGCGAGTA	23	1, 6c, S27, S28
DNA	M5	Cy5-CATGTCAACAATTCCTGCGAGTA	23	6c, S27, S28
DNA	M6	Cy5-CATGTCAAGAGTTCCTGCGAGTA	23	6c, S27, S28
DNA	M7	Cy5-CATGTCATGCGTTCCTGCGAGTA	23	6c, S27, S28
DNA	M8	Cy5-CATGTCTTCCGTTCCTGCGAGTA	23	6c, S27, S28
DNA	DsRed2	6-FAM-ACCGTGAAGCTGAAGGTGACCAAGGG	26	7, S32–S37
DNA	mCherry	Cy5-ACCGCCAAGCTGAAGGTGACCAAGGG	26	7, S32–S37
RNA	DsRed2	6-FAM-ACCGUGAAGCUGAAGGUGACCAAGGG	26	7, 8, S32, S33, S38–S40
RNA	mCherry	Cy5-ACCGCCAAGCUGAAGGUGACCAAGGG	26	7, 8, S32, S33, S38–S40
RNA	DsRed2 mRNA	—ACCGUGAAGCUGAAGGUGACCAAGGG—	817	8, S33
RNA	mCherry mRNA	—ACCGCCAAGCUGAAGGUGACCAAGGG—	1367*	8, S33
DNA	$DsRed2_{alt}$	6-FAM-TCATCACCGAGTTCATGCGCTTCAAG	26	S41–S43
DNA	$mCherry_{alt}$	Cy5-TCATCAAGGAGTTCATGCGCTTCAAG	26	S41–S43
RNA	$DsRed2_{alt}$	6-FAM-UCAUCACCGAGUUCAUGCGCUUCAAG	26	S44–S46
RNA	$mCherry_{alt}$	Cy5-UCAUCAAGGAGUUCAUGCGCUUCAAG	26	S44–S46

*This target is an mCherry–human histone H2B mRNA fusion.

Table S2. Shielded covalent probe sequences. Toehold in green. ^{CNV}K nucleoside in red.

Material	Target	Toehold (nt)	Sequence (5' to 3')	Figures
DNA	C	2	CAACCGTTCCTGCGAGTAGCTGGATACTCGCAKGAACGGTTGAC	6, S10, S21–S24
DNA	C	3	CAACCGTTCCTGCGAGTAGCTGGATACTCGCAKGAACGGTTGACA	4, 6, S10, S15, S17, S21–S24
DNA	C	4	CAACCGTTCCTGCGAGTAGCTGGATACTCGCAKGAACGGTTGACAT	5, 6, S10, S17, S21–S24, S25, S27
DNA	C	5	CAACCGTTCCTGCGAGTAGCTGGATACTCGCAKGAACGGTTGACATG	3, 6, S3–S11, S13, S17, S21–S24, S26, S28, S34
DNA	C	6	CAACCGTTCCTGCGAGTAGCTGGATACTCGCAKGAACGGTTGACATGG	6, S10, S21–S24
DNA	C	7	CAACCGTTCCTGCGAGTAGCTGGATACTCGCAKGAACGGTTGACATGGC	6, S10, S21–S24
DNA	C	8	CAACCGTTCCTGCGAGTAGCTGGATACTCGCAKGAACGGTTGACATGGCT	6, S10, S21–S24
DNA	C	9	CAACCGTTCCTGCGAGTAGCTGGATACTCGCAKGAACGGTTGACATGGCTT	6, S10, S21–S24
RNA	C	3	CAACCGUUCCUGCGAGUAGCUGGAUACUCGCAKGAACGGUUGACA	S29–S31
RNA	C	4	CAACCGUUCCUGCGAGUAGCUGGAUACUCGCAKGAACGGUUGACAU	S16, S18, S29–S31
RNA	C	5	CAACCGUUCCUGCGAGUAGCUGGAUACUCGCAKGAACGGUUGACAUG	3, 4, 5, S11, S13, S15, S18, S29–S31
RNA	C	6	CAACCGUUCCUGCGAGUAGCUGGAUACUCGCAKGAACGGUUGACAUGG	S29–S31
RNA	C	7	CAACCGUUCCUGCGAGUAGCUGGAUACUCGCAKGAACGGUUGACAUGGC	S29–S31
DNA	DsRed2	4	CACCTTCAKCTTCACGGTCCCTTTACCGTGAAGCTGAA	S35, S37, S38, S40
DNA	DsRed2	5	TCACCTTCAKCTTCACGGTCCCTTTACCGTGAAGCTGAA	S35, S37, S38, S40
DNA	DsRed2	6	GTCACCTTCAKCTTCACGGTCCCTTTACCGTGAAGCTGAA	7, S32, S35, S37, S38, S40
DNA	DsRed2	7	GGTCACCTTCAKCTTCACGGTCCCTTTACCGTGAAGCTGAA	8, S33, S35, S37, S38, S40
DNA	DsRed2	8	TGGTCACCTTCAKCTTCACGGTCCCTTTACCGTGAAGCTGAA	S35, S37, S38, S40
DNA	DsRed2	9	TTGGTCACCTTCAKCTTCACGGTCCCTTTACCGTGAAGCTGAA	S35, S37, S38, S40
DNA	mCherry	5	TCACCTTCAKCTTGGCGGTCCCTTTACCGCCAAGCTGAA	S36, S37, S39, S40
DNA	mCherry	6	GTCACCTTCAKCTTGGCGGTCCCTTTACCGCCAAGCTGAA	7, S32, S36, S37, S39, S40
DNA	mCherry	7	GGTCACCTTCAKCTTGGCGGTCCCTTTACCGCCAAGCTGAA	8, S33, S36, S37, S39, S40
DNA	DsRed2$_{alt}$	3	CATGAACTCGGTGAKGAACTAAATCATCACCGAGTTC	S41, S43, S44, S46
DNA	DsRed2$_{alt}$	4	GCATGAACTCGGTGAKGAACTAAATCATCACCGAGTTC	S41, S43, S44, S46
DNA	DsRed2$_{alt}$	5	CGCATGAACTCGGTGAKGAACTAAATCATCACCGAGTTC	S41, S43, S44, S46
DNA	DsRed2$_{alt}$	6	GCGCATGAACTCGGTGAKGAACTAAATCATCACCGAGTTC	S41, S43, S44, S46
DNA	DsRed2$_{alt}$	7	AGCGCATGAACTCGGTGAKGAACTAAATCATCACCGAGTTC	S41, S43, S44, S46
DNA	DsRed2$_{alt}$	8	AAGCGCATGAACTCGGTGAKGAACTAAATCATCACCGAGTTC	S41, S43, S44, S46
DNA	mCherry$_{alt}$	4	GCATGAACTCCTTGAKGAACTAAATCATCAAGGAGTTC	S42, S43, S45, S46
DNA	mCherry$_{alt}$	5	CGCATGAACTCCTTGAKGAACTAAATCATCAAGGAGTTC	S42, S43, S45, S46
DNA	mCherry$_{alt}$	6	GCGCATGAACTCCTTGAKGAACTAAATCATCAAGGAGTTC	S42, S43, S45, S46

Table S3. Unstructured probe sequences. ^{CNV}K nucleoside in red.

Material	Target	Length (nt)	Sequence (5' to 3')	Figures
DNA	C	23	TACTCGCAKGAACGGTTGACATG	S5, S14
RNA	C	23	UACUCGCAKGAACGGUUGACAUG	S20

Table S4. Helper strand sequences.

Material	Target	Length (nt)	Sequence (5' to 3')	Figures
RNA	DsRed2 5'-flank	30	GUUGUGGCCCUCGUAGGGGCGGCCCUCGCC	8, S33
RNA	DsRed2 3'-flank	30	AGGAUGUCCCAGGCGAAGGGCAGGGGGCCG	8, S33
RNA	mCherry 5'-flank	30	CUGGGUGCCCUCGUAGGGGCGGCCCUCGCC	8, S33
RNA	mCherry 3'-flank	30	AGGAUGUCCCAGGCGAAGGGCAGGGGGCCA	8, S33

S2 Equilibrium models and calculations for probe/target hybridization

In this section, we explore the equilibrium properties of probe/target hybridization, including the effects of operating in more or less permissive conditions, and tabulate the predicted reaction free energy changes and measured capture yields for the probes and targets used in the mechanism studies of Figure 6.

S2.1 Equilibrium model for probe/target hybridization

Consider the binding of a probe P to a target T to form the complex PT. We wish to calculate the fraction of probe bound as a function of the standard free energy of binding, stoichometry, and temperature. Let the total concentration of target molecules be denoted T_0 and the total concentration of probe molecules be denoted P_0 with $r \equiv P_0/T_0$. At equilibrium, we have:

$$K_{eq} \equiv \frac{[PT]}{[P][T]} = \frac{[PT]}{(rT_0 - [PT])(T_0 - [PT])} = e^{\frac{-\Delta_r G^0}{R\Theta}} \tag{S1}$$

with

$$\Delta_r G^0 \equiv \Delta G^0(PT) - \Delta G^0(P) - \Delta G^0(T) \tag{S2}$$

the standard molar Gibbs free energy change of reaction and R is the molar gas constant and Θ is temperature. Solving

$$0 = K_{eq}[PT]^2 - (T_0(r+1)K_{eq} + 1)[PT] + K_{eq} r T_0^2$$

for $[PT]$ yields one solution with the proper bound ($\lim_{K_{eq}\to 0}[PT] = 0$):

$$[PT] = \frac{(r+1)T_0K_{eq} + 1}{2K_{eq}} - \frac{\sqrt{(r-1)^2T_0^2K_{eq}^2 + 2(r+1)T_0K_{eq} + 1}}{2K_{eq}}$$

The fraction of target bound at equilibrium is thus

$$F \equiv \frac{[PT]}{T_0} = \frac{(r+1)T_0K_{eq} + 1}{2T_0K_{eq}} - \frac{\sqrt{(r-1)^2T_0^2K_{eq}^2 + 2(r+1)T_0K_{eq} + 1}}{2T_0K_{eq}} \tag{S3}$$

This is plotted vs. $-\Delta_r G^0$ in Figure S1 for target concentrations of 1 M and 1 μM at 22 °C. The concentration and equilibrium constant ($K_{eq} = e^{-\Delta_r G^0/R\Theta}$) only appear as a product in Eq. S3; as a result, hybridization at other concentrations is equivalent to translation along the free energy axis. As seen from the plot, at least ≈8 kcal/mol difference in binding free energy between the correct and mismatched targets ($\Delta\Delta G$) is needed to simultaneously bind the complement with high yield and the mismatch with low yield. In practice, larger changes are desirable to ensure that fluctuations in temperature or other environmental variables do not result in either loss of the desired target or hybridization of mismatches. At room temperature, $\Delta\Delta G$ values are somewhat smaller than 8 kcal/mol for 2-nt DNA mismatches and smaller still for SNP mismatches,[4] implying that durable binding and sequence selectivity cannot be achieved simultaneously using DNA hybridization alone.

S2.2 Effect of target concentration on hybridization equilibrium

Probe-target hybridization is a bimolecular reaction and the equilibrium is therefore sensitive to the concentration of the species, as shown in Figure S1. Typically, the probe concentration is under control of the experimenter but the target concentration is unknown and may be very small relative to that of the probe. To examine hybridization under these conditions, we can rewrite the equilibrium condition of Eq. S1 in terms of the fraction of targets hybridized $F \equiv [PT]/T_0$:

$$K_{eq} = \frac{F}{T_0(1-F)(r-F)} \tag{S4}$$

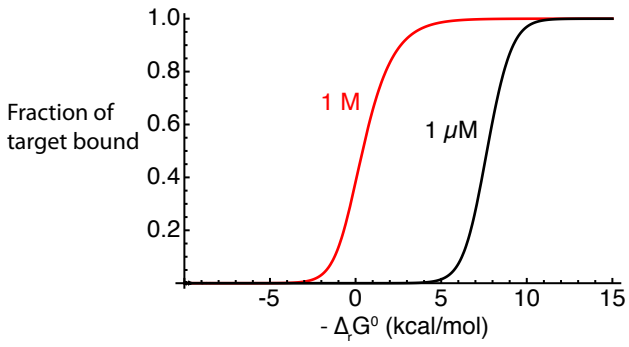


Figure S1. Yield curve for two-state equilibrium model of probe/target hybridization. Two concentrations at 22 °C.

When the target concentration is very small relative to that of the probe ($r \gg 1 \geq F$),

$$K_{eq} \approx \frac{F}{rT_0(1-F)} \tag{S5}$$

yielding

$$F \approx \frac{rT_0K_{eq}}{1+rT_0K_{eq}} = \frac{P_0K_{eq}}{1+P_0K_{eq}}, \qquad r \gg 1. \tag{S6}$$

Hence, the fraction of target bound at equilibrium becomes independent of the target concentration once the target concentration drops substantially below the probe concentration. The same behavior can be observed by plotting Eq. S3 vs. total target concentration (shown in Figure S2 for four values of $\Delta_r G^0$). For fixed P_0 and approximately $r > 10$, the equilibrium fraction of targets bound is determined only by the reaction free energy change (or equivalently K_{eq}), not by the target concentration. Note that for a given $\Delta_r G_0$ (i.e., a given probe/target pair), F asymptotes to a constant from below as T_0 is decreased. Hence, if a probe achieves high-yield capture with a given target concentration, the yield is predicted to be as high or higher if the target concentration is decreased. In practical terms, this means that probes can be designed without knowing the exact target concentration in advance, and that there is no lower limit on target concentration at which one should expect deviation from the designed behavior. Furthermore, to switch to a different probe concentration and maintain a given yield, $\Delta_r G^0$ can be adjusted by changing the probe toehold length (or loop length), providing the flexibility to use SC probes at a desired concentration. Experimental results demonstrating the insensitivity of SC probe capture yield to target concentration are shown in Supplementary Section S6.10.

S2.3 Calculated hybridization reaction free energies and measured capture yields

Equation S3 predicts the equilibrium fraction of target bound as a function of the reaction stoichometry and the reaction equilibrium constant (or reaction free energy, per Eq. S1). In order to compare this prediction to free energies calculated using NUPACK[1] (as in Fig. 1), it is necessary to take into account the fact that NUPACK's calculations are performed using mole fractions rather than molarity. For a reaction $\sum_{i=1}^{L} \nu_i A_i \rightleftharpoons \sum_{j=1}^{M} \upsilon_j B_i$ involving L reactants A_i and M products B_j, the NUPACK equilibrium constant K_{NUPACK} is given in terms of the mole fractions χ of each species by the following expression:

$$K_{\text{NUPACK}} = \frac{\prod_{j=1}^{M} \chi_B^{\upsilon_j}}{\prod_{i=1}^{L} \chi_A^{\nu_i}}$$

In the dilute solutions for which the nearest-neighbor model applies, all species are present at concentrations much lower than that of water ($[H_2O] \equiv \rho_W \approx 55.35$ mol/L at 295 K) and the NUPACK and standard equilibrium

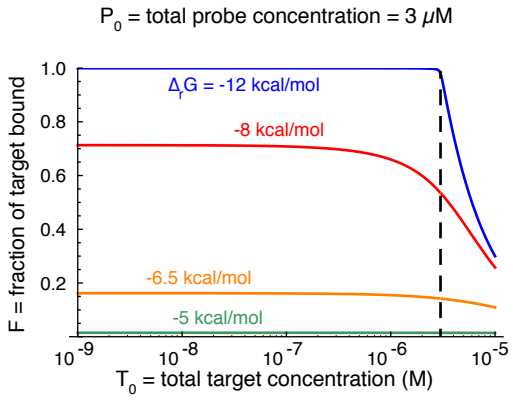


Figure S2. Calculated equilibrium fraction of target bound (F) as a function of total target concentration (T_0) for four values of $\Delta_r G^0$ with total probe concentration (P_0) fixed at 3 μM. Temperature at 22°C.

constants are related by:

$$K_{\text{NUPACK}} = \frac{\prod_{j=1}^{M} \chi_B^{\upsilon_j}}{\prod_{i=1}^{L} \chi_A^{\nu_i}} \cong \frac{\prod_{j=1}^{M} ([B]/\rho_W)^{\upsilon_j}}{\prod_{i}^{L} ([A]/\rho_W)^{\nu_i}} = \rho_W^{(\sum_{i=1}^{L} \nu_i - \sum_{j=1}^{M} \upsilon_j)} K_{eq}$$

For the probe/target binding reaction, we have $L = 2$, $M = 1$, and $K_n = \rho_W K_{eq}$. Similarly, the standard (molar) reaction free energy change, $\Delta_r G^0$, is related to the NUPACK reaction free energy change, ΔG_{NUPACK}, by

$$\Delta G^0 = \Delta G_{\text{NUPACK}} + (\sum_{i=1}^{L} \nu_i - \sum_{j=1}^{M} \upsilon_j) RT \ln \rho_W .$$

For our probe/target reaction, this is simply

$$\Delta G^0 = \Delta G_{\text{NUPACK}} + RT \ln \rho_W , \tag{S7}$$

so at T = 295 K, the correction ($RT \ln \rho_W = 2.36$ kcal/mol) is on the order of 1–2 DNA base pairs.

Tables S5–S7 show the calculated reaction free energies and measured crosslinking yields for experiments in which complementary and mismatched DNA targets (Table S1) were hybridized with DNA SC probes with 3–5 nt toeholds (Table S2). The reaction free energy change, $\Delta_r G^0$ is computed using NUPACK (Equations (S2) and (S7) at 295 K and [Na^+] = 0.195 M (to approximate the 1× SSC hybridization buffer). For these calculations, the maximum complex size was set to 3 strands, but in no case was any ternary complex predicted to form to an appreciable degree.

Target	5-nt toehold $\Delta_r G^0$ (kcal/mol)	Yield (%)
DNA C	-14.0	90.7 ± 3.4
DNA M1	-7.9	< 0.1
DNA M2	-9.5	< 0.1
DNA M3	-7.3	0.1 ± 0.1

Table S5. Crosslinking yield (mean $\pm$ standard deviation, $N = 3$) and calculated reaction free energy for 29-nt DNA complementary and 2-nt mismatched targets (C and M1–M3, Table S1) and a 5-nt toehold DNA probe (Table S2). See Figure 3.

Target	5-nt toehold $\Delta_r G^0$ (kcal/mol)	Yield (%)	4-nt toehold $\Delta_r G^0$ (kcal/mol)	Yield (%)	3-nt toehold $\Delta_r G^0$ (kcal/mol)	Yield (%)
DNA C	-13.6	96.4 ± 1.2	-12.3	95.1 ± 1.1	-11.6	35.1 ± 0.9
DNA S1	-9.1	32.0 ± 2.5	-7.8	1.0 ± 0.1	-7.1	0.3 ± 0.2
DNA S2	-9.7	68.4 ± 1.6	-8.4	3.6 ± 0.1	-7.7	0.4 ± 0.3
DNA S3	-10.2	61.7 ± 1.3	-8.9	3.2 ± 0.6	-8.2	0.4 ± 0.3
DNA S4	-10.6	87.0 ± 0.2	-9.2	12.9 ± 1.4	-8.6	1.1 ± 0.7
DNA S5	-9.7	76.8 ± 1.4	-8.4	5.3 ± 0.5	-7.7	0.6 ± 0.4
DNA S6	-10.8	91.4 ± 1.0	-9.6	29.2 ± 1.4	-8.8	2.2 ± 1.0

Table S6. Crosslinking yield (mean $\pm$ standard deviation, $N = 3$) and calculated reaction free energy for 23-nt DNA complementary and SNP targets (C and S1–S6, Table S1) and 3–5 nt toehold DNA probes (Table S2). See Figures 4 and 6bc.

Target	5-nt toehold $\Delta_r G^0$ (kcal/mol)	Yield (%)	4-nt toehold $\Delta_r G^0$ (kcal/mol)	Yield (%)
DNA C	-13.6	97.1 ± 0.6	-12.3	95.5 ± 1.0
DNA 2-1	-8.0	1.1 ± 0.2	-6.7	0.1 ± 0.1
DNA 2-2	-7.3	0.4 ± 0.1	-6.0	0.1 ± 0.1
DNA 2-3	-7.5	2.5 ± 0.7	-6.2	0.1 ± 0.1
DNA 2-4	-7.6	0.5 ± 0.2	-6.3	0.1 ± 0.1
DNA 2-5	-8.4	14.0 ± 4.3	-7.1	0.5 ± 0.1

Table S7. Crosslinking yield (mean $\pm$ standard deviation, $N = 3$) and calculated reaction free energy for 23-nt DNA complementary and 2-nt mismatched targets (C and M1–M3, Table S1) and 4–5 nt DNA probes (Table S2). See Figure 6c.

S3 SC probe synthesis and purification

S3.1 SC probe synthesis

Probes were synthesized using a MerMade 6 oligonucleotide synthesizer (BioAutomation) with phosphoramidites (including the ^{CNV}K photoactive nucleoside analog) and reagents from Glen Research. DNA probe synthesis was carried out using standard solid-phase protocols, using acetyl-protected dC and dimethylformamidine-protected dG amidites, which enable "UltraFAST" cleavage and deprotection. It was not known whether the ^{CNV}K crosslinker could tolerate the UltraFAST conditions. To investigate this, we synthesized two 5-nt toehold SC probes in parallel and compared the results of deprotection with the UltraFAST protocol (10 min at 55 °C in 1:1 ammonium hydroxide : 40% aqueous methylamine (AMA)) to conventional deprotection conditions (16 hr at RT in ammonium hydroxide). In both cases, deprotection was followed by DMT-affinity chromatography (GlenPak cartridges, Glen Research) to remove truncated syntheses. The reactions were then analyzed by reversed-phase HPLC using an XBridge OST C18 column (2.5 μm, 4.6 $\times$ 50 mm, Waters Corp.) and a 5–15% gradient of acetonitrile (ACN, 0.5%/min) in 0.1 M triethylamine acetate (TEAA, pH 7.0). The chromatograms of Figure S3 show that a secondary species is present in the UltraFAST-deprotected probes. Isolation of the two species shows that 1) the secondary species crosslinks with much lower ($\approx 20\times$) activity than the main product, and 2) the secondary species' mass is ≈58 D larger than the main product by MALDI-TOF MS. This mass difference is larger than the mass of any of the species in the AMA mixture, but is similar to what might be expected from addition of acrylonitrile (produced during removal of the cyanoethyl phosphate protection groups) to ^{CNV}K. We then examined the effect of various deprotection conditions on the fraction of secondary species produced. As Table S8 shows, deprotection with ammonium hydroxide at room temperature appears to avoid generation of the side product. If the probes must be exposed to AMA (as for some RNA synthesis protocols, see below), it appears preferable to perform the reaction at higher temperature for a shorter time rather than at RT.

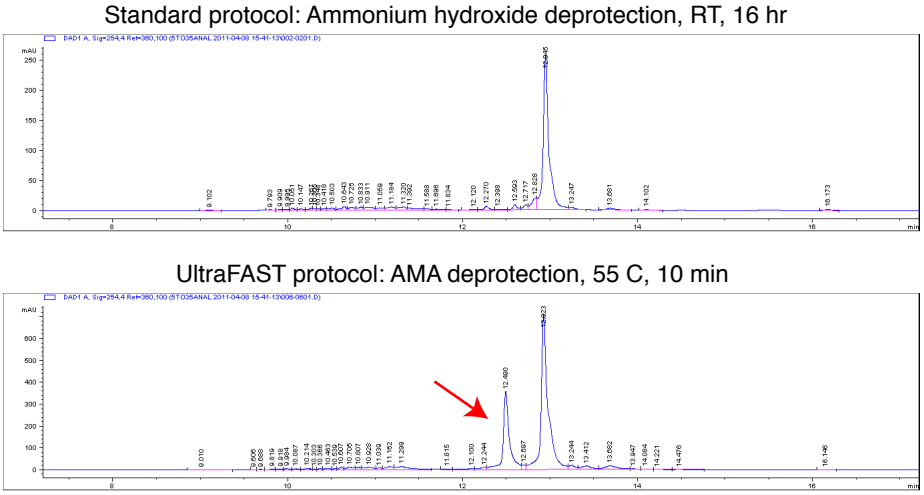


Figure S3. Deprotection in AMA using the UltraFAST protocol produces a secondary species that is resolved by HPLC (Waters OST C18 column, TEAA:ACN gradient, 1.5 %/min). 5-nt toehold DNA SC probe targeting C (Table S2).

RNA probes were also synthesized according to standard protocols, using 2'-TOM-protected monomers (Glen Research). These monomers afford improved yield and purity for longer oligos such as our hairpins (47 nt for the 5-nt toehold probes), but must be deprotected using AMA, which reduced yield of functional probe as described above. As a result of our investigations, the bases were deprotected at elevated temperature, followed by rapid evaporation of the volatile deprotection solution using a blow-down evaporator at room temperature. The sugars were deprotected using triethylamine-buffered TEA.3HF per standard protocols.

Solution	Temperature	Time	% of probe in side peak
NH_4OH	RT	16 hr	≈0
NH_4OH	55 °C	4 hr	6.5
NH_4OH	55 °C	8 hr	11.6
NH_4OH	55 °C	16 hr	23.1
AMA	RT	2 hr	41.3
AMA	55 °C	10 min	31.1

Table S8. Effect of deprotection conditions on probe purity. NH_4OH: ammonium hydroxide; AMA: 1:1 ammonium hydroxide : 40% aqueous methylamine. 5-nt toehold DNA SC probe targeting C (Table S2).

S3.2 SC probe purification

Following removal of truncated species, DNA and RNA probes were purified by RP-HPLC with a XBridge OST C18 column (2.5 μm, 10 x 50 mm, Waters Corp.), using 50 mM ammonium formate (pH 9.2) and acetonitrile as the mobile phases. The collected fractions were evaporated to reduce volume (typically to < 1 mL, $\approx 10\times$ reduction) and remove the volatile salts used in the separation. We found that complete drying produced a small ($< 5\%$) but noticeable amount of covalent probe dimers, so the purification process was completed by butanol extraction of the remaining liquid, followed by resuspension of the probes in the desired buffer (SSC for the DNA probes, RNA Storage Solution (Life Technologies) for the RNA probes).

Interestingly, we found that purification using the ammonium formate : acetonitrile (ACN) mobile phase mixture described by Yoshimura et al.[5] resulted in notably better probe performance than did purification with 0.1M triethylamine acetate (TEAA, pH 7.0) : ACN, which is a more-commonly used mobile phase pair for oligonucleotide purification. Specifically, SC probes purified with TEAA:ACN displayed a "floor" of ≈3% for crosslinking, even with short toeholds and/or large mismatches. Analytical-scale HPLC separations with the two mobile phase pairs showed that TEAA- and ammonium formate-purified probes appeared identical when separated on a TEAA:ACN gradient but TEAA-purified probes displayed a small trailing shoulder when separated on an ammonium formate gradient. Repurification of TEAA-purified probes with ammonium formate resulted in good performance, and repurification of ammonium formate-purified probes with TEAA did not reduce their performance, which suggests that residual salt is not the cause of the difference. Additionally, any residual salt would be present in negligible quantity relative to the storage and reaction buffers. No significant difference was observed when TEAA- and ammonium formate-purified probes were analyzed side-by-side on a denaturing gel, suggesting that whatever difference there is is very subtle. Finally, denaturing gel purification of probes did not produce improved performance relative to TEAA:ACN HPLC purification, and produced noticeable amounts of dimerized or covalently-closed probes, likely due to UV exposure while cutting the bands out of the gel. When purified with ammonium formate RP-HPLC as described above, we estimate that the probes are at least 97% pure prior to UV exposure, as measured by analytical scale RP-HPLC.

S4 ^{CNV}K crosslinker studies

S4.1 Crosslinker durability in highly stringent conditions

To test the durability of the ^{CNV}K photoadduct, we crosslinked a 5-nt toehold SC probe to a Cy5-labeled complementary target and incubated it in a 50% formamide solution at 95 °C, progressively removing aliquots as shown in Table S9. A slight decrease in yield is observed over the course of the experiment. This may be due to degradation of the Cy5 fluorophore or to reversal of the crosslink in these harsh conditions. Typical hybridization and wash conditions are significantly milder, suggesting that bonds formed by SC probes with ^{CNV}K crosslinkers are not labile under practical conditions.

Time (min) at 95 °C	Crosslink yield (%)
0	93.3
8	93.3
13	92.2
18	91.9
28	91.9
48	90.8

Table S9. Crosslink yield as a function of incubation time in 50% formamide at 95 °C. 29-nt complementary DNA target (C in Table S1) and 5-nt toehold DNA SC probe (Table S2).

S4.2 Absence of thermal activation

To test for thermal activation of crosslinking, we annealed SC probes in SSC hybridization buffer and used denaturing polyacrylamide gels to look for covalent closure of the hairpins after 1–4 hours at 0, 50, or 95 °C. No activation was observed under any of these conditions.

S4.3 Light sources for photo-activation

Yoshimura et al.[5,6] activate ^{CNV}K crosslinking using a high-intensity 366 nm LED (UV-LED, Omron, 1600 mW/cm^2) and observe full activation within 25 seconds of exposure, depending on the target and neighboring bases (dT/U is crosslinked faster than dC). We were not able to procure this lamp in the US, but found a near-equivalent (LED-100, Electro-Lite) with 1500 mW/cm^2 nominal irradiance at 365 nm. The beam emitted by this source is highly divergent, but Figure S4 shows that it activates crosslinking of SC probes to complementary targets extremely rapidly (< 10 s) when positioned immediately above a microplate well containing the sample. Evaporative losses are negligible during this short exposure, but only one reaction can be irradiated at a time by this source.

UV time (sec)	Crosslink yield (%)
0	0.7
1	64.5
2	77.6
5	88.0
10	95.0
20	95.5
60	95.9
180	96.3

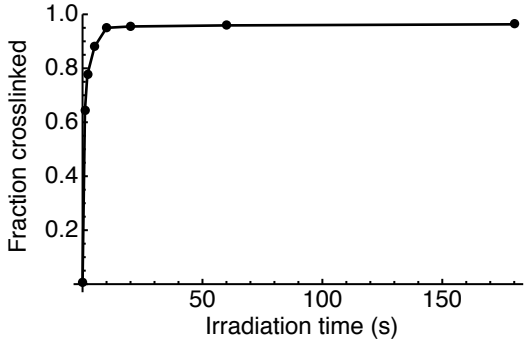


Figure S4. ^{CNV}K crosslinking activation with an LED-100 365 nm lamp. 29-nt complementary DNA target (C in Table S1) and 5-nt toehold DNA probe (Table S2).

We investigated several alternative UV-A sources as well, including a large-bulb mercury arc lamp (B-100A, UVP) which is capable of illuminating a number of reactions simultaneously with an irradiance of ≈ 40 mW/cm^2 (measured with a UVX-36 365 nm photometer, UVP) at a distance of 1 inch. As shown in Figure S5, this lamp also rapidly activates crosslinking, and does so at the same rate for both hairpin SC probes and single-stranded probes hybridized to complementary targets. We also tested two ThorLabs LED sources which can be mounted in standard optical mounts or microscope ports. When fitted with an aspherical collimating lens, the 365 nm LED (M365L2) fully activated crosslinking of an SC probe:target pair in < 2 minutes when set to an irradiance of 20 mW/cm^2 (Figure S6). A 405 nm LED (M405L2, irradiance 1.3 mW/cm^2 measured with the 365 nm photometer) also successfully activated crosslinking, though much more slowly.

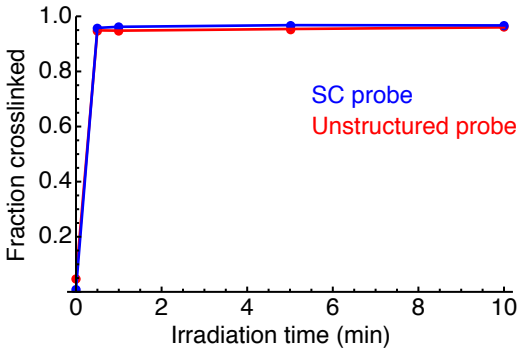


Figure S5. ^{CNV}K crosslinking activation with a UVP B-100A mercury arc lamp and a 365 nm filter. 29-nt DNA complementary target (C in Table S1) and 5-nt toehold DNA SC probe (Table S2) or 23-nt DNA unstructured probe (Table S3).

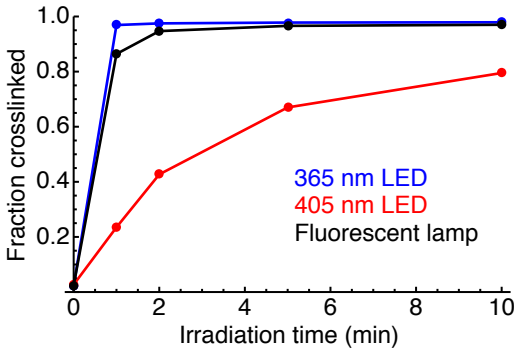


Figure S6. ^{CNV}K crosslinking activation with ThorLabs LEDs, UV-A fluorescent lamp. 29-nt complementary DNA target (C in Table S1) and 5-nt toehold DNA probe (Table S2).

The final light source we tested was a low-intensity UV-A fluorescent lamp (Nishujie-1593 V-1D, 9W electrical power) used for curing acrylic fingernail polish. This lamp is extremely inexpensive ($10) and is capable of illuminating several rows of a microplate or all of a small petri dish at once. The irradiance at 365 nm was measured to be 5.5 mW/cm^2, and SC probe crosslinking was fully activated after ≈6 minutes of exposure (Fig. S6).

Taking the three data sets (Hg arc, 365 nm LED, fluorescent lamp) together, it appears that a dose of roughly 2 J/cm^2 is sufficient to fully activate ^{CNV}K crosslinking in the context of hybridized SC probes and targets. Each of the light sources investigated has its own strengths and weaknesses. The LED-100 allows very short exposure times, which may be useful for kinetically-resolved measurements and minimizes evaporative losses when irradiation is carried out at elevated temperatures. As mentioned above, the ThorLabs LEDs are easily integrated into standard optical systems, and the 405 nm LED may be useful for systems that are particularly UV-sensitive. For most applications, however, the fluorescent lamp may be the best choice, given its price and versatiilty, provided that long exposure times can be tolerated.

We conducted a preliminary experiment to measure whether this dose would result in phototoxic effects to cultured cells. Cultured Ewing's sarcoma TC71 cells[7,8] at ≈50% confluency were exposed to either the high-intensity LED-100 or the low-intensity fluorescent light under conditions more severe than required for crosslinking activation (3 minute exposure with the LED, 15 minute exposure for the fluorescent light) at RT. No sign of morphological or growth rate change could be observed over the subsequent 24 hours. Given that UV-A photons do not have sufficient energy to trigger pyrimidine dimerization,[9] this is not surprising.

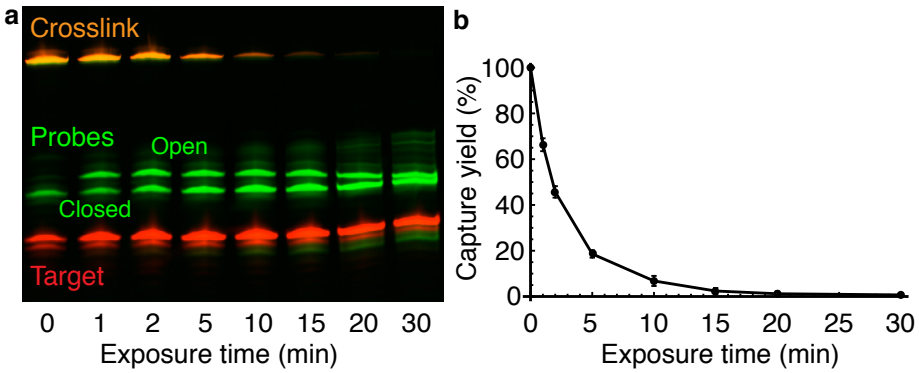


Figure S7. Photo-reversal of SC probe/target crosslinks. Irradiation with 311 nm UV light in denaturing conditions reverses the $^{\mathrm{CNV}}$K crosslink, releasing the target. 29-nt complementary DNA target (Table S1) and 5-nt toehold SC DNA probe (Table S2) from Figure 3. a) Denaturing polyacrylamide gel. Fluorescent channels: Cy5 (red) and SYBR Gold post-stain (green). After target release, the probe is found in a mixture of covalently closed and open states, implying that light at this wavelength promotes both photo-reversal and photo-crosslinking. b) Dependence of probe/target crosslink yield on irradiation time (mean $\pm$ standard deviation, $N = 3$).

S4.4 Photo-reversal studies

We wished to establish reversal of $^{\mathrm{CNV}}$K crosslinks using low-intensity UV-B light sources, as high-intensity irradiation at these wavelengths is associated with photodamage due to pyrimidine dimerization and/or reactive oxygen generation.[9] Figure S7 shows the results of irradiating crosslinked probe-target complexes with low-intensity UV-B light, as described in the main text Methods Summary. Reversal manifests as disappearance of the crosslink band (orange) and strengthening of the probe (green) and target (red) bands. The fraction of crosslinked probes (quantified as the ratio of Cy5 fluorescence in the crosslink band to the total) is tabulated in Figure S8. Probe-target crosslinking is 99% reversed after 20 minutes of exposure, and indistinguishable from background after 30 minutes. Additional faint bands can be discerned above the probe and below the target bands at the longest exposures. These may correspond to photodamaged species, but are present in extremely low quantity ($< 1\%$). We were not able to observe or isolate these species by RP-HPLC.

Previous reports of $^{\mathrm{CNV}}$K crosslinking and photo-reversal have described these two processes separately, with crosslinking being driven by UV-A irradiation (365 nm) and reversal by UV-B (311 nm) irradiation, implying distinct action spectra for the two reactions.[5,6] This behavior is observed for psoralen,[10] a well-characterized [2+2] photoadduct crosslinker, and its derivatives. If this understanding is accurate, we would expect first-order (single-exponential) kinetics for photo-reversal, since our analyses are carried out under denaturing conditions and are therefore sensitive only to whether the probe and target are covalently linked. As shown in Figure S8, however, a double-exponential model fits the observed data better than a single-exponential model does. This implies the existence of at least two steps in the reaction to reverse the covalent bond between the SC probe and complementary target. We also observed that photo-reversal is both faster and more complete in denaturing buffers (lower salt, added urea or formamide, or organic cosolvents), and that, after reversal, the probes are found in a mixture of covalently closed and open states (Figure S7, lower green band increases in intensity after irradiation). Neither of these observations is consistent with simple photo-reversal.

A general kinetic scheme meeting the two-step mechanism requirement is

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} B \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} C,$$

where A is a state with the probe and target crosslinked, B is some non-crosslinked intermediate, and C is a state with the two molecules completely dissociated. In principle, both steps of the reaction are reversible, but the observation that, in denaturing conditions, the photo-reversal reaction proceeds essentially to completion means that either $k_1 \gg k_{-1}$ or $k_2 \gg k_{-2}$. Since we measure the amount of A remaining as a function of time, $k_1 \gg k_{-1}$ would result in

UV time (min)	Crosslink yield (%)
0	100.0
1	66.3 ± 2.8
2	45.7 ± 2.5
5	18.6 ± 1.6
10	6.8 ± 2.2
15	2.4 ± 1.3
20	1.2 ± 1.0
30	0.7 ± 0.8

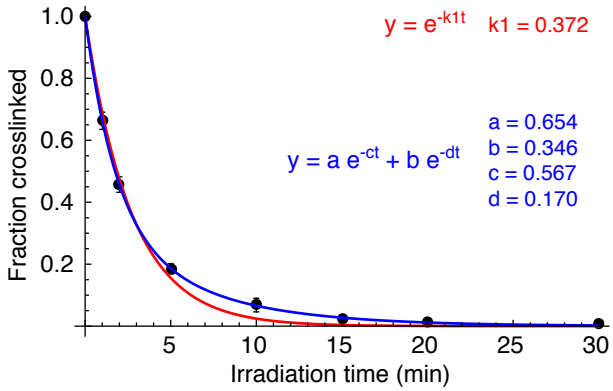


Figure S8. Photo-reversal kinetic data and model fits. Yields are normalized to 0 minute sample (before exposure to UV-B irradiation).

single-exponential kinetics, which is contradicted by the data. Thus, the simplest kinetic scheme that can explain the data is

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} B \xrightarrow{k_2} C.$$

This scheme produces the following rate law for $A(t)$:

$$A(t) = ae^{-ct} + be^{-dt}, \tag{S8}$$

with $a + b = 1$ and $c, d = \frac{1}{2}(k_1 + k_{-1} + k_2 \pm \sqrt{(k_1 + k_{-1} + k_2)^2 - 4k_1k_2})$. The unconstrained fit shown in the figure appears consistent with this rate law, with $a + b = 1.0005$. More data are required in order to fully understand the kinetics of SC probe photo-reversal. However, with the identification of B as the state with the probe and target hybridized but no longer crosslinked, this simple scheme appears consistent with the observed kinetics, as well as the requirement for denaturing conditions (so that $k_2 \gg k_{-2}$) and the observed mixture of covalently-closed and open probes after 311 nm irradiation ($k_1 \approx k_{-1}$).

S4.5 Wavelength-dependence of photo-activation and photo-reversal processes

In order to investigate the photophysics of crosslinking and reversal more fully, we performed an experiment in which a 5-nt toehold DNA SC probe was annealed with its complement and then exposed to light at one of three wavelengths (365 nm, which is expected to drive crosslinking, 311 nm, expected to produce photo-reversal, or 254 nm, which drives photo-reversal in psoralen) in native (1× SSC) conditions to look for crosslinking. A crosslinked probe/target complex was exposed to the same three wavelengths in denaturing (2M urea, 50% ACN) conditions to look for photo-reversal. The 254 nm light was provided by a standard UV-C transilluminator (UVP). The results of this experiment are shown in Figure S9: all three wavelengths drive both the forward (crosslinking) and reverse reactions, though crosslinking appears to be favored as the wavelength increases. These results imply very broad action spectra for both processes, a phenomenon which is not observed for psoralen.[11] Further investigation of this interesting photophysical behavior would be useful.

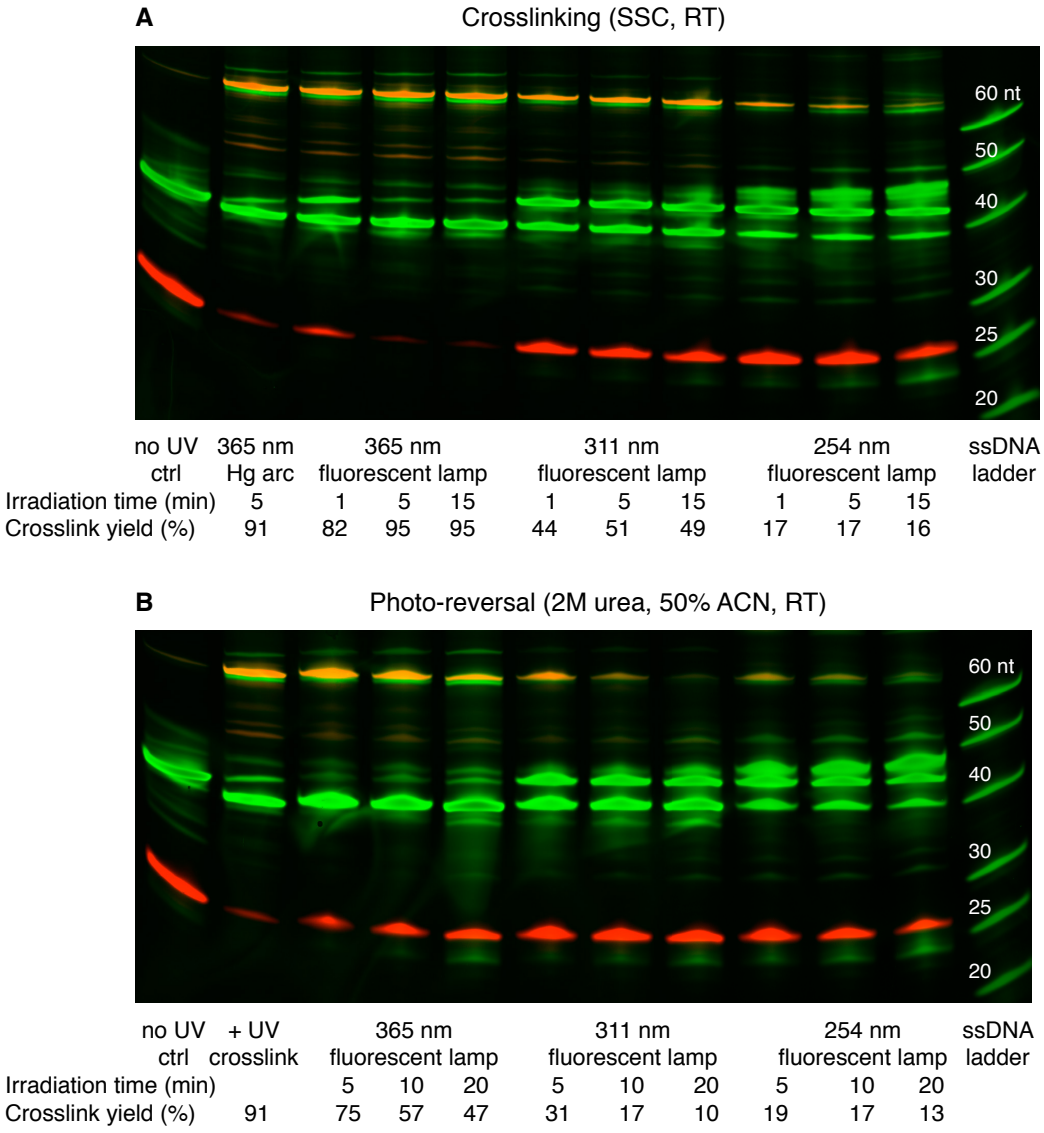


Figure S9. ^{CNV}K crosslinking and photo-reversal as a function of wavelength. Fluorescent channels: Cy5 (red) and SYBR Gold post-stain (green). a) Crosslinking: 29-nt complementary DNA target (Table S1) and 5-nt toehold DNA SC probe (Table S2) were annealed in SSC buffer (lane 1), then exposed to light at three different wavelengths for the indicated time. Crosslinking is observed at all three wavelengths. An additional probe band is observed after 254 nm irradiation, suggesting photo-damage. b) Photo-reversal. Crosslinked probe and target (lane 2) were exposed to the same three wavelengths in denaturing conditions (2M urea, 50% ACN). Photo-reversal of crosslinks is seen at all wavelengths, including 365 nm.

S5 Quantification of capture yields and characterization of uncertainties

S5.1 Gel analysis

The primary method used to measure crosslinking yields in this article is separation of reaction products by denaturing gel electrophoresis followed by quantification of fluorescence (typically Cy5, $\lambda_{\max} = 645$ nm) from the crosslinked and unmodified target bands (Section S6, e.g. Fig. S13). 2.5 μL of crosslinking reaction (7.5 pmol probe, 4.5 pmol dye-labeled target) were mixed with an equal volume of loading buffer and loaded into 15% (tris/borate/EDTA plus 7M urea) polyacrylamide gels, along with a single-stranded DNA marker (10/60 Length Standard, Integrated DNA Technologies). The gels were run at 35 V/cm for 2 hours, then stained with SYBR Gold (Life Technologies) and imaged with a Fujifilm FLA-5100 fluorescence scanner. Crosslinking yields are computed as the ratio of (integrated signal in the (+ UV) crosslink area - integrated signal in the (-UV) crosslink area) to total signal in the lane (approximated by the sum of the integrated signals from the crosslink and monomer areas). The average of the three mismatch (- UV) signals is used as the background for the complementary target signal to avoid error due to spurious activation of crosslinking from ambient light. We find that the reliable dynamic range of this method with our apparatus is roughly 3 orders of magnitude, limited on the high end by band opacity and on the low end by autofluorescence of the gel and scanning plate. As a result, we place a "floor" on yield values at 0.1%, as seen in Figure 3 and elsewhere. Quantification of the gels appears highly repeatable, with re-quantification with integration areas of different sizes, etc. resulting in relative yield fluctuations ($\frac{\Delta Y}{Y}$) of $\approx 1\%$ or less. Repeating gel separations for the same reactions also produces negligible fluctuations in the measured yield values. The reactions were also analyzed on 12% native polyacrylamide gels (run at 14 V/cm for 2 hours, then stained with SYBR Gold) to monitor hybridization (Section S6, e.g. Fig. S13).

In the case of experiments with full-length mRNAs (Figs 8 and S33), labeling the target is not useful since the target mobility shift due to crosslinking a short probe to a long ($\approx$1 kb) target is negligible. As a result, we take an alternative approach and reverse the stoichometry so that the probe becomes the limiting reagent (main text Methods Summary). Crosslinking then results in depletion of the covalently-closed probe band after irradiation. In principle, the yield can be quantified by subtracting the band intensity in the "probe + target" reaction from that of an otherwise identical "probe only" reaction. If this analysis is performed for the data shown in Figure 8, crosslinking yields of $>$95% are observed for the DsRed2 probe and the complementary oligo or the complementary mRNA, and crosslinking yields of $>$75% are observed for the mCherry probe and the complementary oligo, with slightly larger yields observed for the complementary mRNA. This analysis technique is prone to substantial variation due to several factors, including pipetting accuracy, non-uniform background autofluorescence (much larger at the SYBR Gold emission wavelength than for the far-red Cy5 fluorophore), and non-quantitative staining of nucleic acids by intercalating dyes such as SYBR Gold. This variation (we estimate $>$10%) is most significant for low-yield reactions, as the depletion is then calculated as the small difference of two large numbers. As a result, we report only qualitative (high/low) results for the crosslinking experiments with full-length mRNAs.

S5.2 HPLC analysis

Additionally, some reactions were analyzed by HPLC in order to confirm the values obtained from the gels: 30–40 μL per crosslinking reaction was loaded onto a 2.5 μm, 4.6 $\times$ 50 mm Waters XBridge OST C18 column and separated on a 5–35% gradient of acetonitrile (1.5% / min.) in 0.1 M triethylamine acetate, pH 7.0. The crosslinking yield was calculated from the ratio of crosslinked to non-crosslinked peak areas at the dye absorbance maximum. Figure S10 shows a typical set of results: the values obtained by the two methods agree closely. HPLC analysis tends to give slightly larger crosslinking yields than the gels as the yields get larger. The HPLC values would lead to slightly higher discrimination ratios than the more conservative gel values that we report.

S5.3 Variation between experiments

Compared to uncertainties due to quantification, we observe larger fluctuations between crosslinking experiments with the same molecules. As a result, measurements are typically reported as the average from three independent

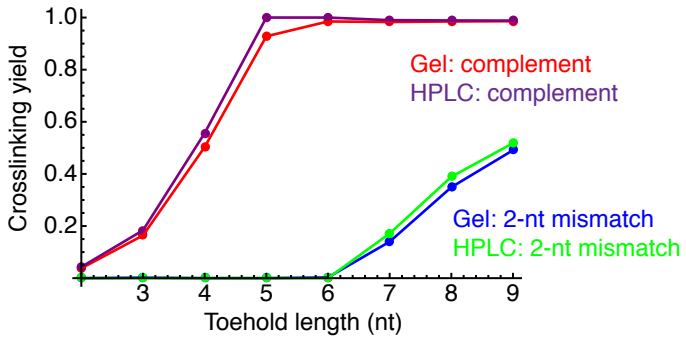


Figure S10. Comparison of gel and HPLC quantification for 29-nt complementary and 2-nt mismatch DNA targets (C and M1 of Table S1) and 2–9 nt toehold DNA SC probes (Table S2). Gels are shown in Figs S21 and S22.

trials. Standard deviations are reported to characterize the degree of variation. Examination of these values shows that the large majority of measurements have fluctuations of 2% or less in absolute yield, and that the measurements with larger fluctuations tend to be those with intermediate yield values (15–70% crosslink yield), where the yield curve is steepest (Fig. S1). We suspect that these fluctuations are therefore likely due to unmeasured differences in environmental conditions from experiment to experiment. In practice, SC probes would likely be designed to avoid this region, so the much smaller deviations observed for low and high yields are more representative of probe performance.

S5.4 Systematic errors due to photobleaching

A final source of error stems from the assumption that the fluorescence signal and number of targets are linearly related. By computing the crosslinking yield as the ratio of integrated fluorescence in the crosslink gel band (or HPLC peak) to the total integrated fluorescence signal, factors such as target labeling efficiency, sample concentration, volume loaded in the lane, etc. are canceled out. Only factors that result in different signal / target molecule ratios between crosslinked and non-crosslinked targets should produce errors. As discussed above, the amount of sample analyzed was chosen so that signals were clearly distinguished from background and did not saturate the detectors. This leaves differential quantum yield or photobleaching between crosslinked and non-crosslinked targets as potential error sources. Measurements are made in denaturing conditions, so quantum yield differences due to fluorophore-base stacking are not anticipated, and we are not aware of any reports of photobleaching rate dependence on oligo hybridization.

Nonetheless, we decided to determine whether either effect was present experimentally as follows: 5-nt toehold DNA and RNA SC probes were annealed with either the corresponding complementary target or a 2-nt mismatch target. Under these conditions, the complement is crosslinked in high yield and the mismatch in low yield (Figure 3, sequences C and M3). Reactions with target but no probe were annealed with the same protocol. The samples were then exposed to super-saturating doses of UV-A radiation with aliquots removed at various time points and analyzed by denaturing gel electrophoresis. If the dye quantum yield is different for hybridized targets, we would expect that at all time points, the total fluorescence for the complementary target would differ depending on whether or not the probe is present. We did not observe this: the target-only and (target + probe) reaction total fluoresence readings agreed within 10% (reasonable, given the multiple pipetting operations) before irradiation. Alternatively, if crosslinked targets photobleach at a different rate than unmodified targets, the total fluorescence should change over time even after crosslinking is saturated, and do so at a different rate for the (probe + complement) reaction than for the other three reactions. As Figure S11 shows, this is what we observed. To our surprise, both DNA and RNA complementary targets photobleached more slowly when crosslinked to a probe than when on their own. At the irradiation time used for the crosslinking experiments (60 s), the non-crosslinked complementary DNA target reaction was 13% less fluorescent than the crosslinked complementary target, and this difference increased over time after the yield saturated. The difference was even larger for RNA, at 33% after 60 s.

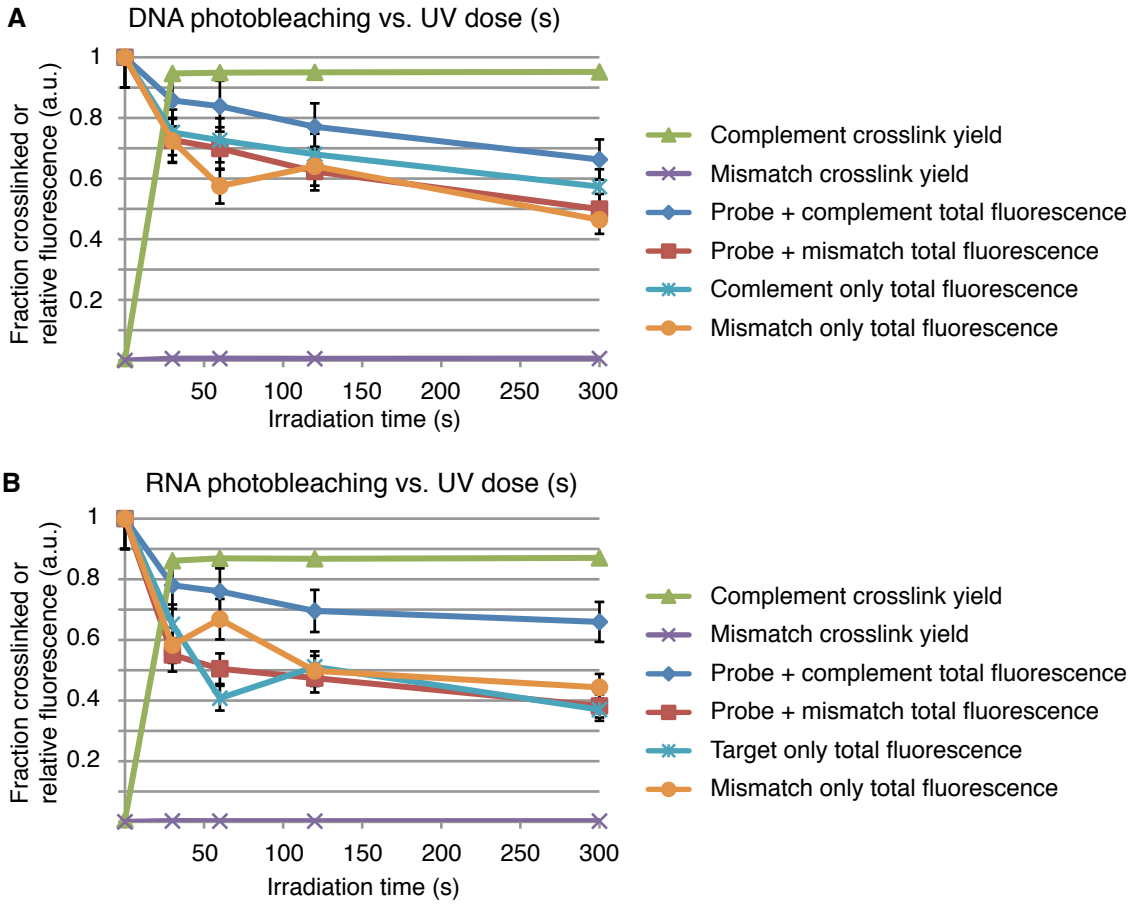


Figure S11. Normalized crosslink yield and normalized total fluorescence vs irradiation time (365 nm LED) for annealed prob/target reactions and for targets alone. Total fluorescence data are normalized to non-irradiated values and are plotted with 10% relative errors for visual reference. 29-nt complementary and 2-nt mismatch DNA and RNA targets (C and M3 of Table S1) and 5-nt toehold DNA and RNA SC probes (Table S2).

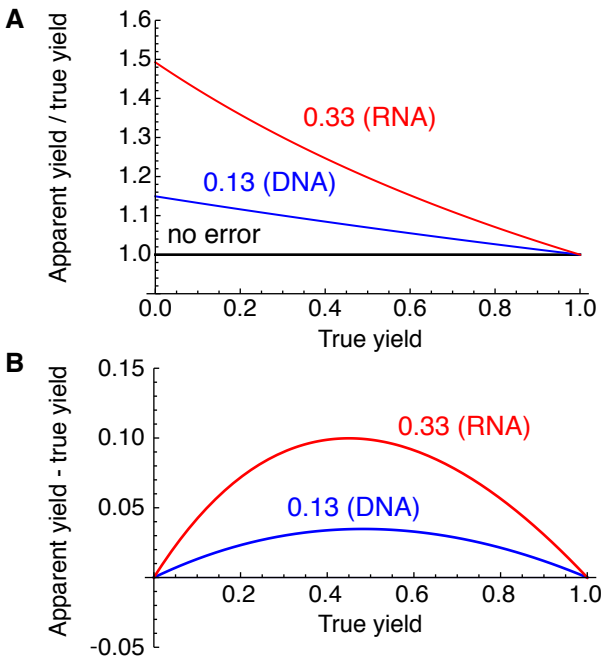


Figure S12. Effect of differential photobleaching on measured crosslinking yields. a) Relative error as a function of true yield for representative values of differential photobleaching (13 % for DNA, 33% for RNA). b) Absolute error as a function of true yield.

The effect of the observed differential photobleaching is to undercount non-crosslinked molecules and therefore to overestimate the crosslinking yield. If (as in our measurements) there are only two fluorescent species of interest, it is straightforward to calculate the effect of differential photobleaching, and this is plotted in Figure S12 based on the quantification approach:

$$\begin{aligned}\text{apparent crosslink yield} &= \frac{\text{measured crosslink fluorescence}}{\text{measured crosslink fluorescence} + \text{measured uncrosslinked target fluorescence}} \\ &= \frac{\text{true crosslink yield}}{\text{true crosslink yield} + (1-\alpha)\text{true uncrosslinked target yield}}\end{aligned}$$

using $\alpha = 0.13$ for DNA and $\alpha = 0.33$ for RNA. The relative error due to this source is largest for small crosslinking yields, while the absolute error peaks at moderate yields. If the photobleaching characteristics of each probe/target pair were known, it would be possible to correct for this effect. Instead, we simply note that: 1) since our probes capture complementary targets with high yield (the regime in which fluorescent measurements overestimate capture yield to the least extent) and capture mismatched targets with low yield (the regime in which fluorescent measurements overestimate capture yield to the greatest extent), the differential photobleaching effect means that our measured discrimination ratios are underestimates of true probe performance, and 2) for DNA probes, the maximum effect is comparable to the observed experiment-to-experiment variation.

S6 Additional hybridization and capture data

The gel data shown in the article are representative gels showing the irradiated reactions only. Here, in support of Figures 3–7, we also show non-irradiated controls, contrast-enhanced Cy5 gels to show mismatch crosslinking, and native (non-denaturing) polyacrylamide gels to show hybridization (see Section S5.1 for methods). Two common features are noticeable in the data. The first is the relationship between crosslinking (seen in the denaturing gels) and hybridization (seen in the native gels). As discussed in the main text, in order to be maximally selective against mismatches, it is necessary for the probe and complementary target to hybridize with only marginal stability, resulting in smeared signals in the native gels due to transient hybridization during migration. Hybridization with the complement becomes more stable as the probe toehold is lengthened (Figures S21–S24), but hybridization and crosslinking to mismatch targets also begin to appear at the longest toehold lengths. Second, examination of the (green) DNA probe bands in the (+) and (-) UV lanes shows a difference in mobility in the denaturing gel but not in the native gel (e.g. Figs S13ac). We believe that this reflects covalent internal crosslinking of the probe hairpins in a closed conformation. When SC probes are synthesized using photoactive crosslinkers, irradiation creates a 'snapshot' of hybridization at that time: unbound probes are inert to subsequent hybridization or crosslinking. This enables kinetic resolution for applications and also ensures that the crosslinkers are unlikely to participate in unwanted side reactions. We expect that similar behavior occurs for the RNA probes as well, but no mobility shift is evident for these probes. Here, we also display the full gel for Figure 8 including mRNA bands.

S6.1 Full gels for 2-nt mismatch studies of Figure 3

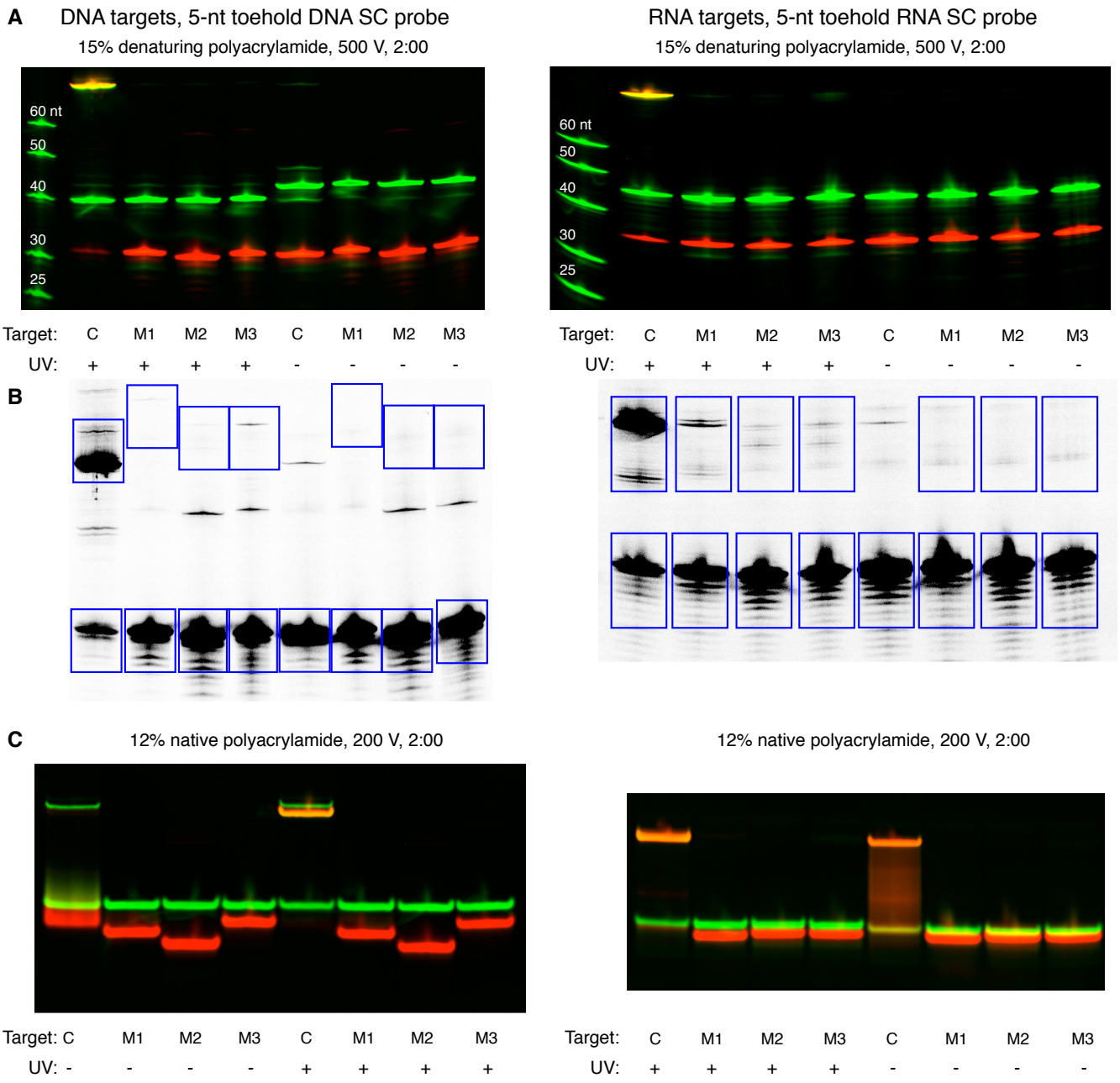


Figure S13. Full gels for 2-nt mismatch studies of Figure 3. 29-nt complementary and 2-nt mismatch DNA and RNA targets (Table S1) and 5-nt toehold DNA and RNA SC probes (Table S2). a) Denaturing gel showing DNA (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). b) Cy5 image, contrast optimized for low-yield crosslinking. Blue rectangles show representative integration areas. c) Native gel analysis of the same reactions.

S6.2 Lack of selectivity for unstructured probe with 2-nt mismatch targets (cf. SC probe data of Fig. 3a)

Figure S14 shows that a 23-nt single-stranded DNA probe (the 5-nt SC probe without loop or shielding section) does not discriminate 2-nt mismatches at RT. Crosslink yields for mismatch targets are 97%, comparable to those obtained for complementary targets with SC probes with $>$ 5-nt toeholds.

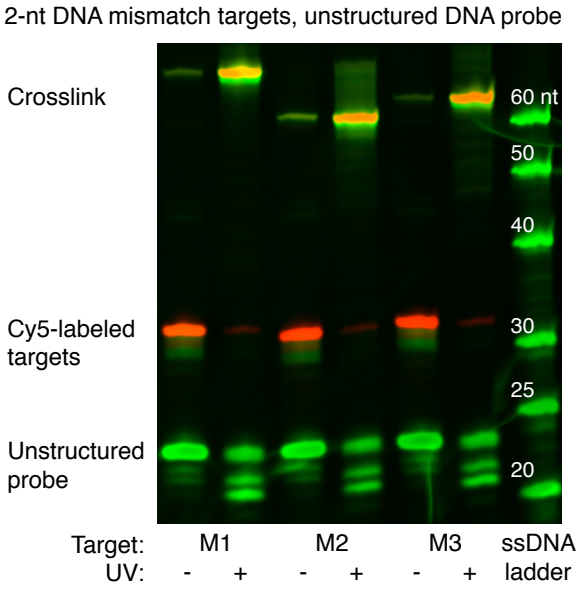


Figure S14. An unstructured covalent DNA probe does not discriminate 2-nt mismatches. 29-nt DNA targets with 2-nt mismatches (M1, M2, M3 of Table S1) and unstructured covalent 23-nt DNA probe (Table S3), cross linked as for SC probes. All three mismatch targets are crosslinked with high yield.

S6.3 Full gels for SNP studies of Figure 4

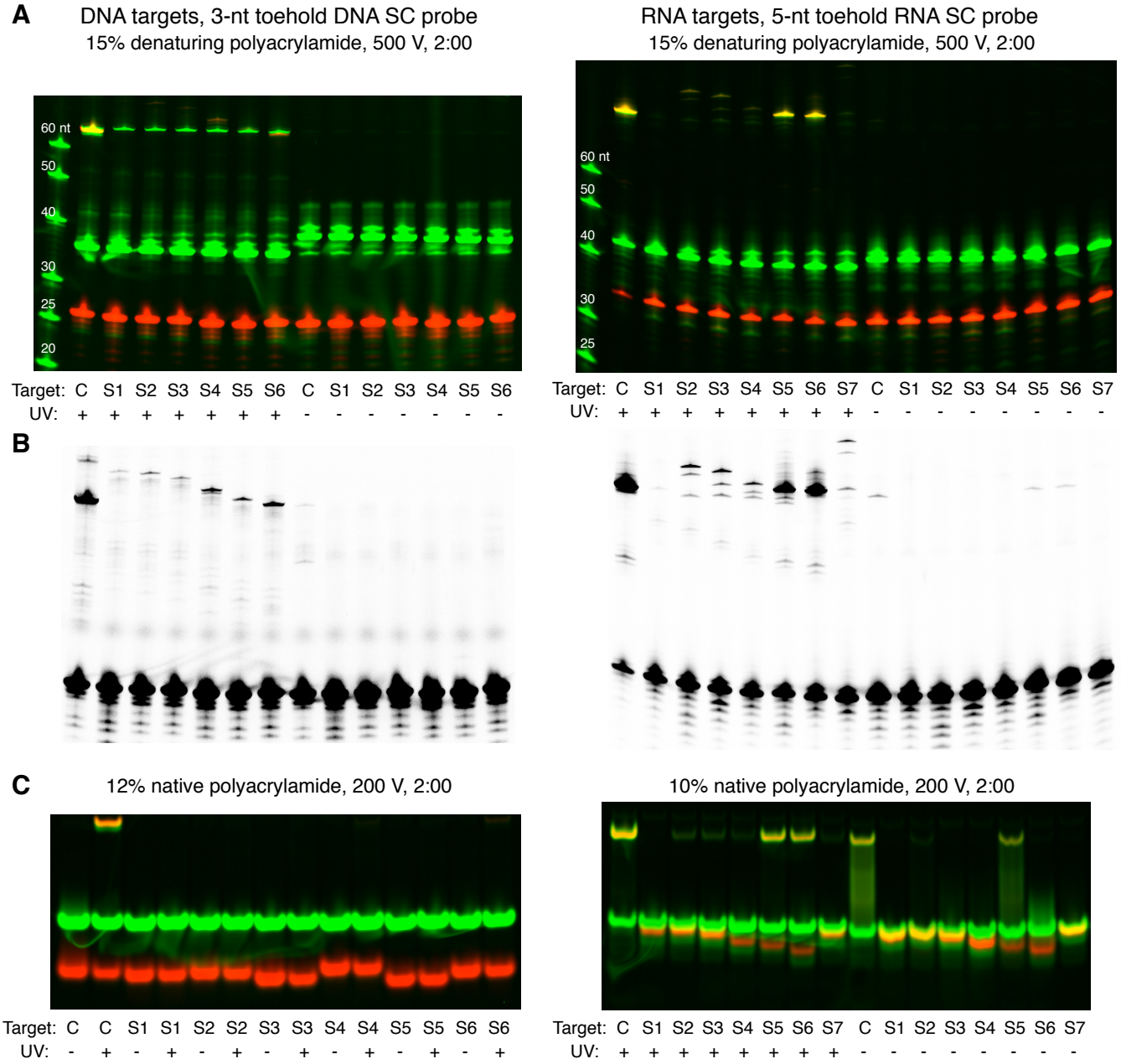


Figure S15. Full gels for SNP studies of Figure 4. Complementary and SNP 23-nt DNA and 29-nt RNA targets (Table S1). 3-nt toehold DNA SC probe and 5-nt toehold RNA SC probe (Table S2). a) Denaturing gel showing DNA (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). b) Cy5 image, contrast optimized for low-yield crosslinking. c) Native gel analysis of the same reactions.

S6.4 Enhanced discrimination of RNA wobble SNPs (cf. Fig. 4)

Shortening the toehold from 5-nt to 4-nt allows an RNA SC probe to discriminate even A to G substitutions (which lead to G·U wobble pairs that are nearly isoenergetic with the A·U pairs they replace). In the context of the thermodynamic hypothesis advanced in the text, this is achieved by reducing the reaction free energy for probe/complement hybridization so that it falls on the steepest part of the yield curve. By operating at room temperature, we avoid reducing the size of the discrimination energy gap, maximizing the achievable discrimination ratio for wobble SNPs.

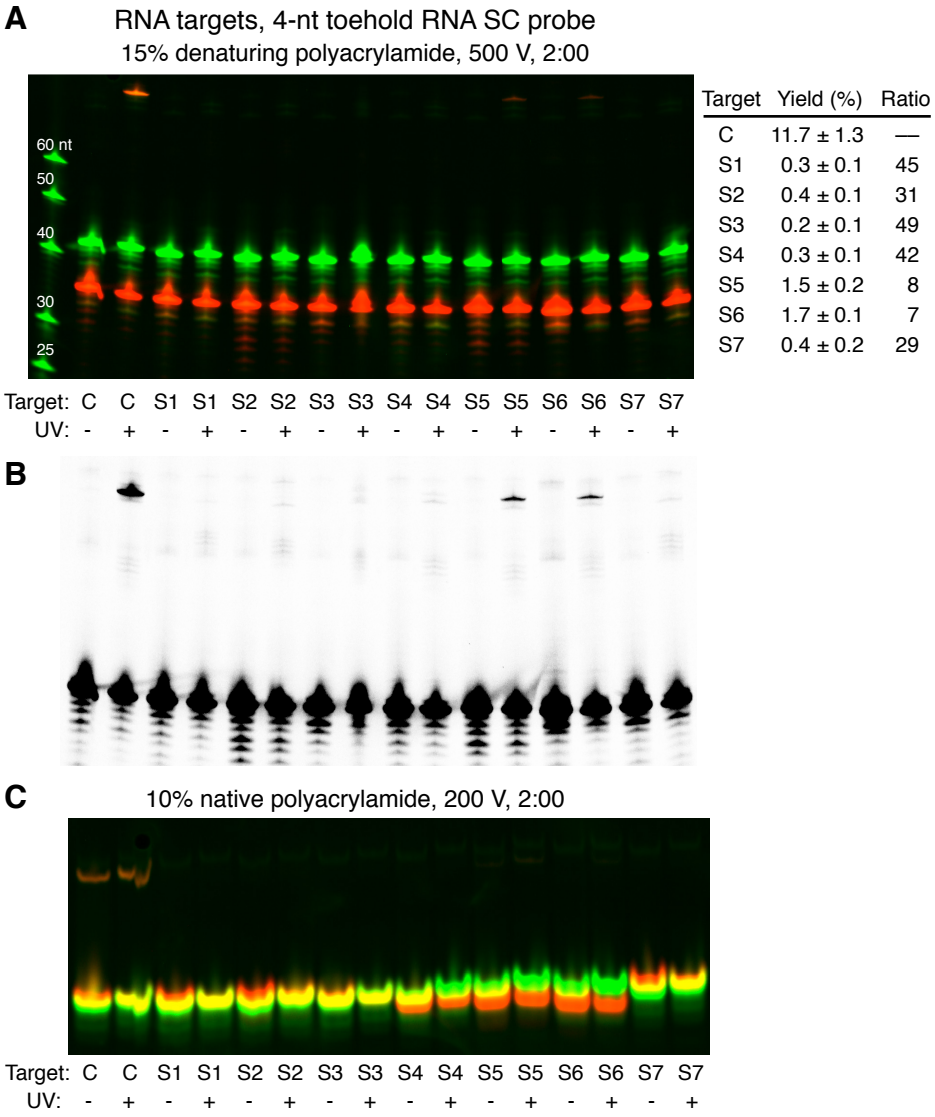


Figure S16. Enhanced discrimination of RNA wobble SNPs using a 4-nt toehold RNA SC probe (cf. the 5-nt toehold data of Fig. 4). 29-nt complementary and SNP RNA targets (Table S1) and 4-nt toehold RNA SC probe (Table S2). a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). Crosslinking yields reported as mean ± standard deviation, $N = 3$. Compared to a 5-nt toehold RNA probe (full gels in Figure S15), capture yield with the complementary target is $\approx 7\times$ lower, but all SNP targets are well-discriminated, including A to G substitutions (S5 and S6). b) Cy5 image, contrast optimized for low-yield crosslinking. c) Native gel analysis of the same reactions.

S6.5 Probe toehold length studies for mismatch pools of Figure 5

Figure 5 demonstrates selective capture of a complementary DNA or RNA target in a large pool of mismatched DNA or RNA targets (24 1-nt mismatches and 72 2-nt mismatches) with each target at approximately the same concentration. Figures S17 and S18 show (-UV) controls and probe toehold length studies for DNA and RNA.

For the RNA study, the covalently captured complementary target appears as a double band in the denaturing gel (Fig. 5). Further examination reveals that the complementary target appears as two bands independent of UV irradiation and the absence/presence of the mismatch target pool (Fig. S19). These results, in combination with the fact that the band splitting appears only in the 6-FAM channel (6-FAM-labeled complementary targets, Cy5-labeled mismatched targets), indicate that the phenomenon is unrelated to the mismatched targets. We observe only one band at the higher target concentration (1.8 μM) used for Figs 3 and 4. Integration of the fluorescence signal over both bands for the 4- and 5-nt toehold RNA probes of Fig. S18 gives similar crosslinking yields to those observed at the higher target concentration (Figs S16 and 3).

We are unsure of the cause of the band splitting when the RNA SC probe is at 3 μM and the complementary RNA target is at 30 nM instead of 1.8 μM. Denaturing PAGE analysis of the 30 nM target with an unstructured RNA covalent probe containing the same hybridization sequence and crosslinker as the hairpin SC probe shows only a single band (Fig. S20), suggesting that the splitting of the target bands in the denaturing gel is unrelated to the ^{CNV}K crosslinker, and may be a gel artifact resulting from incomplete denaturation of the hairpin SC probes. HPLC analysis of the reactions with the RNA SC probe shows only single target peaks before and after UV irradiation at both target concentrations, although the signal-to-background ratio of this assay is limited at the lower target concentration (data not shown).

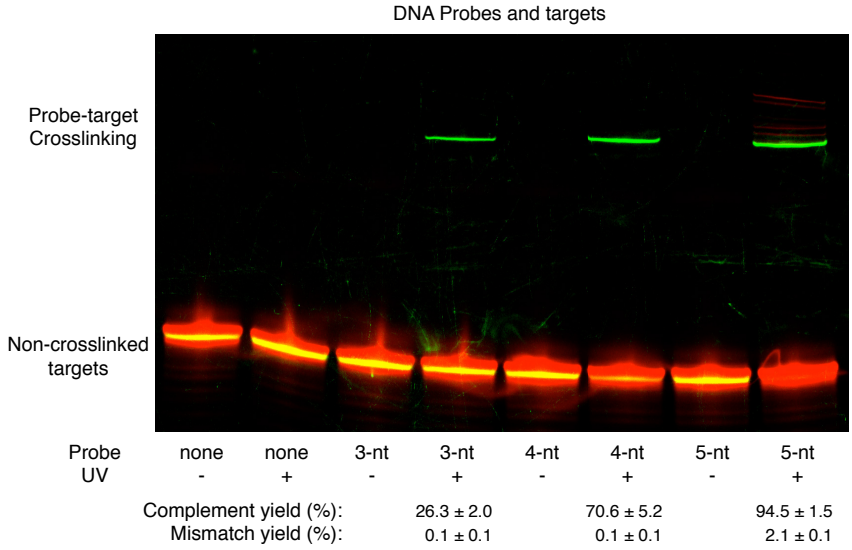


Figure S17. DNA SC probe toehold length studies for the DNA mismatch pool of Figure 5. 29-nt complementary DNA target and pool of 96 1-nt and 2-nt mismatched DNA targets (C and P1–P8 of Table S1) with 3–5 nt toehold DNA SC probes (Table S2). Each target at ≈30 nM, probe at 3 μM. Denaturing polyacrylamide gel. Fluorescent channels: 6-FAM-labeled complementary target (green), Cy5-labeled mismatched targets (red; contrast enhanced to show mismatch bands). Crosslinking yield (mean ± standard deviation, $N = 3$). Within the detection limit of the assay, mismatches are rejected quantitatively by the 3- and 4-nt toehold probes; the 5-nt toehold probe captures mismatches at low yield and the complementary target with near-quantitative yield.

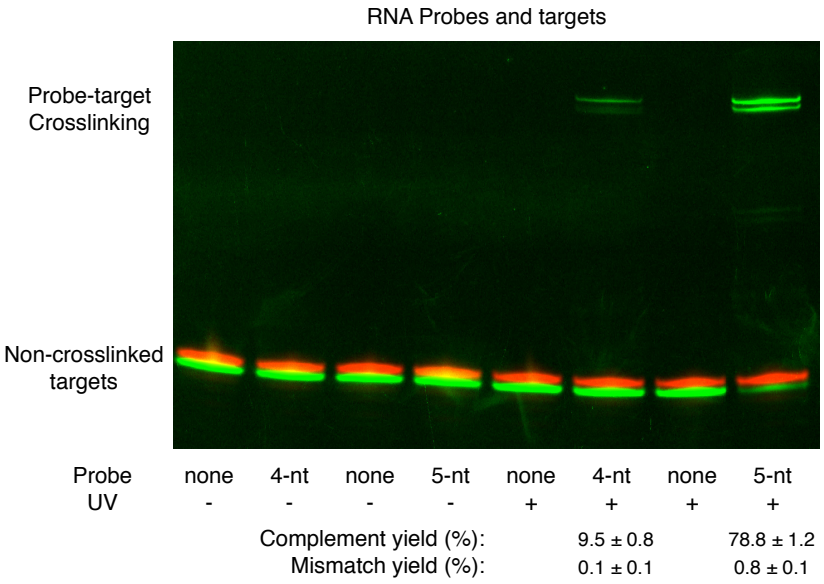


Figure S18. RNA SC probe toehold length studies for the RNA mismatch pool of Figure 5. 29-nt complementary RNA target and pool of 96 1-nt and 2-nt mismatched RNA targets (C and P1–P8 of Table S1) with 4–5 nt toehold RNA SC probes (Table S2). Each target at ≈30 nM, probe at 3 μM. Denaturing polyacrylamide gel. Fluorescent channels: 6-FAM-labeled complementary target (green), Cy5-labeled mismatched targets (red). Crosslinking yield (mean ± standard deviation, $N = 3$).

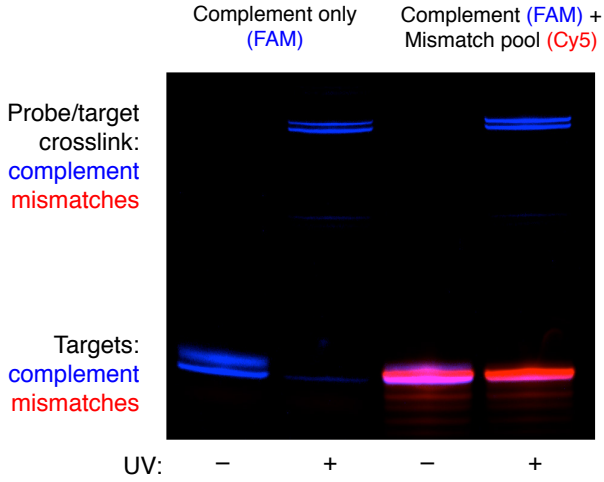


Figure S19. For the RNA mismatch pool study of Figure 5, two bands are observed for the complementary RNA target independent of the absence/presence of the mismatch target pool and independent of UV irradiation. 29-nt complementary RNA target in the absence/presence of a pool of 96 1-nt and 2-nt mismatched RNA targets (C and P1–P8 of Table S1) with 5 nt toehold RNA SC probe (Table S2). Each target at ≈30 nM, probe at 3 μM. Denaturing polyacrylamide gel. Fluorescent channels: 6-FAM-labeled complementary target (blue), Cy5-labeled mismatched targets (red).

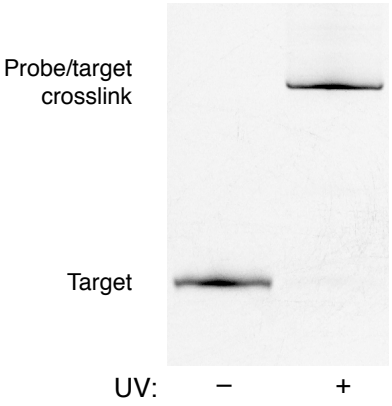


Figure S20. For the complementary RNA target used in the mismatch pool study of Figure 5, only one band is observed using an unstructured probe containing the same hybridization sequence and crosslinker as the hairpin SC probe of Figure S19. 29-nt complementary RNA target (C of Table S1) with 23-nt unstructured RNA covalent probe (Table S3). Target at 30 nM, probe at 3 μM. Denaturing polyacrylamide gel. Fluorescent channel: 6-FAM-labeled complementary target.

S6.6 DNA probe toehold length studies of Figure 6

Figure 6a examines complementary and 2-nt mismatch DNA targets (29-nt targets C and M1 of Fig. 2) using 2–9 nt DNA SC probes. Experiments are performed using either isothermal hybridization (Figs S21 and S22) or annealing (Figs S23 and S24) prior to crosslinking. Capture yields are tabulated in Table S10. Figures 6b and 6C examine complementary and SNP DNA targets (23-nt targets C and S1–S6 of Fig. 4) using 3–5 nt toehold DNA SC probes (Figs S15, S25 and S26). Figure 6c also examines 2-nt mismatch DNA targets (23-nt targets M4–M8 of Table S1) using 2–9 nt toehold DNA SC probes (Figs S27 and S28).

Note that the capture yields for the 23-nt complementary target of Figs S27 and S28 are higher than those observed using the 29-nt complementary target (which includes a 6-nt linker between the fluorophore and the 5' end of the target sequence) of Figures 3 and 6A (see the 4- and 5-nt toehold data of Figure S21 and Table S10). This may be due to stabilization of the 23-nt probe/target duplex by stacking of the Cy5 dye at the terminus.[12] Alternatively, the effect could result from unexpected destabilization of the 29-nt probe/target duplex by the presence of the 6-nt spacer.[13] For consistency, Figure 6c employs 23-nt complementary, SNP and 2-nt mismatch targets.

Table S10. Crosslink yield (mean $\pm$ standard deviation, $N = 3$; $N = 1$ for 2-nt probes) for complementary and 2-nt mismatch DNA targets with 2–9 nt toehold DNA SC probes, using either isothermal or annealed hybridization (gels in Figures S21–S24)

Toehold length (nt)	Complementary target isothermal yield (%)	2-nt mismatch target isothermal yield (%)	Complementary target annealed yield (%)	2-nt mismatch target annealed yield (%)
2	3.8	< 0.1	9.2	< 0.1
3	14.4 ± 3.5	0.1 ± 0.1	28.2 ± 6.2	0.1 ± 0.1
4	50.0 ± 4.5	< 0.1	67.4 ± 5.0	0.1 ± 0.1
5	92.9 ± 3.5	0.1 ± 0.1	94.8 ± 1.0	0.8 ± 1.1
6	97.9 ± 0.6	0.3 ± 0.1	97.4 ± 0.8	1.2 ± 0.4
7	96.6 ± 2.6	12.3 ± 1.6	95.8 ± 3.7	26.0 ± 4.3
8	98.2 ± 0.4	32.5 ± 5.9	99.0 ± 1.0	59.4 ± 6.7
9	97.6 ± 1.1	47.9 ± 1.1	99.0 ± 0.6	79.1 ± 1.7

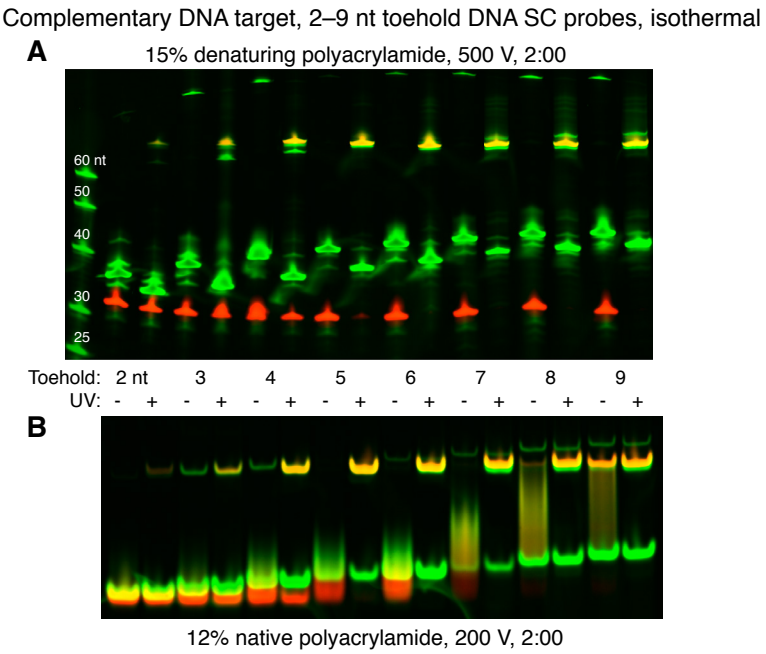


Figure S21. DNA SC probe toehold length studies of Figure 6a: complementary target, isothermal hybridization. 29-nt complementary DNA target (C in Table S1) and 2–9 nt toehold DNA SC probes (Table S2). a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). b) Native gel analysis of the same reactions. Slow-migrating green-only bands are likely probe dimers crosslinked once (- UV) or twice (+ UV) to each other.

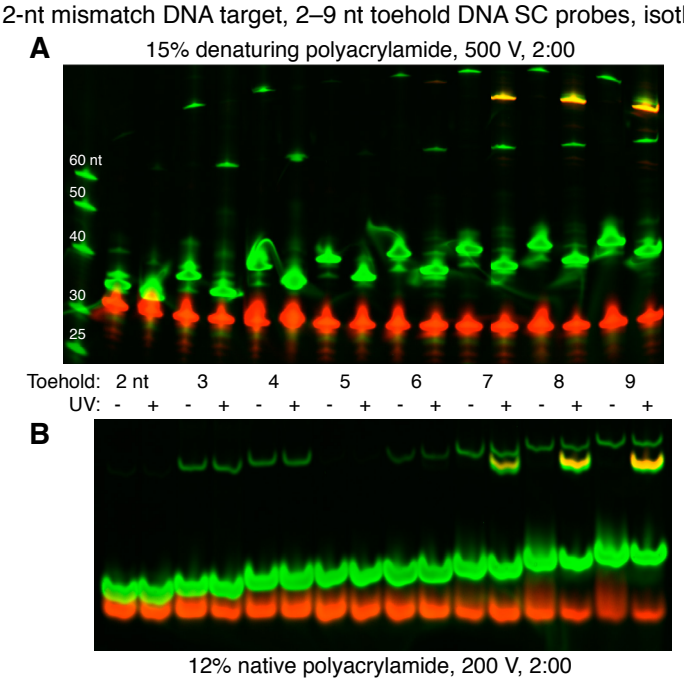


Figure S22. DNA SC probe toehold length studies of Figure 6A: 2-nt mismatch target, isothermal hybridization. 29-nt DNA target with 2-nt mismatch (M1 in Table S1) and 2–9 nt toehold DNA SC probes (Table S2). a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). b) Native gel analysis of the same reactions.

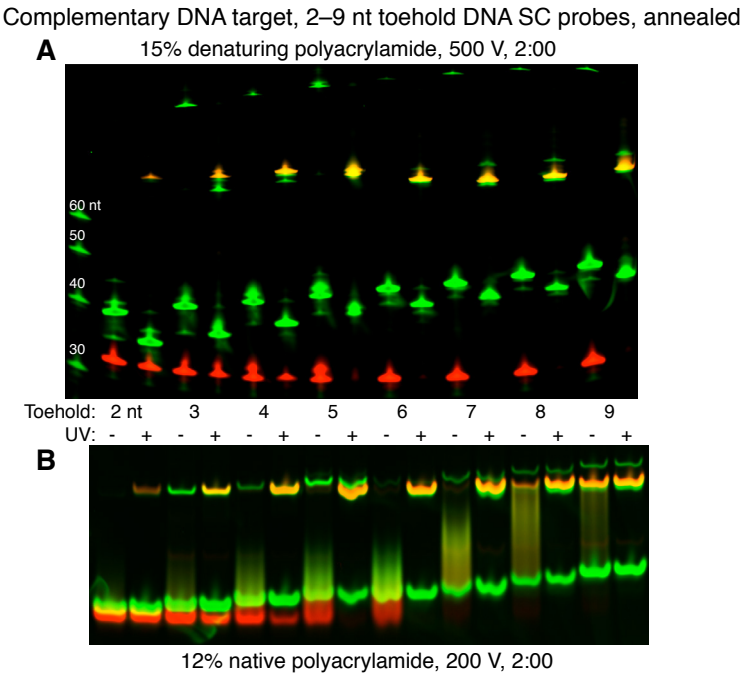


Figure S23. DNA SC probe toehold length studies of Figure 6A: complementary target, annealed hybridization. 29-nt complementary DNA target (C in Table S1) and 2–9 nt toehold DNA SC probes (Table S2). a) Denaturing gel showing DNA (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). b) Native gel analysis of the same reactions.

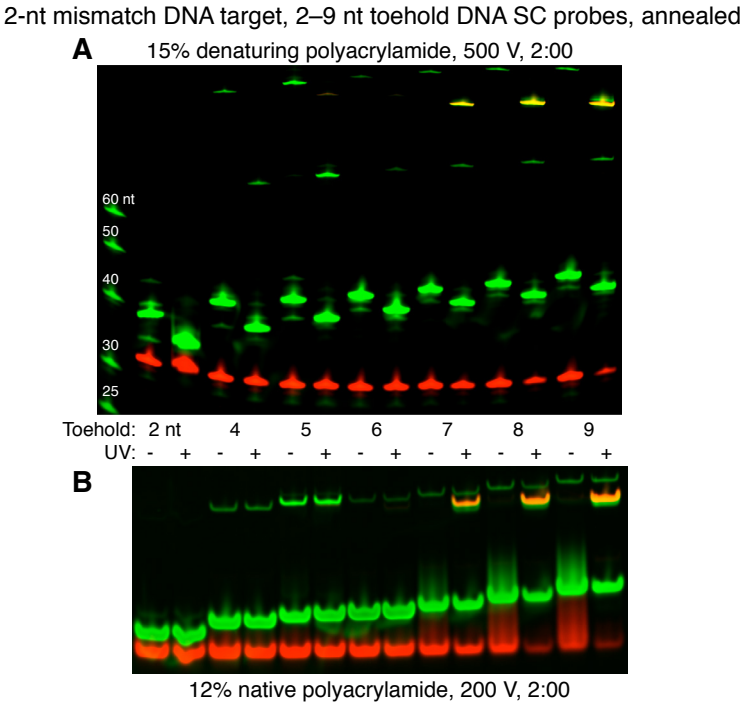
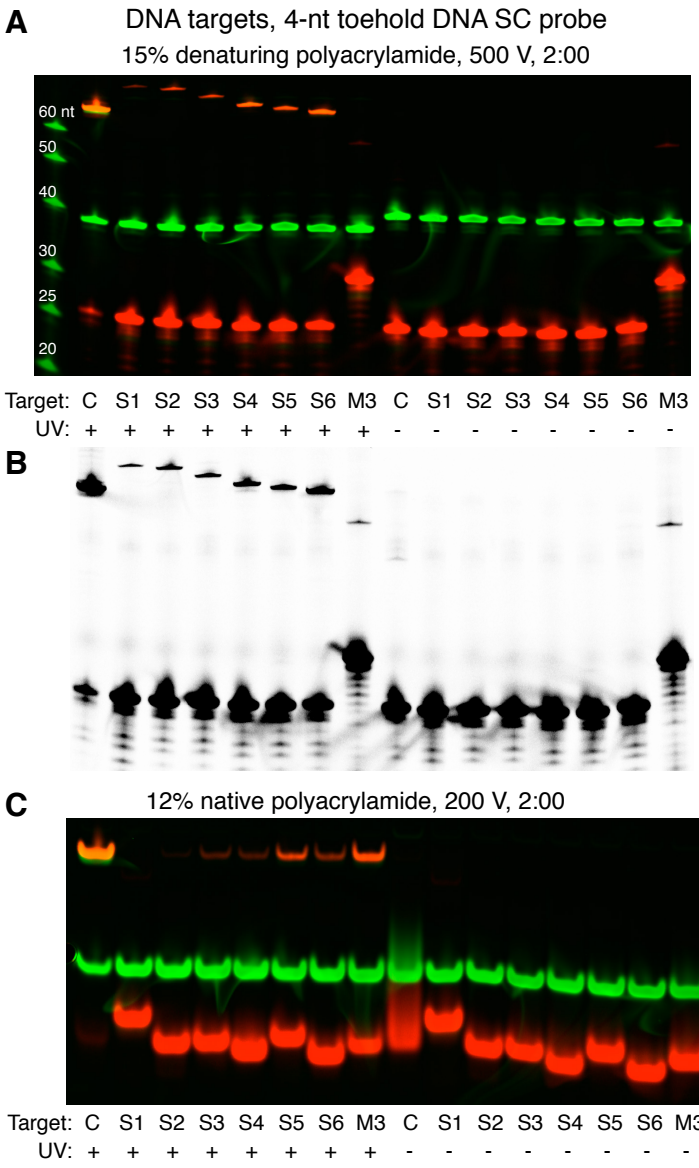


Figure S24. DNA SC probe toehold length studies of Figure 6A: 2-nt mismatch target, annealed hybridization. 29-nt DNA target with 2-nt mismatch (M1 in Table S1) and 2–9 nt toehold DNA SC probes (Table S2). Note: 3-nt toehold probe omitted from these gels. a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). b) Native gel analysis of the same reactions.



Target	Yield (%)	Ratio
C	95.1 ± 1.1	—
S1	1.0 ± 0.1	95
S2	3.6 ± 0.1	26
S3	3.2 ± 0.6	30
S4	12.9 ± 1.4	7
S5	5.3 ± 0.5	18
S6	29.2 ± 1.4	3

Figure S25. DNA SNP mismatch discrimination with 4-nt toehold DNA SC probe (for Figure 6b). 23-nt complementary and SNP mismatch DNA targets (C and S1–S6 in Table S1) and 4-nt toehold DNA SC probe (Table S2). M3 is a 2-nt mismatch (Figure 3) included for reference. a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). Capture yields reported as mean ± standard deviation, $N = 3$. Compared to 5 nt and 3 nt toehold results, yield with complementary target is still high, and some SNPs are discriminated effectively. b) Cy5 image, contrast optimized for low-yield crosslinking. c) Native gel analysis of the same reactions.

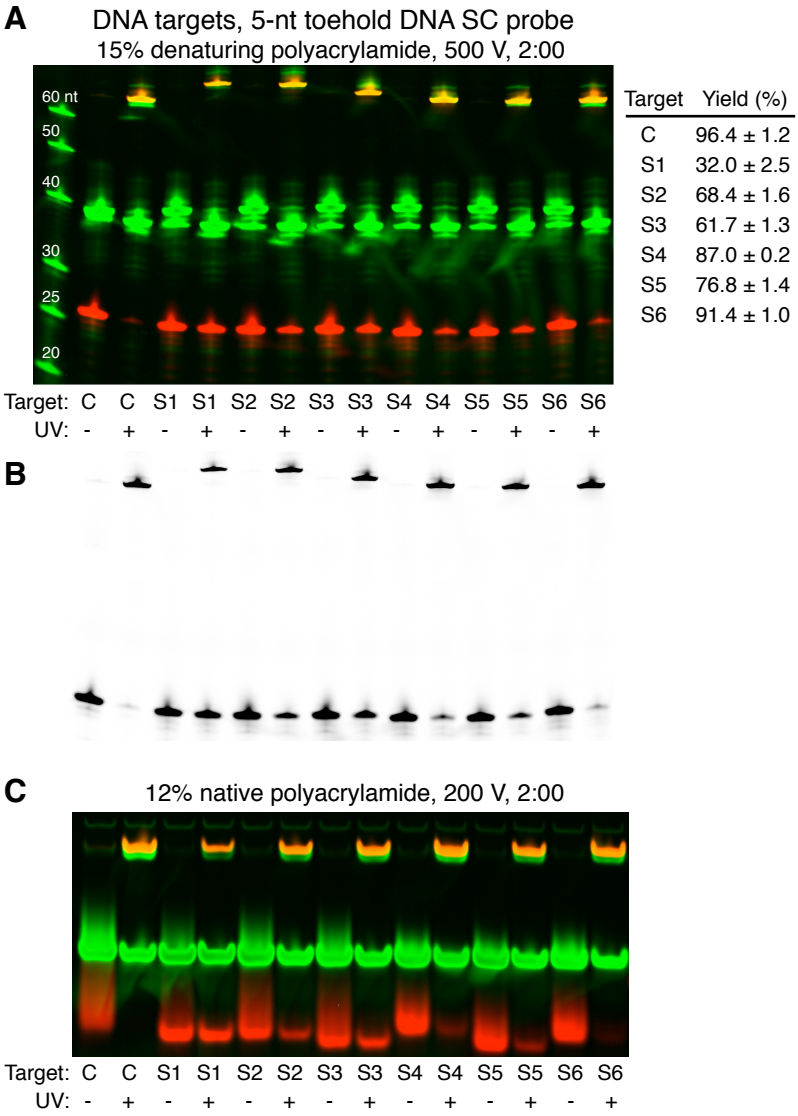
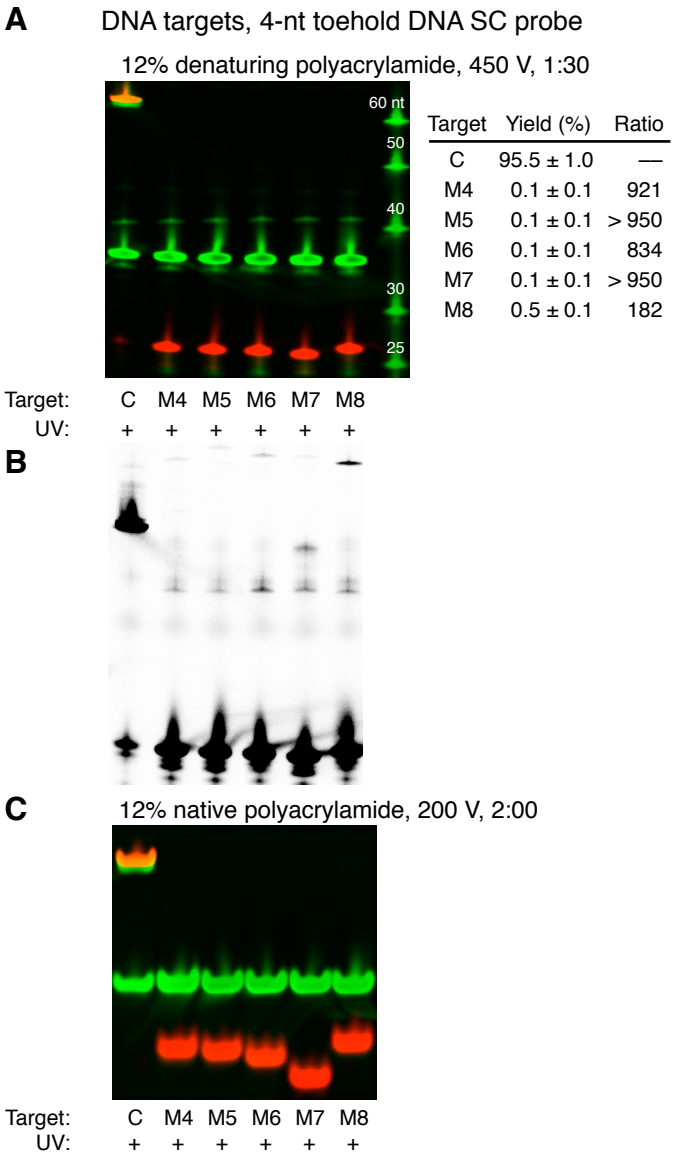


Figure S26. DNA SNP mismatch discrimination with 5-nt toehold DNA SC probe (for Figure 6b). 23-nt complementary and SNP mismatch DNA targets (C and S1–S6 in Table S1) and 5-nt toehold DNA SC probe (Table S2). a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). Capture yields reported as mean ± standard deviation, $N = 3$. Yield with complementary target is high, but SNPs are not discriminated effectively. b) Cy5 image (targets only). c) Native gel analysis of the same reactions.



Target	Yield (%)	Ratio
C	95.5 ± 1.0	—
M4	0.1 ± 0.1	921
M5	0.1 ± 0.1	> 950
M6	0.1 ± 0.1	834
M7	0.1 ± 0.1	> 950
M8	0.5 ± 0.1	182

Figure S27. DNA 2-nt mismatch discrimination with 4-nt toehold DNA SC probe (for Figure 6c). 23-nt complementary and 2-nt mismatch DNA targets (C and M4–M8 in Table S1) and 4-nt toehold DNA SC probe (Table S2). a) Denaturing gel analysis. (+ UV) reactions only. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). Capture yields reported as mean ± standard deviation, $N = 3$. b) Cy5 image (targets only), contrast adjusted to show low yield crosslinking. c) Native gel analysis of the same reactions.

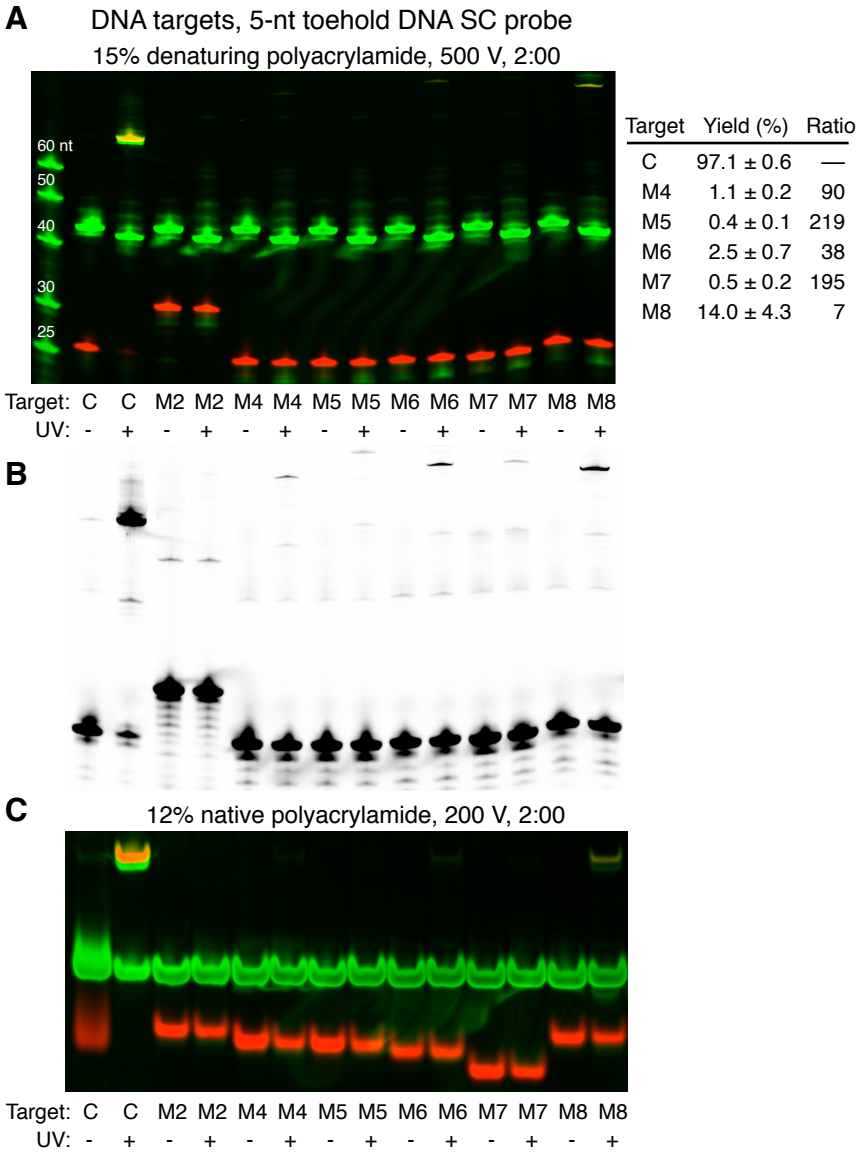


Figure S28. DNA 2-nt mismatch discrimination with 5-nt toehold DNA SC probe (for Figure 6c). 23-nt complementary and 2-nt mismatch DNA targets (C and M4–M8 in Table S1) and 5-nt toehold DNA SC probe (Table S2). a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). Capture yields reported as mean ± standard deviation, $N = 3$. b) Cy5 image (targets only). c) Native gel analysis of the same reactions.

S6.7 RNA probe toehold length studies (cf. the DNA studies of Fig. 6a)

A limited set of toehold studies was also performed with RNA SC probes (Figs S29 and S30). The measured yields appear to lie on a sigmoidal curve, as expected (Fig. S31). Unlike the DNA data of Figure 6a, little difference is observed between isothermal and annealed hybridization prior to crosslinking.

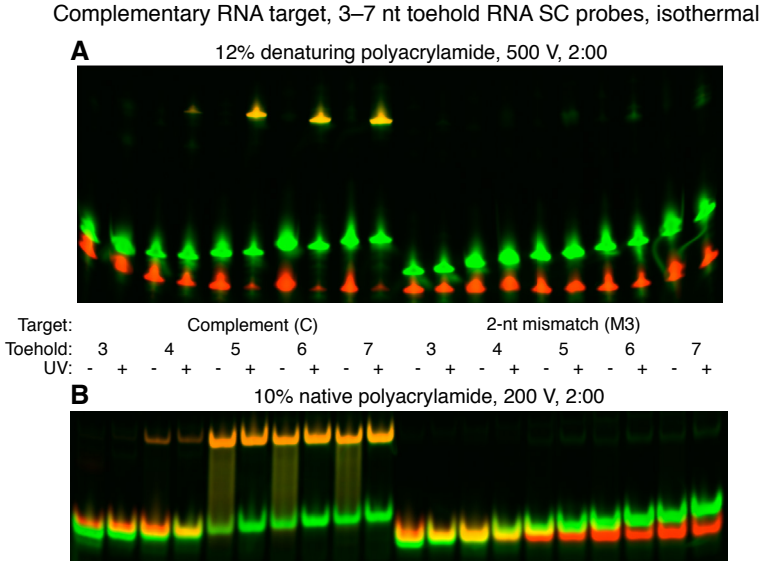


Figure S29. RNA SC probe toehold length studies with isothermal hybridization. 29-nt complementary and 2-nt mismatch RNA targets (C and M3 in Table S1) and 3–7 nt toehold RNA SC probe (Table S2). a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). b) Native gel analysis of the same reactions.

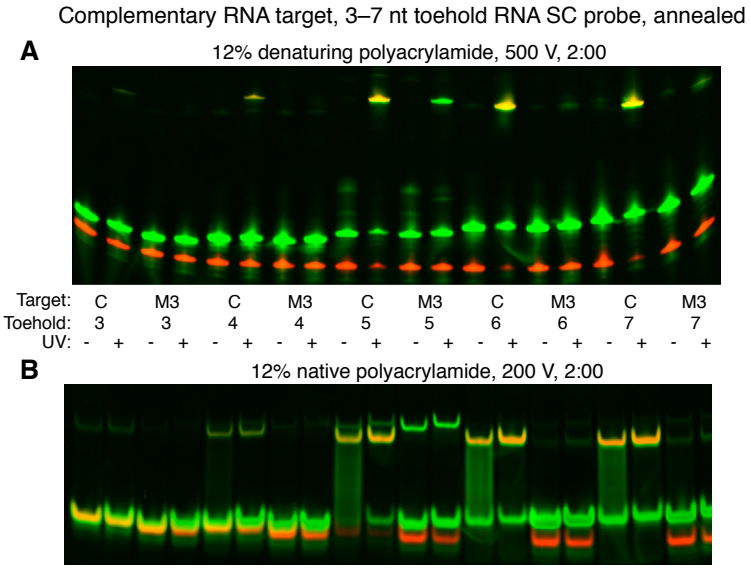


Figure S30. RNA SC probe toehold length studies with annealed hybridization. 29-nt complementary and 2-nt mismatch RNA targets (C and M3 in Table S1) and 3–7 nt toehold RNA SC probe (Table S2). a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). b) Native gel analysis of the same reactions.

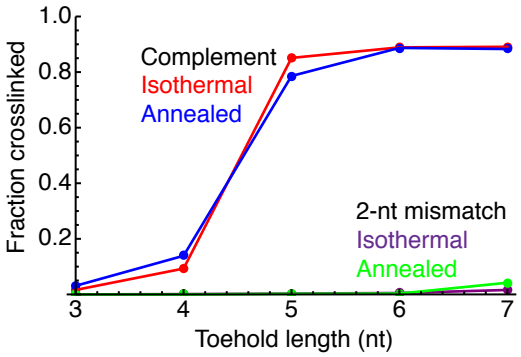


Figure S31. RNA SC probe toehold length studies with isothermal and annealed hybridization ($N = 1$). 29-nt complementary and 2-nt mismatch RNA targets (C and M3 in Table S1) and 3–7 nt toehold RNA SC probe (Table S2). Gels are shown in Figs S29 and S30. Isothermal and annealed hybridization give similar results over the range of toeholds studied.

S6.8 Full gels for fluorescent protein sequence discrimination studies of Figure 7

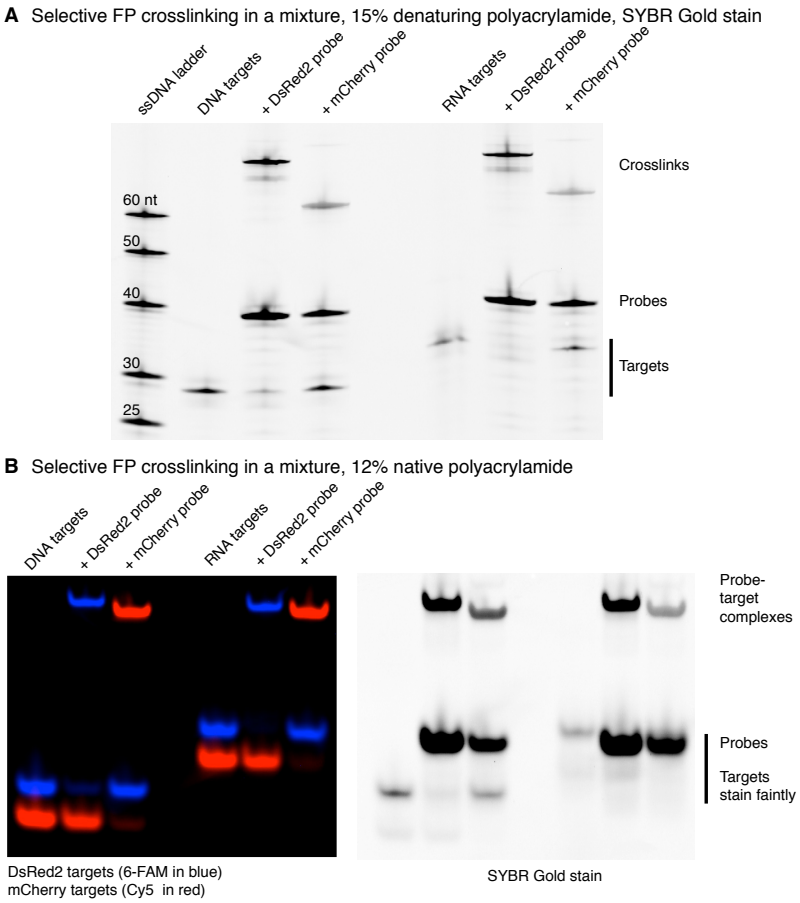


Figure S32. Full gels for fluorescent protein sequence discrimination studies of Figure 7. 26-nt DsRed2 and mCherry DNA targets (Table S1) and 6-nt toehold DsRed2 and mCherry DNA probes (Table S2). a) SYBR Gold stain of denaturing crosslink assay. b) Native gel analysis of the same reactions. Note that the single-stranded targets are not well-stained by SYBR Gold. The SYBR signal is partially quenched by the Cy5-labeled mCherry targets.

S6.9 Full gels for full-length mRNA crosslinking studies of Figure 8

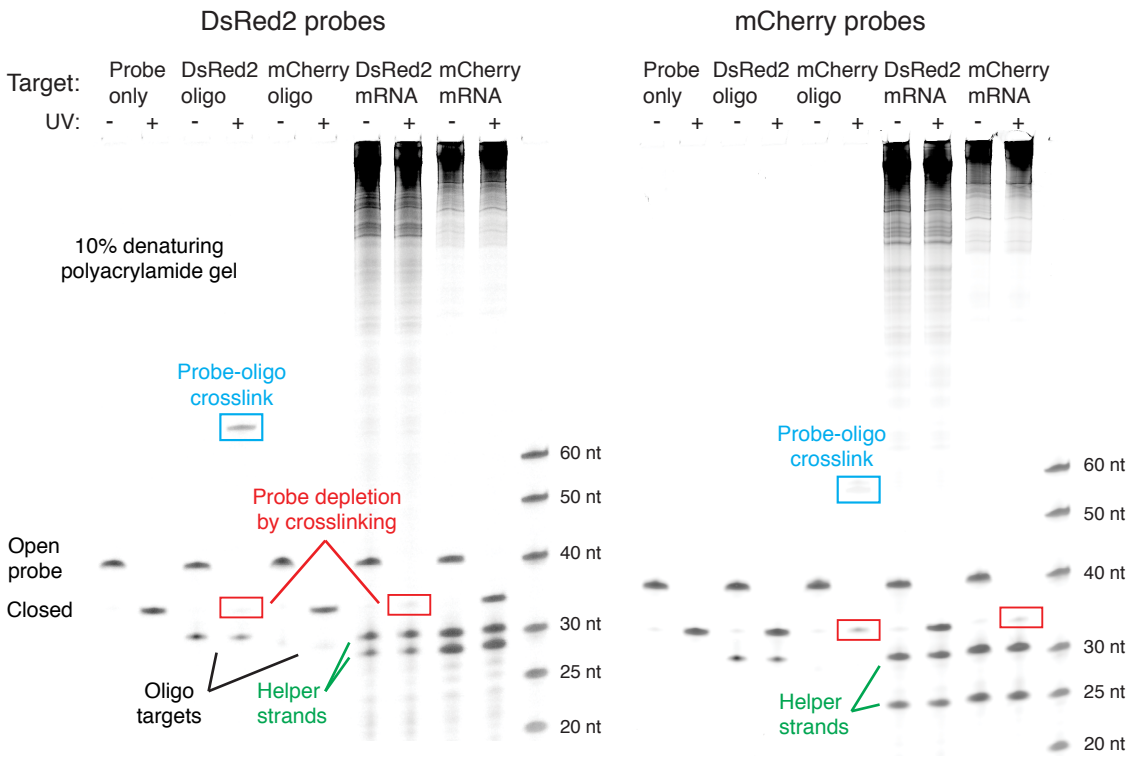
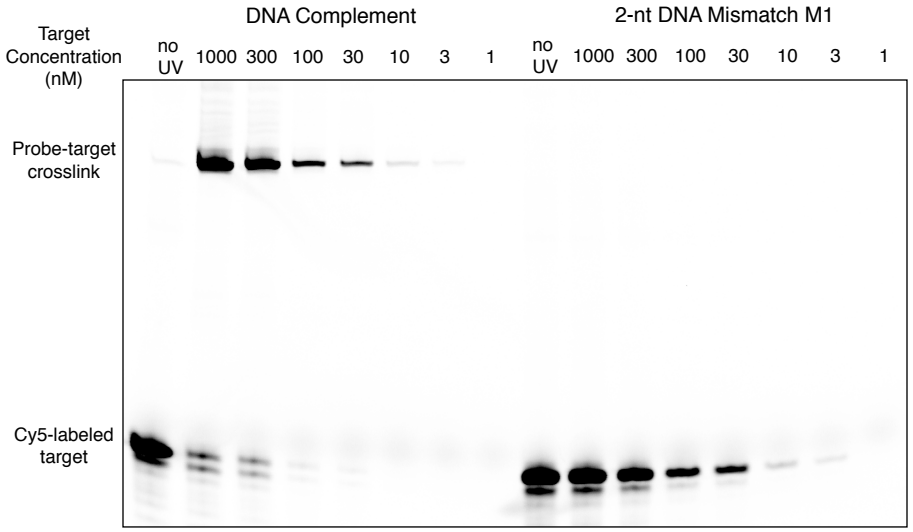


Figure S33. Full gels for full-length mRNA crosslinking studies of Figure 8. 26-nt DsRed2 and mCherry RNA oligo targets, 817-nt DsRed2 and 1367-nt mCherry mRNA targets (Table S1), and 7-nt toehold DsRed2 and mCherry DNA probes (Table S2). For mRNA experiments, the DsRed2 probe is accompanied by DsRed2 helper strands and the mCherry probe is accompanied by mCherry helper strands (Table S4) to combat native secondary structure. Denaturing polyacrylamide gel post-stained with SYBR Gold. The covalently-closed probe band is depleted for complementary RNA oligo and full-length mRNA targets (red boxes). There is no measurable depletion of the covalently-closed probe band for 2-nt mismatched RNA oligo and full-length mRNA targets. The mCherry oligo target is labeled with Cy5, which quenches SYBR fluorescence, resulting in low signal. Contrast has been optimized to show probe and oligo target staining, causing mRNA bands to appear very dark.

S6.10 Sensitive and selective capture as target concentration is decreased

The mechanism studies of Fig. 6 suggest that SC probe capture yields are consistent with equilibrium probe/target hybridization yields. Hence, based on the equilibrium calculations of Supplementary Section S2.2, we expect the probe performance (yield and selectivity) to be independent of the target concentration for target concentrations much smaller than the probe concentration. To test this prediction, we measured capture yield for a 5-nt toehold DNA SC probe at 3 μM with 29-nt complementary and 2-nt mismatched targets (c.f. Figure 3) at concentrations from 1 μM to 1 nM. Complement yield and mismatch rejection were both preserved over 3 orders of magnitude in target concentration (Fig. S34). At 1 nM (10 fmol of target), target fluorescence was only marginally distinguishable from background, manifested by the larger variation in measured complement yield. Readout of lower concentrations of captured targets should be possible using post-capture signal amplification. SC probe capture yield is expected to remain constant to arbitrarily low target concentrations, so any sensitivity limitation is expected to derive from target readout, not from target capture.



Concentration (nM)	1000	300	100	30	10	3	1
Complement yield (%)	93.9 ± 0.5	94.3 ± 0.1	94.5 ± 0.5	94.6 ± 1.2	98.5 ± 1.6	95.6 ± 4.4	93.9 ± 7.3
Mismatch yield (%)	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	< 0.1	< 0.1	< 0.1

Figure S34. Effect of target concentration on measured covalent capture yields. 29-nt complementary and 2-nt mismatched Cy-5 labeled DNA targets (C and M1 of Table S1) with 5-nt toehold DNA SC probe (Table S2). Fluorescent scan of a denaturing polyacrylamide gel (Cy5-labeled DNA targets). Contrast is optimized to show the middle of the concentration range (10–100 nM), leading to blooming of the high-concentration target bands. Crosslinking yield (mean ± standard deviation, $N = 3$).

S7 Probe design and capture studies for fluorescent protein sequences

This section describes crosslinking studies carried out with two pairs of target sites drawn from the mCherry and DsRed2 sequences[14,15] (pair 1: mCherry and DsRed2 targets used in Figure 7, pair 2: $\text{mCherry}_{\text{alt}}$ and $\text{DsRed2}_{\text{alt}}$). DNA probes were tested with both DNA and RNA targets (target and probe sequences in Tables S1 and S2).

To design DNA probes for the DNA targets, we began by calculating reaction free energy changes for hybridization of probes with varying toehold lengths using the same method described in Section S2.3; these values are shown in Tables S11 and S12. The equilibrium binding curve (Eq. S3, plotted in Fig. 6c for our experimental concentrations and temperature) suggests that hybridization yield will be saturated for $-\Delta_r G^0$ values of $>$11–12 kcal/mol, and the experimental data of Figure 6c are in reasonable agreement with this prediction. Furthermore, our experience with the probes and targets studied in Figures 3–6 suggested that 2-nt mismatches produce a sufficient discrimination gap ($\Delta\Delta G$) to be efficiently discriminated by probes whose yield with complementary targets is marginally saturated. We therefore surmised that, for effective discrimination of 2-nt mismatches, probe toehold length should be chosen such that $-\Delta_r G^0$ for hybridization to the complement is predicted to be approximately 11–12 kcal/mol. For the mCherry and DsRed2 targets (pair 1), this occurs for 6-nt toehold probes; for the $\text{mCherry}_{\text{alt}}$ and $\text{DsRed2}_{\text{alt}}$ targets (pair 2), this occurs for 5-nt toehold probes (highlighted in Tables S11 and S12). Accordingly, we synthesized mCherry and $\text{mCherry}_{\text{alt}}$ probes with the predicted optimal toehold length $\pm$ 1 nt, and DsRed2 and $\text{DsRed2}_{\text{alt}}$ probes with 6 different toehold lengths in order to also verify the toehold length results of Figure 6a with two additional sequences.

An alternative approach would be to design probes based on the predicted yields for both the complementary and mismatch targets. We chose to design probes based on the complement $\Delta_r G^0$ alone because: a) it does not require a specific mismatch sequence for design, b) the data underlying the nearest-neighbor thermodynamic model parameters are more complete for hybridization of perfect duplexes than for mismatches, leading to higher confidence in the predicted free energies for arbitrary sequences. As will be seen below, our design approach worked well for both the sequences tested.

After synthesis and purification, the probes were then hybridized with each target (separately), irradiated, and the crosslinking yields determined. Note that our gel scanner is not able to spectrally distinguish the FAM label on the DsRed2 targets from SYBR Gold fluorescence, so the gels containing these targets were scanned for fluorescence first, then stained with SYBR Gold and scanned again.

High-yield crosslinking of the complementary target and efficient rejection of the 2-nt mismatch target was observed for both DNA target sites. In agreement with our calculations, optimal discrimination was observed using 6-nt toehold probes for mCherry and DsRed2 DNA targets (Fig. S37), and 5-nt probes for $\text{mCherry}_{\text{alt}}$ and $\text{DsRed2}_{\text{alt}}$ targets (Fig. S43). These DNA studies suggest that the -12 kcal/mol design criterion (reaction free energy for probe/complement hybridization) may be useful for the design of other SC probes.

These DNA probes were also tested with RNA targets without further design. High-yield crosslinking of the complementary target and efficient mismatch discrimination were observed for RNA versions of the mCherry and DsRed2 targets (Fig. S40). For the alternate site, high yield and efficient discrimination were observed with the $\text{DsRed2}_{\text{alt}}$ probe, but the complementary target yield was much lower for the $\text{mCherry}_{\text{alt}}$ probe (Fig. S46). Examination of the native gels suggests that this is due to poor hybridization, and NUPACK indeed predicts that this target has significantly more secondary structure than either the other DsRed2 target or the DNA target with the same sequence. Despite the lower yield with the $\text{mCherry}_{\text{alt}}$ target, a high discrimination ratio was still observed.

S7.1 Calculated reaction free energy changes

Probe	Toehold (nt)	DsRed2 target $\Delta_r G^0$ (kcal/mol)	mCherry target $\Delta_r G^0$ (kcal/mol)
DsRed2	4	−10.0	−6.7
DsRed2	5	−10.8	−7.5
DsRed2	6	−12.3	−8.9
DsRed2	7	−14.7	−11.4
DsRed2	8	−15.6	−12.3
DsRed2	9	−16.4	−13.1
mCherry	5	−7.2	−10.8
mCherry	6	−8.7	−12.2
mCherry	7	−11.2	−14.7

Table S11. Calculated reaction free energy for DsRed2 and mCherry DNA targets and DsRed2 and mCherry DNA SC probes (Tables S1 and S2). Toehold lengths selected for the differential crosslinking experiment of Figure 7 highlighted in blue.

Probe	Toehold	$\text{DsRed2}_{\text{alt}}$ target $\Delta_r G^0$ (kcal/mol)	$\text{mCherry}_{\text{alt}}$ target $\Delta_r G^0$ (kcal/mol)
$\text{DsRed2}_{\text{alt}}$	3	−7.3	−3.8
$\text{DsRed2}_{\text{alt}}$	4	−9.6	−5.9
$\text{DsRed2}_{\text{alt}}$	5	−12.0	−8.2
$\text{DsRed2}_{\text{alt}}$	6	−14.9	−11.1
$\text{DsRed2}_{\text{alt}}$	7	−16.1	−12.3
$\text{DsRed2}_{\text{alt}}$	8	−16.9	−13.1
$\text{mCherry}_{\text{alt}}$	4	−6.0	−9.7
$\text{mCherry}_{\text{alt}}$	5	−8.3	−12.0
$\text{mCherry}_{\text{alt}}$	6	−11.2	−14.9

Table S12. Calculated reaction free energy for $\text{DsRed2}_{\text{alt}}$ and $\text{mCherry}_{\text{alt}}$ DNA targets and $\text{DsRed2}_{\text{alt}}$ and $\text{mCherry}_{\text{alt}}$ DNA SC probes (Tables S1 and S2). Optimal discrimination for DNA targets was obtained using blue-highlighted toehold lengths.

S7.2 DNA DsRed2 and mCherry targets of Figure 7

For DNA targets, probe toehold length studies are shown for DsRed2 probes in Figure S35 and for mCherry probes in Figure S36, with capture yields plotted in Figure S37.

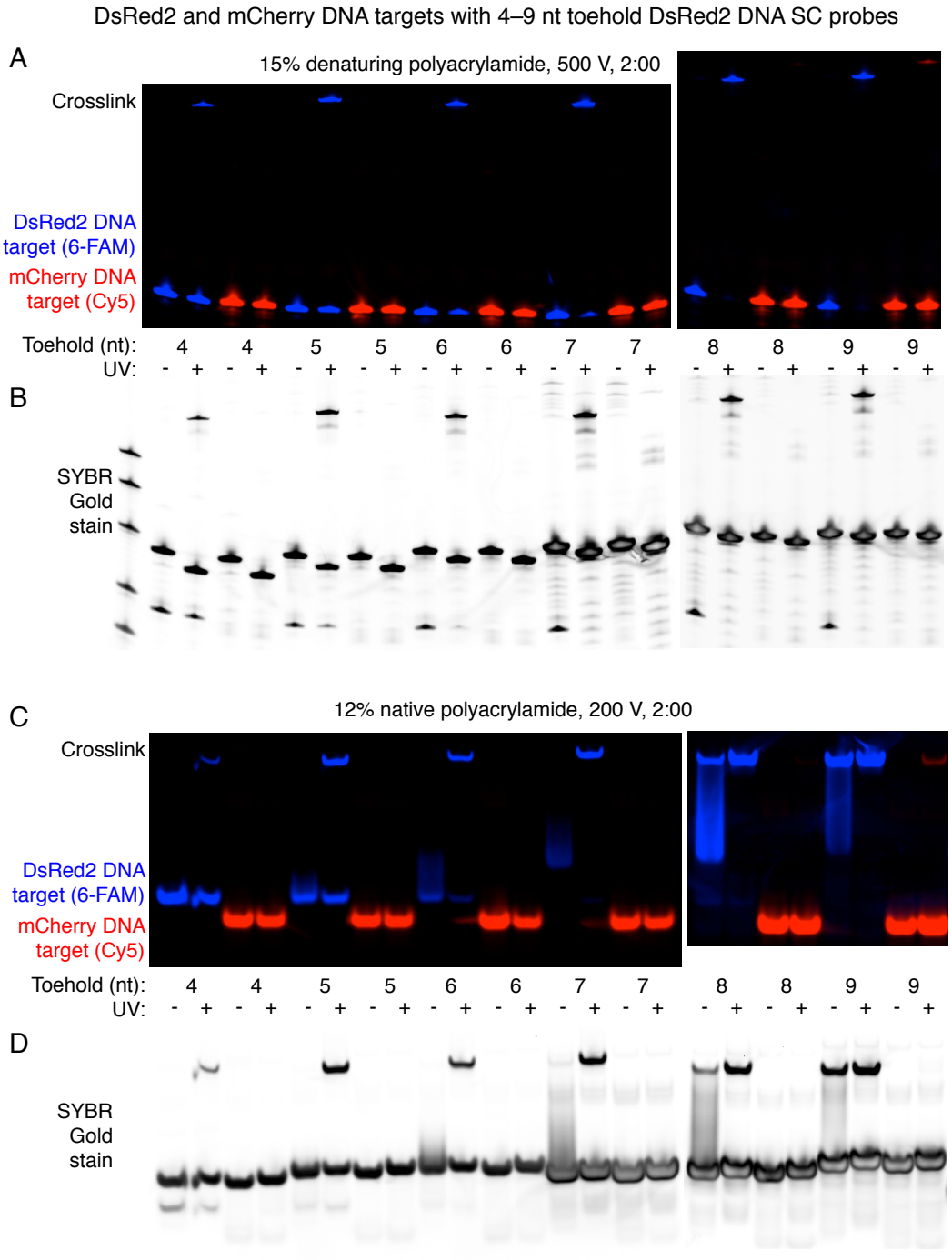


Figure S35. DNA DsRed2 probe toehold length studies with DsRed2 and mCherry DNA targets (for Fig. 7). a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled mCherry targets (red), 6-FAM-labeled DsRed2 targets (blue). b) SYBR Gold post-stain of denaturing gel. c) Native gel analysis of the same reactions. d) SYBR Gold post-strain of native gel.

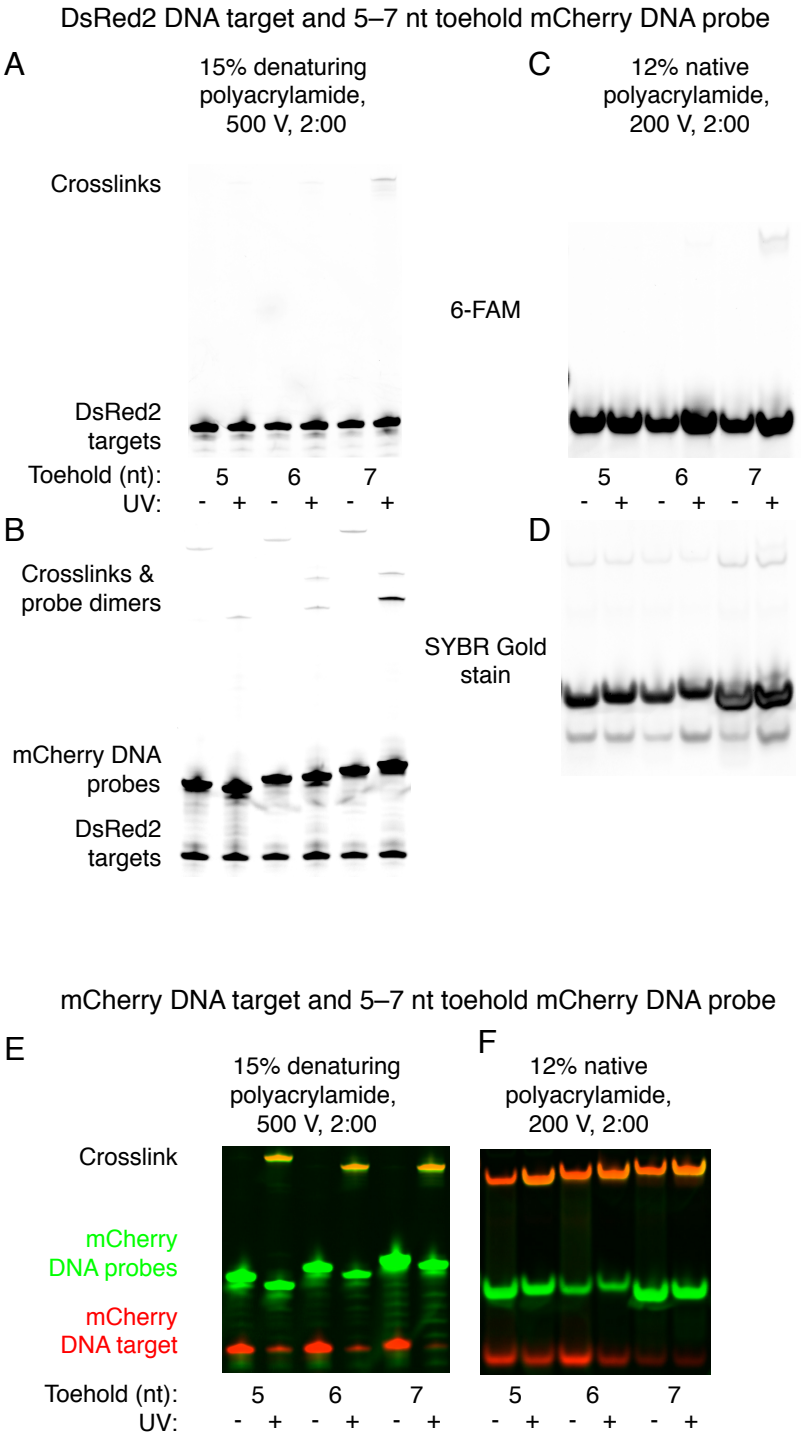


Figure S36. DNA mCherry probe toehold length studies with DsRed2 and mCherry DNA targets (for Fig. 7). a) Denaturing gel showing (+ UV) reactions and (- UV) controls. 6-FAM-labeled targets. b) SYBR Gold post-stain of denaturing gel. c) Native gel analysis of the same reactions. 6-FAM-labeled targets. d) SYBR Gold post-strain of native gel. e) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). f) Native gel analysis of the same reactions.

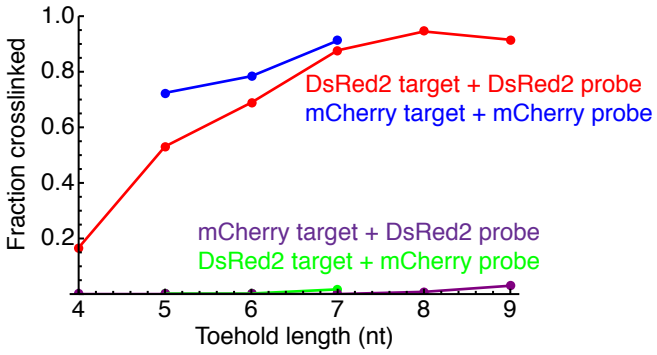


Figure S37. DNA DsRed2 and mCherry probe toehold length studies with DNA DsRed2 and mCherry targets ($N = 1$). Gels are shown in Figs S35 and S36. Yields increase with toehold length. The maximum discrimination ratio is observed for 6-nt toehold DsRed2 and mCherry DNA probes.

S7.3 RNA DsRed2 and mCherry targets of Figure 7

For RNA targets, probe toehold length studies are shown for DsRed2 probes in Figure S38 and for mCherry probes in Figure S39, with capture yields plotted in Figure S40.

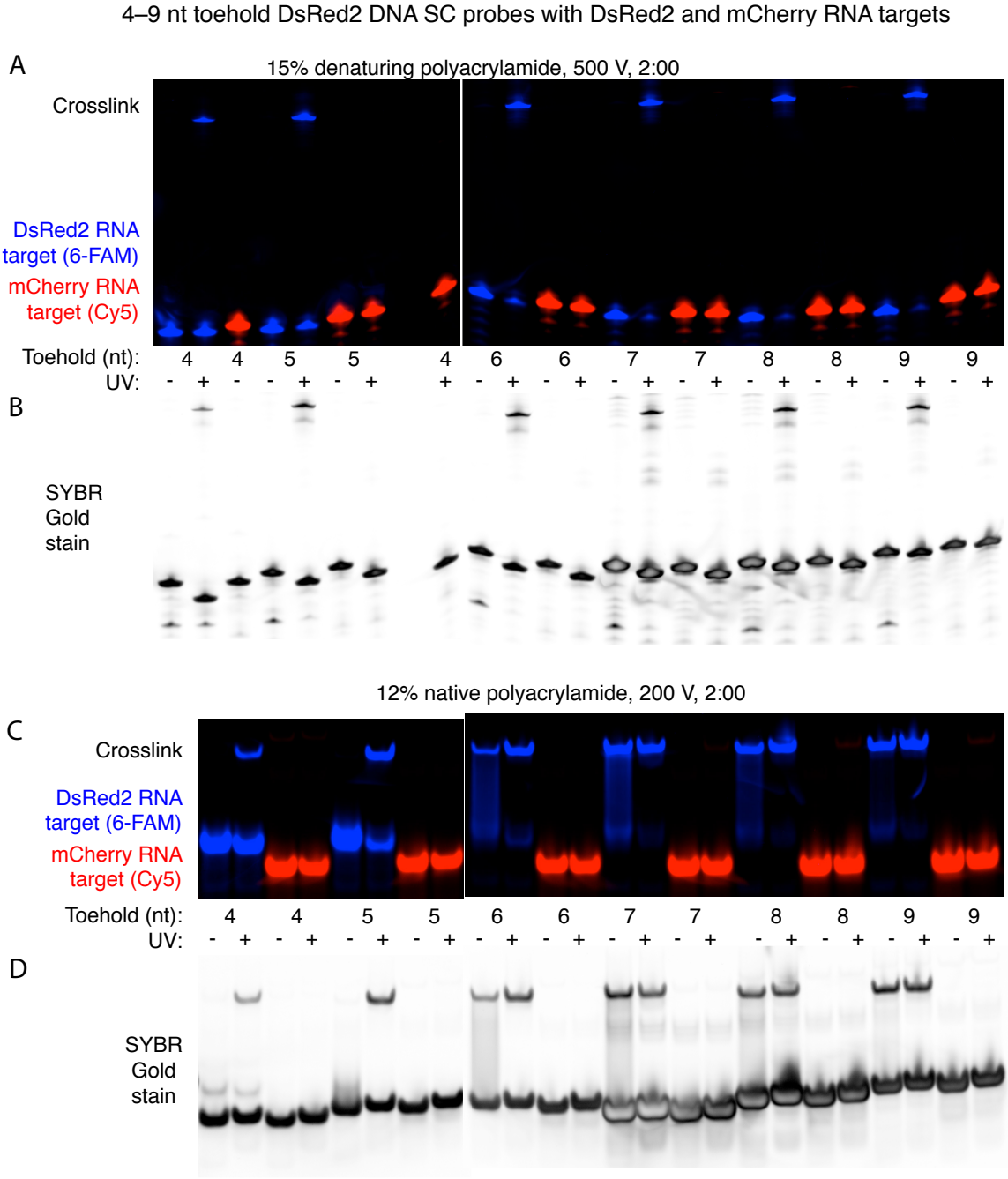


Figure S38. DNA DsRed2 probe toehold length studies with DsRed2 and mCherry RNA targets (for Fig. 7). a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled mCherry targets (red), 6-FAM-labeled DsRed2 targets (blue). b) SYBR Gold post-stain of denaturing gel. c) Native gel analysis of the same reactions. d) SYBR Gold post-strain of native gel.

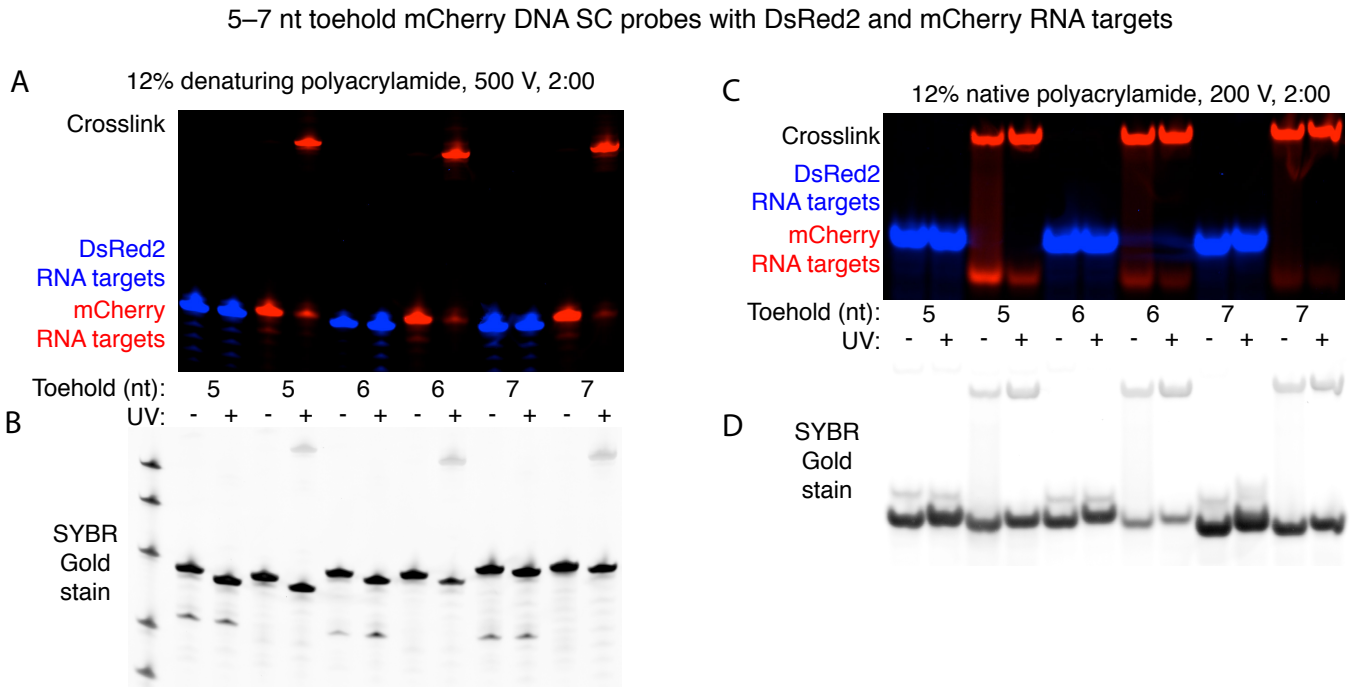


Figure S39. DNA mCherry probe toehold length studies with DsRed2 and mCherry RNA targets (for Fig. 7). a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled mCherry targets (red), 6-FAM-labeled DsRed2 targets (blue). b) SYBR Gold post-stain of denaturing gel. c) Native gel analysis of the same reactions. d) SYBR Gold post-strain of native gel.

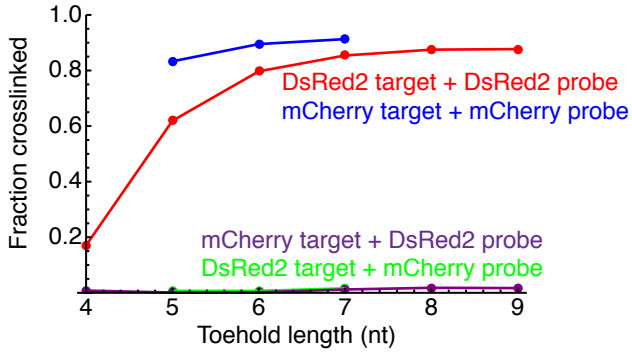


Figure S40. DNA DsRed2 and mCherry probe toehold length studies with RNA DsRed2 and mCherry targets ($N = 1$). Gels are shown in Figs S38 and S39. Yields increase with toehold length. As for DNA targets, the maximum discrimination ratio is observed for 6-nt toehold DsRed2 and mCherry DNA probes.

S7.4 Alternate DNA DsRed2 and mCherry targets

For alternate DNA targets, probe toehold length studies are shown for $DsRed2_{alt}$ probes in Figure S41 and for $mCherry_{alt}$ probes in Figure S42, with capture yields plotted in Figure S43.

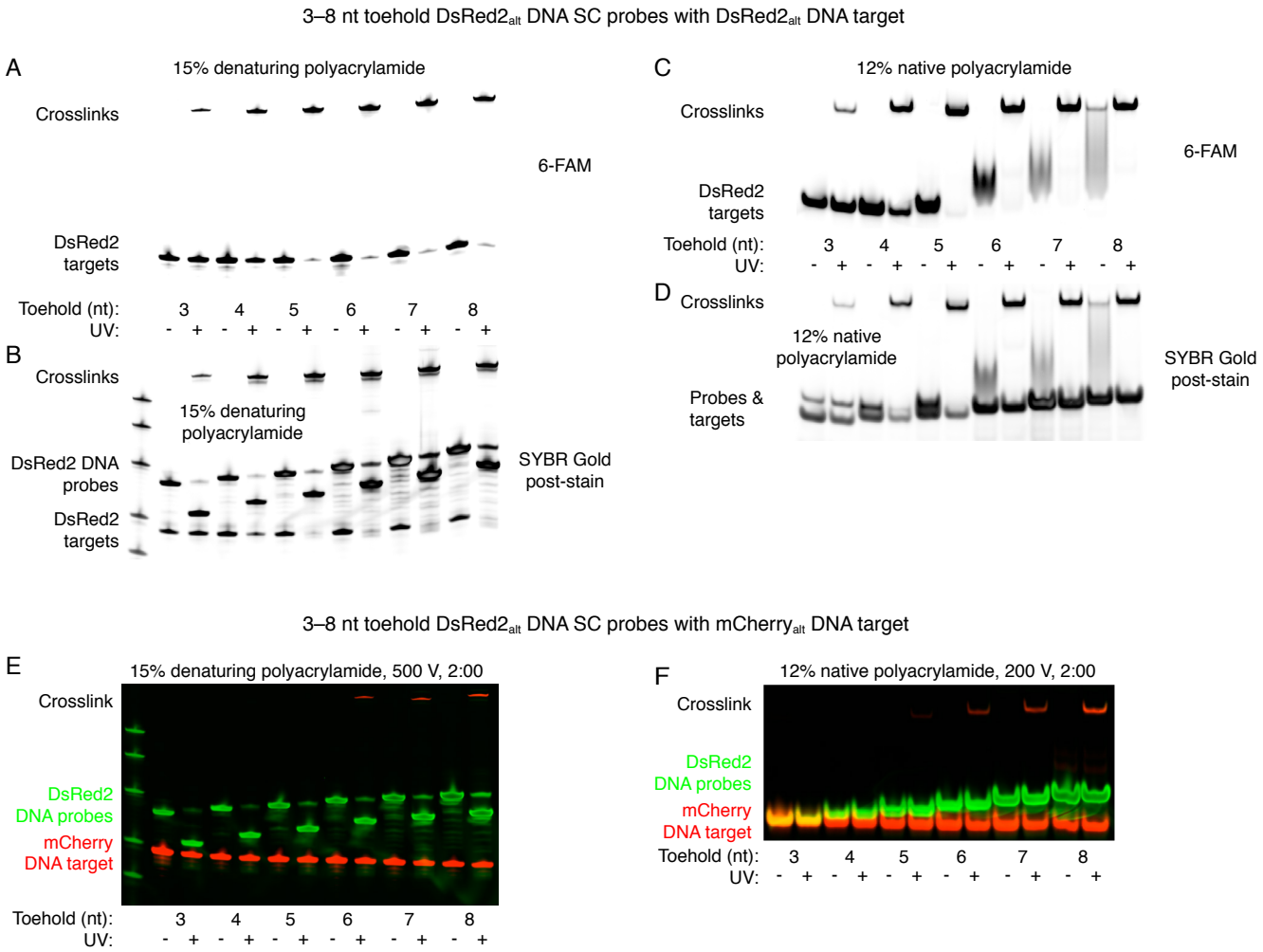


Figure S41. DNA $DsRed2_{alt}$ probe toehold length studies with $DsRed2_{alt}$ and $mCherry_{alt}$ DNA targets. a) Denaturing gel showing (+ UV) reactions and (- UV) controls. 6-FAM-labeled targets. b) SYBR Gold post-stain of denaturing gel. c) Native gel analysis of the same reactions. 6-FAM-labeled targets. d) SYBR Gold post-strain of native gel. e) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). f) Native gel analysis of the same reactions.

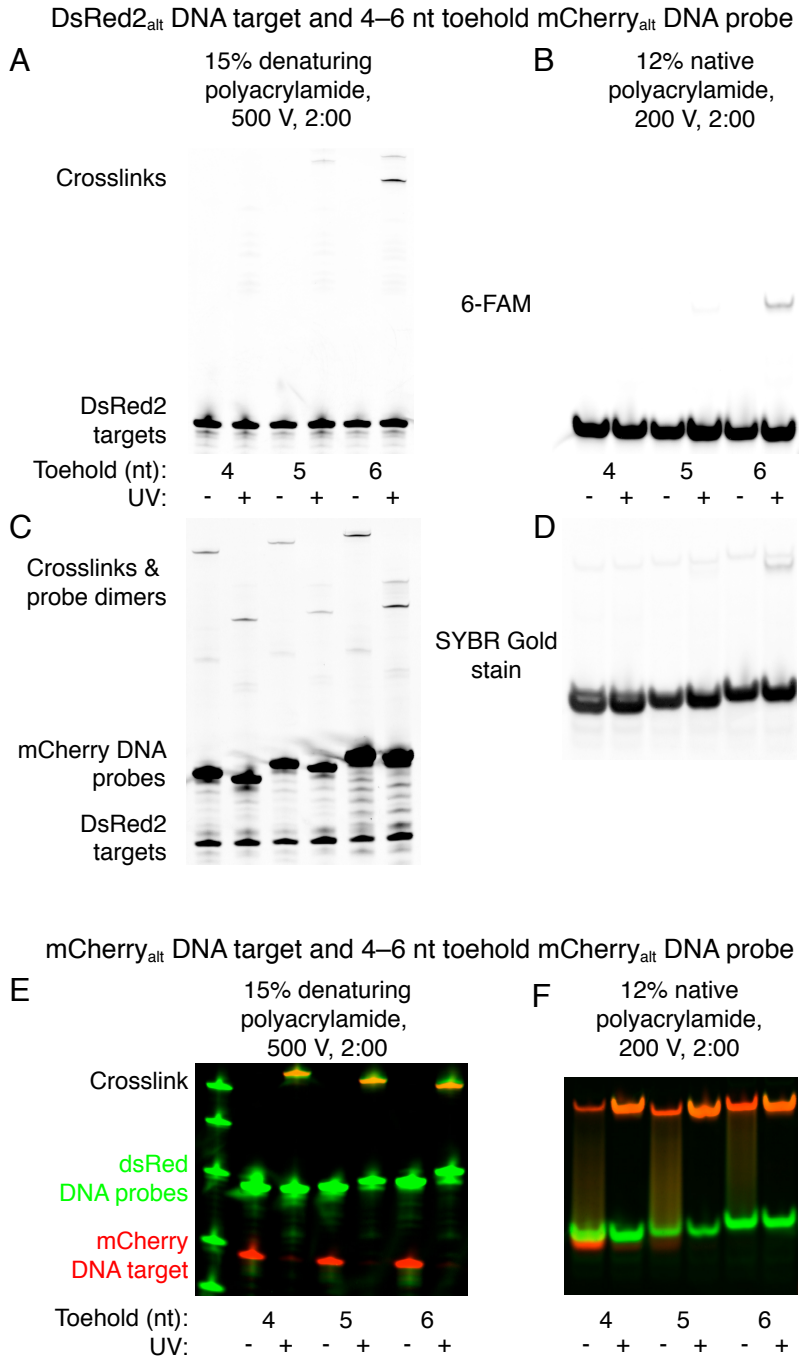


Figure S42. DNA mCherry$_{alt}$ probe toehold length studies with DsRed2$_{alt}$ and mCherry$_{alt}$ DNA targets. a) Denaturing gel showing (+ UV) reactions and (- UV) controls. 6-FAM-labeled targets. b) SYBR Gold post-stain of denaturing gel. c) Native gel analysis of the same reactions. 6-FAM-labeled targets. d) SYBR Gold post-strain of native gel. e) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled targets (red), SYBR Gold post-stain (green). f) Native gel analysis of the same reactions.

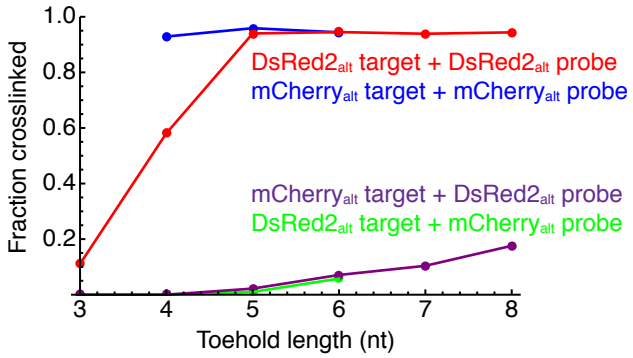


Figure S43. DNA DsRed2$_{alt}$ and mCherry$_{alt}$ probe toehold length studies with DNA DsRed2$_{alt}$ and mCherry$_{alt}$ targets ($N = 1$). Gels are shown in Figs S41 and S42. Yields increase with toehold length. The maximum discrimination ratio is observed for 5-nt toehold DsRed2$_{alt}$ and mCherry$_{alt}$ DNA probes.

S7.5 Alternate RNA DsRed2 and mCherry targets

For alternate RNA targets, probe toehold length studies are shown for $DsRed2_{alt}$ probes in Figure S44 and for $mCherry_{alt}$ probes in Figure S45, with capture yields plotted in Figure S46.

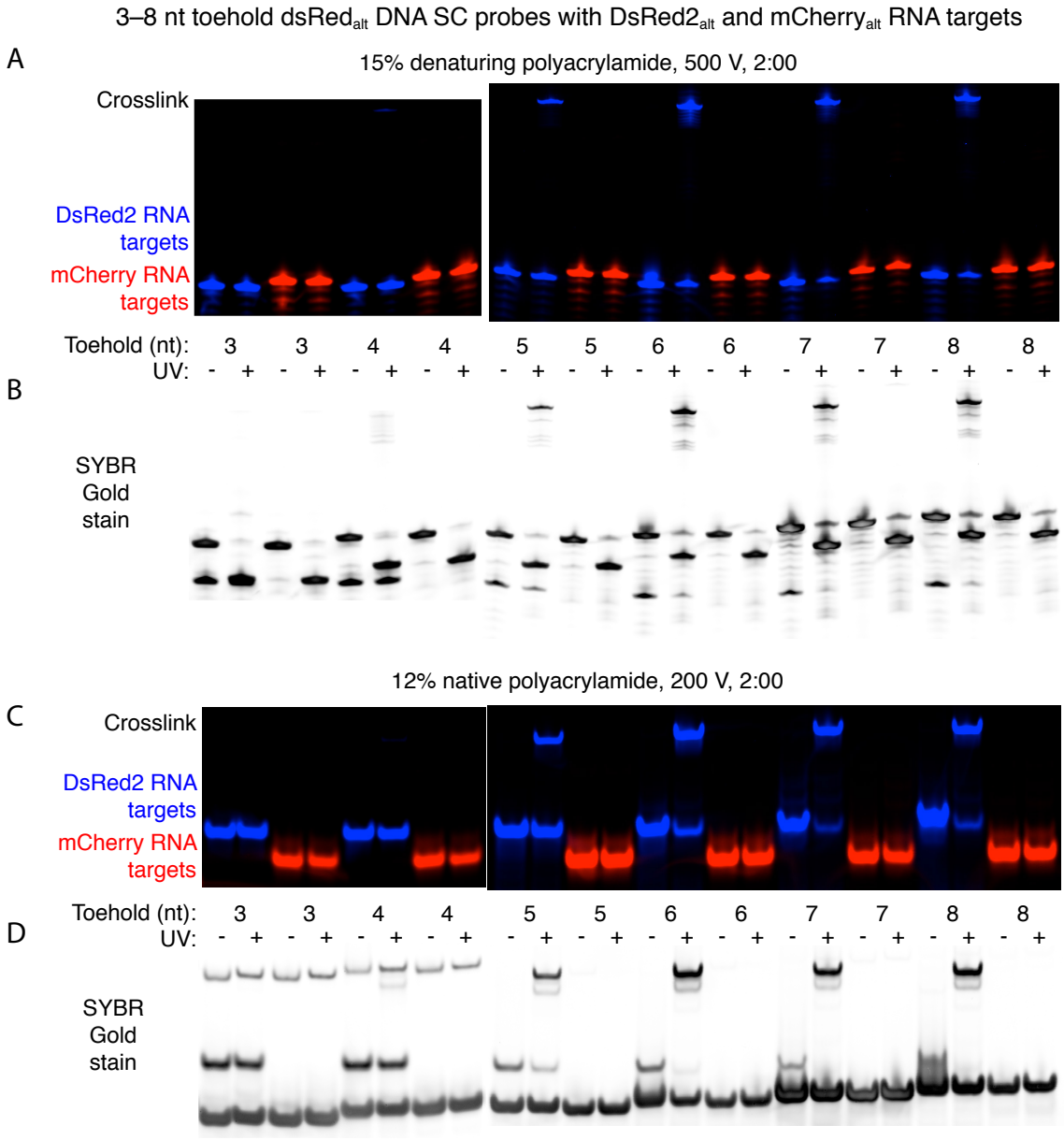


Figure S44. DNA $DsRed2_{alt}$ probe toehold length studies with $DsRed2_{alt}$ and $mCherry_{alt}$ RNA targets. a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled mCherry targets (red), 6-FAM-labeled DsRed2 targets (blue). b) SYBR Gold post-stain of denaturing gel. c) Native gel analysis of the same reactions. d) SYBR Gold post-strain of native gel.

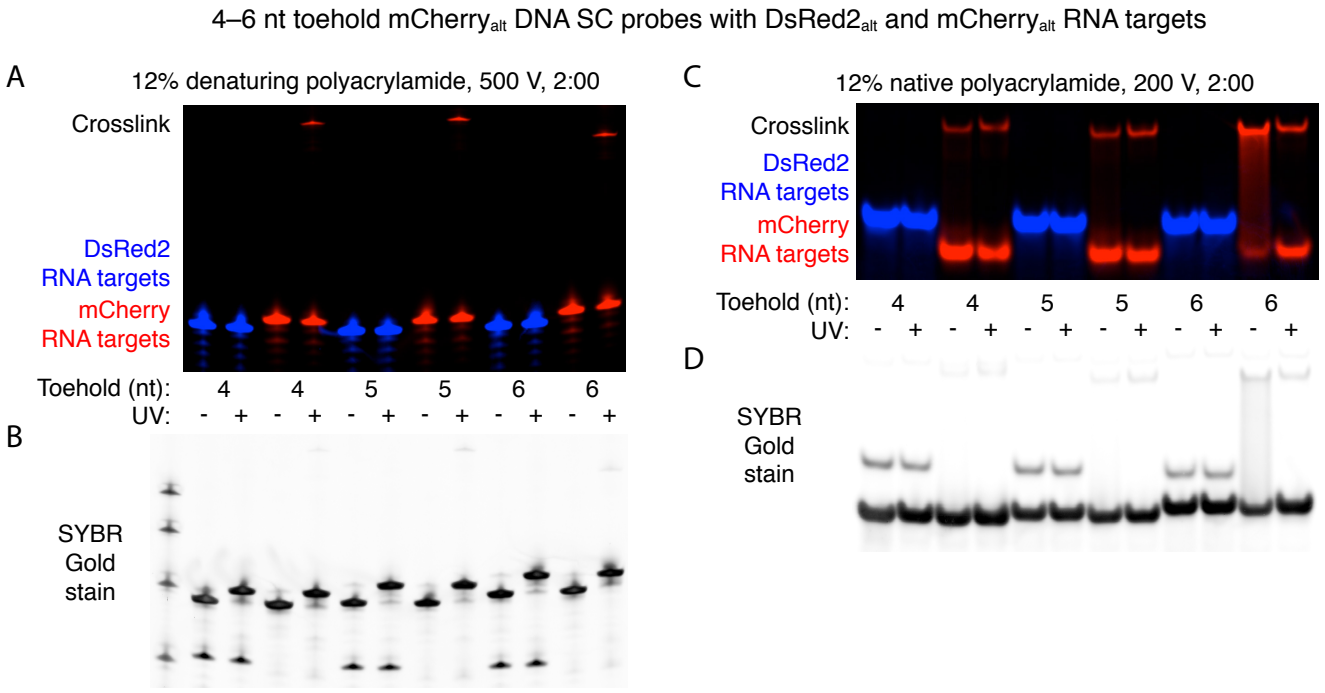


Figure S45. DNA mCherry$_{alt}$ probe toehold length studies with DsRed2$_{alt}$ and mCherry$_{alt}$ RNA targets. a) Denaturing gel showing (+ UV) reactions and (- UV) controls. Fluorescent channels: Cy5-labeled mCherry targets (red), 6-FAM-labeled DsRed2 targets (blue). b) SYBR Gold post-stain of denaturing gel. c) Native gel analysis of the same reactions. d) SYBR Gold post-strain of native gel.

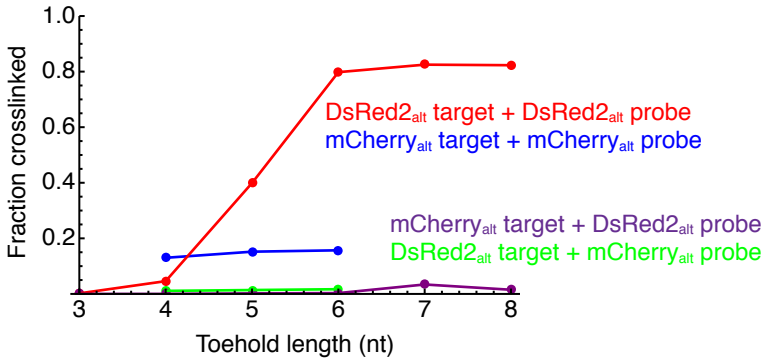


Figure S46. DNA DsRed2$_{alt}$ and mCherry$_{alt}$ probe toehold length studies with RNA DsRed2$_{alt}$ and mCherry$_{alt}$ targets ($N = 1$). Gels are shown in Figs S44 and S45. Yields increase with toehold length. The yield with the mCherry$_{alt}$ probe and target is low, though a high discrimination ratio is still achieved.

References

[1] Zadeh, J. N.; Steenberg, C. D.; Bois, J. S.; Wolfe, B. R.; Pierce, M. B.; Khan, A. R.; Dirks, R. M.; Pierce, N. A. *J. Comput. Chem.* **2011**, 32, 170–173.

[2] Zadeh, J.N.; Wolfe, B.R.; Pierce, N.A. *J. Comput. Chem.* **2011**, 32, 439–452.

[3] Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. *Bioinformatics* **2007**, 23, 2947–2948.

[4] SantaLucia, J., Jr.; Hicks, D. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, 33, 415–440.

[5] Yoshimura, Y.; Fujimoto, K. *Org. Lett.* **2008**, 10, 3227–3230.

[6] Yoshimura, Y.; Ohtake, T.; Okada, H.; Fujimoto, K. *ChemBioChem* **2009**, 10, 1473–1476.

[7] Whang-Peng, J.; Triche, T. J.; Knutsen, T.; Miser, J.; Douglass, E. C.; Israel, M. A. *N. Engl. J. Med.* **1984**, 311, 584–585.

[8] Hu-Lieskovan, S.; Heidel, J. D.; Bartlett, D. W.; Davis, M. E.; Triche, T. J. *Cancer Res.* **2005**, 65, 8984–8992.

[9] Kielbassa, C.; Epe, B. *Method Enzymol* **2000**, 319, 436–445.

[10] Cimino, G. D.; Shi, Y. B.; Hearst, J. E. *Biochemistry* **1986**, 25, 3013–3020.

[11] Hearst, J. *Annu. Rev. Biophys. Bioeng.* **1981**, 10, 69–86.

[12] Moreira, B. G.; You, Y.; Behlke, M. A.; Owczarzy, R. *Biochem. Biophys. Res. Commun.* **2005**, 327, 473–484.

[13] Zhang, D. Y.; Chen, S. X.; Yin, P. *Nature Chem.* **2012**, 4, 208–214.

[14] Matz, M. V.; Fradkov, A. F.; Labas, Y. A.; Savitsky, A. P.; Zaraisky, A. G.; Markelov, M. L.; Lukyanov, S. A. *Nat. Biotechnol.* **1999**, 17, 969–973.

[15] Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. *Nat. Biotechnol.* **2004**, 22, 1567–1572.