SUPPLEMENTARY INFORMATION

The *Francisella novicida* Cas12a is sensitive to the structure downstream of the terminal repeat in CRISPR arrays

Chunyu Liao[1,2], Rebecca A. Slotkowski[1], Tatjana Achmedov[2], Chase L. Beisel[1,2,3†]

[1]Department of Chemical and Biomolecular Engineering

North Carolina State University, Raleigh, NC 27695, USA

[2]Helmholtz Institute for RNA-based Infection Research

Josef-Schneider-Str. 2 / D15, D-97080 Würzburg, Germany

[3]Faculty of Medicine

University of Würzburg, D-97080 Würzburg, Germany

Table of Contents

SUPPLEMENTARY TABLES

Table S1. List of microbial strains, plasmids, and oligos used in this work. See the file Table S1.xlsx.

SUPPLEMENTARY FIGURES

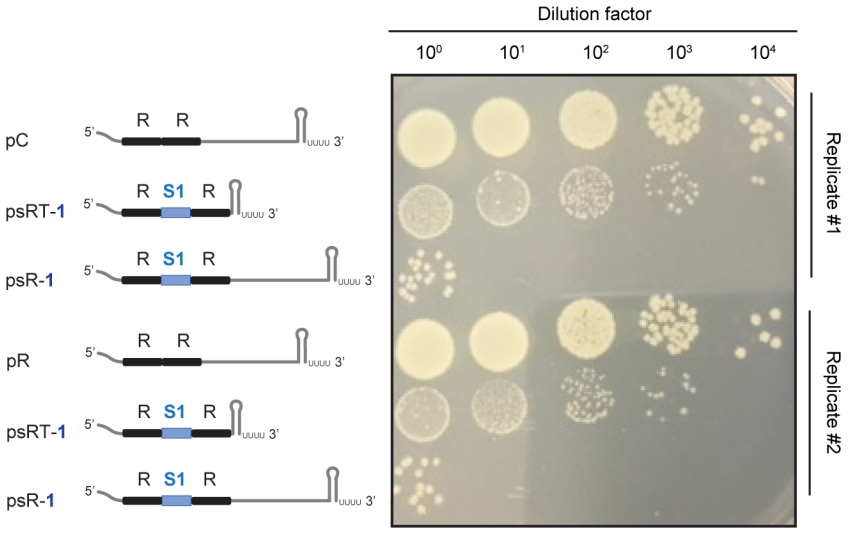

Figure S1. Plated cells from the plasmid clearance assays. *E. coli* cells harboring the FnCas12a plasmid and the targeted reporter plasmid were transformed with one of three CRISPR array constructs. The constructs contained a no-spacer CRISPR array (pC); a single-spacer CRISPR array with spacer S1 and a synthetic Rho-independent terminator immediately adjacent to the 3' repeat (psRT-1); or a single-spacer CRISPR array with spacer S1, no obvious adjacent secondary structure, and the *rrnB* terminator whose designated region begins ~88 nts downstream of the 3' repeat (psR-1). Transformed cells were plated at the indicated dilution factors and incubated at 37°C for ~16 hours. A representative plate with two independent replicates is shown.

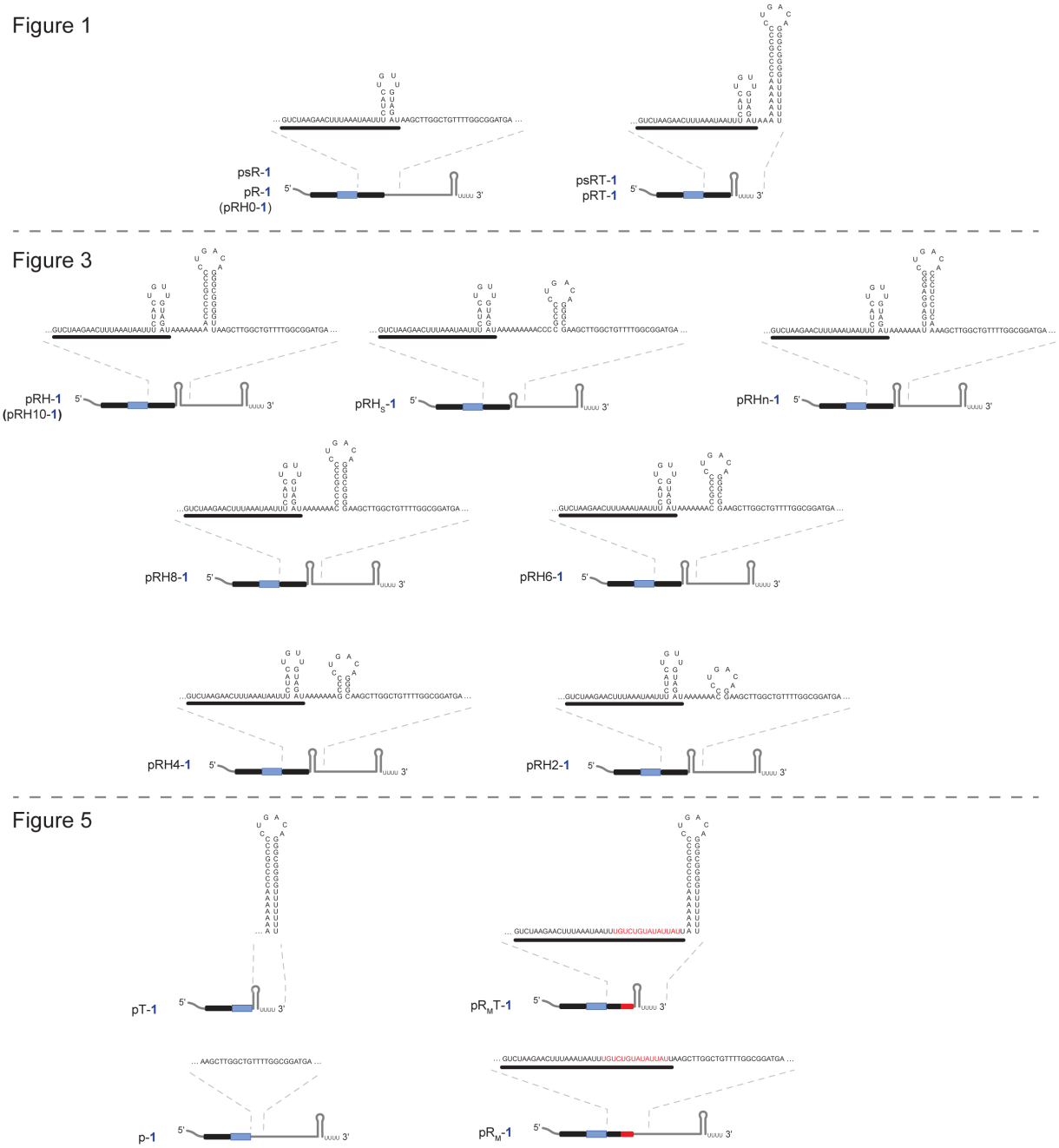

Figure S2. Sequence representations of CRISPR array constructs used in this study. For the targeting constructs, black bars represent direct repeats, colored bars represent spacers, gray hairpins are stem-loop structures, and gray hairpins followed by multiple U's are Rho-independent *rrnB* terminators. The nt sequence is illustrated for regions contained within the dashed lines. Direct repeats are underlined with a black bar, while unique sequences within the native terminal repeat for FnCas12a are red.

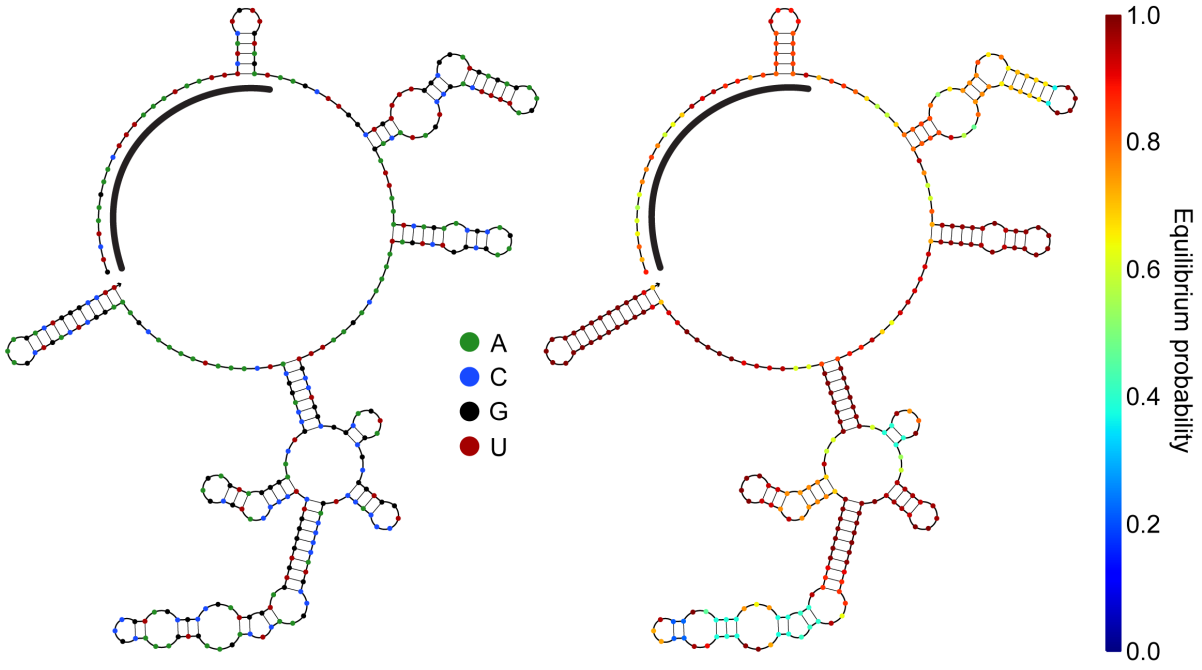

Figure S3. No appreciable secondary structure predicted between the 3' repeat and the downstream *rrnB* terminator in the array constructs without a synthetic Rho-independent terminator. The RNA folding prediction software NUPACK (www.nupack.org) was used to predict the secondary structure from the beginning of the 3' repeat to the end of the *rrnB* transcriptional terminator using standard parameters. The sequence of the minimal-free energy structure is shown on the left, while the probability of a formed base-pairing interaction is shown on the right. The region spanning the repeat is underlined with a black bar. The closest predicted hairpin has fewer consecutive base pairs and has moderate base-pairing probabilities.

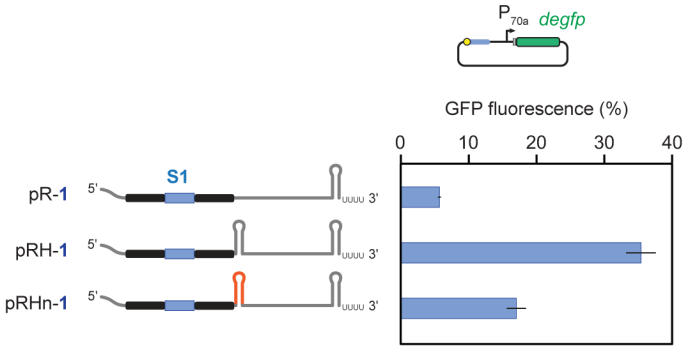

Figure S4. Inhibition of DNA cleavage by different hairpin sequences flanking the 3' repeat. Single-spacer arrays encoding spacer S1 and lacking any strong secondary structure (pR-1) or flanked by one of two distinct hairpins (pRH-1, pRHn-1) were assessed in the TXTL-based cleavage assay with FnCas12a. See Figure S2 for sequences of the region immediately downstream of the 3' repeat. The TXTL-based cleavage assays were conducted as described in Figure 1C using a deGFP reporter plasmid with the target sequence inserted upstream of the reporter expression construct. Fluorescence values are reported as the GFP fluorescence of the tested array divided by that of the no-spacer array 5 hours after the addition of the reporter plasmid. Values and error bars represent the mean and S.D., respectively, from triplicate experiments conducted on different days.

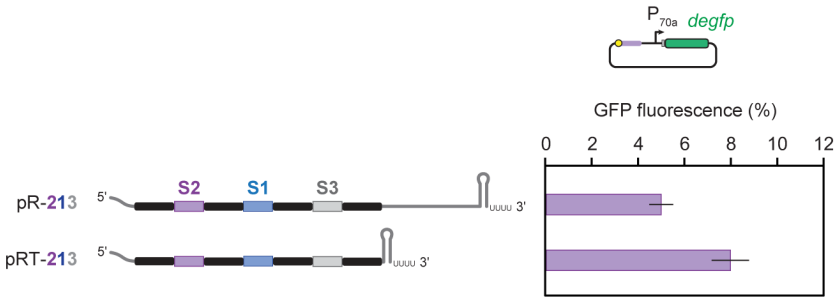

Figure S5. The flanking hairpin can partially impair DNA cleavage by the upstream-most spacer in a three-spacer array. Three-spacer arrays were constructed to contain, from 5' to 3', spacers S2, S1, and S3. The arrays were flanked immediately by a synthetic Rho-independent terminator (pRT-213) or a far downstream *rrnB* terminator (pR-213). The arrays were then subjected to the TXTL-based cleavage assay with FnCas12a as described in Figure 1C using a deGFP reporter plasmid with the target sequence inserted upstream of the reporter construct. Fluorescence values are reported as the GFP fluorescence of the tested array divided by that of the no-spacer array 5 hours after the addition of the reporter plasmid. Values and error bars represent the mean and S.D., respectively, from triplicate experiments conducted on different days.