Supplementary Material

The molecular mechanism of dsRNA processing by a bacterial Dicer

Lan Jin[1], He Song[1], Joseph E. Tropea[1], Danielle Needle[1], David S. Waugh[1], Shuo Gu[2,*] and Xinhua Ji[1,*]

[1] Macromolecular Crystallography Laboratory, National Cancer Institute, Frederick, MD 21702, USA
[2] RNA Biology Laboratory, National Cancer Institute, Frederick, MD 21702, USA

* To whom correspondence should be addressed. Tel: 1-301-846-5447 (S.G.) or 1-301-846-5035 (X.J.); Fax: 1-301-846-6073 (X.J.); Email: shuo.gu@nih.gov (S.G.) or jix@mail.nih.gov (X.J.)

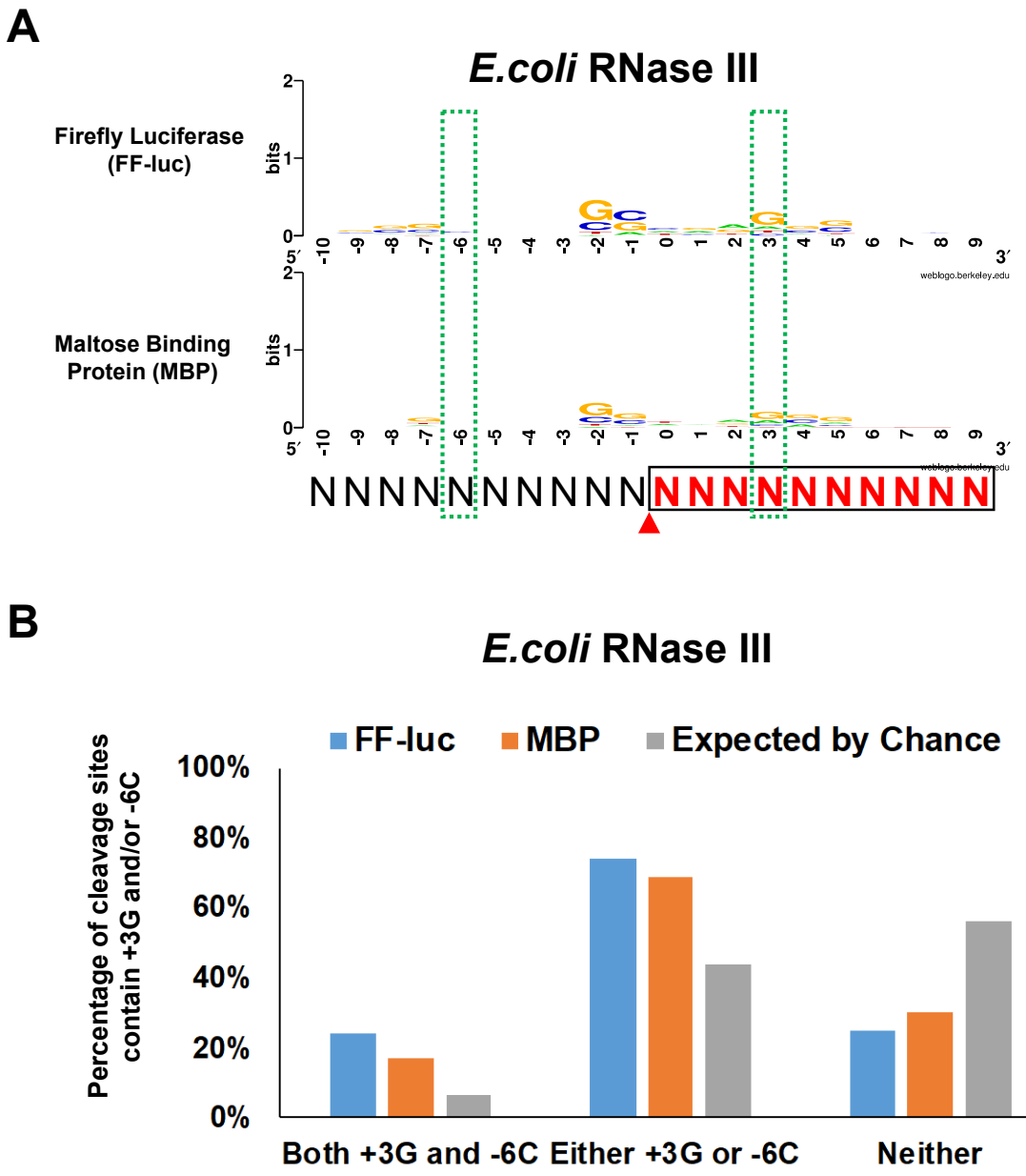

Figure S1. RNase III recognizes +3G in cleavage site selection (**A**) Sequence logos surrounding preferred *Escherichia coli* RNase III cleavage sites on Firefly luciferase (FF-luc) and maltose-binding protein (MBP) are presented. Red arrowheads indicate the cleavage site (between positions 0 and -1), which is inferred from the end of cleavage products. Cleavage product upstream of cleavage site is in black whereas cleavage products downstream of cleavage site is in red and boxed. The dashed-line boxes in green indicate two consensus nucleotides +3G and -6C. (**B**) Percentage of cleavage sites that contain +3G and/or -6C is higher than that expected by chance.

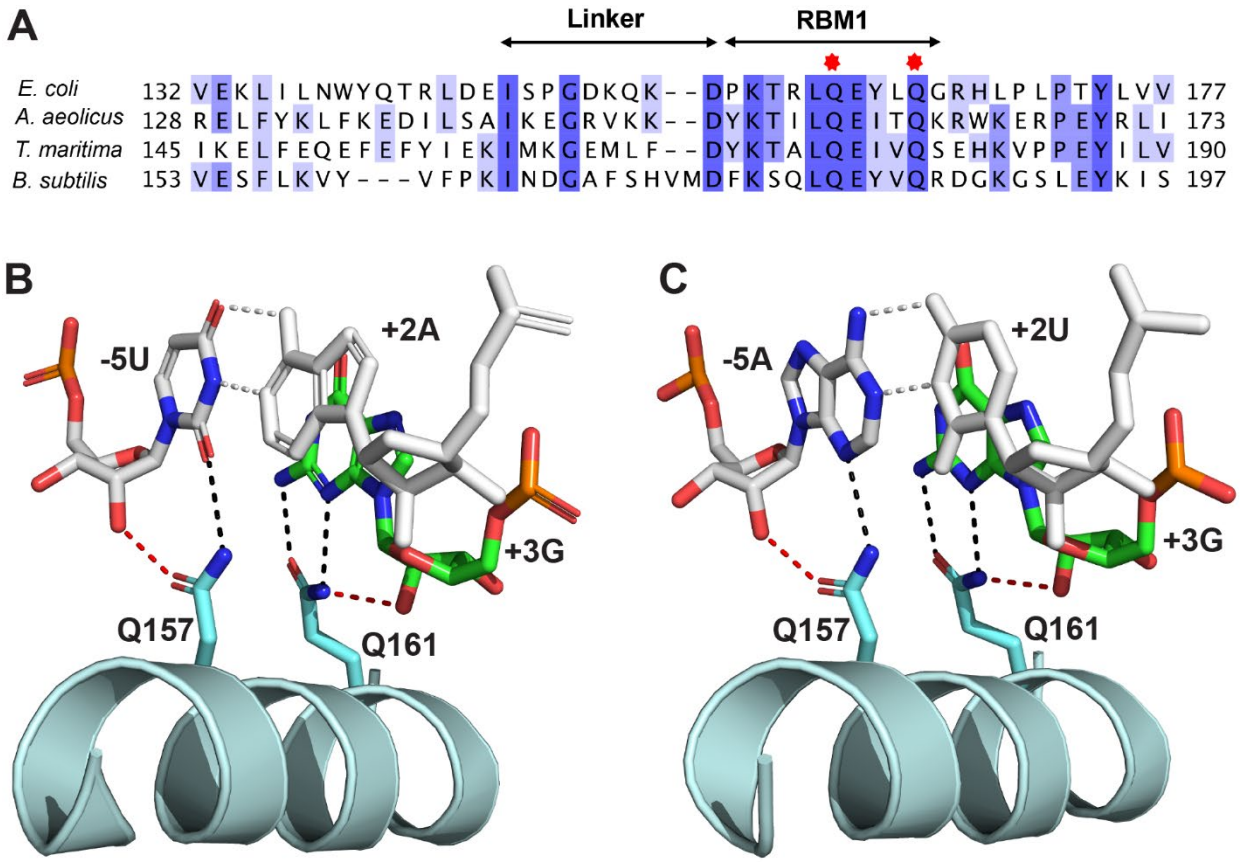

Figure S2. Involvement of RNase III residues Q157 and Q161 in the recognition of RNA substrate. (**A**) Sequence alignment of selected bacterial RNase III enzymes from *Escherichia coli*, *Aquifex aeolicus* (Ac), *Thermotoga maritima*, and *Bacillus subtilis*. Identical and similar residues are shaded in blue and light blue. The red asterisks indicate the two conserved glutamine residues, Q157 and Q161 in AaRNase III, in RBM1 (RNA-binding motif 1), i.e., the first α-helix of double-stranded RNA-binding domain (dsRBD). The linker between RIIID (the specialized RNase III domain) and dsRBD (dsRNA-binding domain) is indicated. **(B**) As revealed by the crystal structure of AaRNase III in complex with dsRNA (PDB entry 2EZ6), the Q157 side chain forms one hydrogen bond with the 2'-hydroxyl and one hydrogen bond with the base of the -5U nucleotide. **(C**) As revealed by the crystal structure of AaRNase III in complex with a different dsRNA (PDB entry 2NUG), the Q157 side chain forms one hydrogen bond with the 2'-hydroxyl and one hydrogen bond with the base of the -5A nucleotide. As such, Q157 recognizes both the AU and UA pair in this position. Unlike Q157, however, Q161 recognizes the +3G nucleotide in both structures by forming one hydrogen bond with the 2'-hydroxyl and two hydrogen bonds with the base. The first α-helix in the dsRBD of AaRNase III is illustrated as a spiral, the amino acid and nucleotide residues are shown as stick models, and hydrogen bonds are indicated with dashed lines. The -6C nucleotide that is paired with the +3G nucleotide is not shown for clarity.

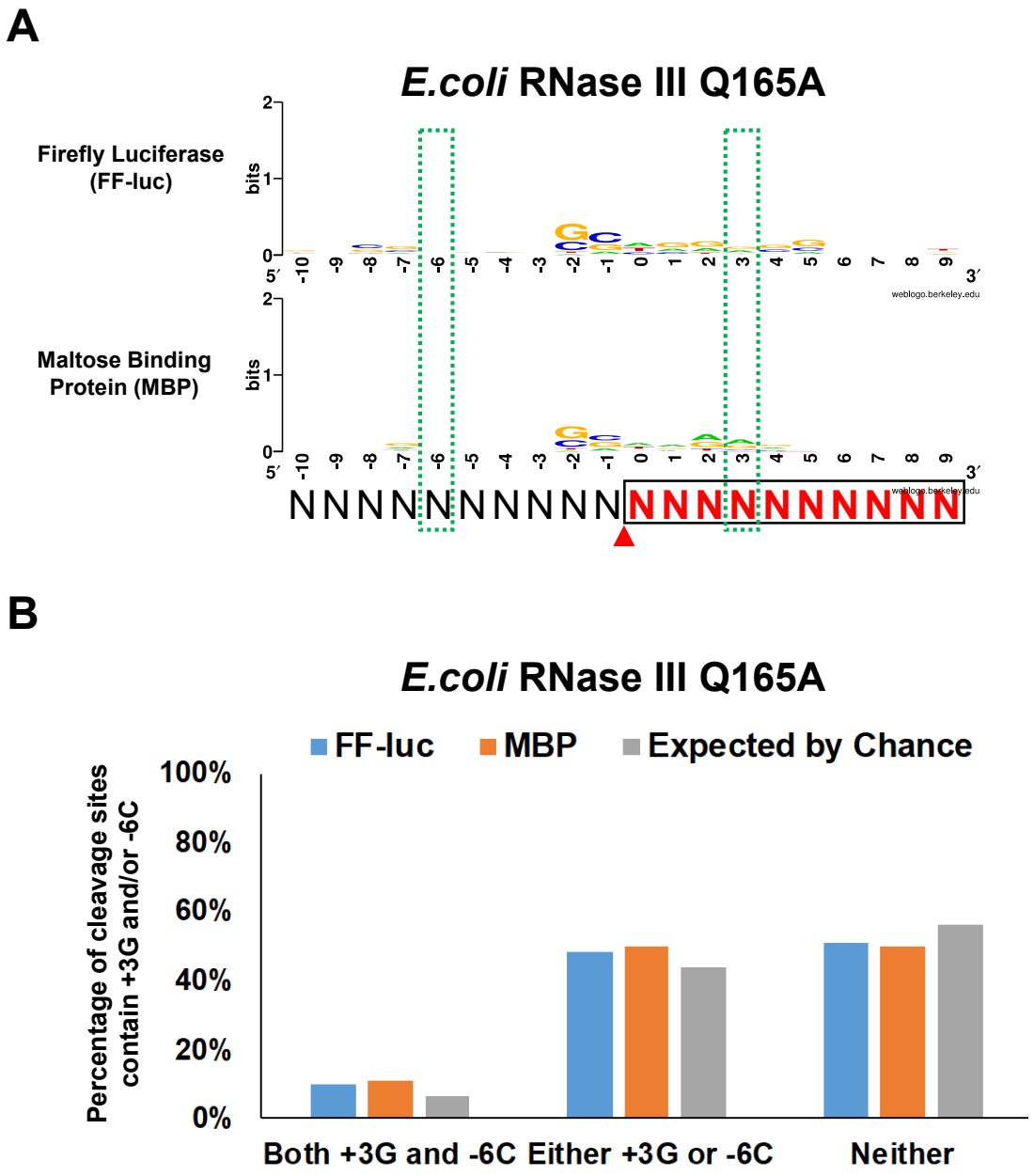

Figure S3. Residue Q165 of *E. coli* RNase III recognizes the +3G near the cleavage site. (**A**) The consensus sequence of cleavage site is abolished when residue 165 was mutated from Q to A. Sequence logos were created as illustrated in Figure S1. (**B**) Percentage of highly preferable cleavage sites of the *E. coli* RNase III Q165A mutant, with or without +3G and/or -6C, is similar to that expected by chance.

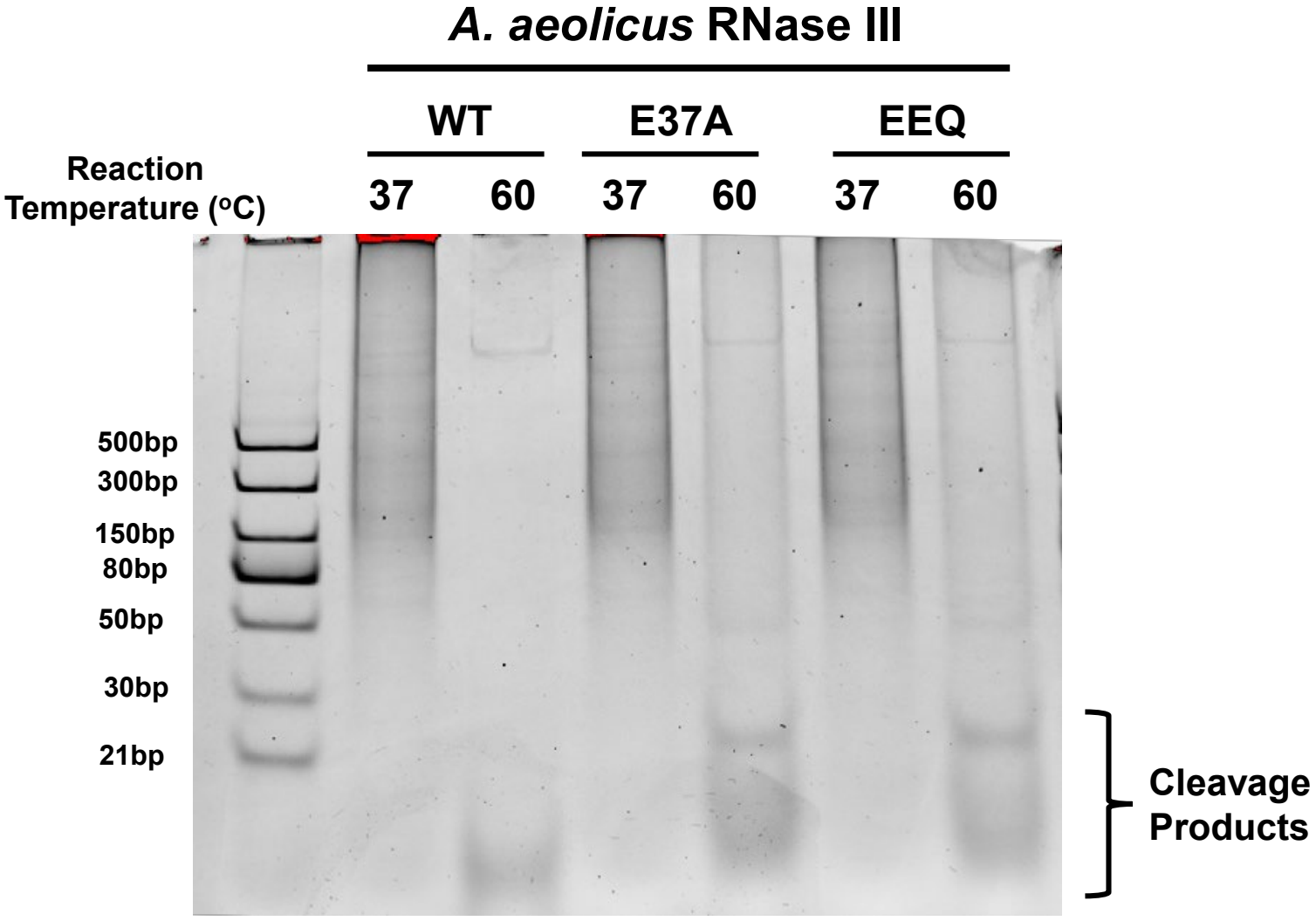

Figure S4. *A. aeolicus* RNase III is inactive at 37°C. dsRNA of FF-luc sequence was cleaved by the wild type (WT), E37A mutant, or EEQ mutant of *A. aeolicus* (Aa) RNase III at either 37°C or 60°C. The cleavage products were separated on 20% polyacrylamide native gels and detected by staining. While WT AaRNase III generates products of smaller size, E37A and EEQ mutants produce additional siRNA-size products at 60°C but are inactive at 37°C. Of note, the substrate dsRNAs (1.6Kb) are too large to run into this polyacrylamide gel and are stuck in the wells.

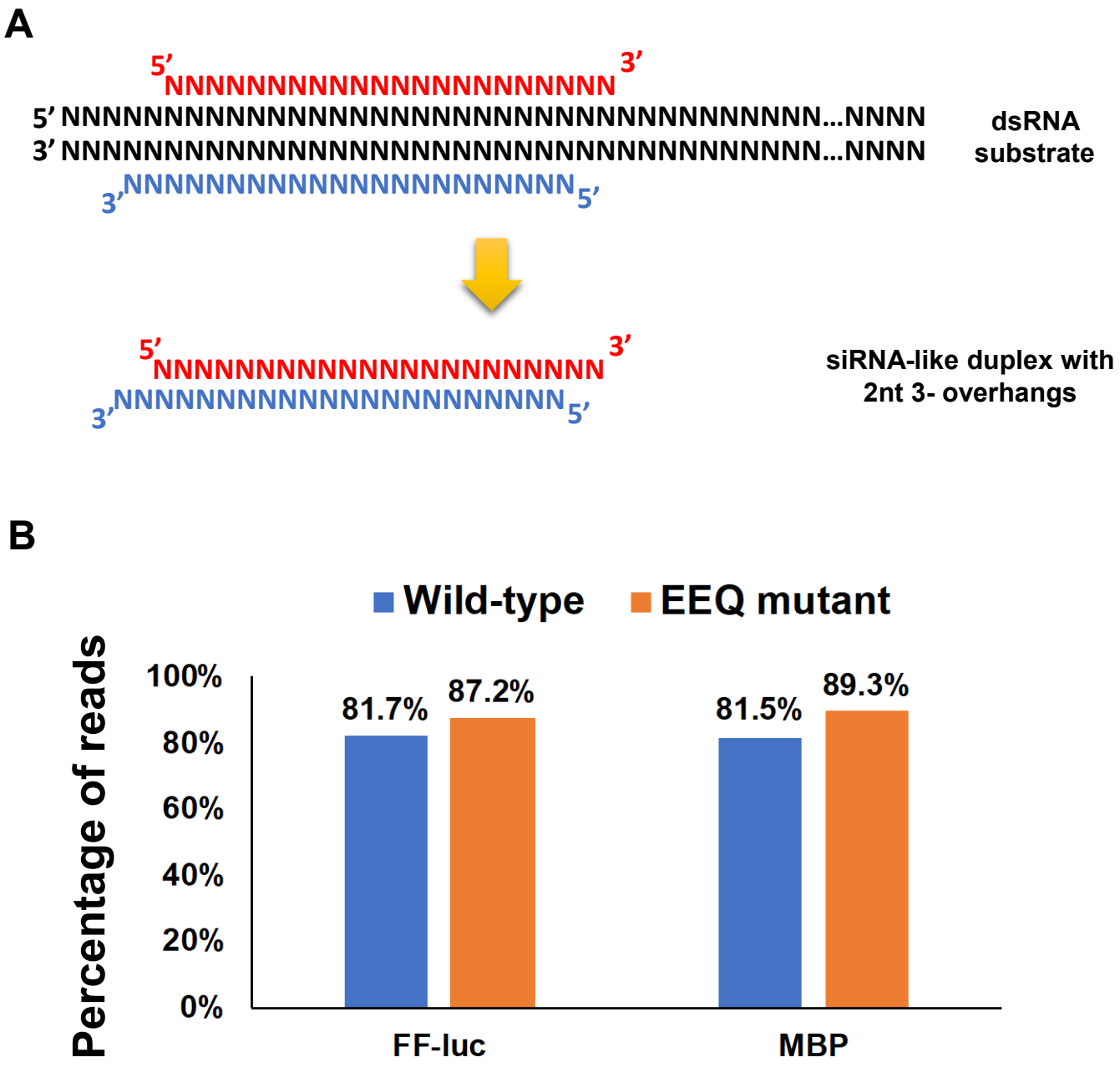

Figure S5. The *E. coli* EEQ mutant generates siRNA-like duplexes. (**A**) Schematic representation of the flow of analysis. Reads of cleavage products were first mapped to the dsRNA substrate. Reads mapped to the sense strand of the substrate were in red; Reads mapped to the antisense strand are in blue. Red and blue reads are counted as paired if they are aligned to form the siRNA-like duplex with 2-nt 3' overhangs at both ends. (**B**) The percentage of 22-mer paired reads to the total number of 22-mer reads was plotted.

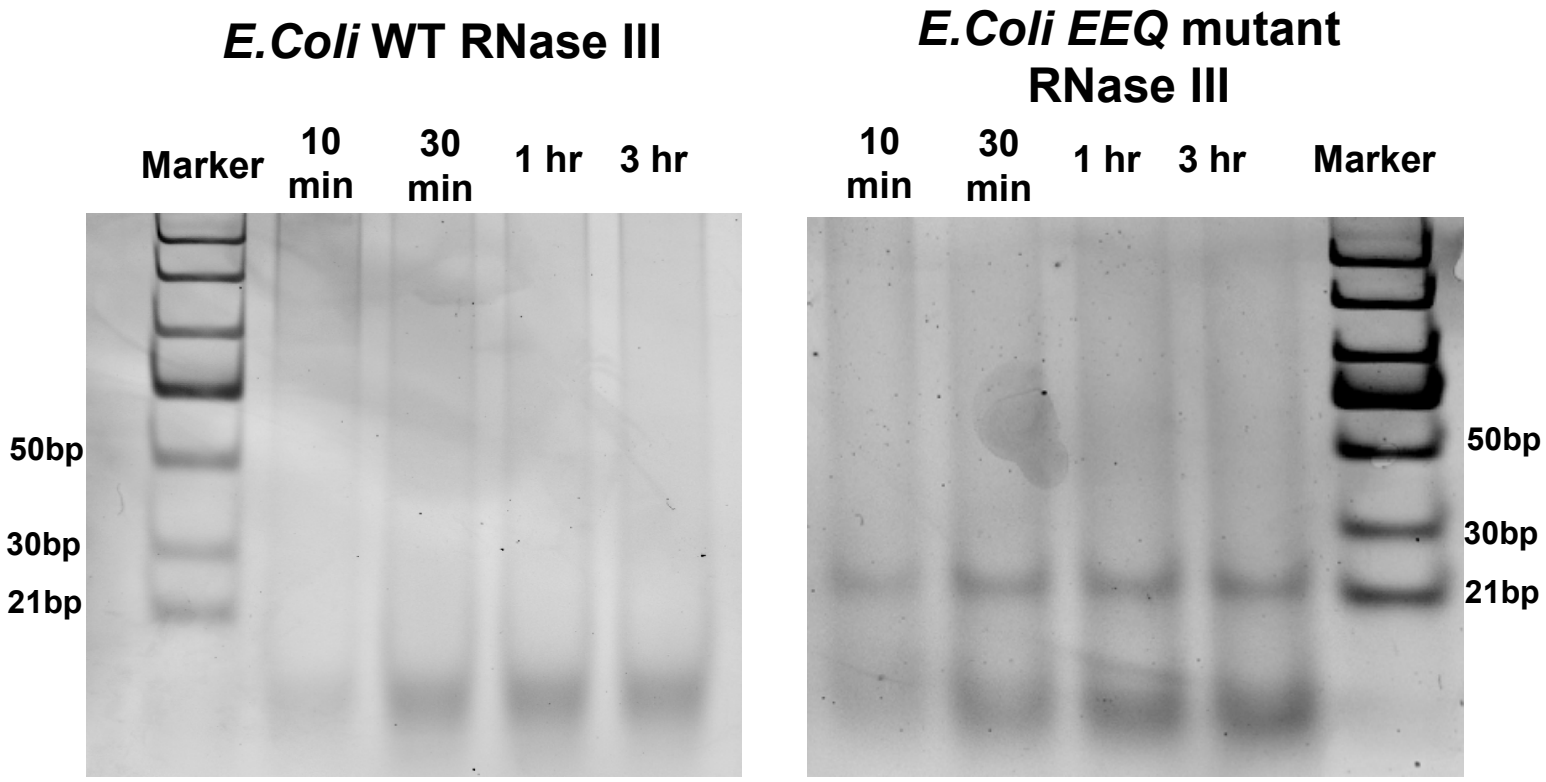

Figure S6. RNase activities of *E. coli* RNase III and its EEQ mutant were measured by a time course study. dsRNA of FF-luc sequence was cleaved by the wild type (WT) or EEQ mutant of *E. coli* RNase III at 37°C. The cleavage products were separated on 20% polyacrylamide native gels and detected by staining.

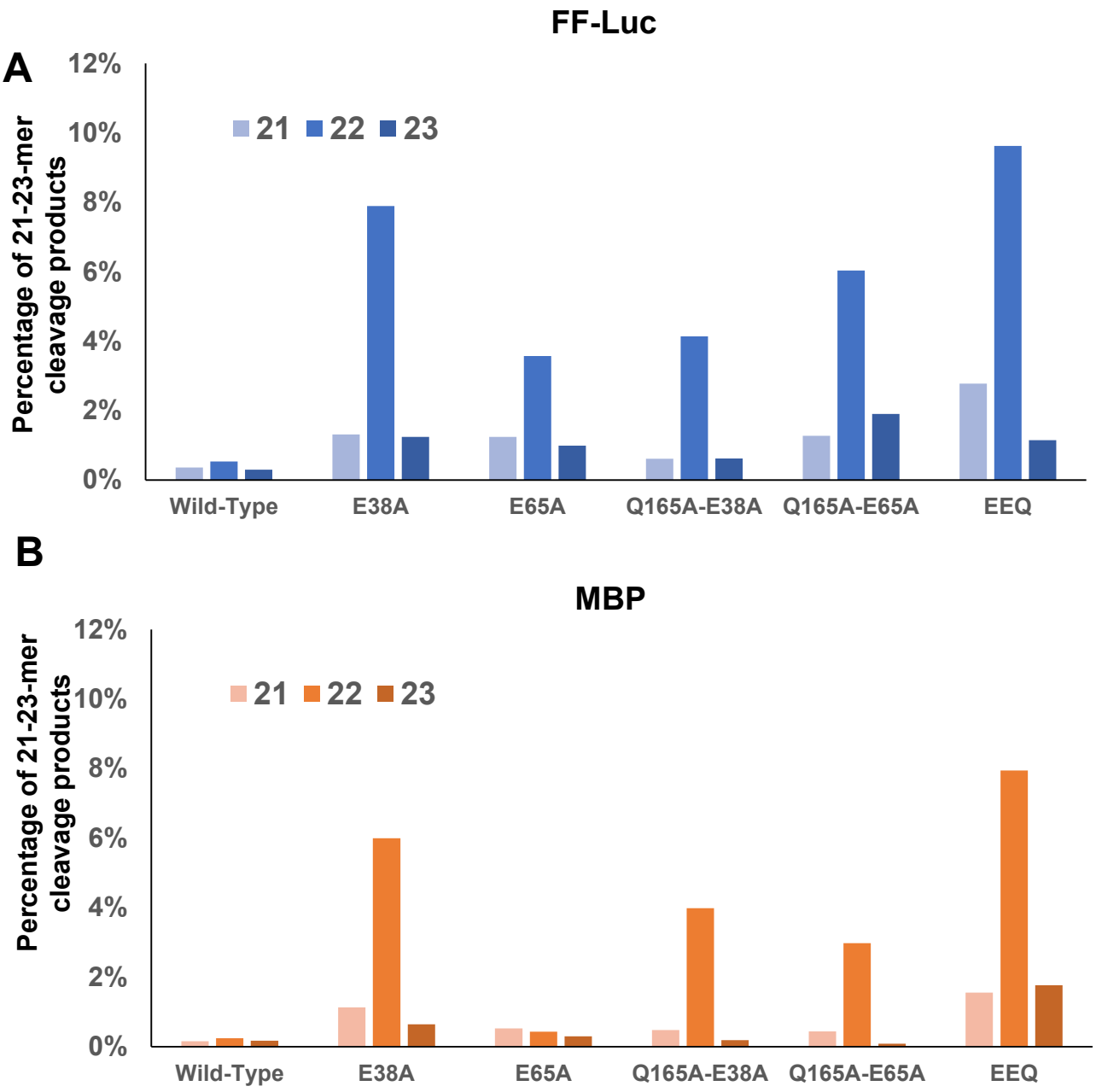

Figure S7. Engineered *E. coli* RNase III cleaves in vitro transcribed dsRNA into a heterogeneous mixture of siRNAs with a narrow size distribution centered at 22 nt. (**A**) The percentage of 21-23-nt products to the total number of reads was plotted for Firefly luciferase (FF-luc). (**B**) The percentage of 22-nt products to the total number of reads was plotted for maltose-binding protein (MBP).