

Title

The Genetic Insulator RiboJ Increases Expression of Insulated Genes

Author List

Authors: Kalen P Clifton^{#1}, Ethan M Jones^{#1}, Sudip Paudel¹, John P Marken², Callan E Monette¹, Andrew D Halleran², Lidia Epp¹, Margaret S Saha^{1*}

[#]These authors equally contributed to the work

^{*}Corresponding Author

Abstract

The self-cleaving ribozyme RiboJ is an insulator commonly used in genetic circuits to prevent unexpected interactions between neighboring parts. These interactions can compromise the modularity of the circuit, impeding the implementation of predictable genetic constructs. Despite its utility as an insulator, a quantitative assessment of the effect of RiboJ on the properties of downstream genetic parts is lacking. Here, we characterized the impact of insulation with RiboJ on expression of a reporter gene driven by a promoter from a library of 24 frequently employed constitutive promoters. We show that depending on the strength of the promoters, insulation with RiboJ increased protein abundance between twofold and tenfold and increased transcript abundance by an average of twofold. This result is the first to demonstrate that genetic insulators can impact the expression of downstream genes, potentially hindering the design of predictable genetic circuits and constructs.

Keywords

RiboJ, Insulation, Characterization, digital droplet PCR, genetic circuit, ribozyme

Main Text

A fundamental goal of synthetic biology is the prediction of the behavior of genetic constructs based on the properties of their constituent genetic parts [1]. This predictability relies on the use of modular parts, whose behaviors are unaffected by other parts in the construct [2]. However, modularity can be compromised when interactions between neighboring parts create unintended functional sequences, altering genetic context [3]. These alterations lead constructs to behave unpredictably, impeding their function.

One significant source of these alterations of genetic context is the use of synthetic promoters containing regulatory sequences downstream of the transcriptional start site [4]. This additional sequence is transcribed, leading to the inclusion of unintended nucleotides termed “RNA leaders” at the 5’ end of the transcript. It has been shown that these RNA leaders can modify the stability and secondary structure of mRNA, which alters the translational properties of genetic constructs [2,4]. The nature of these alterations is specified by the interactions between a given RNA leader and the downstream sequence of the transcript. Thus, changes to a construct’s behavior will depend on both the specific promoter used and the composition of a construct.

To circumvent the effects of RNA leaders, constructs can be designed to include genetic insulators, which isolate parts from unwanted interactions with their neighboring regions. One routinely used genetic insulator is the synthetic self-cleaving ribozyme RiboJ [4,5,6,7]. RiboJ is a 75 nucleotide (nt) sequence which is inserted in a construct at the junction of a promoter and its downstream sequence. After transcription, RiboJ self-cleaves, removing the RNA leader and leaving behind a short sequence from its uncleaved region. This elimination of RNA leaders

standardizes the behavior of promoters across constructs insulated with RiboJ, aiding the design of predictable genetic constructs.

Although RiboJ is frequently used as an insulator in genetic constructs, there has been limited characterization of what, if any, effect this insulation has on downstream genetic parts. Since the construction of accurate and robust genetic circuits could be impeded if insulation with RiboJ has any unanticipated effects on gene expression, we characterized the impact of RiboJ insulation on the gene expression of a set of constitutive promoter constructs.

We assembled two sets of reporter constructs that differed only by the presence RiboJ. Each set contained 24 constructs, which each expressed a superfolder green fluorescent protein (sfGFP) [4] reporter with a different synthetic *Escherichia coli* (*E. coli*) constitutive promoter part. The collection of promoter parts spans a wide range of transcriptional strengths and includes the well characterized and commonly used Anderson Promoter Library [8], as well as synthetic hybrid promoters such as R0011 (pLlacO-1) (**Supplementary Table 1**). Each construct was transformed into BL21 *E. coli* and expression of the fluorescent reporter was measured using flow cytometry [9]. Detailed methods can be found in the supporting information.

We found that for each pair of promoter parts, the construct insulated by RiboJ had greater absolute fluorescence than the corresponding construct without RiboJ (**Figure 1a**, **Supplemental Figure 1**), with increases ranging from twofold to tenfold (**Figure 1b**). We found that the fold change in expression with RiboJ exhibited bimodality between “strong” and “weak” promoters. Across 18 of our assay’s 19 strongest promoters, insulation with RiboJ increased absolute fluorescence by an average of sevenfold, ranging from threefold to tenfold. For the remaining promoters, we found that the magnitude of the increase was lower, with an average increase of fivefold, ranging from twofold to eightfold. While we found that RiboJ exhibited bimodality between “stronger” and “weaker” promoters, we did not find a monotonic relationship between promoter strength and fold change in protein expression (**Supplemental Figure 2**).

Since the increase in protein expression with RiboJ could be attributed to either differential transcription or translation of insulated genes, we characterized the effect of insulation with RiboJ on the relative abundances of sfGFP transcripts. We used reverse transcription digital droplet qPCR (ddPCR) with Uroporphyrinogen-III C-methyltransferase (CysG) serving as an endogenous reference [9]. We found that insulation of constructs with RiboJ, increased sfGFP transcript abundance by an average of twofold, while there was no change in transcript abundance of our reference gene CysG on average (**Supplemental Figures 3, 4**). The mean fold change for sfGFP transcript counts was greater than for a null distribution and for an endogenous reference gene, which indicates that the observed increase in the transcript abundance of sfGFP is indeed due to insulation with RiboJ (**Figure 2, Supplemental Figure 5**).

While the cause of the increased expression of RiboJ-insulated genes is unknown, the increased expression levels of protein and mRNA could be driven mainly by an increase in mRNA concentration. This increase in mRNA concentration could be due to increased mRNA stability from the terminal hairpin formed by the remaining RiboJ sequence post-cleavage [4, 11]. We found that sfGFP transcript abundance was well correlated to sfGFP protein concentration (**Supplemental figure 6**), which could imply that increased protein abundance is driven by a RiboJ-associated increase in mRNA abundance. However, the fold change in transcript abundance for the collection of constructs did not correlate well with fold change in protein (**Supplemental figure 7**), which suggests that increases in protein abundance due to RiboJ could be due to translational processes as well as increases in transcript

abundance. Additional investigation is required to determine the causes of these increases in gene expression under insulated conditions.

The genetic insulator RiboJ is a valuable tool that aids the implementation of predictable genetic circuits by allowing promoter characterization to be standardized across genetic constructs. While insulation with RiboJ is widely used, there has been no comprehensive characterization of its effects on the expression of insulated genes. Here we provide a quantitative characterization of the effect of insulation with RiboJ on a collection of promoter constructs. We determined that insulating a construct with RiboJ leads to an increase in protein expression and transcript abundance. Since unanticipated increases in gene expression can decrease circuit performance by increasing cellular metabolic strain or by causing decreases in the dynamic range of portions of the circuit [12], this characterization provides valuable information for the design and implementation of genetic constructs and circuits.

Supporting Information

SI.docx: Supporting figures S1-7, Supplemental methods, Supporting Table 1, Construct design
Data.csv: Experimental Data
Analysis.ipynb: Analysis methods

Abbreviations:

ddPCR digital droplet quantitative polymerase chain reaction
sfGFP superfolder GFP

Author Information:

1. Department of Biology, Integrated Science Center, The College of William and Mary, 540 Landrum Drive, Williamsburg, VA 23185, United States.
2. Division of Biology and Bioengineering, California Institute of Technology, Pasadena, CA 91125, United States.

Author Contributions:

ADH designed constructs. EMJ and KPC cloned constructs with help of ADH. EMJ, KPC and SP isolated RNA. LE performed reverse transcription and digital droplet qPCR. EMJ and KPC performed flow cytometry measurements and analysis. EMJ, KPC and JPM analyzed data. KPC and JPM created figures. MSS supervised project. EMJ, KPC wrote manuscript. KPC, ADH EMJ, CEM, JPM, SP, MSS edited manuscript.

Conflict of interest:

Authors declare no known conflict of interest.

Acknowledgements:

We would like to thank Joseph Maniaci for his assistance in the initial cloning of the construct library. We would also like to thank Vice Provost Dennis Manos for his intellectual support throughout the project.

Funding Sources:

Research was funded by NSF Grant 1257895 to MSS and NIH Grant 1R15HD077624-01 to MSS. Project was also supported by funding from the office of the Vice Provost for Research and Graduate/Professional Studies.

References:

1. Davidsohn, N., Beal, J., Kiani, S., Adler, A., Yaman, F., Li, Y., Xie, Z., and Weiss, R. (2015) Accurate Predictions of Genetic Circuit Behavior from Part Characterization and Modular Composition. *ACS Synth. Biol.* 4, 673–681. doi: 10.1021/sb500263b
2. Del Vecchio D, Ninfa AJ, Sontag ED (2008) Modular cell biology: retroactivity and insulation. *Mol Syst Biol* 4: 161 doi: 10.1038/msb4100204
3. Yeung, E., Dy, A. J., Martin, K. B., Ng, A. H., Del Vecchio, D., Beck, J. L., Collins, J. J., and Murray, R. M. (2017) Biophysical Constraints Arising from Compositional Context in Synthetic Gene Networks. *Cell Syst.* 5, 11–24. doi: 10.1016/j.cels.2017.06.001
4. Lou, C., Stanton, B. C., Chen, Y.-J., Munsky, B., and Voigt, C. A. (2012) Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotechnol.* 30, 1137–1142. doi: 10.1038/nbt.2401

5. Courbet A, Endy D, Renard E, Molina F, Bonnet J. (2015). Detection of pathological biomarkers in human clinical samples via amplifying genetic switches and logic gates. *Science Translational Medicine*, 7(289). doi:10.1126/scitranslmed.aaa3601
6. Bai C, Zhang Y, Zhao X, Hu Y, Xiang S, Miao J, Zhang L. (2015). Exploiting a precise design of universal synthetic modular regulatory elements to unlock the microbial natural products in *Streptomyces*. *Proceedings of the National Academy of Sciences*, 112(39), 12181-12186. doi:10.1073/pnas.1511027112
7. Rubens J. R, Selvaggio G, Lu T K (2016). Synthetic mixed-signal computation in living cells. *Nature Communications*, 7, 11658. doi:10.1038/ncomms11658
8. Anderson promoter collection. Available: <http://parts.igem.org/Promoters/Catalog/Anderson>.
9. Castillo-Hair, S. M., Sexton, J. T., Landry, B. P., Olson, E. J., Igoshin, O. A., and Tabor, J. J. (2016) FlowCal: A User-Friendly, Open Source Software Tool for Automatically Converting Flow Cytometry Data from Arbitrary to Calibrated Units. *ACS Synth. Biol.* 5, 774–780. doi: 10.1021/acssynbio.5b00284
10. Peng S, Stephan R, Hummerjohann J, Tasara T. (2014). Evaluation of three reference genes of *Escherichia coli* for mRNA expression level normalization in view of salt and organic acid stress exposure in food. *FEMS Microbiology Letters*, 355(1), 78-82. doi:10.1111/1574-6968.12447
11. Carrier, T. A., and Keasling, J. D. (1997) Engineering mRNA stability in *E. coli* by the addition of synthetic hairpins using a 5' cassette system. *Biotechnol. Bioeng.* 55, 577–580. doi: 10.1002/(SICI)1097-0290(19970805)55:3<577::AID-BIT16>3.0.CO;2-D
12. Ceroni F, Algar R, Stan G. B, Ellis T. (2015) Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* 12, 415–418. doi: 10.1038/nmeth.3339

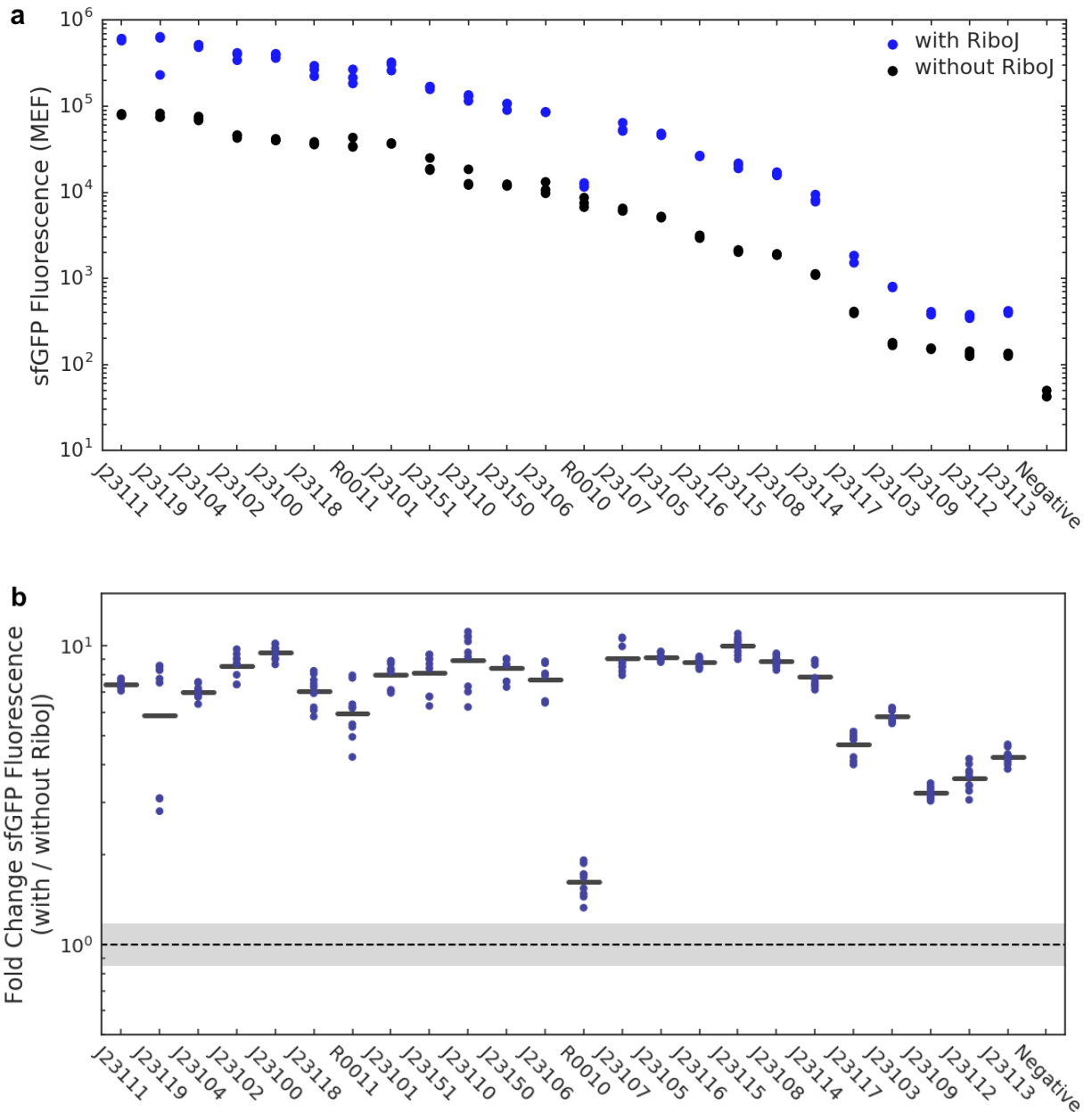


Figure 1: (a) Absolute fluorescence of constructs denoted by BioBrick ID (**Supplementary Table 1**), with (blue) and without (black) RiboJ insulation as measured by calibrated flow cytometry. Each dot represents the geometric mean fluorescence of $n > 10,000$ cells. **(b)** Fold change in fluorescence of constructs when insulated with RiboJ. Bars represent the fold change in the mean fluorescence across replicates, and dots represent all pairwise fold changes between replicates. The dashed line and grey region indicate one geometric SD factor around the geometric mean of a null fold change distribution computed from the fluorescence data (**Supplemental Methods**).

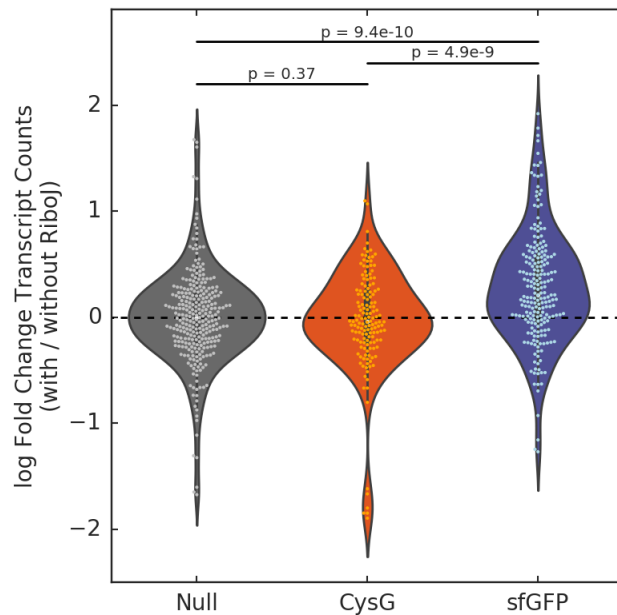


Figure 2: Fold change in the transcript abundance of CysG, sfGFP, and null distribution when promoter constructs are insulated with RiboJ. P values were calculated using Welch's one-tailed t-tests with hypotheses $sfGFP > Null$ ($p = 9.41e-10$) and $sfGFP > CysG$ ($p = 4.94e-09$). For the comparison of Null and CysG, Welch's two-tailed t-test was used ($p = 0.37$). Dots represent all pairwise ratios of replicates (9 dots per promoter). The null distribution was calculated by comparing all pairwise ratios of sfGFP transcript counts within a RiboJ condition (with/with or without/without) for a given promoter, excluding identity ratios (where a sample is divided by itself) (**Supplementary Methods**).