Engineered bacterial M1GS ribozyme efficiently cleaves the most abundant ribosomal RNA in a human cancer cell line

Neha Priyadarshini[1†], Navinchandra Venkatarama Puppala[1†], Harikrishnan Poiyamozhi[1], Swati Biswas[2], and Gireesha Mohannath[1*]

[1]Department of Biological Sciences, Birla Institute of Technology & Science-Pilani, Hyderabad Campus, Hyderabad, Telangana, India.

[2]Department of Pharmacy, Birla Institute of Technology & Science-Pilani, Hyderabad Campus, Hyderabad, Telangana, India.

† These authors contributed equally to this work
*Correspondence:
Email: gireesha@hyderabad.bits-pilani.ac.in

Abstract

Ribonuclease P (RNase P), a ribozyme conserved across all domains of life, is involved in the tRNA 5′ maturation. RNase P recognizes precursor tRNA based on its structure, not the tRNA sequence. This feature had been exploited to engineer RNase P to selectively target and cleave any RNA as a gene inactivation strategy. Of these, a strategy called M1GS involves tethering an *Escherichia coli* M1 RNA to a short stretch of guide sequence (GS), which is complementary to the RNA targeted for cleavage. Despite its simplicity and versatility, M1GS tool appears to be underutilized compared to other gene inactivation strategies. Perhaps one of the reasons is that employment of the M1GS strategy requires prior knowledge about the requirements of the M1GS target sites. To facilitate its broader use, we have developed a Python script-based user-friendly bioinformatic tool built based on the requirements of M1GS to predict its target sites for any given RNA (using either DNA or RNA sequence as an input). In this study, we first demonstrate the utility of the bioinformatic tool in predicting M1GS target sites for human 28S rRNA and then we show that the customized M1GS-mediated downregulation of 28S rRNA in a human cancer cell line. We further validate the bioinformatic tool by predicting M1GS target sites for two previously targeted RNAs. Lastly, we discuss the utility of M1GS ribozyme-mediated rRNA downregulation as a potential anticancer modality in cancers where rRNAs are upregulated.

Keywords
RNase P ribozyme, M1GS target site prediction, RNA cleavage, human rRNA, gene regulation

1. Introduction

Targeted RNA cleavage or degradation has commonly been employed as a reverse-genetic tool to understand genes' functions and as a defense modality against pathogens, including viruses. The commonly used tools involve RNA interference (RNAi)/antisense RNA (Arechavala-Gomeza & Garanto, 2022; Chery, 2016; Davidson & McCray Jr, 2011; Isenmann et al., 2023), engineered ribozymes (Puerta-Fernández et al., 2003), and recently developed CRISPR-Cas-based genome editing tools (Gunitseva et al., 2023; Terns, 2018).

The Ribonuclease P (RNase P) RNA is among the first discovered ribozymes, and it is a ribonucleoprotein (RNP) complex found in all three domains of life (Gopalan et al., 2018; Gopalan et al., 2002). Although RNase P has a shared function in the 5' tRNA maturation, the composition of RNase P RNP complex is evolutionarily divergent (Esakova & Krasilnikov, 2010; Evans et al., 2006; Jarrous & Gopalan, 2010; Klemm et al., 2016; Lai et al., 2010; Liu & Altman, 2010). In all three domains of life, RNase P RNP constitutes a single catalytic RNase P RNA (RPR) subunit with a shared common ancestry and the RNase P protein (RPP) component that varies in number and identity. A single RPP subunit in Bacteria, up to five RPPs in Archaea, and up to ten RPPs in Eukarya have been reported (Gopalan et al., 2002; Klemm et al., 2016). Interestingly, recent studies have demonstrated the co-existence of both protein-only and RNP-based RNase P in diverse eukaryotes (Gobert et al., 2010; Holzmann et al., 2008; Lai et al., 2011; Wang et al., 2010).

In *Escherichia coli*, RNase P is made up of a catalytic RNA component called M1 RNA and a protein cofactor called C5 (Altman et al., 1986; Kole & Altman, 1979; Vioque & Altman, 2001). It is involved in tRNA 5′ maturation (Altman, 1989; Esakova & Krasilnikov, 2010; Evans et al., 2006; Frank & Pace, 1998; Gopalan et al., 2018; Gopalan et al., 2002). M1 RNA possesses catalytic activity *in vitro* without a protein cofactor (Guerrier-Takada et al., 1983). RNase P recognizes precursor tRNA based on its structure, not the tRNA sequence, and this feature was explored to engineer RNase P to selectively target and cleave any RNA molecule as a gene inactivation strategy (Forster & Altman, 1990; Gopalan et al., 2002; Jarrous & Liu, 2023; Trang et al., 2004).

M1GS is an RNase P-based gene inactivation strategy, where in *E. coli* M1 RNA is fused through its 3' end to a guide sequence (GS), which is complementary to the target RNA. When the M1GS is expressed in cells, the GS hybridizes to the target RNA which will then be cleaved by the tethered M1 RNA at immediately upstream of the binding site (Fig. 1) (Liu & Altman, 1995, 1996, 2010). If M1GS is used to target a mRNA (usually the 5' end), it typically leads to the mRNA degradation, resulting in the reduced expression or a knock-down of the

targeted gene. However, M1GS has certain sequence requirements for effective binding to and cleavage of the target RNA, which are described below in the Results section.

The M1GS strategy has been successfully used to downregulate genes both in prokaryotes and eukaryotes, demonstrating its utility as an effective tool for disrupting genes' function(s) (Mao et al., 2014; Trang, Lee, et al., 2000; Ujor et al., 2021). Moreover, the M1GS strategy is simple, entailing the expression of a single exogenous M1 RNA fused to a few other short sequences, including GS, with a maximum total length of ~ 450 nt. Further, it can be expressed using various types of inducible and constitutive promoters (Gao et al., 2021; Henriksen et al., 2007). Multiple studies have included an inactive M1GS carrying enzymatically deficient mutant version of M1 RNA in their studies to demonstrate that the phenotype observed is due to the enzymatically active M1GS-mediated cleavage, not due to the GS-target RNA base pairing or any other random effects M1GS expression (Liu & Altman, 1995; Trang, Kilani, et al., 2000; Trang, Lee, et al., 2000; Ujor et al., 2021). However, despite its simplicity, specificity, and the effectiveness of the engineered M1GS ribozyme system in selectively cleaving an RNA, it seems to be underutilized. Perhaps one of the reasons is that some prior knowledge of the M1GS target site requirements is necessary for its employment, discouraging its frequent usage. Towards enabling facile employment of the M1GS strategy, we designed a Python script-based user-friendly bioinformatic tool to predict M1GS target sites for any given RNA using either DNA or RNA sequence as an input.

Further, we used this bioinformatic tool to predict M1GS target sites in human 28S ribosomal RNA (28S rRNA), which is part of human 45S rRNA genes. In the nucleolus, RNA polymerase I (RNA Pol I) transcribes 45S rRNA genes to produce pre-45S rRNA, which is then processed to 18S, 5.8S, and 28S rRNAs, which form the catalytic cores of ribosomes, the cellular protein synthesising machinery (Fig. 2). 45S ribosomal RNA genes are upregulated in several cancers (Tsoi et al., 2017; Uemura et al., 2012; Zhou et al., 2016) Therefore, RNA Pol I transcription has been therapeutically targeted these cancers (Bywater et al., 2012; Drygin et al., 2011; El Hassouni et al., 2019; Ferreira et al., 2020; Hannan et al., 2013; Low et al., 2019; Wei et al., 2018)

In this study, we first demonstrate the utility of our bioinformatic tool in predicting M1GS targets sites for human 28S rRNA and then we demonstrate that the M1GS ribozyme, customized to target 28S human rRNA successfully cleaves the target rRNA in a lung cancer cell line, resulting in significant rRNA downregulation. We further validate our tool by predicting M1GS target sites for two previously published M1GS targeted RNAs. Lastly, we discuss the potential advantages of M1GS ribozyme and utility of M1GS ribozyme-mediated rRNA downregulation as a potential anticancer modality in cancers where rRNA are upregulated.

Materials and Methods

Prediction of the target RNA secondary structure

If the secondary structure of the target RNA is not provided by the user in the Vienna format in the form of ".", "(", ")") as an input, the tool uses the secondary structure predictions provided by the RNAfold webserver (http://rna.tbi.univie.ac.at/cgibin/RNAWebSuite/RNAfold.cgi) (Gruber et al., 2008). The default settings of RNAfold are used for predicting the secondary structure and those settings are listed in Figure S1.

The web interface

The web interface for the tool was developed using Python3 and Django 3.1.6. The web interface mainly consists of two sections: the inputs and the output. To find the target sites, the interface runs the program in the background and the time taken for this task varies depending on the size of the target RNA. Once this task is over, the tool loads the output above 'the inputs' section. The bioinformatic tool M1GS target site finder can be accessed at the following link http://m1gstargetfinder.bits-hyderabad.ac.in:8000/m1gs/

Plasmid Construction

Designed M1GS expression cassettes consisted of H1 promoter, followed by wildtype (active) or mutant (inactive) M1 RNA, a linker region, a 15 bp guide sequence (GS) (5'-GCCAATCCTTATCCC-3') to target human 28S rRNA (Fig. S2), and a self-cleaving hammerhead ribozyme (HH). The GS designed was to cleave 28S rRNA at between nts 2862 and 2863, a target site predicted by the bioinformatic tool. These expression cassettes were commercially synthesised (ABclonal), which were then cloned into the pRS retroviral vector using EcoRI and XhoI restriction sites. Plasmid DNA from the positive clones was isolated using *QIAprep Spin Miniprep Kit* (Qiagen) and the presence of the M1GS constructs was verified through PCR amplification using primers (Fig. S3) specific for active and inactive M1 RNA.

Human Cell Culture

The human lung cancer cell line A549 (ATCC CCL-185) was purchased from the American Type Culture Collection (Manassas, VA, USA). The cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (HiMedia), supplemented with 10% (v/v) fetal bovine serum (HiMedia) and 1% (w/v) Penicillin streptomycin (HiMedia), at 37°C in a humidified incubator with 5% CO_2. The medium was changed every 2 days until the cells reached 70–80% confluence. For detaching the cells during passages, 0.25% (w/v) Trypsin-EDTA (HiMedia) was used.

Transfection

Transient transfection of A549 cells was performed using jetPRIME® kit (PolyPlus). Once the cells in 6-well plate reached 70-80% confluency, a transfection master mix consisting of 2 μg plasmid DNA containing active (pRS-M1GS^{28SrRNA}) or inactive M1GS (pRS-ΔM1GS^{28SrRNA}) and 4 µl of JetPrime reagent (1:2 ratio) in 200 µL of JetPrime® buffer (PolyPlus) was added for each well and then the cells were incubated at 37°C for 24 h. For transfection, the cells were seeded at a density of 2,50,000 cells per well in a 6-well plate. After transfection, cells were supplemented with additional media. The cells were harvested 24 h post-transfection and processed immediately for RNA and DNA isolation.

DNA and RNA isolation and semiquantitative reverse transcription (RT)-PCR

Total DNA and RNA were isolated from the control and the transfected cell lines using *AllPrep DNA/ RNA/Protein Mini Kit* (Qiagen). To confirm the presence of the M1GS constructs in the transfected cells, PCR amplification of M1GS constructs and the endogenous glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene (as an internal control) was carried using initial denaturation for 2 min at 95°C, followed by 30 cycles of 30 sec at 95°C, 30 sec at 58°C, and 30 sec at 72°C, and a final extension for 3 min at 72°C, in the presence of *GoTaq® Green Master Mix* (Promega).

The isolated total RNA was treated with *Turbo DNA-free* kit (Invitrogen) to eliminate any contaminating genomic DNA. 1 μg of DNA-free RNA was reverse transcribed to cDNA using *SuperScript III First Strand cDNA synthesis* kit (Invitrogen). One μl (cDNA from ~50 ng DNA-free RNA) of reverse transcribed product was then used for PCR amplification of active and inactive M1GS ribozymes, HH ribozyme, and GAPDH control with initial denaturation for 2 min at 95°C, followed by 25 cycles of 30 sec at 95 °C, 30 sec at 58 °C, and 30 sec at 72 °C, and a final extension for 3 min at 72 °C. The same PCR profile was used for amplification of 28S rRNA but for 23 cycles. For all the semiquantitative RT-PCR reactions, RT- negative controls were included.

Quantitative PCR

Quantitative PCR (qPCR) was performed using a *Step One Plus* real-time PCR system (Applied Biosystems) in the presence of *TB Green® Premix Ex Taq™* (Tli RNase H Plus) (Takara). The qPCR amplification was performed using an initial denaturation at 95 °C for 10 sec, followed by 40 cycles of 15 sec at 95 °C and 60 sec at 60 °C. The relative gene expression of 28S RNA was calculated using the 2^(-△△CT) method (Livak & Schmittgen, 2001) and the expression levels were normalised to the expression of GAPDH as a reference gene. The qPCR was performed in duplicates for each sample. Primer sequences used to amplify 28S

rRNA span the M1GS cleavage site (2862). All the primers used in qPCR reactions are listed in Figure S3.

Statistical analysis

Data from two technical replicates from two independent batches of cell lines were used for all the calculations and to generate error bars. The qPCR data were used to generate a graph using boxplot analysis. The standard two-tailed t-test was used to analyze significant differences in gene expression levels between the control (non-transfected) and the transfected cells expressing active or inactive M1GS. A p of < 0.05 was considered statistically significant.

Alignments:

Alignment of the active and the inactive M1 RNA sequences was carried out using nucleotide BLAST (blastn) tool from NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_SPEC=GeoBlast&PAGE_TYPE=BlastSearch)

Other Softwares used:

Figures were made using the free version of Figma and the licensed version of PowerPoint.

Results

Since identifying potentially the best possible M1GS target sites would help achieve a higher success rate in cleaving the target RNA, we developed a bioinformatic tool with a user-friendly web interface for automated identification of all the possible M1GS cleavage sites for any target RNA. It provides the flexibility to customize the need-based predictions with various input options. A flow chart listing various steps of our *M1GS target finder* tool is given in Figure 3. The several input options that an end user can provide are described below.

Essential inputs required for M1GS target site finder

Sequence of the target RNA

M1GS can be used to cleave both coding and non-coding RNA. If mRNA of any protein-coding gene is targeted, targeting the region towards 5' end of the mRNA is preferred because single cleavage in the 5' end results in the loss of 5' protective guanosine cap and subsequent degradation of the mRNA by ribonucleases. On the other hand, if 3' end of mRNA is targeted, most of the target mRNA might still be available for translation to occur at some levels. To identify the sites throughout the target RNA, as the first step, the sequence of the target RNA or DNA should be entered in the given space in the tool which says 'Sequences of target

RNA/DNA'. If you provide DNA sequence as input, the tool will automatically convert it to the corresponding RNA sequence.

The nucleotide at -1 position

The 5' nucleotide with respect to the M1GS cleavage site (-1 position) is a pyrimidine (Liu & Altman, 1996, 2010). Uracil is the most conserved preference at the -1 position of the pre-tRNA in bacteria and archaea (Zahler et al., 2003). However, some bacterial species also have adenine, cytosine, or guanine at the -1 position, demonstrating that cleavage is also possible with nucleotides other than Uracil at -1. Several studies have demonstrated successful customized M1GS-mediated cleavage of the target mRNA with varying nucleotides at the -1 position (Cobaleda & Sánchez-Garcıa, 2000; Liu et al., 2023; Toumpeki et al., 2018). Therefore, we designed our tool to provide a choice of all the four nucleotides at the -1 position. The end users may prefer a nucleotide of their choice.

The nucleotide at +1 position

The identity of the base at the +1 position is a salient determinant for the M1GS cleavage site selection. When a purine is at the +1 position, cleavage occurs exactly between the -1 and +1 sites, regardless of whether the +1 nucleotide is base-paired or not (Liu & Altman, 1996). In contrast, when a pyrimidine is at the +1 position and is not base-paired, results in an aberrant cleavage (Liu & Altman, 1996). Moreover, cleavage of a substrate with a purine at the +1 position is more efficient than with a pyrimidine at +1 position. Therefore, a purine nucleotide at the +1 position is a dominant determinant for M1GS cleavage site selection. This observation is consistent with the notion that a guanine at the +1 position, which has been found in most bacterial pretRNAs, serves as the guide nucleotide for M1 RNA and RNase P cleavage in *E. coli* (Liu & Altman, 1996). Therefore, we have provided in this tool an option to choose either A or G at the +1 position.

Length of the desired GS

The optimal length of the GS can vary from 12-19 nt (Liu & Altman, 1995). Longer the length of the GS, higher the specificity of the targeting but this may come at a cost of lower turnover number for the M1GS ribozyme due to tighter binding of the M1GS to the target mRNA. By contrast, shorter GS could lead to an enhanced M1GS turnover number but could result in more off-target cleavages. In addition to the length of GS, the GC% of GS also contributes to the M1GS ribozyme cleavage efficiency. Based on the length of the guide sequence, target sites of an equal length in the target RNA will be predicted. Note that the length of GS does not change the positions or number of the M1GS target sites in the target RNA.

Desirable length of the GS-binding target RNA length from the -1 position to be in a single-stranded configuration

Although it is desirable to target the specific sequence of the target RNA which likely exists in the single-stranded form *in vivo*, it is often not possible to find such a sequence based on the *in silico* secondary structure prediction. M1GS could still bind and cleave such RNAs (Liu & Altman, 1995). Nevertheless, the end users can still try to find such a desirable target RNA sites by providing desirable length of the GS-binding target RNA length from the -1 position to be in a single-stranded configuration. For example, if the length of the GS is 14 nt, one can start with choosing the desirable length of the target RNA in single-stranded form as 14 nt. However, if enough target sites or no target sites are found, the end user is recommended to keep reducing this number till finding a minimum number of desired target sites. In case, if the softwares used predicts that no region of the input RNA sequence exists in single-stranded configuration, the input for desirable length of the single-stranded target RNA should be “0” and M1GS could still cleave such targets, as described above. This is because all the secondary structures are just predictions, and the guide sequence might still bind to these target sites for the cleavage to occur (Liu & Altman, 1995). Moreover, there is a possibility that the secondary structures of RNA could be weak and could be disrupted in the presence of the customized M1GS, favoring the base pairing between the target RNA and the M1GS.

Optional inputs

Secondary structure of the target RNA:

For efficient M1GS cleavage, it is preferrable to choose the region of the target RNA that is likely to exist in single-stranded form. To identify regions in the target RNA that are easily accessible (single-stranded regions) for the guide sequence to efficiently bind, several studies have used chemical methods (Liu & Altman, 1995; Trang, Kilani, et al., 2000) and prediction softwares (Mao et al., 2014; Ujor et al., 2021). For this tool, we use secondary structures predicted by publicly available RNAfold software tool (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). Providing the secondary structure of the target RNA is an optional field in the tool. If this field is left empty, by default the tool uses the RNAFold web server to predict the secondary structure of the target RNA based on the given input sequence. If a custom secondary structure is available, it should be entered in the Vienna format in this field. For example, another RNA folding prediction software "Mfold" gives the secondary structure output in multiple formats, including the Vienna format, which can be given as the input. Often, different prediction softwares give different predicted secondary structures for the same RNA sequence, as demonstrated below in Results and Discussion section. Therefore, we provided the flexibility in the tool to choose either the default RNAfold or any other software of users’ choice.

Unpaired nucleotide at -1 position

This is an additional feature added to choose if the nucleotide at the -1 position according to the predicted secondary structure is base-paired or free. This is not a compulsory feature. The idea is to keep the cleavage site on the target RNA as free as possible. However, cleavage can occur when the nucleotide at -1 position is paired (Mao et al., 2014).

Unpaired CCA-3'

To mimic *E. coli* pre-tRNA substrates, an unpaired NCCA is added at the 3' end of the M1GS RNA (Liu & Altman, 2010). The 3' CCA sequence is important for cleavage efficiency but not essential for cleavage (Liu & Altman, 1996). Hence, we provide an option in our tool to select unpaired 3' CCA where possible. However, if having an unpaired 3' CCA becomes a limiting factor in finding M1GS target sites, this feature may not be chosen.

Avoiding off-target effects

The provided targeted sites need to be unique to avoid off-target effects. To avoid such off-target effects, the user is recommended to verify that the chosen target sites do not match any other part of the genome of interest using any of the available bioinformatic tools as deemed appropriate.

Running the tool

Upon entering input RNA sequence (with or without secondary structure data) and choosing the desirable features of the GS and the target RNA, the user can run the tool by clicking 'find target sites'.

Output

The output (results) provides a list of all the possible M1GS target sites as per the inputs provided. The list consists of three components: The position of -1 site in the target RNA, target site sequence (from +1), and the GC content of the target site sequence, as described below.

Prediction of M1GS target sites for human 28S rRNA

To demonstrate that the bioinformatic tool we designed works as described, we ran the program using a human 28S ribosomal RNA sequence (GenBank accession number: KY962518.1) (Figure S2) as an input. We started with stringent input options. For the nucleotide at -1 position, pyrimidines (C and U) were selected and for the nucleotide at +1 position, purine (G) was selected. Lastly, an unpaired CCA at the 3' end and an unpaired nucleotide at the -1 position was opted. The output listed multiple target sites (Fig. 4A), of

which one at the position 2862 (highlighted by a red rectangle in Fig. 4A) was used to generate a guide sequence of 15 bp for cleaving 28S rRNA. Further, we demonstrate that the tool will provide more target sites if we relax the input options. For example, when we excluded the requirement of unpaired CCA at 3' end and unpaired nucleotide at the -1 position, the output listed a more number of target sites compared to the output obtained when the same two features were included as requirements (compare output of Fig. 4A with Fig. 4B).

Demonstration of M1GS-mediated downregulation of 28S rRNA in a lung cancer cell line

A guide sequence of 15 nt, as described above, was designed to target 28S rRNA using a wild-type M1 RNA ($M1GS^{28SrRNA}$) or enzymatically inactive mutant M1 RNA ($\Delta M1GS^{28SrRNA}$), as a negative control. The mutant M1 RNA mutant (ΔM1) is used as a control to delineate the antisense effects, if any, that could arise due to base pairing of the GS with the target RNA (Gopalan et al., 2002; Liu & Altman, 1995). A cartoon depicting the different elements of the designed M1GS constructs is given in Figure 5A. Sequences used in the M1GS constructs, sequence alignment of the active and the inactive M1 RNA, and a PCR assay to differentiate the active and the inactive M1 RNA are given in Figure S4.

M1GS can be expressed either with 3' overhang (Fig. 5B) or without a 3' overhang sequence by inserting a self-cleaving hammerhead ribozyme at the 3' end of GS (Fig. 5C). Usage of hammerhead ribozyme at the 3' end of the GS has been recommended for an efficient cleavage because additional sequences downstream from the 3' CCA sequence had been shown to lower the rate of cleavage of substrates by M1 RNA (Liu & Altman, 1995). Therefore, we introduced a self-cleaving hammerhead (HH) ribozyme (Ujor et al., 2021), downstream of the M1GS to ensure a free 3′ end for M1GS RNA. We transfected the human lung cancer cell line A549 with $M1GS^{28SrRNA}$ and $\Delta M1GS^{28SrRNA}$ in two independent batches. Following transfection, DNA-free RNA was obtained and used to carry out quantitative PCR (qPCR) and semiquantitative RT-PCR assays.

Results from the qPCR assays showed a statistically significant reduction in the levels of 28S rRNAs in $M1GS^{28SrRNA}$-expressing cells compared to $\Delta M1GS^{28SrRNA}$-expressing and the non-transfected control cells (Fig. 6A). We recapitulated similar results in a semiquantitative RT-PCR assay (Fig. 6B). To ensure that the reduction of 28S rRNA in the transfected cells expressing the active $M1GS^{28SrRNA}$ compared to the cells expressing $\Delta M1GS^{28SrRNA}$ or the non-transfected cells, we carried out PCR or RT-PCR assays, to establish successful transfection of all the plasmids at the DNA level, and expressions of both the active and inactive M1GS only in the transfected cells at the RNA level (Fig. 6B). In the transfected cells, we also demonstrated self-cleavage of HH ribozyme that was part of the M1GS constructs (Fig. 6B).

Further validation of the bioinformatic tool

Lastly, we demonstrate that our tool can sucessfully predict target sites for two previously published target RNAs that were cleaved by customized M1GS. One of the targets is mRNA sequence of the thymidine kinase (TK) of human herpes simplex virus 1 (HSV-1) (Liu & Altman, 1995). The sequence of this gene and a screenshot of the output with the predicted M1GS targets sites are given in Figure S5.

The other target is the 5' UTR of hepatitis C virus (HCV) (Mao et al., 2014). For this target, we used both RNAfold and Mfold separately to predict the secondary structure of the HCV RNA. The results indicate that different softwares provide different secondary structure predictions for the same RNA sequence (Fig. S6). For a 13 nt GS with a requirement of 6 nt to be in the single-stranded form, when the secondary structure as predicted by the Mfold software was fed to the tool, it provided one target site which was used to design GS to cleave the target RNA (Fig. S6.B) (Mao et al., 2014). However, for the same input options, when the secondary structure was predicted by RNAfold software, no M1GS target sites were identified by the tool because accroding to the RNAfold, stretch of 6 nt is not predicted to exist in a single-stranded form (Fig. S6.C). However, when the input of requirement of single-stranded form target RNA was set to '0' the tool predicted three M1GS target sites including the one used in the study (Fig. S6.C) (Mao et al., 2014).

Discussion

We describe here a user-friendly bioinformatic tool to predict M1GS cleavage sites in any target RNA. Ideally, the target RNA region should exist in a single-stranded form for effective M1GS binding, although this feature is not necessary. Therefore, we have designed this tool to subject any input RNA (or DNA) sequence for RNA secondary structure analysis by linking the tool to publicly available RNAfold software (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). Alternatively, the users can provide RNA secondary structure input in Vienna format obtained from any other RNA secondary structure prediction tool of their choice. The tool is designed to choose various features of the GS as per the users' requirements. Having flelxibility in input options would enhance the chances of obtaining M1GS target sites for various target RNA which vary both in the length and the squence. In this manuscript, we demonstrated that the bioinformatic tool can successfully predict M1GS target sites by using either DNA or RNA as an input (Figs. 3, 4, S5, and S6). We also demonstrated that changing the stringency of M1GS requirements would accordingly change the number of predicted M1GS target sites (Fig. 4).

Compared to RNA interference (RNAi) and other gene-targeting strategies, ribozymes have some advantages. One is that they are ribonucleic acids with a catalytic activity that can make or break covalent bonds within a nucleic acid, often without the necessity of protein

cofactors (Grassi & Marini, 1996; Guerrier-Takada et al., 1983; Kruger et al., 1982). On the other hand, RNAi, for example, requires several cellular factors such as Exportin V, Dicers, Argonaute proteins to process small interfering RNAs (siRNAs) and to mediate gene silencing (Wilson & Doudna, 2013). Moreover, several viruses have evolved counter-defense mechanisms to block RNAi-mediated host defense responses (Csorba et al., 2015; Li & Ding, 2005). By contrast, viruses lack counter-defense mechanisms against engineered ribozymes as they don't encounter them naturally during infections. Further, modulation of expression levels of ribozymes is lot easier as it requires expression of only one component, compared to a few other systems, including CRISPR-Cas-based systems, which require simultaneous modulation of multiple components. Customized ribozymes seldom have been used as a gene targeting tool or antiviral agent, despite their versatile utility (Mao et al., 2014; Shin et al., 2004; Trujillo et al., 2019). We hope that the availability of a user-friendly bioinformatic tool to predict M1GS target sites would encourage more users to employ M1GS tool in RNA-cleaving/reverse genetic experiments.

We also demonstrated that customized M1GS strategy can be used to downregulate nucleolar rRNA in human cells (Figs 5 and 6). Further experiments are needed to determine if M1GS-mediated downregulation of rRNA possesses any anti-cancer activity in cancers where rRNA levels are upregulated. Direct downregulation of rRNA, if proven successful in exhibiting anticancer activity, may potentially serve as a better strategy compared to targeting RNA polI because RNA polI may be necessary for other cellular functions (Wong et al., 2024). Previous studies have demonstrated systemic delivery of plasmids containing M1GS constructs in mouse models by using attenuated bacteria and successfully cleaved the target RNA (Bai et al., 2011; Bai et al., 2010). Several attenuated bacterial strains have been explored as components of gene therapeutic strategies (Vassaux et al., 2006).

Acknowldegements

The authors are also thankful to all those from the biological sciences and other departments at BITS-Pilani Hyderabad campus, who have supported this work in one way or another.

Author Contributions: N.P., N.V.P, S.B., and G.M. designed research, N.P., N.V.P., and H.P. carried out research, N.P., N.V.P., H.P., and G.M. analyzed data, N.P., N.V.P., and G.M. wrote the manuscript. All authors reviewed the manuscript. The authors are profusely thankful to Dr. Venkat Gopalan and Dr. Lien Lai, The Ohio State University (USA), for their valuable support in sharing reagents, expertise, and critique of the manuscript.

FUNDING

G.M. and N.P. are thankful to the Birla Institute of Science and Technology (BITS) Pilani, Hyderabad campus, for the research grant they received as part of Centre for Human Disease Research and an Institute Fellowship to N.P. G.M. is also thankful to the Science and Engineering Research Board (SERB), Government of India, for the Ramanujan Fellowship Research Grant (SB-S2-RJN-062-2017). N.P. is thankful to the Indian Council of Medical Research (ICMR), Government of India, for providing Senior Research Fellowship (F.No.2021-13491/GTGE-BMS) to work on this project. N.V.P. is thankful for the internship received as part of the Scientific Social Responsibility (SSR) component of CRG grant (CRG/2020/002855) awarded to G.M.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Altman, S. (1989). Ribonuclease P: an enzyme with a catalytic RNA subunit. *Adv Enzymol Relat Areas Mol Biol*, *62*, 1-36.

Altman, S. (2011). Ribonuclease P. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1580), 2936-2941.

Altman, S., Baer, M., Guerrier-Takada, C., & Vioque, A. (1986). Enzymatic cleavage of RNA by RNA. *Trends in biochemical sciences*, *11*(12), 515-518.

Arechavala-Gomeza, V., & Garanto, A. (2022). Antisense RNA therapeutics: a brief overview. *Antisense RNA Design, Delivery, and Analysis*, 33-49.

Bai, Y., Gong, H., Li, H., Vu, G.-P., Lu, S., & Liu, F. (2011). Oral delivery of RNase P ribozymes by Salmonella inhibits viral infection in mice. *Proceedings of the National Academy of Sciences*, *108*(8), 3222-3227.

Bai, Y., Li, H., Vu, G.-P., Gong, H., Umamoto, S., Zhou, T., Lu, S., & Liu, F. (2010). Salmonella-mediated delivery of RNase P-based ribozymes for inhibition of viral gene expression and replication in human cells. *Proceedings of the National Academy of Sciences*, *107*(16), 7269-7274.

Bywater, M. J., Poortinga, G., Sanij, E., Hein, N., Peck, A., Cullinane, C., Wall, M., Cluse, L., Drygin, D., & Anderes, K. (2012). Inhibition of RNA polymerase I as a therapeutic strategy to promote cancer-specific activation of p53. *Cancer cell*, *22*(1), 51-65.

Chery, J. (2016). RNA therapeutics: RNAi and antisense mechanisms and clinical applications. *Postdoc journal: a journal of postdoctoral research and postdoctoral affairs*, *4*(7), 35.

Cobaleda, C., & Sánchez-Garcıa, I. (2000). In vivo inhibition by a site-specific catalytic RNA subunit of RNase P designed against the BCR-ABL oncogenic products: a novel approach for cancer treatment. *Blood, The Journal of the American Society of Hematology*, *95*(3), 731-737.

Csorba, T., Kontra, L., & Burgyán, J. (2015). viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. *Virology*, *479-480*, 85-103. https://doi.org/https://doi.org/10.1016/j.virol.2015.02.028

Davidson, B. L., & McCray Jr, P. B. (2011). Current prospects for RNA interference-based therapies. *Nature Reviews Genetics*, *12*(5), 329-340.

Deng, Q., Liu, Y., Li, X., Yan, B., Sun, X., Tang, W., Trang, P., Yang, Z., Gong, H., & Wang, Y. (2019). Inhibition of human cytomegalovirus major capsid protein expression and replication by ribonuclease P–associated external guide sequences. *Rna*, *25*(5), 645-655.

Drygin, D., Lin, A., Bliesath, J., Ho, C. B., O'Brien, S. E., Proffitt, C., Omori, M., Haddach, M., Schwaebe, M. K., & Siddiqui-Jain, A. (2011). Targeting RNA polymerase I with an oral small molecule CX-5461 inhibits ribosomal RNA synthesis and solid tumor growth. *Cancer research*, *71*(4), 1418-1430.

El Hassouni, B., Mantini, G., Immordino, B., Peters, G. J., & Giovannetti, E. (2019). CX-5461 inhibits pancreatic ductal adenocarcinoma cell growth, migration and induces DNA damage. *Molecules*, *24*(24), 4445.

Esakova, O., & Krasilnikov, A. S. (2010). Of proteins and RNA: the RNase P/MRP family. *Rna*, *16*(9), 1725-1747.

Evans, D., Marquez, S. M., & Pace, N. R. (2006). RNase P: interface of the RNA and protein worlds. *Trends in biochemical sciences*, *31*(6), 333-341.

Ferreira, R., Schneekloth Jr, J. S., Panov, K. I., Hannan, K. M., & Hannan, R. D. (2020). Targeting the RNA polymerase I transcription for cancer therapy comes of age. *Cells*, *9*(2), 266.

Forster, A. C., & Altman, S. (1990). External guide sequences for an RNA enzyme. *Science*, *249*(4970), 783-786.

Frank, D. N., & Pace, N. R. (1998). Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annual review of biochemistry*, *67*(1), 153-180.

Gao, Z., van der Velden, Y. U., Fan, M., van der Linden, C. A., Vink, M., Herrera-Carrillo, E., & Berkhout, B. (2021). Engineered miniature H1 promoters with dedicated RNA polymerase II or III activity. *J Biol Chem*, *296*, 100026. https://doi.org/10.1074/jbc.RA120.015386

Gobert, A., Gutmann, B., Taschner, A., Gössringer, M., Holzmann, J., Hartmann, R. K., Rossmanith, W., & Giegé, P. (2010). A single Arabidopsis organellar protein has RNase P activity. *Nat Struct Mol Biol*, *17*(6), 740-744. https://doi.org/10.1038/nsmb.1812

Gopalan, V., Jarrous, N., & Krasilnikov, A. S. (2018). Chance and necessity in the evolution of RNase P. *Rna*, *24*(1), 1-5. https://rnajournal.cshlp.org/content/24/1/1.full.pdf

Gopalan, V., Vioque, A., & Altman, S. (2002). RNase P: variations and uses. *Journal of Biological Chemistry*, *277*(9), 6759-6762.

Grassi, G., & Marini, J. C. (1996). Ribozymes: structure, function and potential therapy for dominant genetic disorders. *Annals of medicine*, *28*(6), 499-510.

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R., & Hofacker, I. L. (2008). The vienna RNA websuite. *Nucleic Acids Res*, *36*(suppl_2), W70-W74.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, *35*(3), 849-857.

Gunitseva, N., Evteeva, M., Borisova, A., Patrushev, M., & Subach, F. (2023). RNA-Dependent RNA Targeting by CRISPR-Cas Systems: Characterizations and Applications. *Int J Mol Sci*, *24*(8). https://doi.org/10.3390/ijms24086894

Ha, S., Zhou, H., Gautam, M., Song, Y., & Wang, C. (2019). Reduced ribosomal RNA expression and unchanged ribosomal DNA promoter methylation in oral squamous cell carcinoma. *Molecular Genetics & Genomic Medicine*, *7*(7), e00783.

Hannan, K. M., Sanij, E., Rothblum, L. I., Hannan, R. D., & Pearson, R. B. (2013). Dysregulation of RNA polymerase I transcription during disease. *Biochimica Et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1829*(3-4), 342-360.

Henriksen, J. R., Løkke, C., Hammerø, M., Geerts, D., Versteeg, R., Flaegstad, T., & Einvik, C. (2007). Comparison of RNAi efficiency mediated by tetracycline-responsive H1 and U6 promoter variants in mammalian cell lines. *Nucleic Acids Res*, *35*(9), e67. https://doi.org/10.1093/nar/gkm193

Hitrik, A., Abboud-Jarrous, G., Orlovetskie, N., Serruya, R., & Jarrous, N. (2016). Targeted inhibition of WRN helicase by external guide sequence and RNase P RNA. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1859*(4), 572-580.

Holzmann, J., Frank, P., Löffler, E., Bennett, K. L., Gerner, C., & Rossmanith, W. (2008). RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell*, *135*(3), 462-474. https://doi.org/10.1016/j.cell.2008.09.013

Isenmann, M., Stoddart, M. J., Schmelzeisen, R., Gross, C., Della Bella, E., & Rothweiler, R. M. (2023). Basic Principles of RNA Interference: Nucleic Acid Types and In Vitro Intracellular Delivery Methods. *Micromachines*, *14*(7), 1321.

Jarrous, N., & Gopalan, V. (2010). Archaeal/eukaryal RNase P: subunits, functions and RNA diversification. *Nucleic Acids Res*, *38*(22), 7885-7894. https://doi.org/10.1093/nar/gkq701

Jarrous, N., & Liu, F. (2023). Human RNase P: overview of a ribonuclease of interrelated molecular networks and gene-targeting systems. *Rna*, *29*(3), 300-307.

Jiang, X., Chen, Y.-C., Gong, H., Trang, P., Lu, S., & Liu, F. (2012). Ribonuclease P-mediated inhibition of human cytomegalovirus gene expression and replication induced by engineered external guide sequences. *RNA biology*, *9*(9), 1186-1195.

Klemm, B. P., Wu, N., Chen, Y., Liu, X., Kaitany, K. J., Howard, M. J., & Fierke, C. A. (2016). The diversity of ribonuclease P: protein and RNA catalysts with analogous biological functions. *Biomolecules*, *6*(2), 27.

Kole, R., & Altman, S. (1979). Reconstitution of RNase P activity from inactive RNA and protein. *Proceedings of the National Academy of Sciences*, *76*(8), 3795-3799.

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., & Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*, *31*(1), 147-157.

Lai, L. B., Bernal-Bayard, P., Mohannath, G., Lai, S. M., Gopalan, V., & Vioque, A. (2011). A functional RNase P protein subunit of bacterial origin in some eukaryotes. *Mol Genet Genomics*, *286*(5-6), 359-369. https://doi.org/10.1007/s00438-011-0651-y

Lai, L. B., Vioque, A., Kirsebom, L. A., & Gopalan, V. (2010). Unexpected diversity of RNase P, an ancient tRNA processing enzyme: challenges and prospects. *FEBS Lett*, *584*(2), 287-296. https://doi.org/10.1016/j.febslet.2009.11.048

Li, H. W., & Ding, S. W. (2005). Antiviral silencing in animals. *FEBS Lett*, *579*(26), 5965-5973. https://doi.org/10.1016/j.febslet.2005.08.034

Liu, F., & Altman, S. (1995). Inhibition of viral gene expression by the catalytic RNA subunit of RNase P from Escherichia coli. *Genes & development*, *9*(4), 471-480.

Liu, F., & Altman, S. (1996). Requirements for cleavage by a modified RNase P of a small model substrate. *Nucleic Acids Res*, *24*(14), 2690-2696.

Liu, F., & Altman, S. (2010). *Ribonuclease P* (S. A. Fenyong Liu, Ed. 1 ed., Vol. 10). Springer New York, NY. https://doi.org/https://doi.org/10.1007/978-1-4419-1142-1

Liu, J., Shao, L., Trang, P., Yang, Z., Reeves, M., Sun, X., Vu, G.-P., Wang, Y., Li, H., & Zheng, C. (2016). Inhibition of herpes simplex virus 1 gene expression and replication by RNase P-associated external guide sequences. *Sci Rep*, *6*(1), 27068.

Liu, Y., Chen, Y.-C., Yan, B., & Liu, F. (2023). Suppressing Kaposi's Sarcoma-Associated Herpesvirus Lytic Gene Expression and Replication by RNase P Ribozyme. *Molecules*, *28*(8), 3619.

Low, J. Y., Sirajuddin, P., Moubarek, M., Agarwal, S., Rege, A., Guner, G., Liu, H., Yang, Z., De Marzo, A. M., & Bieberich, C. (2019). Effective targeting of RNA polymerase I in treatment-resistant prostate cancer. *The Prostate*, *79*(16), 1837-1851.

Mao, X., Li, X., Mao, X., Huang, Z., Zhang, C., Zhang, W., Wu, J., & Li, G. (2014). Inhibition of hepatitis C virus by an M1GS ribozyme derived from the catalytic RNA subunit of Escherichia coli RNase P. *Virology journal*, *11*, 1-9.

Priyadarshini, N., Puppala, N. V., Jayaprakash, J. P., Khandelia, P., Sharma, V., & Mohannath, G. (2023). Downregulation of ribosomal RNA (rRNA) genes in human

head and neck squamous cell carcinoma (HNSCC) cells correlates with rDNA promoter hypermethylation. *Gene*, *888*, 147793.

Puerta-Fernández, E., Romero-López, C., Barroso-delJesus, A., & Berzal-Herranz, A. (2003). Ribozymes: recent advances in the development of RNA tools. *FEMS microbiology reviews*, *27*(1), 75-97.

Rangarajan, S., Raj, M. S., Hernandez, J. M., Grotewold, E., & Gopalan, V. (2004). RNase P as a tool for disruption of gene expression in maize cells. *Biochemical Journal*, *380*(3), 611-616.

Raval, A., Sridhar, K. J., Patel, S., Turnbull, B. B., Greenberg, P. L., & Mitchell, B. S. (2012). Reduced rRNA expression and increased rDNA promoter methylation in CD34+ cells of patients with myelodysplastic syndromes. *Blood, The Journal of the American Society of Hematology*, *120*(24), 4812-4818.

Shin, K.-S., Sullenger, B. A., & Lee, S.-W. (2004). Ribozyme-mediated induction of apoptosis in human cancer cells by targeted repair of mutant p53 RNA. *Molecular therapy*, *10*(2), 365-372.

Terns, M. P. (2018). CRISPR-based technologies: impact of RNA-targeting systems. *Mol Cell*, *72*(3), 404-412.

Toumpeki, C., Anastasakis, D., Panagoulias, I., Stamatopoulou, V., Georgakopoulos, T., Kallia-Raftopoulos, S., Mouzaki, A., & Drainas, D. (2018). Construction of an M1GS ribozyme for targeted and rapid mRNA cleavage; application on the Ets-2 oncogene. *Medicinal Chemistry*, *14*(6), 604-616.

Trang, P., Kilani, A., Kim, J., & Liu, F. (2000). A ribozyme derived from the catalytic subunit of RNase P from Escherichia coli is highly effective in inhibiting replication of herpes simplex virus 1. *Journal of molecular biology*, *301*(4), 817-826.

Trang, P., Kim, K., & Liu, F. (2004). Developing RNase P ribozymes for gene-targeting and antiviral therapy. *Cellular microbiology*, *6*(6), 499-508.

Trang, P., Lee, M., Nepomuceno, E., Kim, J., Zhu, H., & Liu, F. (2000). Effective inhibition of human cytomegalovirus gene expression and replication by a ribozyme derived from the catalytic RNA subunit of RNase P from Escherichia coli. *Proceedings of the National Academy of Sciences*, *97*(11), 5812-5817.

Trujillo, A. J., Myers, J. M., Fayazi, Z. S., Butler, M. C., & Sullivan, J. M. (2019). A discovery with potential to revitalize hammerhead ribozyme therapeutics for treatment of inherited retinal degenerations. Retinal Degenerative Diseases: Mechanisms and Experimental Therapy,

Tsoi, H., Lam, K., Dong, Y., Zhang, X., Lee, C., Zhang, J., Ng, S., Ng, S., Zheng, S., & Chen, Y. (2017). Pre-45s rRNA promotes colon cancer and is associated with poor survival of CRC patients. *Oncogene*, *36*(44), 6109-6118.

Uemura, M., Zheng, Q., Koh, C. M., Nelson, W. G., Yegnasubramanian, S., & De Marzo, A. M. (2012). Overexpression of ribosomal RNA in prostate cancer is common but not linked to rDNA promoter hypomethylation. *Oncogene*, *31*(10), 1254-1263.

Ujor, V. C., Lai, L. B., Okonkwo, C. C., Gopalan, V., & Ezeji, T. C. (2021). Ribozyme-Mediated downregulation uncovers DNA integrity scanning protein A (DisA) as a solventogenesis determinant in Clostridium beijerinckii. *Frontiers in bioengineering and biotechnology*, *9*, 669462.

Vassaux, G., Nitcheu, J., Jezzard, S., & Lemoine, N. R. (2006). Bacterial gene therapy strategies. *J Pathol*, *208*(2), 290-298. https://doi.org/10.1002/path.1865

Vioque, A., & Altman, S. (2001). *Ribonuclease P*. Oxford: Elservier Science.

Wang, G., Chen, H. W., Oktay, Y., Zhang, J., Allen, E. L., Smith, G. M., Fan, K. C., Hong, J. S., French, S. W., McCaffery, J. M., Lightowlers, R. N., Morse, H. C., 3rd, Koehler, C. M., & Teitell, M. A. (2010). PNPASE regulates RNA import into mitochondria. *Cell*, *142*(3), 456-467. https://doi.org/10.1016/j.cell.2010.06.035

Wei, T., Najmi, S. M., Liu, H., Peltonen, K., Kucerova, A., Schneider, D. A., & Laiho, M. (2018). Small-molecule targeting of RNA polymerase I activates a conserved transcription elongation checkpoint. *Cell reports*, *23*(2), 404-414.

Wilson, R. C., & Doudna, J. A. (2013). Molecular mechanisms of RNA interference. *Annu Rev Biophys*, *42*, 217-239.

Wong, L. H., Brettingham-Moore, K. H., Chan, L., Quach, J. M., Anderson, M. A., Northrop, E. L., Hannan, R., Saffery, R., Shaw, M. L., Williams, E., & Choo, K. H. A. (2024). Corrigendum: Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. *Genome Res*, *34*(6), 979-980. https://doi.org/10.1101/gr.279693.124

Zahler, N. H., Christian, E. L., & Harris, M. E. (2003). Recognition of the 5′ leader of pre-tRNA substrates by the active site of ribonuclease P. *Rna*, *9*(6), 734-745.

Zeng, W., Vu, G.-P., Bai, Y., Chen, Y.-C., Trang, P., Lu, S., Xiao, G., & Liu, F. (2013). RNase P-associated external guide sequence effectively reduces the expression of human CC-chemokine receptor 5 and inhibits the infection of human immunodeficiency virus 1. *Biomed Res Int*, *2013*.

Zhou, H., Wang, Y., Lv, Q., Zhang, J., Wang, Q., Gao, F., Hou, H., Zhang, H., Zhang, W., & Li, L. (2016). Overexpression of ribosomal RNA in the development of human cervical cancer is associated with rDNA promoter hypomethylation. *PLoS One*, *11*(10), e0163340.

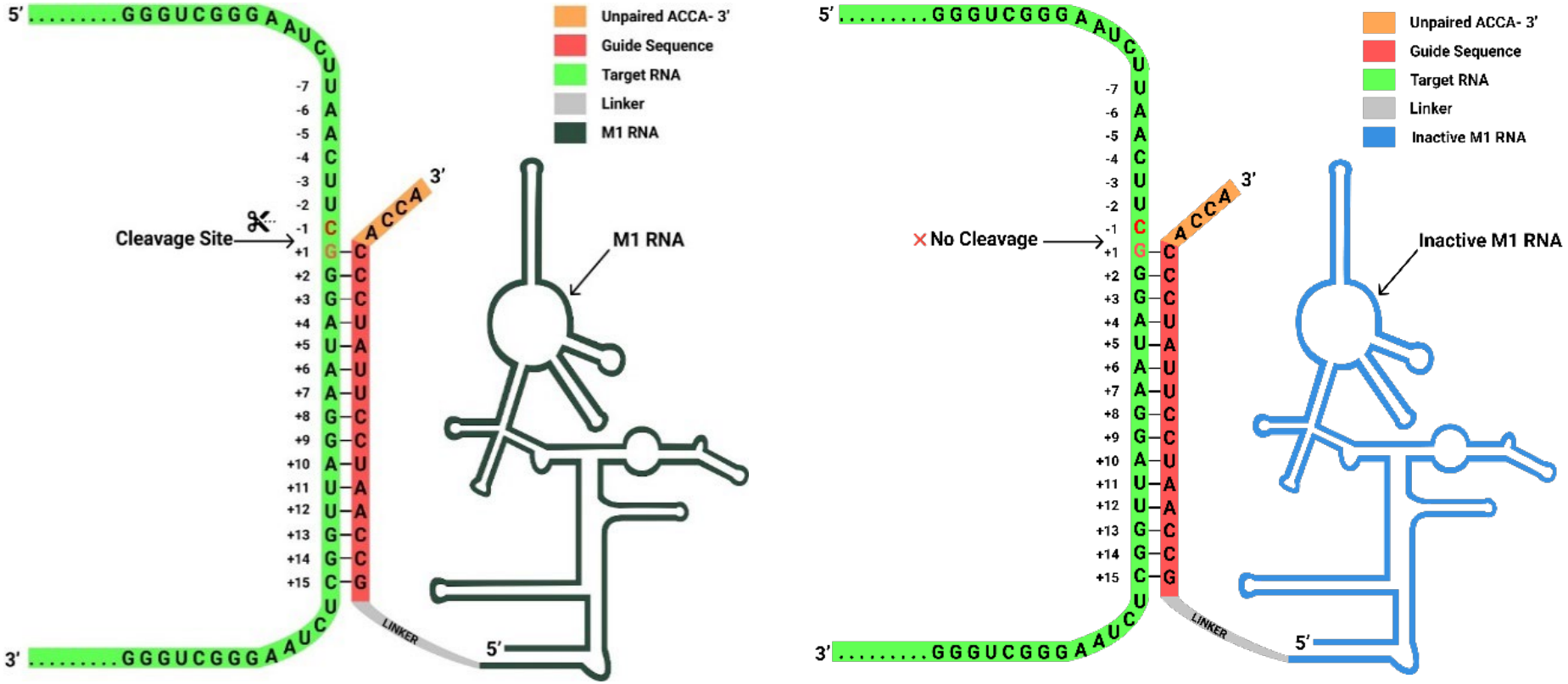


Figure 1: A schematic representation of the M1GS strategy. The cartoon shows how the engineered RNase P ribozyme M1GS cleaves the target RNA by hybridizing to it through a complementary guide sequence (GS). A random RNA as a target RNA and the corresponding complementary guide sequence (GS) are shown in the diagram. The diagram on the left represents M1GS with an active M1 RNA and the one on the right represents M1GS with an inactive, mutant M1 RNA.

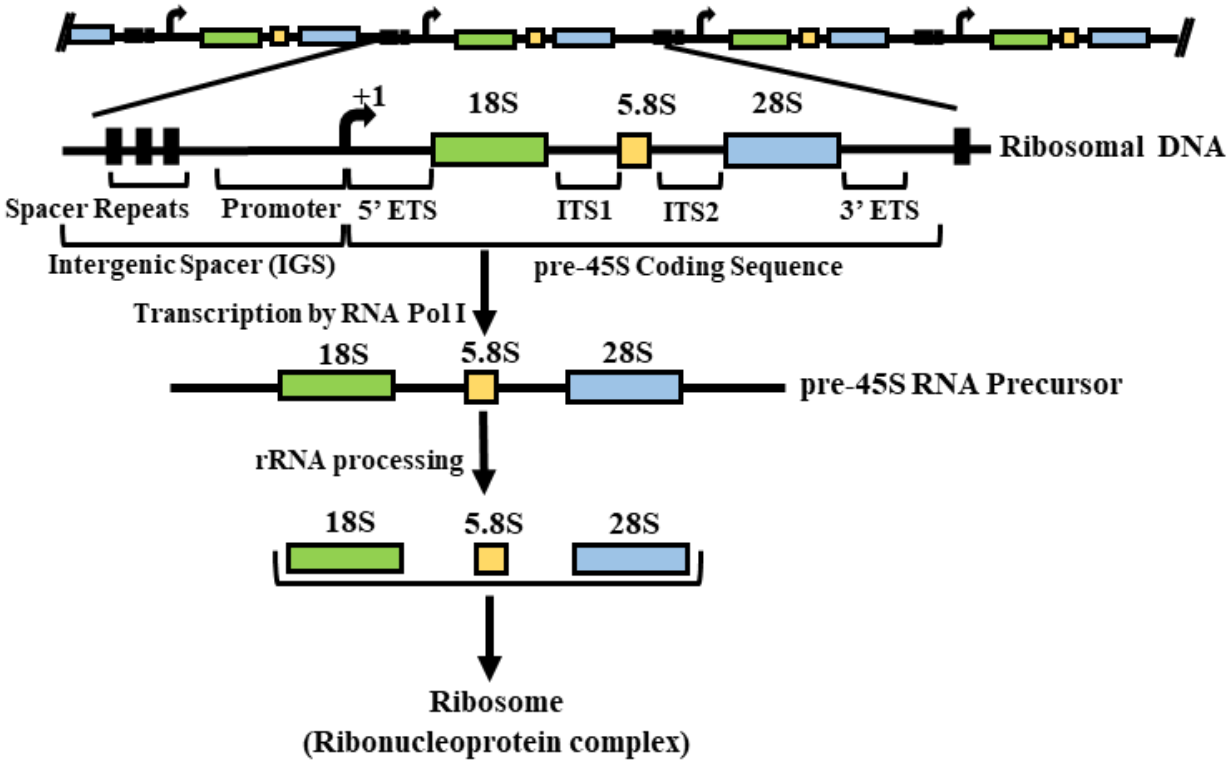


Figure 2: Organization of human 45S rRNA genes, their transcription, and ribosome assembly. The diagram shows different human 45S rRNA gene organization, components of a typical human 45S rRNA gene including the structural rRNA which become components of functional ribosomes.

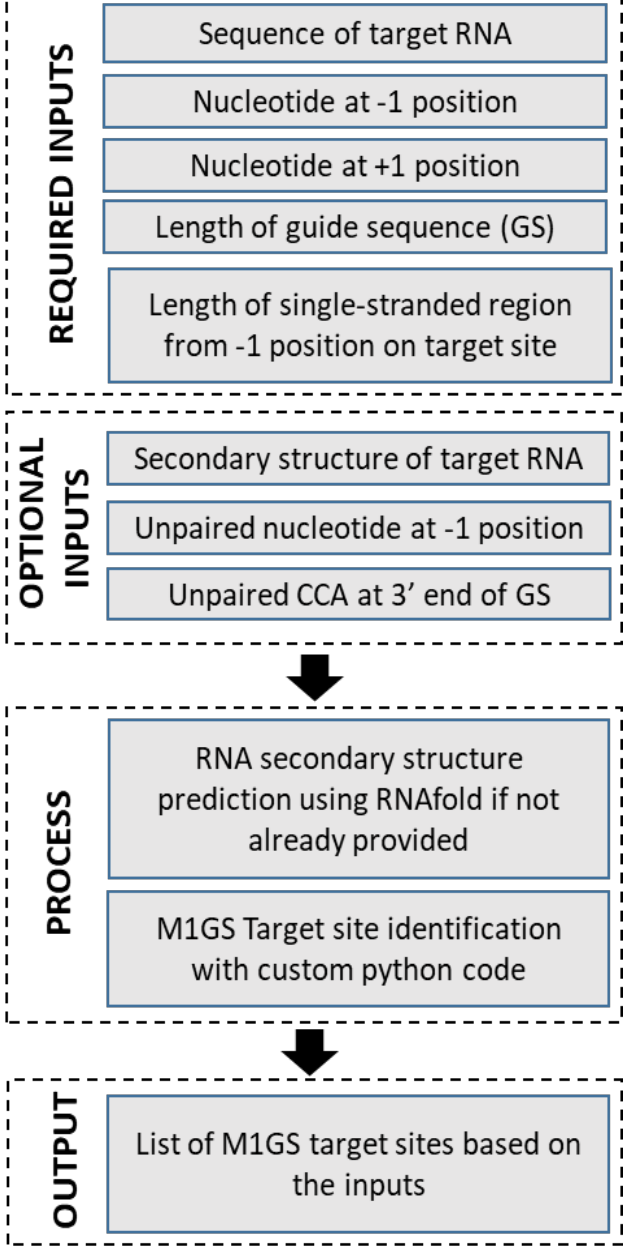


Figure 3: The flowchart depicts the operational modality of the M1GS Target site finder tool.

A Screenshot depicting the inputs and outputs with unpaired CCA-3' and nucleotide -1

Input

Sequences of target DNA/RNA

CGCGACCUCAGAUCAGACGUGGCGACCCGCUGAAUUUAAGCAUAUUAGUCAGCGGAGGAAAAGAAACUAACCAGGAUUCCCUCAGUAACGGCGAGU
GAACAGGGAAGAGCCCAGCGCCGAAUCCCCGCCCCGCGGCGGGGCGCGGGACAUGUGGCGUACGGAAGACCCGCUCCCCGGCGCCGCUCGUGGG
GGGCCCAAGUCCUUCUGAUCGAGGCCCAGCCCGUGGACGGUGUGAGGCCGGUAGCGGCCCCCGGCGCGCCGGGCCCGGGUCUUCCCGGAGUCG

Secondary sequence of target RNA

Enter sequence here

Nucleotide at -1 Position: C U A G

Nucleotide at +1 Position: G A

Unpaired CCA-3'

Unpaired Nucleotide at -1 according to secondary structure

Desirable length of target RNA from -1 position that is predicted to exist in single-stranded form as per the RNAFold.: 7

Length of Guide Sequence: 15

Results:

Target sites with C at -1

Position of -1 site	Target Site Sequence (from +1)	GC Content
1359	5'-GCACGAAAGCCGCCG-3'	73.33333333333333
2862	5'-GGGAUAAGGAUUGGC-3'	53.333333333333336
4675	5'-GCCUCUAAGUCAGAA-3'	46.666666666666664

Target sites with U at -1

Position of -1 site	Target Site Sequence (from +1)	GC Content
1974	5'-GGAAGUCGGAAUCCG-3'	60.0

B Screenshot depicting the inputs and outputs without unpaired CCA-3' and nucleotide -1

Input

Sequences of target DNA/RNA

CGCGACCUCAGAUCAGACGUGGCGACCCGCUGAAUUUAAGCAUAUUAGUCAGCGGAGGAAAAGAAACUAACCAGGAUUCCCUCAGUAACGGCGAGU
GAACAGGGAAGAGCCCAGCGCCGAAUCCCCGCCCCGCGGCGGGGCGCGGGACAUGUGGCGUACGGAAGACCCGCUCCCCGGCGCCGCUCGUGGG
GGGCCCAAGUCCUUCUGAUCGAGGCCCAGCCCGUGGACGGUGUGAGGCCGGUAGCGGCCCCCGGCGCGCCGGGCCCGGGUCUUCCCGGAGUCG

Secondary sequence of target RNA

Enter sequence here

Nucleotide at -1 Position: C U A G

Nucleotide at +1 Position: G A

Unpaired CCA-3'

Unpaired Nucleotide at -1 according to secondary structure

Desirable length of target RNA from -1 position that is predicted to exist in single-stranded form as per the RNAFold.: 7

Length of Guide Sequence: 15

Results:

Target sites with C at -1

Position of -1 site	Target Site Sequence (from +1)	GC Content
1359	5'-GCACGAAAGCCGCCG-3'	73.33333333333333
2691	5'-GGUUCCGGCGGCGUC-3'	80.0
2854	5'-GUAACUUCGGGAUAA-3'	40.0
2862	5'-GGGAUAAGGAUUGGC-3'	53.333333333333336
3840	5'-GAAACCACAGCCAAG-3'	53.333333333333336
4675	5'-GCCUCUAAGUCAGAA-3'	46.666666666666664
4943	5'-GCUAAACCAUUCGUA-3'	40.0

Target sites with U at -1

Position of -1 site	Target Site Sequence (from +1)	GC Content
1314	5'-GAAACACGGACCAAG-3'	53.333333333333336
1974	5'-GGAAGUCGGAAUCCG-3'	60.0
1998	5'-GUGUAACAACUCACC-3'	46.666666666666664
3939	5'-GUAGAAUAAGUGGGA-3'	40.0
4272	5'-GAUCUUGAUUUUCAG-3'	33.33333333333333
4421	5'-GAUCCUUCGAUGUCG-3'	53.333333333333336
4671	5'-GAACGCCUCUAAGUC-3'	53.333333333333336

Figure 4: Input and output in the M1GS target site finder tool. Screenshots from M1GS target site finder tool show input options can be chosen and the output obtained for 28S rRNA as the target. (A) Shows results for more stringent input options. (B) Shows results for less stringent input options.

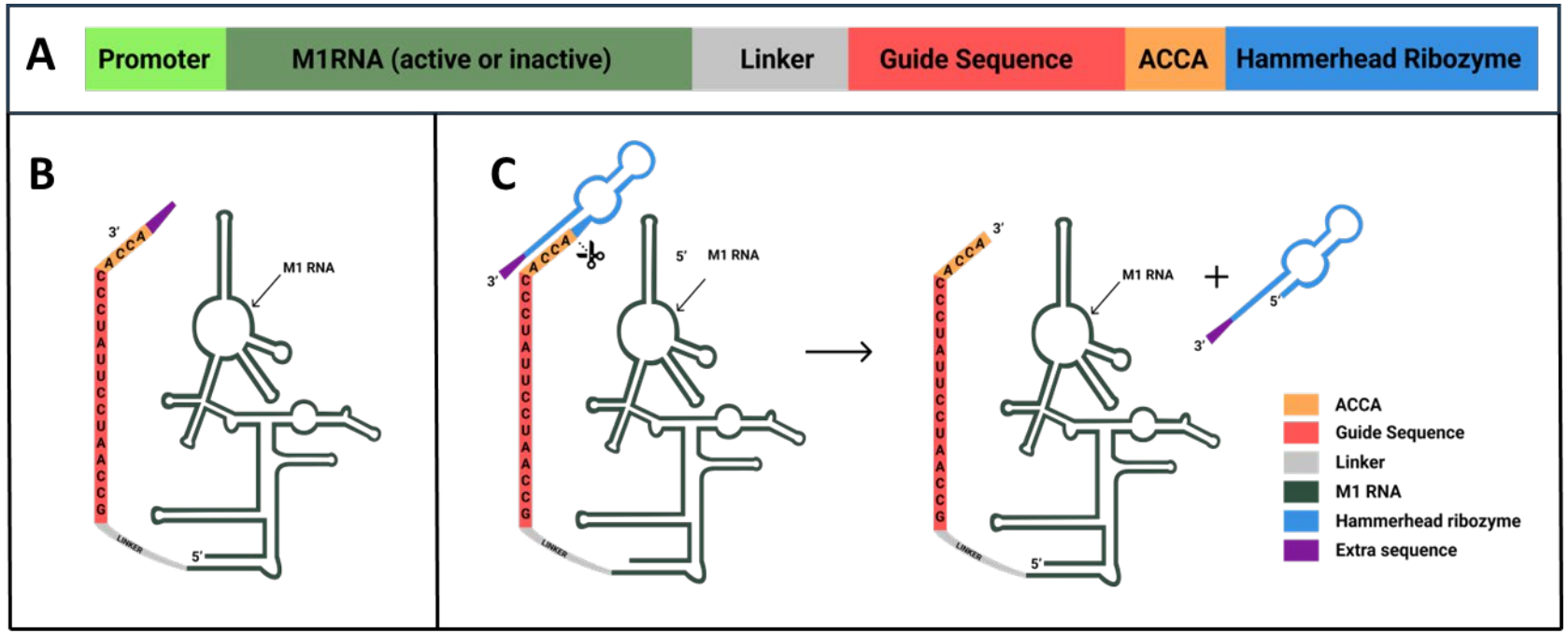


Figure 5: **Structural components of M1GS construct and M1GS cleavage strategy employed. (**A) The cartoon depicts the different components of the M1GS constructs used in this study. (B) M1GS cleavage strategy with a 3' overhang. **C.** M1GS cleavage strategy for generating M1GS constructs without 3' overhang by introducing a self-cleaving hammerhead ribozyme at the 3' end.

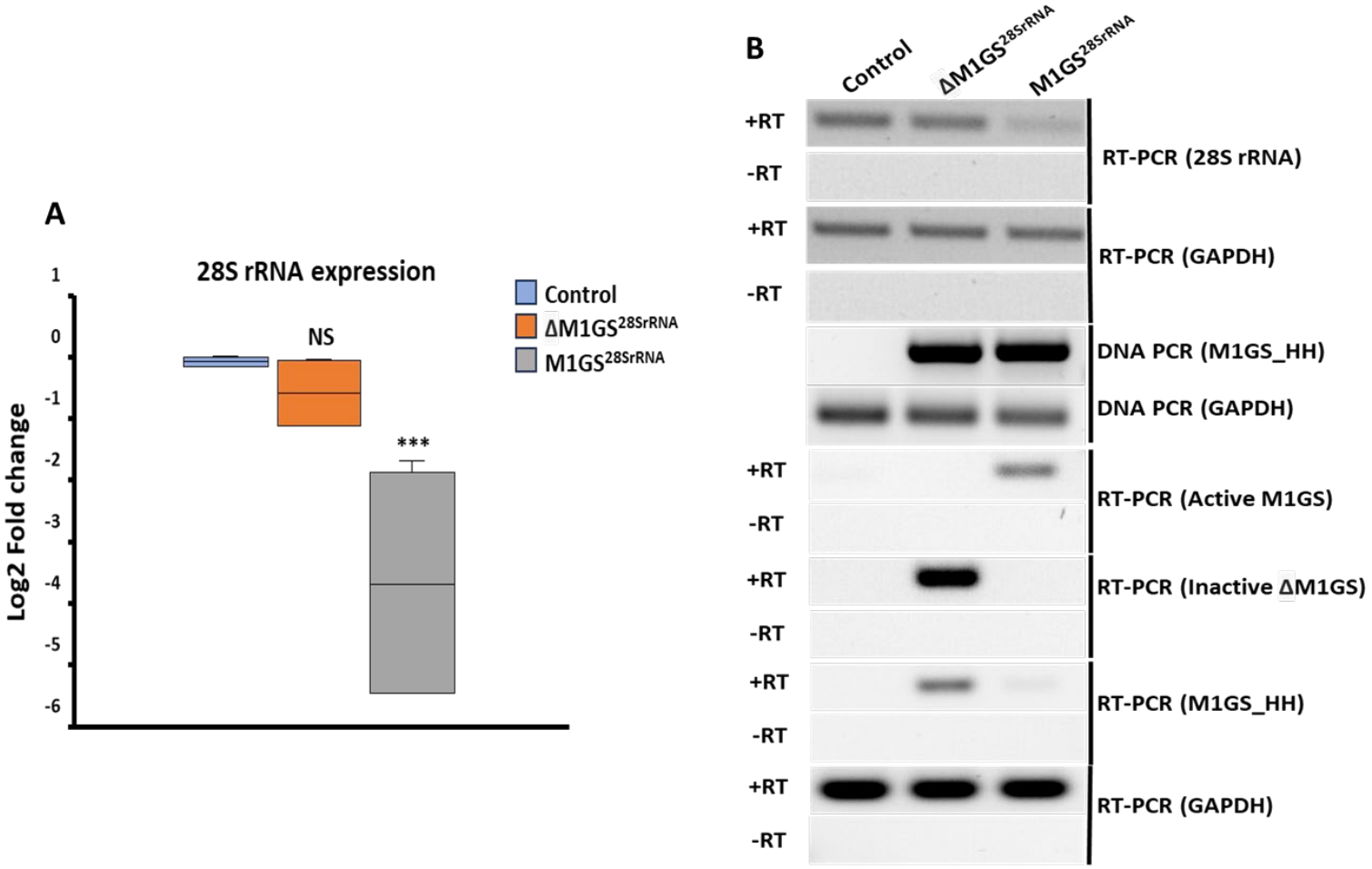


Figure 6: M1GS-mediated downregulation of 28S rRNA in human lung cancer cell line. (A) The box plot analysis of qPCR data reveals relative 28S rRNA levels in human A549 cell lines expressing different M1GS constructs and no-transfection controls. (B) Top to bottom: Gel images show semiquantitative RT-PCR analysis of 28S rRNA, PCR analysis of M1GS constructs, RT-PCR analysis of M1GS, GAPDH, and the HH ribozyme. For all RT-PCR analyses involving M1GS, primers specific to active and inactive M1GS were used.

Fold algorithms and basic options:

- minimum free energy (MFE) and partition function
- avoid isolated base pairs

Advanced options:

- Dangling end options: dangling energies on both sides of a helix in any case
- Energy Parameters: RNA parameters (Turner model, 2004)
- SHAPE reactivity data: None
- rescale energy parameters to a given temperature (C) - 37

Figure S1: Default settings of RNAFold used for secondary structure prediction.

Figure S2: Human 28S rRNA sequence

Target RNA: 28S ribosomal RNA gene

Target RNA used in the paper: 5'-GGGAUAAGGAUUGGC-3'

>KY962518.1:7921-12971 Homo sapiens 28S ribosomal RNA gene

CGCGACCTCAGATCAGACGTGGCGACCCGCTGAATTTAAGCATATTAGTCAGCGGAGGAAAAGAAACTAACCAGGATTCCCTCAGT
AACGGCGAGTGAACAGGGAAGAGCCCAGCGCCGAATCCCCGCCCCGCGGCGGGGCGCGGGACATGTGGCGTACGGAAGACCCGCTC
CCCGGCGCCGCTCGTGGGGGGCCCAAGTCCTTCTGATCGAGGCCCAGCCCGTGGACGGTGTGAGGCCGGTAGCGGCCCCCGGCGCG
CCGGGCCCGGGTCTTCCCGGAGTCGGGTTGCTTGGGAATGCAGCCCAAAGCGGGTGGTAAACTCCATCTAAGGCTAAATACCGGCA
CGAGACCGATAGTCAACAAGTACCGTAAGGGAAAGTTGAAAAGAACTTTGAAGAGAGAGTTCAAGAGGGCGTGAAACCGTTAAGAG
TAAACGGGTGGGGTCCGCGCAGTCCGCCCGGAGGATTCAACCCGGCGGCGGGTCCGGCCGTGTCGGCGGCCCGGCGGATCTTTCCC
GCCCCCCGTTCCTCCCGACCCCTCCACCCGCCCTCCCTTCCCCCGCCGCCCCTCCTCCTCCTCCCCGGAGGGGGCGGGCTCCGGCG
GGTGCGGGGGTGGGCGGGCGGGGCCGGGGGTGGGGTCGGCGGGGGACCGTCCCCCGACCGGCGACCGGCCGCCGCCGGGCGCATTT
CCACCGCGGCGGTGCGCCGCGACCGGCTCCGGGACGGCTGGGAAGGCCCGGCGGGGAAGGTGGCTCGGGGGGCCCCGTCCGTCCGT
CCGTCCGTCCTCCTCCTCCCCCGTCTCCGCCCCCCGGCCCCCGCGTCCTCCCTCGGGAGGGCGCGCGGGTCGGGGCGGCGGCGGCGG
TGGCGGCGGCGGCGGCGGCGGCGGGACCGAAACCCCCCCGAGTGTTACAGCCCCCCCGGCAGCAGCACTCGCCGAATCCCGGGGCG
AGGGAGCGAGACCCGTCGCCGCGCTCTCCCCCCTCCCGGCGCCCACCCCCGCGGGGAATCCCCCGCGAGGGGGGTCTCCCCCGCGG
GGGCGCGCCGGCGTCTCCTCGTGGGGGGGCCGGGCCACCCCTCCCACGGCGCGACCGCTCTCCCACCCCTCCTCCCCGCGCCCCCG
CCCCGGCGACGGGGGGGGTGCCGCGCGCGGGTCGGGGGGCGGGGCGGACTGTCCCCAGTGCGCCCCGGGCGGGTCGCGCCGTCGGG
CCCGGGGGGGAGGTTCTCTCGGGGCCACGCGCGCGTCCCCCGAAGAGGGGGACGGCGGAGCGAGCGCACGGGGTCGGCGGCGACGTC
GGCTACCCACCCGACCCGTCTTGAAACACGGACCAAGGAGTCTAACACGTGCGCGAGTCGGGGGCTCGCACGAAAGCCGCCGTGGC
GCAATGAAGGTGAAGGCCGGCGCGCTCGCCGGCCGAGGTGGGATCCCGAGGCCTCTCCAGTCCGCCGAGGGCGCACCACCGGCCCG
TCTCGCCCGCCGCGCCGGGGAGGTGGAGCACGAGCGCACGTGTTAGGACCCGAAAGATGGTGAACTATGCCTGGGCAGGGCGAAGC
CAGAGGAAACTCTGGTGGAGGTCCGTAGCGGTCCTGACGTGCAAATCGGTCGTCCGACCTGGGTATAGGGGCGAAAGACTAATCGA
ACCATCTAGTAGCTGGTTCCCTCCGAAGTTTCCCTCAGGATAGCTGGCGCTCTCGCAGACCCGACGCACCCCCGCCACGCAGTTTT
ATCCGGTAAAGCGAATGATTAGAGGTCTTGGGGCCGAAACGATCTCAACCTATTCTCAAACTTTAAATGGGTAAGAAGCCCGGCTC
GCTGGCGTGGAGCCGGGCGTGGAATGCGAGTGCCTAGTGGGCCACTTTTGGTAAGCAGAACTGGCGCTGCGGGATGAACCGAACGC
CGGGTTAAGGCGCCCGATGCCGACGCTCATCAGACCCCAGAAAAGGTGTTGGTTGATATAGACAGCAGGACGGTGGCCATGGAAGT
CGGAATCCGCTAAGGAGTGTGTAACAACTCACCTGCCGAATCAACTAGCCCTGAAAATGGATGGCGCTGGAGCGTCGGGCCCATAC
CCGGCCGTCGCCGGCAGTCGAGAGTGGACGGGAGCGGCGGGGGCGGCGCGCGCGCGCGCGTGTGGTGTGCGTCGGAGGGCGGCG
GCGGCGGCGGGGGTGTGTGGGGTCCTCCCCCGCCCCCCCCCACGCCTCCTCCCCTCCTCCCGCCCACGCCCCGCTCCCCGCCCCCG
GAGCCCCGCGGACGCTACGCCGCGACGAGTAGGAGGGCCGCTGCGGTGAGCCTTGAAGCCTAGGGCGCGGGCCCGGGTGGAGCCGC
CGCAGGTGCAGATCTTGGTGGTAGTAGCAAATATTCAAACGAGAACTTTGAAGGCCGAAGTGGAGAAGGGTTCCATGTGAACAGCA
GTTGAACATGGGTCAGTCGGTCCTGAGAGATGGGCGAGCGCCGTTCTGAAGGGACGGGCGATGGCCTCCGTTGCCCTCGGCCGATC
GAAAGGGAGTCGGGTTCAGATCCCCGAATCCGGAGTGGCGGAGATGGGCGCCGCGAGGCGTCCAGTGCGGTAACGCGACCGATCCC
GGAGAAGCCGGCGGGAGCCCCGGGGAGAGTTCTCTTTTCTTTGTGAAGGGCAGGGCGCCCTGGAATGGGTTCGCCCCGAGAGAGGG
GCCCGTGCCTTGGAAAGCGTCGCGGTTCCGGCGGCGTCCGGTGAGCTCTCGCTGGCCCTTGAAAATCCGGGGGAGAGGGTGTAAAT
CTCGCGCCGGGCCGTACCCATATCCGCAGCAGGTCTCCAAGGTGAACAGCCTCTGGCATGTTGGAACAATGTAGGTAAGGGAAGTC
GGCAAGCCGGATCCGTAACTTCGGGATAAGGATTGGCTCTAAGGGCTGGGTCGGTCGGGCTGGGGCGCGAAGCGGGGCTGGGCGCG
CGCCGCGGCTGGACGAGGCGCCGCCGCCCCCCCCACGCCCGGGGCACCCCCCTCGCGGCCCTCCCCCGCCCCACCCCGCGCGCGCC
GCTCGCTCCCTCCCCACCCCGCGCCCTCTCTCTCTCTCTCCCCCGCTCCCCGTCCTCCCCCCTCCCCGGGGGAGCGCCGCGTGG
GGGCGGCGGCGGGGGGAGAAGGGTCGGGGCGGCAGGGGCCGGCGGCGGCCGCCGCGGGGCCCCGGCGGCGGGGGCACGGTCCCCCG
CGAGGGGGGCCCGGGCACCCGGGGGGCCGGCGGCGGCGGCGACTCTGGACGCGAGCCGGGCCCTTCCCGTGGATCGCCCCAGCTGC
GGCGGGCGTCGCGGCCGCCCCCGGGGAGCCCGGCGGGCGCGGCGCGCCCCCCCCCACCCCACGTCTCGTCGCGCGCGCGTCCGC
TGGGGGCGGGGAGCGGTCGGGCGGCGGCGGTCGGCGGGCGGCGGGGCGGGGCGGTTCGTCCCCCCGCCCTACCCCCCCGGCCCCGT
CCGCCCCCCGTTCCCCCCTCCTCCTCGGCGCGCGGCGGCGGCGGCGGGCGGCGGAGGGGCCGCGGGCCGGTCCCCCCCGCCGGGTC
CGCCCCCGGGGCCGCGGTTCCGCGCGGCGCCTCGCCTCGGCCGGCGCCTAGCAGCCGACTTAGAACTGGTGCGGACCAGGGGAATC
CGACTGTTTAATTAAAACAAAGCATCGCGAAGGCCCGCGGCGGGTGTTGACGCGATGTGATTTCTGCCCAGTGCTCTGAATGTCAA
AGTGAAGAAATTCAATGAAGCGCGGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAG
TGACGCGCATGAATGGATGAACGAGATTCCCACTGTCCCTACCTACTATCCAGCGAAACCACAGCCAAGGGAACGGGCTTGGCGGA
ATCAGCGGGGAAAGAAGACCCTGTTGAGCTTGACTCTAGTCTGGCACGGTGAAGAGACATGAGAGGTGTAGAATAAGTGGGAGGCC
CCCGGCGCCCCCCCGGTGTCCCCGCGAGGGGCCCGGGGCGGGGTCCGCCGGCCCTGCGGGCCGCCGGTGAAATACCACTACTCTGA
TCGTTTTTTCACTGACCCGGTGAGGCGGGGGGGCGAGCCCCGAGGGGCTCTCGCTTCTGGCGCCAAGCGCCCGGCCGCGCGCCGGC
CGGGCGCGACCCGCTCCGGGGACAGTGCCAGGTGGGGAGTTTGACTGGGGCGGTACACCTGTCAAACGGTAACGCAGGTGTCCTAA
GGCGAGCTCAGGGAGGACAGAAACCTCCCGTGGAGCAGAAGGGCAAAAGCTCGCTTGATCTTGATTTTCAGTACGAATACAGACCG
TGAAAGCGGGGCCTCACGATCCTTCTGACCTTTTGGGTTTTAAGCAGGAGGTGTCAGAAAAGTTACCACAGGGATAACTGGCTTGT
GGCGGCCAAGCGTTCATAGCGACGTCGCTTTTTGATCCTTCGATGTCGGCTCTTCCTATCATTGTGAAGCAGAATTCACCAAGCGT
TGGATTGTTCACCCACTAATAGGGAACGTGAGCTGGGTTTAGACCGTCGTGAGACAGGTTAGTTTTACCCTACTGATGATGTGTTG
TTGCCATGGTAATCCTGCTCAGTACGAGAGGAACCGCAGGTTCAGACATTTGGTGTATGTGCTTGGCTGAGGAGCCAATGGGGCGA
AGCTACCATCTGTGGGATTATGACTGAACGCCTCTAAGTCAGAATCCCGCCCAGGCGGAACGATACGGCAGCGCCGCGGAGCCTCG
GTTGGCCTCGGATAGCCGGTCCCCCGCCTGTCCCCGCCGGCGGGCCGCCCCCCCCTCCACGCGCCCCGCGCGCGCGGGAGGGCGCG

Primer	*Forward (5'-3')*	*Reverse (5'-3')*
28S rRNA	ATGTAGGTAAGGGAAGTCGGCAAGCC	ACCGACCCAGCCCTTAGAGCCAATC
M1GS_HH	TAACAGTCCGTGGCACGGTAAACTCC	CAGGACATCGGGTTTCAGATGTCGAC
Active M1GS	TAACAGTCCGTGGCACGGTAAACTCC	GACCGATAAGCCGGGTTCTGT
Inactive M1GS	TAACAGTCCGTGGCACGGTAAACTCC	GTGAAACTGACCGAGATACATCTCGCTAG
GAPDH	CTCCTGCACCACCAACTGCT	GGGCCATCCACAGTCTTCTG

Figure S3: List of primers used in the study

Figure S4: Sequences of different components of M1GS constructs. (A) H1TetO2 promoter (B) Linker sequence (C) Hammerhead ribozyme (D) The alignment of the active vs inactive M1 RNA sequences. The arrows depict the reverse primer binding sites used to differentiate the active from the inactive M1 RNA (E) An agarose gel image shows PCR amplification of the active and the inactive M1 RNA.

A H1TetO2 promoter (human H1 RNA)

```
GAATTCAATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAATCCCTATCAGTGA
TAGAGACTTATAAGTTCCCTATCAGTGATAGAGATCC
```

B Linker sequence

```
AAGCTTGGCGTAATCATGGTCATA- 24 bp
```

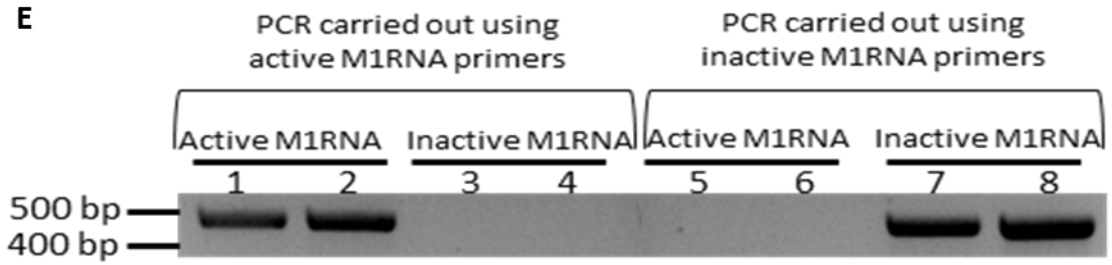
C Hammerhead ribozyme sequence

```
GTCGACATCTGAAACCCGATGTCCTGATGAGCTCAAGGTAGAGCGAAACTGGTCTCGAG
```

D Sequence alignment of active and inactive M1 RNA

```
  Active   1    GAAGCTGACCAGACAGTCGCCGCTTCGTCGTCGTCCTCTTCGGGGGAGACGGGCGGAGGG  60
                ||||||||||||||||||||||||||| |||||||||||||||| |||||||| |||||||
Inactive   1    GAAGCTGACCAGACAGTCGCCGCTTCATCGTCGTCCTCTTCGGAGGAGACGGACGGAGGG  60

  Active   61   GAGGAAAGTCCGGGCTCCATAGGGCAGGGTGCCAGGTAACGCCTGGGGGGGAAACCCACG  120
                ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Inactive   61   GAGGAAAGTCCGGGCTTCATAGGGCAGGGTGCCAGGTAACGCCTGGGGGGGAAACCCACG  120

  Active   121  ACCAGTGCAACAGAGAGCAAACCGCCGATGGCCCGCGCAAGCGGGATCAGGTAAGGGTGA  180
                ||||||||||||||||||||||||||||||||||||||||||||||||| || ||||||||
Inactive   121  ACCAGTGCAACAGAGAGCAAACCGCCGATGGCCCGCGCAAGCGGGATCGGGCAAGGGTGA  180

  Active   181  AAGGGTGCGGTAAGAGCGCACCGCGCGGCTGGTAACAGTCCGTGGCACGGTAAACTCCAC  240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Inactive   181  AAGGGTGCGGTAAGAGCGCACCGCGCGGCTGGTAACAGTCCGTGGCACGGTAAACTCCAC  240

  Active   241  CCGGAGCAAGGCCAAATAGGGGTTCATAAGGTACGGCCCGTACTGAACCCGGGTAGGCTG  300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Inactive   241  CCGGAGCAAGGCCAAATAGGGGTTCATAAGGTACGGCCCGTACTGAACCCGGGTAGGCTG  300
                                                                  <.......................
  Active   301  CTTGAGCCAGTGAGCGATTGCTGGCCTAGATGAATGACTGTCCACGACAGAACCCGGCTT  360
                |||| |||||||||||||||||||||||||||||||||||||||||||  ||      |  |
Inactive   301  CTTGGGCCAGTGAGCGATTGCTGGCCTAGATGAATGACTGTCCACGCTAGCGAGATGTAT  360
                                                                  <------------
                ..............
  Active   361  ATCGGTCAGTTTCACCT  377
                |||||||||||||||||
Inactive   361  CTCGGTCAGTTTCACCT  377
                ----------------
```

E



A Sequence of thymidine kinase (TK) of human herpes simplex virus 1 (HSV-1)- Liu & Altman,1996

Target RNA : 5'-GUCUGCGUUCGACCAGGCU-3'

>OR232881.1 Human alphaherpesvirus 1 isolate SP1.UL23.14 thymidine kinase gene, complete cds

ATGGCTTCGTACCCCTGCCATCAACACGCGTCTGCGTTCGACCAGGCTGCGCGTTCTCGCGGCCATAGCAACCGACGTACGG
CGTTGCGCCCTCGCCGGCAGCAAGAAGCCACGGAAGTCCGCCTGGAGCAGAAAATGCCCACGCTACTGCGGGTTTATATAGA
CGGTCCTCACGGGATGGGGAAAACCACCACCACGCAACTGCTGGTGGCCCTGGGTTCGCGCGACGATATCGTCTACGTACCC
GAGCCGATGACTTACTGGCAGGTGCTGGGGGCTTCCGAGACAATCGCGAACATCTACACCACACAACACCGCCTCGACCAGG
GTGAGATATCGGCCGGGGACGCGGCGGTGGTAATGACAAGCGCCCAGATAACAATGGGCATGCCTTATGCCGTGACCGACGC
CGTTCTGGCTCCTCATATCGGGGGGGAGGCTGGGAGCTCACATGCCCCGCCCCCGGCCCTCACCCTCATCTTCGACCGCCAT
CCCATCGCCGCCCTCCTGTGCTACCCGGCCGCGCGATACCTTATGGGCAGCATGACCCCCCAGGCCGTGCTGGCGTTCGTGG
CCCTCATCCCGCCGACCTTGCCCGGCACAAACATCGTGTTGGGGGCCCTTCCGGAGGACAGACACATCGACCGCCTGGCCAA
ACGCCAGCGCCCCGGCGAGCGGCTTGACCTGGCTATGCTGGCCGCGATTCGCCGCGTTTACGGGCTGCTTGCCAATACGGTG
CGGTATCTGCAGGGCGGCGGGTCGTGGCGGGAGGATTGGGGACAGCTTTCGGGGACGGCCGTGCCGCCCCAGGGTGCCGAGC
CCCAGACCAACGCGGGCCCACGACCCCATATCGGGGACACGTTATTTACCCTGTTTCGGGCCCCCGAGTTGCTGGCCCCCAA
CGGCGACCTGTATAACGTGTTTGCCTGGGCCTTGGACGTCTTGGCCAAACGCCTCCGTCCCATGCACGTCTTTATCCTGGAT
TACGACCAATCGCCCGCCGGCTGCCGGGACGCCCTGCTGCAACTTACCTCCGGGATGGTCCAGACCCACGTCACCACCCCAG
GCTCCATACCGACGATCTGCGACCTGGCGCGCACGTTTGCCCGGGAGATGGGGGAGGCTAACTGA

Figure S5: M1GS target site prediction for thymidine kinase of HSV-1. (A) Sequence of thymidine kinase (TK) and the GS used to target it. The highlighted sequence in yellow indicate the target sequence used to design GS.
(B) A screenshot shows M1GS target sites predicted for the TK RNA. The target site enclosed by a red rectangle in the diagram was the GS used in that study (Liu & Altman, 1996).

B Screenshot depicting the list of predicted M1GS target sites (output)

Results:

Target sites with C at -1

Position of -1 site	Target Site Sequence (from +1)	GC Content
8	5'-GUACCCCUGCCAUCAACAC-3'	57.89473684
27	5'-GCGUCUGCGUUCGACCAGG-3'	68.42105263
29	5'-GUCUGCGUUCGACCAGGCU-3'	63.15789474
35	5'-GUUCGACCAGGCUGCGCGU-3'	68.42105263
39	5'-GACCAGGCUGCGCGUUCUC-3'	68.42105263
50	5'-GCGUUCUCGCGGCCAUAGC-3'	68.42105263
52	5'-GUUCUCGCGGCCAUAGCAA-3'	57.89473684
58	5'-GCGGCCAUAGCAACCGACG-3'	68.42105263
60	5'-GGCCAUAGCAACCGACGUA-3'	57.89473684
73	5'-GACGUACGGCGUUGCGCCC-3'	73.68421053
76	5'-GUACGGCGUUGCGCCCUCG-3'	73.68421053
80	5'-GGCGUUGCGCCCUCGCCGG-3'	84.21052632
83	5'-GUUGCGCCCUCGCCGGCAG-3'	78.94736842
88	5'-GCCCUCGCCGGCAGCAAGA-3'	73.68421053
94	5'-GCCGGCAGCAAGAAGCCAC-3'	68.42105263
97	5'-GGCAGCAAGAAGCCACGGA-3'	63.15789474
113	5'-GGAAGUCCGCCUGGAGCAG-3'	68.42105263
121	5'-GCCUGGAGCAGAAAAUGCC-3'	57.89473684
143	5'-GCUACUGCGGGUUUAUAUA-3'	42.10526316
151	5'-GGGUUUAUAUAGACGGUCC-3'	47.36842105
165	5'-GGUCCUCACGGGAUGGGGA-3'	68.42105263
174	5'-GGGAUGGGGAAAACCACCA-3'	57.89473684
197	5'-GCAACUGCUGGUGGCCCUG-3'	68.42105263
221	5'-GCGCGACGAUAUCGUCUAC-3'	57.89473684
223	5'-GCGACGAUAUCGUCUACGU-3'	52.63157895
225	5'-GACGAUAUCGUCUACGUAC-3'	47.36842105
228	5'-GAUAUCGUCUACGUACCCG-3'	52.63157895
234	5'-GUCUACGUACCCGAGCCGA-3'	63.15789474
240	5'-GUACCCGAGCCGAUGACUU-3'	57.89473684
246	5'-GAGCCGAUGACUUACUGGC-3'	57.89473684
251	5'-GAUGACUUACUGGCAGGUG-3'	52.63157895
282	5'-GAGACAAUCGCGAACAUCU-3'	47.36842105
291	5'-GCGAACAUCUACACCACAC-3'	52.63157895
293	5'-GAACAUCUACACCACACAA-3'	42.10526316
316	5'-GCCUCGACCAGGGUGAGAU-3'	63.15789474
321	5'-GACCAGGGUGAGAUAUCGG-3'	57.89473684
338	5'-GGCCGGGGACGCGGCGGUG-3'	89.47368421
342	5'-GGGGACGCGGCGGUGGUAA-3'	73.68421053
348	5'-GCGGCGGUGGUAAUGACAA-3'	57.89473684
350	5'-GGCGGUGGUAAUGACAAGC-3'	57.89473684
353	5'-GGUGGUAAUGACAAGCGCC-3'	57.89473684

Position of -1 site	Target Site Sequence (from +1)	GC Content
342	5'-GGGGACGCGGCGGUGGUAA-3'	73.68421053
348	5'-GCGGCGGUGGUAAUGACAA-3'	57.89473684
350	5'-GGCGGUGGUAAUGACAAGC-3'	57.89473684
353	5'-GGUGGUAAUGACAAGCGCC-3'	57.89473684
369	5'-GCCCAGAUAACAAUGGGCA-3'	52.63157895
399	5'-GUGACCGACGCCGUUCUGG-3'	68.42105263
405	5'-GACGCCGUUCUGGCUCCUC-3'	68.42105263
408	5'-GCCGUUCUGGCUCCUCAUA-3'	57.89473684
411	5'-GUUCUGGCUCCUCAUAUCG-3'	52.63157895
429	5'-GGGGGGGAGGCUGGGAGCU-3'	78.94736842
458	5'-GCCCCCGGCCCUCACCCUC-3'	84.21052632
464	5'-GGCCCUCACCCUCAUCUUC-3'	63.15789474
483	5'-GACCGCCAUCCCAUCGCCG-3'	73.68421053
487	5'-GCCAUCCCAUCGCCGCCCU-3'	73.68421053
498	5'-GCCGCCCUCCUGUGCUACC-3'	73.68421053
501	5'-GCCCUCCUGUGCUACCCGG-3'	73.68421053
518	5'-GGCCGCGCGAUACCUUAUG-3'	63.15789474
522	5'-GCGCGAUACCUUAUGGGCA-3'	57.89473684
524	5'-GCGAUACCUUAUGGGCAGC-3'	57.89473684
526	5'-GAUACCUUAUGGGCAGCAU-3'	47.36842105
558	5'-GUGCUGGCGUUCGUGGCCC-3'	73.68421053
566	5'-GUUCGUGGCCCUCAUCCCG-3'	68.42105263
570	5'-GUGGCCCUCAUCCCGCCGA-3'	73.68421053
584	5'-GCCGACCUUGCCCGGCACA-3'	73.68421053
587	5'-GACCUUGCCCGGCACAAAC-3'	63.15789474
597	5'-GGCACAAACAUCGUGUUGG-3'	52.63157895
609	5'-GUGUUGGGGGCCCUUCCGG-3'	73.68421053
626	5'-GGAGGACAGACACAUCGAC-3'	57.89473684
642	5'-GACCGCCUGGCCAAACGCC-3'	73.68421053
646	5'-GCCUGGCCAAACGCCAGCG-3'	73.68421053
658	5'-GCCAGCGCCCCGGCGAGCG-3'	89.47368421
664	5'-GCCCCGGCGAGCGGCUUGA-3'	78.94736842
669	5'-GGCGAGCGGCUUGACCUGG-3'	73.68421053
672	5'-GAGCGGCUUGACCUGGCUA-3'	63.15789474
676	5'-GGCUUGACCUGGCUAUGCU-3'	57.89473684
699	5'-GCGAUUCGCCGCGUUUACG-3'	63.15789474
701	5'-GAUUCGCCGCGUUUACGGG-3'	63.15789474
706	5'-GCCGCGUUUACGGGCUGCU-3'	68.42105263
709	5'-GCGUUUACGGGCUGCUUGC-3'	63.15789474
711	5'-GUUUACGGGCUGCUUGCCA-3'	57.89473684
717	5'-GGGCUGCUUGCCAAUACGG-3'	63.15789474

Position of -1 site	Target
734	5'-GGUGC
739	5'-GGUAU
753	5'-GGCGG
756	5'-GGGUC
761	5'-GUGGC
766	5'-GGGAG
788	5'-GGGGA
794	5'-GGCCG
798	5'-GUGCC
803	5'-GCCCC
816	5'-GAGCC
831	5'-GCGGG
833	5'-GGGCC
841	5'-GACCC
852	5'-GGGGA
860	5'-GUUAU
877	5'-GGGCC
885	5'-GAGUU
903	5'-GGCGA
906	5'-GACCU
918	5'-GUGUU
939	5'-GUCUU
952	5'-GCCUC
958	5'-GUCCC
969	5'-GUCUU
987	5'-GACCA
995	5'-GCCCG
999	5'-GCCGG
1002	5'-GGCUG
1009	5'-GGGAC
1014	5'-GCCCU
1035	5'-GGGAU
1053	5'-GUCAC
1076	5'-GACGA
1079	5'-GAUCU
1086	5'-GACCU
1094	5'-GCGCA
1096	5'-GCACG
1100	5'-GUUUG
1108	5'-GGGAG

A Sequence of 5' untranslated region (UTR) of hepatitis C virus (HCV) - Mao et al., 2014

Target RNA: 5'-GCAGAAAGCGUCU-3'

>AB828701.1 Hepatitis C virus subtype 1b genomic RNA, complete genome, strain: Pt1-98_Q2

(ONLY FIRST 341 NUCLEOTIDES)

GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCCTGTGAGGAACTACTGTCTTCACGCAGAAAGCGTCTAG
CCATGGCGTTAGTATGAGTGTCGTGCAGCCTCCAGGACCCCCCCTCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGT
ACACCGGAATTGCCAGGACGACCGGGTCCTTTCTTGGATCAACCCGCTCAATGCCTGGAGATTTGGGCGTGCCCCCGCGAGA
CTGCTAGCCGAGTAGTGTTGGGTCGCGAAAGGCCTTGTGGTACTGCCTGATAGGGTGCTTGCGAGTGCCCCGGGAGGTCTCG
TAGACCGTGCACC

Figure S6: M1GS target site prediction for hepatitis C subtype 1b genomic RNA.

(A) The hepatitis C subtype 1b genomic RNA sequence used in the study and the highlighted sequence in yellow indicate the target sequence used to design GS.

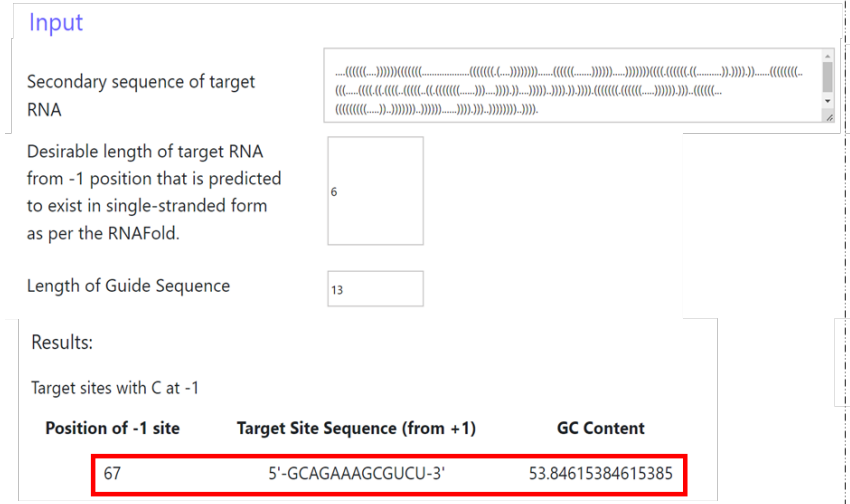
(B) A screenshot showing the output when secondary structure for the target RNA in the Vienna format as predicted by Mfold software was given as an input for the tool.

(C) A screenshot showing the output when secondary structure for the target RNA was predicted by RNAfold software for the same input as (A)

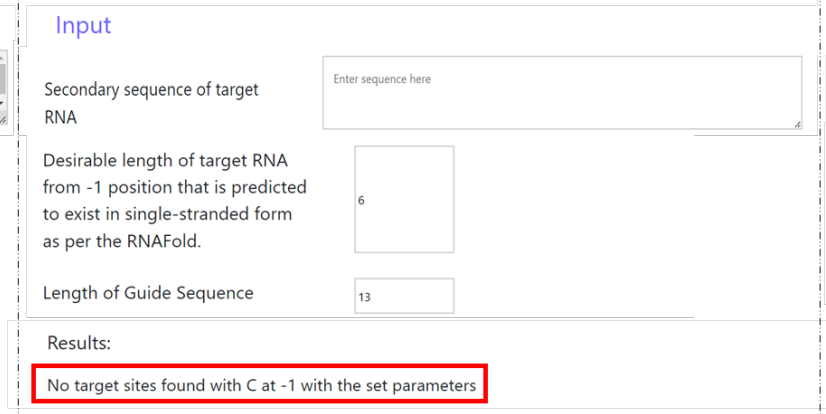
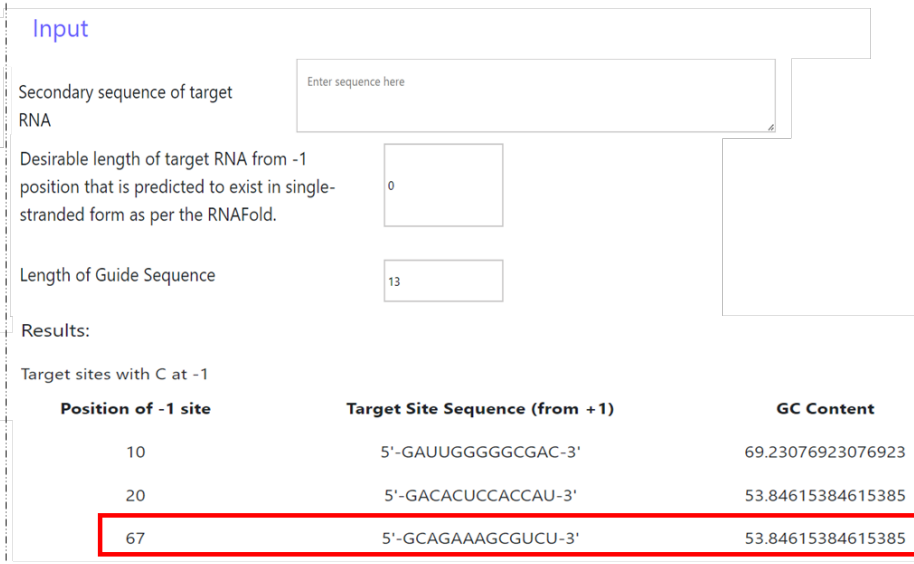
(D) A screenshot showing the output when secondary structure for the target RNA was predicted by RNAfold software when the input was changed.

In (B-D), red rectangles enclose the target site for cleaving the target RNA in the study (Mao et al. 2014)

B Secondary structure predicted by mfold



C Secondary structure predicted by RNAfold



D Secondary structure predicted by RNAfold