# Inducing gene expression by targeting promoter sequences using small activating RNAs

**Ji Wang**[#1], **Robert F. Place**[#1], **Victoria Portnoy**[1], **Vera Huang**[1], **Moo Rim Kang**[1], **Mika Kosaka**[1], **Maurice Kwok Chung Ho**[2,‡], and **Long-Cheng Li**[1,3,*]

[1] Department of Urology and Helen-Diller Comprehensive Cancer Center, University of California San Francisco, San Francisco, CA 94158, USA

[2] Biotechnology Research Institute and Division of Life Science, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

[3] Laboratory of Molecular Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[#] These authors contributed equally to this work.

## Abstract

Vector-based systems comprised of exogenous nucleic acid sequences remain the standard for ectopic expression of a particular gene. Such systems offer robust overexpression, but have inherent drawbacks such as the tedious process of construction, excluding sequences (e.g. introns and untranslated regions) important for gene function and potential insertional mutagenesis of host genome associated with the use of viral vectors. We and others have recently reported that short double-stranded RNAs (dsRNAs) can induce endogenous gene expression by targeting promoter sequences in a phenomenon referred to as RNA activation (RNAa) and such dsRNAs are termed small activating RNAs (saRNAs). To date, RNAa has been successfully utilized to induce the expression of different genes such as tumor suppressor genes. Here, we describe a detailed protocol for target selection and dsRNA design with associated experiments to facilitate RNAa in cultured cells. This technique may be applied to selectively activate endogenous gene expression for studying gene function, interrogating molecular pathways and reprogramming cell fate.

## Keywords

RNAa; saRNA; transcriptional activation; gene regulation

## BACKGROUND

Small double-stranded RNA (dsRNA) molecules including small interfering RNAs (siRNAs), microRNAs (miRNAs), and piwi-interacting RNAs (piRNAs) have been recognized as master regulators of gene expression. Short dsRNAs are the canonical trigger of RNA interference (RNAi)—a well-conserved mechanism of gene silencing in which dsRNAs inhibit translation or degrade complementary mRNA sequences [1,2]. RNAi is frequently implemented as a molecular tool to knockdown and study gene function both *in vitro* and *in vivo* [1-3]. It is also being developed in the clinic as means to inhibit production of disease-causing proteins [4]. Recently, small dsRNAs have also been used to target non-coding regulatory sequences such as promoters and 3' gene termini to either silence or activate gene transcription in phenomena known respectively as a transcriptional gene silencing (TGS) [5-7] and RNA activation (RNAa) [7-11].

RNAa can be achieved by targeting selected promoter sequences with short dsRNAs, which have been termed small activating RNAs (saRNAs) or antigene RNAs (agRNAs) to distinguish them from siRNAs and other RNA duplexes (**Fig. 1**). Like RNAi, RNAa depends on Argonaute (Ago) proteins, in particular Ago2, for saRNA maturation which occurs in the cytoplasm [12] and RNAa activity in the nucleus [8,13-15]. However, RNAa possesses unique kinetics that typically requires ~48 h before gene activation is readily detectable [8,16,17] (**Fig. 1D**) and can subsequently last for over a week [8,17](**Fig. 1E**). Published studies have demonstrated the presence of RNAa in human, mouse, and monkey cells, suggesting that it is an evolutionarily conserved mechanism in at least mammalian species [14,16,18].

Traditionally, gain-of-function studies often require the use of an exogenous DNA construct for ectopic expression. Such systems do not typically resemble natural genes, which are often cloned from cDNA libraries or PCR amplicons that lack regulatory elements such as introns and untranslated regions (UTRs) [19,20]. These regions are involved in multiple processes such as alternative splicing, post-transcriptional modification and transcript stability which can influence gene function [19,21-23]. In the clinic, gene overexpression systems have been problematic requiring viral-based systems to drive delivery of exogenous genes. Such approaches can have detrimental effects on host genome integrity and undesired immunological consequences [24]. An alternative approach of gene overexpression is in demand for safer gene therapy development.

RNAa offers a new method to direct gene overexpression. As a laboratory tool, it has the unique ability to enhance transcription of an endogenous targeted gene offering a more natural approach at analyzing gene function. Moreover, RNAa could be a new option for gene therapies. The ability to selectively upregulate gene expression in the absence of exogenous DNA can have far-reaching impacts in the realms of basic research and therapeutics development. In this protocol, we provided step-by-step procedures for performing RNAa experiments which involve the following 3 major steps: i) saRNA target selection, ii) picking saRNAs compatible with Ago2 loading, and iii) testing saRNAs by transfection and gene expression analysis.

# EXPERIMENTAL CONSIDERATIONS

## Promoter analysis and gene selection

Accurate identification and thorough analysis of promoter sequences are critical for saRNA design and target identification in gene promoters [15]. Unfortunately, the NCBI Nucleotide database contains only a limited number of promoter sequences, most of which are poorly annotated. As such, it is not an ideal or comprehensive primary resource for retrieving promoter sequences. Instead, major genome databases such as the UCSC genome browser (UGB) and the Ensembl genome browser (EGB) provide convenient tools for identifying and extracting promoter sequences. For instance, promoter regions can be approximated by isolating sequences prior to the first exon of gene transcripts defined within the genome databases. Databases containing information on expressed sequence tags (EST) (e.g dbEST) [25] and transcription start sites (TSS) (e.g. dbTSS) [26] can serve as additional resources for retrieving upstream sequences to map gene promoters. Ambiguity in determining the correct promoter in some genes can occur due to existence of alternative promoters. In such situations, examining promoter sequences within genome browsers (e.g. UGB or EGB) in conjunction with data tracks from genome-wide location analysis of RNA polymerase II (RNAPII) occupancy and histone markers can greatly improve prediction of TSS.

Once the promoter sequence for a target gene is identified, it is necessary to run basic bioinformatics sequence analyses against the sequence. Many promoters contain repetitive elements (e.g. simple repeats, interspersed repeats, etc.) that may have multiple copies scattered throughout the genome. As such, these elements should be excluded to avoid off-target effects (**Fig. 2A**). CpG islands, stretches of DNA with unusually high GC content and high frequency of CpG sites found in about 60% of human genes [27] should also be avoided. These sequences can be identified by running a CpG island prediction program (e.g. MethPrimer) [28]. The rationale for avoiding these specific sequences is two-fold: (i) RNA duplex with high GC content may hinder Ago processing of the passenger strand [29] and (ii) targeting promoter CpG islands with dsRNAs can cause TGS through DNA methylation-dependent [5] or independent [30] epigenetic mechanisms. For this reason, perhaps, CpG islands are preferable targets for inducing TGS.

It is also important to take into consideration basal levels of gene expression when searching for suitable target genes. First and foremost the gene itself must be present in the model cell line. For instance, cancer cells frequently contain deletions or mutations resulting in null status of gene activity. Absence of promoter sequences and/or the target gene would render RNAa ineffective. To improve susceptibility and magnitude of gene induction by RNAa, targetable genes should be poised for transcription and manifest low to intermediate levels of mRNA expression c. For instance, genes anticipated or identified to be downregulated (e.g. p21, KLF4, etc.) are suitable targets in cancer cells [32,33]. Genes that are already overexpressed tend to be refractory to RNAa because they are most likely already poised for maximal transcription. It is important to note that genes completely silenced by DNA methylation also tend to respond poorly to RNAa [8]. However, treatments with epigenetic modifying agents such as DNA demethylation agents and histone deacetylase inhibitors have been shown to restore susceptibility to or enhance RNAa [8,34] (**Fig. 1F**).

### saRNA design and target location

Success of saRNA design is currently limited by a lack of complete understanding of RNAa mechanism. Unbiased tiling design leads to the identification of saRNAs [9,11,34]; however, this approach can be time-consuming, expensive, and inefficient. Based on empirical observations, we have generated a set of rules that have improved the chance of identifying functional saRNA targets [18]. These rules can be divided into two major groups: (i) duplex-related rules, which define the sequence and chemical properties of a saRNA *per se*, and (ii) location-related rules that determine target location in a promoter. Because RNAa depends on Argonaute proteins much like RNAi [8,13,16], duplex-related rules follow similar guidelines as siRNA design including (i) duplex size of ~19 nt, (ii) 2-nt overhangs on the 3'-ends of the RNA strands, (iii) GC content between 40-60%, and (iv) asymmetric thermodynamic stability in base-pairing at the 5' end to drive strand selection by Argonaute proteins (**Fig. 2A-C**). In regards to target location, we generally design saRNAs targeting promoter sequence between −100 to −1000 bp relative to the TSS with the most responsive sites usually occurring within the −200 to −500 bp region of the promoter. We tend to avoid the region immediate upstream (−1 to −100 bp) of the TSS for two reasons: (i) many genes have multiple TSSs originated from a wide area that can span ~100 bp over the core promoter [35], and (ii) sequences proximal to the TSS usually have high GC content (e.g. GC boxes) [36]. Design of saRNAs in this region may erroneously target mRNA for degradation and/or possess undesirable GC levels. We normally favor saRNAs that target the 'preferred' region (−500 bp/−200 bp) in gene promoters. If possible, we favor duplexes that possess an A/T-rich sequence immediately flanking the target site on the 3'-end of the sense DNA strand relative to the direction of targeted gene transcription (**Fig. 2B**). A/T stretches have been reported to correlate with depleted nucleosome occupancy, improving accessibility of promoter DNA by regulatory proteins [37,38].

To improve success of identifying functional saRNAs, we suggest selecting 4–6 high-scoring targets from different areas in a gene promoter (**Fig. 2A**). In some cases, certain areas of a gene promoter support 'hotspots' in which several neighboring saRNAs induce variable levels of gene expression. Such regions can be identified by creating additional candidate saRNAs targeting overlapping or neighboring sequence [34]. These duplexes should also adhere to design rules as best as possible, although they may likely be less desirable by design standards. This approach can lead to the identification of multiple functional saRNAs creating several possible candidates for gene analysis and/or drug discovery.

### Methods for introducing saRNAs into cells

RNAa has been performed in cell lines [8,9,39] and primary cells [18,39] utilizing synthetic short dsRNAs [8,9,18] or expressed short hairpin RNA (shRNA) *in vivo* [14]. Any method that can introduce nucleic acids effectively into cells (especially into the nucleus) with limited cytotoxicity should be adequate to facilitate RNAa. Nucleofection is a method of choice for delivering nucleic acids directly into the nucleus with high efficiency and offers an option for delivery into primary cells, which can be refractory to transfection by many liposomal-based reagents [40]. Viral-based expression of shRNAs is another alternative method for performing RNAa in hard-to-transfect cells and has been utilized for *in vivo*

application in mouse models [14]. Liposomal transfection products (e.g. Lipofectamine, etc.) are a convenient and effective approach for delivery of saRNAs into many cell lines. Reverse or forward transfection protocols may offer slightly variable effects on transfection efficiency. A fluorescent-labeled dsRNA can be used to validate transfection or nucleofection efficiency of cell lines if needs be.

### Cell preparation for transfection

Healthy and exponentially growing cells are easily manageable for saRNA transfections. Cell stress should be minimized by avoiding overdigestion with trypsin and/or high speed centrifugation upon passage. Additionally, seeding density of cells for transfection is critical for detecting RNAa. Transfection of cells that are overcrowded or sparsely plated can result in inconsistent output and interfere with detection of maximal RNAa activity. As a rule of thumb, cells should be seeded at a concentration that prevents overcrowding at time of harvest. Although growth rates can vary between different cell types, seeding densities generally fall in the range of 40-60% for most cell lines. Activation of certain genes by RNAa can also interfere with cell growth rates and/or viability. In such situations, cells should be plated at densities in which the control treatments are not overcrowded at harvest. This will allow for visual monitoring of phenotypic changes associated/predicted to accompany activation of the target gene (**Fig. 3C**). However, if certain assays (e.g. cell cycle analysis) require normalization for growth potential and/or equal cell density at time of harvest, compensation can be performed during passage prior to transfection. For instance, if activation of a gene impairs cell growth (e.g. tumor suppressors), cells may need to be plated at higher densities in saRNA treatment groups to compensate for growth inhibition so cell content at time of harvest roughly equals those found in control treatments.

### saRNA concentrations and related kinetics

saRNAs have been reported to have $EC_{50}$ values around 1-10 nM [11,17,41]. It is generally higher than that of RNAi, which can fall into the picomolar or femtomolar range [42]. The elevated $EC_{50}$ may be partially attributed to its nuclear nature [16] in which greater concentrations of a saRNA would be required to compensate for the inherent nuclear exclusion properties attributed to duplex RNAs [43]. As such, optimal transfection concentrations for initial studies may range from 10 to 100 nM in order to measure RNAa activity. However, dose response experiments should be performed in order to determine the lowest concentration of a saRNA required for maximal induction of gene activation.

A unique feature of RNAa is that the onset of gene activation is delayed by ~48 h (**Fig. 1D**). The underlying reason for this observation is not entirely known, but may result from a more complicated mechanism with additional rate-limiting steps in comparison to RNAi [16]. A tentative explanation is that the entry of saRNAs into the nucleus is coupled to cell division. It has been recently shown that duplex RNAs are loaded by Ago2 in the cytoplasm before entering the nucleus [12]. It is possible that nuclear entry of RNA-Ago2 complex is a passive process which occurs during cell division when nuclear envelope breaks down. In contrast to RNAi, maximal activity of RNAa occurs within a window of ~72–96 h [16]. As such, analysis of gene expression should occur ~72 h after saRNA treatment in order to validate RNAa activity.

### Control treatments and experimental replication

It is important to include adequate controls with RNAa experiments:

**1. Control saRNAs**—Control saRNAs can be composed of unrelated or scrambled sequence in relationship to a functional saRNA. It should be noted that RNAa does not require perfect complementarity between the saRNA and its target sequence to function [8]. As such, controls should noticeably deviate from functional saRNA sequence. Additionally, control duplexes should be checked for potential homology with sequence in the genome of the same organism in order to limit putative function. BLAST searches can be performed by selecting the RefSeq RNA or the reference genome database. A sequence that returns hits with significant homology (   80%), particularly if it matches known transcripts or sequence neighboring the targeted promoter, should be rejected as controls. Ideally, control duplexes should possess similar GC content and terminal thermodynamic asymmetry as the functional saRNA. Testing RNAa activity against multiple control duplexes also enriches validity of saRNA experiments.

**2. Transfection controls**—Transfections in absence of duplex RNA can be a very useful control for identifying any effects on cell viability or gene expression related to the use of transfection reagents.

Both technical and biological repeats should be performed in order to acquire statistically significant results. For each transfection experiment, at least duplicate samples should be transfected and analyzed for changes in gene expression. Each experiment should then be repeated at least 3 times. The average of all replicates and repeats represents the final level of gene expression. Changes in expression levels should be quantified and evaluated against control treatments for statistical significance.

### Data interpretation

Cautions should be taken when interpreting results to exclude potential non-specific and off-target effects that may be rampant and difficult to distinguish from on-target effects. Several assays may be necessary to validate target specificity of gene activation and associated phenotypic changes.

**1. Assessment of transcriptional gene activation**—The tech nique of RNAa described in this report occurs at the level of transcription. To distinguish from post-transcriptional mechanisms that can lead to elevated mRNA levels, there are several experimental options that may be performed to confirm transcriptional activation. For instance, nuclear runon is considered the gold standard for accessing transcriptional upregulation at a specific target gene. Additionally, enrichment of RNAPII can be quantified at the target gene promoter by ChIP analysis as a marker for transcriptional activation. Measuring levels of pre-mRNA by RT-PCR can also be used as a method to assess changes in gene transcription [44].

**2. Multiple saRNAs targeting the same promoter**—Identifying multiple saRNAs that activate gene expression at different positions (e.g. dsEcad-215 and −302) improves the

likelihood that gene induction is dependent on the targeted promoter (**Fig. 1A-C**). Each activating duplex likely possesses a different 'seed' sequence and target non-overlapping regions. As such, the probability for each duplex functioning by suppression of non-specific upstream regulators becomes increasing lower. Additionally, if changes in molecular and/or cellular downstream phenotypes are conserved between multiple saRNAs, it correlates with gene activation being on-target.

**3. Verifying target specificity by vector-mediated overexpression—**Ectopic expression of genes using vector-based systems remains the standard method for gain-of-function study. The specificity of downstream gene modulation and functional consequences following RNAa can be verified by vector-mediated overexpression of the same gene. Recapitulation of RNAa by an ectopic system can validate the phenotypic changes facilitated by saRNAs [32,45].

**4. Evaluating candidate off-target gene expression—**Even carefully selected saRNAs or siRNAs cannot avoid partial sequence homology with other coding and non-coding sequences. Partial homology with off-target coding sequence may trigger miRNA-like mechanisms of post-transcriptional gene silencing. As such, expression of putative off-target genes should be screened to remove the potential for non-specific mechanisms of gene activation.

### Limitations

In our previous studies, not every gene tested was responsive to promoter targeting when testing 1–2 saRNAs per gene [8,18]. Although the design rules have since improved and the recommended number of saRNAs utilized in screenings has increased, there still remains the possibility that some targeted promoters may not respond to the saRNAs. It is unclear if all genes are susceptible to activation if all possible targets on their promoters are exhausted. It has been shown that genes completely silenced by DNA methylation [8] and/or genes silenced in a tissue- or cell type-specific manner are less accessible to RNAa [34]. In such cases, treating cells with subtoxic doses of demethylating agents (e.g. 5-AzaC) [8] or histone deacetylase inhibitors [e.g. valproic acid (VPA)]) [34] may restore/promote susceptibility to RNAa (**Fig. 1F**). As such, RNAa activity can vary between different cells types/lines in regards to target gene expression. One saRNA that works in one cell type may not work with equal efficacy in another [9]. Differences in basal levels of expression, promoter environment, chromatin state, target availability, etc. all likely influence susceptibility to a given saRNA.

## MATERIALS

### Reagents

✓ Synthetic short duplex RNA (desalted or HPLC purified)

✓ Diethyl pyrocarbonate (DEPC, Sigma, cat. no. 40718). CAUTION: Toxic. Use personal protective equipment and avoid breathing vapors, mist or gas

✓ RPMI-1640 medium (Invitrogen, cat. no. 11875-093)

✓ Opti-MEM® I Reduced Serum Medium (Invitrogen, cat. no. 31985-062)

✓ 1×Trypsin-EDTA (0.05% Trypsin with EDTA 4Na) 1× (Invitrogen, cat. no. 25300-054)

✓ Lipofectamine™ RNAiMax™ transfection reagent (Invitrogen, cat. no. 13778-150)

✓ Lipofectamine 2000™ transfection reagent (Invitrogen, cat. no. 11668-019)

✓ Silencer® Cy™3 labeled Negative Control No. 1 siRNA (Applied Biosystems, cat. no. AM4621)

✓ GelRed (Biotium, cat. no. 41003)

✓ Power SYBR Green PCR Master Mix (2X) (Applied Biosystems, cat. no. 4367659)

✓ TaqMan® Fast Universal PCR Master Mix (Applied Biosystems, cat. no. 4352042)

✓ Reagents for standard cell lysis and protein quantification

✓ RIPA buffer with EDTA (Boston BioProducts, cat. no. BP-115D)

✓ Protease inhibitor cocktail (Sigma, cat. no. P8340)

✓ BCA protein assay reagent (bicinchoninic acid; Thermo Scientific, cat. no. 23227)

✓ Reagents for standard SDS-PAGE and western blotting assay:

✓ 30% Acrylamide/Bis solution, 29:1 (Bio-Rad, cat. no. 161-0156)

✓ APS (ammonium persulfate; Bio-Rad, cat. no. 161-0700)

✓ TEMED (N,N,N′,N′-tetramethylethylenediamine; Sigma, cat. no. T22500)

✓ Trizma base (Sigma, cat. no. T6066)

✓ Glycine (USB, cat. no. 16407)

✓ Methanol (Fisher Scientific, cat. no. A412)

✓ SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, cat. no. 34087)

**Equipment**

✓ Sterile 6-well cell culture plates (Nunc, cat. no. 140685)

✓ Sterile 15 ml conical tubes (Santa Cruz Biotechnology, cat. no. sc-200250)

✓ 1.5-ml microcentrifuge tubes (USA Scientific, cat. no. 1615-5510)

✓ Standard cell culture equipment including laminar flow hood, humidified tissue culture incubators at 37°C and 5% $CO_2$, water bath, refrigerator and −20°C freezer.

✓ Refrigerated centrifuge with swing-bucket rotor accommodating 15-ml conical tubes.

✓ Hemocytometer

✓ Inverted microscope with combined ocular-objective magnification of 40×, 100× and 250× and camera

✓ 7500 Fast Real-Timer PCR System (Applied Biosystems, cat. no. 4351107)

✓ Standard molecular biology plasticware and glassware

✓ UV transilluminator

✓ Agarose gel electrophoresis apparatus

✓ Single-channel pipettes (2 μl, 10 μl, 100 μl and 1000 μl) with arousal resistant plastic pipette tips

✓ NanoDrop 2000 (Thermo Scientific)

✓ Personal computer with Internet access and web browser

✓ Sequence analysis and manipulation program, or text editor

✓ Equipment for standard cell lysis, SDS-PAGE and western blotting assay including refrigerated centrifuge such as centrifuge Legend RT (Thermo Scientific, cat. no. 75-004-377), spectrophotometer such as Multiskan MCC/340 (Thermo Scientific, cat. no. 14-386-26), vertical electrophoresis systems such as Mini Protean Tetra Cell (Bio-Rad, cat. no. 166-0828EDU), shaker such as Nutator Mixer (BD, cat. no. 421105), nitrocellulose membranes such as NitroBind (GE Water & Process Technologies, EP4HY00010), film cassette such as autoradiography cassettes (Fisher Scientific, cat. no. FB-CS-810), and autoradiography film such as HyBlot CL (Denville scientific, cat. no. E3018).

### Reagent setup

Stock dsRNA solution (20 μM): To make 20 μM dsRNA stock solutions, reconstitute lyophilized synthetic dsRNA with desired amount of DEPC-treated water. Mix well and store at –80°C in 200–250 μl aliquots.

DEPC-treated water: Add 1 ml of 0.1% DEPC to 1000 ml distilled or nanopure water. Mix well by rigorous shaking and let the solution rest at room temperature for at least 1 h. Autoclave for 15 min to inactive DEPC and cool to room temperature prior to use.

### Equipment setup

A standard cell culture facility should be established and properly maintained. Contamination of cells by mycoplasmas, fungi, and bacteria should be routinely screened and disinfected as needed in order to sustain pure cultures. A standard molecular biology work bench with RNase-free environment is required when handling duplex RNAs and RNA extracts.

## PROCEDURE

### Promoter identification and sequence retrieval (1 h)

**1.** Using the Ensembl Genome Browser (EGB)

    **1.1** Go to EGB at http://www.ensembl.org

**1.2** Choose the desired genome in the search box and type the name or symbol of the intended target gene.

**1.3** On the result page, choose the correct gene by clicking "Gene" under the category "By Feature type".

**1.4** On the "Result in Detail" page, click the title of the correct gene.

**1.5** Click "Sequence" on the left column, which will lead to the marked-up sequence page. On the lower half of the page, a 600-bp (by default) 5' flanking sequence, exons, introns, 3' UTR, and 3' flanking sequence are displayed with exons in red text.

**1.6** To display 1 kb of the 5' flanking sequence of the gene, click "Configure this page" on the lower left column and a new window will pop up. In the text box for "5' Flanking sequence (upstream)", type "1000" and click the check ( ) sign on the upper right hand corner of the window to save the change and close the window.

**1.7** The marked-up sequence page will reload to reflect the change by displaying 1 kb 5' flanking sequence at the beginning followed by the genomic sequence for the rest of the gene.

**1.8** Copy the first 1 kb sequence which is the putative promoter sequence of the gene and paste it into a text editing program.

**2.** Using the UCSC Genome Browser (UGB)

**2.1** Go to UGB at http://genome.ucsc.edu/cgi-bin/hgGateway

**2.2** Choose the desired genome by first selecting clade and then genome in the dropdown menu of the search box.

**2.3** In the box "position or search term", type the name or symbol of the intended target gene and click "submit".

**2.4** In the result page, under RefSeq genes or UCSC genes, click the correct gene. The next page will display the genomic structure of the gene along with many different annotation tracks.

**2.5** Click "View" then "DNA" on the top menu bar, and type "1000" in the box for "Add [ ] extra bases upstream (5')" and then click "get DNA".

**CAUTION:** If the gene is transcribed from the '-' strand of the DNA, check "Reverse complement" in the Sequence Formatting Options to get '-' strand sequence.

**2.6** Copy the first 1 kb sequence as the promoter sequence of the gene and paste into a text editing program.

**CAUTION:** There are sometimes discrepancies in TSS annotation among the databases. They are often caused by the existence of alternative TSSs upstream or downstream the major TSS. In such a case, data from dbTSS should help accurately determine the TSS. In addition, many annotation tracks derived from bioinformatics

sequence analyses and experimental data are available in both EGB and UGB and may also help determine the correct TSS. These annotations include tracks of CpG islands, TSS mapping data, ESTs, histone markers and RNAPII occupancy.

### Promoter sequence analysis (1 h)

**3** Identify CpG islands in the retrieved promoter sequence using UGB CpG island annotation or other online programs such as MethPrimer at http://www.urogene.org/cgi-bin/methprime, EMBOSS CpGPlot/CpGReport at http://www.ebi.ac.uk/Tools/emboss/cpgplot.

**4** Identify repetitive sequences and sequence variants by examining the sequence in UGB's "Variation and Repeats tracks" or using the RepeatMasker Web Server at http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker.

**TIP:** If a repetitive sequence or part of it is unique to the target promoter, it can still be targetable.

### saRNA target selection (1-2 h)

**5** Manually saRNA target selection

**5.1** Paste the 1-kb promoter sequence into a sequence analysis program or any text editor.

**5.2** Starting from −100 bp, visually examine the sequence toward its 5' end and select targets based on the following rules (**Table 1**):

**5.2.1** Use the sense DNA sequence of the promoter as the template for saRNA design.

**5.2.2** Targets can be selected within a promoter region between at −100 and −1000 bp upstream the TSS.

**5.2.3** Targets should be 19-nt in size.

**5.2.4** Targets should have a GC content of 40-60%, GC-rich regions or CpG islands should be avoided.

**5.2.5** The corresponding saRNA duplexes should have lower thermodynamic stability at the 3' end than the 5' end (**Fig. 2**).

**5.2.6** The 18th and 19th positions counted from the 5' end of targets should be "A/T"s, preferably "A"s.

**5.2.7** Avoid sequences that have 5 or more consecutive nucleotides.

**5.2.8** Avoid simple repeat sequences such as di- or tri-nucleotide repeats.

**5.2.9** The 20-23th nucleotides (flanking the 3' end of a target) should preferably be "A"s or "T"s.

**5.3** Repeat Step 5 to choose 4-5 targets.

**6** Select saRNA targets using an Excel program which implements the above rules (**File S1**).

**6.1** Download the Excel program (**File S1**) and save it on a local computer.

**6.2** Open the Excel file and choose "Allow Macro" when prompted.

**6.3** Paste the 1-kb promoter sequence into the DNA sequence input box.

**6.4** Optional: Specify the location of the TSS if it is known.

**6.5** Click the button "Click here to find the best saRNAs".

**6.6** Choose 4–5 saRNA targets based on overall quality score.

**7** For each candidate target, run a BLAST search by going to http://blast.ncbi.nlm.nih.gov/Blast.cgi.

**8** Use the 19-nt target sequence as query sequence, "Reference genome sequences" as "Search Set", and "Optimize for Highly similar sequences (megablast)" as "Search Program".

**9** Use the 19-nt target sequence along with 1 nt 5' or 3' flanking sequence as query sequence to run a BLAT search (http://genome.ucsc.edu/cgi-bin/hgBlat) to identify SNPs within the target and reject the target if it contains SNPs.

**CAUTION:** Since the acceptable minimal length of query sequence for BLAT is 20 nt, it is necessary to add at least one nt flanking sequence to the 19-nt target sequence to run the search.

**10** Once a 19-nt target is selected, create a sense RNA sequence based on the target sequence by converting all 'T's to 'U's in the target sequence and an antisense RNA sequence by reverse complementing the sense RNA sequence.

**11** Add either [dT][dT] or [U][U] overhangs to the 3'-end of both RNA strands.

**12** Perform chemical synthesis of the RNA strands and anneal them to make a dsRNA duplex, or order the dsRNA duplex from a reliable vendor.

**CAUTION:** In some promoters the existence of long CpG islands leaves little room for picking ideal saRNA targets. In this case, targets can still be selected in less CpG-rich areas within a CpG island [32].

### saRNA transfection of cells using a reverse transfection protocol (3–4 d)

We routinely use Lipofectamine RNAiMAX to transfect saRNA into cultured cells as we found it an efficient delivery reagent with minimal cytotoxicity in most cell types in comparison to other commercial reagents.

**Preparing saRNA and RNAiMAX transfection complex for 6-well format**

13   Prepare one tube (Tube A, for diluting saRNAs) for each well to be transfected, and one master mixture tube (Tube B, for diluting transfection reagent) for all wells.

14   Remove RNAiMAX from the refrigerator and let it warm up to room temperature.

15   Before use, gently mix RNAiMAX solution by tapping on the tube.

16   Add 250 μl of Opti-MEM medium and 5 μl RNAiMAX to tube B, mix gently by tapping on the tube or pipetting up and down a few times. A master mixture can be prepared in Tube B for all treatments.

17   Add 250 μl Opti-MEM medium to each of Tube As.

18   Add 6.25 μl of 20 μM saRNA to each of Tube As and mix well. For mock transfection, omitting dsRNA.

19   Add 250 μl of the diluted RNAiMAX from tube B into each of Tube As. Mix well by pipetting or inverting the tubes several times.

20   Let the tube containing both diluted saRNA and diluted RNAiMAX set at room temperature for 10–20 min to allow transfection complex to form.

21   If a transfection reagent other than RNAiMAX is used, follow manufacturer's manual.

**CAUTION:** For each dsRNA to be transfected, replicates (duplicate or triplicate) should be done as well as control saRNA transfections and mock transfections.

**Preparing cells for transfection—**While waiting for the transfection complex to form, prepare cells for transfection.

22   Remove medium from exponentially growing cells in a 10-cm dish.

23   Wash cells with 1–2 ml PBS and remove the PBS.

24   Add 1 ml 0.05% trypsin into the dish and rock the dish to allow for even coverage of the surface by trypsin.

25   Return the dish to a $CO_2$ incubator and wait for 5 min.

26   While waiting, add 10 ml fresh complete medium into a 15-ml conical tube.

27   Remove the dish from the incubator after 5 min (or shorter to prevent over-digestion). Pipet up and down the trypsin solution containing digested cells to disperse cells into single cell suspension.

28   Transfer cells to the 15-ml conical tube containing 10 ml complete medium.

29   Mix the cells with the complete medium by pipetting.

30   Take 100 μl of the mixed cell suspension and add it to a 1.5-ml microcentrifuge tube containing 900 μl PBS to dilute the cells, mix well for cell counting.

**31** Centrifuge the remaining cells in the 15-ml tube at $800 \times g$ and 4°C for 5 min.

**32** While waiting, count the cells using a hemocytometer and calculate the total number of cells in the 15-ml tube.

**33** When centrifugation is done, remove the supernatant from the tube and resuspend cells in the desired amount of medium without antibiotics that gives the desired cell concentration.

**TIP:** The number of cells to be seeded in each well of 6-well plates should be pre-determined so that they will be 50–60% confluent 24 h after plating. It varies with different types of cells depending on cell doubling time.

**34** Mix the cells well and add 2 ml of the cells suspension to each well.

**35** Gently swing the plate to distribute the cells evenly in the wells.

**36** Return the plate to the incubator if the transfection mixture is still incubating.

**Adding transfection complex to the cells**

**37** Add the transfection mixture to each well drop-wisely and evenly to all areas of the well.

**38** Check under a microscope to make sure that the cells are evenly distributed across the surface of every well; otherwise, swing the plate again to distribute the cells evenly.

**39** Return the plate to the incubator.

**40** Optional: Change the transfection medium with fresh complete medium 24 h later.

**41** Observe and document cell morphology using an inverted microscope equipped with a camera starting from 48 h when RNAa effect begins to appear.

**TIP:** The success of RNAa experiments highly depend on transfection efficiency. When working with a new cell line with unknown dsRNA transfection efficiency, the transfection of a fluorescence-labeled dsRNA is recommended. Generally, RNAiMax produces acceptable transfection efficiency in many cell types with minimum cytotoxicity. For hard-to-transfect cells, transfect cells with Lipofectamine® 2000 by following the forward transfection protocol included in its manual, and change the transfection medium with fresh complete medium after 5–6 h to alleviate potential cytotoxicity of Lipofectamine® 2000.

**RNA isolation and reverse transcription (RT) reaction (1 d)**

**42** Isolate total cellular RNA using a method or kit of choice.

**43** Treat the RNA with DNase I to remove potential DNA contamination.

**44** Accurately measure RNA concentrations using Nanodrop 2000 or a UV spectrophotometer.

**45** Use 100–1000 ng of the isolated RNA in RT reaction along with oligo(dT) or random hexamer primers, RNase inhibitor, dNTP, reverse transcriptase with compatible buffer in a 25–30 µl RT reaction to convert the RNA into cDNA.

**46** Dilute the resulted cDNA (1:4) using nuclease-free water.

**CAUTION:** Precautions and standard practice should be used to avoid RNase contamination when handling RNA. Accurate measurement of RNA concentration and subsequent use of equal quantity of RNA in RT reaction are critical for reliably evaluating changes in gene expression. Big differences in the amount of starting RNA among samples cannot be accurately normalized by simultaneous detection of the mRNA levels of a housekeeping gene. It is recommended not to pause the procedure after RNA isolation, and RT reaction should be performed immediately without freezing and thawing the RNA. It is also important to dilute the cDNA sample before PCR amplification for the following considerations: to dilute out any impurities that may inhibit DNA polymerase and to allow for detection of maximum differences in mRNA levels of the target gene caused by treatments.

### Analysis of gene activation at the mRNA level by semi-quantitative and quantitative RT-PCR (1 d)

**47** mRNA expression analysis by semi-quantitative RT-PCR.

**47.1** Using the diluted cDNA as template to amplify mRNA of the target gene using gene specific RT-PCR primers by semi quantitative RT-PCR in a PCR thermocycler. A housekeeping gene should also be amplified as RNA loading controls.

**CAUTION:** The proper number of PCR cycles should be determined by experiments for both target gene and housekeeping gene to make sure that the amplification of cDNA is within the exponential phase.

**47.2** Use 1–2% agarose gel prestained with GelRed^tm Stain to separate the PCR products.

**47.3** Visualize and document the gel image using a UV transilluminator at 254 nm.

**48** mRNA expression analysis by quantitative RT-PCR (qRT-PCR). To quantify mRNA expression levels of the target gene and a housekeeping gene, use standard SYBR® green or TaqMan® based detection using a real-time thermal cycler.

**CAUTION:** Appropriate RT-PCR primer design is critical for detecting gene activation. The primers should span introns or exon boundary so that potential DNA contamination in RNA can be easily identified by the size differences in PCR products. If the target gene has multiple isoforms resulted from alternative splicing or use of TSS, primers should be designed that amplify most of the isoforms or the major isoforms only transcribed from the target promoter (**Fig. 3A-B**) Appropriate selection of a housekeeping gene is important because the expression of some commonly used housekeeping genes such as ACTB (ß-

actin) can be affected by certain treatments [46]. If GAPDH (glyceraldehyde 3-phosphate dehydrogenase) is used as the housekeeping gene, the RT-PCR primers should be carefully designed for there exist multiple GAPDH pseudogenes and its mRNA level detected by primers that can also amplify pseudogenes could be severely skewed by minute amount of DNA contamination in RNA preparations.

### Analysis of gene activation at the protein level by standard western blotting assay (1–2 d)

**49** Collect cells by scraping, wash cell pellet twice with PBS and lyse cells with RIPA buffer with protease inhibitor cocktail for 30 min on ice. Centrifuge lysates for 15 min at 14,000 rpm 4°C and collect supernatants.

**50** Optional: If activation of a target gene is expected to cause a decrease in cell proliferation than control treatments such as the activation of p21 gene [33], combine multiple wells of saRNA treated cells to obtain enough cell lysate.

**51** Determine protein concentration by BCA protein assay.

**52** Load equal amounts of total protein in lysates (30–50 μg) onto a SDS-PAGE gel and transfer the separated proteins to a nitrocellulose membrane.

**53** Block the membrane in 5% nonfat milk and incubate with primary antibodies at 4°C overnight. After incubation with horseradish peroxidase-conjugated secondary antibodies, detect signals by chemiluminescent reagents.

## TIMING

Steps 1–2, Promoter identification and sequence retrieval: 1 h.

Steps 3–4, Promoter sequence analysis: 1 h.

Steps 5–12, saRNA target selection: 1–2 h.

Steps 13–41, dsRNA transfection of cells: 3–4 d.

Steps 42–46, RNA isolation and RT reaction: 1 d.

Steps 47–48, Analysis of gene activation at the mRNA level: 1 d.

Steps 49–53, Analysis of gene activation at the protein level: 2 d

## ANTICIPATED RESULTS

Transfecting cultured cells with promoter-targeted dsRNAs may result in three possible outcomes: (i) no change, (ii) upregulation, or (iii) downregulation of target gene expression. If the design rules described in Step 6 are followed, it is expected that 15–20% of the designed ds-RNAs will significantly activate target gene expression, while 70% will have no effect and 5–10% will suppress gene expression. If 4–6 targets are selected the probability at least one will activate gene expression is significantly improved.

It is anticipated that induction of gene expression after transfection of saRNAs should follow the typical kinetics of RNAa, which is delayed by ~48 h. Maximal levels of gene induction may require 3–4 days. However, RNAa can persist for over 10 days and, in some cases, last

up to 3–4 weeks after a single transfection. The length at which saRNAs activate gene expression may likely depend on the rate of cell division and dilution of intracellular saRNA concentrations.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

## Abbreviations used

| | |
|---|---|
| **saRNA** | small activating RNA |
| **RNAa** | RNA activation |
| **dsRNA** | double-stranded RNA |
| **TSS** | transcription start site |
| **5-AzaC** | 5-aza-2′-deoxycytidine |

## References

1. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, et al. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 1998; 391:806–811. doi: 10.1038/35888. PMID: 9486653. [PubMed: 9486653]

2. Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature. 2001; 411:494–498. doi: 10.1038/35078107. PMID: 11373684. [PubMed: 11373684]

3. Song E, Lee S, Wang J, Ince N, Ouyang N, et al. RNA interference targeting Fas protects mice from fulminant hepatitis. Nat Med. 2003; 9:347–351. doi: 10.1038/nm828. PMID: 12579197. [PubMed: 12579197]

4. Davis ME, Zuckerman JE, Choi CHJ, Seligson D, Tolcher A, et al. Evidence of RNAi in humans from systemically administered siRNA via targeted nanoparticles. Nature. 2010; 464:1067–1070. doi: 10.1038/nature08956. PMID: 20305636. [PubMed: 20305636]

5. Morris KV, Chan SW, Jacobsen SE, Looney DJ. Small interfering RNA-induced transcriptional gene silencing in human cells. Science. 2004; 305:1289–1292. doi: 10.1126/science.1101372. PMID: 15297624. [PubMed: 15297624]

6. Janowski BA, Huffman KE, Schwartz JC, Ram R, Hardy D, et al. Inhibiting gene expression at transcription start sites in chromosomal DNA with antigene RNAs. Nat Chem Biol. 2005; 1:216–222. doi: 10.1038/nchembio725. PMID: 16408038. [PubMed: 16408038]

7. Yue X, Schwartz JC, Chu Y, Younger ST, Gagnon KT, et al. Transcriptional regulation by small RNAs at sequences downstream from 3' gene termini. Nat Chem Biol. 2010; 6:621–629. doi: 10.1038/nchembio.400. PMID: 20581822. [PubMed: 20581822]

8. Li LC, Okino ST, Zhao H, Pookot D, Place RF, et al. Small dsRNAs induce transcriptional activation in human cells. Proc Natl Acad Sci U S A. 2006; 103:17337–17342. doi: 10.1073/pnas.0607015103. PMID: 17085592. [PubMed: 17085592]

9. Janowski BA, Younger ST, Hardy DB, Ram R, Huffman KE, et al. Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. Nat Chem Biol. 2007; 3:166–173. doi: 10.1038/nchembio860. PMID: 17259978. [PubMed: 17259978]

10. Place RF, Li LC, Pookot D, Noonan EJ, Dahiya R. MicroRNA-373 induces expression of genes with complementary promoter sequences. Proc Natl Acad Sci U S A. 2008; 105:1608–1613. doi: 10.1073/pnas.0707594105. PMID: 18227514. [PubMed: 18227514]

11. Matsui M, Sakurai F, Elbashir S, Foster DJ, Manoharan M, et al. Activation of LDL receptor expression by small RNAs complementary to a noncoding transcript that overlaps the LDLR promoter. Chem Biol. 2010; 17:1344–1355. doi: 10.1016/j.chembiol.2010.10.009. PMID: 21168770. [PubMed: 21168770]

12. Gagnon KT, Li LC, Chu Y, Janowski BA, Corey DR. RNAi factors are present and active in human cell nuclei. Cell Rep. 2014; 6:211–221. doi: 10.1016/j.celrep.2013.12.013. PMID: 24388755. [PubMed: 24388755]

13. Chu Y, Yue X, Younger ST, Janowski BA, Corey DR. Involvement of argonaute proteins in gene silencing and activation by RNAs complementary to a non-coding transcript at the progesterone receptor promoter. Nucleic Acids Res. 2010; 38:7736–7748. doi: 10.1093/nar/gkq648. PMID: 20675357. [PubMed: 20675357]

14. Turunen MP, Lehtola T, Heinonen SE, Assefa GS, Korpisalo P, et al. Efficient regulation of VEGF expression by promoter-targeted lentiviral shRNAs based on epigenetic mechanism: a novel example of epigenetherapy. Circ Res. 2009; 105:604–609. doi: 10.1161/CIRCRESAHA. 109.200774. PMID: 19696410. [PubMed: 19696410]

15. Gagnon KT, Corey DR. Argonaute and the nuclear RNAs: new pathways for RNA-mediated control of gene expression. Nucleic Acid Ther. 2012; 22:3–16. doi: 10.1089/nat.2011.0330. PMID: 22283730. [PubMed: 22283730]

16. Portnoy V, Huang V, Place RF, Li LC. Small RNA and transcriptional upregulation. Wiley Interdiscip Rev RNA. 2011; 2:748–760. doi: 10.1002/wrna.90. PMID: 21823233. [PubMed: 21823233]

17. Place RF, Noonan EJ, Földes-Papp Z, Li LC. Defining features and exploring chemical modifications to manipulate RNAa activity. Curr Pharm Biotechnol. 2010; 11:518–526. PMID: 20662764. [PubMed: 20662764]

18. Huang V, Qin Y, Wang J, Wang X, Place RF, et al. RNAa is conserved in mammalian cells. PLoS One. 2010; 5 doi: 10.1371/journal.pone.0008848. PMID: 20107511.

19. Brinster RL, Allen JM, Behringer RR, Gelinas RE, Palmiter RD. Introns increase transcriptional efficiency in transgenic mice. Proc Natl Acad Sci U S A. 1988; 85:836–840. PMID: 3422466. [PubMed: 3422466]

20. Clark AJ, Archibald AL, McClenaghan M, Simons JP, Wallace R, et al. Enhancing the efficiency of transgene expression. Philos Trans R Soc Lond B Biol Sci. 1993; 339:225–232. doi: 10.1098/rstb.1993.0020. PMID: 8097052. [PubMed: 8097052]

21. Stemmler MP, Hecht A, Kemler R. E-cadherin intron 2 contains cis-regulatory elements essential for gene expression. Development (Cambridge, England. 2005; 132:965–976. doi: 10.1242/dev. 01662. PMID: 15673570.

22. Breitbart RE, Nguyen HT, Medford RM, Destree AT, Mahdavi V, et al. Intricate combinatorial patterns of exon splicing generate multiple regulated troponin T isoforms from a single gene. Cell. 1985; 41:67–82. PMID: 2986851. [PubMed: 2986851]

23. Leff SE, Rosenfeld MG, Evans RM. Complex transcriptional units: diversity in gene expression by alternative RNA processing. Annu Rev Biochem. 1986; 55:1091–1117. doi: 10.1146/annurev.bi. 55.070186.005303. PMID: 3017190. [PubMed: 3017190]

24. Hedman M, Hartikainen J, Yla-Herttuala S. Progress and prospects: hurdles to cardiovascular gene therapy clinical trials. Gene Ther. 2011; 18:743–749. doi: 10.1038/gt.2011.43. PMID: 21490683. [PubMed: 21490683]

25. Boguski MS, Lowe TM, Tolstoshev CM. dbEST--database for "expressed sequence tags. Nat Genet. 1993; 4:332–333. PMID: 8401577. [PubMed: 8401577]
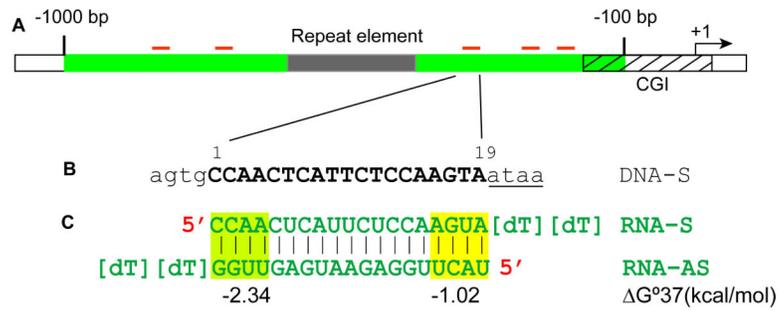
26. Wakaguri H, Yamashita R, Suzuki Y, Sugano S, Nakai K. DBTSS: database of transcription start sites, progress report 2008. Nucleic Acids Res. 2008; 36:97–101. doi: 10.1093/nar/gkm901. PMID: 17942421.

27. Larsen F, Gundersen G, Lopez R, Prydz H. CpG islands as gene markers in the human genome. Genomics. 1992; 13:1095–1107. PMID: 1505946. [PubMed: 1505946]

28. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002; 18:1427–1431. PMID: 12424112. [PubMed: 12424112]

29. Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, et al. Rational siRNA design for RNA interference. Nat Biotechnol. 2004; 22:326–330. doi: 10.1038/nbt936. PMID: 14758366. [PubMed: 14758366]

30. Ting AH, Schuebel KE, Herman JG, Baylin SB. Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation. Nat Genet. 2005; 37:906–910. doi: 10.1038/ng1611. PMID: 16025112. [PubMed: 16025112]

31. Xia T, SantaLucia J Jr, Burkard ME, Kierzek R, Schroeder SJ, et al. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. Biochemistry. 1998; 37:14719–14735. doi: 10.1021/bi9809425. PMID: 9778347. [PubMed: 9778347]

32. Wang J, Place RF, Huang V, Wang X, Noonan EJ. Prognostic Value and Function of KLF4 in Prostate Cancer: RNAa and Vector-Mediated Overexpression Identify KLF4 as an Inhibitor of Tumor Cell Growth and Migration. Cancer Res. 2010; 70:10182–10191. doi: 10.1158/0008-5472.CAN-10-2414. PMID: 21159640. [PubMed: 21159640]

33. Chen Z, Place RF, Jia Z, Pookot D, Dahiya R, et al. Antitumor effect of dsRNA-induced p21(WAF1/CIP1) gene activation in human bladder cancer cells. Mol Cancer Ther. 2008; 7:698–703. doi: 10.1158/1535-7163.MCT-07-2312. PMID: 18347154. [PubMed: 18347154]

34. Wang J, Huang V, Ye L, Bárcena A, Lin G, et al. Identification of Small Activating RNAs that Enhance Endogenous OCT4 Expression in Human Mesenchymal Stem Cells. Stem Cells Dev. 2014; 24:345–353. doi: 10.1089/ scd.2014.0290. PMID: 25232932. [PubMed: 25232932]

35. Yamashita R, Sathira NP, Kanai A, Tanimoto K, Arauchi T, et al. Genome-wide characterization of transcriptional start sites in humans by integrative transcriptome analysis. Genome Res. 2011; 21:775–789. doi: 10.1101/gr.110254.110. PMID: 21372179. [PubMed: 21372179]

36. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci U S A. 2006; 103:1412–1417. doi: 10.1073/pnas.0510310103. PMID: 16432200. [PubMed: 16432200]

37. Anderson JD, Widom J. Poly(dA-dT) promoter elements increase the equilibrium accessibility of nucleosomal DNA target sites. Mol Cell Biol. 2001; 21:3830–3839. doi: 10.1128/MCB. 21.11.3830-3839.2001. PMID: 11340174. [PubMed: 11340174]

38. Field Y, Kaplan N, Fondufe-Mittendorf Y, Moore IK, Sharon E, et al. Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. PLoS Comput Biol. 2008; 4 doi: 10.1371/journal.pcbi.1000216. PMID: 18989395.

39. Chen R, Wang T, Rao K, Yang J, Zhang S, et al. Up-regulation of VEGF by small activator RNA in human corpus cavernosum smooth muscle cells. J Sex Med. 2011; 8:2773–2780. doi: 10.1111/j. 1743-6109.2011.02412.x. PMID: 21819543. [PubMed: 21819543]

40. Nakayama A, Sato M, Shinohara M, Matsubara S, Yokomine T, et al. Efficient transfection of primarily cultured porcine embryonic fibroblasts using the Amaxa Nucleofection system. Cloning Stem Cells. 2007; 9:523–534. doi: 10.1089/clo.2007.0021. PMID: 18154513. [PubMed: 18154513]

41. Kang MR, Yang G, Place RF, Charisse K, Epstein-Barash H, et al. Intravesical delivery of small activating RNA formulated into lipid nanoparticles inhibits orthotopic bladder tumor growth. Cancer Res. 2012; 72:5069–5079. doi: 10.1158/0008-5472.CAN-12-1871. PMID: 22869584. [PubMed: 22869584]

42. Vaishnaw AK, Gollob J, Gamba-Vitalo C, Hutabarat R, Sah D, et al. A status report on RNAi therapeutics. Silence. 2010; 1:14. doi: 10.1186/1758-907X-1-14 PMID: 20615220. [PubMed: 20615220]

43. Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev. 2003; 17:3011–3016. doi: 10.1101/gad.1158803. PMID: 14681208. [PubMed: 14681208]

44. Schwartz JC, Corey DR. Practical considerations for analyzing antigene RNAs (agRNAs): RNA immunoprecipitation of argonaute protein. Methods Mol Biol. 2011; 764:301–315. doi: 10.1007/978-1-61779-188-8_20. PMID: 21748649. [PubMed: 21748649]

45. Wang X, Wang J, Huang V, Place RF, Li LC. Induction of NANOG expression by targeting promoter sequence with small activating RNA antagonizes retinoic acid-induced differentiation. Biochem J. 2012; 443:821–828. doi: 10.1042/ BJ20111491. PMID: 22339500. [PubMed: 22339500]

46. Selvey S, Thompson EW, Matthaei K, Lea RA, Irving MG, et al. Beta-actin--an unsuitable internal control for RT-PCR. Mol Cell Probes. 2001; 15:307–311. doi: 10.1006/mcpr.2001.0376. PMID: 11735303. [PubMed: 11735303]
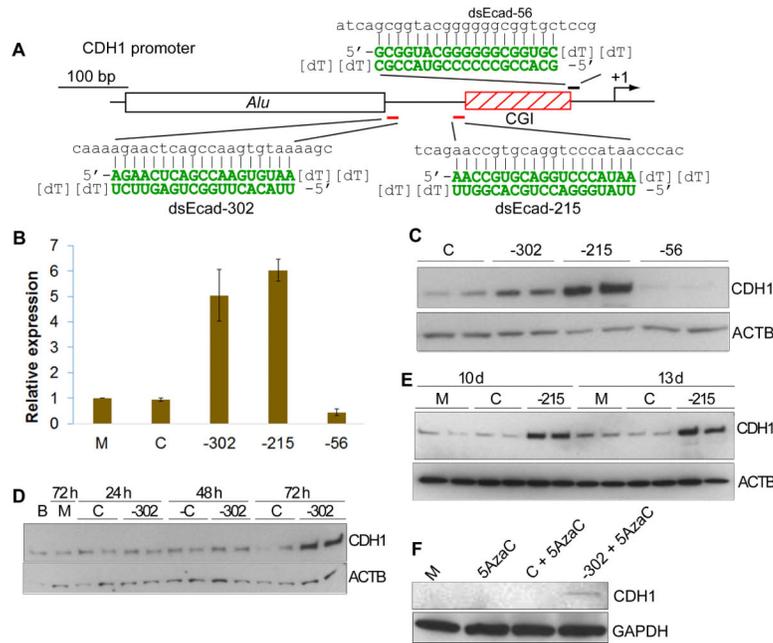
**Figure 1. Activation of the CDH1 gene by RNAa and its related kinetics**
**A.** Schematic representation of the CDH1 promoter. Indicated are its CpG island and Alu repetitive sequence. The CDH1 TSS (+1) is denoted by a bent arrow. Two saRNAs (marked in red) targeting sequences −302 (dsEcad-302) and −215 (dsEcad-215) relative to the TSS were generated according to design rules, whereas dsEcad-56 (marked in black) deviated from our design criteria. The sequence of each duplex is shown in green next to its derived target sequence in the CDH1 promoter. **B.** Induction of CDH1 expression by saRNAs. PC-3 cells were transfected at 50 nM of the indicated saRNAs for 72 h. mRNA expression of CDH1 in the transfected cells was assessed by quantitative RT-PCR. Results represent means ± SD of two independent transfections. **C.** Regional effect of RNAa on the CDH1 promoter. PC-3 cells were transfected as in (B). Protein levels of CDH1 and ACTB were analyzed by immunoblot analysis. ACTB served as a loading control. **D.** Delay in the onset of CDH1 induction by saRNA. PC-3 cells were transfected as in (B). Protein levels of CDH1 and ACTB were evaluated at 24, 48, and 72 h by immunoblot analysis. **E.** Prolonged activation of CDH1 by RNAa. PC-3 cells were transfected as in (B), split after ~4–5 days of culture and harvested on days 10 and 13 for analysis. CDH1 and ACTB levels were detected by immunoblot analysis. **F.** DNA demethylation restores RNAa. HeLa cells were treated at 1 μM 5-aza-2′-deoxycytidine (5-AzaC) with or without 50 nM of the indicated saRNAs for 72 h. Protein was isolated from the treated cells and subject to immunoblot analysis using a CDH1 or GAPDH specific antibody. Blank (B) denotes untreated cells. Mock (M) samples were transfected in the absence of saRNA. Control (C) treatments utilized a non-specific RNA duplex.

**A.** -1000 bp    Repeat element    -100 bp    +1    CGI

**B.**    1    19
agtg**CCAACTCATTCTCCAAGTA**<u>ataa</u>    DNA-S

**C.**    5' CCAACUCAUUCUCCAAGUA[dT][dT]    RNA-S
|||||||||||||||||||
[dT][dT]GGUUGAGUAAGAGGUUCAU 5'    RNA-AS
-2.34    -1.02    ΔG°37(kcal/mol)

**Figure 2. Guidelines for saRNA design on gene promoters**
**A.** A hypothetical promoter with a CpG island (CGI, hatched box) and a repeat element (grey box) are shown. Five saRNAs (red bars) are designed within a region ranging from −100 to −1000 bp (green zone) relative to the TSS. The repeat element and CpG island were avoided as targetable sequence. **B.** An exemplary target DNA site (sense, DNA-S) consisting of 4 flanking nucleotides and 19-nt target site complementary to its cognate saRNA. Underlined is an A/T-rich sequence immediately flanking the 19-nt target site. **C.** The resulted 19-nt saRNA duplex with 3' terminal [dT][dT] overhangs added. The duplex has lower thermodynamic stability ($\Delta$G°$_{37}$) at the 5' end of the antisense RNA (RNA-AS) strand. The terminal free energy ($\Delta$G°$_{37}$) values of the first 4 base pairs at either end of the duplex were calculated by an expanded nearest-neighbor model for RNA duplexes with Watson-Crick base pairing as previously described in Xia et al. [31].
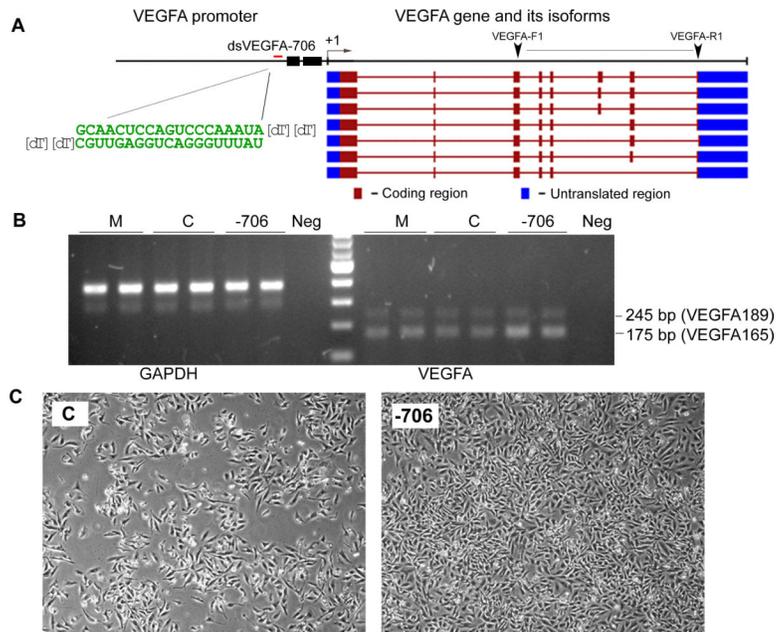
**Figure 3. RNAa of the VEGFA gene**
**A.** Schematic presentation of the VEGFA gene promoter and its various transcribed isoforms. VEGFA TSS (+1) is denoted by the bent arrow. The black boxes denote two CpG islands located within its promoter. A VEGFA saRNA (dsVEGFA-706) was designed to target a site (red bar) located at −706 relative to the TSS. RT-PCR primers (VEGFA-F1 and VEGFA-R1) were carefully designed to differentially amplify two major isoforms of VEGFA (VEGFA189 and VEGFA165). **B.** HeLa cells were transfected with the indicated dsRNAs at 50 nM/L for 72 h. Mock (M) samples were transfected in the absence of saRNA. Control (C) treatments utilized a non-specific RNA duplex. Expression levels of VEGFA and GAPDH were determine by PCR and visualized on an agarose gel stained with ethidium bromide. Amplification using only $H_2O$ as template material for PCR reactions served as a negative (Neg) control. Both major VEGFA isoforms amplified by VEGFA-F1 and VEGFA-R1 are indicated. **C.** Representative contrast phase images of HeLa cells taken at 72 h following transfection with control (C) dsRNA or dsVEGF-706. Note dsVEGFA-706 treated cells appeared denser than the control treatment indicative of enhanced proliferation.

**Table 1**

saRNA design rules.

| Rule name | Criterion |
|---|---|
| Distance from TSS | −100 - −1000 |
| CpG islands | avoid |
| Target size | 19 nt |
| Target GC content | 40-60% |
| Consecutive nucleotides | <= 5 |
| Simple repeats (di or tri-nucleotide) | avoid |
| CpGs in target | <= 1 |
| 1st and 2nd nucleotides | preferably G or C |
| 18th nucleotide | A or T |
| 19th nucleotide | A |
| 20-23th nucleotides | preferably A or T |

**Table 2**

Troubleshooting table.

| Step | Problem | Possible reason | Solution |
|---|---|---|---|
| 14-53 | Low or no RNAa activity | Transfection efficiency is too low | • Use a fluorescent-labeled dsRNA to determine transfection efficiency.<br>• Transfect a standard siRNA and detect RNAi activity.<br>• Choose an appropriate transfection agent for a particular cell type.<br>• Use a reverse transfection protocol.<br>• Transfect exponential growing cells.<br>• Disperse cells into single cell solution before transfection.<br>• Use Nucleofection™ (Lonza) for primary cells. |
| | | Transfection duration is too short. | • The optimal window of assessing gene activation is between 3-5 days. Do not perform analyses earlier than 48 h. |
| | | Use of a suboptimal number of cells. | • Transfect cells at a density that at harvest, no treatments should reach 100% confluency. Low cell density has negative impact on cell growth. |
| | | RNA degradation | • Check RNA quality by denaturing or non-denaturing agarose gel electrophoresis before RT reaction. If RNA is degraded, make every effort to avoid RNase contamination when working with RNA. |
| | | Inappropriate target design | • Redesign saRNA targets. |
| | | Gene silencing by epigenetic mechanisms especially DNA methylation | • Use subtoxic dose of epigenetic reagents such as 5-AzaC and VPA in combination with saRNA transfection. |
| 13-53 | RNAa is not reproducible | dsRNA degradation | • Dissolve dsRNA in DEPC-treated water, make small aliquots and store them at −80°C. Avoid multiple freeze-thaw cycles. |
| | | Different passages of primary cells or cell lines are used | • Use primary cells from similar early passages. Revive earlier passages of cells when cell lines have been cultured for a long period of time. |
| 49-53 | RNAa cannot be confirmed by protein expression analysis | Use un-validated primary antibodies for target protein detection. | • Determine specificity of un-validated primary antibodies by overexpressing the target gene using a cDNA expressing vector or knocking down the target gene with its specific siRNA. |
| | | Protein degradation | • Use protease inhibitors during protein isolation. |
| | | Primary antibody does not recognize the specific isoform induced by promoter-targeted saRNAs. | • Choose an antibody that recognizes the major isoform of protein encoded by the activated mRNA. |