Available online at www.sciencedirect.com
ScienceDirect
Antiviral Research 79 (2008) 37–48
Antiviral Research
www.elsevier.com/locate/antiviral

Structure and sequence motifs of siRNA linked with *in vitro* down-regulation of morbillivirus gene expression

Renata Servan de Almeida *, Djénéba Keita, Geneviève Libeau, Emmanuel Albina

CIRAD, UPR Contrôle des Maladies, Montpellier F-34398, France

Received 22 March 2007; accepted 23 January 2008

Abstract

The most challenging task in RNA interference is the design of active small interfering RNA (siRNA) sequences. Numerous strategies have been published to select siRNA. They have proved effective in some applications but have failed in many others. Nonetheless, all existing guidelines have been devised to select effective siRNAs targeting human or murine genes. They may not be appropriate to select functional sequences that target genes from other organisms like viruses. In this study, we have analyzed 62 siRNA duplexes of 19 bases targeting three genes of three morbilliviruses. In those duplexes, we have checked which features are associated with siRNA functionality. Our results suggest that the intramolecular secondary structure of the targeted mRNA contributes to siRNA efficiency. We also confirm that the presence of at least the sequence motifs U13, A or U19, as well as the absence of G13, cooperate to increase siRNA knockdown rates. Additionally, we observe that G11 is linked with siRNA efficacy. We believe that an algorithm based on these findings may help in the selection of functional siRNA sequences directed against viral genes.

Keywords: siRNA; RNAi; Morbillivirus; Sequence motifs; Secondary structure; Target sequence selection

1. Introduction

Small interfering RNAs (siRNAs) can achieve post-transcriptional gene silencing by degradation of homologous mRNA through a process known as RNA interference (RNAi) (Fire et al., 1998; Hammond et al., 2001; Hannon, 2002). RNAi pathway is triggered, in an initiation phase, by longer double-stranded precursor RNA (dsRNA), which is cleaved in fragments of 21–25 nt in length by the RNAs III-like protein DICER. The cleavage products are subsequently incorporated into the RNA-induced silencing complex (RISC) (Hammond et al., 2000; Bernstein et al., 2001). This RISC, which consists of several proteins and a single strand siRNA, recognizes, combines, and degrades the target mRNA (Zamore et al., 2000; Martinez et al., 2002). siRNAs are naturally produced when transposons or endogenous genes form long dsRNAs in vegetal or animal cells (Fire et al., 1998; Zamore et al., 2000; Hammond et al., 2001). RNAi is a natural antiviral defense mechanism in plants (Zamore et al., 2000). Likewise, it may also occur in animal cells in defense against viral infections. In mammalian cells, delivery of chemically synthesized siRNAs into the cytoplasm can silence the expression of a desired gene, circumventing the non-specific neutralization of mRNA by interferon response (Elbashir et al., 2001).

Recent studies have shown that the most critical point in applying RNAi is the identification of potent siRNAs and that their functionality is affected by the duplex nucleotide composition and the accessibility of target RNA (Elbashir et al., 2001; Amarzguioui and Prydz, 2004; Gilmore et al., 2004; Khvorova et al., 2003; Luo and Chang, 2004; Reynolds et al., 2004; Ui-Tei et al., 2004; Yoshinari et al., 2004; Heale et al., 2005; Huesken et al., 2005; Overhoff et al., 2005; Schubert et al., 2005; Westerhout et al., 2005). This limitation represents an important drawback since experiments to verify whether a given siRNA is effective are both time-consuming and expensive. To solve this problem, various guidelines or algorithms to optimize the design of effective siRNAs have been published (Amarzguioui and Prydz, 2004; Ding et al., 2004; Levenkova et al., 2004; Reynolds et al., 2004; Ui-Tei et al., 2004; Wang and Mu, 2004; Arziman et al., 2005; Patzel et al., 2005; Yiu et al., 2005). These strategies have proved effective under certain circumstances, but they have failed in many others. In the case of viruses, to our knowledge,

* Corresponding author. Tel.: +33 4 67593724; fax: +33 4 67593798.
E-mail addresses: renata.almeida@cirad.fr (R.S. de Almeida), djeneba.keita@cirad.fr (D. Keita), genevieve.libeau@cirad.fr (G. Libeau), emmanuel.albina@cirad.fr (E. Albina).

0166-3542/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.antiviral.2008.01.159

there is no comprehensive procedure for the selection of active interfering RNA molecules homologous to virus genes which has been validated by extensive in vitro biological assays. Naito et al. (2006) have developed a computational tool to design optimal antiviral siRNA sequences. However, their software incorporates guidelines previously devised for sequences targeting mammalian genes. Although there are probably no inherent differences between viral and non-viral mRNA sequences, it could be interesting to evaluate whether the already published rules for the inactivation of mammalian genes are also relevant for viral genes.

Here, we carried out an analysis on 62 siRNA duplexes directed against three genes of three different morbilliviruses. We evaluated the participation of various published features previously reported necessary for siRNA functionality, defined here as a minimum reduction of 70% of protein expression. We have evaluated several published algorithms as siVirus (http://sivirus.rnai.jp), E-RNAi (http://www.dkfz.de/signaling2/e-rnai/), SDS-siRNA Design Software (http://i.cs.hku.hk/~sirna/software/sirna.php), siRNA Selection Program (http://jura.wi.mit.edu/bioc/siRNAext/siRNA_search.cgi?tasto=680346486), Sfold (http://sfold.wadsworth.org/sirna.pl), Deqor (http://cluster-1.mpi-cbg.de/Deqor/deqor.html), siSeach–siRNA design (http://sonnhammer.cgb.ki.se/siSearch/siSearch_1.7.html), and guidelines (Schwarz et al., 2003; Amarzguioui and Prydz, 2004; Ding et al., 2004; Henschel et al., 2004; Pekarik, 2005; Reynolds et al., 2004; Ui-Tei et al., 2004; Yoshinari et al., 2004; Huesken et al., 2005; Jagla et al., 2005; Schubert et al., 2005; Shabalina et al., 2006). However, none of these algorithms were able to find the functional sequences. Moreover, we have detected that several criteria previously recommended in these guidelines are not needed for siRNA functionality and other ones, which had not been reported before, and probably contribute to the siRNA knockdown rates. The important features related with functionality were specific sequence determinants and secondary of the mRNA target.

2. Materials and methods

2.1. siRNA design

Sixty-two 19-bases long siRNAs were designed in the laboratory against three morbilliviruses, the human measles virus (MV), the bovine rinderpest virus (RPV), and the small ruminant "peste des petits ruminants" virus (PPRV). The design was made by using private software (Cenix algorithm, Ambion) for 26 molecules or in-house criteria for the other 36. Among these siRNAs, 20 were specific for the PPRV nucleoprotein (N) gene of PPRV, 12 specific for MV-N gene, 14 specific for the RPV-N gene, 11 specific for the PPRV matrix (M) gene, and 5 specific for the RPV fusion (F) gene. All siRNA sequences and their position in corresponding genes are listed in Table 1 ordered by gene target. They were all synthesized and purified by Ambion. Twenty-eight of these siRNAs were variations of functional siRNAs defined by displacement in 3′ and 5′ directions (Servan de Almeida et al., 2007; Keita et al., submitted). For the design of siRNA, we first selected conserved regions of the genome on different multiple alignments of morbilliviruses, including 208 MV, 11 RPV, 6 PPRV, 51 canine distemper viruses and 3 phocine distemper viruses. The conservation was assessed on successive stretches of 20 nucleotides covering the full sequence. In the multiple alignment of 208 N gene of MV, the minimum and maximum percentages of identities were 86% and 100%, respectively. In the rest of the study, the selection of conserved segments for further siRNA design was based on a minimum conservation of 93%. siRNA sequences were finally derived from conserved regions of the following strains MV (GenBank/EMBL accession numbers U03668), RPV (Z30697), PPRV (X74443). The 36 siRNA were designed without particular attention to their sequences, but some sequences were rejected because they lacked necessary characteristics defined by Amarzguioui and Prydz (2004), Reynolds et al. (2004), and Ui-Tei et al. (2004). We therefore considered features like GC content of the duplex ranging between 32% and 58%, low internal stability of the 3′ terminal of the sense strand, and some desired bases in strategic positions, as S1 and W19. We limited our study to the 19mer duplex region of the siRNA and excluded the 3′ overhangs. However, more than 75% of sequences have dTdT 3′ overhang in both strands. Irrelevant GAPDH siRNA (Ambion, Silencer^{TM} GAPDH siRNA Human Control, sense: 5′-AAGGUCAUCCAUGACAACUtt-3′, antisense: 5′-AGUUGUCAUGGAUGACCUUtt-3′), cells that have received lipofectamine without siRNA and after were infected (lipo + siRNA-virus + cells) and cells that have received siRNA without lipofectamine and were subsequently infected (lipo-siRNA + virus + cells) were used as controls.

2.2. Cell culture and transfection

Vero cells (ECACC 84113001) with 48 h of growth were trypsinized and plated at 1×10^5 cells per well in 24-well plates. Twenty-four hours later, different final concentrations of siRNAs (100, 50, 25 and 12,5 nM) were complexed with 2 μg per well of Lipofectamine 2000 (Invitrogen) in 100 μl of OptiMEM I (Invitrogen) serum-free. After 20 min of incubation at room temperature, the siRNA/Lipofectamine 2000 complexes were added to wells containing 200 μl of MEM serum-free. Plates were incubated for 6 h at 37 °C in an atmosphere containing 5% of CO_2. Afterwards, the cell supernatant was replaced with 1 ml of MEM supplemented with 5% of fetal bovine serum (FBS). Twenty-four hours after transfection, cells were infected with PPRV, RPV, or MV, using a multiplicity of infection of 0.1. Four days later, the siRNA silencing effect was evaluated by scoring the reduction of cytopathic effect (CPE) and by flow cytometry quantification of the decrease in virus protein expression. Reductions of 50% and 70%, respectively for CPE and protein expression, were considered as the cut-off for functionality.

2.3. Immunostaining and flow cytometry

For flow cytometry analysis, viral proteins in infected cells were labeled using specific monoclonal antibodies (Mabs) against N protein of PPRV and RPV, M protein of PPRV, F

Table 1
Targeted positions and siRNA sequences

	siRNA	Target gene	Position (mRNA)
NPPRV1	5'- GGAUCAACUGGUUUGAGAAtt -3' 3'- ttCCUAGUUGACCAAACUCUU -5'	N PPRV	480-498
NPPRV1-1	5'- AGGAUCAACUGGUUUGAGAtt -3' 3'- ttUCCUAGUUGACCAAACUCU -5'	N PPRV	479-497
NPPRV1-2	5'- AAGGAUCAACUGGUUUGAGtt-3' 3'- ttUUCCUAGUUGACCAAACUC -5'	N PPRV	478-496
NPPRV1-3	5'- AAAGGAUCAACUGGUUUGAtt -3' 3'- ctUUUCCUAGUUGACCAAACU -5'	N PPRV	477-495
NPPRV1+1	5'- GAUCAACUGGUUUGAGAACtt -3' 3'- tcCUAGUUGACCAAACUCUUG -5'	N PPRV	481-499
NPPRV1+2	5'- AUCAACUGGUUUGAGAACAtt -3' 3'-ccUAGUUGACCAAACUCUUGU –5'	N PPRV	482-500
NPPRV5	5'- GAGAACUCAAUUCAGAACAtt -3' 3'- ttCUCUUGAGUUAAGUCUUGU -5'	N PPRV	1001-1019
NPPRV6	5'- GGCGGUUCAUGGUAUCUCUtt -3' 3'- ttCCGCCAAGUACCAUAGAGA -5'	N PPRV	741-759
NPPRV6-1	5'- CGGCGGUUCAUGGUAUCUCtt -3' 3'- ttGCCGCCAAGUACCAUAGAG -5'	N PPRV	740-758
NPPRV6-2	5'- UCGGCGGUUCAUGGUAUCUtt -3' 3'- ttAGCCGCCAAGUACCAUAGA -5'	N PPRV	739-757
NPPRV6+1	5'- GCGGUUCAUGGUAUCUCUCtt -3' 3'- ttCGCCAAGUACCAUAGAGAG -5'	N PPRV	742-760
NPPRV6+2	5'- CGGUUCAUGGUAUCUCUCAtt -3' 3'- ttGCCAAGUACCAUAGAGAGU -5'	N PPRV	743-761
NPPRV7	5'- GCAUUAGGCCUUCACGAGUtt -3' 3'- ttCGUAAUCCGGAAGUGCUCA -5'	N PPRV	899-917
NPPRV7-1	5'- UGCAUUAGGCCUUCACGAGtt -3' 3'- ttACGUAAUCCGGAAGUGCUC -5'	N PPRV	898-916
NPPRV7-2	5'- CUGCAUUAGGCCUUCACGAtt -3' 3'- ttGACGUAAUCCGGAAGUGCU -5'	N PPRV	897-915
NPPRV7+1	5'- CAUUAGGCCUUCACGAGUUtt -3' 3'- ttGUAAUCCGGAAGUGCUCAA -5'	N PPRV	900-918
NPPRV7+2	5'- AUUAGGCCUUCACGAGUUCtt -3' 3'- ttUAAUCCGGAAGUGCUCAAG -5'	N PPRV	901-919
NPPRV8	5'- GUAUCAACAGCUAGGAGAGtt -3' 3'- ttCAUAGUUGUCGAUCCUCUC -5'	N PPRV	958-976
NPPRV9	5'- GAACUUUGGCAGGUCAUAUtt -3' 3'- ttCUUGAAACCGUCCAGUAUA -5'	N PPRV	1102-1120
NPPRV10	5'- CGCCAGUUUCAUUCUUACUtt -3' 3'- ttGCGGUCAAAGUAAGAAUGA -5'	N PPRV	850-868
NRPV1	5'- AGUCUUACUGGUUUGAGAAtt -3' 3'- ttUCAGAAUGACCAAACUCUU -5'	N RPV	480-498
NRPV1-1	5'- CAGUCUUACUGGUUUGAGAtt -3' 3'- ttGUCAGAAUGACCAAACUCU -5'	N RPV	479-497
NRPV1-2	5'- GCAGUCUUACUGGUUUGAGtt -3' 3'- ttCGUCAGAAUGACCAAACUC -5'	N RPV	478-496
NRPV1+1	5'- GUCUUACUGGUUUGAGAAUtt-3' 3'- ttCAGAAUGACCAAACUCUUA -5'	N RPV	481-499
NRPV6	5'- GCAGAUUUAUGGUGGCAUUtt -3' 3'- ttCGUCUAAAUACCACCGUAA -5'	N RPV	741-759
NRPV6-1	5'- CGCAGAUUUAUGGUGGCAUtt -3' 3'- ttGCGUCUAAAUACCACCGUA -5'	N RPV	740-758
NRPV6-2	5'- ACGCAGAUUUAUGGUGGCAtt -3' 3'- ttUGCGUCUAAAUACCACCGU -5'	N RPV	739-757

Table 1 (*Continued*)

NRPV6+1	5'- CAGAUUUAUGGUGGCAUUGtt -3' 3'- ttGUCUAAAUACCACCGUAAC -5'	N RPV	742-760
NRPV6+2	5'- AGAUUUAUGGUGGCAUUGAtt -3' 3'- ttUCUAAAUACCACCGUAACU -5'	N RPV	743-761
NRPV7	5'- GCACUGGGCCUGCAUGAAUtt -3' 3'- ttCGUGACCCGGACGUACUUA -5'	N RPV	899-917
NRPV7-1	5'- AGCACUGGGCCUGCAUGAAtt -3' 3'- ttUCGUGACCCGGACGUACUU -5'	N RPV	898-916
NRPV7-2	5'- CAGCACUGGGCCUGCAUGAtt -3' 3'- ttGUCGUGACCCGGACGUACU -5'	N RPV	897-915
NRPV7+1	5'- CACUGGGCCUGCAUGAAUUtt -3' 3'- ttGUGACCCGGACGUACUUAA -5'	N RPV	900-918
NRPV7+2	5'- ACUGGGCCUGCAUGAAUUCtt -3' 3'- ttUGACCCGGACGUACUUAAG -5'	N RPV	901-919
NMMV1	5'- GGUUCGGAUGGUUCGGGAAtt -3' 3'- ttCCAAGCCUACCAAGCUCUU-5'	N MMV	535-553
NMMV1-1	5'- AGGUUCGGAUGGUUCGAGAtt -3' 3'- ttUCCAAGCCUACCAAGCUCU-5'	N MMV	534-552
NMMV1-2	5'- CAGGUUCGGAUGGUUCGAGtt -3' 3'- ttGUCCAAGCCUACCAAGCUC -5'	N MMV	533-551
NMMV1+1	5'- GUUCGGAUGGUUCGAGAACtt -3' 3'- ttCAAGCCUACCAAGCUCUUG -5'	N MMV	536-554
NMMV6	5'- GCCGAUUCAUGGUCGCUCUtt -3' 3'- ugCGGCUAAGUACCAGCGAGA -5'	N MMV	796-814
NMMV6-1	5'- CGCCGAUUCAUGGUCGCUCtt -3' 3'- ugGCGGCUAAGUACCAGCGAG -5'	N MMV	795-813
NMMV7	5'- GCUCUUGGACUGCAUGAAUtt-3' 3'- gaCGAGAACCUGACGUACUUA -5'	N MMV	954-972
NMMV7-2	5'- CUGCUCUUGGACUGCAUGAtt-3' 3'- gaGACGAGAACCUGACGUACU -5'	N MMV	952-970
NMMV7-3	5'- CCUGCUCUUGGACUGCAUGtt-3' 3'- gaGGACGAGAACCUGACGUAC -5'	N MMV	951-969
N1MMV	5'- GGAGCUUAGCAUUGUUCAAtt -3' 3'- ttCCUCGAAUCGUAACAAGUU -5'	N MMV	17-35
N2MMV	5'- CCGGUUGGUGAGGUUAAUUtt -3' 3'- ctGGCCAACCACUCCAAUUAA -5'	N MMV	159-177
N3MMV	5'- GCUGAUUCGGAGCUAAGAAtt -3' 3'- gtCGACUAAGCCUCGAUUCUU -5'	N MMV	565-583
MPPRV1	5'- AUACUCUGCUGAUUAUUGCaa -3' 3'- caUAUGAGACGACUAAUAACG -5'	M PPRV	725-743
MPPRV2	5'- GCUGAUUAUUGCAAAAUGAag -3' 3'- gaCGACUAAUAACGUUUUACU -5'	M PPRV	732-750
MPPRV4	5'- GCCAUCAGUACCCCAAGAAtt -3' 3'- guCGGUAGUCAUGGGGUUCUU -5'	M PPRV	862-880
MPPRV5	5'- GGCGCACGGCAGGACUAAAtt-3 3'- ttCCGCGUGCCGUCCUGAUUU -5'	M PPRV	337-355
MPPRV6	5'- GGCCAGAAGAACUACUAAGtt -3' 3'- ttCCGGUCUUCUUGAUGAUUC -5'	M PPRV	289-307
MPPRV8	5'- GAAGCACAGGGAAAAUGAGca -3' 3'- agCUUCGUGUCCCUUUUACUC -5'	M PPRV	814-832
MPPRV9	5'- GAAGAUUGAAAAGAUGGGUtt -3' 3'- ttCUUCUAACUUUUCUACCCA -5'	M PPRV	749-767
MPPRV11	5'- GGUAUUUUACAACAACACUtt-3' 3'- ttCCAUAAAAUGUUGUUGUGA -5 '	M PPRV	366-384
MPPRV 75/1 1	5'- CAAGAUGACCGAGAUCUACtt -3' 3'- ctGUUCUACUGGCUCUAGAUG -5'	M PPRV	29-47

Table 1 (*Continued*)

MPPRV 75/1 2	5'- UGUCAAAGGGUCAAUUGCUtt -3' 3'- ctACAGUUUCCCAGUUAACGA -5'	M PPRV	71-89
MPPRV 75/1 3	5'- GGAGUGAUUGAGGAUAACGtt -3' 3'- atCCUCACUAACUCCUAUUGC -5'	M PPRV	201-219
FPPRV1	5'- GAUACUGAAGACUGUUAAAtt -3' 3'- gtCUAUGACUUCUGACAAUUU -5'	F PPRV	2067-2085
FPPRV2	5'- CCUUUCAGUGGCAAUAUAUtt -3' 3' – atGGAAAGUCACCGUUAUAUA -5'	F PPRV	2092-2110
FPPRV3	5'- CGUGAGAUCACUGUAGUUAtt -3' 3' - atGCACUCUAGUGACAUCAAU -5'	F PPRV	2259-2277
FPPRV4	5'- UCCCUAGGGCUUGUCACAUtt -3' 3' - atAGGGAUCCCGAACAGUGUA -5'	F PPRV	2140-2158
FPPRV5	5'- ACCUUAUCUGAGAUCAAGGtt -3' 3' - gtUGGAAUAGACUCUAGUUCC -5'	F PPRV	1480-1498

The functional sequences are in green.

protein of RPV, and FITC antibody conjugated as secondary antibody. In brief, cells were trypsinized and plated on 96-well plates. Monoclonal antibodies specific for each of the viral proteins were appropriately diluted and added to each well. After 30 min of incubation at 4 °C and two washing steps, cells were incubated for 30 min at 4 °C with a second FITC-conjugated antibody (Bio Rad). Finally, cells were washed twice and fixed with 1% paraformaldehyde. All dilutions and washings were carried out using a PBS solution containing 0.025–0.015% saponine for cell permeabilization and intracellular staining. The silencing of the viral protein production by siRNAs was measured by flow cytometry using a FacSort (Becton Dickinson).

2.4. Validation of positively correlated features with siRNA functionality

From our siRNA efficacy results and based on previously published data (Bohula et al., 2003; Luo and Chang, 2004; Yoshinari et al., 2004; Heale et al., 2005; Overhoff et al., 2005; Schramm and Ramey, 2005; Schubert et al., 2005; Westerhout et al., 2005; Yiu et al., 2005) we analyzed the secondary structure of siRNA targets and the intrinsic sequence of siRNAs to define a procedure that would optimize the identification of effective siRNA. Secondary structures of mRNA targets were defined using the Mfold software (http://www.bioinfo.rpi.edu/applications/mfold/old/rna/) and all predicted structures were analyzed. Moreover, we applied a computational statistical approach proposed by Luo and Chang (2004) that is based on the average number of hydrogen bonds formed between nucleotides in the target region, a parameter called "the hydrogen bond (H–b) index". The H–b index scores duplex siRNAs according to their secondary structure and distinguishes between functional and non-functional molecules. In addition, we examined the accessibility of our siRNA sequences using the application STarMir described by Long et al. (2007). This application evaluates the potential effects of the secondary structure of a target mRNA on its sensitivity to miRNA regulation by energy calculations for miRNA-target hybridization and is available at http://sfold.wadsworth.org (for more details see Long et al., 2007). Last, the frequency of the 12 criteria compiled from the literature and related to siRNA activity, was determined on active versus non-active siRNAs. The list of these criteria and the respective positions on which they applied in the duplex sense strand are given in Table 2. Statistical analyses were done using the Chi-square test on discrete variables, testing the null hypothesis H_0 (no difference between the active and non-active siRNA populations) and using the R language (http://www.r-project.org/).

3. Results

3.1. Determination of siRNA functionality

To assess the efficacy of the sequences, we considered that a reduction of 70% of protein expression was the minimal requirement for functionality, provided that a minimum level of 80% of viral protein expression was obtained using cells treated with the irrelevant GAPDH siRNA, lipo + siRNA-virus + cells and lipo-siRNA + virus + cells controls. All these controls have shown no reduction of viral protein expression. In total, 12 out of 62 (19%) siRNA duplexes were functional according to this criterion. The summary of results for all siRNAs is presented in Table 3. We identified four siRNAs able to knock down the N gene of PPRV, five active on the N gene of RPV and three active on the N gene of MMV.

3.2. Secondary structure and siRNA functionality

Using the Mfold software, we obtained several predicted RNA folded structures for the same mRNA. A compilation of the secondary structure of mRNA targets of the siRNA set is shown in Table 4. We found that the hairpin structures and target sites located between two very close branches possibly contributed to the inefficacy of siRNAs. In contrast, our results do not support a relation of stem doubled-stranded targets, or H–b index, proposed by Luo and Chang (2004), with a decrease in siRNA activity. The method proposed by Long et al. (2007), which searches for potential target sites for a defined

Table 2
Criteria evaluated in the siRNA sequences

Criteria	Authors
1. At least four out of seven A/U residues in the 5′ terminal of the antisense strand* and four out of seven G/C residues in the 3′ terminal of the antisense strand	Ui-Tei et al. (2004)
2. Absence G, C stretches >4 nt	Ding et al. (2004), Henschel et al. (2004)
3. S1	Schwarz et al. (2003), Amarzguioui and Prydz (2004), Jagla et al. (2005)
4. A3	Reynolds et al. (2004), Schubert et al. (2005), Pekarik (2005)
5. A6	Amarzguioui and Prydz (2004)
6. U10	Amarzguioui and Prydz (2004), Reynolds et al. (2004), Yoshinari et al. (2004), Pekarik (2005), Shabalina et al. (2006)
7. W19	Amarzguioui and Prydz (2004), Reynolds et al. (2004), Huesken et al. (2005), Pekarik (2005), Shabalina et al. (2006)
8. U13	Pekarik (2005), Huesken et al. (2005)
9. Absence of G13	Arziman et al. (2005) Pekarik (2005)
10. S11	Jagla et al. (2005)
11. G16	Hsieh et al. (2004)
12. U18	Huesken et al. (2005)
13. Evaluation of secondary structure of mRNA target	Bohula et al. (2003), Luo and Chang (2004), Yoshinari et al. (2004), Heale et al. (2005), Overhoff et al. (2005), Schramm and Ramey (2005), Schubert et al. (2005), Westerhout et al. (2005), Yiu et al. (2005)

mi(si)RNA sequence according to its secondary structure, was able to select 4 out of 12 (∼33%) of our functional sequences and 19 out of 50 (38%) of our non-functional sequences. Nonetheless, this method has identified a probable off-target site for the functional NPPRV7 siRNA which opens the possibility that this set of siRNAs interacts and regulates the mRNA expression in an alternative way. At the end, 100% of our functional siRNAs targeted mRNA positions which were accessible in more than 60% of the predicted secondary RNA structures. In contrast, 68% of non-functional siRNAs targeted positions with less than 60% of accessible secondary mRNA structures. The frequency of the accessible secondary structures in active siRNAs was significantly higher ($p \leq 0.05$) than in non-active siRNAs.

3.3. Sequence stability and siRNA functionality

Low internal stability of the 3′ terminal may be used to enhance RNAi activity (Khvorova et al., 2003; Schwarz et al., 2003). As a measure of stability asymmetry, we considered a minimum of four out of seven A/U residues in the 3′ terminal of the sense strand, according to Ui-Tei et al. (2004) and a minimum of four out of seven G/C residues in the 5′ terminal of the sense strand. We also estimated the stability asymmetry of the first four nucleotide pairs of the 5′ end duplex by calculating the thermodynamic energy and considering a minimum energy difference of −0.3 to guarantee this asymmetry, as proposed by Hutvagner (2005).

Nine out of our twelve functional siRNAs (75%) had a high incidence of A/U at the 3′ end of the sense whereas 38 out of 50 (76%) non-functional siRNAs also showed this criterion. Moreover, only 50% of functional siRNAs showed four out of seven G/C residues in the sense 5′ end, whereas 54% of non-functional siRNAs followed this rule. Finally, we found a minimum value of −0.3 kcal/mol for the energy difference between the 5′ ends of the duplex in 8 out of 12 (67%) functional siRNAs. However, a higher percentage (76%) of the non-functional siRNA also showed this minimal difference.

3.4. Specific sequence motifs and siRNA functionality

We next investigated if some nucleotides in specific positions previously described as suitable for functional siRNA duplex were also associated with functional siRNAs in our study (Table 2). The incidence of these nucleotides was determined for functional (*f*) and non-functional (*nf*) siRNA groups. Results are given in Table 5. The most frequent motifs found in the set of functional siRNAs were (i) U10 ($f = 50\%$; $nf = 14\%$), (ii) S11 ($f = 75\%$; $nf = 40\%$) where S is G or C, (iii) U13 ($f = 75\%$; $nf = 26\%$), (iv) absence of G13 ($f = 100\%$; $nf = 58\%$), and (v) W19 ($f = 100\%$; $nf = 60\%$) where W is A or U. The occurrence of these motifs was significantly higher in active siRNAs ($p \leq 0.01$ except for S11 where $p \leq 0.05$). The motif A3 ($f = 42\%$; $nf = 22\%$) was moderately represented in the functional siRNAs and the less frequent motifs were (i) G16 ($f = 16\%$; $nf = 20\%$), (ii) S1 ($f = 75\%$; $nf = 72\%$), (iii) A6 ($f = 17\%$; $nf = 22\%$), and (iv) U18 ($f = 17\%$; $nf = 22\%$). However, the difference between active versus non-active siRNAs was not significant ($p \geq 0.05$). Additionally, the incidence of the four nucleotides was determined at each position, and the functional and non-functional sequences were compared (see Fig. 1). This analysis confirmed the strong dominance of motif U13 (differential incidence of 49%), the absence of G13 (−42%), U10 (36%) and A/U19 (40%) in our functional sequences. In contrast to other reports, motifs S1 (C = −15%; G = 18%), A3 (22%), A6 (−5%), and U18 (4%) were not dominant in the functional siRNAs in this study. However, a new important motif was identified. Indeed, motif G11 (differential incidence of 59%) and a strong statistical difference ($p \leq 0.0001$) was over-represented in functional siRNAs and under-represented in non-functional sequences. Addition-

Table 3
siRNA functionality measured by flow cytometry

siRNA	Percentages of functionality (Gene target)		
	N	M	F
PPRV			
NPPRV1	9–13% ($n \geq 10$)		
NPPRV1 − 1	19.5–22% ($n = 5$)		
NPPRV1 − 2	83–88% ($n = 7$)		
NPPRV1 − 3	89–93% ($n = 8$)		
NPPRV1 + 1	75–78% ($n = 6$)		
NPPRV1 + 2	68–74% ($n = 7$)		
NPPRV5	83.7–84.3% ($n = 5$)		
NPPRV6	19–21% ($n \geq 10$)		
NPPRV6 − 1	64.1–65.9% ($n = 2$)		
NPPRV6 − 2	85.5–90.4% ($n = 2$)		
NPPRV6 + 1	42.3–45.7% ($n = 10$)		
NPPRV6 + 2	57.6–64.5% ($n = 10$)		
NPPRV7	12.7–13.3% ($n \geq 10$)		
NPPRV7 − 1	54.6–57.3% ($n = 4$)		
NPPRV7 − 2	64.2–65.8% ($n = 2$)		
NPPRV7 + 1	35.3–38.7% ($n = 5$)		
NPPRV7 + 2	58–59.9% ($n = 3$)		
NPPRV8	65.3–66.7% ($n = 2$)		
NPPRV9	64.9–66% ($n = 5$)		
NPPRV10	36.5–40.4% ($n = 6$)		
MPPRV1		66.6–71.3% ($n = 2$)	
MPPRV2		66.5–68.5% ($n = 4$)	
MPPRV4		79–81% ($n = 4$)	
MPPRV5		89–93% ($n = 2$)	
MPPRV6		65.9–69% ($n = 2$)	
MPPRV8		54.3–57.7% ($n = 3$)	
MPPRV9		64.6–67.4% ($n = 3$)	
MPPRV11		68.9–75.1% ($n = 3$)	
MPPRV75/1 1		95.3–100% ($n = 2$)	
MPPRV 75/1 2		96.7–99.3% ($n = 2$)	
MPPRV 75/1 3		44–48% ($n = 10$)	
FPPRV1			72.8–79.2% ($n = 2$)
FPPRV2			82.9–91.2% ($n = 3$)
FPPRV3			93–100% ($n = 2$)
FPPRV4			94.5–100% ($n = 2$)
FPPRV5			99.5–100% ($n = 2$)
RPV			
NRPV1	25–30.2% ($n \geq 10$)		
NRPV1 − 1	2.3–4.2% ($n \geq 10$)		
NRPV1 − 2	65.5–71% ($n = 6$)		
NRPV1 + 1	54–60% ($n = 4$)		
NRPV6	14.4–15.6% ($n \geq 10$)		
NRPV6 − 1	97.7–100% ($n = 2$)		
NRPV6 − 2	87.8–90.1% ($n = 2$)		
NRPV6 + 1	63.1–64.9% ($n = 3$)		
NRPV6 + 2	58.5–59.5% ($n = 4$)		
NRPV7	23.5–24.5% ($n \geq 10$)		
NRPV7 − 1	61.7–66.2% ($n = 4$)		
NRPV7 − 2	81.8–82.2% ($n = 2$)		
NRPV7 + 1	14.7–15.2% ($n \geq 10$)		
NRPV7 + 2	50.2–51.8% ($n = 4$)		
MV			
NMMV1	9.5–11% ($n \geq 10$)		
NMMV1 − 1	13–18% ($n = 10$)		
NMMV1 − 2	62–66% ($n = 6$)		
NMMV1 + 1	67–75% ($n = 8$)		
NMMV6	45.4–48.5% ($n = 5$)		
NMMV6 − 1	75.4–78.6% ($n = 2$)		
NMMV7	45.9–48.1% ($n = 5$)		
NMMV7 − 2	65.5–72.4% ($n = 3$)		

Table 3 (*Continued*)

siRNA	Percentages of functionality (Gene target)		
	N	M	F
NMMV7 − 3	52.2–57.5% ($n = 2$)		
N1MMV	14.6–15.3% ($n = 2$)		
N2MMV	61.5–62% ($n = 2$)		
N3MMV	57.6–58.5% ($n = 2$)		
GAPDH	97.2–100% ($n \geq 10$)	98–99% ($n \geq 10$)	98–100% ($n \geq 10$)

The cut-off established for functionality was a reduction to a minimum of 70% of the N protein expression. The level of viral protein expression was measured using the data obtained by transfection of the irrelevant GAPDH siRNA (see Section 2), non-treated + infected cells and non-lipofectamine + infected cells, used as controls. These controls were considered when a minimum level of 80% of viral protein expression was obtained. The minimum and maximum level of viral protein expression and the number of replicates for each siRNA are shown for each siRNA.

ally, other motifs were moderately over-represented like G2 (30%), U4 (34%), G7 (26%), C9 (30%), G16 (30%) and G15 (26%) or under-represented like A5 (−26%), U8 (−26%), G9 (−30%) and A15 (−30%).

4. Discussion

Previously published works have correlated several specific motifs intrinsic to the siRNA sequences as well as particular secondary structure of the target site with efficient siRNA silencing (Elbashir et al., 2001; Amarzguioui and Prydz, 2004; Gilmore et al., 2004; Khvorova et al., 2003; Luo and Chang, 2004; Reynolds et al., 2004; Ui-Tei et al., 2004; Yoshinari et al., 2004; Heale et al., 2005; Huesken et al., 2005; Overhoff et al., 2005; Schubert et al., 2005; Westerhout et al., 2005). The application of these features has proven to be effective in certain cases but has failed to distinguish between functional and non-functional siRNAs in many others. To address the conditions that determine siRNA functionality, we analyzed a panel of 62 siRNA sequences targeting morbillivirus genes. The features associated with siRNA activity comprised mainly the secondary structure of target mRNA and at least five motifs strategically located in the 19 base-long siRNA sequences.

Target mRNA secondary structure is a criterion that cannot be easily defined nor confidently scored. Using the Mfold software, secondary structure can be obtained as sets of 10 predicted RNA folded structures for the same mRNA. It is difficult to know which of the probable structures represents the real or

Table 4
Secondary structure of mRNA targets

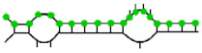
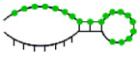
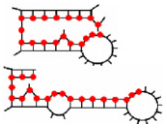
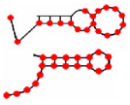
Structure	Description	References that agree with the results of this work
	Stem with lateral or asymmetric internal bulges and paired extremities	
	Stem with asymmetric internal bulges and unpaired 5′ end	
	Paired extremities	Luo and Chang (2004), Schramm and Ramey (2005), Schubert et al. (2005)
	Unpaired extremities	Schubert et al. (2005)
	Target site located between two branches separated by two bases and unpaired 3′ end	Schubert et al. (2005)
	Target sites located between two very close branches	Schubert et al. (2005), Yiu et al. (2005)
	Hairpin and unpaired 3′ end	Luo and Chang (2004), Schramm and Ramey (2005), Westerhout et al. (2005)
	Hairpin and unpaired extremities	Luo and Chang (2004), Schramm and Ramey (2005), Westerhout et al. (2005)

Table 5
The incidence of each criterion for the functional and non-functional siRNA groups

Criteria	Functional siRNAs	Non-functional siRNAs
1	8/12 (>66%)	24/50 (48%)
2	12/12 (100%)	47/50 (94%)
3	9/12 (75%)	36/50 (72%)
4	5/12 (>41%)	11/50 (22%)
5	7/12 (>58%)	30/50 (60%)
6	6/12 (50%)	7/50 (14%)
7	12/12 (100%)	30/50 (60%)
8	9/12 (75%)	13/50 (26%)
9	12/12 (100%)	29/50 (58%)
10	9/12 (75%)	20/50 (40%)
11	2/12 (>16%)	10/50 (20%)
12	11/12 (>91%)	32/50 (64%)

All incidence differences between functional and non-functional groups were statistically calculated by Chi-square test on discrete variables.

the most frequent folding into cell. Luo and Chang (2004) proposed to use a statistical approach to solve this limitation by introducing a single parameter (H–b index) to try to reflect the overall probability of the target siRNA sites to form doubled-stranded complexes, which are considered not accessible by several studies (Schubert et al., 2005; Heale et al., 2005). Nevertheless, our results do not support the lower activity of siRNA targeting doubled-stranded mRNA stems or siRNA sequences having a high H–b index (data not shown). In the same way, the Sfold algorithm (Ding et al., 2004) for RNA folding and prediction of target accessibility rejected all of our active siRNAs. On the other hand, we found that hairpin structures and target sites located between two very close branches may contribute to the inefficacy of siRNAs, as previously reported (Luo and Chang, 2004; Schramm and Ramey, 2005; Westerhout et al., 2005; Yiu et al., 2005). Moreover, 100% of the functional siRNAs, whereas 68% of the non-functional siRNAs targeted mRNA positions were defined as accessible in more than 60% of the predicted secondary RNA structures. Westerhout and Berkhout (2007) also found that hairpin structures in target sequences can obstruct the access of siRNA. However, these authors showed that, although the tightest RNA hairpins were completely resistant to RNAi, an inverse correlation between the overall target hairpin stability and RNAi-mediated inhibition existed. Although, the structure stability could determine the inefficacy of our siRNAs sequences, a threshold of paired bases that confer this stability has to be determined. Ideally, the correlation between the stability level of the target structure and RNAi efficiency has to be done by a mutational analysis of one target instead of comparing different siRNAs with intrinsically different RNAi efficacy. Consequently, this kind of evaluation was not possible with our set of data since we have analyzed siRNAs with different targets and nucleotide compositions.

For the selective incorporation of the antisense strand into RISC complex, the two strands of siRNA duplex should have a weak binding affinity at the 5′ end of the antisense strand (Khvorova et al., 2003; Schwarz et al., 2003). This weak affinity can be obtained by selecting target sequences rich in A/U content at the 5′ end of the antisense (Khvorova et al., 2003; Schwarz et al., 2003). Seventy-five percent of our functional siRNAs but also 76% of our non-functional siRNAs showed a minimum of four out of seven A/U residues at the 5′ end of the antisense. Although several studies suggested that the thermodynamic instability of the 5′ end of the antisense plays an important role in siRNA activity, our results show many exceptions to this rule, also found by Beale et al. (2003), Bohula et al. (2003), Ui-Tei et al. (2004) and Gilmore et al. (2004). Moreover, the requirement for efficacy of four out of seven G/C residues in the 5′ end of the sense was not confirmed in our siRNA set,

(a)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
A	5%	-9%	22%	-12%	-26%	-5%	1%	15%	5%	-20%	-18%	-18%	-8%	-1%	-30%	-13%	20%	2%	26%
U	-8%	-16%	-13%	34%	4%	4%	-7%	-26%	-5%	36%	-17%	10%	49%	18%	10%	-20%	-24%	-3%	14%
C	-15%	-5%	-9%	-11%	19%	-1%	-20%	15%	30%	1%	-24%	-12%	1%	-8%	-5%	3%	-4%	-12%	-18%
G	18%	30%	1%	-11%	3%	3%	26%	-5%	-30%	-17%	59%	20%	-42%	-9%	26%	30%	7%	13%	-22%

(b)

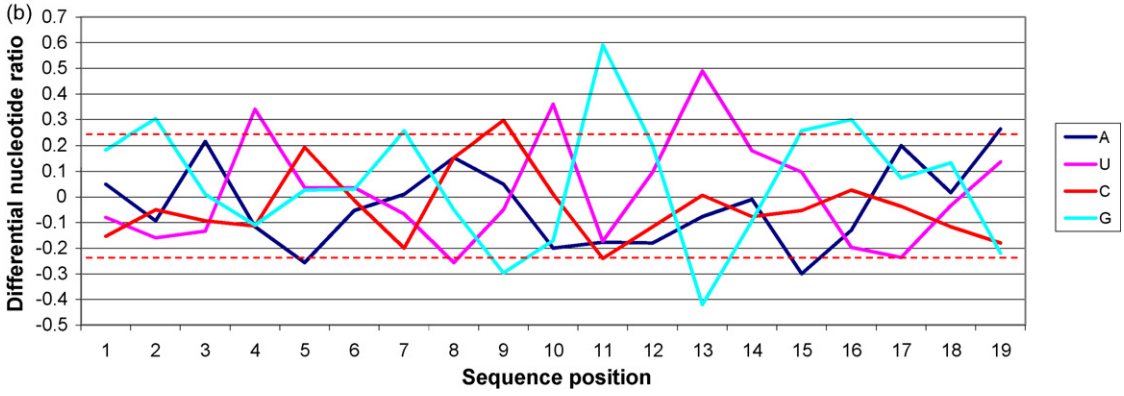


Fig. 1. Differential nucleotide incidence in functional vs. non-functional siRNAs. For each of the 19 positions, the relative percentages of functional siRNAs with the four nucleotides were determined. Percentage differences between functional and non-functional siRNAs are reported in table (a). Differences higher than 25% (positive cut-off) were considered over-representations of the corresponding nucleotides in functional siRNAs. Conversely, results lower than −25% (negative cut-off) were considered under-representations. Results >25% or <−25% are shown in gray boxes. In plot (b), the results are shown with curves, and cut-offs are materialized by horizontal dashed lines.

since only 50% of functional siRNAs showed this characteristic whereas 54% of non-functional siRNAs followed this rule. Hutvagner (2005) evaluated the stability asymmetry of the 5′ ends of the duplex by calculating thermodynamic energy and found that an energy difference of −0.3 kcal/mol was the minimum to guarantee the incorporation of the antisense strand into the RISC. We found a relatively high percentage (67%) of functional siRNAs satisfying this requirement. However, 76% of the non-functional siRNAs also fulfilled this criterion. Interestingly, we observed that 8 out of 12 (>66%) of functional siRNAs had simultaneously a higher A/U content at the 5′ end of the antisense strand and a lower A/U content at the 5′ end of the sense strand while 24 out of 50 (48%) of non-functional siRNAs did not. Although the difference was not significant ($p \geq 0.05$), we still believe that this parameter could probably be more relevant for siRNA activity than the strict requirement of a minimum number of A/U or G/C nucleotide pairs at the siRNA extremities (Ui-Tei et al., 2004) or a minimal value of thermodynamic energy (Hutvagner, 2005). Our prediction is also supported by Amarzguioui and Prydz (2004).

This work also shows that the parameter previously published that most contributes to the activity of our siRNAs was the presence of a uracil in position 13. Using this criterion alone we were able to increase substantially the likelihood of selecting a functional siRNA in our siRNA set because 75% of these siRNA have a U13 against 26% in non-functional siRNAs. We have evaluated the occurrence of G or C at position 11 suggested by Jagla et al. (2005) but we only confirmed that the G nucleotide is over-represented at position 11 of active siRNAs and under-represented in inactive siRNAs. Consequently, this motif, which had not been reported previously, deserves special attention. Our analyses also revealed that motifs U10, A and U at position 19 and absence of G13 are dominant, although less intensively, in functional siRNAs as reported by others (Amarzguioui and Prydz, 2004; Reynolds et al., 2004; Huesken et al., 2005; Pekarik, 2005; Shabalina et al., 2006). We can speculate that these motifs contribute to the overall performance of the siRNA directed against viral genes. Our findings also show that G/C at position 1, A at positions 3 and 6, and U at position 18 are not clearly linked with siRNA efficacy against the N gene of morbillivirus. Additionally, the results suggest that some other new motifs might play a role in functional siRNAs, such as the presence of G2, U4, G7, C9, G15 and G16 and absence of A5, U8, G9, and A15. However, the dominance of G2 in the functional sequences might have resulted from an artifact since at least 4 out of 12 sequences had been selected using this criterion to guarantee a higher stability at the 5′ end of the sense strand. In addition, G7, C9, G15 and G16, although they exceed the positive cut-off selected in this study, are not so distant from the other nucleotides, as visualized in Fig. 1. Therefore, the interest of these motifs as selection criteria for siRNA efficacy is presently not obvious. In contrast, U4 is clearly individualized from the other nucleotides. Its functional importance needs to be validated on a larger set of functional sequences.

To optimize the identification of effective siRNA, firstly the sequences were selected using as eliminatory criterion an mRNA accessibility of at least 60%. The requirement for G2, U10, G11, U13, absence of G13 and W19 and a stability asymmetry between the 5′ and 3′ termini of the sense strand were confronted to the selected ones using an inclusive cut-off score of any four out of these seven features. This strategy allowed the selection of >83% functional siRNAs while only 8% of non-functional siRNAs could not be rejected (Fig. 2).

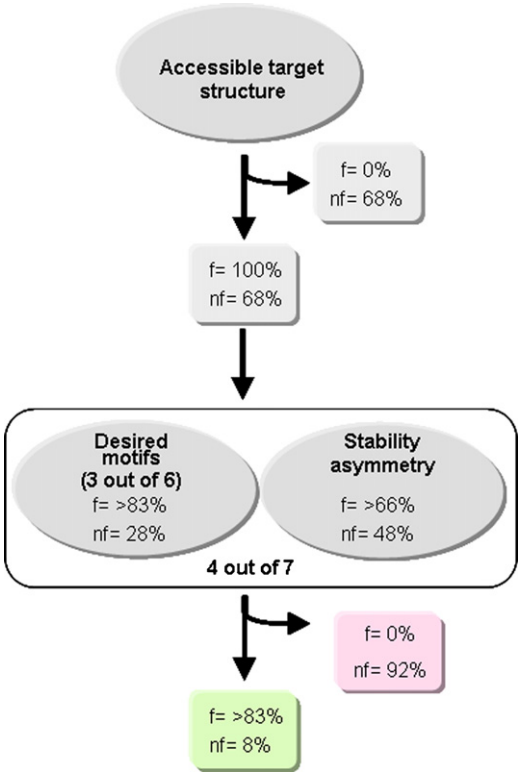


Fig. 2. Strategy for identification of effective siRNA. The procedure considers as potentially functional firstly the 19-nt siRNAs that target accessible sequences in at least 60% of the predicted mRNA secondary structures. Subsequently, the screened siRNAs are evaluated based on the stability asymmetry between the 5′ and 3′ termini of the sense strand (higher A/U content in the 3′ termini and a lower A/U content in the 5′ termini) and the desirable motifs (G2, U10, G11, U13, absence of G13 and W19). The siRNAs that fulfill four out of seven of these latter criteria are finally selected. The percentages of our sequences that showed these attributes are in the green area. Abbreviations: f, functional siRNA; *nf*, non-functional siRNA.

As discussed above, our results suggest that the intramolecular secondary structure of the targeted mRNA has an essential contribution to siRNA efficiency according to several previous publications (Bohula et al., 2003; Luo and Chang, 2004; Yoshinari et al., 2004; Heale et al., 2005; Overhoff et al., 2005; Schramm and Ramey, 2005; Schubert et al., 2005; Westerhout et al., 2005; Yiu et al., 2005; Köberle et al., 2006; Ameres et al., 2007; Long et al., 2007; Westerhout and Berkhout, 2007). The findings from Long et al. (2007) for example indicate a potent effect of target structure on target recognition by miRNA and based on a defined model of interaction between an miRNA and a target (for further details, see Long et al., 2007). They have developed the STarMir application which searches for potential target sites for a defined mi(si)RNA sequence according to its secondary structure. The STarMir application has found that only 4 out of our 12 functional siRNAs (∼33%) have an acces-

sible target against 19 out of our 50 non-functional sequences (38%). In the end, there is a higher number of our active siRNAs that match inaccessible targets as predicted by the Long's method compared to our method including eight criteria, leading to the elimination of more than 66% of them. Our method has, therefore, a better output for the identification of active siRNA. With this method, we are able to select 14 potentially active sequences of which 10 (>71%) have been confirmed to be active in vitro (our method did not select the two remaining functional siRNAs). Interestingly, when the Long's application is used after the pre-selection of active siRNAs according to our method, then a more restrictive selection is obtained with 4 out of 14 pre-selected siRNAs being retained and three of them (75%, NRPV1-1, NRPV6 and N1MMV) being functional in vitro. Two out of these three selected active sequences yield an inhibitory effect higher than 85%, the average degree of knockdown observed with all of our functional siRNAs (85%). By this combination of methods (first ours and second Long's method), the success rate for the selection of active siRNA is increased, thus reducing the cost for synthesis and *in vitro* testing, but this is in detriment of a high number of other functional ones. Therefore, we recommend using that combination if the purpose is to identify only few active sequences with a high probability of success and stronger activity. In the other cases, our method may offer more active sequences but with a higher number of non-active sequences. In addition, the Long's application can be secondarily applied for searching potential off-target sites of in vitro functional siRNA. Indeed, the Long's application has found that one of the active siRNA (NPPRV7) has a high probability of having an off-target effect. Interestingly, this sequence is one of our two functional sequences that were not predicted as active using our method because it has not the minimal number of criteria required. Therefore, to check if the activity of the siRNA NPPRV7 results from an off-target site hybridization, mutational studies in the off-target and/or target sites of the corresponding mRNA are required.

Further validation on other functional siRNA sequences possibly by other groups is needed, but we already believe that our selection strategy might be useful to increase the probability of selecting functional siRNA sequences against viral gene and could serve in the development of an algorithm for computerized design of siRNAs. Although generated on siRNA targeting virus genes, we anticipate that these rules would also be applicable to siRNA against mammalian mRNA sequences since both uses the same conserved cell machinery for RNA interference.

5. Conclusions

In this study, some specific related features of siRNA functionality are confirmed as being important for the design of active siRNAs against viral genes. The sequence that we can finally recommend as a model for the selection of functional siRNAs directed to virus mRNAs is a RNA duplex of 19 nucleotides with a 2 nucleotides 3′ overhang and the sense strand consisting of 5′-N_1, G_2, $N_{3\text{-}9}$, U_{10}, G_{11}, N_{12}, $U(\text{-}G)_{13}$, $N_{14–18}$ and W_{19}-dTdT-3′ (where subscripted numbers indicate the nucleotide positions and S; N; and W are G or C; an arbitrary nucleotide; and A or U, respectively as defined by the International Union of Pure and Applied Chemistry: IUPAC, http://www.iupac.org. U(-G) at position 13 means the presence of U13 and absence of G13. Moreover, our data show that the intramolecular secondary structure of mRNA target contributes to siRNA efficacy.

Acknowledgements

We thank Sophie Sagnier for technical assistance. We would also like to acknowledge Catherine Cetre-Sossah and Aurélie Perrin for helpful discussions and Samir Messad for statistical support. This research was supported by a Marie Curie International Fellowship within the 6th European Community Framework Program. This study was also partially granted by the European Union Network of Excellence EPIZONE, no. FOOD-CT-2006-016236.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.antiviral.2008.01.159.

References

Amarzguioui, M., Prydz, H., 2004. An algorithm for selection of functional siRNA sequences. Biochem. Biophys. Res. Commun. 316, 1050–1058.

Ameres, S.L., Martinez, J., Schroeder, R., 2007. Molecular basis for target RNA recognition and cleavage by human RISC. Cell 130, 101–112.

Arziman, Z., Horn, T., Boutros, M., 2005. E-RNAi: a web application to design optimized RNAi constructs. Nucleic Acids Res. 33, W582–W588.

Beale, G., Hollins, A.J., Benboubetra, M., Sohail, M., Fox, S.P., Benter, I., Akhtar, S., 2003. Gene silencing nucleic acids designed by scanning arrays: anti-EGFR activity of siRNA, ribozyme and DNA enzymes targeting a single hybridization-accessible region using the same delivery system. J. Drug Target. 11, 449–456.

Bernstein, E., Caudy, A.A., Hammond, S.M., Hannon, G.J., 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409, 363–366.

Bohula, E.A., Salisbury, A.J., Sohail, M., Playford, M.P., Riedemann, J., Southern, E.M., Macaulay, V.M., 2003. The efficacy of small interfering RNAs targeted to the type 1 insulin-like growth factor receptor (IGF1R) is influenced by secondary structure in the IGF1R transcript. J. Biol. Chem. 278, 15991–15997.

Ding, Y., Chan, C.Y., Lawrence, C.E., 2004. Sfold web server for statistical folding and rational design of nucleic acids. Nucleic Acids Res. 32, W135–W141.

Elbashir, S.M., Martinez, J., Patkanoiwska, A., Lendeckel, W., Tuschl, T., 2001. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. EMBO J. 20, 6877–6888.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391, 806–811.

Gilmore, I.R., Fox, S.P., Hollins, A.J., Sohail, M., Akhtar, S., 2004. The design and exogenous delivery of siRNA for post-transcriptional gene silencing. J. Drug Target. 12, 315–340.

Hammond, S.M., Bernstein, E., Beach, D., Hannon, G.J., 2000. An RNA directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. Nature 404, 293–296.

Hammond, S.M., Caudy, A.A., Hannon, G.J., 2001. Post-transcriptional gene silencing by doubled-stranded RNA. Nat. Rev. Genet. 2, 110–119.

Hannon, G.J., 2002. RNA interference. Nature 418, 244–251.

Heale, B.S.E., Soifer, H.S., Bowers, C., Rossi, J.J., 2005. siRNA target site secondary structure predictions using local stable substructures. Nucleic Acids Res. 33, e30.

Henschel, A., Buchholz, F., Habermann, B., 2004. DEQOR: a web-based tool for the design and quality control of siRNAs. Nucleic Acids Res. 32, W113–W120.

Hsieh, A.C., Bo, R., Manola, J., Vazquez, F., Bare, O., Khvorova, A., Scaringe, S., Sellers, W.R., 2004. A library of siRNA duplexes targeting the phosphoinositide 3-kinase pathway: determinants of gene silencing for use in cell-based screens. Nucleic Acids Res. 32, 893–901.

Huesken, D., Lange, J., Mickanin, C., Weiler, J., Asselbergs, F., Warner, J., Meloon, B., Engel, S., Rosenberg, A., Cohen, D., Labow, M., Reinhardt, M., Natt, F., Hall, J., 2005. Design of a genome-wide siRNA library using an artificial neural network. Nat. Biotechnol. 23, 995–1001.

Hutvagner, G., 2005. Small RNA asymmetry in RNAi: function in RISC assembly and gene regulation. FEBS Lett. 579, 5850–5857.

Jagla, B., Aulner, N., Kelly, P.D., Song, D., Volchuk, A., Zatorski, A., Shum, D., Mayer, T., De Angelis, D.A., Ouerfelli, O., Rutishauser, U., Rothman, J.E., 2005. Sequence characteristics of functional siRNAs. RNA 11, 864–872.

Khvorova, A., Reynolds, A., Jayasena, S.D., 2003. Functional siRNAs and miRNAs exhibit strand bias. Cell 115, 209–216.

Köberle, C., Kaufmann, S.H.E., Patzel, V., 2006. Selecting effective siRNAs based on guide RNA structure. Nat. Protoc. 1, 1832–1839.

Levenkova, N., Gu, Q., Rux, J.J., 2004. Gene specific siRNA selector. Bioinformatics 20, 430–432.

Long, D., Lee, R., Williams, P., Chan, C.Y., Ambros, V., Ding, Y., 2007. Potent effect of target structure on microRNA function. Nat. Struct. Mol. Biol. 14, 287–294.

Luo, K.Q., Chang, D.C., 2004. The gene-silencing efficiency of siRNA is strongly dependent on the local structure of mRNA at the targeted region. Biochem. Biophys. Res. Commun. 318, 303–310.

Martinez, J., Patkaniowska, A., Urlaub, H., Lührmann, R., Tuschl, T., 2002. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell 110, 563–574.

Naito, Y., Ui-Tei, K., Nishikawa, T., Takebe, Y., Saigo, K., 2006. siVirus: web-based antiviral siRNA design software for highly divergent viral sequences. Nucleic Acids Res. 34, W448–W450.

Overhoff, M., Alken, M., Far, R.K., Lemaitre, M., Lebleu, B., Sczakiel, G., Robbins, I., 2005. Local RNA target structure influences siRNA efficacy: a systematic global analysis. J. Mol. Biol. 348, 871–881.

Patzel, V., Rutz, S., Dietrich, I., Köberle, C., Scheffold, A., Kaufmann, S.H.E., 2005. Design of siRNA producting unstructured guide-RNAs results in improved RNA interference efficiency. Nat. Biotechnol. 23, 1440–1444.

Pekarik, V., 2005. Design of shRNA for RNAi—a lesson from pre-miRNA processing: possible clinical applications. Brain Res. Bull. 68, 115–120.

Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W.S., Khvorova, A., 2004. Rational siRNA design for RNA interference. Nat. Biotechnol. 22, 326–330.

Servan de Almeida, R., Keita, D., Libeau, G., Albina, E., 2007. Control of ruminant morbillivirus replication by small interfering RNA. J. Gen. Virol. 88, 2307–2311.

Schramm, G., Ramey, R., 2005. siRNAs design including secondary structure target site prediction. Nat. Methods 2, 1–2 (Application note).

Schubert, S., Grünweller, A., Erdmann, V.A., Kurreck, J., 2005. Local RNA target structure influences siRNA efficacy: systematic analysis of intentionally designed binding regions. J. Mol. Biol. 348, 883–893.

Schwarz, D.S., Hutvágner, G., Du, T., Xu, Z., Aronin, N., Zamore, P.D., 2003. Asymmetry in the assembly of the RNAi enzyme complex. Cell 115, 199–208.

Shabalina, S.A., Spiridonov, A.N., Ogurtsov, A.Y., 2006. Computational models with thermodynamic and composition features improve siRNA design. BMC Bioinf. 7, 65.

Ui-Tei, K., Naito, Y., Takahashi, F., Haraguchi, T., Ohki-Hamazaki, H., Juni, A., Ueda, R., Saigo, K., 2004. Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. Nucleic Acids Res. 32, 936–948.

Wang, L., Mu, F.Y., 2004. A web-based design center for vector-based siRNA and siRNA cassette. Bioinformatics 20, 1818–1820.

Westerhout, E.M., Ooms, M., Vink, M., Das, A.T., Berkhout, B., 2005. HIV-1 can escape from RNA interference by evolving an alternative structure in its RNA genome. Nucleic Acids Res. 33, 796–804.

Westerhout, E.M., Berkhout, B., 2007. A systematic analysis of the effect of target RNA structure on RNA interference. Nucleic Acids Res. 35, 4322–4330.

Yoshinari, K., Miyagishi, M., Taira, K., 2004. Effects on RNAi of the tight structure, sequence and position of the targeted region. Nucleic Acids Res. 32, 691–699.

Yiu, S.M., Wong, P.W.H., Lam, T.W., Mui, Y.C., Kung, H.F., Lin, M., Cheung, Y.T., 2005. Filtering of ineffective siRNAs and improved siRNA design tool. Bioinformatics 21, 144–151.

Zamore, P.D., Tuschl, T., Sharp, P.A., Bartel, D.P., 2000. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell 101, 25–33.