www.nature.com/cgt



REVIEW

Precise and efficient siRNA design: a key point in competent gene silencing

E Fakhr^{1,2}, F Zare^{1,2} and L Teimoori-Toolabi²

RNA interference-related strategies have become appealing methods in various fields of research. Exact sequence design of these small molecules is an essential step in the silencing procedure. Numerous researchers have tried to define some algorithms in order to increase the chance of short interfering RNA's (siRNA's) success. In recent decades, online designing software has aimed at promoting the quality of siRNA designing based on the most cited algorithms. According to our previous experiments, a combination of different criteria would be helpful. That is, siRNAs suggested by a combination of tools seem to be more efficient. Furthermore, different factors such as distance of target region to transcription start site, nucleotide composition, absence of off-target effects and secondary structures in the target site and siRNA and the presence of asymmetry and energy valley within the siRNA will increase the efficiency of siRNAs. Despite application of different online tools and fulfilling the criteria, there is no guarantee for designing an effective siRNA. However, meticulous designing of siRNAs according to the suggested algorithms and scoring systems and using different siRNAs for targeting the same gene would lead to improved silencing outcome. In this review, we focus on common algorithms and online software, and introduce a new scoring system used in our experiments.

Cancer Gene Therapy (2016) 23, 73-82; doi:10.1038/cqt.2016.4; published online 18 March 2016

INTRODUCTION

The reliable knockout or knockdown of the genes related to different biological pathways is an essential technique in the molecular studies. This aim can be achieved by RNA interference (RNAi) technique. ^{1,2} This technique exploits an intrinsic system to knock down genes and, compared with other antisense strategies, has a remarkably higher efficiency. RNAi is a promising tool to investigate various molecular pathways underlying numerous diseases and natural processes at the posttranscriptional level. RNA silencing phenomenon was first seen in Caenorhabditis elegans in which degradation of par-1 mRNA was followed by introducing double-stranded RNA to this nematode worm.^{3,4} Antisense procedure in knocking down gene expression has become an appealing technique for decreasing the virulence of pathogens such as bacteria, fungi and viruses^{5,6} in one hand and for studying the role of individual proteins by silencing different mRNA isoforms and investigating novel and effective therapeutic modalities for various disorders such as cancer^{7,8} on the other hand.

Transfection procedure for short interfering RNA (siRNA) can be transient or stable. While in the stable transfection, the exogenous double-stranded DNA expressing short hairpin RNA (shRNA) is exposed to the cells, in transient one the double-stranded RNA is introduced to the target cells by different methods. The stable procedure tends to be more effective in mammalian cells. In stable transfection, the double-stranded DNA will be integrated to the cell's genome and transcribed to shRNA by RNA polymerase II/III. Then the processed shRNA is exported from the nucleus to the cytoplasm. A ribonuclease III enzyme, called Drosha, initiates RNAi and microRNA (miRNA) processing. 10

Dicer, a member of RNaselll family, in cooperation with TAR RNAbinding protein and Argonaute (Ago), produces a 21–25 nucleotide siRNA with two nucleotide 3' overhangs. 11,12 This molecule becomes a part of RNA-induced silencing complex (RISC), which recognizes the target mRNA. Complementation of target mRNA with siRNA can lead to different outcomes varying from repression of translation to degradation of mRNA. The attachment of siRNA and 3' untranslated region (UTR) of mRNA with the help of Ago1, Ago3 and Ago4 will lead to translational repression; however, if the siRNA becomes complement with the coding sequence (CDS) completely, the target mRNA will be cleaved by the endonucleolytic activity of Ago2 13,14 (Figure 1).

Efficiency of siRNA molecules, which depends on different factors, including target availability, secondary structures of mRNA, position of matching and intrinsic characteristics of siRNA and mRNA, can mostly be accomplished by means of applying a good designing method.

Precise design of siRNAs is a critical step owing to the fact that only few changes in the nucleotides within the sequence can alter its functionality. Numerous studies have tried to introduce uniform and practical algorithms for selecting the most efficient siRNAs in which suitable design can lead to a more efficient silencing. In this study, we review various algorithms for designing an efficient siRNA and the advantages and disadvantages of different online software. We also introduce our own scoring system for designing siRNAs based on different methods, including our own studies, for selecting the most efficient siRNAs.

BASIC CRITERIA FOR DESIGNING SIRNAS

The first step in designing a siRNA is assessing the target gene which is going to be silenced. Length of the siRNA, its nucleotide content and specificity of the designed siRNA are other important

¹Department of Medical Biotechnology, Faculty of Advanced Medical Technologies, Tehran University of Medical Sciences, Tehran, Iran and ²Molecular Medicine Department, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran. Correspondence: Dr L Teimoori-Toolabi, Molecular Medicine Department, Biotechnology Research Center, Pasteur Institute of Iran, 69th Pasteur Avenue, Kargar Street, Tehran 13169-43551, Tehran, Iran. E-mail: Iteimoori@pasteur.ac.ir



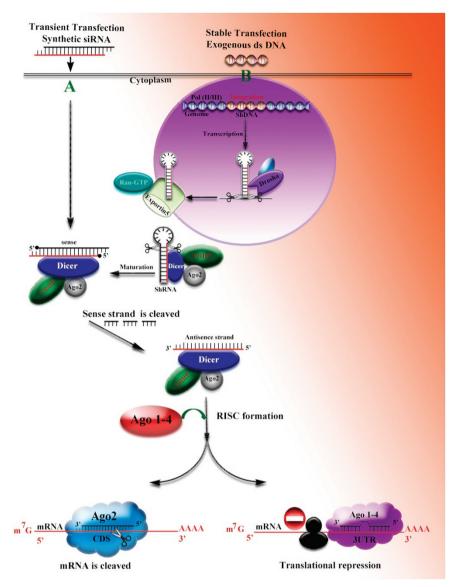


Figure 1. Cellular pathways of RNA interference. Transfection methods can be transient (**a**) or stable (**b**), and both of them cause short interfering RNA (siRNA) formation and the silencing process. After unwinding the siRNA duplex, antisense strand will become a part of RNA-induced silencing complex (RISC). Other types of proteins also contribute to the formation of RISC. Antisense strand is complementary with target gene, and it leads to cleavage of mRNA or translational repression. Ago, Argonaute; CDS, coding sequence; dsDNA, double-stranded DNA; shDNA/shRNA, short hairpin DNA/RNA; TRBP, TAR RNA-binding protein; UTR, untranslated region.

factors influencing the functionality of these small molecules. In what follows, details about these essential elements for designing appropriate siRNAs are presented.

Target site

Single-nucleotide polymorphism (SNP) locations are excluded as the target site for siRNA designing as they might cause variation in the silencing efficiency of different cell lines with different SNPs. Based on the study by Elbashir *et al.*, ¹⁸ it is better to avoid considering intronic parts and 5'UTR or 3'UTR as the sites for siRNA designing. UTRs and sequences close to start codon are more prone to regulatory proteins that can interfere with RISC complex and its silencing effect. Regions about 50–100 nucleotides downstream of start codon in the open reading frame of the target gene are the best target sites to be silenced. ¹⁸ Our own experiences showed that siRNAs closer to the start codon tend to be more effective than farther ones. ¹⁹ Searching for motifs such as

5'-AA(N19)TT, 5'-AA(N21) or 5'-NA(N21) on the target sequence is worthwhile as guanidine–cytosine (GC) content of these regions is < 50%; also, they facilitate the function of RISC complex. 20,21 Recent studies have proved that the accessibility of the target site owing to the secondary structures of the mRNA is another determinant of siRNA's functionality and any variations in partial base pairing of the target site will influence the effectiveness of siRNA. 22,23 Native mRNAs have intrinsic secondary structures that can be predicted by the software such as Mfold (http://mfold.rna. albany.edu/?q = mfold). 24

Gredell et al.²⁵ showed target sites, including four consecutive unpaired bases (within mRNA) at their 5′-or 3′-ends are more prone to the silencing effect of siRNAs than target sites comprising unpaired regions in their central region. As the accessibility of the target site affects the siRNA efficiency, Tafer et al.²⁶ constructed a software called RNAxs (http://rna.tbi.univie.ac.at/cgi-bin/RNAxs), which is a siRNA designing tool mainly based on mRNA target site accessibility.



Length of siRNA

There is a controversy over the best length of siRNAs. Although some groups, including ours, have worked with 19 nucleotide ones²⁷⁻³¹ and obtained very good results, others have used longer siRNAs ranging from 21 to 29 nucleotides.^{32–34} Although shorter siRNAs may lead to an unspecific binding, siRNAs from 19 to 25 nucleotides have shown the same efficiency in silencing. 35,36 However, smaller siRNAs are better to use for mammalian cells as longer siRNAs can induce mammalian immune response.³⁷

Specificity checking

After designing siRNAs with different methods, both sense and antisense strands should be checked via blast with reference sequence database (Refseq-RNA database) of the desired organism to reduce the risk of silencing unintended genes. As their alignment may not be the result of chance, blast's results with smaller E-values should not be overlooked. Less than 78% guery coverage with other genes, ≤ 15 nucleotides out of 19 matching with the respective siRNA, is believed to be tolerable.³⁸ Yet, it is worthwhile to mention that there is always a probability of unpredictable off-target effects for siRNAs.

Nucleotide content of siRNA

Several studies have attempted to establish uniform and practical algorithms in order to design the most effective siRNAs. Nevertheless, some of the criteria that are introduced by different algorithms overlap with each other. ^{39–42} Table 1 presents a summary of their essential parameters.

GC content is an obvious and basic parameter in the efficiency of siRNAs. This is due to the fact that low GC content leads to unspecific and weak binding, while high GC content may hinder unwinding the siRNA duplex by helicase and RISC complex.⁴³ Different studies have demonstrated different acceptable intervals for GC content. Amarzguioui *et al.*⁴¹ suggested that the GC content should range from 31.6% to 57.9%.⁴¹ Besides, Reynolds et al.³⁹ found the most functional siRNAs, which means up to 95%, having GC content of 36–52%.³⁹ Ui-Tei *et al.*⁴⁰ explored more exactly and showed that, in the antisense strand, GC percentages between the second to the seventh and the eighth to the eighteenth nucleotides were desirable to be 19% and 52%, respectively. In addition, functional siRNAs have an unstable region (lower GC content than others) between the ninth to the fourteenth nucleotides, called energy valley, which is a vital criterion for the selection of siRNAs. This internal instability increases the RISC complex functionality by inducing the most desirable conformation during mRNA cleavage.44

Sense and antisense strands should have the ability to form the duplex properly and avoid any secondary structures that can be prohibitive (Figure 2). SiRNA sequences must be screened for internal repeats and palindromes. However, siRNAs in which the internal secondary structures have Tm (melting temperature) values < 20 °C, can be tolerated when the body temperature reaches to 37 °C, which, in turn, unwinds the respective secondary structures. Evaluation of secondary structure can be carried out by Mfold.

The most substantial criterion for siRNA selection is low internal stability at 5'-end of the antisense strand, which is probably an important factor for proper unwinding of siRNA duplex and entering into the RISC complex.⁴⁵

This instability can be easily assessed by the number of A/U nucleotides. Ui-Tei et al.40 showed that, for having an effective siRNA, at least four out of the seven nucleotides in the 5'-end of antisense strand should be A or U.46 In the same line, Reynolds et al.³⁹ proposed that at least 3 A/U bases between the thirteenth and the nineteenth nucleotides of the sense strand were needed. However, Amarzquioui et al.41 indicated that A/U duplex differential between two strands, especially in the last

Table 1. Summary of essential parameters from different algorithms for designing a functional siRNA

Algorithms
Amarzguioui et al.41
Reynolds et al. ³⁹
Ui-Tei <i>et al</i> . ⁵¹
Amarzguioui et al.41
Ui-Tei <i>et al</i> . ⁵¹
Jagla et al. ⁴²
Reynolds et al. ³⁹
Ui-Tei <i>et al</i> . ⁵¹
Amarzguioui et al. ⁴¹
Reynolds et al. ³⁹
Reynolds et al. ³⁹
Jagla et al. ⁴²
Reynolds <i>et al.</i> ³⁹ Jagla <i>et al.</i> ⁴²

three base pairs, correlates with functional siRNAs. Jaqla et al.⁴² also confirmed that the presence of >3 A/Us at the thirteenth to nineteenth position is important for siRNA's functionality. In addition, high internal stability or high G/C content at 3'-end of the antisense strand is influential in the efficiency of siRNA.

Sequences such as GGGG or CCCC should be avoided as they increase the risk of hairpin structures. 46 More than three A or U nucleotides can also be detrimental because RNA polymerase III tends to end the transcription at poly U site. 47 The presence of purine at the 5'-end of the sense strand makes the initiation of transcription, by RNA polymerase III, more efficient. siRNAs lacking internal repeats and palindromes are strongly preferred.

Other parameters in sequence were considered to be important for the efficiency of a siRNA. Amarzguioui et al.41 found that the sixth nucleotide of the antisense strand is preferred to be A. Conversely, U1 and G19 in the antisense strand decrease the efficiency of siRNA.

Furthermore, algorithm by Reynolds et al.³⁹ pays attention to the sequence preferences of the sense strand. Based on the studies by Reynolds *et al.*,³⁹ the presence of A at the nineteenth position, A at the third position and U at the tenth position, besides the absence of G or C at the nineteenth and G at the thirteenth nucleotide of sense strand are influential.

Among all these suggested nucleotides, U10 showed the strongest correlation with the efficiency of siRNA owing to the fact that RISC cleaves target mRNA between nucleotide 10 and 11 and, similar to most of endonucleases, prefers to cut 3' of U rather than other bases⁴⁵ (Figure 2).

Moreover, Jagla and his research team analyzed 600 chemically synthesized siRNAs⁴² and revealed four sets of rules based on the base composition of the sense strand. According to their findings, the first rule is the best one that increases the chance of functional siRNA prediction up to 90%. Here are the rules:

Rule 1: Presence of A/U at position 19, GC at the first position, A/U at position 10, >3 A/Us at position 13–19.

Rule 2: Presence of A/U at position 19, GC at the first position, GC at position 10, > 3 A/Us at position 13–19.

Rule 3: Presence of G/C at position 19, GC at position 1, GC at position 11, >6 A/Us at position 5–19.

Rule 4: Presence of A/U at position 19, A/U at position 1, > 3 A/Us at position 13-19.

Two nucleotides can be added chemically to the 3'-end of sense and antisense strands and these siRNA duplexes tend to be more efficient than duplexes with \geqslant 3 '-overhanging nucleotides.⁴⁸



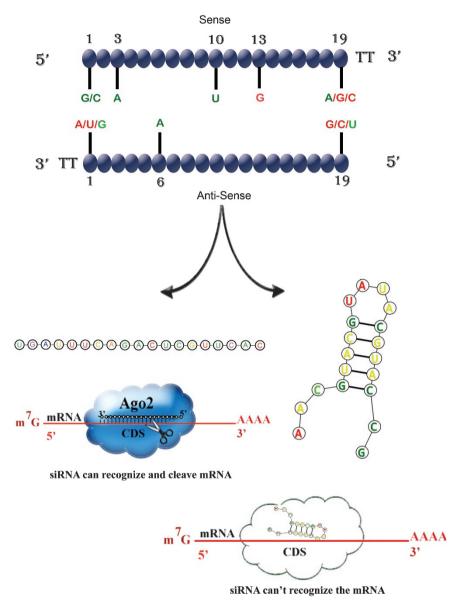


Figure 2. Preferences of bases in different positions of sense and antisense strands of short interfering RNA (siRNA); green bases are favorable in that positions, whereas red ones should be avoided. Although nucleotide composition of the sense and antisense strands is important, the sequences producing secondary structures should not be ignored owing to the fact that such secondary structures are prohibitive. Ago, Argonaute; CDS, coding sequence.

Among different base compositions of the overhangs, TT overhangs may be more beneficial than others as they reduce the expense of RNA synthesis and can increase the persistence of siRNA duplexes by enhancing the resistance of siRNA duplexes to RNAase.⁴⁹

ONLINE SOFTWARE FOR DESIGNING SIRNAS

There are lots of online tools for siRNA designing. Some of them are updated regularly whereas others are not supported any more. In this part, we introduce some of the best online software and their characteristics. Their advantages and limitations are shown in Table 2.

siDirect

SiDirect (http://sidirect2.rnai.jp/) is an online tool for designing mammalian siRNA using the algorithm by Ui-Tei *et al.*⁵¹ by default.

Users can choose other algorithms such as by Reynolds *et al.*, ³⁹ Amarzguioui *et al.* ⁴¹ and various combinations of these algorithms provided in the software, resulting in more definite and flexible designing of siRNAs. ⁵⁰

The most important parameter of siDirect is its capability to create and use a non-redundant sequence set of genes in order to decrease the effect of alternative splicing. By this non-redundant database, users can design a siRNA that is specific for just one transcript variant of the desired gene. This database is a qualified form of UniGene and RefSeq databases and includes sequences of duplicate-free exons and exon–exon junctions. However, non-redundant databases are limited to *Homo sapiens*, *Mus Musculus* and *Rattus Norvegicus*.

Both sense and antisense strands are evaluated against nonredundant database as it cannot be overlooked that sense strand can be probably combined with the RISC complex and may cause off-target silencing.



Online softwares	Advantages	Limitations	Offered loops
siDirect	(a) Using non-redundant database (b) Using different algorithms and their combinations (c) Evaluating seed duplex Tm for reducing off-target effects (d) Presenting the sequence of sense and antisense strands and exact siRNA position	(a) Limited to non-redundant database (b) Just 21-nucleotide siRNAs are designed (c) Ignoring AA overhangs on target sequence	None
siDESIGN Center	(a) Isoform-specific siRNA design and siRNA design for conserved genes (b) Designing cross-species siRNAs (c) Considering off-target effects by evaluating the seed sequence	(a) The sequence of siRNA is not displayed (b) Just 19-nucleotide siRNAs are designed (c) Databases for seed sequence are limited to human, mouse and rat	None
Genscript	(d) Minimum mismatches of sense and antisense strands with target sequence evaluated by Blast (a) Comprehensive filtering parameters (b) The facilities of siRNA Construct building and siRNA scrambler (c) Off-target effect can be reduced sufficiently up to 95%	(a) Databases are limited to human, mouse and rat (b) Designing only 21-nucleotide siRNAs (c) Coding squence of the target gene should	(a) TTGATATCCG
Asi designer	(a) Designing siRNAs for targeting different isoforms of desired gene simultaneously (b) Showing two nucleotide overhangs at 3' and 5'-ends. Users can set types of overhangs, such as AA, NN and NA (c) Reporting self-alignment energy of siRNAs and EP score (binding energy profile score according to their algorithm) melting temperature (°C), identity limit for off-target filtering (d) SNP database is included (e) Off-target filtering by NCBI BLAST and analyzing siRNA	be used (a) All isoforms of different genes are not included (b) Designing only 19-nucleotide siRNAs (c) siRNA's position from the transcription start site is not reported exactly and antisense strand is not shown	None
BLOCK-iT RNAi Designer	secondary structure folding by Mfold (a) Highly guaranteed (b) Facility of shRNA design with various vectors (c) Including databases of different species for doing blast (d) Introducing four groups of siRNAs (e) Presenting the blast's results separately	(a) Not showing different algorithms and various parameters to choose (b) Not including splice variants for the genes (c) Only designing 19-nucleotide siRNAs (d) Not showing Antisense strand in the results	(a) GAGA (b) AACG (c) CGAA
RNAi explorer	(a) Designing a wide range of siRNAs; (b) shRNA designing and offering different loops	(a) Not including different algorithms and parameters (b) Only designs 19-nucleotide siRNAs (c) No guarantee for siRNAs' functionality	(a) TTCG (b) TCAAGAG (c) GAAGCTTG
RNAi codex	(a) Query can be entered in any format, even the name of the gene (b) The shRNA information displayed on the screen can be downloaded too (c) Antisense, sense and linker are shown in the result section (d) Being linked with lots of NCBI's related databases	(a) Not showing any parameters for the users (b) No guarantee for the shRNAs (c) Only designing 21 mere siRNAs	(a) Human-mir -30 (b) Human-mir -1 (user cannot change the sequence of loop
SiRNA at Whitehead	(a) User friendly results and explanations (b) Covering different algorithms and parameters and considering 3' overhangs (c) Being linked with lots of NCBI's related databases (d) Presenting sense, antisense and even mRNA sequence in the results	(a) The website is not maintained and is only being updated for blast databases (b) No guarantee for the shRNAs (c) Only designing 21 mere siRNAs	None
Oligowalk	(a) Comprehensive thermodynamics results (b) Considering the probability of being efficient siRNA (c) Sending the results via e-mail	(a) Not updated since 2008 and no guarantee for its siRNAs	None
SiRNA Design (IDT)	(a) User friendly results and explanations (b) Covering different pattern of bases and asymmetrical end stability (c) Including a new form of siRNA design with sufficient guide and explanations	(a) The interactive tutorial is not updated (b) Only designing 25-nt siRNAs in custom approaches	None
RNA wizard	(d) Showing blast related links in the results (a) Designing siRNAs in various sizes (b) shRNA design and offering efficient and novel vectors (c) Performing two kinds of blast: (1) blast against mRNA database and (2) blast against miRNA database (d) Considering optimal internal stability of siRNAs (e) Excluding immunogenic sequences	(a) Not including various algorithms and parameters (b) Databases limited to human, rat and mouse (c) No guarantee for siRNAs' functionality (d) Only the sequence of the guery is accepted	(a) TCAAGAG

Blast has some obstacles to find high similarities between small sequences owing to its default parameters; it ignores some of the off-target candidates. As a solution, 'mismatch tolerance' is used on siDirect. As mismatch tolerance shows the number of mismatches between siRNA sequence and unintended targets, a higher mismatch tolerance is desirable. It was seen that sense and antisense strands with $\geqslant 3$ mismatch tolerances have high specificity.

Ui-Tei et al.⁵¹ found that off-target effects are highly related to the thermodynamic properties of duplex formed between siRNA's

seed region (the second to the eighth nucleotides from the 5'-end of the guide strand) and 3'UTR of target mRNA. Then, for having a functional siRNA with reduced off-target effects, melting temperature of seed duplex should be $< 21.5\,^{\circ}\text{C}$.

Other parameters such as GC content, number of consecutive C/G or A/T and target region can be determined in this tool. In addition, the users are able to define their own pattern of siRNAs.

Searching for AA overhangs on target sequence is not included in this software as it would limit the number of predicted siRNAs.



siDESIGN Center

Dharmacon is efficient online software for designing siRNAs and is based on the algorithms by Elbashir et al. 18 and Reynolds et al. 39 (http://dharmacon.gelifesciences.com/design-center/). can use accession numbers or FASTA format of the target sequences. If designing a specific siRNA against a variant of mRNA is preferred, the accession number should be used. By providing the accession number, different variants of desired target will be displayed and some of them can be excluded. Besides, Dharmacon can design cross-species siRNA, in addition to siRNA, against the conserved gene family members among different species. If the nucleotide sequence is provided, the seed region specificity of siRNAs can be evaluated. The exact matches of the seed region of siRNA to the 3'UTR of the genes are determined and, then the lower frequency of seed region in the target genome proves the specificity of siRNA. This analysis can only be carried out for human, mouse and rat. Open reading frames are the best regions for siRNA designing because of being more conserved. In the cases that a few number of siRNA candidates is suggested in CDS owing to sequence limitations, or the criteria for a good target are not fulfilled, UTRs can be considered as target sites for siRNAs. As a result, Dharmacon suggests three options for target sites: 3'UTR, CDS, and 5'UTR. The sequences of siRNAs are not shown in the output. The minimum mismatches in sense and antisense strands to the target gene are also reported.

Genscript

GenScript USA Inc. (Piscataway, NJ, USA) has a complete online software collection for siRNA design and is mainly based on the algorithms by Reynolds *et al.*³⁹ It consists of three online software:

siRNA Target Finder siRNA Construct Builder siRNA Sequence Scrambler

The users are able to design siRNA, build the construct based on their insert and use Scrambler to select a suitable negative control for the desired siRNAs.

GenScript siRNA Target Finder (https://www.genscript.com/ssl-bin/app/rnai) is based on different algorithms and empirical parameters. This software is unique because of its filtering parameters and controlled off-target effects. Its filtering parameters include different organisms, different organs and the size of seed region; however, designing is only supported for human, rat and mouse genome. It also has a functional alignment filter based on the asymmetry of siRNA and its effect on RICS formation.

As some researches have shown that some sequences can stimulate immune response, these types of sequences are also filtered. 52,53 siRNAs with a lot of consecutive base repetition. stringent secondary structures and location on SNP sites are deleted. GenScript only confirms the siRNAs with overlaps < 16 nucleotides with other transcripts. It can also be restricted to the tenth and the twelfth nucleotides. This approach will lead to a remarkable reduction of off-target effects. As some siRNAs with high functionality may be removed from the final list, because of possible off-target effects, users can choose 'tradeoff' option to save those siRNAs in the list. In this software, the only acceptable sequence is cDNA. ΔE for each siRNA mainly shows the specificity of target region. Also, siRNAs with unfavorable thermodynamic parameters are eliminated. After siRNA designing, users can build the construct with selected siRNAs. In addition, the loop sequences can be changed.

Asi designer

Young-Joo Kim and his group provided online designing software based on their novel algorithm (http://sysbio.kribb.re.kr/AsiDe

signer). Kim's algorithm is exon based and is mainly concerned with alternative splicing.³⁸ The most important parameter of this website is to consider different isoforms of the desired gene by mapping the genome and designing siRNAs capable of silencing more than one mRNA at the same time. By means of this approach, users can choose the isoforms to be silenced. Databases of three organisms, H. sapiens, M. musculus and R norvegicus are included in this online tool. The guery can be put as a form of sequence or its accession number. If the isoforms related to desired gene are found, they will be shown. Asi designer can design siRNA sequences in three formats: NN (N19), NA (N19), and AA (N19); that is, only 19 nucleotide siRNAs can be designed by this tool. Users can define a range in target mRNA sequence for siRNA designing, Besides, three options of CDS, internal CDS (100 bp after translation start site) and full transcript are provided. For general purposes, CDS and internal CDS are the best choices. The users are able to either change the number of consecutive bases or simply leave the default number unchanged. The number of siRNA candidates can be limited too.

BLOCK-iT RNAi designer

Thermo Fisher Scientific Inc. (Waltham, MA, USA) web site has a comprehensive online designing tool, called BLOCK-iT, which highly guarantees its designed siRNAs (https://rnaidesigner.life technologies.com). In this software, basic criteria and blast databases for different species are provided. Users can choose default or Tuschl's motif pattern for siRNA's sequences. This software, also, has the facilities of shRNA construction. BLOCK-iT introduces four categories of siRNA, three of them being chemically modified for obtaining the best results.

Chemically modified siRNAs

Silencer select siRNA: it is mainly recommended for *in vitro* experiment. Latest improvements are considered in its designing process.

Ambion *In vivo* siRNA: These groups are chemically modified in order to increase serum stability, so they will be beneficial for *in vivo* experiments.

Stealth RNAi: This kind of siRNAs is meticulously designed and modified. It is suitable for both *in vitro* and *in vivo* experiments. Stealth siRNAs have four unique characteristics: (1) effective knockdown, (2) higher specificity, (3) greater stability, and (4) less cellular toxicity (by suppressing the initiation of Protein kinase R/ interferon response pathway)

Unmodified siRNAs: They are the first generation of BLOCK-iT siRNAs and are mainly suggested for *in vitro* experiments.

RNAi explorer

Gene Link Inc. (Westchester, NY, USA) introduced an online designing software based on basic and simple criteria (http://www.genelink.com/sirna/RNAicustomorder.asp) that does not guarantee the experimental results of siRNAs. In this tool, parameters such as GC content and consecutive bases are included. RNAi explorer introduces lots of candidates that can be limited by the user to top 40s. It only designs 19 nucleotide siRNAs, and while its position on target mRNA is shown, antisense strand is not illustrated. Although RNAi explorer contains shRNA design section that offers different options and various loops, its results are not as clear as other shRNA designing software.

RNAi codex

Cold Spring Harbor Laboratory introduced an online software (http://codex.cshl.edu/.) that mainly designs shRNAs. Its database is limited to human, mouse, rat and fruit fly but its going to be expanded and include other organisms.



SiRNA at Whitehead

SiRNA at Whitehead is another useful program developed by Whitehead Institute for Biomedical Research (http://sirna.wi.mit. edu/home.php). This software is an attempt to cover various algorithms and parameters carefully. It restricts the Blast's parameters and suggests two databases for performing Blast: RefSeq database and Ensemble transcript. Also, a special Blast is provided in which seed region of siRNA is blasted against seed region of miRNAs. The results and explanations of siRNA at Whitehead are user friendly and complete. Right now, the website is no more maintained and is only available for blast.

Oligowalk

Oligo walk was designed by Mathew's group at the University of Rochester Medical Center (http://rna.urmc.rochester.edu/cgi-bin/server_exe/oligowalk/oligowalk_form.cgi). It mainly selects functional siRNAs based on thermodynamic hybridization properties and the algorithm by Reynolds *et al.*³⁹

SiRNA design (IDT)

Integrated DNA Technologies Inc. (San Diego, CA, USA) offers a siRNA design tool that covers various parameters and patterns to design efficient and functional siRNAs (http://eu.idtdna.com/Scitools/Applications/RNAi/RNAi.aspx?source = menu).

Custom RNAi Design tool is an user-friendly software and has provided an interactive tutorial for novice users. However, some updates and changes are not included in its tutorial. This software allows the users to select and introduce their own patterns based on base composition. Asymmetrical end stability of antisense strand is also considered. A unique option of IDT's siRNA design is the introduction of a new format of siRNA design called Dicer Substrate RNAi Design. This approach is mainly dedicated to the interaction of Dicer with RNA duplexes and its role in proper RISC formation. It seems that, with the intelligent design of 27nucleotide RNA duplexes, they can interact properly with Dicer and, as a result, the efficiency of silencing will increase.⁵⁴ The criteria for designing such duplexes are explained clearly in design help section. Therefore, the users can choose one of the two classes of RNAi duplexes: (1) custom designed and (2) DsiRNA 27 mer. After presenting the results, blast can be carried out by users and the exact position of duplexes on the target sequence will be included.

RNA wizard

In vivo gene offers a siRNA designing tool called RNA wizard (http://www.sirnawizard.com/design_advanced.php). RNA wizard consists of three sections: (1) Find siRNA sequence, (2) Scramble siRNA (for generating negative control of siRNA) and (3) Design hairpin insert. These sections are connected with each other in the result page of designed siRNAs.

In the siRNA designing section, basic criteria are considered. Here, unlike most siRNA designing software, users are able to determine the size of their siRNA. RNA wizard can perform two kinds of Blast: (1) blast against mRNA database and (2) blast against miRNA database. However, the databases are limited to human, mouse and rat. The blast results are presented with sufficient information and expression profiles of related homologous genes including UniGene's link. RNA wizard offers *in vivo* gene's psiRNA vectors for *in vitro* and *in vivo* usage. Also, some of these vectors can be used for simultaneous expression of two shRNAs. Similar to Genscript software, the sequences known to incite immune response such as GTCCTTCAA, TGTGT and CTGAATT are excluded. 52,55

After precise design of siRNAs including sense and antisense strands with respective designing software, short hairpin siRNAs should be ordered for construction by online software, such as

Genscript, BLOCK-iT RNAi Designer, RNAi explorer, RNAi codex and RNA wizard, which enjoy this capability. For constructing any short hairpin siRNA, there is an important part called loop that links sense and antisense strands. Based on in vivo molecular biology studies, loops ranging from the third to the ninth nucleotide have the same effectiveness. It is observed that linkers such as CCACC, CCACACC,²⁹ AAGCTT, CTCGAG³⁴ and TTCAAGAGA³⁰ are also practical. In different software, some loops are offered and users can also change the sequence of loop (Table 2). For assessing transfection efficiency in any RNAi related experiment, designing proper non-targeting controls, called scrambled siRNAs, is needed and an arbitrary DNA sequence can be used for these negative control siRNAs. Scrambled siRNAs must be filtered in order to decrease the probability of their attachment to other mRNA's sequences and miRNA-related seed regions. Online software such as RNA wizard and Genscript will help the researchers to design an appropriate scrambled siRNA. Some of these scrambled siRNAs can be obtained commercially. Although some of them could be considered scrambled for all mammalian mRNAs, others are negative controls for only a set of siRNAs and could be used in one set of experiment. The latter type of scrambled siRNAs could be designed by scrambling the nucleotide sequences of designed siRNA and blasting them against the target genome such as Refseq-RNAs.

THE SCORING SYSTEM FOR SIRNA DESIGN

There are various online software and algorithms for precise design of siRNAs. However, while none of them guarantees 100% efficiency of siRNAs, they all strive to increase the probability of siRNA's functionality. As a result, in our experiments, we tried to create our own scoring system based on combination of different algorithms and using online software. We have used this scoring system in order to design siRNA in our previous experiments. 5,19,56–58 Hereby, we introduce this scoring system step by step:

(1) Using various online software:

As described in the previous parts, all online designing tools have their own advantages and limitations. Therefore, siRNAs designed by more than one software/tool seem to fulfill most criteria and are more favorable.

(2) Doing Blast via NCBI Blast tool and interpreting the results: Most of the designing software has an internal blast. Besides, a more precise blast against refseq-RNA database for both sense and antisense strands is mandatory. As Blast has some limitations in alignment of small sequences, some parameters of blast's algorithms should be changed. Blast has three different programs: (1) mega blast, (2) discontinuous mega blast, and (3) Blastn. Blastn program is more suitable for smaller queries as it allows assessing smaller word sizes. It should be better to set the word size to seven in order to have more precise alignment. In Blast program, the Expect threshold is set 10 by default. For doing a more stringent specificity checking, Expect threshold should be as high as 1000 or 3000, similar to primer blast program. In the blast section of IDT's siRNA designing tool, the Expect threshold is considered to be 1000 but it can be ≥ 3000.

The queries with >78% coverage with the subject are considered as a risk factor for off-target effects. In addition, based on our previous experiments, 19,56 the strands of RNA duplex, which are homologous with >7 nucleotides in undesired genes or with a query coverage $\le 78\%$, is recommended to be discarded. It must be mentioned that the attachment of inner nucleotides of siRNA to unwanted genes is more detrimental than the attachment of 3′- or 5′-ends of it.

(3) Filter siRNAs based on parameters of different algorithms: SiRNAs are filtered through a scoring system. Almost all basic criteria and parameters of different algorithms are considered in this scoring system.



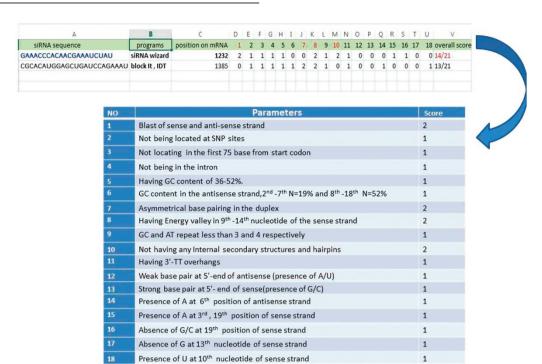


Figure 3. Scoring system for selecting the best short interfering RNA (siRNA) among different siRNA candidates: siRNAs with scores > 14 can be considered as good candidates.

Most of the parameters get one score in our scoring system. These criteria include:

Target site criteria:

Not being in SNP sites

Not being in the first 75 bases from the start codon

Not being in the intron.

Nucleotide content of siRNA:

GC content of 36-52%

GC content in the second to seventh nucleotides of the antisense strand = 19% and in the eighth to eighteenth nucleotides of antisense strand = 52%

Frequency of GC and AT repeats less than 3 and 4, respectively TT overhangs in 3'-end

Weak base pairing at 5'-end of the antisense strand (presence of A/U)

Strong base pairing at 5'-end of the sense strand (presence of G/C) Presence of A nucleotide at the sixth position of the antisense strand

Presence of A nucleotide at the third and the nineteenth position of the sense strand

Absence of G/C at the nineteenth position of the sense strand Absence of G nucleotide at the thirteenth position of the sense strand

Presence of U nucleotide at the tenth position of the sense strand.

In our scoring system, important criteria get two scores. These criteria include:

Very good blast of sense and antisense strands (if just one strand is good, give one point)

Asymmetrical base pairing in the duplex (More A/U at 5' of antisense sense strand and more G/C at 5' of sense strand)

Presence of energy valley in the ninth to fourteenth nucleotide of the sense strand

Lack of internal secondary structures and hairpins, which can be studied by the oligoanalyzer software (Molecular Biology Insights, Inc., Cascade, CO, USA) or Generunner (Hastings software, Inc., Hastings, NY, USA). Therefore, any siRNAs that meet the defined parameters will get one score or two scores, and if it does not fulfill the criteria, it will not receive any score. This scoring system can be managed by an MS Excel file (Figure 3). Thereafter, the maximum overall score will be 21, and siRNAs acquiring higher total score will be considered as good candidates.

Finally, for achieving the best silencing results, it is suggested that at least three siRNAs should be designed for any experiment. This is due to the fact that the efficiency of a single siRNA is not guaranteed, and different siRNAs have different off-target effects decreasing the bias of using one siRNA.

CONCLUSION

By exact silencing of the genes via siRNAs, we will be able to study the role and the effect of the genes in biological pathways profoundly and obtain a comprehensive vision of biological processes. In this review, we mainly focused on different components of siRNA designing, including its basic criteria, different algorithms and online designing software, and we finally suggested a new scoring system for selecting efficient siRNAs. Although precise design of siRNA is vital in RNAi technology, it is not enough for an efficient experiment. Indeed, in order to have an efficient silencing outcome, efficiency of transferring these siRNAs to the cells is another substantial factor. Also, various other studies have examined safe delivery of therapeutics siRNAs, their role in stimulating immunological systems in vivo and finally their efficiency in the body. 59,60 Finally, it is suggested that at least three different siRNAs for each gene should be designed to achieve the desired results in gene silencing. Taken together, by conducting a meticulous design of siRNAs using a reliable and



suitable delivery method, the silencing of the target gene will lead to the desired results.

ABBREVIATIONS

Ago, Argonaute; CDS, coding sequence; GC content, guanidinecytosine content; miRNA micro RNA; Refseg database, Reference sequence database; RISC, RNA-induced silencing complex; RNAi, RNA interference; shRNA, short hairpin RNA; siRNA short interfering RNA; UTR, untranslated region.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- 1 Matzke MA, Birchler JA. RNAi-mediated pathways in the nucleus. Nat Rev Genet 2005; **6**: 24-35.
- 2 Huppi K, Martin SE, Caplen NJ. Defining and assaying RNAi in mammalian cells. Mol Cell 2005; 17: 1-10.
- 3 Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature
- 4 Guo S. Kemphues KJ, par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. Cell 1995: **81**: 611-620.
- 5 Moazeni M, Khoramizadeh MR, Kordbacheh P, Sepehrizadeh Z, Zeraati H, Noorbakhsh F et al. RNA-mediated gene silencing in Candida albicans: inhibition of hyphae formation by use of RNAi technology. Mycopathologia 2012; 174:
- 6 Fischer LT, James QY. RNAi, a new therapeutic strategy against viral infection. Cell Res 2004; 14: 460-466.
- 7 de Fougerolles A, Vornlocher H-P, Maraganore J, Lieberman J. Interfering with disease: a progress report on siRNA-based therapeutics. Nat Rev Drug Discov 2007;
- 8 Yin JQ, Gao J, Shao R, Tian WN, Wang J, Wan Y. siRNA agents inhibit oncogene expression and attenuate human tumor cell growth. J Exp Ther Oncol 2003; 3:
- 9 Ambesajir A, Kaushik A, Kaushik JJ, Petros ST. RNA interference: a futuristic tool and its therapeutic applications. Saudi J Biol Sci 2012; 19: 395-403.
- 10 Filippov V, Solovyev V, Filippova M, Gill SS. A novel type of RNase III family proteins in eukaryotes. Gene 2000; 245: 213-221.
- 11 Paddison PJ. RNA interference in mammalian cell systems. Curr Top Microbiol Immunol 2008; 320: 1-19.
- 12 Ji X. The mechanism of RNase III action: how dicer dices. Curr Top Microbiol Immunol 2008; 320: 99-116.
- 13 Rivas FV, Tolia NH, Song J-J, Aragon JP, Liu J, Hannon GJ et al. Purified Argonaute2 and an siRNA form recombinant human RISC. Nat Struct Mol Biol 2005; 12:
- 14 Su H, Trombly MI, Chen J, Wang X. Essential and overlapping functions for mammalian Argonautes in microRNA silencing. Genes Dev 2009; 23: 304-317.
- 15 Harborth J, Elbashir SM, Vandenburgh K, Manninga H, Scaringe SA, Weber K et al. Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing. Antisense Nucleic Acid Drua Dev 2003: 13: 83-105.
- 16 Taxman DJ, Livingstone LR, Zhang J, Conti BJ, locca HA, Williams KL et al. Criteria for effective design, construction, and gene knockdown by shRNA vectors. BMC Biotechnol 2006; 6: 7.
- 17 Takasaki S, Kotani S, Konagaya A. Selecting effective siRNA target sequences for mammalian genes. RNA Biol 2005; 2: 21-27.
- 18 Elbashir SM, Harborth J, Weber K, Tuschl T. Analysis of gene function in somatic mammalian cells using small interfering RNAs. Methods 2002; 26: 199-213.
- 19 Teimoori-Toolabi L, Hashemi S, Azadmanesh K, Eghbalpour F, Safavifar F, Khorramizadeh MR. Silencing the wild-type and mutant K-ras increases the resistance to 5-flurouracil in HCT-116 as a colorectal cancer cell line. Anticancer Drugs 2015; 26: 187-196.
- 20 Ding Y, Lawrence CE. 8 Rational design of siRNAs with the Sfold software. RNA Interference Technology. In: Krishnarao A (ed). From Basic Science to Drug Development. Cambridge University Press: UK, 2005, pp 129–138.
- 21 Van Vranken D, Weiss G. Introduction to Bioorganic Chemistry and Chemical Biology. Garland Science: USA, 2012, pp 155-156.
- 22 Westerhout EM, Berkhout B. A systematic analysis of the effect of target RNA structure on RNA interference. Nucleic Acids Res 2007; 35: 4322-4330.

- 23 Yoshinari K, Miyaqishi M, Taira K. Effects on RNAi of the tight structure, sequence and position of the targeted region. Nucleic Acids Res 2004; 32: 691-699.
- 24 Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 2003: 31: 3406-3415.
- 25 Gredell JA, Berger AK, Walton SP. Impact of target mRNA structure on siRNA silencing efficiency: a large-scale study. Biotechnol Bioena 2008: 100: 744-755.
- 26 Tafer H, Ameres SL, Obernosterer G, Gebeshuber CA, Schroeder R, Martinez J et al. The impact of target site accessibility on the design of effective siRNAs. Nat Biotechnol 2008: 26: 578-583.
- 27 Yu J-Y, DeRuiter SL, Turner DL. RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. Proc Natl Acad Sci 2002; 99:
- 28 Paul CP, Good PD, Winer I, Engelke DR. Effective expression of small interfering RNA in human cells. Nat Biotechnol 2002; 20: 505-508.
- 29 Jacque J-M, Triques K, Stevenson M. Modulation of HIV-1 replication by RNA interference. Nature 2002: 418: 435-438.
- 30 Brummelkamp TR, Bernards R, Agami R. A system for stable expression of short interfering RNAs in mammalian cells. Science 2002; 296: 550-553.
- 31 Mivagishi M. Taira K. U6 promoter-driven siRNAs with four uridine 3' overhangs efficiently suppress targeted gene expression in mammalian cells. Nat Biotechnol 2002: 20: 497-500.
- 32 Paddison PJ, Caudy AA, Bernstein E, Hannon GJ, Conklin DS. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. Genes Dev 2002; **16**: 948-958
- 33 Lee NS, Dohjima T, Bauer G, Li H, Li M-J, Ehsani A et al. Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. Nat Biotechnol 2002; 20: 500-505.
- 34 Sui G, Soohoo C, Affar EB, Gay F, Shi Y, Forrester WC et al. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. Proc Natl Acad Sci 2002; 99: 5515-5520.
- 35 Yu J-Y, Taylor J, DeRuiter SL, Vojtek AB, Turner DL. Simultaneous inhibition of GSK3α and GSK3β using hairpin siRNA expression vectors. Mol Ther 2003; 7:
- 36 Song E, Lee S-K, Wang J, Ince N, Ouyang N, Min J et al. RNA interference targeting Fas protects mice from fulminant hepatitis. Nat Med 2003; 9: 347-351.
- 37 Naito Y, Ui-Tei K, Nishikawa T, Takebe Y, Saigo K, siVirus; web-based antiviral siRNA design software for highly divergent viral sequences. Nucleic Acids Res 2006; 34 (suppl 2): W448-W450
- 38 Kim YJ. Computational siRNA design considering alternative splicing. Methods Mol Biol 2010: 623: 81-92.
- 39 Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Khvorova A. Rational siRNA design for RNA interference. Nat Biotechnol 2004; 22: 326-330.
- 40 Ui-Tei K, Naito Y, Takahashi F, Haraguchi T, Ohki-Hamazaki H, Juni A et al. Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. Nucleic Acids Res 2004; 32: 936-948.
- 41 Amarzguioui M, Prydz H. An algorithm for selection of functional siRNA sequences. Biochem Biophys Res Commun 2004; 316: 1050-1058.
- 42 Jagla B, Aluner N, Kelly PD, Song D, Volchuka A, Zatorski A et al. Sequence characteristics of functional siRNAs. RNA 2005; 11: 864-872.
- 43 Ishizuka A, Siomi MC, Siomi H. A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. Genes Dev 2002: 16: 2497-2508.
- 44 Khvorova A, Reynolds A, Jayasena SD. Functional siRNAs and miRNAs exhibit strand bias. Cell 2003: 115: 209-216.
- 45 Donis-Keller H. Site specific enzymatic cleavage of RNA. Nucleic Acids Res 1979; 7: 179-192
- 46 Hardin CC, Watson T, Corregan M, Bailey C. Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d (CGCG3GCG). Biochemistry 1992; 31: 8334.
- 47 Nielsen S, Yuzenkova Y, Zenkin N. Mechanism of eukaryotic RNA polymerase III transcription termination. Science 2013; 340: 1577-1580.
- 48 Elbashir SM, Martinez J, Patkaniowska A, Lendeckel W, Tuschl T. Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. EMBO J 2001; 20: 6877-6888.
- 49 Elbashir SM, Lendeckel W, Tuschl T. RNA interference is mediated by 21-and 22nucleotide RNAs. Genes Dev 2001; 15: 188-200.
- 50 Naito Y, Yamada T, Ui-Tei K, Morishita S, Saigo K. siDirect: highly effective, targetspecific siRNA design software for mammalian RNA interference. Nucleic Acids Res 2004: 32(Suppl 2): W124-W129.
- 51 Ui-Tei K, Naito Y, Nishi K, Juni A, Saigo K. Thermodynamic stability and Watson-Crick base pairing in the seed duplex are major determinants of the efficiency of the siRNA-based off-target effect. Nucleic Acids Res 2008; 36: 7100-7109.
- 52 Judge AD, Sood V, Shaw JR, Fang D, McClintock K, MacLachlan I. Sequencedependent stimulation of the mammalian innate immune response by synthetic siRNA. Nat Biotechnol 2005; 23: 457-462.



- 53 Hornung V, Guenthner-Biller M, Bourquin C, Ablasser A, Schlee M, Uematsu S et al. Sequence-specific potent induction of IFN-α by short interfering RNA in plasmacytoid dendritic cells through TLR7. Nat Med 2005; 11: 263–270.
- 54 Kim D-H, Behlke MA, Rose SD, Chang M-S, Choi S, Rossi JJ. Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat Biotechnol* 2005; **23**: 222–226.
- 55 Lan T, Putta MR, Wang D, Dai M, Yu D, Kandimalla ER et al. Synthetic oligoribonucleotides-containing secondary structures act as agonists of Toll-like receptors 7 and 8. Biochem Biophys Res Commun 2009; 386: 443–448.
- 56 Gheidari F, Bakhshandeh B, Teimoori-Toolabi L, Mehrtash A, Ghadir M, Zeinali S. TCF4 silencing sensitizes the colon cancer cell line to oxaliplatin as a common chemotherapeutic drug. *Anticancer Drugs* 2014; 25: 908–916.
- 57 Moazeni M, Khoramizadeh MR, Teimoori-Toolabi L, Noorbakhsh F, Rezaie S. The effect of EFG1 gene silencing on down-regulation of SAP5 gene, by use of RNAi technology. *Acta Med Iran* 2014; **52**: 9–14.
- 58 Moazeni M, Khorramizadeh MR, Teimoori-Toolabi L, Noorbakhsh F, Fallahi AA, Rezaie S. Down-regulation of the ALS3 gene as a consequent effect of RNA-mediated silencing of the EFG1 gene in *Candida albicans. Iran Biomed J* 2012; **16**: 172.
- 59 Gomes-da-Silva LC, Fernández Y, Abasolo I, Schwartz S Jr, Ramalho JS, Pedroso de Lima MC *et al.* Efficient intracellular delivery of siRNA with a safe multitargeted lipid-based nanoplatform. *Nanomedicine (Lond)* 2013; **8**: 1397–1413.
- 60 Kapoor M, Burgess DJ. Efficient and safe delivery of siRNA using anionic lipids: formulation optimization studies. Int J Pharm 2012; 432: 80–90.