


In Silico Design of siRNAs for Silencing CLEC5A Receptor as a Potential Therapeutic Approach Against Dengue and Japanese Encephalitis Virus Infection in Human

Bioinformatics and Biology Insights
Volume 16: 1–14
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/11779322221142122


Tahirah Yasmin, Maisha Adiba, Abdullah Al Saba and AHM Nurun Nabi

Laboratory of Population Genetics, Department of Biochemistry and Molecular Biology, University of Dhaka, Dhaka, Bangladesh.

ABSTRACT: Dengue and Japanese encephalitis virus (JEV) are mosquito-borne RNA viruses that can cause severe illness leading to death in the tropics and subtropics. Both of these viruses interact directly with the C-type lectin domain family 5, member A receptor (CLEC5A) on human macrophages which stimulates the release of proinflammatory cytokines. Since blockade of this interaction has been shown to suppress the secretion of cytokines, CLEC5A is considered a potential target for the development of new treatments to reduce virus-induced brain damage. Developing a vaccine against dengue is challenging because this virus can cause disease through 4 different serotypes. Therefore, the vaccine must immunize against all 4 serotypes to be effective, while unvaccinated people still contract JEV and suffer from its complications. Small interfering RNAs (siRNAs) play an important role in regulating gene expression by causing the degradation of target mRNAs. In this study, we attempted to rationally design potential siRNA molecules using various software, targeting the *CLEC5A* gene. In total, 3 siRNAs were found to be potential candidates for CLEC5A silencing. They showed good target accessibility, optimum guanine-cytosine (GC) content, the least chance of off-target effects, positive energy of folding, and strong interaction with Argonaute2 protein as denoted by a negative docking energy score. In addition, molecular dynamics simulation of the siRNA-Ago2-docked complexes showed the stability of the complexes over 1.5 nanoseconds. These predicted siRNAs might effectively downregulate the expression of the CLEC5A receptor and thus prove vital in the treatment of dengue and JEV infections.

KEYWORDS: CLEC5A, dengue, in silico, Japanese encephalitis, siRNA

RECEIVED: May 09, 2022. **ACCEPTED:** November 09, 2022.

TYPE: Original Research Article

FUNDING: The author(s) received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Tahirah Yasmin, Department of Biochemistry and Molecular Biology, University of Dhaka, Dhaka 1000, Bangladesh. Email: tahirah.yasmin@du.ac.bd

Introduction

The dengue virus (DV), which is becoming a major source of illness and death in tropical and subtropical countries, is now a threat to approximately half of the world's population.¹ This single-stranded RNA virus belongs to the Flaviviridae family² and is transmitted to humans by *Aedes* species mosquitoes. Each year, as many as 400 million people get infected with DV, of which around 100 million show symptoms and yet there is no specific therapy or vaccine available to prevent the infection.³ Developing a vaccine against the virus is challenging since with 4 different serotypes of the virus that can cause the disease, the vaccine must immunize against all 4 types to be effective.⁴ In addition, primary infection by a particular serotype provides life-long immunity only against the infecting serotype and not the other ones. More importantly, secondary infections are associated with an increased risk of developing dengue hemorrhagic fever. It was discovered by Chen et al⁵ that CLEC5A (C-type lectin domain family 5, member A; also known as myeloid DAP12-associating lectin [MDL-1] receptor on monocytes and macrophages) interacts directly with the dengue virion, acting as a signaling receptor for the release of cytokines. The pathogenesis of dengue infection involves massive secretion of proinflammatory cytokines

from macrophages. They also showed that the knockdown of the *CLEC5A* gene alone was sufficient to suppress the release of cytokines from DV-infected macrophages. Thereby it was recognized that CLEC5A silencing could be a promising strategy to avoid the fatal consequences of dengue hemorrhagic fever and dengue shock syndrome.⁵

The same researchers discovered in a second investigation that CLEC5A interacts with the Japanese encephalitis virus (JEV), a flavivirus that is spread by mosquitoes and is thought to be the main cause of viral encephalitis in Asia, where there are an estimated 68 000 clinical cases per year.^{6,7} Although vaccines are available to prevent JEV, unvaccinated individuals are at risk of contracting the infection. During JEV infection, central nervous system (CNS) inflammation is induced which involves neuronal cell death, inhibition of the proliferation and differentiation of neural progenitors, and disruption of the blood-brain barrier (BBB). Anti-CLEC5A antibodies have been demonstrated to protect JEV-infected mice against BBB disruption and limit the infiltration of inflammatory myeloid cells into the brain. This is because the CLEC5A receptor interacts with JEV to activate monocytes and macrophages.⁷ It is critical to establish an efficient curative therapy for JE patients because the case-fatality rate among those with encephalitis can be as high as 30%, and 30% to 50% of those



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

who survive suffer long-term neurological sequelae.⁸ CLEC5A can be considered a potential target to reduce virus-induced brain damage.⁷

Furthermore, a study in 2010 marked CLEC5A as a key regulator of synovial injury and bone erosion during autoimmune joint inflammation. It has been demonstrated that activation of CLEC5A increases the migration of inflammatory neutrophils and macrophages to the joint and encourages bone disintegration, whereas functional inhibition of the receptor decreases the clinical manifestations of autoimmune joint inflammation. These findings suggest that CLEC5A may also be a therapeutic target for the treatment of immune-mediated skeletal disorders.⁹

In 2016, it was also revealed that the CLEC5A receptor interacts with the hemagglutinin protein of influenza viruses and plays once again an important role in the pathogenesis of the virus by enhancing the inflammatory response in myeloid cells. This observation made the receptor a potential therapeutic target against influenza infections as well.¹⁰

RNA interference (RNAi) is an evolutionarily conserved naturally occurring eukaryotic process by which double-stranded RNA-like siRNA (small interfering RNA) and miRNA (microRNA) trigger posttranscriptional gene silencing.¹¹ Small interfering RNA is 20 to 25 base pairs in length and functions by causing mRNA to be broken down after transcription¹² resulting in no translation. Using nanoparticles like polyethyleneimine (PEI) and other methods, RNAi can be introduced into the cell for the knockdown of a gene of interest¹³ and thus may be used for therapeutic purposes like chemical drugs.¹⁴ The gene silencing efficiency of PEI nanoparticles is similar to the commercially available transfecting agent lipofectin but with reduced cytotoxicity.¹⁵ Besides, such nonviral vectors are very effective and safe for targeting diverse tissues in vivo. Since CLEC5A is a human receptor, its expression is only to be transiently silenced as it can be sometimes beneficial to the host in fighting bacterial infections. Therefore, siRNA is preferred to shRNAs (short hairpin RNAs), another regulatory RNA molecule, as siRNAs are transiently expressed in cells, while shRNAs get stably integrated leading to long-term knockdown of the target gene.¹⁶

Therefore, in this study, we set on to rationally design potential siRNA molecules for *CLEC5A* gene silencing. For accomplishing the task of designing potential RNAi candidates, multiple computational web-based tools such as similarity search, secondary structure prediction, RNA interaction evaluation, thermodynamics analysis, molecular docking, and molecular dynamics (MD) simulation were conducted. After initial prediction followed by rigorous filtration and authentication, we finally designed the 3 most potential siRNAs against the CLEC5A receptor which should transiently prevent the translation of the receptor and thus protect against the fatal symptoms of dengue and JEV infections.

Materials and Methods

CLEC5A sequence retrieval and prediction of siRNA

The coding sequence of the human CLEC5A receptor was retrieved from the National Center for Biotechnology Information (NCBI) protein portal (<http://www.ncbi.nlm.nih.gov/>). The entire sequence was used for designing functional and target-specific siRNAs, using the siDirect 2.0 web server (<http://siDirect2.RNAi.jp/>) which is an updated version of the web-based software siDirect.¹⁷ The study of Naito et al¹⁷ has concluded that when the candidate functional siRNAs could form seed-target duplexes with T_m values below 21.5°C, and their 19-nt regions spanning positions 2 to 20 of both strands have at least 2 mismatches to any other nontargeted transcripts, siDirect 2.0 can design at least one qualified siRNA for >94% of human mRNA sequences in RefSeq. Therefore, the maximum melting temperature (T_m value) for the formation of the seed duplex was kept below 21.5°C (the default) for reducing off-target effects and guanine-cytosine (GC) content was set between 30% and 64%.¹⁸ In the server, 3 rules including Ui-Tei,¹⁹ Amarzguoui,²⁰ and Reynolds²¹ based algorithms were selected as parameters for siRNA selection (Table 1).

The predicted siRNAs were cross-validated by the i-Score (inhibitory-Score) Designer server²² in which siRNA prediction scores were calculated from 9 different algorithms including both first-generation algorithms and second-generation algorithms. The second-generation algorithms include i-Score,²² s-Biopredsi,²³ and DSIR²⁴ which are considered more reliable than the other algorithms. In this server, the siRNAs predicted from siDirect were used as input and those which ranked in the top 10 predictions made by any of the following algorithms (s-Biopredsi, i-Score, and DSIR) in i-Score Designer analysis, were selected for further analysis. In this server, the GC content of individual siRNAs was also given as part of the prediction results. Afterwards, the GC content of each of the candidate siRNAs was calculated in OligoCalc.²⁵

Off-target similarity search and filtering siRNAs with palindromic sequence

Both sense and antisense strands of the siRNAs should avoid any secondary structures like palindrome that can be prohibitive to the formation of the duplex properly.²⁶ Therefore, the predicted siRNA sequences were screened for palindromic sequences using the EMBOSS palindrome online tool (www.bioinformatics.nl/cgi-bin/emboss/palindrome) in which the minimum length of the palindrome was set at 4 nucleotides. Next, both strands of the siRNAs were checked via blastN at the NCBI database against the Refseq-RNA database of humans to reduce the risk of silencing unintended genes. While the parameters were adjusted for short-input sequences, word size was set at 7, and the expected threshold was set at 100 to cover random matches as well.

Table 1. Algorithms/rules applied in siDirect 2.0 server for rational designing of siRNA molecules.

UI-TEI RULES	REYNOLDS RULES	AMARZGUIOUI RULES
According to Ui-Tei rules, highly effective RNA interference can occur in mammalian cells if siRNA satisfying the 4 following sequence conditions at the same time is used.	According to Reynolds rules, a score of 6 out of an 8-component algorithm defines the cutoff for selecting siRNAs. The 8 criteria are the following conditions.	According to Amarzguioui rules, a score of 3 represents the lower-end cut-off for selecting siRNAs. The algorithm incorporates the following features.
1. The presence of A/U base at the 5' end of the antisense strand	1. 30%-52% of G/C content	1. A/U content differential between the terminal 3 nucleotides of both ends
2. Presence of G/C at the 5' end of the sense strand	2. The presence of at least 3 A or U bases at positions 15-19 of the sense strand	2. The absence of U at position 1 of the sense strand
3. At least 5 A/U residues in the 5' terminal one-third of the antisense strand	3. The absence of internal repeats	3. The presence of A at position 6 of the sense strand
4. Absence of any GC stretch of more than 9 nucleotides in length	4. The presence of an A base at position 19 of the sense strand	4. Strong attachment of 5' sense strand
	5. The presence of an A base at position 3 of the sense strand	5. Weak binding of 3' sense strand
	6. The presence of a U base at position 10 of the sense strand	6. The absence of G at position 19 of the sense strand
	7. The absence of G or C at position 19 of the sense strand	7. The presence of G or C at position 1 of the sense strand
	8. The absence of G at position 13 of the sense strand	8. The presence of A or U at position 1 of the sense strand
	Satisfaction with criteria I, III, IV, V, and VI earns 1 point. Failure to satisfy the negative criteria (criteria VII and VIII) results in a 1-point decrease. For criterion II, 1 point is added for each A or U base present in positions 15-19 (potential 5 points total)	The presence of a positively correlated feature earns 1 point, while a negatively correlated feature results in a 1-point decrease

Abbreviation: siRNA, small interfering RNA.

The guide and passenger strands of the siRNAs were also screened by the Genome-wide Enrichment of Seed Sequence matches (GESS) online tool (www.flyrnai.org/gess/)²⁷ to further check for any potential off-target effects. The seed regions of the siRNAs were analyzed to identify potential off-targeted transcripts that could produce false-positive results. The input file contained only active RNAi reagents, the region was selected 3' UTR and the reference database was the built-in human database. In a final step, whether the siRNAs contained any immunostimulatory motifs including "GUCCUCAA"²⁸ and "UGUGU"²⁹ that could result in toxicities were also checked.

Secondary structure prediction and thermodynamics study of the siRNA-target gene interaction

The function of a regulatory RNA molecule is determined by its structure mainly. RNA tertiary interactions involve secondary structure elements and are substantially weaker than secondary interactions. Thus, to a large extent, the free energies in the secondary structure represent the thermodynamics of RNA folding.³⁰ In this study, the secondary structures of the guide strand of the siRNAs and their target site as well as their free energy of folding were determined using the MaxExpect

program³¹ of the RNA structure web server (<https://rna.urmc.rochester.edu/RNA>). Here, the RNAi structures with higher values of energy were considered better candidates as those molecules had less chance of folding.

The thermodynamics of RNA-RNA interaction is another major parameter for siRNA efficiency where stronger interaction denoted by lower binding energy indicates better efficacy. For this, the interaction between the guide strand and the target sequence was analyzed by the DuplexFold program of the RNAstructure web server in which 2 RNA sequences are folded into their lowest hybrid free energy conformation along with their free energy of binding being calculated. Afterwards heat capacity and concentration plots for the predicted siRNAs were generated by the DINAMelt application (<http://mfold.rna.albany.edu/?q=DINAMelt/Hybrid2>).³² In the heat capacity plot, the collective heat capacity of the RNA duplex (C_p) is plotted as a function of temperature, and the melting temperature, T_m (C_p) is denoted by the local maximum of the heat capacity curve. The contributions of each species to the ensemble heat capacity are shown in another detailed plot. On the contrary, in the concentration plot, the relative concentration of 5 species (1 heterodimer, 2 homodimers, and 2 single strands) is plotted as a function of temperature, and T_m (conc) indicates

the point at which the concentration of a double-stranded molecule is one-half of its maximum value.

Prediction of siRNA target site accessibility within target mRNA

The siRNA application module of S-fold (<https://sfold.wadsworth.org/cgi-bin/sirna.pl>) was applied to determine the target site accessibility of the predicted siRNA molecules. The target sequences of the siRNAs were given as input and the accessibility of target sites within the target mRNA for specific siRNAs was then assessed by the server through multiple parameters. Several siRNA parameters are being provided by S-fold, including target site accessibility score, differential stability of siRNA duplex ends (DSSE), average internal stability at the cleavage site (AIS), total stability of the siRNA duplex, and so on.

Modeling and docking of argonaute 2 protein and the siRNAs

The 3-dimensional structure of argonaute 2 (Ago2) protein was predicted using SWISS-MODEL (<https://swissmodel.expasy.org/interactive>)³³ based on the crystal structure of RNA-bound human AGO2 protein (Protein Data Bank [PDB] identifier: 4Z4D) via homology modeling. The model was further refined by the GalaxyRefine (<http://galaxy.seoklab.org/refine>) server³⁴ which can improve structure quality beyond the scope of template information, through MD simulation. Next, the quality of the models was checked using ERRAT³⁵ based on the quality scores from ERRAT and GalaxyRefine, and one of the refined models was finally selected.

On the contrary, the 3D structure of the guide RNA was modeled in BIOVIA Discovery Studio 2021 in which the RNA structure bound to the Ago2 protein (PDB identifier: 4Z4D) was taken as a template. The sequence of the template guide RNA was mutated according to the predicted siRNA sequence so that similar conformation could be retained. Afterwards, the geometry of the siRNAs was optimized in Avogadro software³⁶ using Merck molecular force field 94 under the steepest descent algorithm for 1000 steps and $10e^{-7}$ convergence.

Finally, molecular docking between the Ago2 and the siRNAs was performed in the HDock server (<http://hdock.phys.hust.edu.cn/>) which is a hybrid docking algorithm employing both free docking and template-based modeling with high efficacy and robustness.³⁷ The resulting complexes with the lowest docking energy scores and closest siRNA binding position to the template were finally selected and visualized for further analysis.

MD simulation of siRNA (guide strand)—Ago2 protein docked complex

Protein complexes can be dynamic in nature as they interconvert between several alternate configurations. Therefore, MD

simulation is considered a reliable tool for studying the stability of macromolecular complexes, especially how biomolecules will respond upon the addition or removal of a ligand.³⁸ In this study, MD simulation of the siRNA-Ago2 docked complexes was performed with the help of the NAMD 2.14 package.³⁹ For this, the input files were generated using the CHARMM GUI server⁴⁰ where the docked protein complexes were solvated and NaCl ions were added to the simulation box. A periodic boundary condition was incorporated to perform the simulation which was retrieved using VMD 1.9.3 software.⁴¹ In this work, the thermodynamic parameter was set at 310 K. NAMD 2.14 software calculates frame step trajectory files (each frame step corresponds to every time step). Finally, the simulations were evaluated through root mean square deviation (RMSD) and root mean square fluctuation (RMSF) analysis through R package Bio3D developed by Grant laboratory.

The workflow of the entire study is depicted in Figure 1.

Results

siDirect 2.0 and i-score designer predicts siRNAs from the retrieved CLEC5A target sequence

The coding sequence of the human CLEC5A receptor was retrieved from the NCBI-Genbank database, and it consists of 567 nucleotides. siDirect 2.0 predicted in total 29 siRNAs based on the algorithms of Ui-Tei, Reynolds, and Amarzguioui as well as seed-duplex stability. Out of these 29 siRNAs, 14 followed all 3 rules (URA). However, 10 siRNAs were finally selected for further analysis as they had optimum GC content in the range of 30% to 64% for siRNA functionality.⁴² To cross-check, the predicted siRNAs were also analyzed by i-Score Designer, and 6 of them were found to rank among the top 10 predictions made by s-Biopredsi, i-Score, or DSIR prediction tools and so these siRNAs (Table 2) were selected for further filtration.

Filtration of siRNAs for the presence of palindromes, off-target effects, and immunostimulatory action results in 3 potential siRNAs

Four out of the 6 siRNAs did not have any palindromic sequences in their guide and passenger strands. These 4 siRNAs were blasted against the reference RNA database of humans, and 3 of them showed less than 78% query coverage with other genes which are considered tolerable. For further confirming this observation, when these RNAs were analyzed by GESS, the results showed that the off-targets were not significant by any of the tests including Fisher, Bonferroni, Bonferroni step-down (Holm) correction, and Benjamini and Hochberg methods. It suggests that these siRNAs would not create undesirable silencing of any other genes. In addition, the absence of the immunostimulatory motifs within these 3 siRNAs showed that the RNAs would be associated with minimal

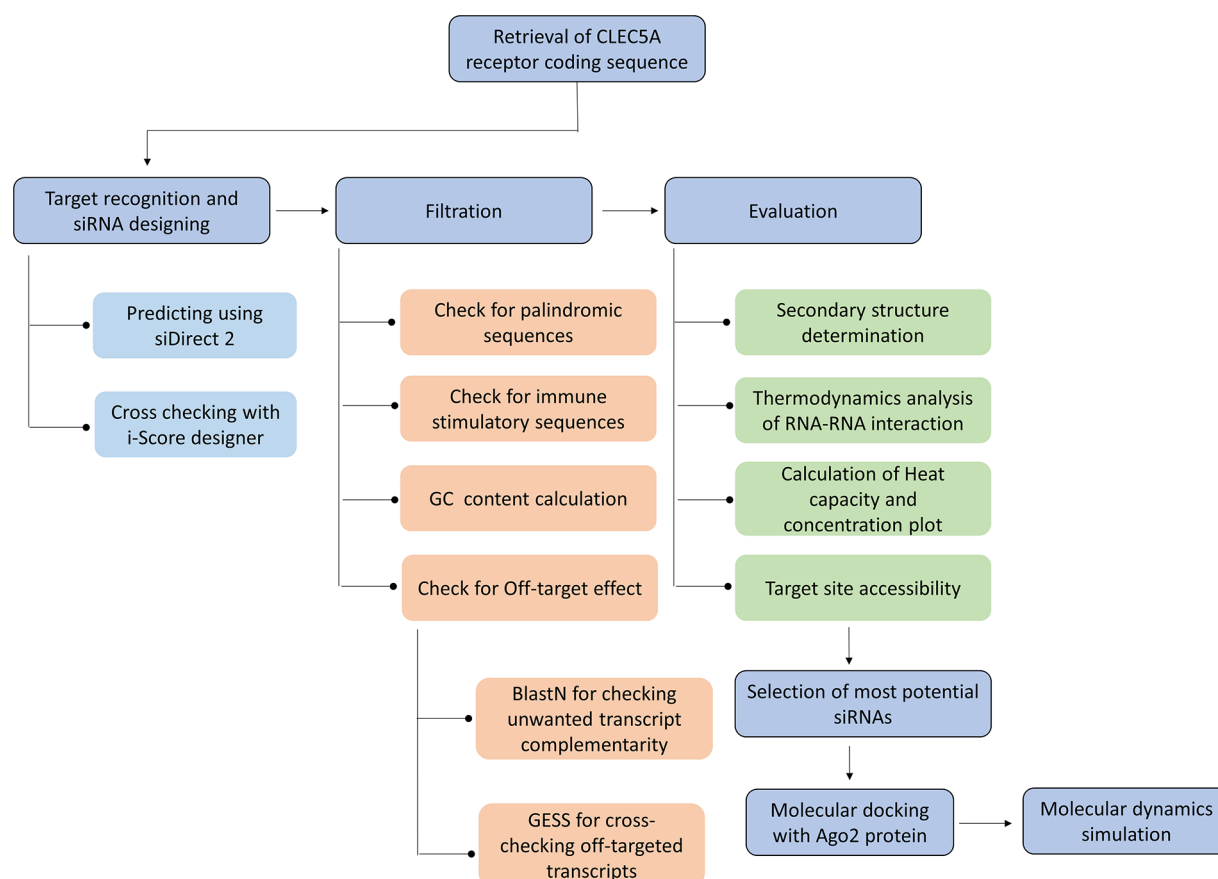


Figure 1. Graphical representation of the workflow used in this study. CLEC5A indicates C-type lectin domain family 5, member A receptor; GESS, genome-wide enrichment of seed sequence matches; siRNA, small interfering RNA.

immune activation. Therefore, these 3 siRNAs were finally selected as potentially most effective siRNAs which were then subjected to further analyses (Table 3).

Secondary structure and thermodynamics study of the candidate siRNAs

The predicted secondary structures of the guide siRNA strands are presented in Figure 2. Their free energy of folding ranged between 1.7 to 1.9 kcal/mol (Table 3). All of the candidate siRNAs displayed promising outcomes, which is consistent with the theory that siRNAs with positive free energies of folding may encourage increased permissibility to the target, improving gene silencing. On the contrary, the binding between the siRNAs and their target gene was also analyzed (Figure 3), and the resulting binding energy ranged between -33.4 and -34.4 kcal/mol (Table 3).

Using the DINAMelt web server,³² the heat capacity plot and concentration plot were also calculated for the RNAi molecules. As temperature increases, the heat capacity of the siRNAs also increases, and the temperature at which the collective heat capacity (C_p) reaches its maximum value is denoted as the melting temperature, T_m (C_p). In the concentration plot, the concentration of the double-stranded molecule decreases with increasing temperature, and the

concentration of folded and unfolded single strands increases with increasing temperature. The temperature at which the concentration of the double-stranded molecule is half of its maximum value is denoted as T_m (Conc). The higher values of these 2 melting temperatures can be considered as an indication of higher efficacy of the siRNAs. From Table 3, it can be seen that among the 3 siRNAs, S1 siRNA showed the highest T_m (C_p) and T_m (Conc) values (87.5°C and 86.6°C , respectively).

Target site accessibility and further RNA duplex thermodynamics study of the siRNAs

The parameters obtained from Sfold output for the 3 finally selected siRNAs are given in Supplementary Table 1. Through multiple experimental approaches, the significance of target accessibility for the function of siRNAs has been proven.⁴³ In this study from Sfold analyses, it was found that S1 siRNA had a target accessibility score of 7, whereas S2 and S3 both had a target accessibility score of 8 (Supplementary Table 1). Furthermore, according to the target accessibility rule, anti-sense siRNA binding energy should be less than or equal to -10 kcal/mol, and from the results, it can be seen that all of the 3 siRNAs fulfilled this criterion ranging between -14.5 and -20.2 kcal/mol.

Table 2. siDirect and i-Score designer predicted overlapped siRNAs along with their different parameters.

TARGET POSITION	TARGET SEQUENCE	GUIDE STRAND	PASSENGER STRAND	SEED-DUPLEX STABILITY TM (°C)		GC CONTENT
				GUIDE STRAND	PASSENGER STRAND	
213-235	CCCCAAGACTGGGAATTTATC	UAAAAUCCCGAGUCUUUGGG	CCAAAGACUGGGAUUUUUAUC	0.4	19.2	43%
214-236	CCCCAAGACTGGGAATTTATCA	AUAAAAUCCCGAGUCUUUGGG	CAAAGACUGGGAUUUUUAUCA	-9.7	17.7	38%
263-285	CTGAATCATCTTGGAAATGAAAGC	UUUCAUUCCCAAGAUGAUUCAG	GAAUCAUCUUGGAAUGAAAGC	16.3	16.2	33%
286-308	AGGGACTTTTGCAGAAAGGAAAAGG	UUUCCUUUUGCAAAAAGUCCCU	GGACUUUUUGCAAAGGAAAAGG	18.7	13.3	38%
437-459	CTGTGTTCAATGGCAATGTTACC	UAACAUUUGCCAUUGAACACAG	GUGUUCAAUGGCAAUUGUUAAC	12.1	20.5	38%
447-469	TGGCAATGTTACCAATCAGAATC	UUCUGAUUGGUACAUAUUGCCA	GCAAUUUUACCAUUCAGAAUC	20.4	12.1	38%

Abbreviation: siRNA, small interfering RNA.

It has also been stated that functional siRNA duplexes tend to have lower stability at the cleavage site as well as on the 5'-antisense end than on the 5'-sense end and the relative stability of the 2 ends influences the degree to which each strand participates in the gene silencing pathway.^{44,45} This difference is shown by the DSSE value (differential stability of siRNA duplex ends, in kcal/mol) which should be more than 0 kcal/mol as the guide strand gets assembled with RNA-induced silencing complex (RISC). The results demonstrate that the values for the proposed siRNAs ranged between 1 and 6 kcal/mol (Supplementary Table 1), demonstrating that ineffective and undesirable mRNA targeting that can occur from the assembly of passenger strands with RISC is avoided. Another score which is siRNA duplex feature score, was found to be 7 for all the siRNAs which can be as low as -2 and as high as 10 and values greater than 6 are acceptable.

Results from Sfold also showed that the AIS values (AIS at the cleavage site) for all the predicted siRNAs were around -6.0 kcal/mol and according to cleavage site instability rule this value should be greater than -8.6 kcal/mol. Average internal stability computes the average internal energy for positions 9 to 14 in the antisense strand (counting from 5' end of antisense strand). In addition, the duplex thermodynamics score of the 3 siRNAs was found to be 2 in which 1 point was attributed to their DSSE values being greater than 0 kcal/mol, and another point was due to their AIS values being greater than -8.6 kcal/mol. Finally, the total score for these siRNA duplexes were found to be either 16 or 17 which surpassed the Sfold-imposed threshold of 12. This score sums up target accessibility score, duplex feature score, and duplex thermodynamics score and its maximum value can be 20. It can be seen that all Sfold filtration criteria have been fulfilled by these 3 siRNAs.

Molecular docking between Ago-2 protein and the siRNAs

The predicted 3D structure of Ago2 protein was refined by GalaxyRefine server which generated 5 models and calculated multiple parameters for each of the model through which the quality of the models could be assessed including MolProbity score, clash score, poor rotamers, Ramachandran favored region, and so on. In addition, the models were checked by ERRAT. All the results are shown in Supplementary Table 2.

Clash score represents the number of atom pairs in a protein model that are unusually close to each other and may result in steric clashes and therefore lower numbers are considered better. Poor rotamers indicate the number of amino acids in the protein with poor side chain geometry and a high percentage of poor rotamers (>1.5%) can be concerning. Another important factor is Ramachandran favored region within the protein which refers to the percentage of amino acids having favored phi/psi angles in the protein backbone. Ideally, this number should be higher than 98%. The next value is MolProbity score which combines all scores regarding atom contacts, geometry,

Table 3. Predicted most effective siRNAs with various parameters.

ALIAS AND GUIDE STRAND SEQUENCE	FREE ENERGY OF FOLDING (KCAL/MOL)	FREE ENERGY OF BINDING (KCAL/MOL)	TM (CONC) (°C)	TM (CP) (°C)
S1(UUUUCCUUUGCAAAAGUCCCU)	1.7	−33.7	86.6	87.5
S2(UAACAUUGCCAUUGAACACAG)	1.8	−33.4	85	86.1
S3(UUCUGAUUGGUAACAUUGCCA)	1.9	−34.4	85.1	85

Abbreviation: siRNA, Small interfering RNA.

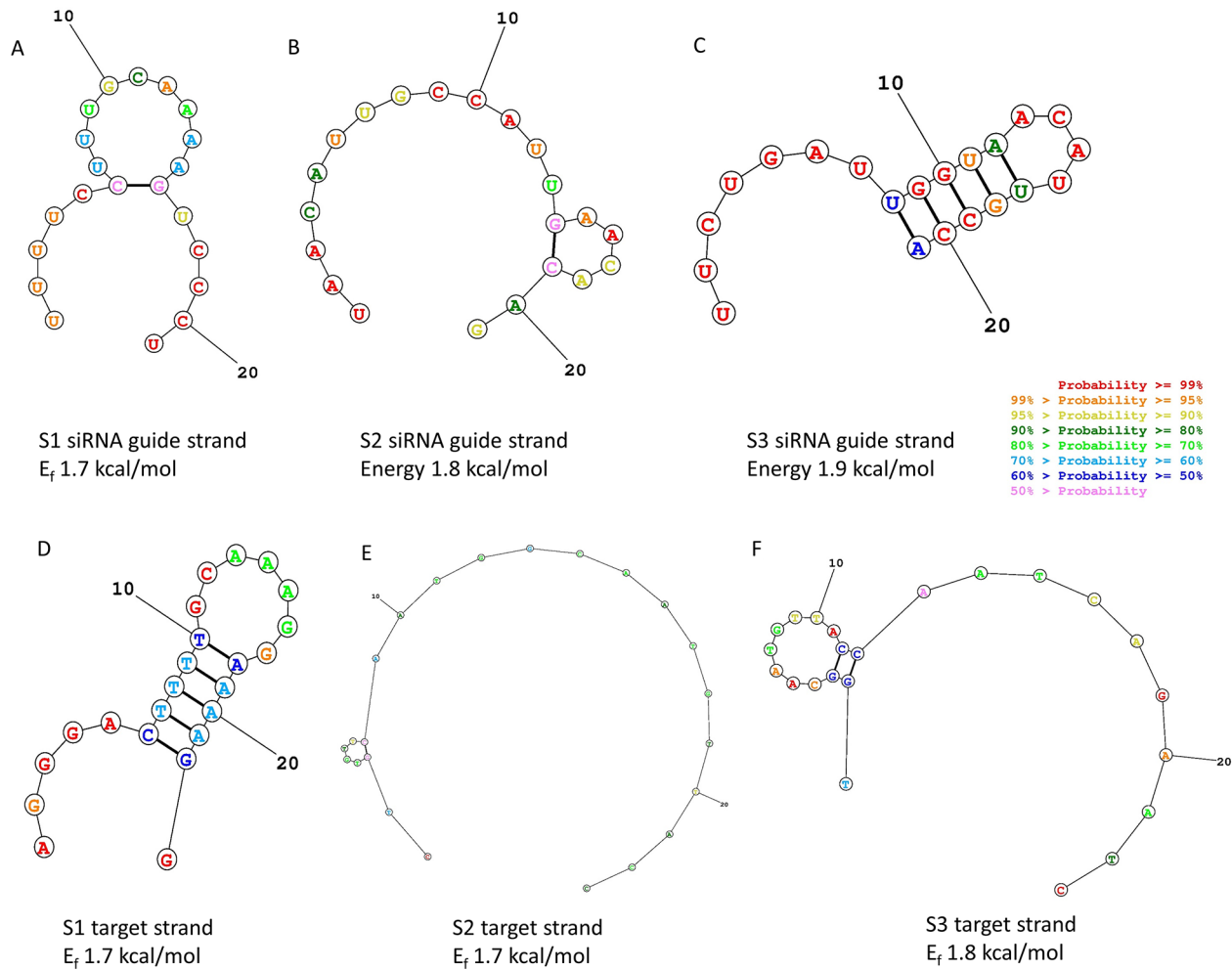


Figure 2. Secondary structures of the guide strands and target sequences of S1, S2, and S3 siRNAs along with their free energy of folding: (A) secondary structure of the guide strand of S1, (B) secondary structure of the guide strand of S2, (C) secondary structure of the guide strand of S3, (D) secondary structure of the target sequence of S1, (E) secondary structure of the target sequence of S2, and (F) secondary structure of the target sequence of S3. siRNA indicates small interfering RNA.

and conformation into a single value to validate structure quality and once again comparatively lower values are desired. Finally, the models were also run through ERRAT program which determines whether the distribution of nonbonded atomic interactions in a candidate structure correlates with the distribution established from a database of reliable, high-resolution structures. Thus, ERRAT can be considered an overall quality factor for nonbonded interactions, with higher scores suggesting higher quality. From the results shown in Supplementary Table 2, it can be seen that model 1 had highest

ERRAT score (95.45%), closest MolProbity score (1.28) to the original model and lowest clash score (11.7) among the 5 models. In addition, it showed less than 1.5% poor rotamers (0.1) and greater than 98% residues in Ramachandran favored region. Therefore, model 1 of Ago2 protein structure was selected for docking with the candidate siRNAs. The structure is given in Supplementary Figure 1.

The docked complexes between Ago2 protein and our candidate siRNAs were screened to find out in which complexes the siRNA was placed inside the PAZ and MID domains of

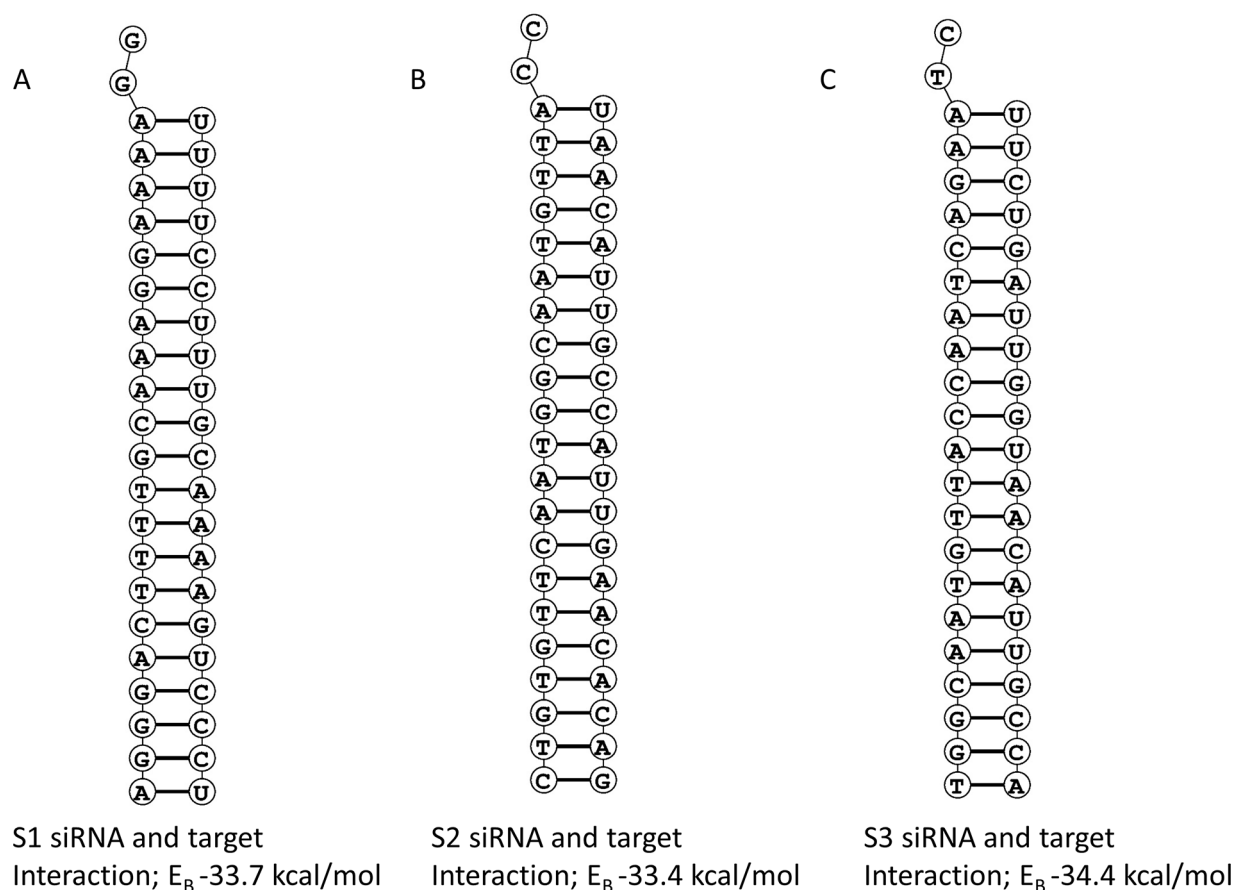


Figure 3. Structure of binding of siRNA (guide strand) and target RNA along with their predicted free energy of binding: (A) structure of interaction between guide strand and target sequence for S1 siRNA, (B) structure of interaction between guide strand and target sequence for S2 siRNA, and (C) structure of interaction between guide strand and target sequence for S3 siRNA. siRNA indicates small interfering RNA.

the Ago2 protein thus maintaining a similar conformation to the PDB 4z4d structure. It was found that for S1 and S3 siRNAs, only one docked complex of each docking analysis showed the involvement of PAZ and MID domains with the siRNAs. On the contrary, for S2 siRNA 2 docked complexes had similar conformation to the 4z4d structure and based on the docking energy score, the complex with the lower docking energy was selected. The docking score, interaction statistics and interactive residues of each of the siRNA docked complexes are presented in Table 4 and the docked complexes are shown in Figure 4.

MD simulation of the siRNA-Ago2 docked complexes

Among the 3 predicted siRNAs, S1 and S2 showed stronger binding interaction with the Ago2 protein compared with S3, although S2 had less hydrogen bond interactions compared to S3 siRNA. S1 and S2 docked complexes were further examined through MD simulations. A 1.5 ns simulation was conducted to observe the deviation of structural dynamics between the S1 and S2 siRNA-Ago2 docked complexes in terms of RMSD and RMSF analysis.

Root mean square deviation represents the measure of deviation from the overlapping of 2 compared structures, and here, each configuration from the MD trajectory file has been compared to the initial configuration structure of the simulation. All the measurements of the RMSD were done based only on the C α backbone atoms. The RMSD graphs are shown in Figure 5 where RMSD values are plotted against 750 frames (each frame corresponds to 2 pico-second [ps]). Ideally, the RMSD should be zero; therefore, the closer the RMSD is to zero, the more equivalent 2 structures are. It can be seen that compared to S2 docked complex, S1 had lower RMSD values. In case of S1, after 1200 ps, the RMSD values started to go above 2.5 Å, while in case of S2, the RMSD started to go beyond 2.5 Å after only 300 ps.

On the contrary, RMSF represents the portions of a protein structure that are fluctuating from their mean structure by analyzing the average deviation of each protein residue over a period of time from a reference position. The RMSF graphs are shown in Figure 6, and it can be seen that once again S1 docked complex was comparatively more stable than S2 docked complex as in S1 complex the RMSF values only reached 3 Å and in S2 complex, the RMSF values went up to almost 5 Å.

Table 4. HDock docking energy score, number of different types of bonds within each docked complex and interactive residues of Ago2 protein in the docked complexes.

SIRNA IN AGO2-SIRNA COMPLEX	HDock DOCKING ENERGY SCORE (KCAL/MOL)	INTERACTION STATISTICS			INTERACTIVE RESIDUES WITHIN PAZ AND MID DOMAIN OF AGO2 PROTEIN
		NUMBER OF HYDROGEN BONDS	NUMBER OF ELECTROSTATIC BONDS	NUMBER OF HYDROPHOBIC BONDS	
S1	-536.15	41	28	12	PAZ: ILE269, ARG277, LYS278, TYR279, TYR311, PHE312, ARG315, HIS316, VAL330, GLY331, GLN332, LYS335, THR337, TYR338, LEU339 MID: LEU522, GLY524, LYS525, THR526, PRO527, TYR529, LYS533, THR544, GLN545, CYS546, VAL547, GLN548, LYS550, ASN551, THR556, GLN558, THR559, ASN562, LEU563, LYS566, LYS570
S2	-464.60	25	23	6	PAZ: TYR279, GLN297, VAL330, GLY331, GLN332, LYS335, THR337 MID: LEU522, GLY524, LYS525, THR526, TYR529, GLN545, CYS546, VAL547, GLN548, LYS550, ASN551, GLN558, THR559, ASN562, LYS566
S3	-432.26	33	23	11	PAZ: ARG277, TYR279, LEU296, TYR311, ARG315, HIS316, VAL330, GLY331, GLN332, LYS335, THR337, LEU339 MID: LEU522, GLY524, LYS525, THR526, PRO527, TYR529, ALA530, LYS533, GLN545, CYS546, VAL547, GLN548, LYS550, ASN551, THR556, GLN558, THR559, ASN562, LEU563, LYS566, LYS570

Abbreviation: siRNA, small interfering RNA.

Discussion

RNA interference is a remarkable endogenous regulatory pathway that can bring about sequence-specific gene silencing. If harnessed effectively, RNAi could result in a targeted therapeutic approach with applications ranging from viral diseases to cancer that are particularly hard to address with existing drugs.^{42,43} The pathways are guided by small RNAs including small interfering RNAs (siRNAs).⁴⁴ In addition to translational inhibition and transcript degradation of target mRNAs that are poorly complementary, these effector molecules also guide sequence-specific cleavage of perfectly complementary mRNAs.⁴⁴ Theoretically, they can be designed for any gene of interest based on its mRNA sequence alone. Moreover, by using nanoparticles like PEI and other approaches, they can be introduced into the cell for specific knockdown of a gene of interest. Such unlimited potential has particularly made RNAi a favorite gene knockdown strategy in mammalian cells.⁴⁵ In this study, we made an attempt through computational methods to rationally design potential siRNA molecules against CLEC5A surface receptor in humans which could lead to new treatment strategies for dengue and JEV infections.

CLEC5A is expressed on the surface of monocytes and macrophages and was first identified by its interaction with DNAX-activating protein of molecular mass 12 kDa (DAP12 or MDL-1).⁵ It contains a C-type lectin-like domain in its C-terminal extracellular region and has recently been shown to

be a macrophage receptor for DV. The interaction results in the activation of macrophages and a significant release of proinflammatory cytokines, which increases the mortality rate of DV-induced infection.⁵ Similar interactions were found between JEV and CLEC5A receptor. After being demonstrated to play a crucial role in the pathogenesis of JEV, CLEC5A is now considered a potential target for new treatments to reduce virus-induced brain damage.⁷

In this study, to narrow the search space and improve prediction accuracy, various criteria for predicting and filtering potential siRNA molecules were applied. For instance, siDirect 2.0 predicts off-target minimized siRNA based on an algorithm that takes into consideration the fact that the capability of siRNA to induce off-target effect is highly correlated to the thermodynamic stability of its seed-target duplex.⁴⁶ For this, the server first selected highly functional siRNA sequences that follow Ui-Tei,¹⁹ Amarzguioui,²⁰ and Reynolds²¹ rules. For instance they had A/U nucleotides at the 5' end of the guide strand; G/C nucleotides at the 5' end of the passenger strand, and at least 4 A/U residues in the 5' terminal 7 base pair of the guide strand.¹⁹ After that, siRNAs whose seed-target T_m was below 21.5°C for both guide and passenger strands, were regarded almost off-target-free seed sequences. Finally, these siRNAs were checked whether they had at least 2 mismatches to any other nontargeted but near-perfect match transcripts, thus further ensuring that no nontargeted gene silencing can

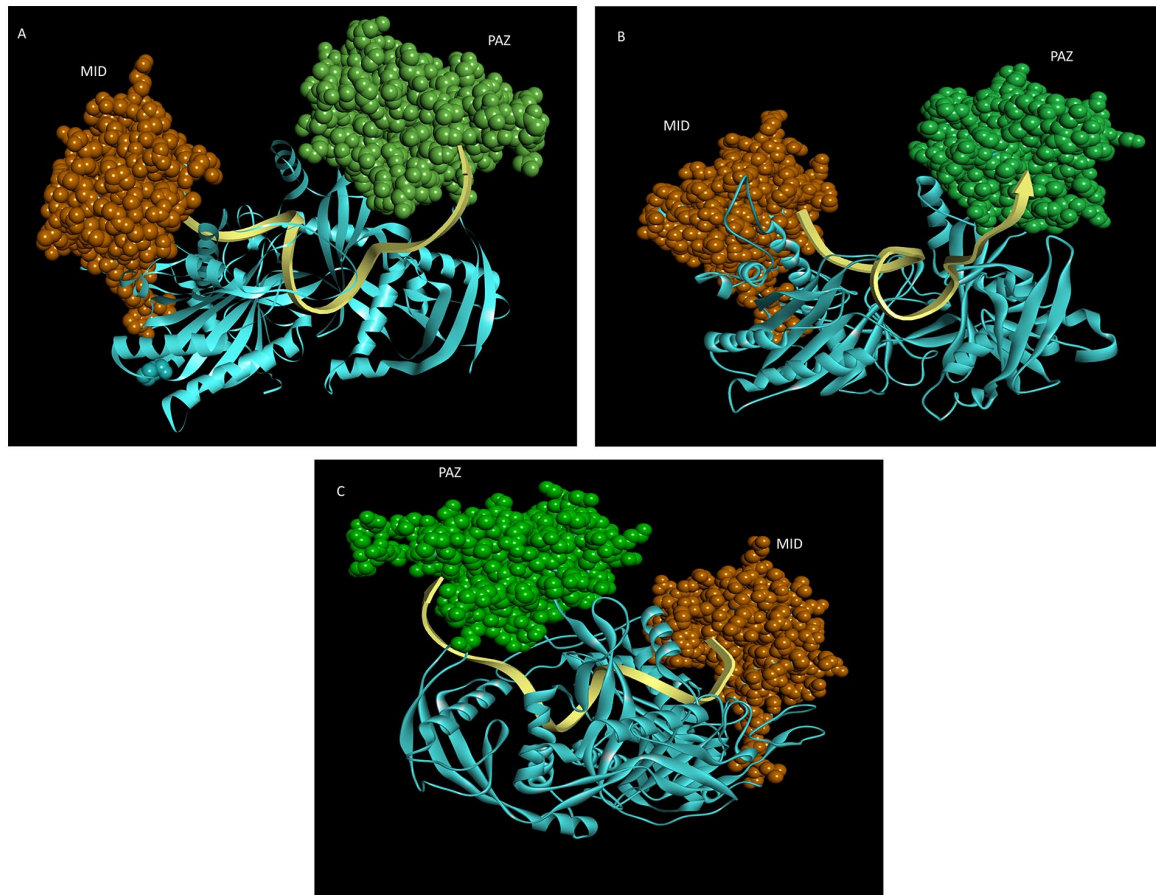


Figure 4. Docked structures of S1, S2, and S3 siRNAs (cartoon view) with Ago2 proteins (surface view). (A) S1 siRNA docked with Ago2 protein (B) S2 siRNA docked with Ago2 protein and (C) S3 siRNA docked with Ago2 protein. The PIWI/Argonaute/Zwille (PAZ) domain and the Middle (MID) domain are colored as green and orange spheres, respectively, and the rest of the protein is denoted by blue color ribbon structure. The siRNAs are represented by yellow arrow structure. siRNA indicates small interfering RNA.

take place. Through these selection steps, siDirect generated and filtrated potential siRNAs.

While these first-generation algorithms heavily depend on base preferences at specific positions, thermodynamic stability, target mRNA secondary structure and several other siRNA features, they can be sometimes not very effective due to their small dataset coverage. However, with the introduction of Huesken's data set that included 2431 experimentally validated siRNAs, several second-generation algorithms came into the scenario that were based on machine learning models and considered more reliable. Three algorithms were included in this work because they were thought to be the most efficient: s-Biopredsi based on artificial neural network,²³ DSIR²² and i-score²⁴ based on linear regression models. Altogether 6 siRNAs were found to give good results in both generations of algorithms and hence predicted to be more efficient during our study.

The GC content of siRNAs is a significant parameter which influences siRNA functionality to a large extent. It has been argued that a high GC content may be a factor that negatively affects the functionality of an siRNA, inhibiting the unwinding of the siRNA duplex, which is necessary for RISC loading.⁴⁷ It might also be associated with a prohibitive secondary structure

of the target mRNA. On the contrary, several studies have also indicated that a low GC content may lead to decreased functionality, presumably due to reduced target affinity and specificity of the siRNA.⁴⁸ For instance, Safari et al⁴⁹ showed that the gene silencing efficiency of an siRNA containing 52.6% GC content was higher than another siRNA with 36.8% GC content. As a result, different ranges of GC content have been suggested by different algorithms such as 31.6% to 57.9%²⁰ and 36% to 52%.²¹ In this work, all the candidate siRNAs showed between 33% and 43% of GC content.

Although a low T_m was set to minimize seed-dependent off-target effects, siRNAs designed by different algorithms might still induce off-target effects. These small RNAs can tolerate several mismatches at the mRNA target and thus can silence those targets with imperfect complementarity. Also as they are nearly identical to the related class of miRNAs, they can recognize and degrade mRNAs that are complementary to their seed region and the rest of the nucleotides' complementarity is insignificant. Therefore, the predicted siRNAs were further checked through BlastN and GESS online tool to ensure their target specificity. BlastN revealed that one of the siRNAs showed 80% query coverage with SMCHD1 gene and hence

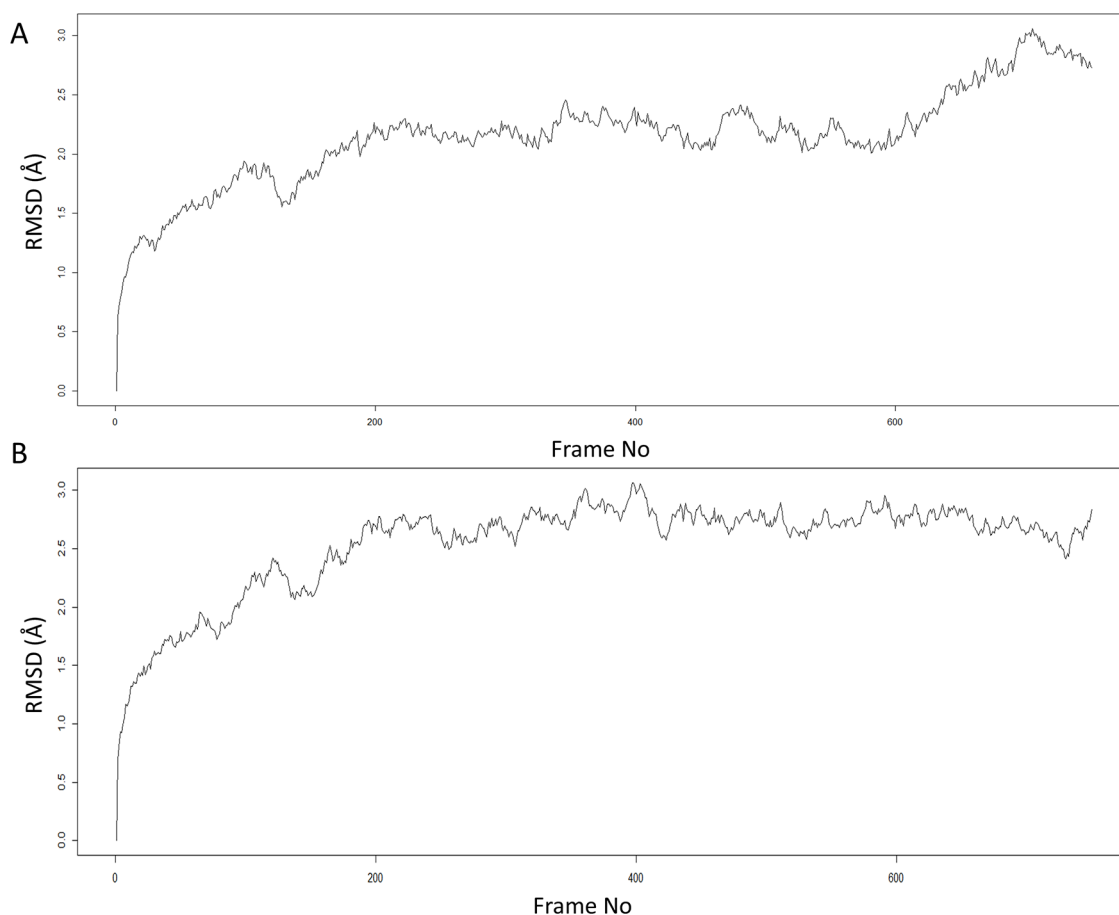


Figure 5. RMSD of C α backbone atoms of the Ago2 protein in siRNA-Ago2 complexes: (A) RMSD plot of S1 siRNA-Ago2 complex and (B) RMSD plot of S2 siRNA-Ago2 complex. Ago2 indicates argonaute 2; RMSD, root mean square deviation; siRNA, small interfering RNA.

filtered out from the candidate list. However, because no 3'UTR regions were discovered to be over-represented for seed region matches, GESS analysis revealed that none of the off-targeted transcripts were significant using the Fisher exact test and Yates chi-square test.²⁷ On the contrary, due to the presence of palindromic motifs within the guide strands, 2 other siRNAs were taken off the list. The remaining final 3 candidates showed they were devoid of immunostimulatory motifs 5'-UGUGU-3' or 5'-GUCCUCAA-3' which could stimulate cytokine production, besides having least possibility of off-target effects and less chance of secondary structure formation like hairpin. When sequences contain "GGGG" or "CCCC" motifs they increase the risk of hairpin structures⁵⁰ and our predicted siRNAs did not have those motifs.

RNA secondary structure significantly influences the efficiency of a siRNA. Self-structure in either the siRNA sequence or the target mRNA at the binding site may prevent gene silencing⁵¹ by inhibiting RISC-mediated cleavage of target. Therefore, predicting secondary structure of the siRNAs and determining free energy of corresponding folding are crucial for designing effective siRNAs and were conducted in this study. Free-energy minimization approach to RNA structure prediction assumes that the lowest free-energy structure will form. All siRNA molecules were discovered to have positive

free energies for folding the guide strands (more than 1.5 kcal/mol), which increased the likelihood that the siRNAs would interact with their targets since they would be less likely to form secondary structures.⁵² To further estimate the efficiency of the predicted siRNAs, their free energy of binding with the target mRNA molecules were also measured. Lower binding energy indicates stronger interaction, and therefore well correlates with a better chance of target inhibition. The effectiveness of the siRNAs is also indicated by the higher values of T_m (Cp) and T_m (Conc) as the thermal stability of the RNA molecules is an important parameter of RNAi activity in its conjugated form.⁵³ Since the T_m values of all the RNAi molecules were above 80°C, it suggested that the molecules were thermally stable and thus efficient. The temperature of the cells may also influence the effectiveness of a particular siRNA since the secondary or local structure of an mRNA can change with temperature changes. One study has found varying effects of the temperature over siRNA efficacy.⁵⁴ While some siRNAs showed no significant changes or slightly increased silencing at lower temperatures, others showed significantly reduced or no silencing when the cells were grown at lower temperatures like 33°C and 35°C. CLEC5A receptor is expressed on tissue macrophages and they are often exposed to elevated tissue temperature during inflammation and fever. Since an increased

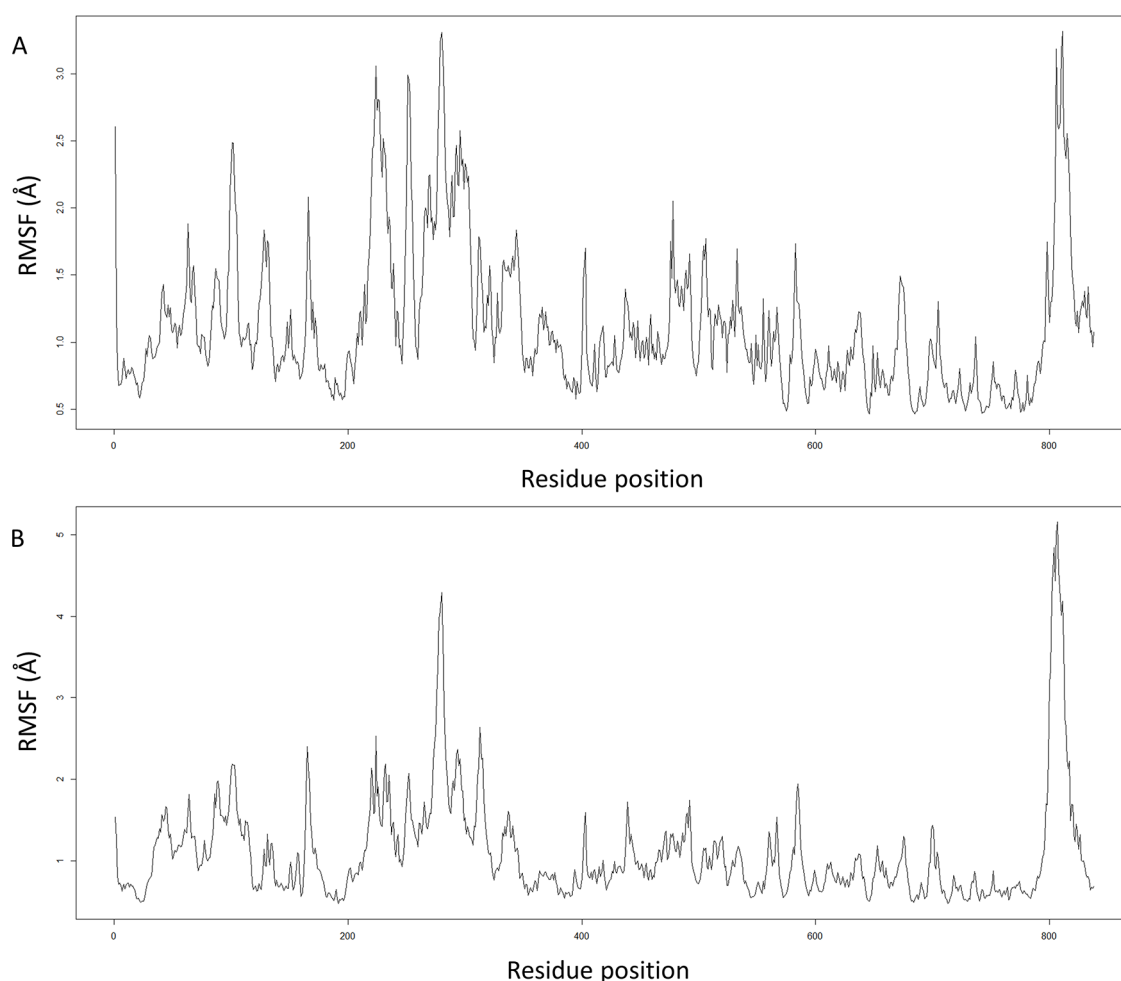


Figure 6. RMSF of each residue of the Ago2 protein in siRNA-Ago2 complexes over 1.5 ns time period: (A) RMSF plot of S1 siRNA-Ago2 complex and (B) RMSF plot of S2 siRNA-Ago2 complex. Ago2 indicates argonaute 2; RMSF, root mean square fluctuation; siRNA, small interfering RNA.

temperature has not been reported to significantly affect gene silencing, it is assumed that the predicted siRNAs from our study would be able to retain their silencing efficiency in the human body.

The downregulation potential of a candidate siRNA crucially depends on its accessibility of the target site within the mRNA molecule.⁵⁵ Hence, this benchmark was also studied in detail in this work using the Sirna module of S-fold. Target site accessibility score ranges between 0 and 8, where 0 indicates antisense siRNA binding energy is greater than -2 kcal/mol and 8 indicates the binding energy is less than -16 kcal/mol. The high target site accessibility score for every siRNA in this study, therefore, indicated stronger target binding (Supplementary Table 1).

The key component of RNAi machinery is the Ago2 protein which plays an important role in the functioning of siRNAs. The interactions between the candidate siRNA molecules and Ago2 protein were analyzed in this study. Human Ago2 protein consists of 4 domains including N-terminal (Asp53-Ser139), PAZ (Pro229-Val347), MID (Gly445-Pro580), and catalytic PIWI (Gln581-Ala859).⁵⁶ The PAZ domain holds the 3'-end of siRNA while the MID domain binds to the 5'-end. It was found that in each of the docked complexes,

ARG714, ARG792, ARG68, SER220, TYR529, CYS546, HIS600, ALA603, SER798, GLN757, LYS566, THR559 residues among others formed hydrogen bonds with the siRNAs. On the contrary, ARG68, ARG97, ARG375, LYS709, ARG761, ARG792, and HIS600 formed electrostatic interactions while LYS525 and CYS66 residues formed hydrophobic bonds with all the siRNAs. Some of these residues including TYR529, Arg761, SER798, ARG792, LYS566, and CYS546 have also been previously reported to interact with other siRNAs and miRNAs.^{56,57}

S1 showed the strongest interaction with the Ago2 protein as revealed by the lowest docking energy score as well as a large number of hydrogen bonds. The score reflects the potential energy change when the protein and ligand come together. A positive scoring function represents higher potential affinity. Since in molecular docking analysis of a complex, the energy of the complex is subtracted from its components, a low and negative score indicates that the complex is more stable than its components by themselves. Therefore, they would have a favorable binding interaction. This means that a very negative score corresponds to a strong binding and a less negative or even positive score corresponds to a weak or nonexistent binding. When the

interaction between S1 siRNA and Ago2 protein was explored, it was found that with the PAZ domain C20 formed 1 and U21 formed 3 hydrogen bonds; G16 formed 1 electrostatic bond and U21 formed 2 hydrophobic bonds. On the contrary, with the MID domain, U1 formed 4 hydrogen bonds, U2 formed 2 hydrogen bonds and U3 formed 1 hydrogen bond, U1 also formed 2 electrostatic and 3 hydrophobic bonds. To further examine the stability of the protein-siRNA complexes, MD simulations were performed on S1 and S2 docked complexes. Performing MD simulations over real experiments allow us to obtain information at the microscopic level and later express it in macroscopic properties.⁵⁸ It was revealed from the RMSD and RMSF analyses that compared to S2, S1 siRNA formed a more stable complex with Ago2 protein which also well correlated with its binding energy and hydrogen bond measurements.

All in all, in this study with the help of multiple computational tools, the most effective and potential siRNA molecules have been designed against the CLEC5A receptor in humans. In the next step, the efficacy of these predicted siRNAs needs to be experimentally validated. The candidate siRNAs may prove instrumental in the downregulation of the viral activity in tissue macrophages through transient silencing of the receptor and therefore may lead to a promising new therapeutic approach against dengue and JEV.

Conclusion

RNA interference molecules were designed through in silico methods against the CLEC5A receptor of humans to silence their expression. We selected 3 siRNAs as final candidates because they showed good target accessibility, positive energy of folding, and strong binding with Argonaute2 protein. The stability of the siRNA-Ago2 docked complexes was also analyzed through MD simulation. The approach may pioneer a vision for the chemical synthesis of novel antiviral therapeutics for the treatment of dengue and JEV at the genomic level. Currently, ONPATRO and GIVLAARI RNAi therapeutics are commercially approved on the market for the treatment of polyneuropathy and acute hepatic porphyria, respectively. We hope this study will help to develop a similar treatment strategy for dengue and JEV infections.

Author Contributions

TY and AHMNN conceived the project. TY collected the data. TY, MA and AAS performed the analyses. TY and AHMNN wrote the manuscript. The manuscript was reviewed by all authors.

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

- Messina JP, Brady OJ, Golding N, et al. The current and future global distribution and population at risk of dengue. *Nat Microbiol*. 2019;4:1508-1515.
- Rodenhuis-Zybert IA, Wilschut J, Smit JM. Dengue virus life cycle: viral and host factors modulating infectivity. *Cell Mol Life Sci*. 2010;67:2773-2786.
- Centers for Disease Control and Prevention (CDC). *Dengue*. Atlanta, GA: CDC; 2021.
- Normile D. Tropical medicine. Surprising new dengue virus throws a spanner in disease control efforts. *Science*. 2013;342:415.
- Chen ST, Lin Y-L, Huang M-T, et al. CLEC5A is critical for dengue-virus-induced lethal disease. *Nature*. 2008;453:672-676.
- WHO. *Japanese encephalitis*. Geneva: WHO; 2019.
- Chen ST, Liu RS, Wu MF, et al. CLEC5A regulates Japanese encephalitis virus-induced neuroinflammation and lethality. *PLoS Pathog*. 2012;8:e1002655.
- Lannes N, Summerfield A, Filgueira L. Regulation of inflammation in Japanese encephalitis. *J Neuroinfl*. 2017;14:158-158.
- Joyce-Shaikh B, Bigler ME, Chao C-C, et al. Myeloid DAP12-associating lectin (MDL)-1 regulates synovial inflammation and bone erosion associated with autoimmune arthritis. *J Exp Med*. 2010;207:579-589.
- Teng O, Chen S-T, Hsu T-L, et al. CLEC5A-mediated enhancement of the inflammatory response in myeloid cells contributes to influenza virus pathogenicity in vivo. *J Virol*. 2017;91:e01813-01816.
- Leung RK, Whittaker PA. RNA interference: from gene silencing to gene-specific therapeutics. *Pharmacol Ther*. 2005;107:222-239.
- Agrawal N, Dasaradhi PV, Mohammed A, Malhotra P, Bhatnagar RK, Mukherjee SK. RNA interference: biology, mechanism, and applications. *Microbiol Mol Biol Rev*. 2003;67:657-685.
- Saengkrit N, Sanitrum P, Woramongkolchai N, et al. The PEI-introduced CS shell/PMMA core nanoparticle for silencing the expression of E6/E7 oncogenes in human cervical cells. *Carbohydr Polym*. 2012;90:1323-1329.
- Geall AJ, Verma A, Otten GR, et al. Nonviral delivery of self-amplifying RNA vaccines. *Proc Natl Acad Sci USA*. 2012;109:14604-14609.
- Nimesh S, Chandra R. Polyethylenimine nanoparticles as an efficient in vitro siRNA delivery system. *Euro J Pharm Biopharm*. 2009;73:43-49.
- Moore CB, Guthrie EH, Huang MT, Taxman DJ. Short hairpin RNA (shRNA): design, delivery, and assessment of gene knockdown. *Methods Mol Biol*. 2010;629:141-158.
- Naito Y, Yoshimura J, Morishita S, Ui-Tei K. siDirect 2.0: updated software for designing functional siRNA with reduced seed-dependent off-target effect. *BMC Bioinformatics*. 2009;10:392.
- Birmingham A, Anderson E, Sullivan K, et al. A protocol for designing siRNAs with high functionality and specificity. *Nat Protoc*. 2007;2:2068-2078.
- Ui-Tei K, Naito Y, Takahashi F, et al. Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. *Nucleic Acids Res*. 2004;32:936-948.
- Amarzguoui M, Prydz H. An algorithm for selection of functional siRNA sequences. *Biochem Biophys Res Commun*. 2004;316:1050-1058.
- Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Khvorovova A. Rational siRNA design for RNA interference. *Nat Biotechnol*. 2004;22:326-330.
- Ichihara M, Murakumo Y, Masuda A, et al. Thermodynamic instability of siRNA duplex is a prerequisite for dependable prediction of siRNA activities. *Nucleic Acids Res*. 2007;35:e123.
- Huesken D, Lange J, Mickanin C, et al. Design of a genome-wide siRNA library using an artificial neural network. *Nat Biotechnol*. 2005;23:995-1001.
- Vert J-P, Foveau N, Lajaunie C, Vandenbrouck Y. An accurate and interpretable model for siRNA efficacy prediction. *BMC Bioinformatics*. 2006;7:520.
- Kibbe WA. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res*. 2007;35:W43-W46.
- Fakhr E, Zare F, Teimoori-Toolabi L. Precise and efficient siRNA design: a key point in competent gene silencing. *Cancer Gene Ther*. 2016;23:73-82.
- Yilmazel B, Hu Y, Sigoillot F, et al. Online GESS: prediction of miRNA-like off-target effects in large-scale RNAi screen data by seed region analysis. *BMC Bioinform*. 2014;15:192.
- Hornung V, Guenther-Biller M, Bourquin C, 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.
- Judge AD, Sood V, Shaw JR, Fang D, McClintock K, MacLachlan I. Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. *Nat Biotechnol*. 2005;23:457-462.
- Ding Y. RNA secondary structure prediction and gene regulation by small RNAs. In: Jianfeng Feng WF, Sun F, eds. *Frontiers in Computational and Systems Biology*. London: Springer- Verlag London Limited; 2010:20-40.
- Lu ZJ, Gloor JW, Mathews DH. Improved RNA secondary structure prediction by maximizing expected pair accuracy. *RNA*. 2009;15:1805-1813.
- Markham NR, Zuker M. DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res*. 2005;33:W577-W581.
- Waterhouse A, Bertoni M, Bienert S, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;46:W296-W303.
- Heo L, Park H, Seok C. GalaxyRefine: protein structure refinement driven by side-chain repacking. *Nucleic Acids Res*. 2013;41:W384-W388.
- Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci*. 1993;2:1511-1519.

36. Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Chem.* 2012;4:17.
37. Yan Y, Zhang D, Zhou P, Li B, Huang S-Y. HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res.* 2017;45:W365-W373.
38. Hollingsworth SA, Dror RO. Molecular dynamics simulation for all. *Neuron.* 2018;99:1129-1143.
39. Phillips JC, Hardy DJ, Maia JDC, et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys.* 2020;153:044130.
40. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem.* 2008;29:1859-1865.
41. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph.* 1996;14:33-827.
42. Gavrilov K, Saltzman WM. Therapeutic siRNA: principles, challenges, and strategies. *Yale J Biol Med.* 2012;85:187-200.
43. Tian Z, Liang G, Cui K, et al. Insight into the prospects for RNAi therapy of cancer. *Front Pharmacol.* 2021;12:644718.
44. Meister G, Tuschl T. Mechanisms of gene silencing by double-stranded RNA. *Nature.* 2004;431:343-349.
45. Aagaard L, Rossi JJ. RNAi therapeutics: principles, prospects and challenges. *Adv Drug Deliv Rev.* 2007;59:75-86.
46. 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.
47. Liu Q, Zhou H, Cui J, Cao Z, Xu Y. Reconsideration of in-silico siRNA design based on feature selection: a cross-platform data integration perspective. *PLoS ONE.* 2012;7:e37879.
48. Liu Q, Zhou H, Zhang K, et al. In silico target-specific siRNA design based on domain transfer in heterogeneous data. *PLoS ONE.* 2012;7:e50697.
49. Safari F, Rahmani Barouji S, Tamaddon AM. Strategies for improving siRNA-induced gene silencing efficiency. *Adv Pharm Bull.* 2017;7:603-609.
50. Hardin CC, Watson T, Corregan M, Bailey C. Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG3GCG). *Biochemistry.* 1992;31:833-841.
51. Mathews DH. Using OligoWalk to identify efficient siRNA sequences. *Methods Mol Biol.* 2010;629:109-121.
52. Chowdhury UF, Sharif Shohan MU, Hoque KI, Beg MA, Sharif Siam MK, Moni MA. A computational approach to design potential siRNA molecules as a prospective tool for silencing nucleocapsid phosphoprotein and surface glycoprotein gene of SARS-CoV-2. *Genomics.* 2021;113:331-343.
53. Berezchnoy A, Brennen R, Bajgelman M, Seales D, Gilboa E. Thermal stability of siRNA modulates aptamer- conjugated siRNA inhibition. *Mol Ther Nucleic Acids.* 2012;1:e51.
54. Russell P, Walsh E, Chen W, Goldwich A, Tamm ER. The effect of temperature on gene silencing by siRNAs: implications for silencing in the anterior chamber of the eye. *Exp Eye Res.* 2006;82:1011-1016.
55. Gilmore IR, Fox SP, Hollins AJ, Sohail M, Akhtar S. The design and exogenous delivery of siRNA for post-transcriptional gene silencing. *J Drug Target.* 2004;12:315-340.
56. Bhandare V, Ramaswamy A. Structural dynamics of human argonaute2 and its interaction with siRNAs designed to target mutant tdp43. *Adv Bioinformatics.* 2016;2016:8792814.
57. Elkayam E, Kuhn C-D, Tocilj A, et al. The structure of human argonaute-2 in complex with miR-20a. *Cell.* 2012;150:100-110.
58. Mulholland AJ. Introduction. Biomolecular simulation. *J Royal Soc Interf.* 2008;5:S169-S172.