Design and Prediction of Potential RNAi (siRNA) Molecules for 3' UTR PTGS of Different Strains of Zika Virus: A Computational Approach

Mohammad Mahfuz Ali Khan Shawan, Md. Mozammel Hossain, Md. Ashraful Hasan, Md. Mahmudul Hasan, Afroza Parvin, Salina Akter, Kazi Rasel Uddin, Subrata Banik, Mahbubul Morshed, Md. Nazibur Rahman and S. M. Badier Rahman

Department of Biochemistry and Molecular Biology, Jahangirnagar University, Savar, Dhaka-1342, Bangladesh
mahfuz_026shawan@yahoo.com, mahfuz_026shawan@juniv.edu

Abstract: Zika virus is an aedes mosquito borne pathogen belonging to the member of Flaviviridae subgroup causes an emerging disease called Zika fever, known as a benign infection usually presenting as influenza like illness with cutaneous rash. Nowadays epidemic outbreak caused by Zika virus is highly contagious and incurable with present technologies; thus considered as a major health risk which need enhanced surveillance. Genetic studies on *Flavivirus* have shown that, the 3' untranslated region (UTR) is consists of seven highly conserved stem loop structure and is important for viral replication and pathogenesis. Therefore, 3' UTR of Zika virus can be utilized as suitable target for controlling Zika virus mediated infection. Viral infection can be reduced by RNA interference (RNAi) technology in which double stranded RNA (siRNA and miRNA) molecules mediate the post transcriptional gene silencing (PTGS) of genes in a sequence specific manner. However genetic variability has been determined in different viral isolates; it is a great challenge to design potential siRNA (small interfering RNA) molecules to repress the expression of respective target gene rather than any other viral gene simultaneously. This work is done using various computational tools to design 21 nucleotides long siRNA sequence on the basis of rational siRNA designing method targeting CDS (coding sequence) of 3' UTR of Zika virus. In this study out of one hundred seventy eight computationally identified siRNAs only four most promising siRNA molecules for gene silencing of 3' UTR of Zika virus were verified using other computer aided tools which might lead to suppress the viral activity. Thus, this approach may provide an insight for chemically synthesized RNA molecules as antiviral agent for Zika virus mediated infection and acts as a foundation stone for an efficient therapeutics at genome level.
[Shawan MMAK, Hossain MM, Hasan MA, Hasan MM, Parvin A, Akter S, Uddin KR, Banik S, Morshed M, Rahman MN and Rahman SMB. **Design and Prediction of Potential RNAi (siRNA) Molecules for 3'UTR PTGS of Different Strains of Zika Virus: A Computational Approach.** *Nat Sci* 2015;13(2):37-50]. (ISSN: 1545-0740). http://www.sciencepub.net/nature. 7

Keywords: 3' UTR, Stem loop, RNAi, PTGS, siRNA, CDS and Computer aided tools

1. Introduction

Zika virus the causative agent of Zika fever is an arbovirus of the *Flavivirius* genus belonging to the Flaviviridae family [1]. It is an enveloped single stranded positive sense RNA virus with 11 kb RNA genome and was first isolated in 1947 from a febrile sentinel rhesus monkey (Rhesus 766) in the Zika Forest of Uganda, Africa [2]. In January 1948 a second isolation was done at the same site from the mosquito *Aedes africanus* [3] while in 1968 isolation was done from human for the first time in Nigeria [1]. Mostly Zika virus infects non-human primates and several *Aedes* species (*Ae.africanus*, *Ae.aegypti* and others) as vectors [2-4] with cyclic epizootics in monkeys [5-8] and maintains a sylvatic cycle. The viral infection can be transmitted to humans by infective mosquito bites which causes fever like syndrome including mild headache, maculopapular rash, fever, malaise, conjunctivitis, and arthralgia [2] however infection may occur through secondary sexual transmission [9]. The pathogenesis of Zika virus involves in infecting dendritic cells near the site of inoculation, and then spread to lymph nodes and the bloodstream. In between 1951-1981, there was evidence of Zika viral infection to human from other African countries including Egypt, Tanzania, Uganda, Gabon, Central African Republic and Sierra Leone as well as in parts of Asia including Thailand, Malaysia, India, Vietnam and Indonesia [1]. In 2007, a large epidemic was caused by Zika virus on Yap Island, Federated States of Micronesia which results in infecting three quarter of local populations [10]. This outbreak shows that, Zika vrus has been detected outside of Africa and Asia, having the potential as an emerging pathogen [9]. The viral illness (Zika fever) is an emerging disease due to the expanding distribution area of Zika virus, confirmed by the recent epidemic affecting French Polynesia, New Caledonian in October, 2013 [11] and Cook Island in February, 2014 [12].

The single stranded positive sense RNA genome of Zika virus is composed of a single open reading frame (ORF) flanked by 5' and 3' UTR. The ORF encodes a polyprotein that is cleaved into three

structural proteins: the capsid (C), premembrane/membrane (prM), and envelope (E); and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, 2K, NS4B, and NS5) [13]. The 5'UTR in flavivirus is 95-101 nucleotides long and are composed of two conserved structural elements, a large stem loop (SLA) and a short stem loop (SLB) [14]. SLA is likely to act as a promoter, and is essential for viral RNA synthesis [15] while SLB mediated interaction between 5'UTR and 3'UTR is crucial for cyclisation of the viral RNA, which is essential for viral replication [16]. The 3'UTR in flavivirus is typically 0.3-0.5 kb long and consist of 7 highly conserved secondary stem loop (SL) structures includes SL-I, SL-II, SL-III, SL-IV, DB1, DB2 and CRE [17, 18]. The SL-II and SL-IV protects the 3'UTR from ribonuclease mediated digestion which is essential for viral replication and virus-induced cytopathicity and pathogenecity [18]. The DB1/DB2 conserved secondary structure play an important role in ensuring efficient translation [17] while CRE (Cis-acting replication element) is essential for the formation of replication complex [19].

RNAi is referred to as an evolutionary conserved gene silencing process that needs double stranded RNA processed into siRNA and miRNA to suppress the expression of homologous RNA (mRNA) targets in a sequence specific manner and thus considered as antiviral therapy [21, 22]. This therapy has already been successful in repressing the expression of genes involved in pathogenic infections as well as genetic disorders. siRNA is a double stranded RNA molecule typically of length between 19 and 25 base pairs with 2 nucleotide overhang on the 3' end. The main role of siRNA is PTGS that mediates the binding of mRNA with siRNA in a sequence specific manner and promote its degradation [23, 24]. The mRNA degradation process induced by siRNA is a highly complex one involving multiple steps initiated with the binding of siRNA with RISC (RNA induced silencing complex) followed by RISC's activation, resulting the recognition of target mRNA and its degradation [25, 26]. As a gene silencing tool for research purpose in the lab, siRNA can also be chemically synthesized and introduced into the cells by direct transfection [27] or delivered into the cells by using nanoparticles and plasmids/viral vectors in the forms of hairpin [28]. After the entry of siRNA duplex within the cell siRNA cleaved by dicer (RNase III-like enzyme) and incorporated into RISC. The passenger (sense) strand is degraded within RISC, while the guide (antisense) strand seeks out and bind with correct target mRNA. The binding of guide strand with appropriate mRNA directs the degradation of mRNA target by using different exo and endo nucleases [29].

Though no remedies are currently available against Zika fever/ clinical Zika viral infection, recent computational approach suggest that Zika viral envelope glycoproteins are most immunogenic and can often be considered as good candidate for vaccine development [2]. Viral infection can be reduced by preventing mosquito bites [20] and the symptoms can only be relieved with non-steroid anti-inflammatories and/or non-salicylic analgesics. So the development of suitable therapeutic molecules against Zika viral infection is very much important and this development has not yet been achieved. As the 3'UTR of Zika virus is very much crucial for viral replication, translation and pathogenecity it can be used as an obligatory target for RNAi mechanism. Therefore, in this study an *in silico* (computational) approach has been made to silence the expression of 3'UTR of Zika virus by the identification of potential siRNA molecules. This *in silico* approach uses a variety of statistical and machine learning programs by help of bioinformatics softwares and machine learning algorithms.

2. Materials and methods

Retrieval and analysis of gene sequence

Thirty seven complete CDS of 3' UTR sequences of different strains of Zika virus were retrieved from viral GenBank database, available at National Centre for Biotechnological Information (http://www.ncbi.nlm.nih.gov/) [30]. All of the genome isolates of Zika virus within the viral database were experimentally established and further used for siRNA designing. The accession number of these thirty seven complete CDS are KF383083.1, KF383082.1, KF383081.1, KF383080.1, KF383079.1, KF383078.1, KF383077.1, KF383076.1, KF383075.1, KF383074.1, KF383073.1, KF383072.1, KF383071.1, KF383070.1, KF383069.1, KF383068.1, KF383067.1, KF383066.1, KF383065.1, KF383064.1, KF383063.1, KF383062.1, KF383061.1, KF383060.1, KF383059.1, KF383058.1, KF383057.1, KF383056.1, KF383055.1, KF383054.1, KF383053.1, KF383052.1, KF383051.1, KF383050.1, KF383049.1, KF383048.1, KF383047.1.

Multiple sequence alignment

All of the thirty seven complete CDS were aligned with each other using clustalW 2.1 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) at GAP open, GAP extension and GAP distances of 10.0, 0.2 and 5.0 respectively [31].

Target recognition and designing of potential siRNA molecules

For target recognition and designing of potential siRNA molecules siDirect 2.0 (http://sidirect2.rnai.jp) tool was used [32] which utilized and considered some rules as parameter like Ui-Tei, Amarzguioui, Renold rules [33] and melting temperature (Tm) should be

below 21.5°C for potential siRNA duplex. Beside these, other factors were also taken on the concept of algorithms Table 1. To verify and confirm the predicted siRNA molecules GeneScript siRNA Target Finder (http://www.genescript.com/index.html), dharma siRNA technology (http://www.dharmacon.com/designcenter/) and siRNA at whitehead (http://sirna.wi.mit.edu/home.php) tools were also used [31].

Table 1. Algorithms or rules for rational design of siRNA molecules

Ui-Tei rules	Amarzguioui rules	Reynolds rules
A/U at the 5' terminus of the sense strand. G/C at the 5' terminus of the antisense strand. At least 4 A/U residues in the 5' terminal 7 bp of sense strand. No GC stretch longer than 9 nt.	Duplex end A/U differential >0. Strong binding of 5' sense strand. No U at position 1. Presence of A at position 6. Weak binding of 3' sense strand. No G at position 19.	Each rule is assigned a score which is summed up to a total duplex score to improve the efficacy of siRNA.

Similarity search and target alignment

To check any off target sequence resemblance in human genomic transcript, blast tool (http://www.ncbi.nlm.nih.gov) [34] was used against whole Genebank database by applying expected threshold value 10 and BLOSUM 62 matrix as parameter. The target sites having similarity of more than 16 adjoining base pairs with human genome were excluded from consideration [35]. Multiple sequence alignment of selected siRNA targets showed that, these sequences were divided into 14 different groups.

Calculation of GC content

GC content of the selected siRNA molecules were analyzed by GC calculator (http://www.genomicsplace.com/cgibin/gc_calculator.pl) [35].

Prediction of secondary structure

Secondary structure and the free energy (ΔG) of folding for selected siRNA molecules at 37°C with target were computed through RNAstructure webserver (http://rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html) [36].

Calculation of RNA-RNA interaction through thermodynamics

To study the thermodynamics of interaction between predicted siRNA (guide strand) and target gene, RNAstructure webserver was used (http://rna.urmc.rochester.edu/RNAstructureWeb/Servers/bifold/bifold.html) [36]. It calculates the hybridization energy and base-pairing form of two RNA sequences. It functions as extension of McCaskill's partition function algorithm to compute probabilities of base pairing, realistic interaction energies and equilibrium concentrations of duplex structures.

Calculation of heat capacity and concentration plot of duplex

For consensus siRNA targets, the heat capacity plot and concentration plot were calculated. The ensemble heat capacity (Cp) was plotted as a function of temperature, with the melting temperature Tm (Cp) indicated. The detailed heat capacity plot also showed the contributions of each species to the ensemble heat capacity, whereas concentration plot-Tm(Conc) is the point at which the concentration of a double stranded molecule of one-half of its maximum value defines the melting temperature Tm(Conc). Both of these were performed by DINAMelt web server (http://mfold.rna.albany.edu/?q=DINAMelt/H ybrid2) [37].

Designing of scrambled siRNA as negative control

Negative control siRNAs (scrambled siRNA) were designed to have no known target in the cells being used. Scrambled siRNA for consensus siRNA molecules were designed by using siRNA Sequence Scrambler (https://www.genscript.com/ssl-bin/app/scramble). Negative control siRNAs were analyzed similarly with siRNA targeting the gene of interest.

Selection of expression vector

The consensus siRNA sequences along with their respective scrambled siRNAs were modified into siRNA inserts by using siRNA Construct Builder (http://www.genscript.com/siRNA_construct_builder.html). The suitable expression vector for the transfer of these siRNAs was selected from Vector-Based siRNAs of siRNA Construct Builder [38]. Using NEBcutter 2 (http://nc2.neb.com/NEBcutter2/), the restriction analysis of selected vector and gene of interest was performed. Flow chart showing the methodology used for screening effective siRNA molecules against 3' UTR of Zika virus Figure 1.

3. Results and discussion

This study was conducted with nucleotide sequence of 3' UTR from thirty seven different strains of Zika virus. Gene sequences available in the viral gene bank database were taken from NCBI and their similarities were analyzed by using clustalW 2.1 which showed that, similarity among sequences was in a range of 81.01-100%. A phylogenetic tree with branch length was established to observe the evolutionary relationship among these strains Figure 2.

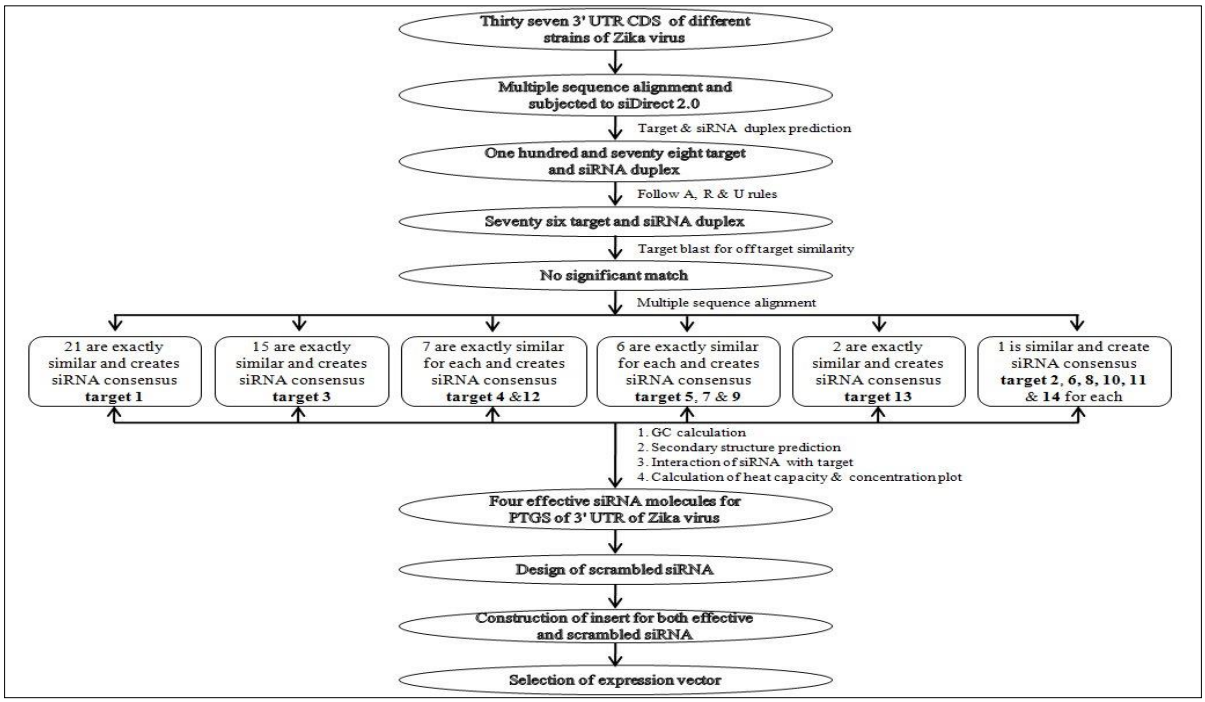


Figure 1. The complete methodology to design and predict effective siRNA molecules against 3' UTR of Zika virus

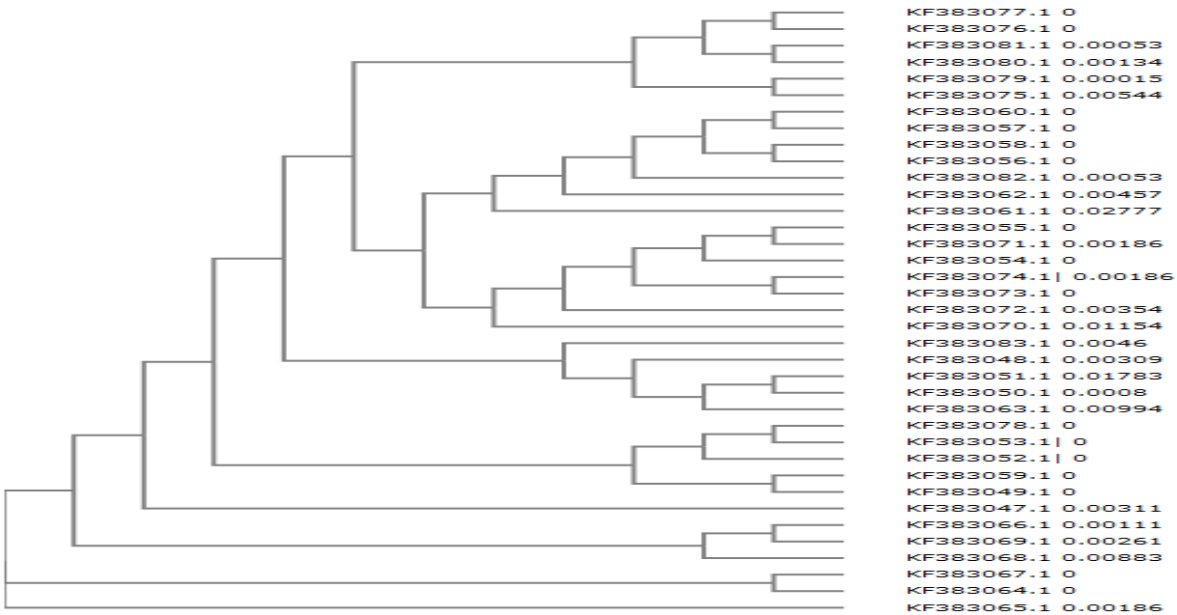


Figure 2. Phylogenetic tree for the nucleotide sequence of 3' UTR from 37 different strains of Zika virus

After that siDirect 2.0 tool was used to identify functional, target-specific siRNA molecules, which significantly reduces off-target silencing. siDirect provided putative siRNA maintaining all rules of Ui-Tei, Amarzguioui and Reynolds. siRNA that follows all the rules and algorithms of Ui-Tei, Amarzguioui and Reynolds are supposed to be most effective [39].

To avoid off target effect Tm should be less than 21.5°C, while the seed-target duplex was calculated using the nearest neighbor model and the thermodynamic parameters for the formation of RNA duplex were also studied [40]. The formula for calculating Tm is:

$\{(1000 \times \Delta H) / (A + \Delta S + R \ln (CT/4))\} - 273.15 + 16.6 \log [Na+]$, Where ΔH (kcal/ mol) is the sum of the nearest neighbor enthalpy change, A is the helix initiation constant (-10.8), ΔS is the sum of the nearest neighbor entropy change [41], R is the gas constant

(1.987 cal/deg/mol) and CT is the total molecular concentration of the strand (100 µM). [Na+] was fixed at 100 mM. Apart from it, to check the accuracy of result Gene script target Finder was also applied and usage statistical modeling method.

According to siDirect result individual 3' UTR gene from thirty seven different strains of Zika virus showed 178 predicted target sequence for siRNA using mixed rule approach i.e. Ui-Tei, Amarzguioui and Reynolds rules. Out of one hundred seventy eight predicted siRNA targets, seventy six were following all the three rules (See supplementary data Table 1). After that, these seventy six siRNA targets were filtered out as possible candidates for further study by subjecting to NCBI Blast tool. Multiple sequence alignment of these seventy six siRNA targets showed that, these sequences were divided into fourteen different groups in which siRNA targets consisting of identical sequence are identified as consensus target Figure 3.

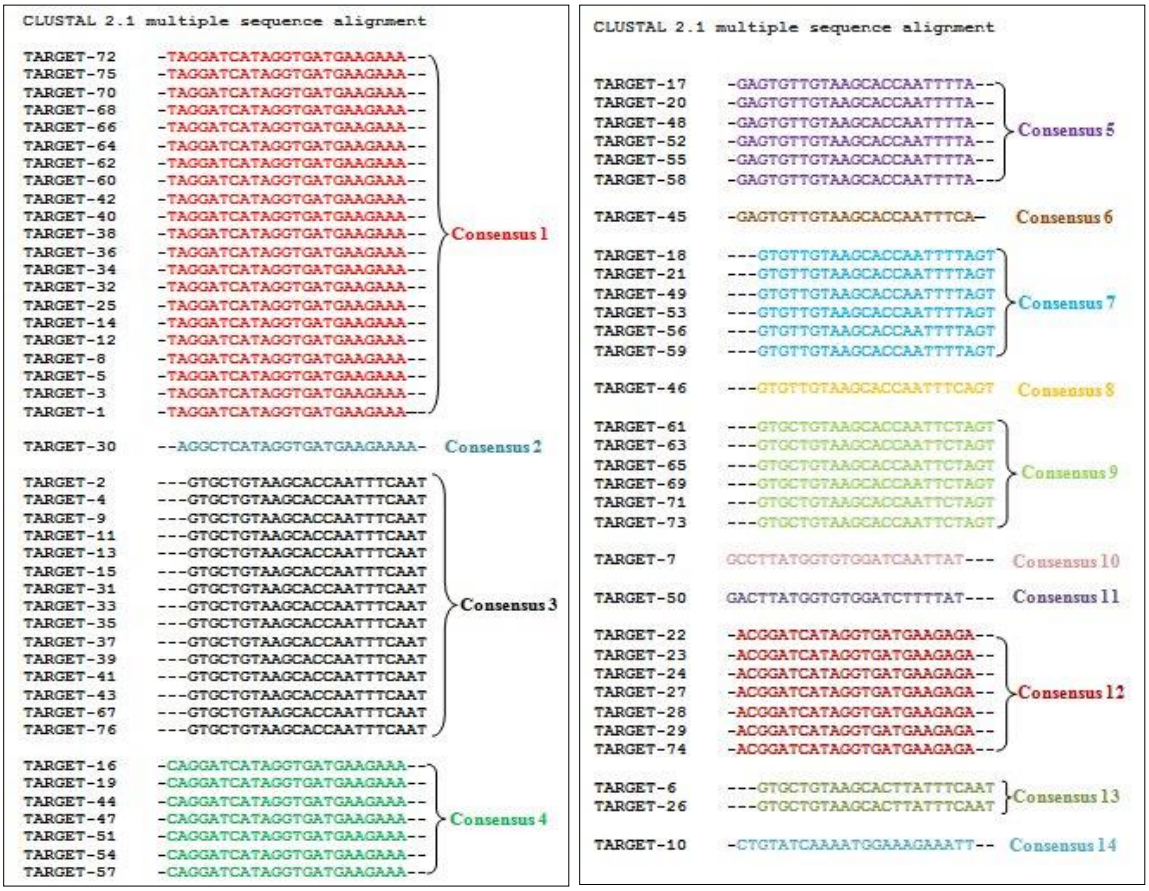


Figure 3. Multiple sequence alignment of predicted siRNA target sequences for 3' UTR of Zika virus showing consensus target 1 to target 14

All the 76 effective siRNA target sequences were selected on the basis of low off target similarity that can be suitable to knockdown the activity of 3' UTR gene and were used for next study with different parameters to determine their perfection. The GC content of the siRNA is an apparent contender for a parameter that might correlate with siRNA functionality. There is a connection between target site accessibility and GC content. It is generally recommended to pick sequences with low GC content (between 31% and 58%), because there is a considerable negative correlation between GC-content and RNAi activity [42-43]. However, there are the incompatible results regarding the effect of GC content and secondary structure on siRNA efficiency. Therefore, these parameters cannot be preferred as a primary determinant of siRNA efficiency. After this analysis among fourteen consensus target (See supplementary data Table 2) only seven of them (Consensus 5, 6, 7, 8, 10, 11 and 13) were holds GC content within 44.4-53.3% and found appropriate for silencing the activity of 3' UTR of Zika virus Table 2.

Though functional RNAi can be determined by predicting the RNA secondary structure, these secondary structures may cause RNAi inhibition and RISC cleavage [44]. For that reason the secondary

structure prediction along with guide strand selection is much more important for effective RNAi [45]. Minimum free energy is considered as benchmark of RNA's secondary structural accuracy [46]. The possible folding and minimum free energy of predicted siRNA molecules (guide strand) for the 3' UTR of Zika virus was done with the online RNAstructure webserver. This server follows most widely used algorithms for RNA secondary structure prediction, which are based on a search for the minimal free energy state [47]. Here, four consensus siRNA target molecules are having more than zero free energy of folding at 37°C Figure 4 and Table 2. Earlier studies have recommended that an RNA molecule should have minimum free energy of folding for their stability [48]. Therefore, the molecule with positive energy may be more accessible for target site and have high potential to bind with target and lead to an effective gene silencing. Other siRNA molecules having less than zero kcal/mol free energy of folding were excluded from the experiment (See supplementary data Table 2).

Computational RNA-RNA binding interaction lies at the center of target prediction algorithm. To predict the thermodynamics of RNA-RNA interaction, the consensus target sequences were subjected to RNAstructure webserver. This server is an ample collection of programs, web services and tools that offer algorithms for RNA folding, assessment and prediction of RNA-RNA interactions and databases related to our work on RNA secondary structures. Bifold one of the important tools of RNAstructure webserver was used to predict free energy of RNA-RNA interaction. It models the binding energy for the interaction at a particular site as, (BE) $\Delta G_{binding} = \Delta G_{uA} + \Delta G_{uB} + \Delta G_h$, Where $\Delta G_{uAB} = (\Delta G_{uA} + \Delta G_{uB})$ is the free energy required to make the binding region in molecule A (target) or B (siRNA) accessible by removing intra-molecular structure. While ΔG_h denotes the free energy gained from forming the intermolecular duplex by the partition function over all structures where the short RNA binds to target region. Calculation of the free energy of interaction (binding) between a siRNA molecule and its target was performed by using the above equation Figure 5, Table 2 and (See supplementary data Table 2). For these 4 consensuses target, multiple sequence alignment and phylogenetic tree with branch length was established to observe the evolutionary relationship within these sequences Figure 6. The entire equilibrium melting profiles were predicted by DINAMelt web server as a function of temperature with melting temperature for a hybridized pair of nucleic acids. The heat capacity (Cp) as a function of temperature and Tm(Conc) along with their heat capacity plot and concentration plot is given in Table 2 and Supplementary data Figure 1 respectively.

Scrambled siRNA sequences for four consensus targets were designed by siRNA Sequence scrambler which uses sense siRNA target sequence. These sequences are important for distinguishing sequence-specific silencing from non-specific effects in the RNAi experiment. Neither the mRNA nor protein level of the experimental gene should be affected by the negative control siRNA. The best scrambled siRNA candidate for Consensus 7 is 5′ AGATACTTACGTCTTAGGTTA 3′, Consensus 8 is 5′ AGCTGTTACGTGTAACATTAC 3′, Consensus 11 is 5′ GGTTATTATATCGTGTGTTCT 3′ and Consensus 14 is 5′ ACAAGAATGTAGACGTTATATAA 3′.

siRNA inserts were generated for four consensus siRNA Figure 7 along with scrambled siRNAs Figure 8 by using siRNA Construct Builder which includes the antisense region, sense loop, termination signal and restriction enzyme digestion site. In effective gene therapy, therapeutic genes can be targeted to the desired cell with the support of safe vectors under the control of suitable promoter [49]. The therapeutic applications of siRNA have been hampered by their instability, poor cellular uptake and mainly the lack of efficient delivery methods. Therefore, development of the carriers with the capacity to stabilize siRNAs and facilitate their uptake by target cell is important [50]. The third generation adeno viral vectors are replication-defective by the deletion of E1 and/or E3 genes which results in the cessation of viral DNA replication and viral capsid protein production. These vectors are constructed by the deletion of other viral genes and the latest have most of the viral genes removed, leaving a vector that is termed "gutless". These gutless vectors contain only the inverted terminal repeats and a packaging sequence surrounding the transgene [51]. In this study, pRNATin-H/1.4-Retro is the expression vector obtained for the integration of siRNA to silence the 3' UTR of Zika virus. It is an adenoviral vector from Genscript siRNA technology containing 8624 bp. Restriction analysis was done by NEBcutter 2 to check the compatibility of gene of interest and vector, which is performed by selecting the enzyme which has minimum restriction site. Xho I and Mlu I are the restriction enzymes used for the insertion of siRNA to the vector. The siRNA insert is located within 5288-5311 bp and very near to the H1.4 promoter, which placed between the promoter and hygromycin gene Figure 9.

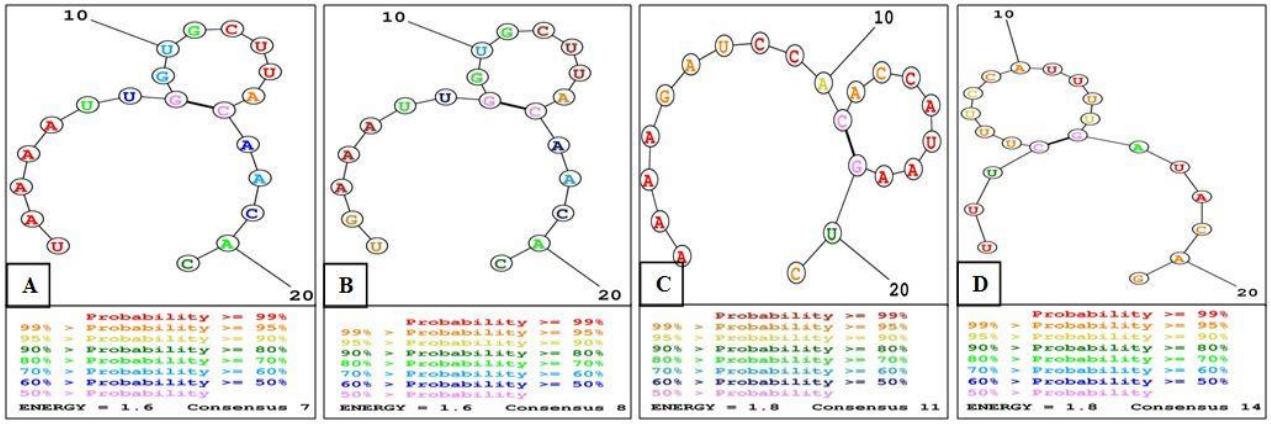


Figure 4. Predicted siRNA 2° structure with possible folding and minimum free energy of (A) Consensus 7, (B) Consensus 8, (C) Consensus 11 and (D) Consensus 14

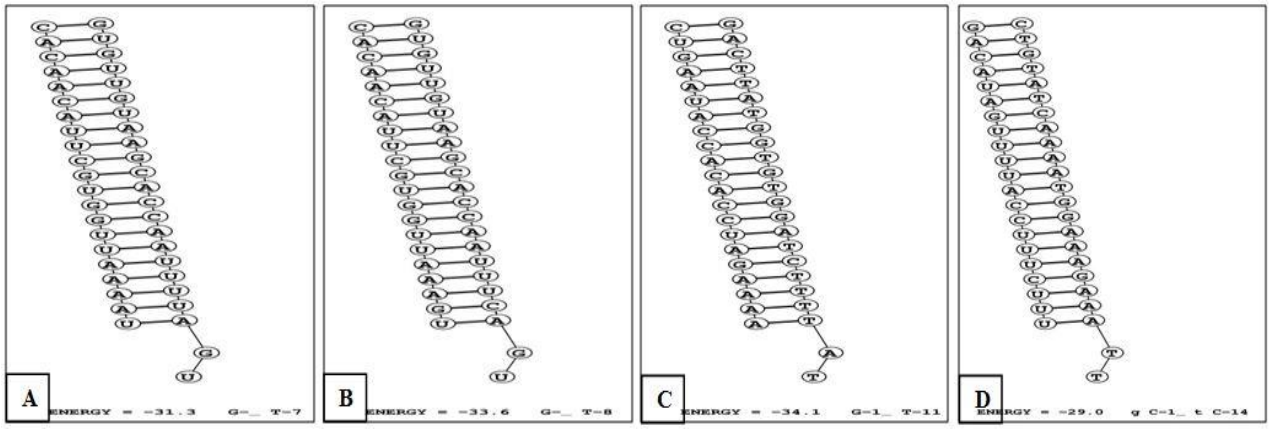
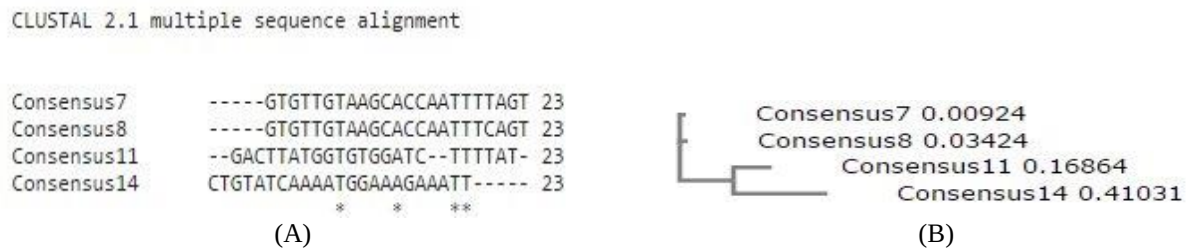


Figure 5. Predicted lowest free energy structure for binding of siRNA (guide strand) and target RNA of (A) Consensus 7, (B) Consensus 8, (C) Consensus 11 and (D) Consensus 14

Table 2. Four effective consensus siRNA molecules for 3' UTR of Zika virus with GC%, free energy of folding and free energy of binding with target

Target	Location of target within mRNA	siRNA target within consensus target 21nt target + 2nt overhang (5′→3′)	Predicted siRNA duplex at 37°C 21nt guide (5′→3′) 21nt passenger (5′→3′)	GC%	ΔG of folding (kcal/ mol)	ΔG of binding (kcal/ mol)	TmCp (°C)	TmConc. (°C)
Consensus Target-7	271-293	GTGTTGTAAGCACCAATTTTAGT	UAAAAUUGGUGCUUACAACAC GUUGUAAGCACCAAUUUUAGU	46.7	1.6	-31.3	80.7	79.5
Consensus Target-8	271-293	GTGTTGTAAGCACCAATTTCAGT	UGAAAUUGGUGCUUACAACAC GUUGUAAGCACCAAUUUCAGU	53.3	1.6	-33.6	81.2	80.2
Consensus Target-11	103-125	GACTTATGGTGTGGATCTTTTAT	AAAAGAUCCACACCAUAAGUC CUUAUGGUGUGGAUCUUUUAU	44.4	1.8	-34.1	82.7	81.4
Consensus Target-14	59-81	CTGTATCAAAATGGAAAGAAATT	UUUCUUUCCAUUUUGAUACAG GUAUCAAAAUGGAAAGAAAUU	60.0	1.8	-29.0	70.1	69.8



(A) (B)

Figure 6. Showing (A) multiple sequence alignment of three predicted siRNA target sequences; (B) phylogenetic tree showing the evolutionary relationship among these three consensus siRNA target sequences

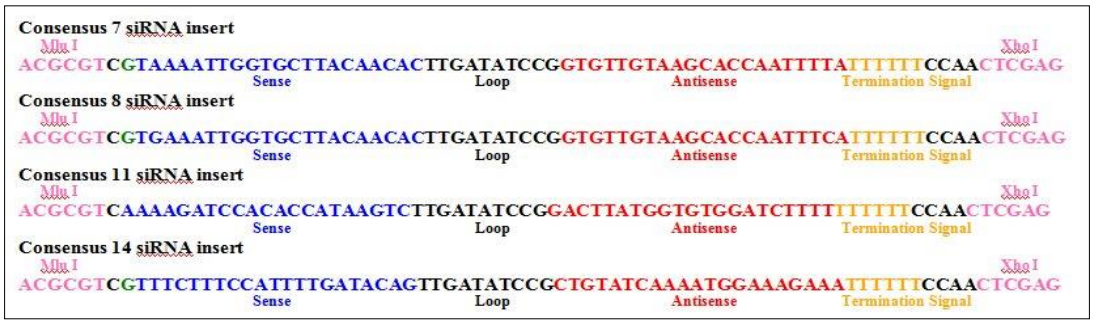


Figure 7. The siRNA inserts of Consensus 7, 8, 11 and 14 generated by siRNA Construct Builder

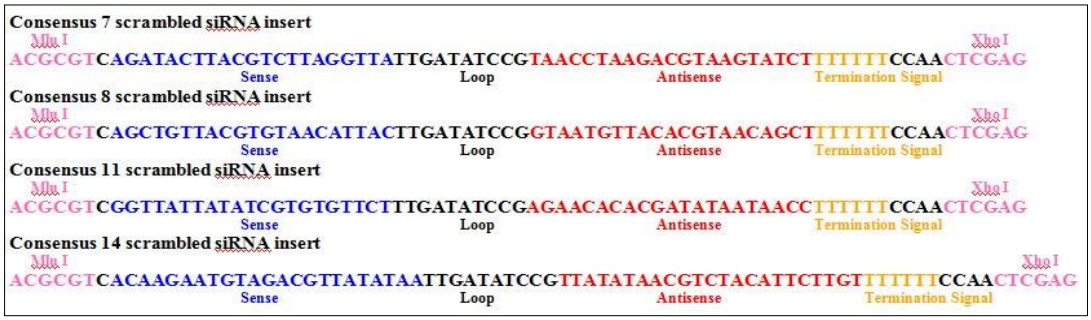


Figure 8. The scrambled siRNA inserts of Consensus 7, 8, 11 and 14 generated by siRNA Construct Builder

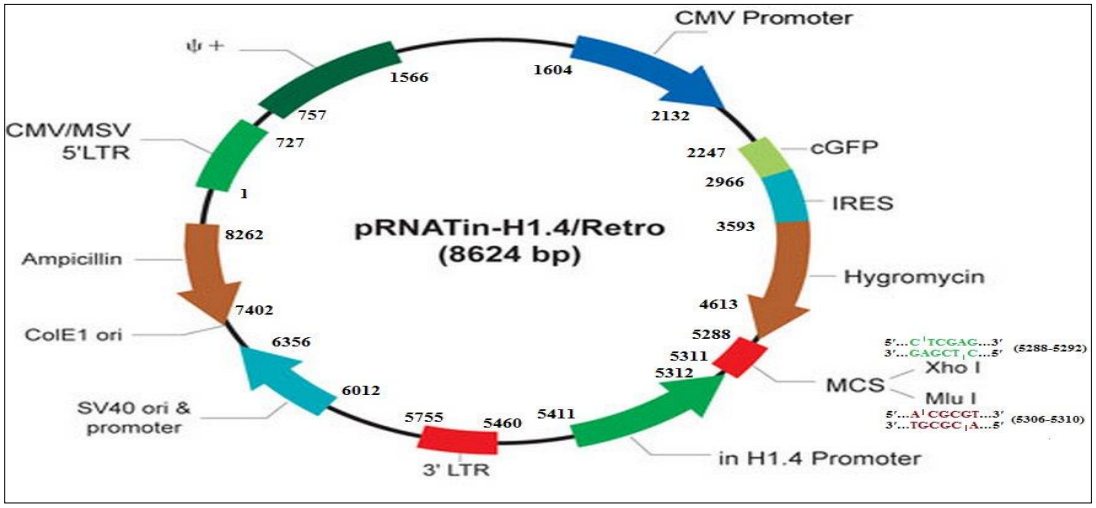


Figure 9. Expression vector pRNATin-H/1.4-Retro for the integration of siRNA against 3' UTR of Zika virus

Supplementary data

Table 1. Predicted siRNA target for 3' UTR of Zika Virus

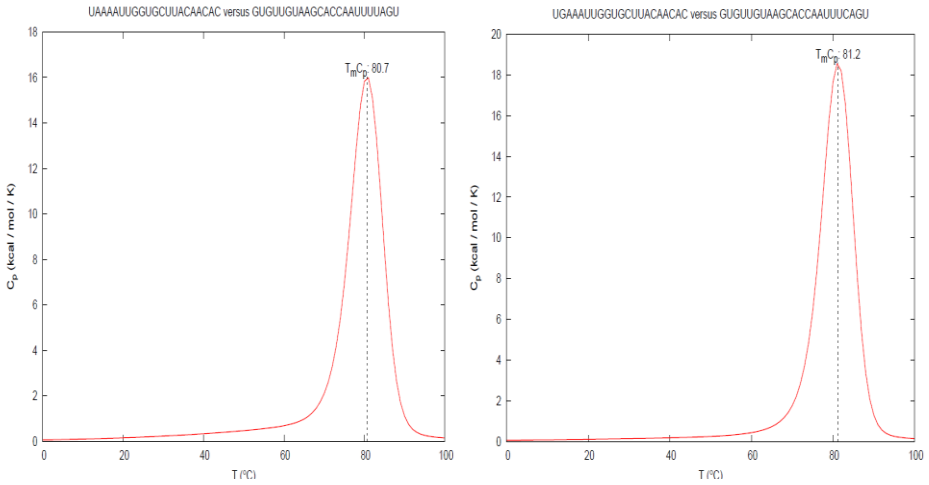
Accession Number	Target	Location of target within the gene	siRNA target sequence within gene
KF383047.1	Target 1	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 2	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383048.1	Target 3	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 4	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383049.1	Target 5	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 6	271-293	GTGCTGTAAGCACTTATTTCAAT
KF383050.1	Target 7	103-125	GCCTTATGGTGTGGATCAATTAT
	Target 8	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 9	271-293	GTGCTGTAAGCACCAATTTCAAT

KF383051.1	Target 10	59-81	CTGTATCAAAATGGAAAGAAATT
	Target 11	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383052.1	Target 12	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 13	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383053.1	Target 14	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 15	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383054.1	Target 16	186-208	CAGGATCATAGGTGATGAAGAAA
	Target 17	269-291	GAGTGTTGTAAGCACCAATTTTA
	Target 18	271-293	GTGTTGTAAGCACCAATTTTAGT
KF383055.1	Target 19	186-208	CAGGATCATAGGTGATGAAGAAA
	Target 20	269-291	GAGTGTTGTAAGCACCAATTTTA
	Target 21	271-293	GTGTTGTAAGCACCAATTTTAGT
KF383056.1	Target 22	186-208	ACGGATCATAGGTGATGAAGAGA
KF383057.1	Target 23	186-208	ACGGATCATAGGTGATGAAGAGA
KF383058.1	Target 24	186-208	ACGGATCATAGGTGATGAAGAGA
KF383059.1	Target 25	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 26	271-293	GTGCTGTAAGCACTTATTTCAAT
KF383060.1	Target 27	186-208	ACGGATCATAGGTGATGAAGAGA
KF383061.1	Target 28	186-208	ACGGATCATAGGTGATGAAGAGA
KF383062.1	Target 29	186-208	ACGGATCATAGGTGATGAAGAGA
KF383063.1	Target 30	187-209	AGGCTCATAGGTGATGAAGAAAA
	Target 31	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383064.1	Target 32	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 33	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383065.1	Target 34	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 35	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383066.1	Target 36	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 37	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383067.1	Target 38	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 39	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383068.1	Target 40	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 41	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383069.1	Target 42	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 43	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383070.1	Target 44	186-208	CAGGATCATAGGTGATGAAGAAA
	Target 45	269-291	GAGTGTTGTAAGCACCAATTTCA
	Target 46	271-293	GTGTTGTAAGCACCAATTTCAGT
KF383071.1	Target 47	186-208	CAGGATCATAGGTGATGAAGAAA
	Target 48	269-291	GAGTGTTGTAAGCACCAATTTTA
	Target 49	271-293	GTGTTGTAAGCACCAATTTTAGT
KF383072.1	Target 50	103-125	GACTTATGGTGTGGATCTTTTAT
	Target 51	186-208	CAGGATCATAGGTGATGAAGAAA
	Target 52	269-291	GAGTGTTGTAAGCACCAATTTTA
	Target 53	271-293	GTGTTGTAAGCACCAATTTTAGT
KF383073.1	Target 54	186-208	CAGGATCATAGGTGATGAAGAAA
	Target 55	269-291	GAGTGTTGTAAGCACCAATTTTA
	Target 56	271-293	GTGTTGTAAGCACCAATTTTAGT
KF383074.1	Target 57	186-208	CAGGATCATAGGTGATGAAGAAA
	Target 58	269-291	GAGTGTTGTAAGCACCAATTTTA
	Target 59	271-293	GTGTTGTAAGCACCAATTTTAGT
KF383075.1	Target 60	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 61	271-293	GTGCTGTAAGCACCAATTCTAGT
KF383076.1	Target 62	186-208	TAGGATCATAGGTGATGAAGAAA

	Target 63	271-293	GTGCTGTAAGCACCAATTCTAGT
KF383077.1	Target 64	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 65	271-293	GTGCTGTAAGCACCAATTCTAGT
KF383078.1	Target 66	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 67	271-293	GTGCTGTAAGCACCAATTTCAAT
KF383079.1	Target 68	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 69	271-293	GTGCTGTAAGCACCAATTCTAGT
KF383080.1	Target 70	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 71	271-293	GTGCTGTAAGCACCAATTCTAGT
KF383081.1	Target 72	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 73	271-293	GTGCTGTAAGCACCAATTCTAGT
KF383082.1	Target 74	186-208	ACGGATCATAGGTGATGAAGAGA
KF383083.1	Target 75	186-208	TAGGATCATAGGTGATGAAGAAA
	Target 76	271-293	GTGCTGTAAGCACCAATTTCAAT

Table 2. Consensus siRNA molecules for for 3' UTR of Zika virus with GC%, free energy of folding and free energy of binding with target

Target	Location of target within mRNA	siRNA target within consensus target 21nt target +2nt overhang (5′→3′)	Predicted siRNA duplex at 37°C 21nt guide (5′→3′) 21nt passenger (5′→3′)	GC%	ΔG of folding (kcal/ mol)	ΔG of binding (kcal/ mol)	TmCp (°C)	TmConc. (°C)
Consensus Target-1	186-208	TAGGATCATAGGTGATGAAGAAA	UCUUCAUCACCUAUGAUCCUA GGAUCAUAGGUGAUGAAGAAA	61.5	-0.2	-35.9	82.9	81.7
Consensus Target-2	187-209	AGGCTCATAGGTGATGAAGAAAA	UUCUUCAUCACCUAUGAGCCU GCUCAUAGGUGAUGAAGAAAA	69.2	-0.7	-37.5	85.1	84.0
Consensus Target-3	271-293	GTGCTGTAAGCACCAATTTCAAT	UGAAAUUGGUGCUUACAGCAC GCUGUAAGCACCAAUUUCAAU	60.0	-3.9	-36.0	84.0	83.9
Consensus Target-4	186-208	CAGGATCATAGGTGATGAAGAAA	UCUUCAUCACCUAUGAUCCUG GGAUCAUAGGUGAUGAAGAAA	69.2	-0.2	-37.2	82.9	81.7
Consensus Target-5	269-291	GAGTGTTGTAAGCACCAATTTTA	AAAUUGGUGCUUACAACACUC GUGUUGUAAGCACCAAUUUUA	53.3	-1.5	-32.9	83.2	82.0
Consensus Target-6	269-291	GAGTGTTGTAAGCACCAATTTCA	AAAUUGGUGCUUACAACACUC GUGUUGUAAGCACCAAUUUCA	53.3	-1.5	-32.9	83.1	81.9
Consensus Target-9	271-293	GTGCTGTAAGCACCAATTCTAGT	UAGAAUUGGUGCUUACAGCAC GCUGUAAGCACCAAUUCUAGU	60.0	-3.9	-36.4	84.9	84.7
Consensus Target-10	103-125	GCCTTATGGTGTGGATCAATTAT	AAUUGAUCCACACCAUAAGGC CUUAUGGUGUGGAUCAAUUAU	52.9	-0.3	-36.4	84.1	82.9
Consensus Target-12	186-208	ACGGATCATAGGTGATGAAGAGA	UCUUCAUCACCUAUGAUCCGU GGAUCAUAGGUGAUGAAGAGA	69.2	-0.2	-37.1	83.4	82.3
Consensus Target-13	271-293	GTGCTGTAAGCACTTATTTCAAT	UGAAAUAAGUGCUUACAGCAC GCUGUAAGCACUUAUUUCAAU	50.0	-4.1	-34.0	81.6	82.4



(A) **(B)**

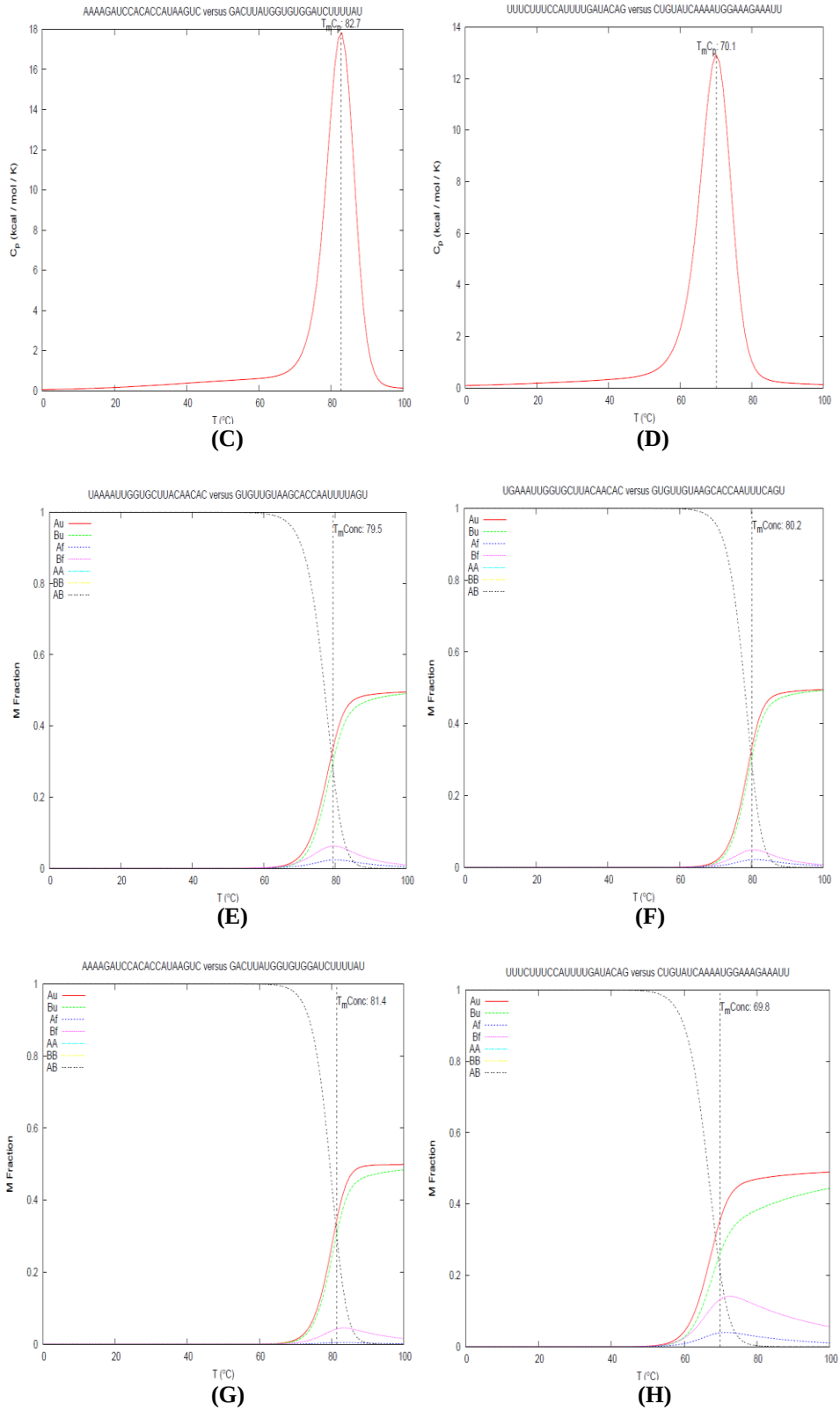


Figure 1. Heat capacity plots (A) Consensus 7, (B) Consensus 8, (C) Consensus 11 and (D) Consensus 14; Concentration plots (E) Consensus 7, (F) Consensus 8, (G) Consensus 11 and (H) Consensus 14

In present study, four potential siRNA molecules against four consensus target (target 7, 8, 11 & 14) were found which fulfill all the parameters, hence support more effective binding of siRNA with target mRNA. Therefore these siRNA molecules may be used as potential candidate in the advanced treatment of Zika fever by silencing the 3' UTR of different strains of Zika virus. This is a high quality finding which helps to meet the demand of same treatment regimen against different viruses. RNAi, a novel post transcriptional gene silencing technique is successfully used against various viral infections such as hepatitis B [52] and C [53] infection. Though both the siRNA and miRNA is the novel constituent of RNAi, the siRNA is comparatively better than miRNA for gene silencing due to the sequence specificity with its target and easier insertion technique into the cell. Recent study showed the possible siRNA mediated therapeutics against metabolic disorder of liver and hypercholesterolemia [54]. Synthetic siRNA targeted against the gene coding structural Env protein of HIV-1 virus can successfully suppress the expression of HIV-1 gene. Thus synthetic siRNA molecules provide a rapid and cost effective tool for new anti-HIV-1 gene therapeutics [55]. Nowadays RNAi is the riding light for the treatment of various diseases.

4. Conclusion

RNAi therapy is a new tool to design a number of potential siRNA molecules for the PTGS of significant gene in various biological systems. The compatibility of these siRNA molecules with their target can be calculated by different bioinformatics tools. The present investigation designed 4 potential siRNA molecules as effective candidate to silence the expression of 3' UTR of different strains of Zika virus. In the battle against viral infection, these synthetic molecules may be used as novel antiviral therapy and provide a basis to the researchers and pharma industry to develop the antiviral therapeutics at genomic level.

Conflicts of interest

The authors declare that they have no conflict of interest.

Corresponding Author
Mohammad Mahfuz Ali Khan Shawan
Department of Biochemistry and Molecular Biology
Faculty of Biological Sciences
Jahangirnagar University
Savar, Dhaka-1342, Bangladesh
Cell No.: +88-01714-451833; +88-01797-547130
E-mail: mahfuz_026shawan@yahoo.com,
mahfuz_026shawan@juniv.edu

References

1. Hayes EB. Zika Virus Outside Africa. Emerging Infectious Diseases 2009; 15(9): 1347–1350.
2. Mohammad Mahfuz Ali Khan Shawan, Hafij Al Mahmud, Md. Mahmudul Hasan, Afroza Parvin, Md. Nazibur Rahman and S. M. Badier Rahman. *In Silico* Modeling and Immunoinformatics Probing Disclose the Epitope Based Peptide Vaccine Against Zika Virus Envelope Glycoprotein. Indian J. Pharm. Biol. Res. 2014; 2(4):44-57.
3. Dick GW, Kitchen SF., Haddow AJ. Zika virus. I. Isolations and serological specificity. Trans R Soc Trop Med Hyg. 1952; 46(5):509-20.
4. Marchette NJ., Garcia R., Rudnick A. Isolation of Zika virus from Aedes aegypti mosquitoes in Malaysia. Am J Trop Med Hyg. 1969; 18(3):411-5.
5. Henderson BE., Hewitt LE., Lule M. Serology of wild mammals. In: Virus Research Institute Annual Report, East African Printer 1968; Nairobi, Kenya. Pp 48–51.
6. Kirya BG., Okia NO. A yellow fever epizootic in Zika Forest, Uganda, during 1972: Part 2: Monkey serology. Trans R Soc Trop Med Hyg. 1977; 71:300–303.
7. McCrae AW., Kirya BG. Yellow fever and Zika virus epizootics and enzootics in Uganda. Trans R Soc Trop Med Hyg. 1982; 76: 552–562.
8. McCrae AW., Kirya BG. and Tukei PM. Summary of an apparent epizootic of Zika virus: Pattern of incidence from Aedes africanus collected from the Zika Forest, 1969–1970. In: Virus Research Institute Annual Report, East African Printer 1970; Nairobi. pp 20–21.
9. Foy BD., Kobylinski KC., Chilson Foy JL., Blitvich BJ., Travassos da Rosa A., Haddow AD. et al. Probable non-vector-borne transmission of Zika virus, Colorado, USA. Emerg Infect Dis. 2011; 17(5):880-2.
10. Duffy MR., Chen TH., Hancock WT., Powers AM., Kool JL. et al. Zika virus outbreak on Yap Island, Federated States of Micronesia. N Engl J Med. 2009; 360: 2536–2543.
11. Haddow AD., Schuh AJ., Yasuda CY., Kasper MR., Heang V., Huy R. et al. Genetic characterization of Zika virus strains: geographic expansion of the Asian lineage. PLoS Negl Trop Dis. 2012; 6(2).
12. Centers for Disease Control and Prevention. Zika fever in Cook Islands. 7 April, 2014.
13. Kuno G., Chang GJ. Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. Arch Viro. 2007; 152: 687–696.

14. Gebhard LG., Filomatori CV., Gamarnik AV. "Functional RNA elements in the dengue virus genome". Viruses 2011; 3(9): 1739–56.
15. Filomatori CV., Lodeiro MF., Alvarez DE., Samsa MM., Pietrasanta L., Gamarnik AV. "A 5' RNA element promotes dengue virus RNA synthesis on a circular genome". Genes Dev. 2006; 20(16): 2238–49.
16. Alvarez DE., Lodeiro MF., Ludueña SJ., Pietrasanta LI., Gamarnik AV. "Long-range RNA-RNA interactions circularize the dengue virus genome". J Virol. 2005; 79(11): 6631–43.
17. Chiu WW, Kinney RM., Dreher TW. "Control of Translation by the 5′- and 3′-Terminal Regions of the Dengue Virus Genome". J. Virol. 2005; 79(13): 8303–15.
18. Pijlman GP., Funk A., Kondratieva N. *et al.* "A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity". Cell Host Microbe. 2008; 4(6): 579–91.
19. Zeng L., Falgout B., Markoff L. "Identification of Specific Nucleotide Sequences within the Conserved 3′-SL in the Dengue Type-2 Virus Genome Required for Replication". J. Virol. 1998; 72(9): 7510–22.
20. Services YSDoH. Zika virus. Information for clinicians and other health professionals. Federated States of Micronesia: Yap State Department of Health; 2007. p. 2
21. Tuschl T., Borkhardt A. Small interfering RNAs: a revolutionary tool for the analysis of gene function and gene therapy. Mol Interv. 2002; 2:158.
22. Ma Y., Chan CY., He ML. RNA interference and antiviral therapy. World J Gastroenterol. 2007; 13(39):5169-79.
23. Elbashir SM., Lendeckel W., Tuschl T. RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev. 2001;15(2):188-200.
24. Zamore PD., Tuschl T., Sharp PA., Bartel DP. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell 2000; 101(1):25-33.
25. Bernstein E., Caudy AA., Hammond SM., Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 2001; 409:363–366.
26. Hammond SM., Bernstein E., Beach D., Hannon GJ. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. Nature 2000; 404:293–296.
27. Sayda M. Elbashir, Jens Harborth, Winfried Lendeckel, Abdullah Yalcin, Klaus Weber, Thomas Tuschl. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 2001; 411: 494-498.
28. Brummelkamp TR., Bernards R., Agami R. A system for stable expression of short interfering RNAs in mammalian cells. Science 2002; 296(5567):550-3.
29. Jackson AL., Linsley PS. Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. Nature Reviews Drug Discovery 2010; 9:57–67.
30. Sarita Singh, Sunil Kumar Gupta, Anuradha Nischal, Sanjay Khattri, Rajendra Nath, Kamlesh Kumar Pant, Prahlad Kishore Seth. Design of potential siRNA molecules for hepatitis delta virus gene silencing. Bioinformation 2012; 8(16): 749–757.
31. Suza Mohammad Nur, Mohammad Al Amin, Rashel Alam, Md Anayet Hasan, Md Amzad Hossain, Adnan Mannan. An *in silico* approach to design potential siRNA molecules for ICP22 (US1) gene silencing of different strains of human herpes simplex 1. J Young Pharm. 2013; 5(2): 46–49.
32. Y. Naito, J. Yoshimura, S. Morishita, K. Ui-Tei. siDirect 2.0: updated software for designing functional siRNA with reduced seed-dependent off-target effect. BMC Bioinformatics 2009; 10, pp.392.
33. Gupta SK., Akhoon BA., Srivastava M., Gupta SK. A Novel Algorithm to Design an Efficient siRNA by Combining the Pre Proposed Rules of siRNA Designing. J Comput Sci Syst Biol. 2010; 3:005-009.
34. Altschul SF., Gish W., Miller W., Myers EW., Lipman DJ. "Basic local alignment search tool". J Mol Biol. 1990; 215:403–410.
35. Hoque KM., Azim MF., Mia MR., Kayesh R., Ali MH., Islam MJ., Shakil SK., Shuvra TM. Design of potential siRNA molecules for T antigen gene silencing of Merkel Cell Polyomavirus. Bioinformation 2012;8(19):924-30.
36. Bellaousov S., Reuter JS., Seetin MG., Mathews DH. RNAstructure: Web servers for RNA Secondary Structure Prediction and Analysis. Nucleic Acids Research 2013; 41**:**W471-W474.
37. Nur SM., Hasan MA., Amin MA., Hossain M., Sharmin T. Design of potential RNAi (miRNA and siRNA) molecules for Middle East respiratory syndrome coronavirus (MERS-CoV) gene silencing by computational method. Interdiscip Sci. 2014; 6:1-9.
38. T.R. Ambili, A. Manimekalan, Bhat D. Abhijith. SIRNA GENE THERAPY FOR ALZHEIMER'S DISEASE TARGETING *APP*

GENE. International Journal of Pharmacy and Pharmaceutical Sciences 2012; 4(3):341-346.
39. Taxman DJ, Livingstone LR., Zhang J. et al. Criteria for effective design, construction, and gene knockdown by shRNA vectors. BMC Biotechnol. 2006; 6:7.
40. 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.
41. Freier SM., Kierzek R., Jaeger JA., Sugimoto N., Caruthers MH., Neilson T., Turner DH. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA 1986; 83:9373-9377.
42. Amarzguioui M., Prydz H. An algorithm for selection of functional siRNA sequences. Biochem Biophys Res commun. 2004; 316:1050-1058.
43. Reynolds A. et al. Rational siRNA design for RNA interference. Nat Biotech. 2004; 22:326–330.
44. Vickers TA., Wyatt JR., Freier SM. Effects of RNA secondary structure on cellular antisense activity. Nucleic Acids Res. 2000; 28(6):1340-7.
45. Bret SE., Harris HS., Bowers SC., Rossi JJ. siRNA target site secondary structure predictions using local stable substructures. Nucleic Acid Res. 2005; 33(3):e30.
46. Mathews DH. Predicting a set of minimal free energy RNA secondary structures common to two sequences. Bioinformatics 2005; 21(10):2246-53.
47. Zuker M. On finding all suboptimal foldings of an RNA molecule. Science 1989; 244:48.
48. Singh S., Gupta SK., Nischal A. et al. Design of potential siRNA molecules for hepatitis delta virus gene silencing. Bioinformation 2012; 8:749-757.
49. Glorioso JC., DeLuca NA., Fink DJ. Development and application of herpes simplex virus vectors for human gene therapy. Annu Rev Microbiol. 1995; 49:675-710.
50. Tanaka K., Kanazawa T., Ogawa T., Takashima Y., Fukuda T., Okada H. Disulfide crosslinked stearoyl carrier peptides containing arginine and histidine enhance siRNA uptake and gene silencing. Int J Pharm. 2010; 398(1-2):219-24.
51. W.C. Russell. Update on adenovirus and its vectors. J. Gen. Virol. 2000; 81:2573– 2604.
52. Chen Y., Cheng G., Mahato RI. RNAi for treating hepatitis B viral infection. Pharm Res. 2008; 25(1):72-86.
53. Motavaf M., Safari S., Alavian SM. Therapeutic potential of RNA interference: a new molecular approach to antiviral treatment for hepatitis C. J Viral Hepat. 2012; 19(11):757-65.
54. Czech MP., Aouadi M., Tesz GJ. RNAi-based therapeutic strategies for metabolic disease. Nat Rev Endocrinol. 2011; 7 (8):473-84.
55. Park WS., Miyano-Kurosaki N., Hayafune M., Nakajima E., Matsuzaki T., Shimada F., Takaku H. Prevention of HIV-1 infection in human peripheral blood mononuclear cells by specific RNA interference. Nucleic Acids Res. 2002; 30(22):4830-5.

2/3/2015