

# Bioinformatics prediction of siRNAs as potential antiviral agents against dengue viruses

Paula M Villegas-Rosales<sup>1</sup>, Alfonso Méndez-Tenorio<sup>2</sup>, Elizabeth Ortega-Soto<sup>1</sup>, Blanca L Barrón<sup>1\*</sup>

<sup>1</sup>Microbiology Department, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Carpio y Plan de Ayala S/N, Casco de Santo Tomás, México D.F. 11340, México; <sup>2</sup>Biochemistry Department, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Carpio y Plan de Ayala S/N, Casco de Santo Tomás, México D.F. 11340, México; Blanca L Barrón\*- Email: bbarron@ipn.mx; Phone: 52-55 57296300; Fax 52-55 57296207; \*Corresponding author

Received June 04, 2012; Accepted June 08, 2012; Published June 16, 2012

## Abstract:

Dengue virus (DENV 1-4) represents the major emerging arthropod-borne viral infection in the world. Currently, there is neither an available vaccine nor a specific treatment. Hence, there is a need of antiviral drugs for these viral infections; we describe the prediction of short interfering RNA (siRNA) as potential therapeutic agents against the four DENV serotypes. Our strategy was to carry out a series of multiple alignments using ClustalX program to find conserved sequences among the four DENV serotype genomes to obtain a consensus sequence for siRNAs design. A highly conserved sequence among the four DENV serotypes, located in the encoding sequence for NS4B and NS5 proteins was found. A total of 2,893 complete DENV genomes were downloaded from the NCBI, and after a depuration procedure to identify identical sequences, 220 complete DENV genomes were left. They were edited to select the NS4B and NS5 sequences, which were aligned to obtain a consensus sequence. Three different servers were used for siRNA design, and the resulting siRNAs were aligned to identify the most prevalent sequences. Three siRNAs were chosen, one targeted the genome region that codifies for NS4B protein and the other two; the region for NS5 protein. Predicted secondary structure for DENV genomes was used to demonstrate that the siRNAs were able to target the viral genome forming double stranded structures, necessary to activate the RNA silencing machinery.

**Keywords:** Dengue virus, siRNA, NS4B and NS5 proteins

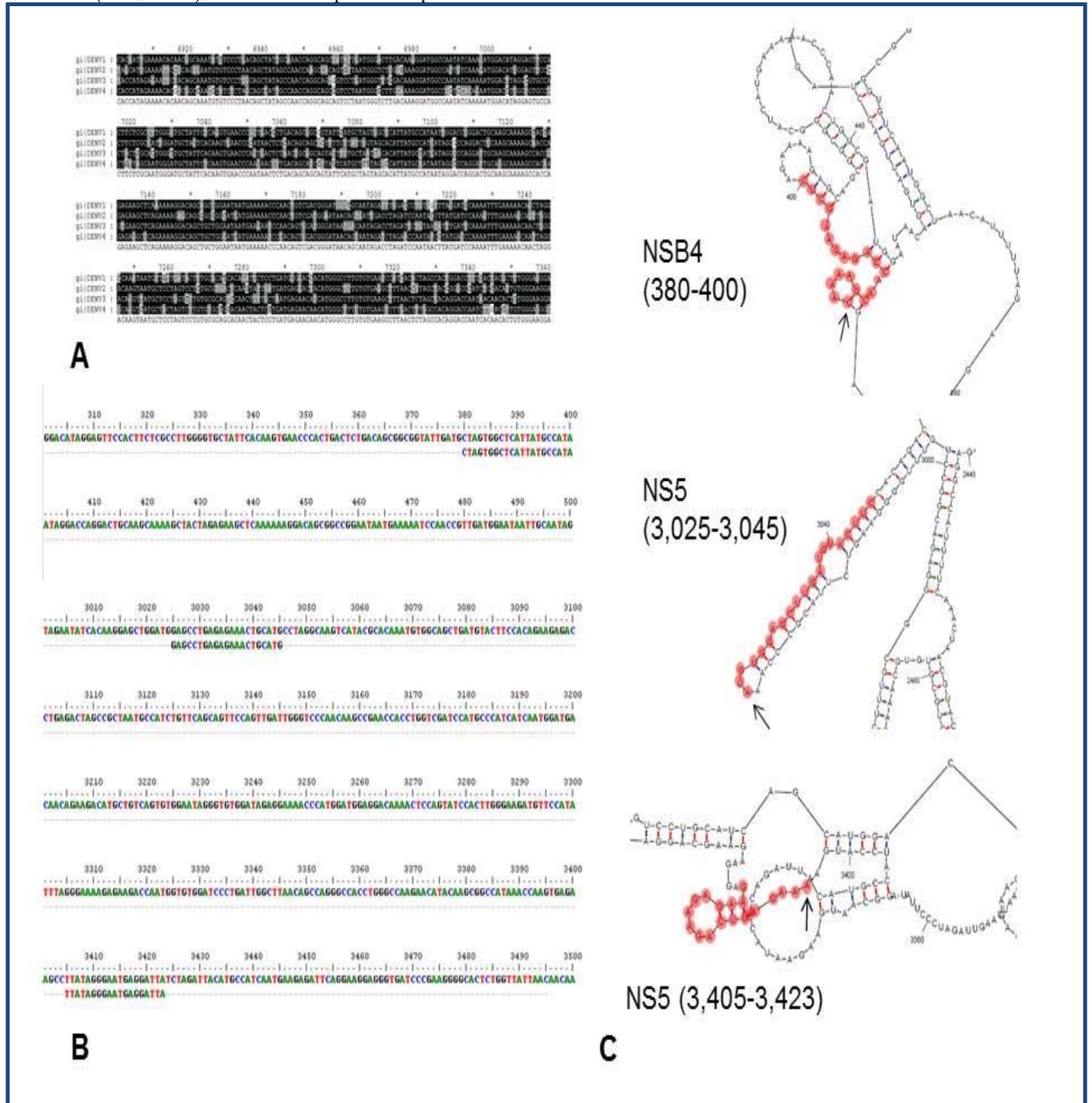
## Background:

Dengue viruses (DENV) are enveloped positive-single stranded RNA viruses, classified in the *Flaviridae* family, *Flavivirus* genus. After virus cell entry, the RNA genome is translated into a single polyprotein which is processed by cellular and virus-derived proteases in three structural proteins (C, M and E) and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) [1-3]. Dengue viruses' infection is the most important viral disease transmitted by arthropods in the world [2-5]. The disease is endemic in more than 100 countries throughout Africa, America, Eastern Mediterranean, South-East Asia and Western Pacific areas. There are four distinct serotypes

of DENV and each of them can cause the same disease symptoms ranking from self-limited febrile illness called dengue fever (DF) to dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS). In fact, infection with one serotype confers protective immunity against that serotype but not against the others, and there is neither a specific treatment nor an approved vaccine for these viral infections [6-10]. Therefore, there is a need to develop therapeutic agents to treat such diseases. We propose the use of short interfering RNAs (siRNAs) to target highly conserved regions among the four DENV serotypes to silence the viral genome. The RNA interference mediated by siRNAs, is now being widely used to

knockdown gene expression in a sequence specific manner, and is a promising new antiviral therapy for many viral infections, such as human immunodeficiency virus (HIV), hepatitis B and C viruses (HBV, HCV), Venezuelan equine encephalitis virus

(VEEV), and some respiratory viruses, including influenza virus, among others [11-13].



**Figure 1:** Design of siRNAs against DENV NS4B and NS5 coding regions. **A)** Multiple alignment of the four DENV serotypes showing a fragment of the highly conserved NS4B encoding region (nt position 6910-7360) depicted in black. **B)** Alignment of the proposed siRNAs with the DENV consensus sequence (obtained by alignment of 215 DENV sequences). **C)** Three regions of DENV-2 genome structure showing the interaction sites with the proposed siRNAs (red circles). The arrows indicate the initial site of interaction.

**Methodology:**

To identify highly conserved regions in the genome of the four DENV serotypes, a multiple alignment with ClustalX program

was carried out, using only the reference sequences for the four serotypes available at NCBI data base (<http://www.ncbi.nlm.nih.gov>). A highly conserved sequence

among the four serotypes was found to be located in the region encoding for the NS4B and NS5 proteins. Moreover a total of 2,893 complete DENV sequences were downloaded. Fingerprints for each sequence was calculated by using the virtual hybridization program [14] and then, clustered in order to discard identical genome sequences [15].

After removing the redundant sequences a total of 220 genomes were left. The complete genomes were edited to select only the encoding region for NS4B and NS5 proteins (nucleotides 6,827-10,217 for DENV-1, 6,826-10,269 for DENV-2, 6,821-10,264 for DENV-3 and 6,828-10,262 for DENV-4). Then, a multiple alignment with ClustalX was built to obtain a consensus sequence using the GeneDoc software. For the siRNA design, three servers were used: Invitrogen RNAi Designer (<https://rna.designer.invitrogen.com/rnaiexpress>), Dharmacon siRNA Design Center (<http://dharmacon.com/sidesign/default.aspx>), and Ambion, Inc. ([http://ambion.com/techlib/misc/siRNA\\_finder.html](http://ambion.com/techlib/misc/siRNA_finder.html)). All the resulting siRNAs were aligned to find out the most prevalent siRNAs; finally the resulting siRNAs were manually analyzed according to the Reynolds' criteria [16] to choose the siRNAs with the higher probability to silence the four DENV genomes. Prediction of the DENV genome secondary structure for each serotype was obtained with the RNAstructure program 4.6 (<http://rna.urmc.rochester.edu/RNAstructure.html>) [17] and GeneBee ([http://www.genebee.msu.su/services/rna2\\_full.html](http://www.genebee.msu.su/services/rna2_full.html)). The genome structure was analyzed as target for the predicted siRNAs.

## Result and Discussion:

To design the siRNAs for DENV genomes silencing, a consensus sequence representing all DENV genomes available at GenBank data base was proposed. For that, the four reference DENV genome serotypes were analyzed by a multiple alignment procedure, finding a highly conserved region among the four DENV serotypes, which was located in the genome region codifying for the NS4B and NS5 proteins (approximately 3,400 nt) (Figure 1). Then, a total of 2,893 complete DENV genomes stored at NCBI GenBank database available as of October 2011, were downloaded: 1,268 genomes were for serotype 1; 872 for serotype 2; 655 for serotype 3 and 83 genomes for serotype 4. The viral genomes were depurated to eliminate identical genomes or redundant sequences, ending with 220 complete DENV genomes, 100 of them were serotype 1, 57 serotype 2, 47 serotype 3 and 16 serotype 4. These genomes were edited to work only with the encoding sequences for NS4B and NS5 proteins, and after carrying out a multiple alignment, a consensus sequence was proposed. The consensus sequence was loaded to design the siRNAs using three servers (Figure 1). A total of 60 siRNAs were obtained. To find out which siRNAs were the most prevalent, all the proposed siRNAs were subject to a multiple alignment, finding 14 different siRNAs.

A final selection of the 14 siRNAs was done by manual analysis based on the Reynolds' criteria [16] for an effective RNA interference with siRNA, 1)30-52% of G/C content, 2) at least 3 A/Us in positions 15-19, 3) absence of internal repetitions, 4) A in position 19, 5) A in position 3, 6) U in position 10, 7) No G/C in position 19, and 8) No G in position 13. At the end, three siRNAs met most of criteria proposed for RNA silencing, one targeted the genome region that codifies for NS4B protein and

the other two; targeted the region for NS5 protein. To assure that the siRNAs were specific for DENV, we performed a BLAST analysis against Homo sapiens and viral databases. To confirm the interaction of the siRNAs with the four DENV genome serotypes, we used as a target the predicted the secondary structure of each DENV genome serotypes obtained with the RNA structure 4.4 and GeneBee programs, finding that the proposed siRNAs were able to form double-chain structures with the viral genome, which are necessary to activate the silencing system. The DENV genome secondary structures we chose was the one for each DENV serotype with the lowest free energy obtained with the RNAstructure program, which it is the most probable structure. Even that the RNAstructure program shows a graphical structure it was unable to analyze a complete DENV genome (approximately 10,700 nucleotides), therefore for this analysis we input only the NS4B and NS5 region (Figure 1). We confirmed that the proposed siRNAs could form a double stranded RNA structure with the target region of DENV genome by analyzing the results provided with the GeneBee program, which does not display a graphical structure, but it takes in account the full genome sequence and gives numerical data indicating positions of loops and stalks. All these results suggest that is possible to silence the four DENV genome serotypes by using siRNAs directed against the genome sequences encoding for the NS4B and NS5 proteins. Moreover, the designed siRNAs seemed to be specific for DENV viruses, according to the BLAST analysis performed. Furthermore, until now, siRNA technology has been mainly used to study DENV infections, mainly by silencing cellular genes which might affect viral infection such as the CD14 monocyte receptor, genes involved in processes of endocytosis, cytoskeletal dynamics and endosome trafficking [18], and also, targeting DENV genome in the highly conserved 5' cyclization sequence (5'CS) region, preM and E genes [19-21], but in this paper we proposed siRNAs as novel therapeutic approach for DENV infections by silencing the encoding region for NS4B and NS5 proteins.

## Conclusion:

Three siRNAs were predicted to be able to silence the four DENV genome serotypes by targeting the viral NS4B and NS5 sequences, forming double-chain structures with the viral genome, which are necessary to activate the silencing system. Therefore, these siRNAs will be synthesized and their inhibitory effect will be tested *in vitro* against the four DENV serotypes.

## Acknowledgment:

We thank Secretaría de Investigacion y Posgrado, Instituto Politecnico Nacional for all the support and facilities to carry out this work.

## References:

- [1] Bollati M *et al.* *Antiviral Res.* 2010 **87**: 125 [PMID: 19945487]
- [2] Rodenhuis-Zybert I *et al.* *Cell Moll Life Sci.* 2010 **67**: 2773 [PMID: 20372965]
- [3] Lindenbach BD & Rice CM, *Adv Virus Res.* 2003 **59**: 23 [PMID: 14696326]
- [4] Erb SM *et al.* *Virology.* 2010 **406**: 328 [PMID: 20708768]
- [5] Mackenzie JS *et al.* *Nat Med.* 2004 **10**: S98 [PMID: 15577938]
- [6] Marianneau P *et al.* *J Virol.* 1999 **73**: 5201 [PMID: 10233989]

- [7] Sampath A *et al.* *Antiviral Res.* 2009 **81**: 6 [PMID: 18796313]
- [8] Schmidt AG *et al.* *PLoS Pathog.* 2010 **6**: e1000851 [PMID: 20386713]
- [9] Noble CG *et al.* *Antiviral Res.* 2010 **85**: 450 [PMID: 20060421]
- [10] Xie X *et al.* *J Virol.* 2011 **85**: 11183 [PMID: 21865382]
- [11] Zhou J & Rossi JJ, *Methods Mol Biol.* 2011 **721**: 67 [PMID: 21431679]
- [12] Seth S *et al.* *Methods Mol Biol.* 2010 **623**: 397 [PMID: 20217566]
- [13] Tripp RA & Tompkins, *Methods Mol Biol.* 2009 **555**: 43 [PMID: 19495687]
- [14] Reyes-López MA *et al.* *Nucleic Acids Res.* 2003 **31**: 779 [PMID:12527788]
- [15] Reyes-Prieto *et al.* *Adv Appl Bioinform Chem.* 2011 **4**: 13 [PMID:21918634]
- [16] Reynolds A *et al.* *Nat Biotechnol.* 2004 **22**: 326 [PMID: 14758366]
- [17] Mathews DH *et al.* *Curr Protoc Nucleic Acid Chem.* 2007 **11**: 11[PMID: 18428968]
- [18] Alhoot MA *et al.* *PLoS Negl Trop Dis.* 2011 **5**: e1410 [PMID: 22140591]
- [19] Stein DA *et al.* *J Virol.* 2011 **85**: 10154 [PMID: 21795337]
- [20] Yue JY *et al.* *Bing Du Xue Bao* 2010 **26**: 373 [PMID: 21043137]
- [21] Ang F *et al.* *Virol J.* 2010 **7**: 24 [PMID: 20122152]

Edited by P Kanguane

Citation: Villegas-Rosales *et al.* *Bioinformation* 8(11): 519-522 (2012)

**License statement:** This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited