Bioinformation by Biomedical Informatics Publishing Group *open access*
www.bioinformation.net **Hypothesis**

Predicted RNA secondary structures for the conserved regions in dengue virus

Pallavi Somvanshi*, Prahlad Kishore Seth

Bioinformatics Centre, Biotech Park, Sector G, Jankipuram, Lucknow 226021, Uttar Pradesh, India
Pallavi Somvanshi - E-mail: psomvanshi@gmail.com; Tel.: +91 522 4012076; fax: +91 522 4012081; *Corresponding author

Received December 03, 2008; revised April 03 2009; accepted April 16, 2009; published August 2, 2009

Abstract
Dengue fever, dengue hemorrhagic fever and dengue shock syndrome are the prevalent mosquito borne viral infections worldwide. The dengue virus belongs to the genus flavivirus with conserved RNA domains peptidase_S7 and dexHc among its members. The secondary structures for RNA domains peptidase_S7 and DexHc are hence predicted and discussed with other known viral RNA structures to glean structural insights through comparison.

Keywords: RNA, dengue; structure; prediction; thermodynamics; pathogenicity

Background:
Dengue viral infection poses a growing public health problem in various tropical and subtropical countries. Dengue was circulated as a quasi-species, which is categorized into four serotypes **[1].** Dengue virus belongs to the genus *Flavivirus* of the family *Flaviviridae* **[2]**. There are four serotypes of dengue virus (DEVI-IV) which causes dengue hemorrhagic fever and dengue shock syndrome **[2, 3].** The infection caused by dengue viruses are widely recognized as a major public health concern, with more than one million cases of dengue hemorrhagic fever (DHF) per year with fatality rates 1 to 10%. The most susceptible to the disease are children and young adults.

Dengue is a positive stranded RNA virus and the genome encodes a single polyprotein **[4].** A comprehensive assessment of phylogenetic relationship of genus *Flavivirus* was determined. Three RNA domains (Peptidase_S7, DEXHc and Flavi_NS5) were found conserved in the genus *Flavivirus*. The conserved RNA domains help in processing of the polyprotein into mature protein subunits. DEXHc contributes in unwinding of nucleic acids for various aspects of RNA metabolism, nuclear transcription, pre mRNA splicing, ribosome biogenesis, nucleo-cytoplasmic transport, translation, RNA decay and organellar gene expression and flavi_NS5 that possess a number of short regions and motifs homologous to other RNA directed RNA polymerases **[5].**

The development of an anti-viral drug or vaccine target is still a non-trivial task for dengue virus. Therefore, it is important to understand RNA structures. RNA structures are characterized with secondary structures and are less complex than protein structures. It is also known that the single stranded RNA structures are often stabilized hydrogen bonds **[6].** MFold is a well known tool used for prediction of RNA secondary structure. RNA secondary structure plays an important role viral multiplication. It is known that the evolution of virus RNA genome is subjected to various structural constrains including RNA structures **[7].** Therefore, it is important to predict and discuss RNA secondary structures for conserved domains of peptidase_S7 and DexHc in dengue virus.

Methodology:
Dengue genome sequence and conserved domain
The complete genome sequence of dengue viruses was downloaded from GenBank at National Center for Biotechnological Information (NCBI) **[8]** and the Universal Virus Database of the International Committee on Taxonomy of Viruses genome database (ICTVdB) **[9].** The complete DNA sequences of dengue virus were used to identify the conserved domain using the conserved domain database (CDD) at NCBI.

RNA secondary structure prediction package
Prediction of RNA folding of the conserved domains of dengue virus was completed using the online MFold package **[10].** MFold is the most widely used algorithms (uses genetic algorithm) for RNA secondary structure prediction, which are based on a search for the minimal free energy state. The algorithm allows structures to be removed at later stages of the simulation if other pairings are found to be more favorable. This is in addition to the possibility of growing of new stems. The algorithm also allows the prediction of certain tertiary interactions (example, RNA pseudo-knots).

Discussion:
The RNA genome of dengue virus contains conserved domains of peptidase_S7 and dexHc among different strains **(Table 1 in supplementary material).** Therefore, it is important to predict their secondary structures **(Figure 1 and Figure 2)**. We used MFold (software package) to predict their secondary structures with energy calculations **(Table 2).** The dengue I stain showed the highest free energy for peptidase_S7 with δG -38.50 and DexH with δG -30.50. However, the dengue IV virus showed the lowest free energy for peptidase_S7 (with δG -42.30) and DexHc (with δG -42.80). We find structural similarity with other predicted RNA structures described henceforth below.

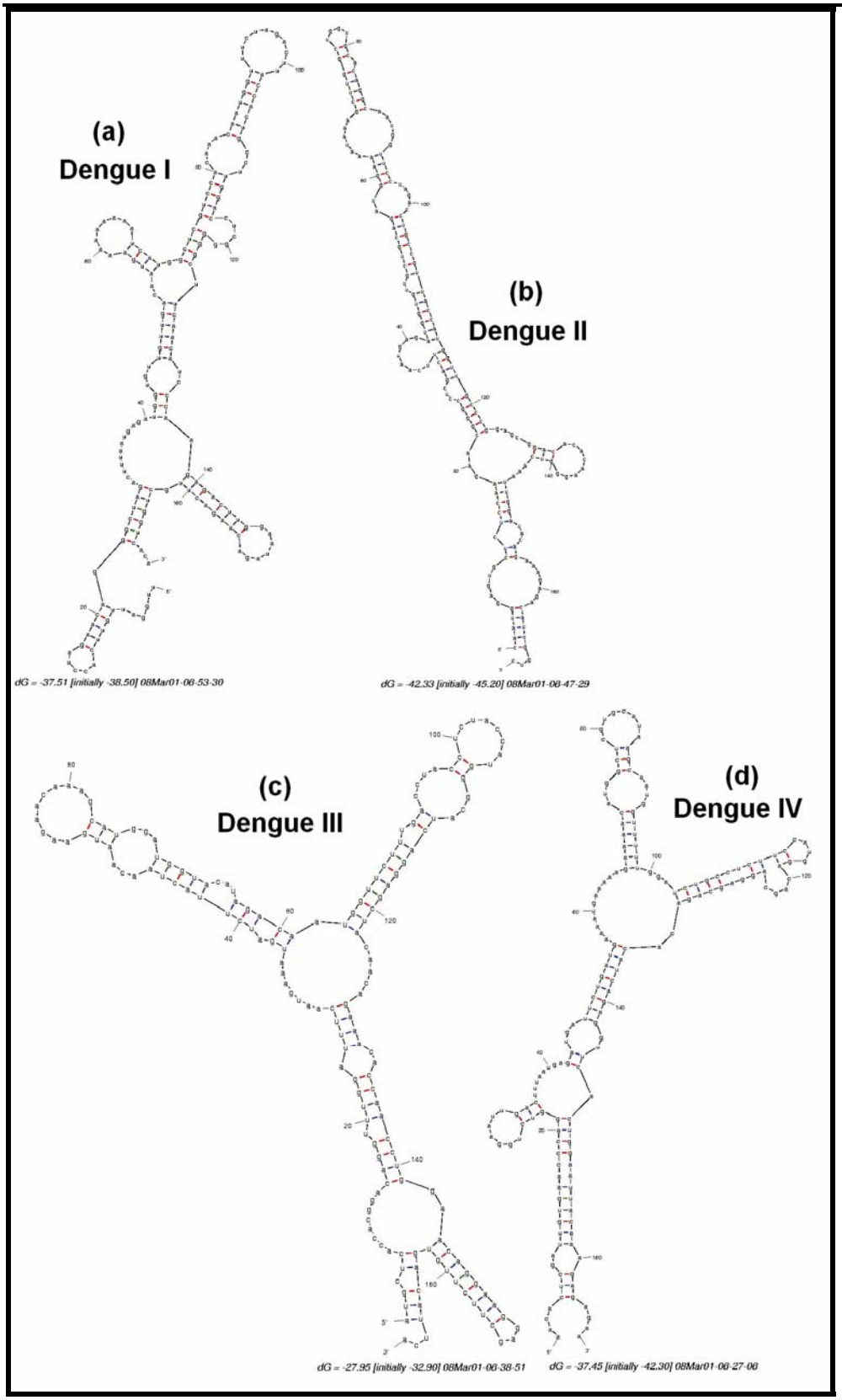

Figure 1: RNA secondary structures of peptidase_S7 conserved domain for dengue serotypes.

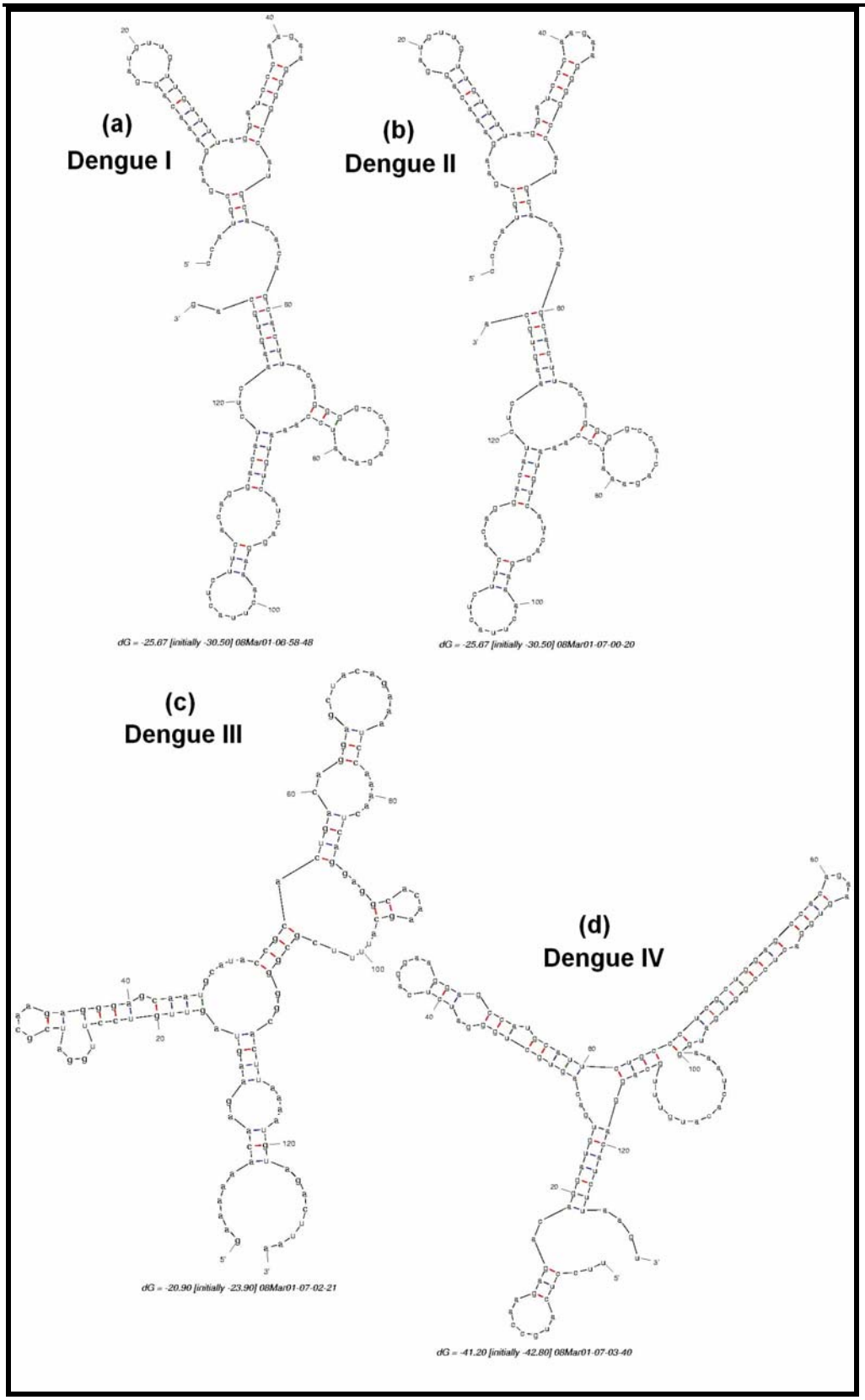

Figure 2: RNA secondary structures of DeXHc conserved domains of dengue serotypes.

Several reports are available on the RNA secondary structure prediction. The genomic diversity of argentine tospo viruses from different geographical areas with several distinct crops has been reported. A 450 nt fragment of the N gene were substantially described. A partial sequence of the N gene was able to classify local isolates within three tospo virus species previously described (Tomato spotted wilt virus, TSWV; tomato chlorotic spot virus, TCSV and groundnut ring spot virus, GRSV) **[11].** Six evolutionarily conserved stem-loop structures in the NS5B encoding region and two in core gene have been reported**.** This observation relates to that found in HCV, GB virus-B (GBV-B) with similar internal base pairing in its coding region **[12].**

Hepatitis C virus (HCV) possesses extensive RNA secondary structure in the core and NS5B-encoding regions of the genome. A program (STRUCTUR_DIST) was developed to the analyses of multiple RNA folding patterns predicted using MFOLD to determine the evolutionary conservation of predicted stem loop structures. This method helps to analyse frequencies of covariant sites in predicted RNA folding between HCV genotypes **[13].** A comparison of the structure of the 3`-untranslated region (3`-UTR) of HGV/GBV-C with the upstream NS5B coding sequence has been shown. The secondary structure predictive algorithms and analysis of covariance between HGV/GBVC genotypes 1 – 4 and the more distantly related HGV/GBV-C chimpanzee variant, show RNA secondary structure formation in both regions. In particular, the NS5B region contained long stem–loop structures of up to 38 internally paired nucleotides which were evolutionarily conserved between human, chimpanzee and HGV/GBV-C variants **[14].** The NS1 gene of Influenza virus is stable during evolution. The computational tool has been used to model its RNA secondary structure (free energy ranged between -222.90 to-251.10 Kcal/mol) for nine different strains of Influenza A virus **[15].**

RNA secondary structure prediction was combined with comparative sequence analysis to construct models of folding for the distal 380 nucleotides of the 3'-untranslated region (3`-UTR) of Yellow fever virus (YFV). A number of structural elements that are thermodynamically stable, conserved in shape, and confirmed by compensatory mutations have been shown **[16].** At the same time structural polymorphisms were observed among strains of YFV. The observation of a strong association between secondary structure of the 3`-UTR and virulence of YFV may help elucidate the molecular mechanisms of virus attenuation. This may lead to the development of new strategies directed towards rational modification of secondary structure of the 3`-UTR.

Prediction of evolutionarily conserved secondary structure motifs in the genomic RNAs of the family *Flaviviridae* has been reported **[17].** This virus family consists of the three genera *Flavivirus, Pestivirus*, *Hepacivirus* and a group of GB virus C / hepatitis G virus (with an uncertain taxonomic classification)**.** RNA secondary structures for 5S rDNA of 37 bacteria have been investigated **[18].** Data show that the lowest free energy of the 5S rDNA is related to the most primitive bacteria and high free energy always indicates less stability during the evolution**.**

Conclusion:
The prediction of RNA secondary structure was performed for peptidase_S7 and DexHc of dengue virus serotypes. The predicted structures of conserved RNA domains provide structural insights for potential function in RNA mediated viral replication.

Acknowledgement:
We thank Indian Council of Medical Research, New Delhi (64/2/07/BIF-BMS) and Department of Biotechnology, New Delhi.

Reference:
[1] J. Mongkolsapaya *et al., J Immunol.,* 176(6): 3821 (2006)
[2] D. J. Gubler *Trends Microbiol*., 10:100 (2002).
[3] S. B. Halstead *Science* 239:476 (1988).
[4] F. Billoir *et al.,* J *Gen. Virol.*, 81:781 (2000).
[5] http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
[6] E. Zamaratski *et al., Nucleosides, Nucleotides & Nucleic Acids* 20: 1219 (2001)
[7] P. Simmonds *et al., RNA* 10:1337 (2004).
[8] http://www.ncbi.nlm.nih.gov/genomes/VIRUSES/viruses.html
[9] http://www.ncbi.nlm.nih.gov/ICTVdb/
[10] M. Zuker *Science* 244:48 (1989).
[11] R.A. Dewey *et al., Acta Horticulturae* 431: (1996)
[12] A. Tuplin *et al., RNA* 8:824 (2002).
[13] A. Tuplin *et al., J. Gen. Virol*. 85:3037 (2004).
[14] N. M. Cuceanu., *et al., J. Gen. Virol*. 82:713 (2001).
[15] P. Somvanshi *et al., J Proteomics Bioinform* 1:219 (2008).
[16] V. Proutski *et al., J. Gen. Virol.* 78: 1543 (1997).
[17] A. Thurner *et al., J. Gen. Virol.* 85: 1113 (2004).
[18] V. Singh *et al., J. mol. Graphics mod.* 29:770(2009)

Edited by P. Kangueane
Citation: Somvanshi & Seth, Bioinformation 3(10): 435-439 (2009)

Bioinformation by Biomedical Informatics Publishing Group *open access*
www.bioinformation.net **Hypothesis**

Supplementary material

Table 1: Conserved regions in RNA of dengue virus

Virus Name	Accession No	Peptidase_S7	DeXHc
Dengue 1	NC_001477	1482 - 1652	1665 - 1794
Dengue 2	NC _001474	1482 - 1652	1664 - 1793
Dengue 3	NC _001475	1480 - 1650	1663 - 1792
Dengue 4	NC_ 002640	1481 - 1648	1663 - 1792

Table 2: Free energies (δG) of different dengue virus strains

S. No.	Strain / conserved domain	base	δG value (Kcal/mol)
1	DEV-I Peptidase_S7	171	-38.50
2	DEV-II Peptidase_S7	170	-38.10
3	DEV-III Peptidase_S7	170	-32.90
4	DEV-IV Peptidase_S7	167	-42.30
5	DEV-I DexHc	129	-30.50
6	DEV-II DexHc	129	-30.50
7	DEV-III DexHc	129	-23.90
8	DEV-IV DexHc	129	-42.80