# RNAVLab: A unified environment for computational RNA structure analysis based on grid computing technology

M. Taufer[1], M.-Y. Leung[2,4], K. L. Johnson[3], A. Licon[1]
[1]Dept. of Computer Science, [2]Dept. of Mathematical Sciences,
[3]Dept. of Biological Sciences, and Bioinformatics Program[4]
University of Texas at El Paso
*{mtaufer, mleung, kljohnson, alicon2}@utep.edu*

## Abstract

*Ribonucleic acid (RNA) molecules play important roles in many biological processes including gene expression and regulation. An RNA molecule is a linear polymer which folds back on itself to form a three dimensional (3D) functional structure. While experimental determination of precise 3D RNA structures is a time consuming and costly process, useful insight into the molecule can be gained from knowing its secondary structure. Structural elements in RNA secondary structures can be separated into two large categories: stem-loops and pseudoknots. The development of mathematical models and computational prediction algorithms for simple stem-loop structures started early in the 1980's. However, building systems that provide the tremendous computer time and memory needed for RNA analysis of both stem-loops and pseudoknots remains a challenge even today. The recently developed grid computing technology can offer a possible solution to this challenge.*

*In this paper we briefly address mathematical problems associated with the grid computing approach to RNA structure prediction. In particular, we introduce models to partition a large RNA molecule into smaller segments to be assigned to different computers on the grid. Based on these models, we formulate a sampling strategy to select RNA segments for computational prediction to maximize prediction consistency. This strategy is under construction as part of RNAVLab, our unified environment for computational RNA structure analysis, i.e., prediction, alignment, comparison, and classification. A first prototype of RNAVLab is presented and used to investigate the possible association of secondary structure types with RNA functions by analyzing secondary structures for a family of nodavirus genomes.*

1-4244-0910-1/07/$20.00 ©2007 IEEE.

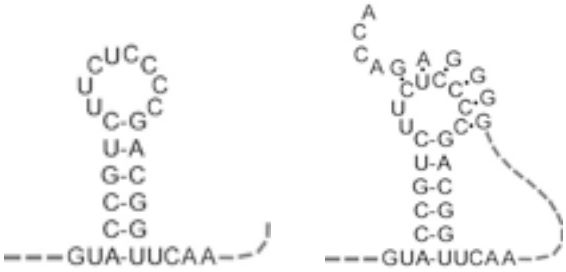
## 1. Introduction

Ribonucleic Acid (RNA) is made up of four types of nucleotide bases: adenine (A), cytosine (C), guanine (G), and uracil (U). A sequence of these bases is strung together to form a long, single-stranded RNA molecule. RNA molecules vary greatly in size, ranging from about twenty nucleotide bases in microRNAs to a long polymer of over 30,000 bases in complete viral genomes [1]. Among the four nucleotide bases, C and G form a complementary base pair by hydrogen bonding, as do A and U. Although an RNA molecule is a linear polymer, it tends to fold back on itself to form a three dimensional (3D) functional structure mostly by pairing between complementary bases. The 3D structure of an RNA molecule is often the key to its function. Because of the instability of RNA molecules, experimental determination of their precise 3D structures is a time consuming and rather costly process. However, useful information about the molecule can be gained from knowing its secondary structure that refers to the collection of hydrogen bonded base pairs in the molecule. Essentially, all RNA secondary structures are made up of elements that can be classified into two basic categories: stem-loops and pseudoknots (see Figure 1). Both kinds of secondary structure elements have been implicated in important biological processes like gene expression and regulation [2, 3, 4, 5]. The development of mathematical models and computational prediction algorithms for stem-loop structures started in the 1980's [6, 7]. Pseudoknots, because of the extra base-pairings, must be represented by more complex models and data structures and require large amounts of memory and computing time to obtain the optimal and suboptimal structures with minimal free energies [8].

To overcome the tremendous demand on computing resources needed by pseudoknot prediction, various alternative algorithms have been proposed [9, 10, 11] which tend to restrict the types of pseudoknots to be

predicted to keep computation time and memory size under control. However, a much larger variety of pseudoknots occurs in reality. The omission of these pseudoknots from computational methods might significantly affect the prediction accuracy.

The need for accurate methods and flexible tools for predictions of a wider range of RNA secondary structures is the key motivation for the research in RNA secondary structures presented in this paper. In particular we are exploring a new approach to RNA secondary structure analysis that combines mathematical methods with grid computing technology. This approach is part of RNAVLab (RNA Virtual Laboratory), a unified computational environment for RNA secondary structure analysis, i.e., prediction, alignment, comparison, and classification. RNAVLab provides the scientists with a computational environment in which they specify the kind of analysis to perform on one or a series of RNA segments and get back final results without requiring the scientists' intervention: scientists are not required to cut and paste results from one portal to another, redirect or reformat output files (e.g., from FASTA to EMBL format) before forwarding them to the next step in multi-step analyses, etc. To computationally address critical questions such as "*what are the mechanisms of nodavirus RNA replication?*", RNAVLab combines different mechanisms and programs in an easy-to-use, automated (i.e., requiring little intervention from the scientist), computationally powerful, and reliable environment.




(a) Stem-loop (b) Pseudoknot

**Figure 1: Examples of stem-loop and pseudoknot.**

The rest of this paper is organized as follows: In Section 2 we present relevant bioinformatics challenges that we are addressing through our RNAVLab; in Section 3 we describe the mathematical methods in RNAVLab and its modular software components for prediction, alignment, comparison, and classification; in Section 4 we show some preliminary results that substantiate and enforce our approach for RNA structure analysis; in Section 5 we conclude and present some work in progress.

## 2. Challenges in Bioinformatics

The critical question that is driving our mathematical and computer science research is whether it is possible to associate stem-loop and pseudoknot types with functions in genomes. According to the specific ways of base pairing observed in currently known structures, pseudoknots can be classified into six different subtypes [12]. Preliminary studies have suggested that the subtype classification may be associated with the pseudoknot function (e.g., viral frameshifting).

As a starting point we are currently investigating the association of stem-loop and pseudoknot structures in genomes with their viral replication mechanisms for nodavirus genomes. In particular, as described in details in Section 4.1, we are targeting the association of anticipated structure at the 3' end of nodavirus RNA2 with their viral replication mechanisms. Nodaviruses provide an excellent model system for the study of RNA replication due to their genetic simplicity, their robust yield of replication products, and the ability of their RNAs to replicate in cells from a wide variety of organisms. As part of ongoing studies to determine the mechanism of nodavirus RNA replication, we are investigating the minimal RNA sequences necessary for replication of nodavirus genomic RNAs, and the RNA secondary structures they adopt, using both the RNAVLab environment presented in this paper and a well-defined reverse-genetic system in yeast cells.

## 3. Mathematical Methods and Grid-based Computational Environment

The idea underlying the possible use of mathematical methods and grid computing is that a large RNA molecule can be cut into shorter segments. The secondary structures of the segments can be predicted individually by different computers and the individual predictions for the small pieces can be assembled to give a predicted structure for the original molecule. The advantage of the grid computing approach is that it can accommodate a variety of existing and new prediction algorithms in a heterogeneous workflow. However, the challenge lies in the necessity of ensuring that the predicted results of the smaller pieces are sufficiently consistent with one another so that they can be assembled to generate a reasonable structure for the original molecule.

### 3.1 Mathematical Methods to Study Inversion Distribution in RNA Sequences

The mathematical component of this paper is motivated by the consistency problem described above. In both stem-loops and pseudoknots, it is necessary to have a stretch of nucleotide sequence (ACCGUC in

Figures 1a and 1b) followed by its inverted complementary sequence (GACGGU) downstream. For simplicity, we refer to this kind of patterns as *inversions*. Based on Markov models with parameters that best fit a set of RNA sequence data (e.g., the genome sequence data of nodaviruses), we can establish probabilistic distributions of inversions in random RNA sequences and statistical criteria for identifying significantly high concentration of inversions using the techniques of Poisson approximation and scan statistics similar to those described by Leung et al. in [13]. All these are based on a basic calculation of the probability of occurrence of an inversion at a particular position of the sequence.

Once the regions of the sequence containing statistically significant high concentrations of inversions are identified, we can formulate a sampling strategy to cut a long RNA sequence into segments so that the most inversions are preserved within the segments. We expect that this would lead to a more consistent prediction of secondary structures. As we are currently building the framework of RNAVLab, we will integrate our statistically based segment sampling strategy into this set of tools allowing a comparison of the predicted structures with already known structures in the database.

### 3.2 Grid-based Environment for the Unified Study of RNA Structures

While past efforts have improved predictions accuracy of sequential prediction programs [8, 9] and prediction efficiency has been improved by using massively parallel machines [14] and heterogeneous local clusters [15], not much is known about prediction systems based on Grid Computing technology (i.e., heterogeneous computers, ranging from supercomputers to clusters and PCs connected to the Internet, spread out at different locations). With their significant computing power, these systems will allow larger spaces of RNA secondary structures to be explored. RNAVLab uses grid computing to build the computing power needed for predictions of large RNA sequences. Grid computing technology has been applied successfully in the past to protein structure prediction [16] and therefore we expect similar achievements for RNA structure prediction.

The modular framework of RNAVLab makes it easy to integrate new features. As shown in Figure 2, RNAVLab has three major components: (1) a segment sampler component (*Sampling*) to sample RNA sequence segments guided by simple heuristics and more sophisticated mathematical methods capable of identifying palindrome distributions; (2) a structure predictor component (*Prediction*) to predict the structures of the sampled segments using different programs on heterogeneous computers; and (3) a structure analysis component (*Analysis*) to evaluate prediction results based on metrics such as energy landscapes built from collected results. Each component is shown in more detail in Figure 3 and described below:

**Segment Sampler Component:** The segment sampler identifies putative segments in RNA sequences and passes them to the structure predictor component. Generally, RNA segments containing a high concentration of close inversions have greater tendency to form local secondary structures because the symmetry facilitates base pairing required in the formation of stem-loops and pseudoknots [17, 18]. Currently RNAVlab includes two sampling strategies: a windowing strategy and a progressive segmentation strategy. The windowing strategy requires the user to input the starting and ending points in the segment in terms of nucleotide base positions. In the progressive sampling approach the user defines a starting point, ending point, and a "step size"; the sampler generates a series of segments by progressively removing "step size" bases from the original segment, whose length is defined by the beginning and ending points given by the user, starting from the starting point. The series of segments with progressively decreasing lengths are forwarded to the prediction component for prediction. Segments can be inverted before being forwarded. The extension of this component to employ more sophisticated statistics-based sampling methods using the distributional properties of close inversion on random RNA sequences as described in Section 3.1 is work in progress.

**Structure Predictor Layer:** The structure predictor deploys heterogeneous computing resourses across the UTEP campus to rebuild RNA secondary structures from RNA segments using different prediction programs. Currently we support the following prediction programs: Pknots-RE [8], Pknots-RG [9], and NuPack [19]. This component of RNAVLab uses a similar approach as that used in the open-source Grid middleware APST, or AppLeS Parameter Sweep Template, [20] to launch application tasks, move data, and discover resources across network domains. This RNAVLab component currently supports Grid environments such as Condor [21] and PBS. Eventually we plan to extend it to support BOINC [22] to allow researchers to deploy desktop and laptop PCs owned by students or administration personnel when their computers are idle. The work in [16] shows that adding idle cycles of PCs significantly increase available computing power.

**Structure Analysis Component:** The structure analysis component evaluates the consistency of the

various predictions collected by the structure predictor. Currently a set of tools allows the end-user to perform classifications as well as different types of comparisons and alignments of secondary structures. An innovative aspect of this component is that it deals with secondary structure rather than nucleotide sequences for classification, alignment, and comparison. To classify, align, and compare secondary structures, we consider them in terms of strings of brackets, i.e., "(" and ")","[" and "]", "{" and "}", dots ".", and colons ":". Two paired nucleotides are represented with a pair of brackets collocated in the string at the same position as the correspondent nucleotides in the input segment. With reference to the classification of secondary structures and more in particular of pseudoknots, we deploy the classification of pseudoknots in [12] that clusters them into five different simple types, i.e., LL-, HLout, HLin-, HH-, HHH-type, and an unclassified type. Note that "H" means hairpin loop, "L" means bulge loop, "in" means internal loop or multiple internal loops, and "out" means external loop or multiple external loops. The tool for classification works on the string of brackets to extract the proper class for pseudoknots. We also align two secondary structures by aligning their two bracket strings using variants of well-known alignment algorithms such as the Smith-Waterman (1981) and Needleman-Wunsh (1970) algorithms. Unlike these algorithms that align string of nucleotides, i.e., "A", "U", "C", and "G", we align strings of brackets, i.e., ":", "(", and "[". Comparison of observed structures and predicted, or predicted structures from different codes are performed on aligned or non-aligned strings of brackets. Three different algorithms can be used for comparisons:

- A variant of the Hamilton algorithm (Hamiltonian Comparison) - we assign a numerical tag to specific nucleotide pairings. This approach is useful when the types of nucleotide bonds are important.
- A stack based algorithm (Comp. Stack) – we use stacks to compare the structures using a stringent comparison method. This approach is useful when exactly alike structures are important.
- A lenient algorithm (Comp. OC) – we use simple string comparisons which allows for similar yet shifted structures to receive high comparison scores. This approach captures similarities that the two previous algorithms would not.

Eventually predictions will be clustered based on energy criteria into energy landscapes. Information on the consistency of the prediction results collected so far will be fed back to the segment sampler component to adjust the sampling strategy and adaptively identify new RNA segments for predictions.

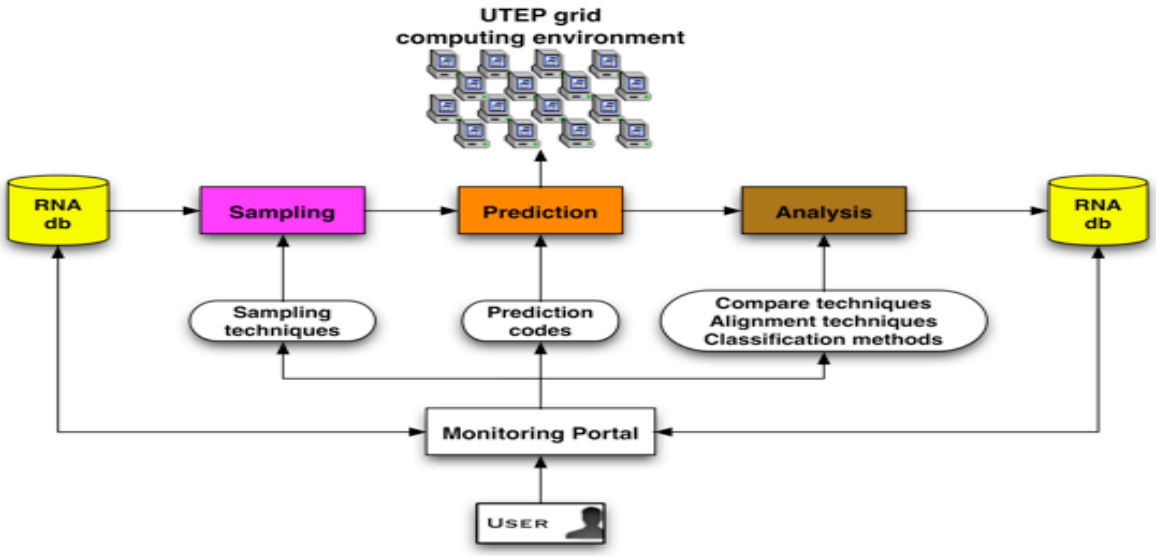



**Figure 2: High level overview of RNAVLab, its software components, and a monitoring portal.**

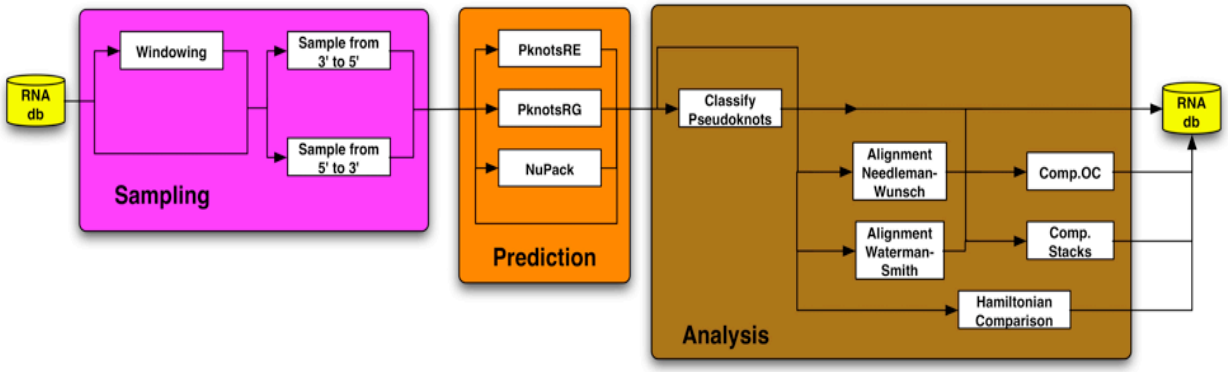



**Figure 3: Detailed overview of the methods in the three major components of RNAVLab.**

## 4. Computational Results

This section provides some insight on how we are deploying RNAVLab for studying the correlation between replication mechanisms of nodaviruses and the secondary structures adopted by the 3' ends of their RNA2 segments, which are hypothesized to play a role in initiation of complementary strand synthesis during RNA replication. The presented preliminary results encourage us to pursue further studies in this field supported by our unified environment.

### 4.1 Short Overview of the Nodavirus Family

The nodavirus genome is divided into two segments of positive-strand RNA: RNA1 encodes the viral RNA-dependent RNA polymerase that replicates both genomic segments, while RNA2 encodes the precursor to the protein that comprises the viral outer coat (capsid). A small subgenomic RNA3 encodes a protein that suppresses host defense mechanisms like RNA interference. The role of RNA secondary structure in the genome replication of other RNA viruses, e.g., members of the plant tombusvirus, potexvirus, and bromovirus families and the animal picornavirus, coronavirus, and flavivirus families, has been well established in the literature. For the nodavirus family, this role has been studied only for one member, *Flock House virus* (FHV). A long-range interaction between two regions of RNA1 was required for synthesis of subgenomic RNA3 [23]. The results of genetic experiments suggest that a similar long-range interaction may be also required for synthesis of the RNA3 of another member of the family, *Nodamura virus* (NoV) as well [24]. However, the role of RNA secondary structure in replication of nodavirus genomic RNAs remains unclear. Defining this role is crucial to understanding the mechanism of nodavirus RNA replication. The predictive approaches deployed in RNAVLab will greatly facilitate our molecular studies by providing a "road map" to elements of possible structural importance, allowing these sequences to be targeted by site-directed mutagenesis.

Previous studies with FHV showed that sequences at the 3' end of RNA2, particularly in the terminal 50 nucleotides, were critical for RNA replication, and could direct replication of chimeric RNAs that contained heterologous core sequences flanked by RNA2 sequences [25]. By replacing the center of RNA2 with the same heterologous sequence, the work in [25] created a family of RNA molecules that differed only at their termini. The different properties of these molecules could be therefore confidently attributed to these termini. This system established a uniform assay for the different RNAs, using a single probe to the common central core region for Northern blot hybridization experiments. Since such chimeric RNA molecules replicate efficiently, they provide an ideal model substrate for secondary structure prediction and analysis.

### 4.2 Preliminary Computational Results

We used RNAVLab to computationally investigate experimental results indicating how the terminal 50 nucleotides in the 3' end of RNA2 are critical for RNA replication [24]. We analyzed predicted RNA secondary structures of progressively shorter lengths from the 3' end of RNA2 from NoV: the original segment of 200 nucleotides was sampled using the progressive segmentation strategy with a step size of 10 nucleotides. Three different prediction programs were used: Pknots-RG, Pknots-RE, and NuPack. All the final predictions, obtained from segments with different lengths and different prediction programs, were crossed-aligned using the variant of the Smith-Waterman algorithm to identify common motifs, i.e., pseudoknots or stem-loops. Due to the dynamic nature of the prediction programs, the final secondary structures are heavily dependent on neighboring structures: having a certain sub-structure present in all

the predictions, independently from the starting and ending points of the segments, may indicate a strong binding that ultimately may be present in nature.

Table 1 presents the results of this computational study. Because of space constraints, we show only the results for the secondary structures with lengths: 100 (from base 100 to base 200), 50 (from base 150 to base 200), 40 (from base 160 to base 200), and 30 (from base 170 to base 200). The omitted results are similar to those shown in the table. Also, note that the character '_' in the strings representing the structures is introduced by the alignment tool to maximize the alignment scores. The prediction time for the several secondary structures ranged from several hours for long segments predicted using Pknots-RE to a couple of seconds for short segments predicted using Pknots-RG and NuPack. The predictions were performed in parallel on a 64 dual-processor Beowulf cluster; RNAVLab managed the distribution of the segments and the collection of the predicted structures through PBS. A performance analysis of the computation time is not in the scope of this paper and therefore not addressed in detail.

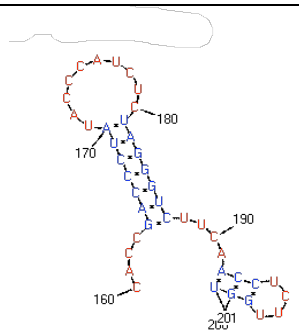
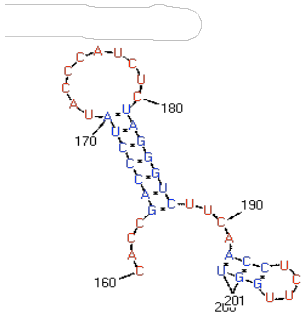
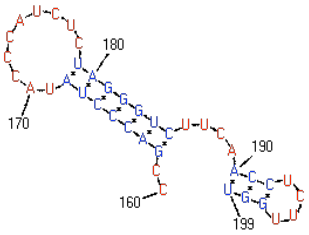
| | Alignments in RNVLab | Stem-loop |
|---|---|---|
| **Pknot-RE** | Pknots-RE<br>CGUUGACGACGCAAAACGUCCUUAAAGCGUUGACGACGCAAAACGUCCCCAAGCUCGUAGCACCGACCCUAUACCCAUCUCUAGGGUCUUCAACCUCUUGGU<br>100 - 200 :((((((((((:::::::::::::::))))))):::((((:::::))))):[[[[[((:::::)):::((((((:::::::::::)))))))):)))))::::]]]]:<br>150 - 200 _____::((:::::)):::((((((:::::::::::)))))))::::(((::::)))<br>160 - 200 _____:::((((((:::::::::::))))))):::::(((::::)))<br>170 - 200 _____::((((____:::::::))))::::(((::::)))_ |  |
| **Pknot-RG** | Pknots-RG<br>CGUUGACGACGCAAAACGUCCUUAAAGCGUUGACGACGCAAAACGUCCCCAAGCUCGUAGCACCGACCCUAUACCCAUCUCUAGGGUCUUCAACCUCUUGGU<br>100 - 200 :(((:((((:((::(((((:::::::)))))):::((((:::::))))):::::)))))))):::(((((((:::::::::::)))))))::::(((::::)))<br>150 - 200 _____((:::::)):::(((((((:::::::::::)))))))::::(((::::)))<br>160 - 200 _____:::(((((((:::::::::::)))))))::::(((::::)))<br>170 - 200 _____(((____:::::::))))::::(((::::))) |  |
| **NuPack** | NuPack<br>CGUUGACGACGCAAAACGUCCUUAAAGCGUUGACGACGCAAAACGUCCCCAAGCUCGUAGCACCGACCCUAUACCCAUCUCUAGGGUCUUCAACCUCUUGGU<br>100 - 200 (((((:(((((((:::[[[[:::::::)))))))):)))))::::]]]]::(((((:::((:(::::(((((((:::::::::::)))))))):):)):)))))):<br>150 - 200 _____((:::::)_):::(((((((:::::::::::)))))))::::(((::::)))<br>160 - 200 _____:::(((((((:::::::::::)))))))::::(((::::)))<br>170 - 200 _____(((____:::::::))))::::(((::::)))_ |  |

**Table 1: The 3' end of NoV's RNA2 segment – predictions of sub-segments with different lengths and using different programs.**

As Table 1 shows, RNAVLab consistently predicted the presence of a stem-loop structure from base 164 to 187 in the segment, within the last 50 nucleotides of the segments. For segments shorter than 40 nucleotides, the stem-loop is no longer present, as shown for the strings with length 30 (from base 170 to base 200). These results are consistent with the work conducted experimentally in [24]. We are currently studying whether the stem-loop is indeed the critical structure that drives the genome replication of the virus. To address this critical question we are combining computational and experimental methods. By using RNAVLab, we are currently computationally searching whether the same structure exists in other nodaviruses. Driven by the computational results, we will ultimately address the question whether the biological relevance of the stream-loop can indeed be experimentally verified.

## 5. Conclusions and Future Work

RNAVLab is a unified environment that facilitates the study of RNA secondary structures, i.e., prediction, alignment, comparison, and classification, through an automated, computationally powerful way: the scientist's intervention is minimized and grid computing technologies are used to address computing intensive tasks such as the prediction of RNA pseudoknots. In this paper we present a first prototype of RNAVLab and an example of its application for the computational study of mechanists that guide nodavirus replication. By predicting RNA secondary structures of progressively shorter lengths from the 3' end o *Nodamura virus* RNA2, RNAVlab indicates that, across prediction programs and with different sampled segments, a stem-loop structure from base 164 to 187 in the segment is consistently predicted. Ongoing work includes studying whether this structure is common to other nodaviruses and whether the role of the stem-loop in the genome replication can indeed be experimentally verified.

Ultimately, the predictive approaches deployed in RNAVLab will greatly facilitate molecular studies by providing a "road map" to elements of possible structural importance, allowing these sequences to be targeted by site-directed mutagenesis.

### Acknowledgements

This material is based in part upon work supported by the Texas Advanced Research Program under Grant No. 003661-0008-2006, and the National Institutes of Health under Grants No. S06GM08012-35, 2G12RR008124-11 and 3T34GM008048-20S1. Financial support through the National Science Foundation (grant # EIA-0080940, "MII: Graduate Education for Minority Students in Computer Science and Engineering: Extending the Pipeline" and grant # SCI-0506429) is acknowledged.

## References

[1] Thiel, V., et al. (2003) Mechanisms and enzymes involved in SARS coronavirus genome expression. *J. Gen. Virol.* 84, 2305-2315.

[2] Petrillo, M., Silvestro, G., Di Nocera, P.P., Boccia, A., and Paolella, G. (2006) Stem-loop structures in prokaryotic genomes. *BMC Genomics* 7,170.

[3] Su, M.-C., et al. (2005) An atypical RNA pseudoknot stimulator and an upstream attenuation signal for -1 ribosomal frameshifting of SARS coronavirus. *Nucleic Acids Res,* 33(13), 4265–4275.

[4] Wilkinson, S.R. and Been, M.D. (2005) A pseudoknot in the 3' non-core region of the glmS ribozyme enhances self-cleavage activity. *RNA* 11,1788–1794.

[5] Yang, H.-S., et al. (2004) A novel function of the MA-3 domains in transformation and translation suppressor pdcd4 is essential for its Bi *Biol.* 24(9), 3894–3906.

[6] Sankoff, D. (1985) Simultaneous solution of the RNA folding, alignment, and protosequence problems. *SIAM J. Appl. Math.* 45, 810–825.

[7] Zuker, M. (1989) Computer prediction of RNA structure. *Methods Enzymol.* 180, 262-288.

[8] Rivas, E., and Eddy, S.R. (1999) A dynamic programming algorithm for RNA structure prediction including pseudoknots. J. Mol. Biol. 285, 2053-2068.

[9] Reeder, J. and Giegerich, R. (2004) Design, implementation and evaluation of a practical pseudoknot folding algorithm based on thermodynamics. BMC Bioinformatics 5, 104.

[10] Ruan, J., Stormo, G.D., and Zhang, W. (2004) ILM: a web server for predicting RNA secondary structures with pseudoknots. Nucleic Acids Res. 32 (Web Server issue):W146-149.

[11] Ren J., Rastegari, B., Condon, A., and Hoos, H.H. (2005) HotKnots: Heuristic prediction of RNA secondary structures including pseudoknots RNA 11, 1494-1504.

[12] Han, K. and Byun,Y. (2003) PseudoViewer2: visualization of RNA pseudoknots of any type. Nucleic Acids Res.,31, 3432–3440.

[13] Leung, M.Y., et al. (2005) Nonrandom Clusters of Palindromes in Herpesvirus Genomes. *J. Computat. Biol.* 12(3), 331-354.

[14] Shapiro, B.A. (1988) An algorithm for comparing multiple RNA secondary structures. Comput Appl. Biosci. 4(3), 387-393.

[15] Shapiro, B.A. and Zhang, K.Z. (1990) Comparing multiple RNA secondary structures using tree comparisons. Comput. Appl. Biosci. 6(4), 309-318.

[16] Taufer, M., et al. (2006) Predictor@Home: A protein structure prediction supercomputer based on global computing. IEEE Transactions on Parallel and Distributed Systems. 17(8), 786- 796.

[17] Batenburg van, F.H.D., et al. (2000) Pseudobase: a database with RNA pseudoknots. Nucleic Acids Res. 28(1), 201-204.

[18] Chew, D.S.H., Choi, K.P., Heidner, H., and Leung, M.Y.

(2004) Palindromes in SARS and other coronaviruses, INFORMS J. Comp. 16, 331-340.
[19] Pierce N. A. (2006) NuPack: A software suite for the analysis and design of nucleic acids. http://www.nupack.org.
[20] Casanova H. and Berman F. (2002) Parameter Sweeps on the Grid with APST. Chapter 33 in Grid Computing: Making the Global Infrastructure a Reality, Wiley Publisher, Inc.
[21] Thain, D., Tannenbaum, T., and Livny, L. (2004) Distributed Computing in Practice: The Condor experience. Concurrency and Computation: Practice and Experience.
[22] Anderson D.P. (2004) BOINC: A System for Public-Resource Computing and Storage. GRID 2004: 4-10.
[23] Lindenbach, B. D., Sgro, J.-Y., and Ahlquist, P. (2002) Long-distance base pairing in Flock House virus RNA1 re76, 3905-3919.
[24] Johnson, K. L., Price, B. D., and Ball, L. A. (2003) Recovery of infectivity from cDNA clonesNodamura virus and identification of small nonstructural proteins. *Virology* 305, 436-451.
[25] Albariño, C. G., et al. (2003) The cis-acting replication signal at the 3' end of Flock House virus RNA2 is RNA3-dependent. *Virology* 311, 181-191.