

# Determining Functional Conformations of Two HDV III Strains

Wojciech Kasprzak,<sup>1</sup> Sarah D. Linnstaedt,<sup>2</sup> John L. Casey,<sup>2</sup> and Bruce A. Shapiro<sup>3</sup>

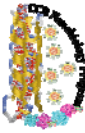
<sup>1</sup>Basic Research Program, SAIC-Frederick, Inc., NCI-Frederick, Frederick, MD

<sup>2</sup>Department of Microbiology and Immunology, Georgetown University Medical Center, Washington, DC

<sup>3</sup>Center for Cancer Research Nanobiology Program, National Cancer Institute at Frederick, Frederick, MD



SAIC SAIC-Frederick, Inc.  
A Laboratory Science Applications  
International Corporation  
From Science to Solutions



## Abstract

Hepatitis Delta virus (HDV) is a sub-viral human pathogen aggravating Hepatitis B virus (HBV) liver infections. The short HDV genome (~1680 nt) is a single stranded, circular RNA encoding only one protein, the hepatitis delta antigen (HDAG). The host enzyme ADAR1 edits the HDV stop codon (UAG) into a tryptophan (W) codon (UGG) enabling expression of the two forms of the protein, short and long, from the same open reading frame. HDAG-S is required for replication, while HDAG-L enables viral particle formation and inhibits replication. The balance between the two forms is crucial and editing must be regulated.

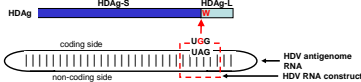
We have applied our programs, MPGAfold and StructureLab, to predict and examine the folding conformations/states of an HDV III construct. This construct includes the editing site (amber/W) and has the editing capabilities of the full HDV III. The predicted secondary structure folding dynamics indicates that the HDV III RNA forms a meta-stable branched structure and a stable rod structure. Both were observed *in vitro*, and the branched structure was identified as the one enabling editing. Computational predictions and the experimental data also indicate that an Ecuadorian strain folds into the editing-capable structures more readily than a Peruvian strain, and we indicate the reasons for the difference. Thus the folding dynamics of HDV III strains appears to strongly influence their RNA editing levels.

Funded in part by NCI Contract N01-CO-12400

## Introduction

### Hepatitis Delta Virus: Background

- Hepatitis delta virus (HDV) increases the severity of liver disease in Hepatitis B virus (HBV) infections.



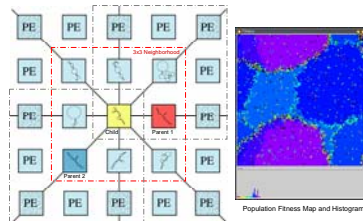
- HDV genome is a single-stranded, circular RNA encoding only the hepatitis delta antigen protein (HDAG).
- RNA editing produces HDAG-S and HDAG-L from the same open reading frame.
- Editing takes place at the amber/W site (ADAR1 deamination from UAG stop codon to UGG-tryptophan (W) codon).
- HDAG-S is required for replication. HDAG-L enables viral particle formation and inhibits replication. Balance is crucial and editing must be regulated.

### Control of Editing Levels in HDV III Strains

- HDV type III Peruvian and Ecuadorian isolates have been examined by computational analysis and *in vitro* and *in vivo* experiments.
- We have been studying how these two strains differ in their ability to distribute their RNA between branched (edited) and unbranched structures, as well as studying the efficiency of editing.
- Both structure and substrate quality were found to contribute to overall editing levels.
- This presentation concentrates on the structural issues responsible for the differences in the levels of editing conformations in HDV III Ecuadorian and Peruvian strains.

## Computational Tools

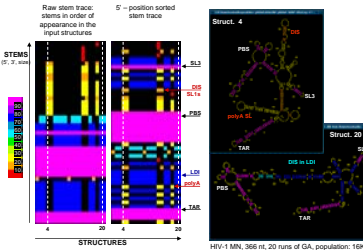
### MPGAfold: Massively Parallel Genetic Algorithm



- Seed a population of a chosen size with initial structure elements (stems from a pre-generated stem pool).
  - Apply random structure mutations (stems) and recombinations (sets of stems) to produce new structures for each generation.
  - Apply the fitness function to the new structures and select for the next generation from those that are most fit (best free energy, including coaxial stem stacking calculations).
  - Repeat steps 2 and 3 for N generations, iterating toward the optimal solution.
- GA is a stochastic algorithm, which requires multiple runs to find the prevailing conformation.

### StructureLab: Stem Trace Data Visualization

- Stem Trace plots all of the unique stems, defined as triplets (5', 3', size), for all of the structures in a solution space

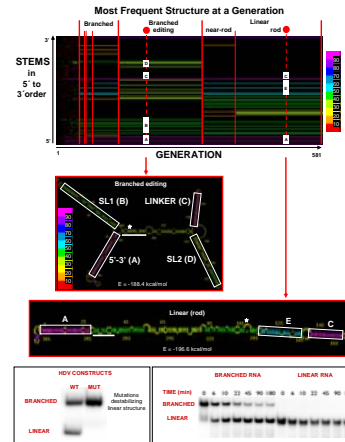


### RNA Structure Prediction and Analysis Tools

- The massively parallel genetic algorithm (MPGAfold) captures RNA folding pathways, including functional intermediates and final states existing in a highly combinatoric solution space.
- A significant amount of information comes from each MPGAfold run, as well as from a set of runs, including variable population runs.
- Interpretation of the results is facilitated by various visualization tools that are part of StructureLab and MPGAfold.
- Each one of these tools views the data from a somewhat different perspective.
- Ultimately, these perspectives are combined to reach an understanding of the folding patterns of the RNA in question.

## HDV III Ecuador

### MPGAfold and Experiments



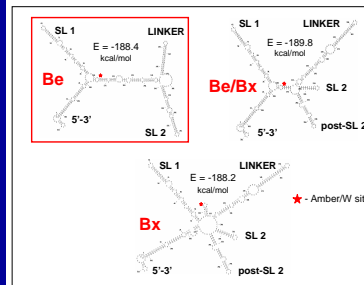
### MPGAfold Statistics

Population	4K	8K	16K	32K	64K
Linear (all)	35	40	57	81	98
Branched (all)	65	60	43	19	2
B. edit final	31	30	16	6	0
B. edit trans.	23	30	41	57	71

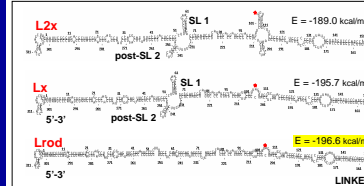
All percentages are based on 100 runs for each MPGAfold population with the output based on the peak histogram structures.

Reference: RNA (2006), 12:1521-1533.

### HDV III Ecuador: Branched Conformations

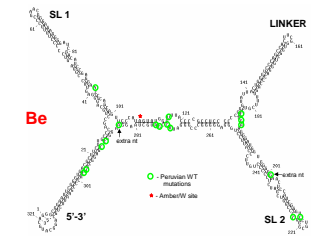


### Linear Conformations



## HDV III Peru

- Folding pathways results for the Peruvian strain has not yet (as of November 2007) been published, and we decided not to show this information at this point (even though it was presented in the poster form at the IMA 2007 meeting).



## Conclusions

- MPGAfold captures the folding states of the HDV III Ecuador, showing the edit conformation (Be) as both the final and a transitional state.
- Agreement with the high levels of Be conformers was observed experimentally *in vitro* and in patients.
- Folding results for the HDV III Peruvian strain reflect experimentally observed differences in the levels of the Be edit structures and their relative stability.

## Selected References

- Linnstaedt, S.D., Kasprzak, W., Shapiro, B.A., and Casey, J.L.: The Role of Metastable RNA Secondary Structure in Hepatitis Delta Virus Genotype III RNA Editing. RNA, 12(8): 1521-1533, 2006.
- Shapiro, B.A., Kasprzak, W., Grunewald, C., and Aman, J.: Graphical Exploratory Data Analysis of RNA Secondary Structure Dynamics Predicted by the Massively Parallel Genetic Algorithm. Journal of Molecular Graphics and Modeling, 25(4): 514-531, 2006.
- Shapiro, B.A., Bengali, D., Kasprzak, W., and Wu, J.-C.: RNA folding pathway functional intermediates: Their prediction and Analysis. Journal of Molecular Biology, 312:27-44, 2001.
- Kasprzak, W. and Shapiro, B.A.: Stem Trace: an interactive visual tool for comparative RNA structure analysis. Bioinformatics, 15(1):16-31, 1999.
- Shapiro, B.A., Wu, J.-C.: Predicting RNA H-type pseudoknots with the massively parallel genetic algorithm. Comput Appl Biosci. 13: 459-71, 1997.
- Shapiro BA, Navetta J.: A massively parallel genetic algorithm for RNA secondary structure prediction. The Journal of Supercomputing. 8: 195-207, 1994.