

forna (force-directed RNA): Simple and Effective Online RNA Secondary Structure Diagrams

Peter Kerpedjiev^{1,*}, Stefan Hammer^{1,2} and Ivo L. Hofacker^{1,2}

¹Institute for Theoretical Chemistry, University of Vienna, Währinger Straße 17/3, A-1090 Vienna, Austria

²Research Group Bioinformatics and Computational Biology, University of Vienna, Währinger Straße 29, A-1090 Vienna, Austria

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Associate Editor: XXXXXXXX

ABSTRACT

Motivation: The secondary structure of RNA is integral to the variety of functions it carries out in the cell and its depiction allows researchers to develop hypothesis about which nucleotides and base pairs are functionally relevant. Current approaches to visualizing secondary structure provide an adequate platform for the conversion of static text-based representations to two dimensional images, but are limited in their offer of interactivity as well as their ability to display larger structures, multiple structures and pseudoknotted structures.

Results: In this paper we present *forna*, a web-based tool for displaying RNA secondary structure which allows users to easily convert sequences and secondary structures to clean, concise and customizable visualizations. It supports, among other features, the simultaneous visualization of multiple structures, the display of pseudoknotted structures, the interactive editing of the displayed structures, and the automatic generation of secondary structure diagrams from PDB files. It requires no software installation apart from a modern web browser.

Availability: The web interface of *forna* is available at <http://rna.tbi.univie.ac.at/forna> while the source code is available on github at www.github.com/pkerpedjiev/forna.

Contact: pkerp@tbi.univie.ac.at

1 INTRODUCTION

The use of secondary structure diagrams is ubiquitous within the field of RNA biology. They convey not only which nucleotides are paired, but also and perhaps more importantly, which are unpaired. The contents and positions of sub-structures such as hairpin loops, interior loops, multiloop junctions and external loops are immediately evident. Such information is of great value to researchers seeking to identify putative mutations to perform when seeking to isolate the structural basis of a biological effect, to find protein binding, and to provide a context for observed behaviour. It is used as both an exploratory as well as a communicative tool. Researchers examine secondary structure diagrams to gain insights about potential functions and mechanisms as well as to describe and disseminate them.

While there are a number of available tools (Darty *et al.*, 2009; Wiese *et al.*, 2005; Hecker *et al.*, 2013; Byun and Han K, 2009)

*to whom correspondence should be addressed

	forna	VARNAs	PV	jVizRNA	RNAfdl
Editing	✓	✓			
Pseudoknots	✓	✓	✓	✓	
PDB Files	✓				
Struct. Prediction	✓				
Probing Data	✓	✓			
Custom Coloring	✓	✓			✓
Color Schemes	✓	✓			
RNA-RNA pairs	✓	✓	✓		
Circular RNA	✓				
Annotations		✓			
Circular Layout		✓		✓	✓

Table 1. Comparison of the features of existing RNA visualization tools (where PV = PseudoViewer). *forna* provides at least three convenient features not found in other programs.

for visualizing the secondary structure of RNA molecules, with the exception of PseudoViewer, none are available online without java and none offer the flexibility in exploring, arranging and manipulating the structure that *forna* does (See Table 1 for an enumerated comparison of features).

2 APPROACH

Our tool, called *forna* (for *f*orce-directed *r*na), consists of a web interface and a server which allows users to input RNA secondary structures as dot-bracket strings, and displays it as a force-directed graph (Screenshot in Figure 1). In a manner first demonstrated by jViz.RNA (Wiese *et al.*, 2005) the user can then position each of the nucleotides and stems by dragging them. Each of the nucleotides is represented as a node, whereas backbone and base-pair bonds are considered links. Connections are treated as springs and a force is calculated to keep them a fixed distance from each other. Hidden helper nodes and extra links help to maintain the familiar RNA secondary structure layout. The initial position of each node (nucleotide) is calculated using the NAView algorithm (Bruccoleri and Heinrich, 1988), but is subsequently optimized by the force calculations. This can (especially for larger molecules) lead to artifacts such as twisted helices and nested loops, but these are easily rectified by dragging the affected nodes to their correct positions.

Input/Output: Users can enter structures in dot-bracket format (See Supplementary Material Section 1.1.3, 1.1.6 and 1.1.7). When done, the diagram can be saved as either a vector (SVG) or raster (PNG) graphic. If one wishes to edit the structure again in the future, it can be saved and reloaded in `forna` using the JSON format.

Dragging to Position Elements: The layout can be rearranged by selecting and dragging single or multiple nodes. The virtual forces then pull the structure toward an RNA-like layout with nearly uniform link distances. This behavior is similar to that available in `jViz.Rna` and valuable for arranging the nucleotides in a relevant, meaningful or simply aesthetically pleasing manner.

Pseudoknots and Custom Links (Supplementary Material Section 1.1.1 and 1.1.6): It is often necessary to display the interaction between two molecules or between different parts of the same molecule (i.e. pseudoknots). While the user can enter pseudoknotted structures in dot-bracket notation (i.e. $((\dots[[\dots)]\dots))$), the pseudoknotted nucleotides in these cases are added as links with no strength. One may also add custom links by holding down shift and dragging from one nucleotide to another. This creates a spring-loaded link which can bring distal portions of a molecule together, or connect separate molecules. Such links are useful in depicting RNA-RNA interactions.

Coloring (Supplementary Material Section 1.1.4): The coloring of nucleotides is essential for overlaying metadata on top of a structure. `forna` provides three default coloring modes: position, structure and sequence which color nucleotides according to their position in the molecule, the type of structural element they are in (i.e. stem, interior, hairpin, multi or exterior loop) or their identity (A,C,G or U). A custom coloring mode is provided where bespoke values (as from probing data) can be entered in a text field.

Integrated Structure Prediction (Supplementary Material Section 1.1.5): To simplify the process of going from sequence to secondary structure, `forna` provides a transparent interface to the Vienna RNA Package (Lorenz et al., 2011) which automatically calculates the minimum free energy for a particular sequence if no secondary structure is provided. This allows one to paste a sequence in the input field and immediately view its predicted secondary structure, without using additional tools.

Tertiary to Secondary Structure (Supplementary Material Section 1.1.2 and 1.1.6): One of the most important and unique features of `forna` is the automatic display of secondary structure information given a 3D structure as a PDB file. `forna` automatically extracts base-pair interaction information using MC-Annotate (Gendron et al., 2001) and displays the canonical secondary structure, which can be explored, manipulated or colored as described in the previous sections. Multiple chains are displayed as disconnected graphs. Proteins are displayed as larger grey nodes and interactions between different chains are represented as dashed lines (see Figure 1).

Reusable Display Container (Supplementary Material Section 1.2): Researchers can effortlessly share RNA structures online by adding a few lines of javascript to their web page and showing a diagram of the secondary structure embedded as an SVG container attached to any specified element in the DOM tree.

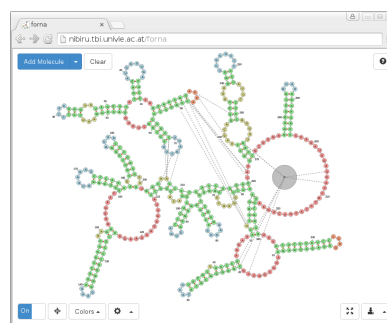


Fig. 1. Screenshot of `forna` web app displaying the "Bacterial Ribonuclease P Holoenzyme in Complex with tRNA" (PDB ID: 3Q1Q). Immediately evident are the regions of the tRNA which are in contact with the ribonuclease, namely the 5' and 3' end nucleotides, as well as the $T\Psi C$ loop. An RNA-binding protein is shown as a grey node in the lower-right hand region of the diagram.

3 CONCLUSION

We provide an easy to use, accessible, free, open-source web tool for RNA secondary structure visualization that produces beautiful, highly customizable plots. Our tool requires no externally installed software and is useful for both the exploration and dissemination of RNA secondary structure.

ACKNOWLEDGEMENTS

Thanks to developers of custom bootstrap elements and other handy tools written in javascript, eg. `saveSvgAsPng.js` `FileSaver.min.js`, `d3.js`.

Funding: This work was funded by the Austrian DK RNA program FG748004, by the Austrian FWF, project "SFB F43 RNA regulation of the transcriptome," and the European Commission under the Environment Theme of the 7th Framework Program for Research and Technological Development (Grant agreement no: 323987).

REFERENCES

- Brucoleri,R.E. and Heinrich,G. (1988) An improved algorithm for nucleic acid secondary structure display. *Computer applications in the biosciences: CABIOS*, **4** (1), 167–173.
- Byun,Y. and Han,K. (2009) PseudoViewer3: generating planar drawings of large-scale RNA structures with pseudoknots. *Bioinformatics*, **25** (11), 1435-1437.
- Darty,K., et al. (2009) VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25** (15), 1974.
- Gendron,P., et al. (2001) Quantitative analysis of nucleic acid three-dimensional structures. *Journal of molecular biology*, **308** (5), 919–936.
- Hecker,N., et al. (2013) RNA secondary structure diagrams for very large molecules: RNAfdl. *Bioinformatics*, **29** (22), 2941–2942.
- Lorenz,R., et al. (2011) ViennaRNA package 2.0. *Algorithms for Molecular Biology*, **6** (1), 26.
- Wiese,K.C. et al. (2005) `jviz. rna-a` java tool for RNA secondary structure visualization. *NanoBioscience, IEEE Transactions on*, **4** (3), 212–218.