Reuter and Mathews *BMC Bioinformatics* 2010, **11**:129
http://www.biomedcentral.com/1471-2105/11/129

BMC Bioinformatics

SOFTWARE Open Access

RNAstructure: software for RNA secondary structure prediction and analysis

Jessica S Reuter, David H Mathews*

Abstract

Background: To understand an RNA sequence's mechanism of action, the structure must be known. Furthermore, target RNA structure is an important consideration in the design of small interfering RNAs and antisense DNA oligonucleotides. RNA secondary structure prediction, using thermodynamics, can be used to develop hypotheses about the structure of an RNA sequence.

Results: RNAstructure is a software package for RNA secondary structure prediction and analysis. It uses thermodynamics and utilizes the most recent set of nearest neighbor parameters from the Turner group. It includes methods for secondary structure prediction (using several algorithms), prediction of base pair probabilities, bimolecular structure prediction, and prediction of a structure common to two sequences. This contribution describes new extensions to the package, including a library of C++ classes for incorporation into other programs, a user-friendly graphical user interface written in JAVA, and new Unix-style text interfaces. The original graphical user interface for Microsoft Windows is still maintained.

Conclusion: The extensions to RNAstructure serve to make RNA secondary structure prediction user-friendly. The package is available for download from the Mathews lab homepage at http://rna.urmc.rochester.edu/RNAstructure.html.

Background

The prediction of RNA structure has received increasing attention over the last decade as the number of known functional RNA sequences, called non-coding RNA (ncRNA), has increased [1]. These new ncRNA sequences range in size from microRNAs to Xist [2,3]. They serve numerous roles, from modulating gene expression [4-6] to catalyzing reactions [7,8].

One of the first steps to understanding the mechanism of action of an RNA is to determine its structure [9]. Secondary structure, defined as the set of canonical base pairs (AU, GC, and GU), can be determined using comparative analysis if a large number of sequences are available [10,11]. In comparative analysis, base pairs are determined when they are conserved in multiple sequences and instances of compensating base pair changes occur. Compensating base pair changes demonstrate the conservation of structure in spite of sequence not being conserved, for example a GC base pair in one sequence being replaced by a homologous AU pair in another sequence. Comparative analysis, however, requires both significant user input and a large number of homologous sequences that can be aligned.

As an alternative to comparative analysis, the secondary structure of an RNA can be predicted for a single sequence using thermodynamics [9]. The thermodynamic methods are based on nearest neighbor rules that predict the stability of a structure as quantified by folding free energy change [12-14]. Often, structure prediction is accomplished by finding the lowest free energy structure, which is the single most probable structure in a folding ensemble [15]. Alternatively, structures can be sampled from the Boltzmann ensemble and a centroid, i.e. representative structure, determined [16,17]. Another alternative method for structure prediction is the prediction of a structure with the highest sum of pairing probabilities, called the maximum expected accuracy structure [18,19].

Single sequence secondary structure prediction is reasonably accurate. On average, for sequences of fewer

* Correspondence: David_Mathews@urmc.rochester.edu
Department of Biochemistry & Biophysics and Center for RNA Biology, University of Rochester Medical Center, 601 Elmwood Avenue, Box 712, Rochester, NY 14642, USA

BioMed Central
© 2010 Reuter and Mathews; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Univ. of Washington EX2061
Guardant Health, Inc v. Univ. of Washington
IPR2022-00817

than 700 nucleotides, the accuracy of predicting known base pairs is as high as 73% [14]. The accuracy, however, benchmarked lower when longer sequences were included [20,21]. Additional sources of information can be used to improve accuracy. For example, base pair probabilities can be determined using a partition function and highly probable pairs are more likely to be correctly predicted pairs [22]. Alternatively, using two or more homologous sequences to determine a conserved structure can result in significantly more accurate structure prediction [23-27]. Experimental data, such as enzymatic cleavage [13], chemical mapping [14], oligonucleotide array binding [28], SHAPE [29], and NMR data [30] can all be used to improve structure prediction accuracy.

In addition to structure prediction, the thermodynamic methods can be applied to other problems. For example, antisense oligonucleotide and siRNA design can be improved using thermodynamic predictions of self-structure in the oligonucleotides and target [31-36]. Sequences can be designed to fold to a specific structure [37,38]. Reverse-PCR primers can be designed to avoid self structure in the template that could prevent hybridization [39]. Novel types of ncRNAs can be found in genomes on the basis of folding stability [40-42].

In this contribution, the RNAstructure software package is described. RNAstructure first appeared in the literature in 1998 as a secondary structure prediction package [43]. At that time, it contained a method to predict the lowest free energy structure and a set of low free energy structures [44,45]. It was subsequently expanded to include bimolecular folding and hybridization thermodynamics with OligoWalk [13,31,33]. It was then expanded to include an algorithm for finding lowest free energy structures common to two sequences, Dynalign [23,41,46]; a partition function algorithm [22]; an alternative prediction method that can determine all low free energy structures for a sequence [28,47]; and stochastic sampling of structures [48]. It provides methods for constraining structures with enzymatic data [13], chemical mapping data [14], SHAPE [29], and NMR data [30]. Finally, recent extensions include PARTS [24], which calculates partition functions for secondary structures common to two sequences and can perform stochastic sampling of common structures [48]; MaxExpect, which finds maximum expected accuracy structures [18]; and a method for removal of pseudoknots, leaving behind the lowest free energy pseudoknot-free structure [49]. Several tutorials exist for using RNAstructure [50-52].

RNAstructure has been publicly available with a user-friendly interface for Microsoft Windows. Testing and development occur in-house using Unix/Linux text-based interfaces, but these have generally been available only upon request. The package is coded in C++. Extensive benchmarks of individual components have been published [14,18,22-24,29-31,46,48,53,54].

Here, three major extensions of RNAstructure are reported. First, a new JAVA-based graphical interface (GUI) is available. This interface functions cross-platform and binaries are available for Apple OS-X and Linux. Second, text interfaces are now available for each component of the package. These interfaces use a standard Unix syntax and include online help. Binaries for the text interfaces are available for Microsoft Windows and source code, including Makefiles, are available for download. Finally, a new class library is available for programmers who want to incorporate the RNAstructure functions into C++ programs. Each of these components is available for download and covered by the GNU Public License, version 2.

Implementation

C++ class library

The C++ class library encapsulates the I/O functions of RNAstructure and also the secondary structure prediction and analysis methods. Four main classes are provided for accessing these functions: RNA, for single sequence structure predictions; Dynalign_object, for Dynalign calculations [23]; HybridRNA, for bimolecular structure prediction [33]; and Oligowalk_object, for OligoWalk [31,33] and OligoScreen [55] calculations. Two other notable classes are utilized by inheritance by the four main classes. Thermodynamics is a class that handles reading and storage of the nearest neighbor parameters. TwoRNA is a class that contains two RNA classes and is inherited by Dynalign_object and instantiated by HybridRNA. A class inheritance diagram is provided as Figure 1.

The classes are designed to be easily included in C++ projects. The classes are compiled to Linux/Unix shared libraries, Windows dynamic link libraries, or Macintosh dynamic shared libraries. An included Makefile provides facility for this. Furthermore, the header (*.h) files are commented using Doxygen-formatted comments that facilitate the creation of a manual http://www.stack.nl/~dimitri/doxygen/. Programmers can either compile their own manuals into html or latex using Doxygen or they can refer to precompiled html manuals included in the source code and posted on the RNAstructure website.

Text interfaces

The new text interfaces are designed to provide the features of RNAstructure for use on the command line and in scripts. Most programs available in the graphical user interfaces are provided with text interfaces. Each of

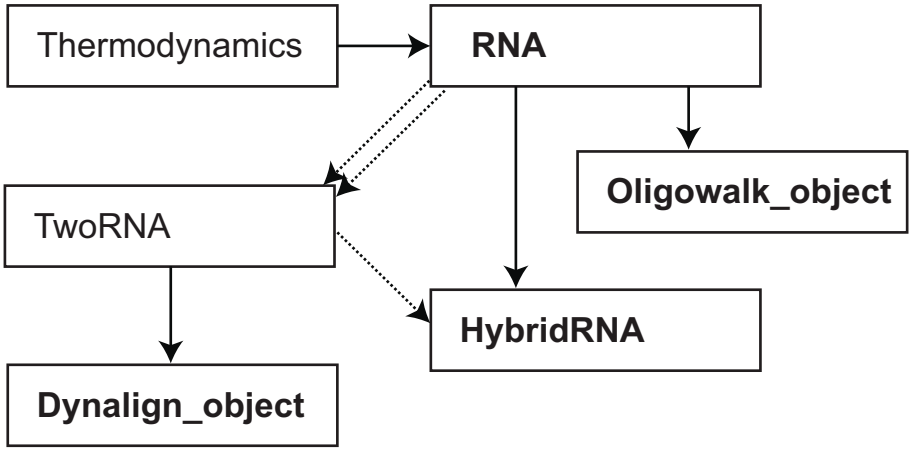

Figure 1 Class inheritance diagram. The four main classes that provide functionality, named in boldface, are RNA, Oligowalk_object, HybridRNA, and Dynalign_object. Solid lines show inheritance and the dotted lines indicate that a class contains an instance of another. TwoRNA contains two instances of RNA and HybridRNA contains an instance of TwoRNA.

these programs provides a brief description of the parameters when invoked without parameters. Additionally, invoking most programs with "-h," "-H," or "–help" will return a more detailed description of the parameters.

The text interfaces are built with ANSI-standard C++. Makefiles are included for compiling the programs in a Unix/Linux/OS X environment. Binaries for Microsoft Windows are also available for download.

The RNAstructure class library is used to implement most of the functions in the text interfaces. Therefore, the text interfaces can be used as tutorials by programmers implementing the RNAstructure algorithms in their own programs.

JAVA Graphical User Interface

The new JAVA GUI is a cross-platform re-implementation of the RNAstructure Windows GUI. This is designed to be user-friendly to make these algorithms accessible to a large audience. Executables are available for Apple OS-X and both 32-bit and 64-bit Linux. A Makefile is provided to facilitate local compilation on Unix/Linux environments.

The JAVA GUI utilizes the new RNAstructure C++ class library to implement the algorithms. The connection between the JAVA front end and the C++ back end is made using SWIG http://www.swig.org/. SWIG wraps the C++ classes, making them accessible to JAVA. The SWIG-wrapped code is made available as part of the RNAstructure code download, so that the JAVA interface can be built on local machines that do not have SWIG installed.

Availability of algorithms

Table 1 shows the availability of the different single-sequence structure prediction and analysis algorithms in RNAstructure. It shows the names of the text interface programs, the corresponding function name and class library, and the menu name in the JAVA and Windows GUIs. Similarly, Table 2 shows the corresponding information for the multi-sequence methods (either for common secondary structure prediction or prediction of nucleic acid hybridization).

Thermodynamic parameters

The algorithms implemented in RNAstructure use nearest neighbor parameters to predict the stability of secondary structures. These include both free energy change parameters at 37°C and enthalpy change parameters to allow prediction of conformation stability at an arbitrary temperature. For RNA, these parameters are those most recent parameters from the Turner group [12,14,54]. For DNA, the parameters are derived from the experimental literature [56-92]. Most algorithms for RNA structure prediction can be invoked for DNA structure prediction using the class, text interfaces, or GUIs. DNA-RNA hybridization parameters are also used by OligoWalk [93].

Unit testing

The RNAstructure package now includes a facility for unit testing. A Makefile automates testing of the text interfaces and comparison with calculation standards. The same tests can be used to test the GUIs, but this

Table 1 Single Sequence Methods

Feature:	Text Interface Program:	JAVA/Windows GUI Menu Item:	Class Library and Function Name:
Free energy minimization structure prediction [14]	Fold	Fold RNA Single Strand	RNA::FoldSingleStrand
Maximum expected accuracy structure prediction [18]	MaxExpect	Predict Maximum Expected Accuracy Structure	RNA:MaximizeExpectedAccuracy
Partition function [22]	partition	Partition Function RNA	RNA::PartitionFunction
Efn2 (energy calculator) [13]	efn2	Efn2 RNA	RNA:: CalculateFreeEnergy
Free energy minimization and generation of all suboptimal structures [28,47]	AllSub	Generate All Suboptimal RNA Structures	RNA::GenerateAllSuboptimalStructures
Stochastic sampling of structures [17]	stochastic	Stochastic RNA Sampling	RNA::Stochastic
Remove Pseudoknots [49]	RemovePseudoknots	Break Pseudoknots	RNA::BreakPeudoknots
Prediction of structures with pairs above specified pairing probability threshold [22]	ProbablePair	Output Probable Structure	RNA::PredictProbablePairs
Drawing secondary structure diagrams	draw	Draw	RNA::DetermineDrawingCoordinates
NAPSS [30]	NAPSS	-	-

Table 2 Multiple Sequence Methods

Feature:	Text Interface Program:	JAVA/Windows GUI Menu Item:	Class Library and Function Name:
Dynalign [23,41,46]	dynalign	Dynalign RNA	Dynalign_object::Dynalign
OligoWalk [31-33]	OligoWalk	OligoWalk	Oligowalk_object::Oligowalk
OligoScreen [55]	oligoscreen	OligoScreen	Oligowalk_object::OligoScreen
Bimolecular structure prediction with intramolecular pairs [33]	bifold	Fold RNA Bimolecular	HybridRNA::FoldBimolecular
PARTS [24,48]	PARTS	-	-
Bimolecular partition function (no intramolecular pairs) [31]	bipartition	Partition Function RNA Bimolecular	HybridRNA::PartitionFunctionBimolecular
Bimolecular structure prediction without intramolecular pairs	DuplexFold	-	HybridRNA::FoldDuplex

requires manual selection of the input. Unit testing is helpful when changes are made to the algorithms to ensure that structure prediction is unchanged. Unit testing is also important when RNAstructure is installed in environments that have not been previously tested, to ensure that the installation worked correctly.

Results

To demonstrate the utility of RNAstructure, an example of secondary structure prediction is provided, showing the input and output. The method utilized is free energy minimization and the example sequence is the 5S rRNA from *Pneumocystis carinii* [94].

Figure 2 shows a screenshot of the "Fold RNA Single Strand" input form from the JAVA GUI. This program predicts the lowest free energy structure for an input sequence and a set of low free energy structures called suboptimal structures. The name of the input sequence is already selected in this screenshot by having clicked the "Sequence File" button. The default output (CT) file and the default parameters for determining suboptimal structures were determined by the program. Save files can be written to accelerate the prediction of an alternative set of suboptimal structures with a subsequent calculation. This view of the GUI shows the checkbox unchecked and so no save file will be written. The three

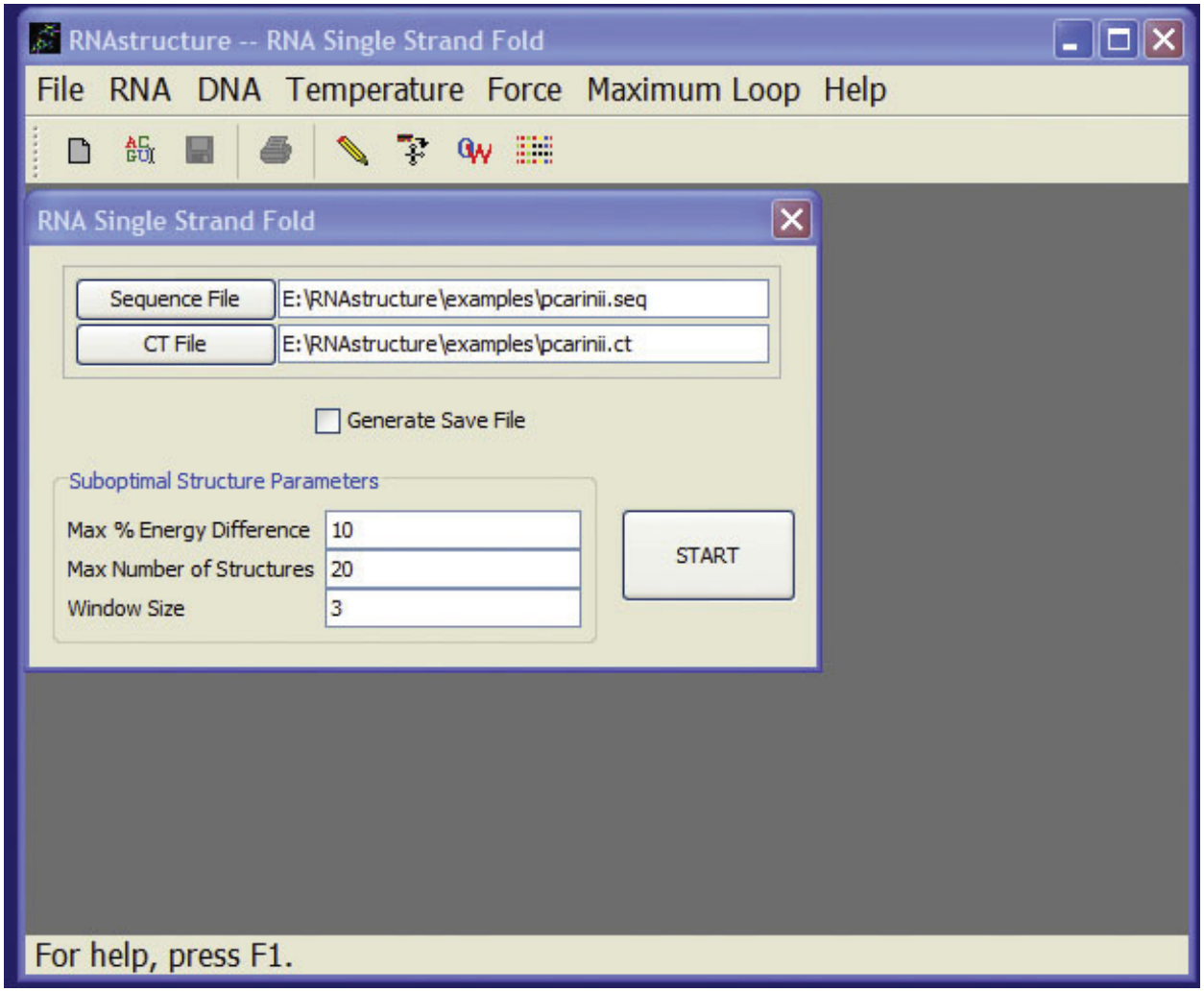

Figure 2 A screenshot of the JAVA GUI for predicting the lowest free energy structure for sequence "pcarinii.seq".

remaining parameters control the prediction of suboptimal structures. "Max % Energy Difference" and "Max Number of Structures" place limits on the number of suboptimal structures predicted, with a maximum of the percentage difference in free energy change above the lowest free energy structure or an absolute limit on the number of structures, respectively. Suboptimal structures will be generated until the percent energy difference above the lowest free energy structure is reached, 10% in this case, or 20 structures have been generated. "Window Size" determines how different the suboptimal structures must be from each other in the set [44]. Zero places no restriction and larger integers become increasingly more stringent in the number of different pairs required between structures.

After clicking "Start," the calculation proceeds and the structure is then drawn as shown in Figure 3. The lowest free energy structure is shown by default and the user can display other structures by selecting the "Draw" menu item or by typing control-up-arrow or control-down-arrow.

The same calculation could have been performed on the command line using the program "Fold." The command line for the same default method is "Fold pcarinii.seq pcarinii.ct." A set of postscript images of the predicted structure can then be rendered using "draw pcarinii.ct pcarinii.ps."

Behind the scenes, the structure prediction is performed using the RNA class. For this example, the class was instantiated using a constructor that reads sequences: RNA::RNA("pcarinii.seq", 2, true). The integer 2 indicates that the file is a sequence file and the bool true indicates that the sequence is RNA (as opposed to DNA). Structure prediction is then

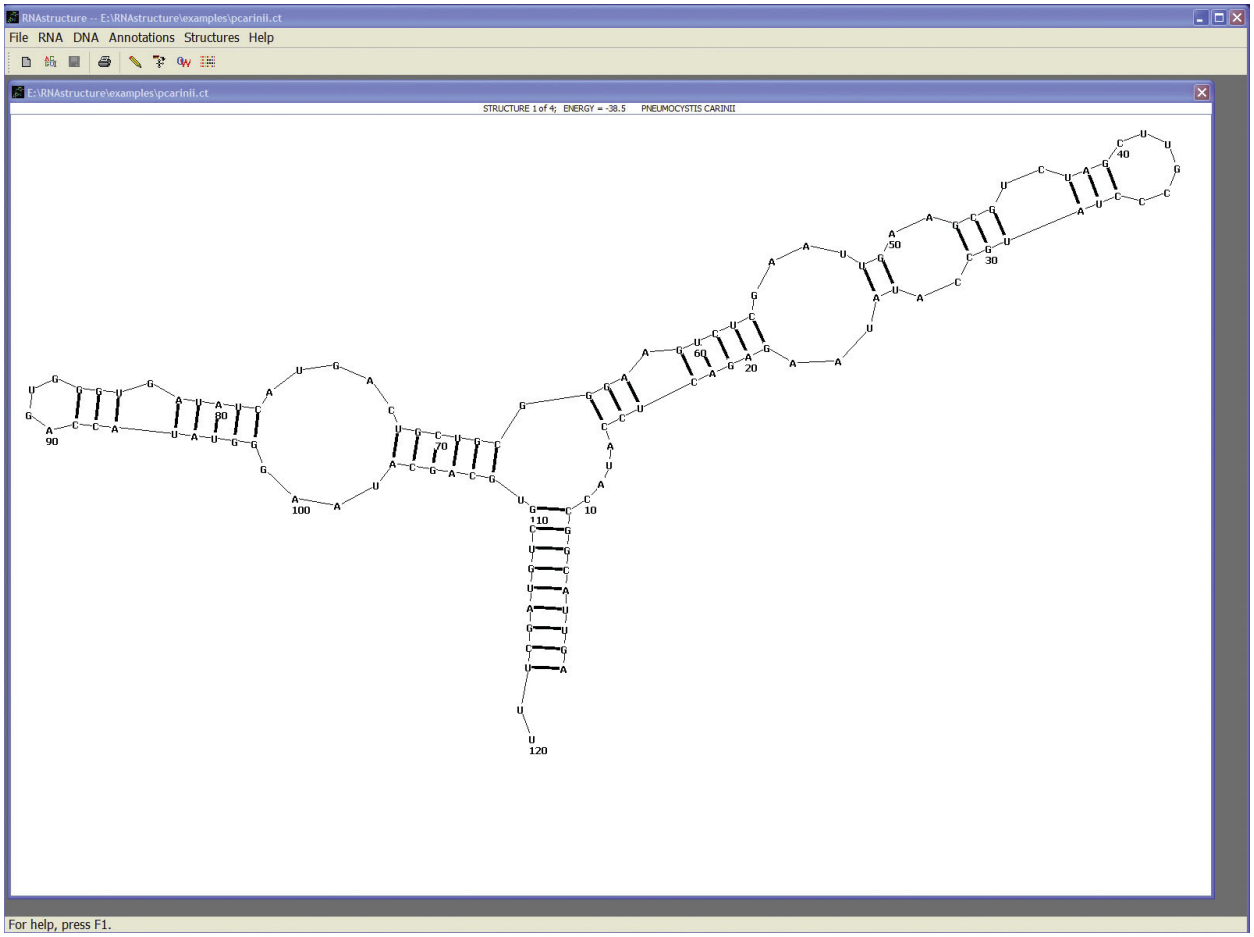

Figure 3 A screenshot of the JAVA GUI displaying the predicted lowest free energy structure for the sequence input as shown in Figure 2.

accomplished using RNA::FoldSingleStrand(10, 20, 3), where the parameters that control suboptimal prediction are the Max % Energy Difference, the Max number of structures, and the Window Size, respectively, a appeared in the GUI (Figure 2). The pairing can then be queried using RNA::GetPair or the drawing coordinates can be determined using RNA::DetermineDrawCoordinates.

Conclusions

RNAstructure is a software package for RNA secondary structure prediction and analysis. It is designed to make algorithms accessible for a variety of user needs. User-friendly GUIs are available for Windows, using native Windows code, and for Linux/Unix and Macintosh OS-X using JAVA. Text interfaces are provided for performing calculations on the command line or for scripting. Finally, a C++ class library is available to implement the algorithms into new programs. The package can be downloaded at http://rna.urmc.rochester.edu/RNAstructure.html.

Availability and requirements

- **Project name:** RNAstructure, version 5.0 and later
- **Project home page:** http://rna.urmc.rochester.edu/RNAstructure.html
- **Operating system(s):**

Text interfaces: Compilation is platform independent. Executables are provided for Microsoft Windows and for 32- and 64-bit Linux.
Class library: Compilation is platform independent.
Windows GUI: Windows XP or later.
JAVA GUI: Compilation is platform independent, but requires Sun JDK 1.6 or higher. Executables are provided for Macintosh OS-X (version 10.5 or later) and Linux.

- **Programming language:**

  Text interfaces, Class library, Windows GUI: C++
  JAVA GUI: JAVA

- **Other requirements:**

  Windows GUI: Compilation requires Microsoft Foundation Classes (MFC) as found in Microsoft Visual Studio 2005 or later and the Intel C++ compiler.
  JAVA GUI: Requires Sun JAVA JDK version 1.6 or later. Recompilation of SWIG interface requires SWIG version 1.3.39 or later.

- **License:** GNU GPL
- **Any restrictions to use by non-academics:** None.

Abbreviations

GUI: graphical user interface; ncRNA: non-coding RNA; siRNA: Small interfering RNA

Acknowledgements

This work was supported by National Institutes of Health grant R01GM076485 to D.H.M. The authors thank those who contributed to the code base.

Authors' contributions

JSR wrote the JAVA interface and the new text interfaces. DHM wrote the C ++ class library. Both authors contributed to the debugging of the code. DHM drafted the manuscript and both authors contributed to revisions.

Received: 23 October 2009 Accepted: 15 March 2010
Published: 15 March 2010

References

1. Eddy SR: **Non-coding RNA genes and the modern RNA world.** *Nature Reviews* 2001, **2**:919-929.
2. Mello CC, Conte D Jr: **Revealing the world of RNA interference.** *Nature* 2004, **431(7006)**:338-342.
3. Chow J, Heard E: **X inactivation and the complexities of silencing a sex chromosome.** *Curr Opin Cell Biol* 2009, **21(3)**:359-366.
4. Wu L, Belasco JG: **Let me count the ways: mechanisms of gene regulation by miRNAs and siRNAs.** *Mol Cell* 2008, **29(1)**:1-7.
5. Storz G, Gottesman S: **Versatile Roles of Small RNA Regulators in Bacteria.** *The RNA World* Cold Spring Harbor: Cold Spring Harbor Laboratory PressGesteland RF, Cech TR, Atkins JF, third 2006, 567-594.
6. Tucker BJ, Breaker RR: **Riboswitches as versatile gene control elements.** *Curr Opin Struct Biol* 2005, **15**:342-348.
7. Rodnina MV, Beringer M, Wintermeyer W: **How ribosomes make peptide bonds.** *Trends Biochem Sci* 2007, **32(1)**:20-26.
8. Doudna J, Cech T: **The chemical repertoire of natural ribozymes.** *Nature* 2002, **418**:222-228.
9. Mathews DH, Turner DH: **Prediction of RNA secondary structure by free energy minimization.** *Curr Opin Struct Biol* 2006, **16(3)**:270-278.
10. Gutell RR, Lee JC, Cannone JJ: **The accuracy of ribosomal RNA comparative structure models.** *Curr Opin Struct Biol* 2002, **12**:301-310.
11. Pace NR, Thomas BC, Woese CR: **Probing RNA structure, function, and history by comparative analysis.** *The RNA World* Cold Spring Harbor Laboratory PressGesteland RF, Cech TR, Atkins JF, 2 1999, 113-141.
12. Xia T, SantaLucia J Jr, Burkard ME, Kierzek R, Schroeder SJ, Jiao X, Cox C, Turner DH: **Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick pairs.** *Biochemistry* 1998, **37**:14719-14735.
13. Mathews DH, Sabina J, Zuker M, Turner DH: **Expanded sequence dependence of thermodynamic parameters provides improved prediction of RNA secondary structure.** *J Mol Biol* 1999, **288**:911-940.
14. Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH: **Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure.** *Proc Natl Acad Sci USA* 2004, **101**:7287-7292.
15. Zuker M, Stiegler P: **Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information.** *Nucleic Acids Res* 1981, **9**:133-148.
16. Ding Y, Chan CY, Lawrence CE: **RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble.** *RNA* 2005, **11(8)**:1157-1166.
17. Ding Y, Lawrence CE: **A statistical sampling algorithm for RNA secondary structure prediction.** *Nucleic Acids Res* 2003, **31(24)**:7280-7301.
18. Lu ZJ, Gloor JW, Mathews DH: **Improved RNA secondary structure prediction by maximizing expected pair accuracy.** *RNA* 2009, **15**:1805-1813.
19. Hamada M, Kiryu H, Sato K, Mituyama T, Asai K: **Prediction of RNA secondary structure using generalized centroid estimators.** *Bioinformatics* 2009, **25(4)**:465-473.
20. Dowell RD, Eddy SR: **Evaluation of several lightweight stochastic context-free grammars for RNA secondary structure prediction.** *BMC Bioinformatics* 2004, **5(1)**:71.
21. Doshi KJ, Cannone JJ, Cobaugh CW, Gutell RR: **Evaluation of the suitability of free-energy minimization using nearest-neighbor energy parameters for RNA secondary structure prediction.** *BMC Bioinformatics* 2004, **5(1)**:105.
22. Mathews DH: **Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization.** *RNA* 2004, **10**:1178-1190.
23. Mathews DH, Turner DH: **Dynalign: An algorithm for finding the secondary structure common to two RNA sequences.** *J Mol Biol* 2002, **317**:191-203.
24. Harmanci AO, Sharma G, Mathews DH: **PARTS: Probabilistic Alignment for RNA joinT Secondary structure prediction.** *Nucleic Acids Res* 2008, **36**:2406-2417.
25. Havgaard JH, Lyngso RB, Stormo GD, Gorodkin J: **Pairwise local structural alignment of RNA sequences with sequence similarity less than 40%.** *Bioinformatics* 2005, **21(9)**:1815-1824.
26. Hofacker IL, Fekete M, Stadler PF: **Secondary structure prediction for aligned RNA sequences.** *J Mol Biol* 2002, **319**:1059-1066.
27. Lück R, Steger G, Riesner D: **Thermodynamic prediction of conserved secondary structure: Application to the RRE element of HIV, the tRNA-like element of CMV and the mRNA of prion protein.** *J Mol Biol* 1996, **258**:813-826.
28. Duan S, Mathews DH, Turner DH: **Interpreting oligonucleotide microarray data to determine RNA secondary structure: application to the 3' end of Bombyx mori R2 RNA.** *Biochemistry* 2006, **45(32)**:9819-9832.
29. Deigan KE, Li TW, Mathews DH, Weeks KM: **Accurate SHAPE-directed RNA structure determination.** *Proc Natl Acad Sci USA* 2009, **106(1)**:97-102.
30. Hart JM, Kennedy SD, Mathews DH, Turner DH: **NMR-assisted prediction of RNA secondary structure: identification of a probable pseudoknot in the coding region of an R2 retrotransposon.** *J Am Chem Soc* 2008, **130(31)**:10233-10239.
31. Lu ZJ, Mathews DH: **Efficient siRNA selection using hybridization thermodynamics.** *Nucleic Acids Res* 2007, **36**:640-647.
32. Lu ZJ, Mathews DH: **Fundamental differences in the equilibrium considerations for siRNA and antisense oligodeoxynucleotide design.** *Nucleic Acids Res* 2008, **36**:3738-3745.
33. Mathews DH, Burkard ME, Freier SM, Wyatt JR, Turner DH: **Predicting oligonucleotide affinity to nucleic acid targets.** *RNA* 1999, **5**:1458-1469.
34. Tafer H, Ameres SL, Obernosterer G, Gebeshuber CA, Schroeder R, Martinez J, Hofacker IL: **The impact of target site accessibility on the design of effective siRNAs.** *Nat Biotechnol* 2008, **26(5)**:578-583.

35. Shao Y, Chan CY, Maliyekkel A, Lawrence CE, Roninson IB, Ding Y: **Effect of target secondary structure on RNAi efficiency.** *RNA* 2007, **13(10)**:1631-1640.
36. Long D, Lee R, Williams P, Chan CY, Ambros V, Ding Y: **Potent effect of target structure on microRNA function.** *Nat Struct Mol Biol* 2007, **14(4)**:287-294.
37. Aguirre-Hernandez R, Hoos HH, Condon A: **Computational RNA secondary structure design: empirical complexity and improved methods.** *BMC Bioinformatics* 2007, **8**:34.
38. Flamm C, Hofacker IL, Maurer-Stroh S, Stadler PF, Zehl M: **Design of multistable RNA molecules.** *RNA* 2001, **7**:254-265.
39. Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW Jr, Swanstrom R, Burch CL, Weeks KM: **Architecture and secondary structure of an entire HIV-1 RNA genome.** *Nature* 2009, **460(7256)**:711-716.
40. Washietl S, Hofacker IL, Stadler PF: **Fast and reliable prediction of noncoding RNAs.** *Proc Natl Acad Sci USA* 2005, **102(7)**:2454-2459.
41. Uzilov AV, Keegan JM, Mathews DH: **Detection of non-coding RNAs on the basis of predicted secondary structure formation free energy change.** *BMC Bioinformatics* 2006, **7(1)**:173.
42. Torarinsson E, Sawera M, Havgaard JH, Fredholm M, Gorodkin J: **Thousands of corresponding human and mouse genomic regions unalignable in primary sequence contain common RNA structure.** *Genome Res* 2006, **16(7)**:885-889.
43. Mathews DH, Andre TC, Kim J, Turner DH, Zuker M: **An updated recursive algorithm for RNA secondary structure prediction with improved thermodynamic parameters.** *Molecular Modeling of Nucleic Acids* American Chemical SocietyLeontis NB, SantaLucia J Jr 1998, 246-257.
44. Zuker M: **On finding all suboptimal foldings of an RNA molecule.** *Science* 1989, **244**:48-52.
45. Steger G, Hofmann H, Fortsch J, Gross HJ, Randles JW, Sanger HL, Riesner D: **Conformational transitions in viroids and virusoids: Comparison of results from energy minimization algorithm and from experimental data.** *J Biomol Struct Dyn* 1984, **2(3)**:543-571.
46. Harmanci AO, Sharma G, Mathews DH: **Efficient pairwise RNA structure prediction using probabilistic alignment constraints in Dynalign.** *BMC Bioinformatics* 2007, **8**:130.
47. Wuchty S, Fontana W, Hofacker IL, Schuster P: **Complete suboptimal folding of RNA and the stability of secondary structures.** *Biopolymers* 1999, **49**:145-165.
48. Harmanci AO, Sharma G, Mathews DH: **Stochastic sampling of the RNA structural alignment space.** *Nucleic Acids Res* 2009, **37**:4063-4075.
49. Smit S, Rother K, Heringa J, Knight R: **From knotted to nested RNA structures: a variety of computational methods for pseudoknot removal.** *RNA* 2008, **14(3)**:410-416.
50. Mathews DH: **RNA secondary structure analysis using RNAstructure.** *Current Protocols in Bioinformatics* New York: John Wiley and Sons, IncBaxevanis AD, Davison DB, Page RDM, Petsko GA, Stein LD, Stormo GD 2006, 12.16.11-12.16.14.
51. Mathews DH: **Predicting the secondary structure common to two sequences with Dynalign.** *Current Protocols in Bioinformatics* New York: John Wiley and Sons, IncBaxevanis AD, Davison DB, Page RDM, Petsko GA, Stein LD, Stormo GD 2005, 12.14.11-12.14.11.
52. Mathews DH, Turner DH, Zuker M: **RNA secondary structure prediction.** *Current Protocols in Nucleic Acid Chemistry* New York: John Wiley and SonsBeaucage SL, Bergstrum DE, Glick GD, Jones RA 2000, **11**:2.1-2.10.
53. Lu ZJ, Mathews DH: **OligoWalk: An online siRNA design tool utilizing hybridization thermodynamics.** *Nucleic Acids Res* 2008, **36**:W104-W108.
54. Lu ZJ, Turner DH, Mathews DH: **A set of nearest neighbor parameters for predicting the enthalpy change of RNA secondary structure formation.** *Nucleic Acids Res* 2006, **34**:4912-4924.
55. Matveeva OV, Mathews DH, Tsodikov AD, Shabalina SA, Gesteland RF, Atkins JF, Freier SM: **Thermodynamic criteria for high hit rate antisense oligonucleotide design.** *Nucleic Acids Res* 2003, **31(17)**:4989-4994.
56. Allawi HT, SantaLucia J Jr: **Thermodynamics and NMR of internal G·T mismatches in DNA.** *Biochemistry* 1997, **36(34)**:10581-10594.
57. Allawi HT, SantaLucia J Jr: **NMR solution structure of a DNA dodecamer containing single G·T mismatches.** *Nucleic Acids Res* 1998, **26(21)**:4925-4934.
58. Allawi HT, SantaLucia J Jr: **Nearest-neighbor thermodynamics of internal A·C mismatches in DNA: sequence dependence and pH effects.** *Biochemistry* 1998, **37(26)**:9435-9444.
59. Allawi HT, SantaLucia J Jr: **Thermodynamics of internal C·T mismatches in DNA.** *Nucleic Acids Res* 1998, **26(11)**:2694-2701.
60. Allawi HT, SantaLucia J Jr: **Nearest neighbor thermodynamic parameters for internal G·A mismatches in DNA.** *Biochemistry* 1998, **37(8)**:2170-2179.
61. Peyret N, Seneviratne PA, Allawi HT, SantaLucia J Jr: **Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A·A, C·C, G·G, and T·T mismatches.** *Biochemistry* 1999, **38**:3468-3477.
62. Plum GE, Grollman AP, Johnson F, Breslauer KJ: **Influence of the oxidatively damaged adduct 8-oxodeoxyguanosine on the conformation, energetics, and thermodynamic stability of a DNA duplex.** *Biochemistry* 1995, **34(49)**:16148-16160.
63. Bolewska K, Zielenkiewicz A, Wierzchowski KL: **Deoxydodecanucleotide heteroduplex d(TTTTATAATAAA). d(TTTATTATAAAA) containing the promoter Pribnow sequence TATAAT. I. Double-helix stability by UV spectrophotometry and calorimetry.** *Nucleic Acids Res* 1984, **12(7)**:3245-3256.
64. Breslauer KJ, Frank R, Blocker H, Marky LA: **Predicting DNA duplex stability from the base sequence.** *Proc Natl Acad Sci USA* 1986, **83(11)**:3746-3750.
65. Leonard GA, Thomson J, Watson WP, Brown T: **High-resolution structure of a mutagenic lesion in DNA.** *Proc Natl Acad Sci USA* 1990, **87(24)**:9573-9576.
66. Tibanyenda N, De Bruin SH, Haasnoot CAG, Marel Van Der GA, Van Boom JH, Hilbers CW: **The effect of single base-pair mismatches on the duplex stability of d(T-A-T-T-A-A-T-A-T-C-A-A-G-T-T-G) · d(C-A-A-C-T-T-G-A-T-A-T-T-A-A-T-A).** *Eur J Biochem* 1984, **139**:19-27.
67. Plum GE, Grollman AP, Johnson F, Breslauer KJ: **Influence of an exocyclic guanine adduct on the thermal stability, conformation, and melting thermodynamics of a DNA duplex.** *Biochemistry* 1992, **31(48)**:12096-12102.
68. Arnold FH, Wolk S, Cruz P, Tinoco I Jr: **Structure, dynamics, and thermodynamics of mismatched DNA oligonucleotide duplexes d (CCCAGGG)2 and d(CCCTGGG)2.** *Biochemistry* 1987, **26(13)**:4068-4075.
69. Ratmeyer L, Vinayak R, Zhong YY, Zon G, Wilson WD: **Sequence specific thermodynamic and structural properties for DNA·RNA duplexes.** *Biochemistry* 1994, **33(17)**:5298-5304.
70. Williams AP, Longfellow CE, Freier SM, Kierzek R, Turner DH: **Laser temperature-jump, spectroscopic, and thermodynamic study of salt effects on duplex formation by dGCATGC.** *Biochemistry* 1989, **28(10)**:4283-4291.
71. Sugimoto N, Honda K, Sasaki M: **Application of the thermodynamic parameters of DNA stability prediction to double-helix formation of deoxyribooligonucleotides.** *Necleosides & Nucleotides* 1994, **13**:1311-1317.
72. Wu P, Nakano S, Sugimoto N: **Temperature dependence of thermodynamic properties for DNA/DNA and RNA/DNA duplex formation.** *Eur J Biochem* 2002, **269(12)**:2821-2830.
73. Hall KB, McLaughlin LW: **Thermodynamic and structural properties of pentamer DNA.DNA, RNA.RNA, and DNA.RNA duplexes of identical sequence.** *Biochemistry* 1991, **30(44)**:10606-10613.
74. Sugimoto N, Nakano S, Yoneyama M, Honda K: **Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes.** *Nucleic Acids Res* 1996, **24(22)**:4501-4505.
75. Bommarito S, Peyret N, SantaLucia J Jr: **Thermodynamic parameters for DNA sequences with dangling ends.** *Nucleic Acids Res* 2000, **28(9)**:1929-1934.
76. Senior M, Jones RA, Breslauer KJ: **Influence of dangling thymidine residues on the stability and structure of two DNA duplexes.** *Biochemistry* 1988, **27(10)**:3879-3885.
77. Aboul-ela F, Koh D, Tinoco I Jr, Martin FH: **Base-base mismatches. Thermodynamics of double helix formation for dCA3XA3G + dCT3YT3G (X, Y = A, C, G, T).** *Nucleic Acids Res* 1985, **13(13)**:4811-4824.
78. Riccelli PV, Hilario J, Gallo FJ, Young AP, Benight AS: **DNA and RNA oligomer sequences from the 3′ noncoding region of the chicken glutamine synthetase gene from intramolecular hairpins.** *Biochemistry* 1996, **35(48)**:15364-15372.
79. Rentzeperis D, Alessi K, Marky LA: **Thermodynamics of DNA hairpins: contribution of loop size to hairpin stability and ethidium binding.** *Nucleic Acids Res* 1993, **21(11)**:2683-2689.
80. Vallone PM, Paner TM, Hilario J, Lane MJ, Faldasz BD, Benight AS: **Melting studies of short DNA hairpins: influence of loop sequence and adjoining base pair identity on hairpin thermodynamic stability.** *Biopolymers* 1999, **50(4)**:425-442.

81. Hilbers CW, Haasnoot CA, de Bruin SH, Joordens JJ, Marel van der GA, van Boom JH: **Hairpin formation in synthetic oligonucleotides.** *Biochimie* 1985, **67(7-8)**:685-695.
82. Antao VP, Tinoco I Jr: **Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops.** *Nucleic Acids Res* 1992, **20**:819-824.
83. Antao VP, Lai SY, Tinoco I Jr: **A thermodynamic study of unusually stable RNA and DNA hairpins.** *Nucleic Acids Res* 1991, **19**:5901-5905.
84. Moody EM, Bevilacqua PC: **Thermodynamic coupling of the loop and stem in unusually stable DNA hairpins closed by CG base pairs.** *J Am Chem Soc* 2003, **125(8)**:2032-2033.
85. Nakano M, Moody EM, Liang J, Bevilacqua PC: **Selection for thermodynamically stable DNA tetraloops using temperature gradient gel electrophoresis reveals four motifs: d(cGNNAg), d(cGNABg), d (cCNNGg), and d(gCNNGc).** *Biochemistry* 2002, **41(48)**:14281-14292.
86. LeBlanc DA, Morden KM: **Thermodynamic characterization of deoxyribooligonucleotide duplexes containing bulges.** *Biochemistry* 1991, **30(16)**:4042-4047.
87. Morden KM, Chu YG, Martin FH, Tinoco I Jr: **Unpaired cytosine in the deoxynucleotide duplex $dCA_3CA_3G{\cdot}dCT_6G$ is outside of the helix.** *Biochemistry* 1983, **22**:5557-5563.
88. Li Y, Zon G, Wilson WD: **Thermodynamics of DNA duplexes with adjacent G.A mismatches.** *Biochemistry* 1991, **30(30)**:7566-7572.
89. Li Y, Agrawal S: **Oligonucleotides containing G·A pairs: effect of flanking sequences on structure and stability.** *Biochemistry* 1995, **34(31)**:10056-10062.
90. Kadrmas JL, Ravin AJ, Leontis NB: **Relative stabilities of DNA three-way, four-way and five-way junctions (multi-helix junction loops): unpaired nucleotides can be stabilizing or destabilizing.** *Nucleic Acids Res* 1995, **23**:2212-2222.
91. Leontis NB, Kwok W, Newman JS: **Stability and structure of three-way DNA junctions containing unpaired nucleotides.** *Nucleic Acids Res* 1991, **19**:759-766.
92. Ladbury JE, Sturtevant JM, Leontis NB: **The thermodynamics of formation of a three-strand, DNA three-way junction complex.** *Biochemistry* 1994, **33(22)**:6828-6833.
93. Sugimoto N, Nakano S, Katoh M, Matsumura A, Nakamuta H, Ohmichi T, Yoneyama M, Sasaki M: **Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes.** *Biochemistry* 1995, **34**:11211-11216.
94. Szymanski M, Barciszewska MZ, Barciszewski J, Erdmann VA: **5S ribosomal RNA database Y2K.** *Nucleic Acids Res* 2000, **28**:166-167.

doi:10.1186/1471-2105-11-129
Cite this article as: Reuter and Mathews: **RNAstructure: software for RNA secondary structure prediction and analysis.** *BMC Bioinformatics* 2010 **11**:129.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central