Published in final edited form as:

Curr Protoc Nucleic Acid Chem. 2007 March ; CHAPTER 11: Unit-11.2. doi:
10.1002/0471142700.nc1102s28.

RNA Secondary Structure Prediction

David H. Mathews, Douglas H. Turner, and Richard M. Watson

University of Rochester, Rochester, New York

Abstract

Protocols are provided for predicting RNA secondary structure with the user-friendly RNAstructure desktop computer program and the RNAstructure Webserver. The minimum free energy structure and a set of suboptimal structures with similar free energies are predicted. Prediction of high-affinity oligonucleotide binding sites to a structured RNA target is also presented.

Keywords

RNA secondary structure prediction; free energy minimization; thermodynamics; binding affinity; RNA folding; partition function

This unit details the steps for predicting the secondary structure of an RNA sequence and for predicting the equilibrium binding affinity of a complementary RNA or DNA oligonucleotide to an RNA target (see Basic Protocol 2). Two protocols are given for secondary structure prediction—one for the desktop computer program RNAstructure (see Basic Protocol 1) and another for the RNAstructure Webserver (see Basic Protocol 3). The RNAstructure Webserver is an internet adaptation of several RNAstructure command-line interfaces. The desktop program is a Java application that can run on Apple OS X, Microsoft Windows, or Linux. It includes OligoWalk for predicting oligonucleotide binding affinities (see Basic Protocol 2). The Webserver is accessed via the internet using a standard web browser. The minimum free energy structure and a set of suboptimal structures with low free energies are predicted by both programs. Both programs have the same structure prediction accuracy and include the ability to guide secondary structure prediction using chemical modification or SHAPE data.

Key References

Mathews et al., 1999b, 2004. See above.

Derives the thermodynamic parameters used by the secondary structure prediction algorithm and tabulates the accuracy of the algorithm with a large database of structures from sequence comparisons.

Zuker, 1989. See above.

Explains the method for predicting suboptimal structures using a dynamic programming algorithm.

Internet Resources

<http://rna.urmc.rochester.edu>

The Mathews Lab homepage is the source for downloading RNAstructure and for running the RNAstructure webserver.

<http://rna.ucsc.edu/rnacenter/xrna/xrna.html>

The XRNA homepage at the Santa Cruz RNA Center is the source of XRNA, which can be used for creating publication-quality RNA secondary structure diagrams.

BASIC PROTOCOL 1: PREDICTING SECONDARY STRUCTURE AND BASE-PAIR PROBABILITIES

Secondary structure prediction by free energy minimization is the core functionality of RNAstructure. The secondary structure prediction algorithm predicts the lowest free energy structure, which is the most probable secondary structure. It also predicts low free energy structures, called suboptimal structures, which suggest possible alternative structures (Zuker, 1989). Low free energy structures can be color-annotated according to base-pair probabilities determined by a partition function calculation, and these probabilities imply confidence in the prediction of pairs (Mathews, 2004).

This protocol describes the basic use of the program on a Microsoft Windows computer. Many other options are available, and these are described in detail in the help manual, which can be accessed from the program by choosing the Program Help item from the Help menu. The protocol requires some basic familiarity with Windows. RNAstructure also runs natively on Apple Mac OS X and GNU/Linux. Instructions for installing RNAstructure on Linux and Mac OS X can be found on the RNAstructure website: <http://rna.urmc.rochester.edu/Overview/>.

Materials

Hardware: A personal computer running Microsoft Windows is required. RNAstructure is compatible with Windows versions 7, 8, and 10. Table 11.2.1 shows the computation times for secondary structure prediction and partition function calculation as a function of sequence length.

Software: The RNAstructure package must be downloaded at <http://rna.urmc.rochester.edu/RNAstructure.html>. User registration is requested so that e-mail can be sent when significant upgrades are available or if significant bugs are found. The list of registered users is not shared with others.

Install RNAstructure

- 1 Download the Windows installer from the RNAstructure website (<http://rna.urmc.rochester.edu/RNAstructure.html>).
Run the downloaded file (**RNAstructureWindows.exe**) to launch the installation routine. During the installation, the user can choose the installation destination. Most users will want to accept the default location. The installation will add RNAstructure to the start menu and will add an icon on the desktop.

Enter a sequence

- 2 Start RNAstructure by choosing it from the Windows start menu or from the desktop icon. Open the sequence editor by choosing **New Sequence** at the top of the **File** menu.

Figure 11.2.1 shows the layout of the sequence editor. The name or a short description of the sequence can be entered in the field labeled

Title. *This information will be displayed with predicted structures. The field labeled **Comment** is available for entering longer comments that are stored with the sequence, but will not be shown with the predicted structure. The sequence is entered in the field labeled **Sequence**.*

- 3 Sequences can be entered manually or pasted from another program. To paste a sequence, copy it to the clipboard, click with the mouse in the Sequence field, then choose **Paste** from the **Edit** menu option or press Ctrl-v (Command-v on Mac). Save the sequence by choosing **Save Sequence** from the **File** menu.

The sequence should consist of A (adenine), C (cytidine), G (guanine), U (uridine), T (treated as uridine in RNA folding), and X (a nucleotide that neither base-pairs nor stacks). Note that nucleotides entered as lowercase characters are forced single-stranded in the structure prediction, so most nucleotides should be uppercase. Spaces and carriage returns are ignored during structure prediction.

*The sequence editor has several features. Clicking the **Format Sequence** button on the editor window will automatically format the sequence into lines of 50 nucleotides with a space after every fifth nucleotide. The sequence is recited aloud over speakers (if available) when the **Speak Sequence** button is clicked. This provides a convenient method for checking the accuracy of a manually entered sequence. Clicking the **Fold as RNA** button saves the sequence and opens the secondary structure prediction window. Clicking **Fold as DNA** saves the sequence and opens a structure prediction window that predicts secondary structure of a DNA sequence. The sequence editor colors characters to indicate their status: standard nucleobases are black, lowercase (unpaired) nucleobases are blue, X (unknown) is green, and any invalid characters are shown in red. The presence of invalid characters in the sequence will result in a warning message upon saving it*

Predict the secondary structure

- 4 After saving the input sequence (above), choose **Fold RNA Single Strand** from the **File** menu. The RNA secondary structure prediction window will open, as shown in Figure 11.2.2.
- 5 Choose the name of the sequence file by clicking the browse (...) button. This will open a standard File Selection dialog for selecting the file. After the file has been selected, the name of the file will appear in the **Sequence File** field, and default values will have been entered in all other fields. The output of the calculation (a predicted secondary structure and a set of low free energy structures) will be stored in a CT (connection table) file. The default CT file

name is the same as the sequence file, but with a .ct file extension. The default name can be changed by clicking the browse button (...) next to the CT File field or by directly editing the **CT File** text field.

*Three parameters control the generation of suboptimal structures. The first parameter is maximum percent energy difference of suboptimal structures as compared to the lowest free energy structure. Larger percent energy differences allow the prediction of more suboptimal structures, whereas a maximum percent of zero allows only prediction of the lowest free energy structure. The maximum percent energy difference can be changed from the default by manually typing in the text box adjacent to **Max % Energy Difference**. The second parameter is the maximum number of structures. This places an absolute limit on the number of suboptimal structures and can be changed manually by typing in the text box adjacent to **Max Number of Structures**. The third parameter is the window size, which specifies how different each suboptimal structure must be, as compared to all other predicted structures. Each structure must have at least window number of pairs separated by at least window nucleotides from all pairs in all other structures. A window size of zero does not place any restriction. Larger window sizes result in structures with greater difference, but also result in fewer predicted structures. The default window parameter can be manually changed by typing in the text box adjacent to **Window Size**.*

*The check box next to **Generate Save File** is checked by default. This will generate a Save file that can be used to show energy dot plots or to predict a different set of suboptimal structures using different suboptimal structure parameters. The online help contains entries labeled **Dot Plot** and **Refolding a Saved Sequence** that explain these functions. The name of the Save file is the same as the CT file, except that the file has a.sav extension instead of.ct.*

- 6 RNAstructure can predict secondary structures with user-specified constraints or restraints. These are entered by choosing the appropriate menu item under the **Force** menu.
 - a. **Force Base Pairs** is used to specify required base pairs in the structure.
 - b. **Chemical Modification** is used to specify nucleotides that are accessible to chemical modification, i.e., single-stranded, at the end of a helix, or in or adjacent to a GU pair.
 - c. **Double Stranded** is used to specify nucleotides that must be base-paired, without specifying to which nucleotide they are paired.

- d. **FMN Cleavage** is used to indicate Us that are in GU base pairs.
- e. **Single Stranded** is used to indicate unpaired nucleotides.
- f. **Prohibit Basepairs** is used to indicate specific base pairs that are not allowed.

*Each of these options opens a dialog box for entering the specified constraints. In the dialog box, **Apply** can be clicked to add the currently-entered constraint and keep the dialog open to enter more constraints, while **Cancel** can be clicked to close the dialog when all constraints have been entered. The entered constraints are displayed on the screen if **Show Current Constraints** is chosen from the **Force** menu. **Reset Current Constraints** removes all entered constraints. **Save Constraints** can be used to save all constraints to a file (with a.con extension). Constraints can be read from a file by choosing **Restore Constraints**.*

- 7 **Read SHAPE Reactivity -- Pseudo-Energy Constraint** under the Force menu option can read normalized SHAPE mapping reactivities. This is the recommended method for using SHAPE mapping data. The SHAPE data file format is specified at http://rna.urmc.rochester.edu/Text/File_Formats.html. It consists of a list of nucleotide index numbers, where nucleotides are numbered consecutively starting at the first nucleotide, followed by the normalized reactivity to SHAPE for each nucleotide. A normalized reactivity of -500 or lower indicates no data was available for a specific nucleotide.
- 8 The default temperature of the prediction, 37°C (310.15 K), can be changed by choosing the **Temperature** menu item. This opens a dialog box in which the desired temperature of structure prediction can be entered in kelvins.

The stability of RNA secondary structure is best determined at 37°C and is poorly determined at temperatures far from 37°C (Lu et al., 2006). For temperatures below 20°C or above 60°C, a more accurate secondary structure can be predicted by using the default folding temperature of 37°C

- 9 Click the **Start** button to begin the secondary structure prediction calculation. A progress bar will appear to provide an estimate of the calculation's progress. When the calculation ends, the program will draw the predicted secondary structure. Figure 11.2.3 shows the secondary structure predicted for the *Drosophila melanogaster* 5S rRNA sequence (with no user-specified constraints). The user can choose whether or not they want the program to automatically draw predicted structures by selecting an option under **File > Draw structures after calculations**. The functionality of the RNAstructure **Draw Window** is detailed in the later steps of this protocol.

Predict base-pair probabilities with a partition function calculation

- 10 Open the partition function window by choosing **Partition Function RNA** from the **RNA** menu. Choose the sequence file by clicking the browse (...) button next to the **Sequence Name** text field. After the calculation, the base-pair probability data are saved to disk in a Partition Function Save (.pfs) file. A default Save file name automatically appears in the field next to the **Save File** button after a sequence has been chosen. The default name can be changed by clicking the browse (...) button or by editing the field directly.
- 11 Enter constraints and change the temperature to that used for secondary structure prediction.

The partition function calculation can accommodate the same structure restraints as secondary structure prediction (step 6). The partition function can also be performed at temperatures other than the default 37°C in the same way as structure prediction (step 7). To get the most information from base pair probabilities, these settings should be set the same as for structure prediction

- 12 Start the calculation by clicking the **Start** button. A progress bar will appear to show the progress of the calculation. When it is completed, a probability dot plot will be displayed that indicates the probability of all valid canonical base pairs (AU, GC, and GU). First, reduce the probability range of base pairs by choosing **Plot Range** under the **Draw** menu option. This will open a dialog box in which a min and max range are entered. Dots are registered by $-\log_{10}$ (Base Pair Probability). Set the maximum to 2, so that all base pairs with pairing probability 0.01 (1%) will be displayed. Click **OK**. The plot will now resemble the plot in Figure 11.2.4.

*The probability dot plot window provides several features for analyzing the dot plot information. By clicking on a dot, the message window at the top of the Draw Window displays the identity of the base pair and $-\log_{10}$ of the base-pair probability. The size of the plot in the window can be changed by choosing **Zoom** under the **Draw** menu. Alternatively, pressing Ctrl-left arrow zooms out and pressing Ctrl-right arrow zooms in. The base-pair probabilities can be written to a tab-delimited text file for analysis with other programs by choosing **Write Dot Plot File** under the **Output Plot** menu option. The text file contains the $-\log_{10}$ of base-pair probability for all base pairs and not just the pairs that are currently displayed on the screen. Secondary structures composed of only probable base pairs can be output by choosing **Output Probable Structure** under the **Output Plot** menu option.*

The probability dot plot window is created from the data stored in the Partition Function Save (.pfs) file. To draw the probability dot plot

again at a later time, choose **Dot Plot Partition Function** from the **File** menu. This will open a File Selection dialog from which the Partition Function Save file can be chosen

Color annotate the predicted secondary structures according to base-pair probabilities

- 13 Return to the structure drawing window that contains the predicted secondary structures for the *D. melanogaster* 5S rRNA sequence. If the window has been closed, the secondary structure can be redrawn using the data stored in the CT file. A new drawing window can be opened by choosing **Draw** under the **File** menu. This will open a File Selection dialog from which the CT file can be chosen.
- 14 To color annotate a predicted secondary structure for which a partition function calculation has been performed, choose **Add Probability Annotation** from the **Annotations** menu with the drawing window open. This will open a File Selection dialog from which the Partition Function Save (.pfs) file can be selected. The secondary structure will then have color annotation with the most probable base pairs in red and the least probable pairs in purple. A legend that indicates the association between color and probability range is shown at the bottom of the image.

Figure 11.2.5 shows the predicted lowest free energy structure for the *D. melanogaster* 5S rRNA with probability annotation.

The drawing window has several functions to facilitate the analysis of secondary structures. When suboptimal structures are included in the CT file, the displayed structure can be changed by choosing **Goto Structure...** under the **Draw** menu. This will open a window that indicates the current structure number and allows that number to be changed. The lowest free energy structure is structure 1, and folding free energy increases with the structure number. Alternatively, pressing Ctrl-up arrow increases the number of the currently displayed structure and pressing Ctrl-down arrow lowers the number of the currently displayed structure. The number of the currently displayed structure is indicated at the top of the window. As indicated in Figures 11.2.3 and 11.2.5, there are a total of four low-energy secondary structures predicted for the *D. melanogaster* 5S rRNA using the default parameters. The size of the structures on the screen can be changed by choosing **Zoom** from the **Draw** menu. Alternatively, pressing Ctrl-left arrow zooms out and pressing Ctrl-right arrow zooms in

- 15 To produce publication-quality drawings of secondary structures, base pairs predicted by RNAstructure can be exported to the drawing program XRNA,

which is available for free download from the Santa Cruz RNA Center at <http://rna.ucsc.edu/rnacenter/xrna/xrna.html>. XRNA is Java-based and will run on most computers. To export helix locations in a file that is readable by XRNA, choose **Write Helix (Text) File** from the **Draw** menu.

BASIC PROTOCOL 2: PREDICTING BINDING AFFINITIES OF OLIGONUCLEOTIDES COMPLEMENTARY TO AN RNA TARGET WITH OLIGOWALK

RNAstructure includes the OligoWalk program for predicting the binding affinity of complementary oligonucleotides to an RNA target. For an RNA sequence of N nucleotides, OligoWalk predicts an overall free energy of binding of all $(N - L + 1)$ oligonucleotides of length L that are complementary to the target. Hence, the binding region considered is walked down the length of the sequence. The overall free energy of binding, ΔG_{37}° overall includes the effects of self structure in the target and self structure in the oligonucleotides.

This protocol uses RNAstructure on a Microsoft Windows platform. RNAstructure also runs natively on Apple Mac OS X and GNU/Linux.. Instructions for installing RNAstructure on Linux and Mac can be found on the RNAstructure website: <http://rna.urmc.rochester.edu/Overview/>.

Materials

Hardware: A personal computer running Microsoft Windows is required. RNAstructure is compatible with Windows 7, 8, and 10. Table 11.2.2 indicates typical calculation times for two different length target sequences.

Software: The RNAstructure package must be downloaded at <http://rna.urmc.rochester.edu>. User registration is requested so that e-mail can be sent when significant upgrades are available or if significant bugs are found. The list of registered users is not shared with others.

Install RNA structure

- 1 Install RNAstructure and predict the secondary structure of the target RNA (see Basic Protocol 1).

Start the OligoWalk calculation

- 2 Open the OligoWalk input window (shown in Fig. 11.2.6) by choosing **RNA OligoWalk** from the **RNA** menu.
- 3 Click on the browse (...) button next to the **CT File** text field and choose the target secondary structure (stored in CT format; see Basic Protocol 1) using the File Selection dialog. A default name is then chosen for output of the calculated free energy change parameters. This name is the same as the CT file, but with the .ct extension changed to .rep.

*The default name can be changed by clicking on the browse button next to the **Report File text field**. The report file stores the calculated parameters as tab-delimited text that can be opened in most spreadsheet programs, such as OpenOffice or Microsoft Excel*

- 4 Choose a Mode for the calculation by clicking the button (see Fig. 11.2.6) adjacent to one of the three options: **Break Local Structure**, **Refold Whole RNA for Each Oligomer**, or **Do Not Consider Target Structure**.

***Do Not Consider Target Structure** is the fastest mode, but is not recommended because the target RNA secondary structure is neglected. **Refold Whole RNA for Each Sequence** predicts a new lowest free energy structure after oligonucleotide binding by predicting a structure with the nucleotides that are bound by the oligonucleotide forced to be unpaired. This mode is the slowest, but best approximates equilibrium. **Break Local Structure** is much faster than **Refold Whole RNA for Each Sequence**, and it calculates the secondary structure formation free energy change of the original structure with any pairs that involve the oligonucleotide-bound nucleotides broken. For most applications, **Break Local Structure** is a good compromise between accuracy and calculation time*

- 5 Below the Mode options (see Fig. 11.2.6), note the check box labeled **Include Target Suboptimal Structures in Free Energy Calculation**. When checked, the cost (in free energy change) of opening base pairs in the target structure is calculated for each suboptimal secondary structure. The contribution made by each suboptimal structure to the total cost of opening target self structure is weighted according to the folding free energy of each structure. In general, it is recommended that this box be checked, so that alternative possible secondary structures are considered.

Input information about the oligonucleotides

- 6 Enter the oligonucleotide length in the text box to the right of **Oligo Length** (see Fig. 11.2.6). Oligonucleotides can be either RNA or DNA; the choice between these is indicated by the radio button in the **Oligomer Chemistry** box. Also choose an oligonucleotide concentration. The concentration units can be changed between mM, μ M, nM, and pM by clicking the down arrow to the right of the displayed units. For the example shown in Figure 11.2.6, DNA nucleotides of length 18 are considered at a concentration of 1 μ M.
- 7 Next, the region for oligonucleotide binding can be reduced by adjusting the **Start** and **Stop** locations (see Fig. 11.2.6, bottom of dialog box). Start refers to the target RNA nucleotide bound to the 3' end of the first oligonucleotide, and Stop refers to the nucleotide bound to the 5' end of the last oligonucleotide in

the walk. By default, these limits are set to the 5' and 3' ends of the target sequence. To adjust the limits, use the up and down arrows next to the value fields. Limiting the area of interest on the target RNA strand reduces the calculation time.

- 8 The temperature of the calculation can be changed from the default of 37°C by choosing the Temperature menu item. This opens a dialog that takes the temperature in kelvins.
- 9 Finally, start the calculation by clicking the **Start** button. A progress bar will appear to show the progress of the calculation. When the calculation is complete, RNAstructure will open the OligoWalk output window as shown in Figure 11.2.7.

Navigating the OligoWalk output window (Fig. 11.2.7)

- 10 The OligoWalk output window provides an interactive method for displaying the calculated results. The target sequence is drawn left to right across the window in a 5' to 3' direction. Red nucleotides are predicted to be base-paired in the lowest free energy structure. Black nucleotides are predicted to be single stranded. The currently displayed oligonucleotide is below the target sequence in a 3' to 5' direction. The position along the target of the currently displayed oligonucleotide is indicated in the upper left-hand corner of the display. Oligonucleotides are numbered according to the 5'-most binding nucleotide in the target; therefore, the oligonucleotides are numbered from 1. Below the current oligonucleotide is the backbone chemistry (RNA or DNA) and concentration of the oligonucleotide.
- 11 The results at the top of the screen are free energy changes at 37°C in kcal/mol. "Overall ΔG_{37}° " is the total free energy change of binding for a structured target and structured oligonucleotide. "Duplex ΔG_{37}° " is the free energy change of duplex formation between the oligonucleotide and target, without the cost of opening self-structure. "Break Target ΔG_{37}° " is the free energy cost of opening target secondary structure. "Oligo-Self ΔG_{37}° " and "Oligo-Oligo ΔG_{37}° " are the free energy costs of opening unimolecular and bimolecular self-structure in the oligonucleotide, respectively. The T_m is the melting temperature of duplex formation in °C, not accounting for self-structure in target or oligonucleotide.

*The graph, by default, shows the "Overall ΔG_{37}° " and "Duplex ΔG_{37}° " profile along the target sequence. The graph color is identical to the text color above (for each calculated free energy change), except that the color for "Overall ΔG_{37}° " for the current oligonucleotide is in red. The free energy term that is graphed can be changed by selecting an option under the **Oligo Graph** menu*

- 12 The currently displayed oligonucleotide can be changed in several ways. The left and right arrow buttons move the displayed oligonucleotide towards the 5' or 3' end, respectively. The currently displayed nucleotide is moved ten nucleotides by clicking the buttons labeled « or ». A navigation window can be shown by clicking the **Go** button. In the navigation window, a specific oligonucleotide for display can be entered, or a button labeled **Most Stable** can be clicked to display the oligonucleotide with the tightest binding.
- 13 For oligonucleotides with self structure, the self structure can be drawn on the screen by double-clicking the oligonucleotide sequence. For oligonucleotides with both bimolecular and unimolecular structure, a window opens to allow user selection of the structure type to display.

BASIC PROTOCOL 3: PREDICTING A SECONDARY STRUCTURE WITH THE RNAstructure WEBSERVER

The RNAstructure Webserver takes advantage of the internet to make RNA secondary structure prediction available to a large audience. It requires no special setup and is more user-friendly than command-line alternatives. This protocol outlines the steps involved in predicting a secondary structure on the RNAstructure Webserver using a standard web browser. This protocol assumes basic familiarity with the internet and web browsing.

Necessary Resources

Software: A Web browser is required for accessing the RNAstructure Web servers

Access the server

- 1 Open the URL <http://rna.urmc.rochester.edu/RNAstructureWeb/> using a web browser. There are multiple structure prediction tools listed on that page. Click the link, **Predict a Secondary Structure**, at the top of the list of tools. This tool takes a sequence, either RNA or DNA, and predicts secondary structures using three methods. This protocol describes the information that must be entered to predict a structure, starting at the top of the page and working down. Be sure to read the information at the top of the page that describes the tool and how to retrieve results.

*The RNAstructure Web servers are organized into two groups: problem-based and program-based. This protocol only discusses use of the **Predict a Secondary Structure** problem-based tool. If a specific program is needed, the program-based tools, can be used instead*

Enter the sequence

- 2 Sequences are either uploaded using a FASTA-formatted text file or by pasting the sequence into the browser. Figure 11.2.8 shows a screen shot of the entry

form. To upload a file, click the button under **Select Sequence File** (the title of this button is browser-dependent). For example, on Google Chrome the button is called “Choose File”, whereas on Microsoft Edge, it is called “Browse”. Then select the file to upload. To paste a sequence, enter a title for the sequence in the **Sequence Title** field and paste the sequence in the **Sequence** field. The Sequence field should contain only A, C, G, U, T, and X. T and U are equivalent. X is a nucleotide that cannot pair or stack. An example sequence can be displayed by clicking on the link directly above the text box labeled “click [here](#) to add example sequences to the box.” This pastes the tRNA sequence RA7680 (Sprinzl and Vassilenko, 2005) into the form.

For the Web server, there is a limit to the length of sequences. The limit is detailed at <http://rna.urmc.rochester.edu/RNAstructureWeb/Information/Limitations.html>. As of this writing, sequences must be 2,500 nucleotides or shorter. The limit is designed to provide users with a reasonable rate of throughput, and it might change as dictated by server demand or as hardware is upgraded. Note that lowercase nucleotides are not allowed to pair in structure prediction. It is therefore important that most nucleotides be uppercase

- 3 Select the nucleic acid backbone. RNA is the default, and this will treat any T in the sequence as uracil. DNA can also be chosen, and this will treat any U in the sequence as thymine.

Select parameters

- 4 Options and parameters that affect the prediction calculation can be set in the **Default Data** section. For most calculations, these parameters can be kept at their default settings.

***Temperature:** The nearest-neighbor parameters are most accurate at 310.15 K. For RNA, they are known to be accurate between 293 and 333 K (Lu et al., 2006). For most calculations, it is reasonable to choose the default temperature.*

***Maximum Loop Size:** To reduce the calculation time, the maximum size of internal and bulge loops can be limited. Traditionally, the limit has been set at 30 unpaired nucleotides. It is unlikely that structure prediction for a biologically relevant sequence would require a change to this parameter.*

***Maximum % Energy Difference, Maximum Number of Structures, and Window Size:** For free energy minimization and maximum expected accuracy (MEA) structure prediction, these control the number and diversity of suboptimal structures, which serve as alternative hypotheses for the predicted structure. The default parameters were chosen to provide a small set of diverse structures*

that can be manually reviewed. Maximum % Energy Difference and Maximum Number of Structures place limits on the total number of suboptimal structures—more structures can be generated by setting these parameters to larger values. Window Size ensures that structures are substantially different. To generate more structures that are more similar, a smaller integer (as low as 0) can be used. Setting Window to a larger integer will generate fewer, more dissimilar structures.

Gamma: *This parameter adjusts the weight on pairing and unpairing in MEA structure prediction. A higher gamma will result in more pairs and a lower gamma will result in fewer pairs in the predicted structure. The default, 1.0, was found in benchmarks to provide a good balance.*

Iterations and Maximum Helix Length: For the ProbKnot program, which is capable of predicting pseudoknots, a single structure is predicted and these parameters control the number of predicted pairs. Increasing the number of iterations (from the default of 1) will predict more base pairs, but usually at the expense of accuracy. The minimum helix length can be reduced from the default of 3 (down to 1) to allow shorter helices or increased to require longer helices. Empirically, 3 was found to provide the best results for a wide range of sequences with known structure

Optional Constraints

- 5 Constraints on the folding can be uploaded as plain text files. To upload a constraints file, click Browse (or Choose File, etc.) at **Select Folding Constraints File**.

Constraints might include information obtained from enzymatic probing of structure (see UNIT 6.1), including nucleotides that cannot pair and nucleotides that must pair (Knapp, 1989). Constraints can also be determined by chemical modification probing of structure, which determines nucleotides that cannot be buried in a helix (Ehresmann et al., 1987). These forms of constraints have both been shown to improve the accuracy of structure prediction (Mathews et al., 1999b, 2004). Other constraints include specifying uridines in GU base pairs, pairs that must occur, and specific pairs that cannot occur. Additionally, SHAPE reactivity data can be provided and these data are found to improve the accuracy of structure prediction (Deigan et al, 2009)

Start the Calculation and Retrieve the Results

- 6 Once all desired parameters and constraints have been added, enter an email address at the bottom of the form and click **Submit Query**.

Entering an email address is optional, but highly recommended to ensure delivery of the calculation results. If an email address is provided, then an email is sent to it when the calculation is complete. The email contains a link to the calculation results. If an e-mail address is not provided, then it is imperative that the browser window remain open to the web page after the calculation is started in order to obtain the results
- 7 After clicking Submit Query, the web page redirects the user to the results page. This page refreshes automatically until the calculation is complete, and the results can be displayed. If the user has closed the webpage during calculation, they can wait for the results email to arrive, and then click the link in that email to view the results page. A screen shot of the output page is shown in Figure 11.2.9. The time to complete a calculation is a function of both the time to complete the calculation and also the number of simultaneous users.
- 8 The results page provides three types of structure predictions. From the top, they are Fold, the predicted lowest free energy structures, MaxExpect, structures composed of highly probable base pairs, and ProbKnot, a method that can predict pseudoknotted structures. All three methods are used because they can provide complementary information. For some sequences, all three programs predict the same structure. In general, on average, there is a trend towards MaxExpect predicting slightly fewer false positives than Fold. Also, ProbKnot is capable of predicting pseudoknots, but this generally comes at the cost of accuracy because it is prone to predicting false pseudoknots (Bellaousov and Mathews, 2010). The prediction by Fold is important because it is the predicted lowest free energy structure, i.e. the most probable structure at equilibrium.
- 9 For each of the three structure predictions, a separate drawing panel is shown. Structures in each of the three drawing panels are color annotated by base pairing probability. Highly probable base pairs (those close to a probability of 1) are more likely to be correctly predicted than low probability pairs (those with probability of less than 0.5). For Fold and MaxExpect, a set of suboptimal structures can also be present. These are structures that do not score as highly as the best, but score well and are therefore alternative hypotheses for the structure. **Previous** and **Next** can be used to cycle through the suboptimal structures, if present. Links at the bottom of the panels can be clicked to download all of the predicted structures (in PDF, Postscript, or CT file formats)

or the currently displayed individual structure (in PDF, Scalar Vector Graphics, Postscript, JPEG, or CT-file formats).

- 10 Below the structure drawing panels is a panel that shows the base pairing probabilities as a dot plot. This is labeled as the result of Partition, the partition function program in RNAstructure. By default, this drawing shows pairs of 0.01 or higher probability, but this threshold can be changed on the input page. The x and y axes are both nucleotide number, and a dot is a specific base pair between two nucleotides. This plot is important because it can reveal locations of pseudoknots, which is the basis of ProbKnot. Also, pairs composed of nucleotides with few competing alternative base pairings tend to have higher pairing probability, and more highly probable pairs are more likely to be correctly predicted. The plot can be downloaded as an image in Scalar Vector Graphic, JPEG, PDF, or Postscript format.

COMMENTARY

Background Information

The RNAstructure Fold program predicts RNA secondary structures on the basis of free energy minimization. The lowest free energy structure is the structure that is predicted most likely to occur at equilibrium. The secondary structure formation free energy change is predicted using a set of empirical nearest-neighbor parameters, determined from optical melting experiments on model systems (Xia et al., 1998; Mathews et al., 1999b, 2004). The partition function in RNAstructure (program Partition) is likewise built from free energy changes for structure formation and implicitly considers all possible secondary structures when calculating base-pair probabilities (Mathews, 2004). Pair probabilities can be used to assemble maximum expected accuracy structures (MaxExpect program in RNAstructure; Lu et al., 2009) or structures composed of mutually maximal pairing probability partners (ProbKnot program in RNAstructure; Bellaousov & Mathews, 2010).

Fold uses a dynamic programming algorithm that guarantees that the predicted lowest free energy structure will be found. Essentially, the structure prediction problem is divided into smaller problems, and recursion builds the complete secondary structure. Two reviews are available that explain dynamic programming in detail (Eddy, 2004; Mathews and Zuker, 2004). The partition function is also calculated with a dynamic programming algorithm. The algorithms used, however, cannot predict pseudoknotted (non-nested) base pairs. Other dynamic programming algorithms exist that can predict pseudoknotted pairs (Rivas and Eddy, 1999; Dirks and Pierce, 2003), but these calculations take considerably longer. ProbKnot is faster, and has similar accuracy to other programs. On average, only 1.4% of base pairs are pseudoknotted in a database of diverse RNA structures (Mathews et al., 1999b), but this percentage can be much higher for some classes of RNA structures, such as RNase P (Brown, 1999) and tmRNA (Williams and Bartel, 1996).

Other computer programs, for example the Vienna RNA package (Hofacker, 2003) and mfold (Zuker, 1989), also predict lowest free energy secondary structures by dynamic programming. RNAstructure, mfold, and the Vienna package differ slightly in the

implementation of the nearest-neighbor parameters for multibranch loops and exterior loops (loops that contain the ends of the sequence). RNAstructure explicitly considers both coaxial stacking of adjacent helices and helices separated by a single mismatch. The Vienna package considers coaxial stacking in free energy minimization, but does not include coaxial stacking in the partition function prediction of base-pair probabilities. In contrast, mfold does not consider coaxial stacking in the dynamic programming algorithm, but a second step, efn2, recalculates the free energy change of folding for each structure, including coaxial stacking of adjacent helices and helices separated by a single mismatch. Because of these differences in the energy models, the programs are not guaranteed to predict the same lowest free energy structure. A benchmark, however, showed that the programs have similar average accuracy (Dowell and Eddy, 2004).

Anticipated Results

Free energy minimization, on average, predicts 73% of known base pairs in the lowest free energy structure for a diverse set of sequences <700 nt and with known secondary structure (Mathews et al., 1999b, 2004). However, using a different set of sequences with known structures, including longer sequences, RNAstructure only predicted 56% of known base pairs (Dowell and Eddy, 2004). Therefore, secondary structure prediction should be viewed as a method for developing structure hypotheses. Suboptimal structures are thus alternative hypotheses for the secondary structure.

Constraints on the possible structures can be specified. RNAstructure can utilize constraints based on enzymatic cleavage (revealing paired or unpaired nucleotides; Knapp, 1989), FMN cleavage (revealing U's in GU pairs; Burgstaller and Famulok, 1997), chemical modification (revealing nucleotides that are unpaired, at the ends of helices, or in or adjacent to GU pairs; Ehresmann et al., 1987), or SHAPE (revealing nucleotides in flexible regions of the RNA; Deigan et al., 2009). *UNIT 6.1* discusses the use of enzymes and chemical reagents to probe RNA structures. For enzymatic cleavage and traditional, gel-read chemical mapping data, it has been shown that the use of constraints based on experimental data improves the accuracy of secondary structure prediction for sequences that would have poorly predicted structures without constraints (Mathews et al., 1999b, 2004). For SHAPE data, generally high structure prediction accuracy can be expected (Deigan et al., 2009).

The base-pair probabilities can be used to determine confidence in a predicted base pair (Mathews, 2004). On average, 66% of predicted base pairs in the lowest free energy structure are in the known structure for a diverse set of sequences. However, when only base pairs with predicted pairing probability ≥ 0.90 are considered, 83% of predicted pairs are in the known structure. For a probability threshold of 0.99, this accuracy increases to 91%. On average, nearly one quarter of predicted base pairs in the lowest free energy structure have pairing probability of at least 0.99. MaxExpect takes advantage of this by constructing structures composed of highly probable pairs, and MaxExpect tends to predict fewer false positives than free energy minimization (Lu et al., 2009).

OligoWalk provides an estimate of binding affinity of oligonucleotides to a structured RNA target (Mathews et al., 1999a). For an oligonucleotide to bind tightly, not only should the duplex free energy be low (more negative), but the magnitude of the cost of opening the

target structure should also be minimized. It has been shown that the duplex formation free energy and oligonucleotide self-structure terms correlate with antisense oligonucleotide efficacy (Matveeva et al., 2003). It is also well accepted that self-structure of RNA targets is an important siRNA design criterion (Bohula et al., 2003; Far and Sczakiel, 2003; Petch et al., 2003; Heale et al., 2005; Lu & Mathews, 2007).

Acknowledgments

The creation of this protocol was supported by grant R01 GM076485 from the National Institutes of Health.

Literature Cited

- Bellaousov S, Mathews DH. ProbKnot: Fast prediction of RNA secondary structure including pseudoknots. *RNA*. 2010; 16:1870–1880. [PubMed: 20699301]
- Bohula EA, Salisbury AJ, Sohail M, Playford MP, Riedemann J, Southern EM, Macaulay VM. The efficacy of small interfering RNAs targeted to the type 1 insulin-like growth factor receptor (IGF1R) is influenced by secondary structure in the IGF1R transcript. *J Biol Chem*. 2003; 278:15991–15997. [PubMed: 12604614]
- Brown JW. The ribonuclease P database. *Nucleic Acids Res*. 1999; 27:314. [PubMed: 9847214]
- Burgstaller P, Famulok M. Flavin-dependent photocleavage of RNA at G.U base pairs. *J Am Chem Soc*. 1997; 119:1137–1138.
- Deigan KE, Li TW, Mathews DH, Weeks KM. Accurate SHAPE-directed RNA structure determination. *Proc Natl Acad Sci USA*. 2009; 106:97–102. [PubMed: 19109441]
- Dirks R, Pierce N. A partition function algorithm for nucleic acid secondary structure including pseudoknots. *J Comput Chem*. 2003; 24:1664–1677. [PubMed: 12926009]
- Dowell RD, Eddy SR. Evaluation of several lightweight stochastic context-free grammars for RNA secondary structure prediction. *BMC Bioinformatics*. 2004; 5:71. [PubMed: 15180907]
- Eddy SR. How do RNA folding algorithms work? *Nat Biotechnol*. 2004; 22:1457–1458. [PubMed: 15529172]
- Ehresmann C, Baudin F, Mougél M, Romby P, Ebel J, Ehresmann B. Probing the structure of RNAs in solution. *Nucleic Acids Res*. 1987; 15:9109–9128. [PubMed: 2446263]
- Far RK, Sczakiel G. The activity of siRNA in mammalian cells is related to structural target accessibility: A comparison with antisense oligonucleotides. *Nucleic Acids Res*. 2003; 31:4417–4424. [PubMed: 12888501]
- Heale BS, Soifer HS, Bowers C, Rossi JJ. siRNA target site secondary structure predictions using local stable substructures. *Nucleic Acids Res*. 2005; 33:e30. [PubMed: 15722476]
- Hofacker IL. Vienna RNA secondary structure server. *Nucleic Acids Res*. 2003; 31:3429–3431. [PubMed: 12824340]
- Knapp G. Enzymatic approaches to probing RNA secondary and tertiary structure. *Methods Enzymol*. 1989; 180:192–212. [PubMed: 2482414]
- Lu ZJ, Gloor JW, Mathews DH. Improved RNA secondary structure prediction by maximizing expected pair accuracy. *RNA*. 2009; 15:1805–1813. [PubMed: 19703939]
- Lu ZJ, Mathews DH. Efficient siRNA selection using hybridization thermodynamics. *Nucleic Acids Res*. 2007; 36:640–647. [PubMed: 18073195]
- 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. [PubMed: 16982646]
- 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. [PubMed: 15272118]
- Mathews, DH.; Zuker, M. Predictive methods using RNA sequences. In: Baxeveinis, A.; Oullette, F., editors. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. 3. John Wiley & Sons; Hoboken, N.J: 2004. p. 143-170.

- Mathews DH, Burkard ME, Freier SM, Wyatt JR, Turner DH. Predicting oligonucleotide affinity to nucleic acid targets. *RNA*. 1999a; 5:1458–1469. [PubMed: 10580474]
- Mathews DH, Sabina J, Zuker M, Turner DH. Expanded sequence dependence of thermodynamic parameters provides improved prediction of RNA secondary structure. *J Mol Biol*. 1999b; 288:911–940. [PubMed: 10329189]
- 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. [PubMed: 15123812]
- 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; 31:4989–4994. [PubMed: 12930948]
- Petch AK, Sohail M, Hughes MD, Benter I, Darling J, Southern EM, Akhtar S. Messenger RNA expression profiling of genes involved in epidermal growth factor receptor signalling in human cancer cells treated with scanning array-designed antisense oligonucleotides. *Biochem Pharmacol*. 2003; 66:819–830. [PubMed: 12948863]
- Rivas E, Eddy SR. A dynamic programming algorithm for RNA structure prediction including pseudoknots. *J Mol Biol*. 1999; 285:2053–2068. [PubMed: 9925784]
- Sprinzl M, Vassilenko KS. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res*. 2005; 33:D139–140. [PubMed: 15608164]
- Walter AE, Turner DH, Kim J, Lyttle MH, Müller P, Mathews DH, Zuker M. Coaxial stacking of helices enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc Natl Acad Sci USA*. 1994; 91:9218–9222. [PubMed: 7524072]
- Williams KP, Bartel DP. Phylogenetic analysis of tmRNA secondary structure. *RNA*. 1996; 2:1306–1310. [PubMed: 8972778]
- 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. [PubMed: 9778347]
- Zuker M. On finding all suboptimal foldings of an RNA molecule. *Science*. 1989; 244:48–52. [PubMed: 2468181]
- Zuker M, Jacobson AB. “Well-determined” regions in RNA secondary structure predictions. Applications to small and large subunit rRNA. *Nucleic Acids Res*. 1995; 23:2791–2798. [PubMed: 7544463]
- Zuker M, Jacobson AB. Using reliability information to annotate RNA secondary structures. *RNA*. 1998; 4:669–679. [PubMed: 9622126]
- Zuker, M.; Mathews, DH.; Turner, DH. Algorithms and thermodynamics for RNA secondary structure prediction: A practical guide. In: Barciszewski, J.; Clark, BFC., editors. *RNA Biochemistry and Biotechnology*. Kluwer Academic Publishers; Boston: 1999. p. 11-43.

Figure 11.2.1.

Screen shot of the RNAstructure sequence editor. The *D. melanogaster* 5S rRNA sequence is entered and formatted using the Format Sequence button. A second sequence editor is shown (to right) with modifications to the sequence to demonstrate the colorization of characters: lowercase (which will be forced to be unpaired; blue), X (a nucleotide that will be unpaired and unstacked; green), and invalid characters (red). The natural sequence (left) was used for subsequent sequence prediction.

Figure 11.2.2.
Screen shot of the RNA secondary structure prediction window.

Figure 11.2.3.

The lowest free energy secondary structure predicted for the *D. melanogaster* 5S rRNA sequence as drawn by RNAstructure.

Figure 11.2.4.

The probability dot plot for base pairs in the *D. melanogaster* 5S rRNA sequence. Each dot in this plot is a single base pair and is associated with a base-pairing probability. For pairs between nucleotides i and j , with $i < j$, i is down the right-hand side of the plot and j is across the top. In RNAstructure, dots are colored to indicate a probability interval, with the most probable pairs in red and the least probable in blue. The current base pairs displayed have pairing probability of 1%. A color key is drawn at the bottom of the window (not shown in this view). After clicking on a dot, the pair and $-\log_{10}(\text{base-pair probability})$ are indicated at the bottom of the RNAstructure window. The current display shows that the base pair between G75 and C102 was clicked and has a $-\log_{10}(\text{base-pair probability})$ of 5.33232×10^{-2} (88.4%). For a color version of this figure, see <http://www.currentprotocols.com>.

Figure 11.2.5.

The lowest free energy secondary structure predicted for the *D. melanogaster* 5S rRNA sequence with color annotation. The color annotation key is shown in the lower right-hand corner. The most probable base pairs are red and least probable are purple. The predicted pairs with the highest base pair probabilities are more likely to be correctly predicted pairs (Mathews, 2004). For a color version of this figure, see <http://www.currentprotocols.com>.

Figure 11.2.6.
The OligoWalk input window.

Figure 11.2.7.

The OligoWalk output display. For a color version of this figure, see <http://www.currentprotocols.com>.

The screenshot displays the RNAstructure webserver interface. At the top, there is a blue header with the 'RNAstructure' logo and the tagline 'Web Servers for RNA Secondary Structure Prediction'. Below the header is a green navigation bar with links: 'Servers Home', 'Contributors', 'Mathews Lab Home', 'RNAstructure Downloads', and 'Help / Contact Us'. The main content area has a green sub-header: 'Welcome to the Predict a Secondary Structure Web Server'. A paragraph explains the server's capabilities: 'The Predict a Secondary Structure server combines four separate prediction and analysis algorithms: calculating a partition function, predicting a maximum free energy (MFE) structure, finding structures with maximum expected accuracy, and pseudoknot prediction. This server takes a sequence, either RNA or DNA, and creates a highly probable, probability annotated group of secondary structures, starting with the lowest free energy structure and including others with varied probabilities of correctness. Other structures are included because the minimum free energy structure may not be the correct one. Note that if SHAPE constraints are specified, the SHAPE constraints are applied to the probability annotated structures. In addition, a second group of SHAPE constrained, SHAPE annotated structures will be generated. This SHAPE structure group is distinct from the probability annotated structure group, and is not probability annotated itself.' Below this, a note states: 'If you plan to leave this page while waiting for calculations to complete, please make sure you have entered a valid email address below. Leaving this page without entering a valid email address will render results inaccessible.' Another note says: 'If you need specific help using the Predict a Secondary Structure server, please click here.' The form itself is titled 'Select Sequence Input:' and has two main sections. The first section, 'Select Sequence File:', contains a 'Choose File' button and the text 'No file chosen'. The second section, 'OR -', instructs the user to 'Enter your sequence title and content below (replaces upload if present). Please enter nucleotides only, no headers or comments in content. Valid nucleotides are A, C, G, T, U, and X (unknown nucleotide). Lower case nucleotides are forced single stranded in calculation.' This section includes a 'Sequence Title:' field with the text 'D. melanogaster 5S rRNA' and a 'Sequence: (Click here to add an example sequence to the box.)' text area containing the sequence: 'GCCAAGACC AUACCAAGCU GAUACAUUG GUUCUGUCC GAUACCGAA AUUAAGCAGC GUUUGGCGCG GUUAGUACUU AGAUUGGGGA CCGUUGGGA ACACCGGUG UUGUUGGCCU'. Below the sequence input is a 'Select Nucleic Acid Type:' section with radio buttons for 'DNA' and 'RNA', where 'RNA' is selected. At the bottom, there is a 'Select Predict a Secondary Structure Options:' section with a 'Default Data' label and a note: 'If a default value is left blank, the value is treated as if it was not changed at all.' This section includes a 'Temperature (K):' field with the value '310.15'.

Figure 11.2.8.

The Fold entry form on the RNAstructure Webserver. For a color version of this figure, see <http://www.currentprotocols.com>.

Figure 11.2.9.

The secondary structure of *D. melanogaster* 5S rRNA as predicted by the RNAstructure Webserver. For a color version of this figure, see <http://www.currentprotocols.com>.

Table 11.2.1

Time Requirements for Secondary Structure Prediction and Base-Pair Probability Prediction Using RNAstructure^a

Sequence	Length (Nucleotides)	Time for Structure Prediction	Time For Base-Pair Probability Calculation
<i>Tetrahymena thermophila</i> group I intron	433	3 sec	2 sec
<i>E. coli</i> 16S rRNA	1542	1 min: 49 sec	1 min: 35 sec
<i>E. coli</i> 23S rRNA	2904	10 min: 35 sec	11 min : 02 sec

^aCalculations were performed on a computer with a 3.4 GHz Intel I7-2600K, 4-core processor, and 8 GB of memory, running Microsoft Windows 7. Calculation times are less with a faster processor or with more memory and slower with a slower processor. The calculation time scales according to $O(N^3)$, where N is the length of the sequence. Therefore, doubling the length of the sequence requires roughly eight times the calculation time.

Table 11.2.2Time Requirements for OligoWalk Calculations^a

TARGET	TARGET LENGTH	OLIGONUCLEOTIDE LENGTH	MODE	TIME
Saccharomyces cerevisiae A5 group II intron	631	10	Break Local Structure; Do Not Include Suboptimal Structures	1 sec
<i>Saccharomyces cerevisiae</i> A5 group II intron	631	20	Break Local Structure; Do Not Include Suboptimal Structures	2 sec
<i>Saccharomyces cerevisiae</i> A5 group II intron	631	20	Break Local Structure; Include Suboptimal Structures	2 sec
<i>Saccharomyces cerevisiae</i> A5 group II intron	631	20	Do Not Consider Target Structure	1 sec
<i>Saccharomyces cerevisiae</i> A5 group II intron	631	20	Refold Whole RNA; Include Suboptimal Structures	2 hr, 6 min
<i>E. coli</i> 16S rRNA	1542	20	Break Local Structure; Include Suboptimal Structures	8 sec

^aCalculations were performed on a computer with a 3.4 GHz Intel I7-2600K, 4-core processor, and 8 GB of memory, running Microsoft Windows 10. For calculations that break local structure or do not consider target structure, the calculation scales $O(NL^3)$ where N is the length of the target and L is the length of the oligonucleotide. Doubling the length of the target only doubles the length of the calculation. A doubling in the length of the oligonucleotide requires eight times as much calculation time. For calculations in which the whole target RNA is refolded for each oligonucleotide, the calculation scales roughly $O(N^4)$. A doubling in the length of the target requires 16 times as much computation time, but lengthening the oligonucleotide results in little change in the calculation time.