



# Visualization of Secondary Structures including Pseudoknots

Christian Heine    Gerek Scheuermann

Image and Signal Processing Group  
University of Leipzig

Bled 2007-02-21

Popular Visualization Methods

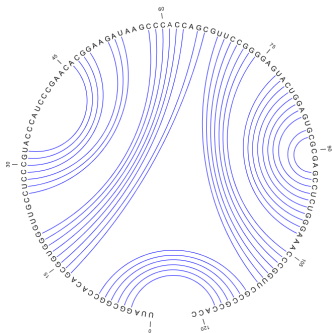
Survey of “Structural” Drawing Programs

Methods for Structural Representation

Improvements

Conclusions

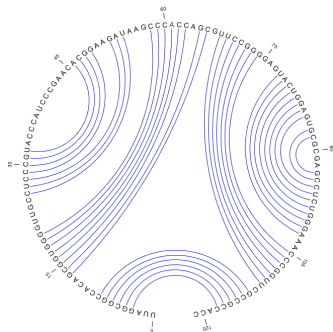
# Secondary Structure



- ▶ list of base pairs
- ▶ no base is part of more than one base pair
- ▶ base pairs do not cross
- ▶ pseudoknot: the list of base pairs can be split in two lists where each list describes a valid secondary structure

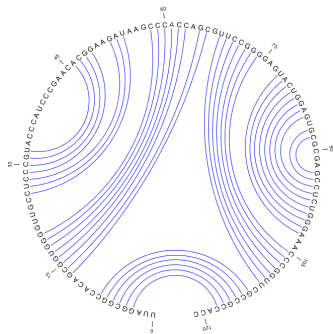


# Secondary Structure



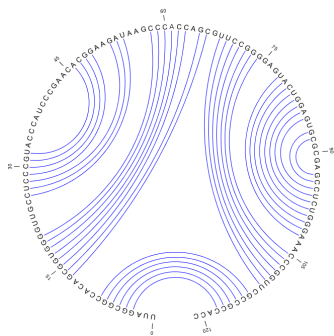
- ▶ list of base pairs
- ▶ no base is part of more than one base pair
- ▶ base pairs do not cross
- ▶ pseudoknot: the list of base pairs can be split in two lists where each list describes a valid secondary structure

## Secondary Structure



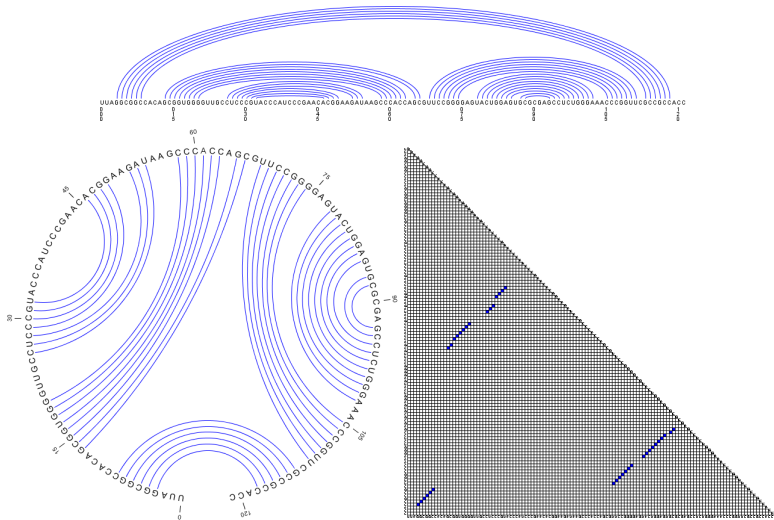
- ▶ list of base pairs
- ▶ no base is part of more than one base pair
- ▶ base pairs do not cross
- ▶ pseudoknot: the list of base pairs can be split in two lists where each list describes a valid secondary structure

# Secondary Structure



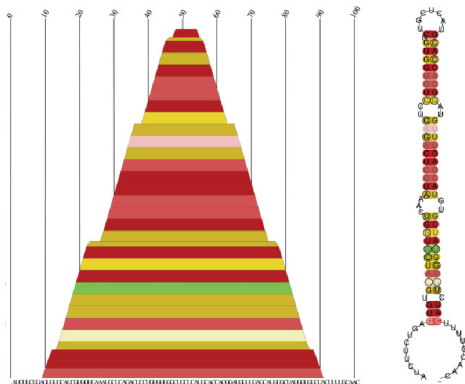
- ▶ list of base pairs
- ▶ no base is part of more than one base pair
- ▶ base pairs do not cross
- ▶ pseudoknot: the list of base pairs can be split in two lists where each list describes a valid secondary structure

# Methods



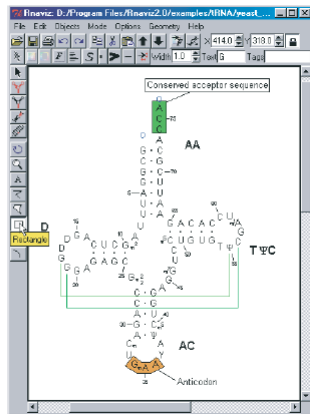
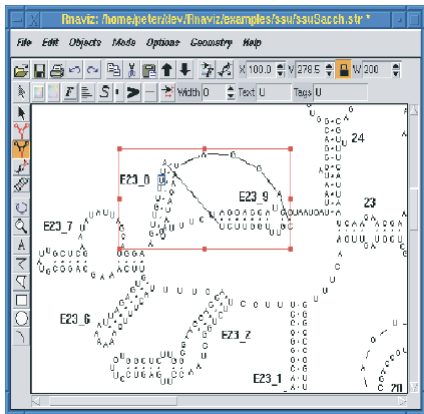


# Methods



- ▶  $A \in O(n^2)$  poor screen usage
- ▶ different drawings emphasize different aspects
- ▶ “structural” drawings emphasize higher level structural components, e.g. helices and loops

## RnaViz



P. de Rijk and R. de Wachter, "RnaViz, a program for the visualization of RNA secondary structure," *Nucleic Acids Research* 25(22), pp 4679-4684, 1997.

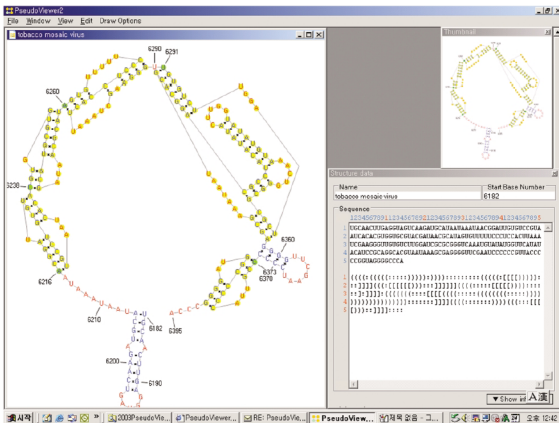
J. Wuyts, P. de Rijk, and R. de Wachter, "RnaViz2, an improved representation of RNA secondary structure," *Bioinformatics* 19, pp 299-300, 2003.

## xRna

The screenshot displays the xRna software interface. At the top, the title bar reads "Survey of 'Structural' Drawing Programs xRna". The main window has a blue border and contains the following elements:

- Main Controls:** A row of buttons including "Pick Strand", "rna single nuc", "Render", "Save", "Find", and "Stop XRNA".
- Import/Export:** A tabbed interface with "Import/Export", "Annotate", "Edit", and "Format" tabs. The "Import/Export" tab is active, showing the text: "Import/Export 'Tet.L.ref' In File Tet.L.ref.xrna In Directory C:\christian:". Below this, there are input fields for "Scale Figure:" (set to 1.0) and "postsript scale:" (set to 1.0).
- Left Panel:** A vertical menu with options: "Choose Input File", "Save as ...", "Write JPG Image", "landscape mode" (with a checkbox), "Print Postscript", and "cartesian coords" (with a checkbox). At the bottom of this panel is a small schematic diagram of the RNA structure.
- Central Display:** A large window showing the RNA secondary structure. The structure is a complex network of stems and loops, with various regions labeled P1 through P10. Nucleotide sequences are shown along the stems, and arrows indicate the 5' and 3' ends. A small schematic diagram is also visible in the bottom-left corner of the main display area.
- System Out:** A text area at the bottom of the window, currently empty.

# PseudoViewer



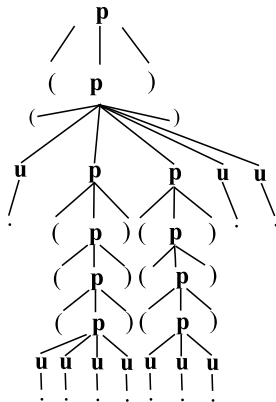
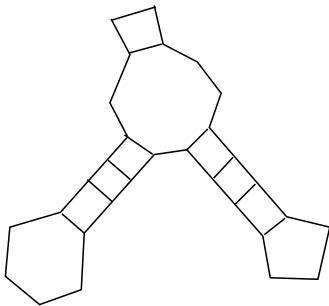
K. Han, Y. Lee, and W. Kim, "PseudoViewer: automatic visualization of RNA pseudoknots," *Bioinformatics* Vol. 18 Suppl. 1, pp S321-328, 2002

K. Han and Y. Byun, "PseudoViewer2: visualization of RNA pseudoknots of any type," *Nucleic Acids Research* 31(13), pp3432-3440, 2003

The screenshot displays the jViz.RNA application window. On the left is a control panel with several sliders: Font Size (0-20), Backbone Tension (2-17), Bond Tension (50-500), Node Transparency (0-225), and Line Thickness (0-15). Below these is a 'Help Contents' section with the heading 'Classical Structure' and instructions on how to pan, zoom, and adjust the drawing. The main window shows a 3D-like representation of an RNA secondary structure with nucleotides labeled with their bases (A, C, G, U) and connected by blue lines representing base pairs. The structure is complex, with multiple stems and loops.

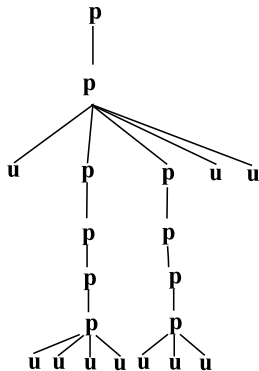
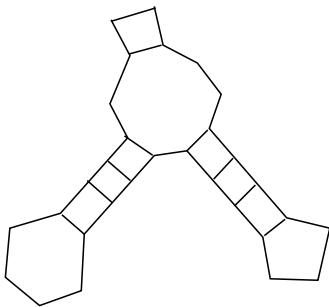
K.C. Wiese and E. Glen, “jViz.Rna – An interactive graphical tool for visualizing RNA Secondary Structure Including Pseudoknots,” Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems (CBMS’06), 2006.

# Structure Tree



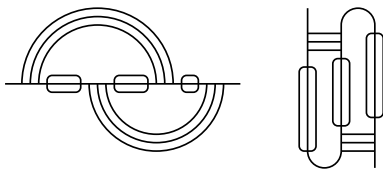
- ▶ Simple to compute the structural tree
- ▶ Simple to lay out this tree
- ▶ Problems: occlusions of subtrees and pseudoknots

# Structure Tree



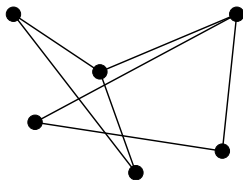
- ▶ Simple to compute the structural tree
- ▶ Simple to lay out this tree
- ▶ Problems: occlusions of subtrees and pseudoknots

# Pseudoknots



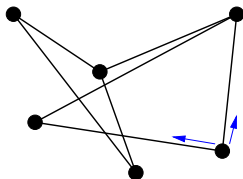


# Force-Directed Methods



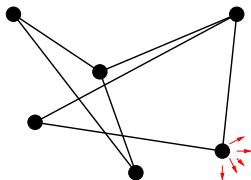
- ▶ general (undirected) graph layout
- ▶ computationally expensive ( $O(n^3)$ )
- ▶ non-deterministic
- ▶ low probability of crossings and occlusions

# Force-Directed Methods



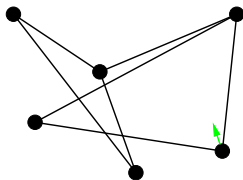
- ▶ general (undirected) graph layout
- ▶ computationally expensive ( $O(n^3)$ )
- ▶ non-deterministic
- ▶ low probability of crossings and occlusions

# Force-Directed Methods



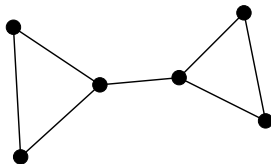
- ▶ general (undirected) graph layout
- ▶ computationally expensive ( $O(n^3)$ )
- ▶ non-deterministic
- ▶ low probability of crossings and occlusions

# Force-Directed Methods



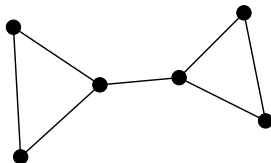
- ▶ general (undirected) graph layout
- ▶ computationally expensive ( $O(n^3)$ )
- ▶ non-deterministic
- ▶ low probability of crossings and occlusions

# Force-Directed Methods



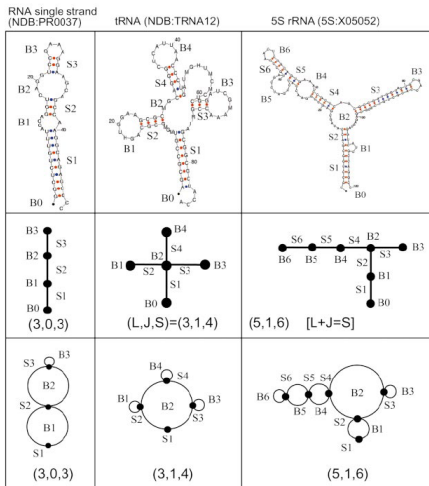
- ▶ general (undirected) graph layout
- ▶ computationally expensive ( $O(n^3)$ )
- ▶ non-deterministic
- ▶ low probability of crossings and occlusions

# Force-Directed Methods



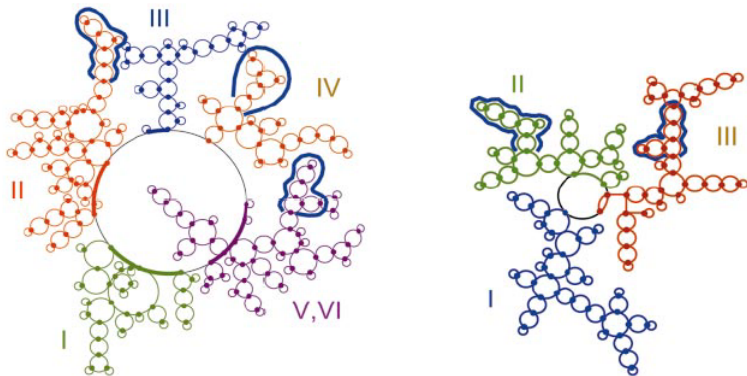
- ▶ general (undirected) graph layout
- ▶ computationally expensive ( $O(n^3)$ )
- ▶ non-deterministic
- ▶ low probability of crossings and occlusions

# Dual Graph



H.H. Gan, S. Pasquali, and T. Schlick, "Exploring the repertoire of RNA secondary motifs using graph theory; implications for RNA design," *Nucleic Acids Research* 31(11), pp 2926-2943, 2003.

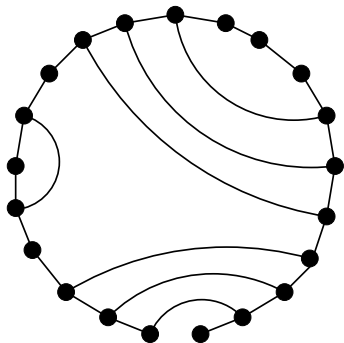
# Dual Graph



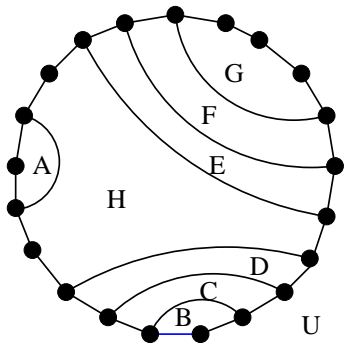
H.H. Gan, S. Pasquali, and T. Schlick, "Exploring the repertoire of RNA secondary motifs using graph theory; implications for RNA design," *Nucleic Acids Research* 31(11), pp 2926-2943, 2003.



# Dual Graph Minus U

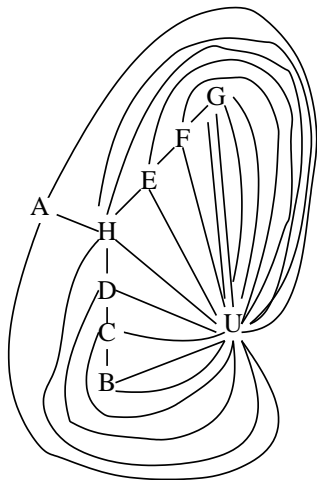
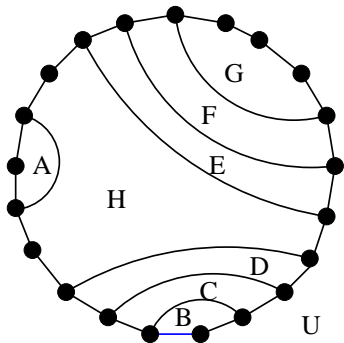


# Dual Graph Minus U

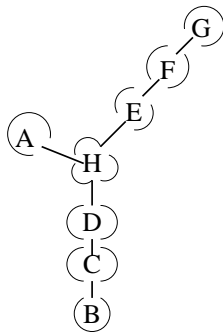
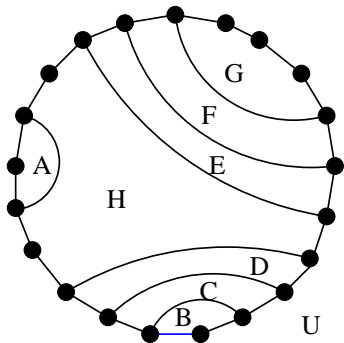




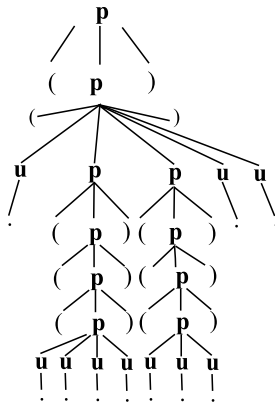
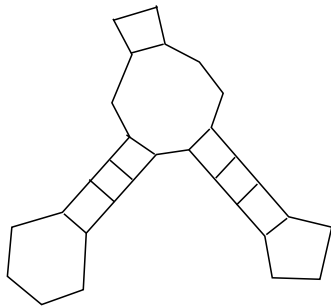
# Dual Graph Minus U



# Dual Graph Minus U

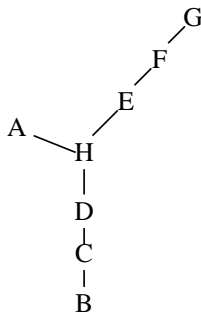
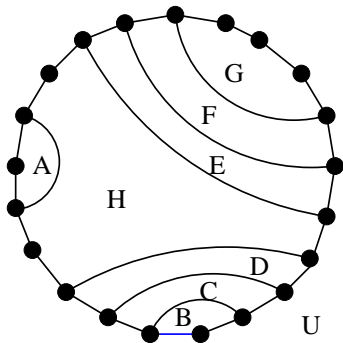


# Structure Tree

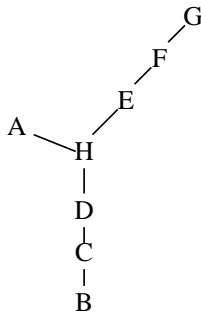
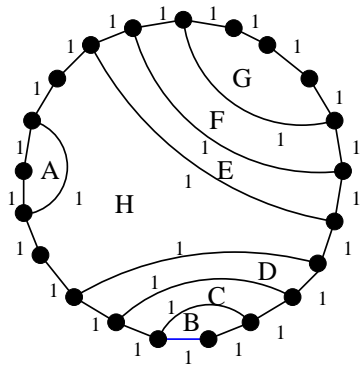


- ▶ Simple to compute the structural tree
- ▶ Simple to lay out this tree
- ▶ Problems: occlusions of subtrees and pseudoknots

# Reduced Dual Graph Minus U

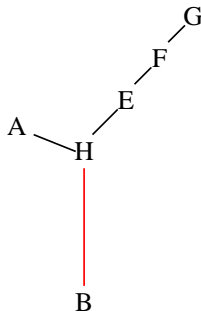
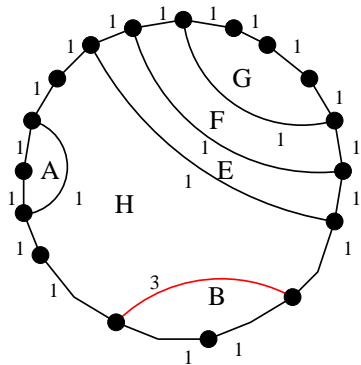


# Reduced Dual Graph Minus U

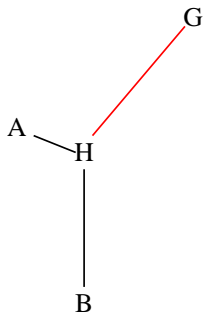
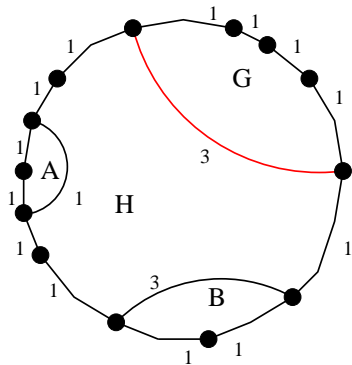




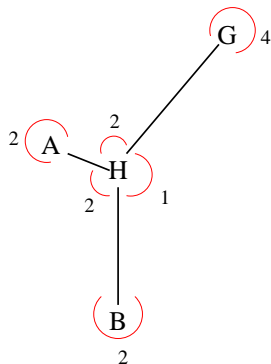
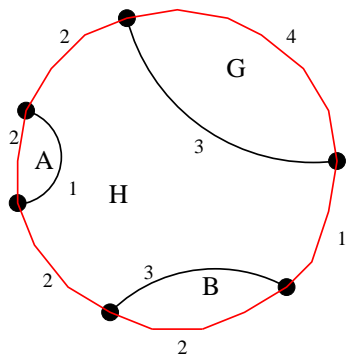
# Reduced Dual Graph Minus U



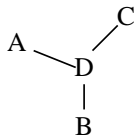
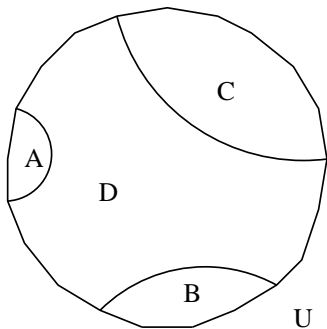
# Reduced Dual Graph Minus U



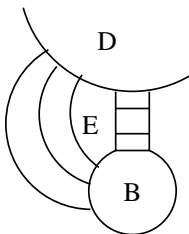
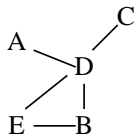
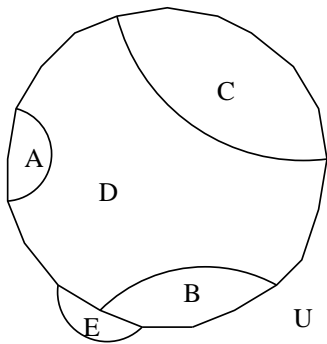
# Reduced Dual Graph Minus U



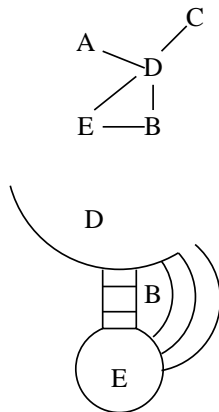
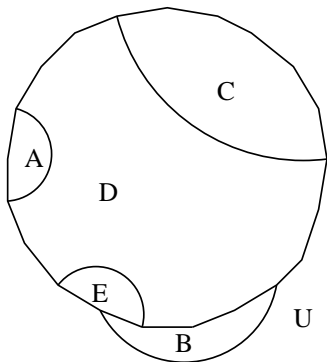
# Dual Graphs with Pseudoknots



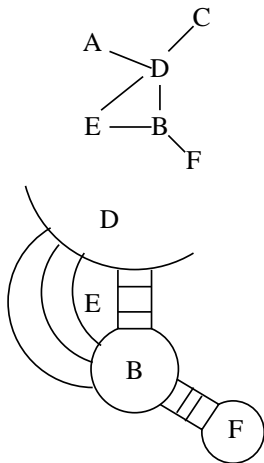
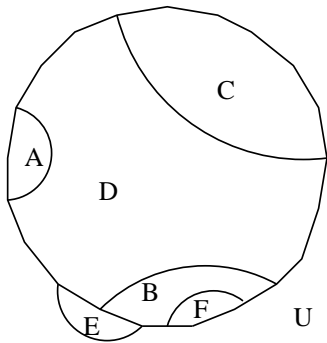
# Dual Graphs with Pseudoknots



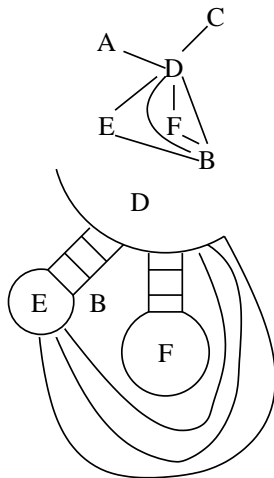
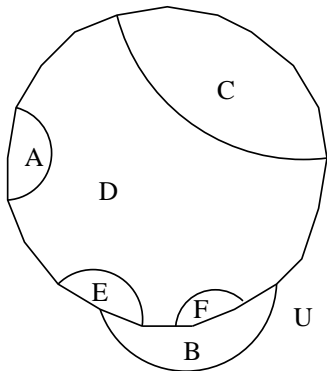
# Dual Graphs with Pseudoknots



# Dual Graphs with Pseudoknots

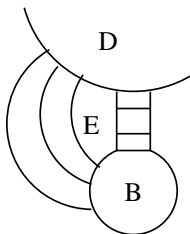
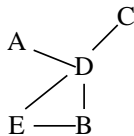
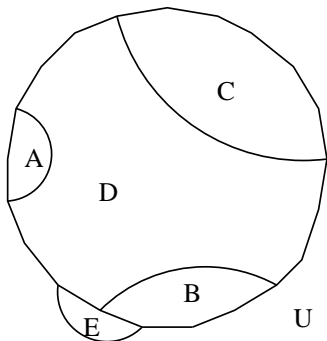


# Dual Graphs with Pseudoknots

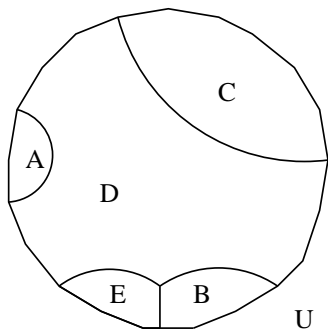




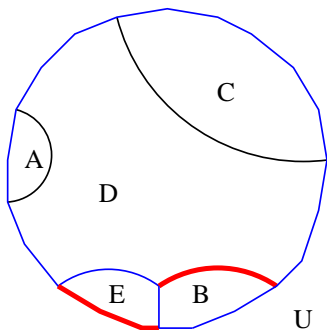
# Dual Graphs with Pseudoknots



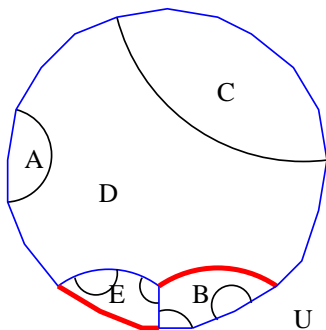
# Dual Graphs with Pseudoknots



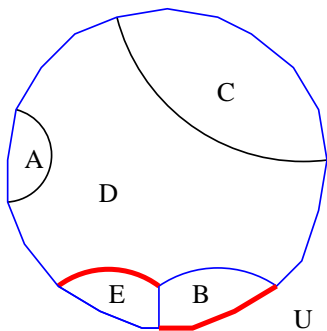
# Dual Graphs with Pseudoknots



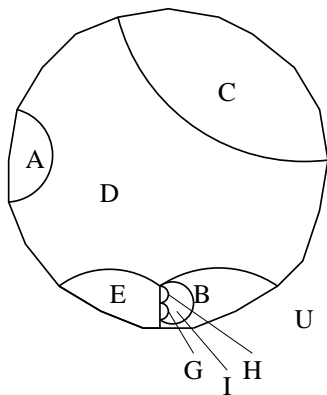
# Dual Graphs with Pseudoknots



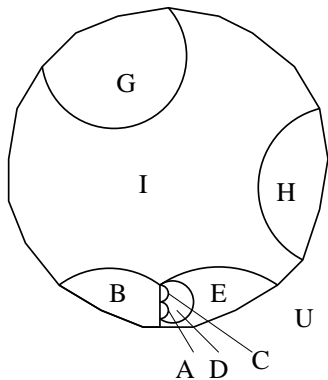
# Dual Graphs with Pseudoknots



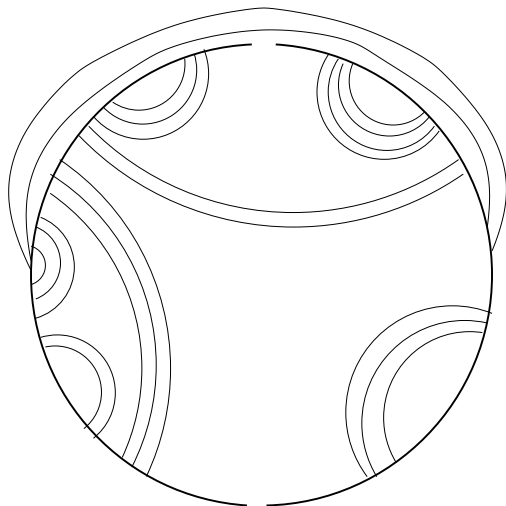
# Dual Graphs with Pseudoknots



# Dual Graphs with Pseudoknots

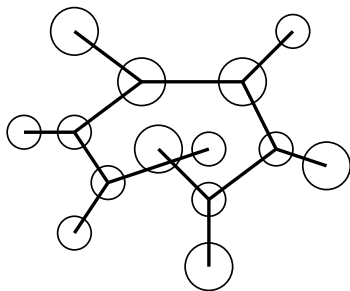


Is the visualization of pseudoknots a solution to visualization cotranscriptional folding?

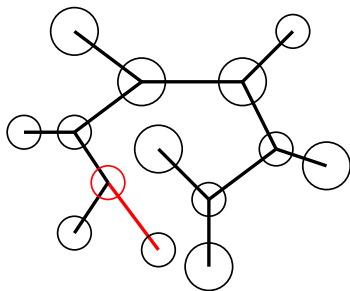




# The Occlusion Problem

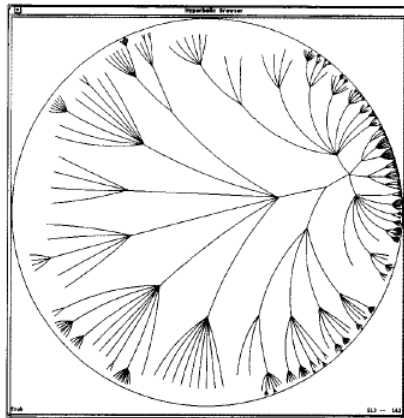
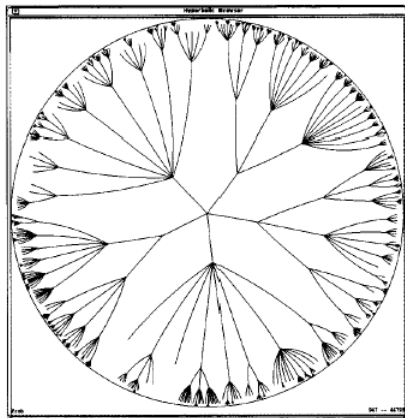


# The Occlusion Problem



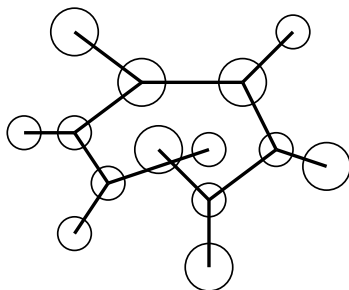
# The Occlusion Problem

- Layout in hyperbolic space



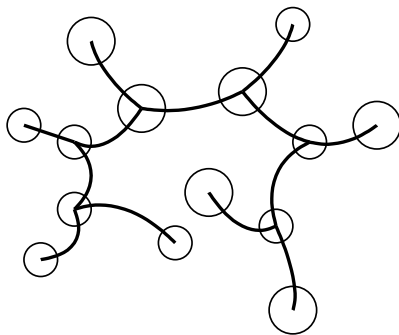
# The Occlusion Problem

- ▶ Force- Directed- “Flexible Stack” Layout.



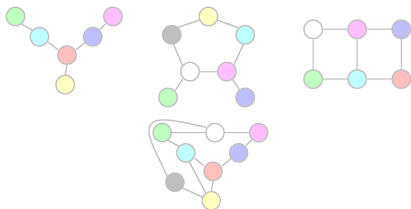
# The Occlusion Problem

- ▶ Force- Directed- “Flexible Stack” Layout.



# Visualization of Consensus Structures

$$G_i = (V_i, E_i)$$



$$G = (\cup V_i, \cup E_i)$$

no layout

---

with layout

$$L = L(G)$$

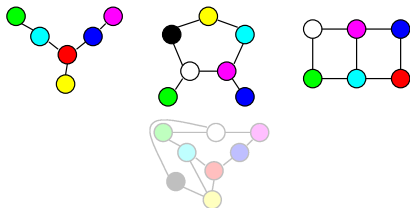


$$L(G_i) = L(G)|_{G_i}$$



# Visualization of Consensus Structures

$$G_i = (V_i, E_i)$$



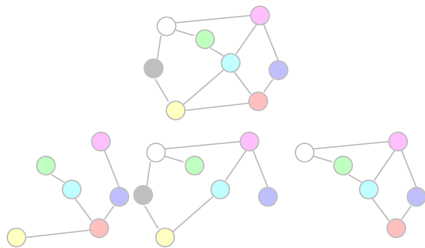
$$G = (\cup V_i, \cup E_i)$$

no layout

---

with layout

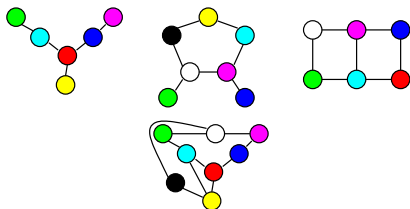
$$L = L(G)$$



$$L(G_i) = L(G)|_{G_i}$$

# Visualization of Consensus Structures

$$G_i = (V_i, E_i)$$



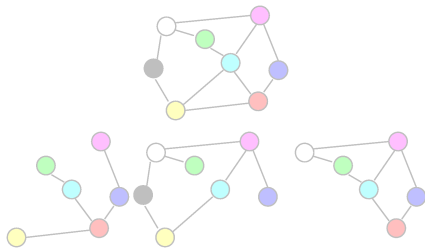
$$G = (\cup V_i, \cup E_i)$$

no layout

---

with layout

$$L = L(G)$$



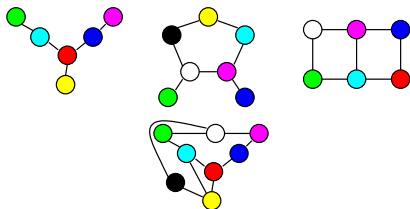
$$L(G_i) = L(G)|_{G_i}$$





# Visualization of Consensus Structures

$$G_i = (V_i, E_i)$$



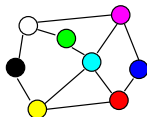
$$G = (\cup V_i, \cup E_i)$$

no layout

---

with layout

$$L = L(G)$$

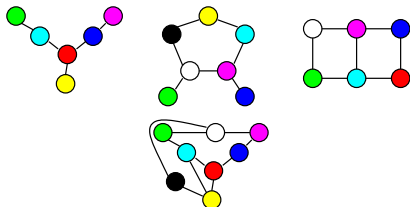


$$L(G_i) = L(G)|_{G_i}$$



# Visualization of Consensus Structures

$$G_i = (V_i, E_i)$$

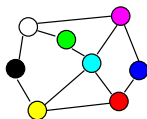


$$G = (\cup V_i, \cup E_i)$$

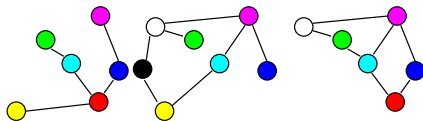
no layout

with layout

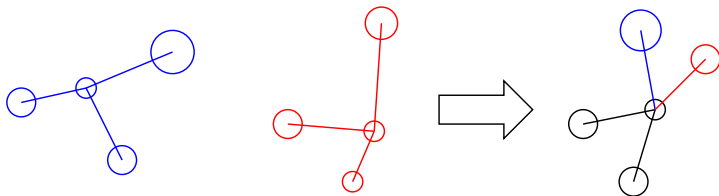
$$L = L(G)$$



$$L(G_i) = L(G)|_{G_i}$$



# Consensus Structure



- ▶ use tree alignment to determine supergraph
- ▶ use consensus dual graph for pseudoknots?

## Conclusions and Future Work

- ▶ High similarity between dual graph (minus U) and structure tree
- ▶ Use of dual graph allows us to describe structures with pseudoknots

### Discussion

- ▶ Is jViz.Rna already suitable for you? Shall we proceed to investigate pseudoknots?

### Future Work

- ▶ Formulate and solve the decision problem, where to place edges of pseudoknots (use dual graph)
- ▶ Force- Directed- “Flexible Stack” Layout
- ▶ Visualization of consensus structures (with pseudoknots)

# Thank You!