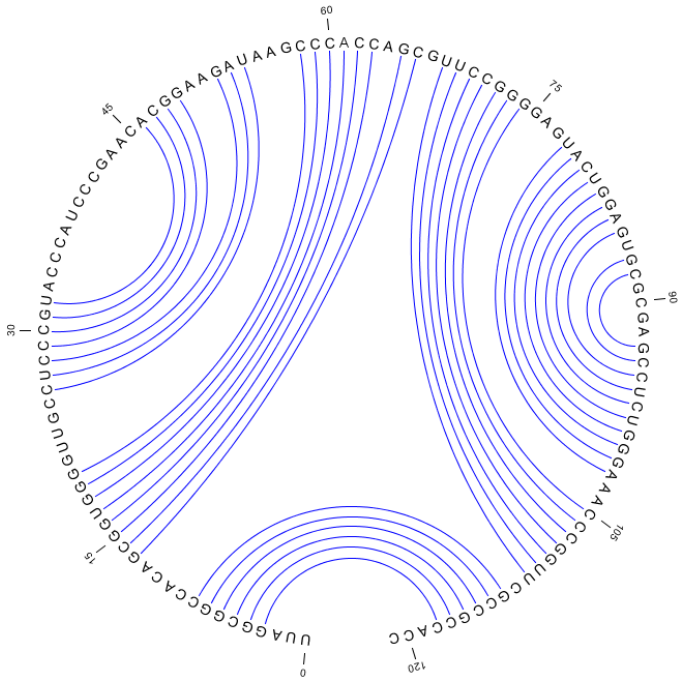
Visualization of Secondary Structures including Pseudoknots

Christian Heine Gerik Scheuermann

Image and Signal Processing Group
University of Leipzig

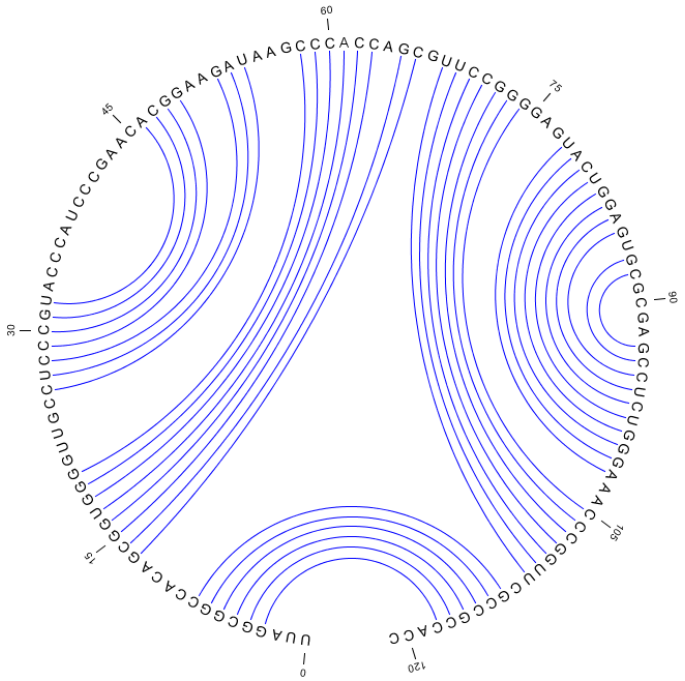
Bled 2007-02-21

Secondary Structure



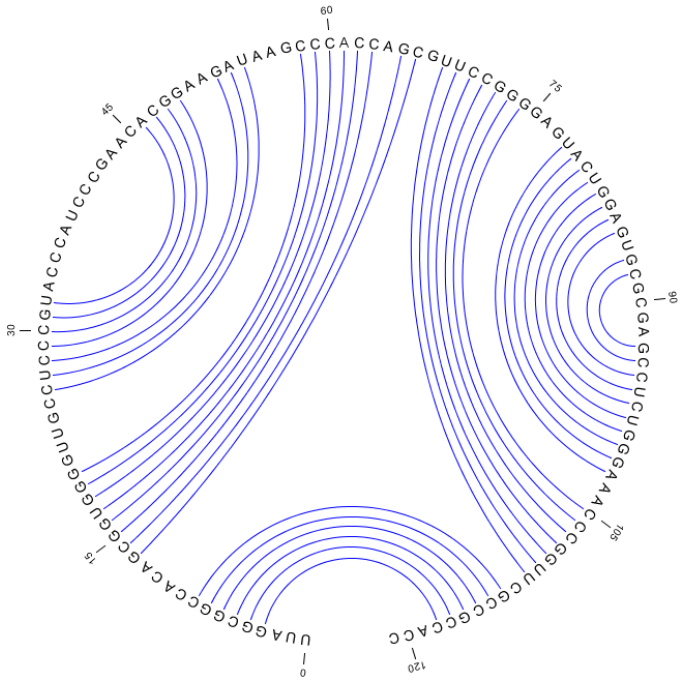
- list of base pairs
- no base is part of more than one base pair
- base pairs do not cross
- pseudoknot: the list of base pairs can be split in two lists where each list describes a valid secondary structure

Secondary Structure



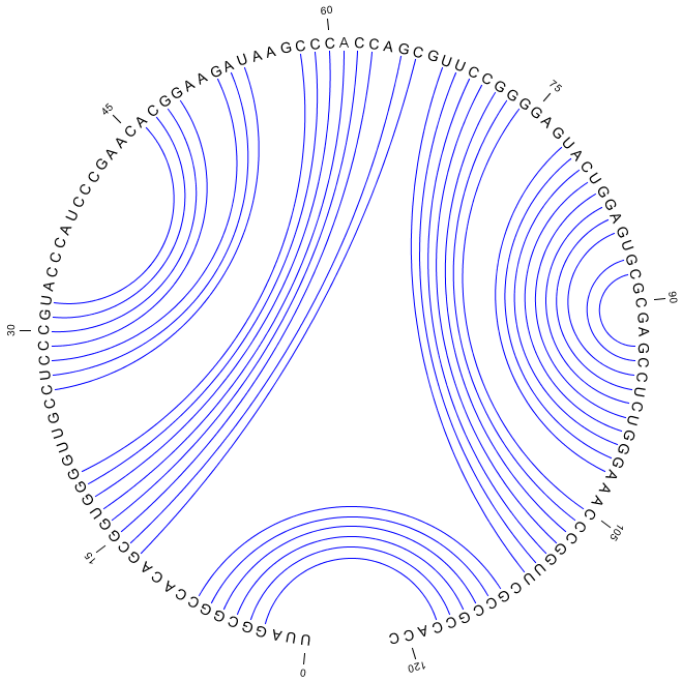
- list of base pairs
- no base is part of more than one base pair
- base pairs do not cross
- pseudoknot: the list of base pairs can be split in two lists where each list describes a valid secondary structure

Secondary Structure



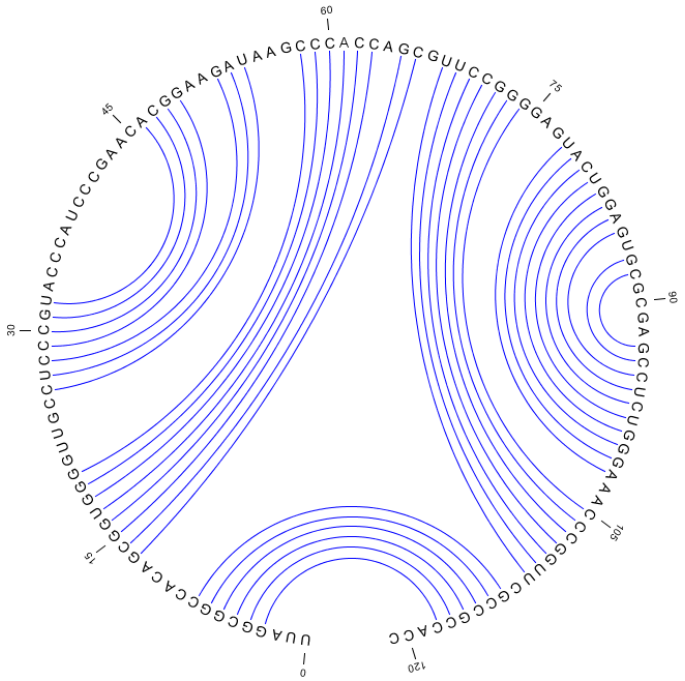
- list of base pairs
- no base is part of more than one base pair
- base pairs do not cross
- pseudoknot: the list of base pairs can be split in two lists where each list describes a valid secondary structure

Secondary Structure



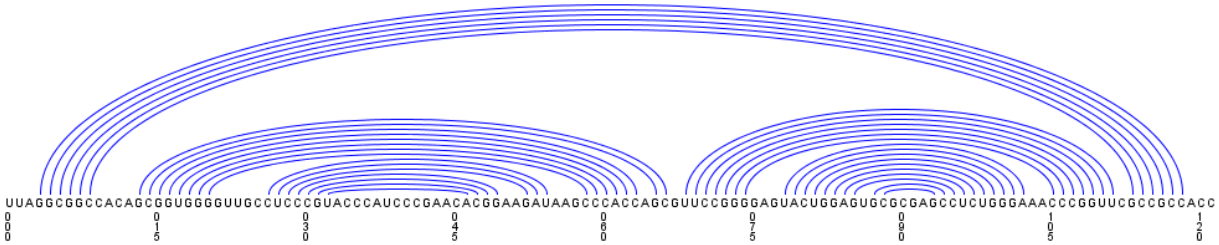
- list of base pairs
- no base is part of more than one base pair
- base pairs do not cross
- pseudoknot: the list of base pairs can be split in two lists where each list describes a valid secondary structure

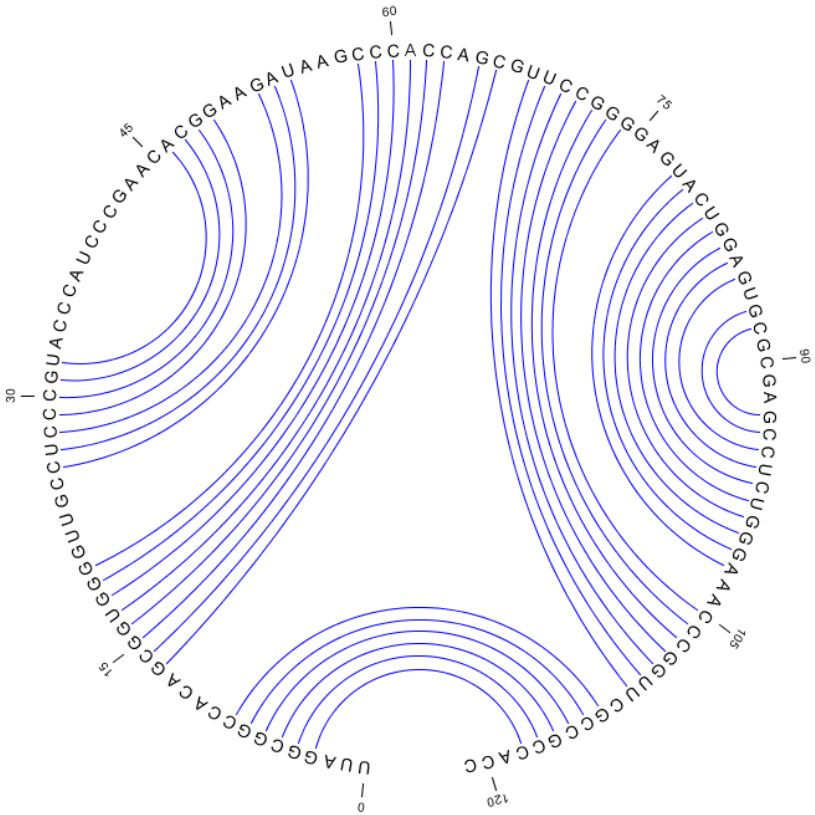
Secondary Structure

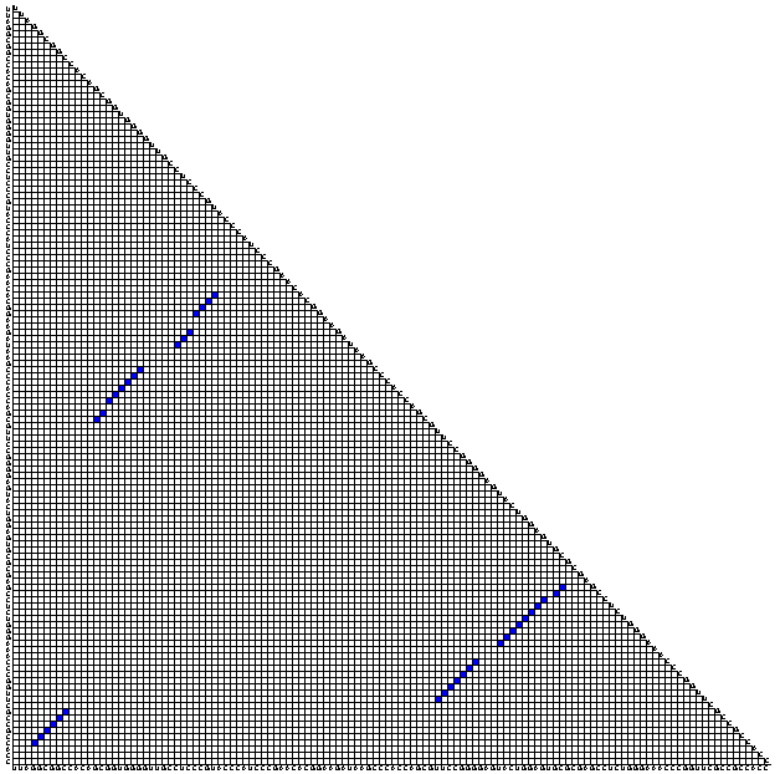


- list of base pairs
- no base is part of more than one base pair
- base pairs do not cross
- pseudoknot: the list of base pairs can be split in two lists where each list describes a valid secondary structure

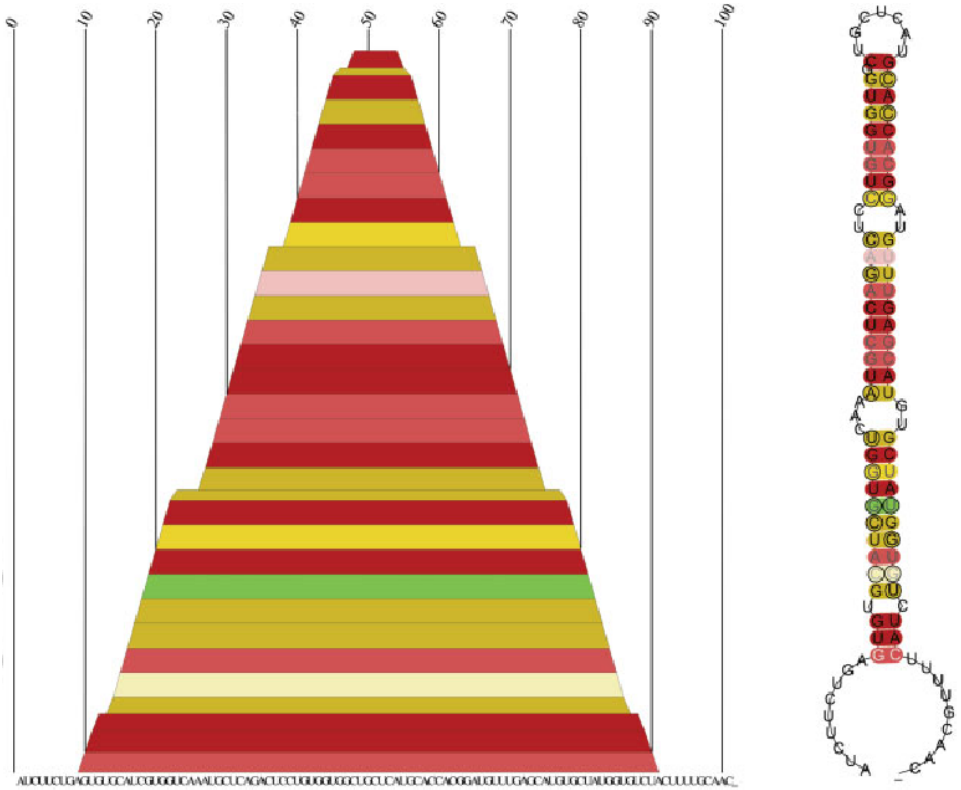
Methods


UUAGGCGGCCACAGCGGUGGGGUUGCCUCCCGUACCCAUCCCGAACACGGAAGAUAAGCCCACCAGCGUUCCGGGGAGUACUGGAGUGCGCGAGCCUCUGGGAAACCCGGUUCGCCGCCACC
0
15
30
45
60
75
90
105
120


UUAGGCGGCCACAGCGGUGGGGUUGCCUCCCGUACCCAUCCCGAACACGGAAGAUAAGCCCACCAGCGUUCCGGGGAGUACUGGAGUGCGCGAGCCUCUGGGAAACCCGGUUCGCCGCCACC
0
15
30
45
60
75
90
105
120

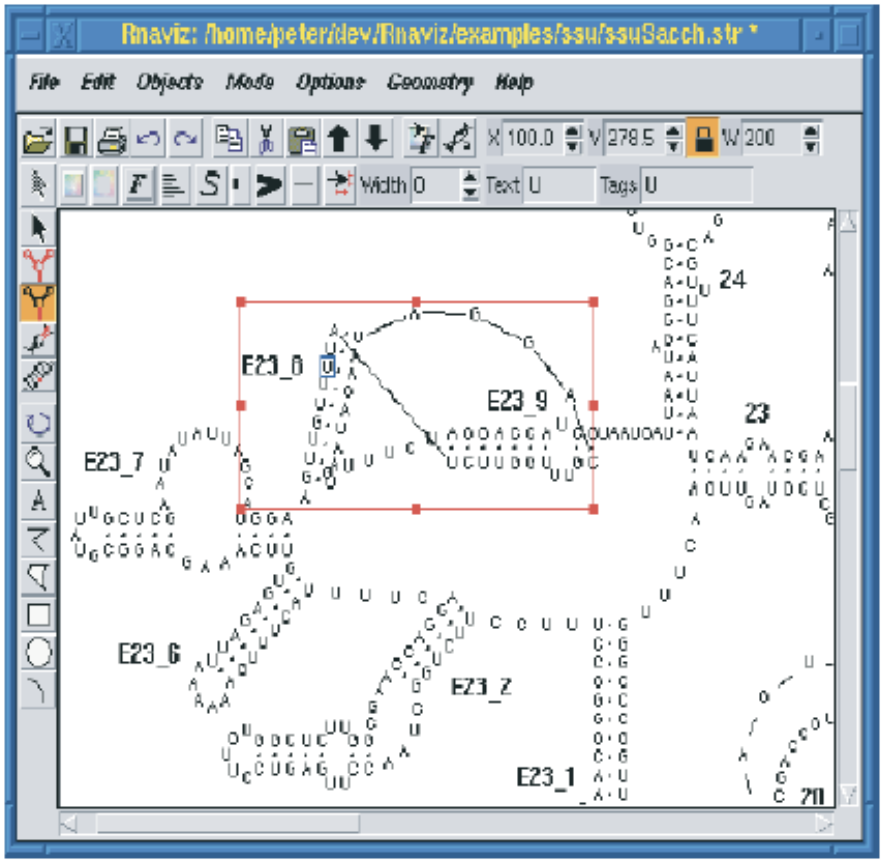


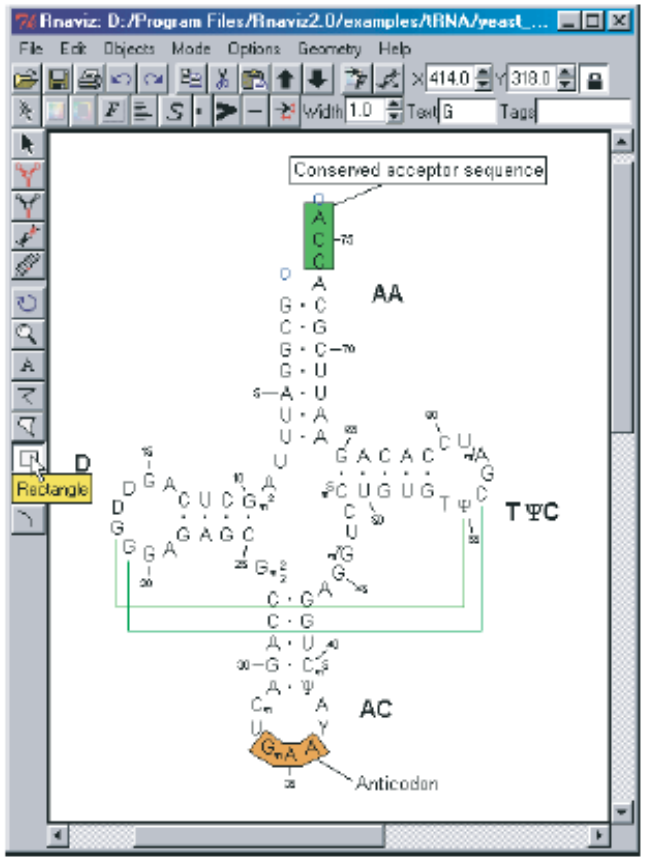
Methods



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

RnaViz

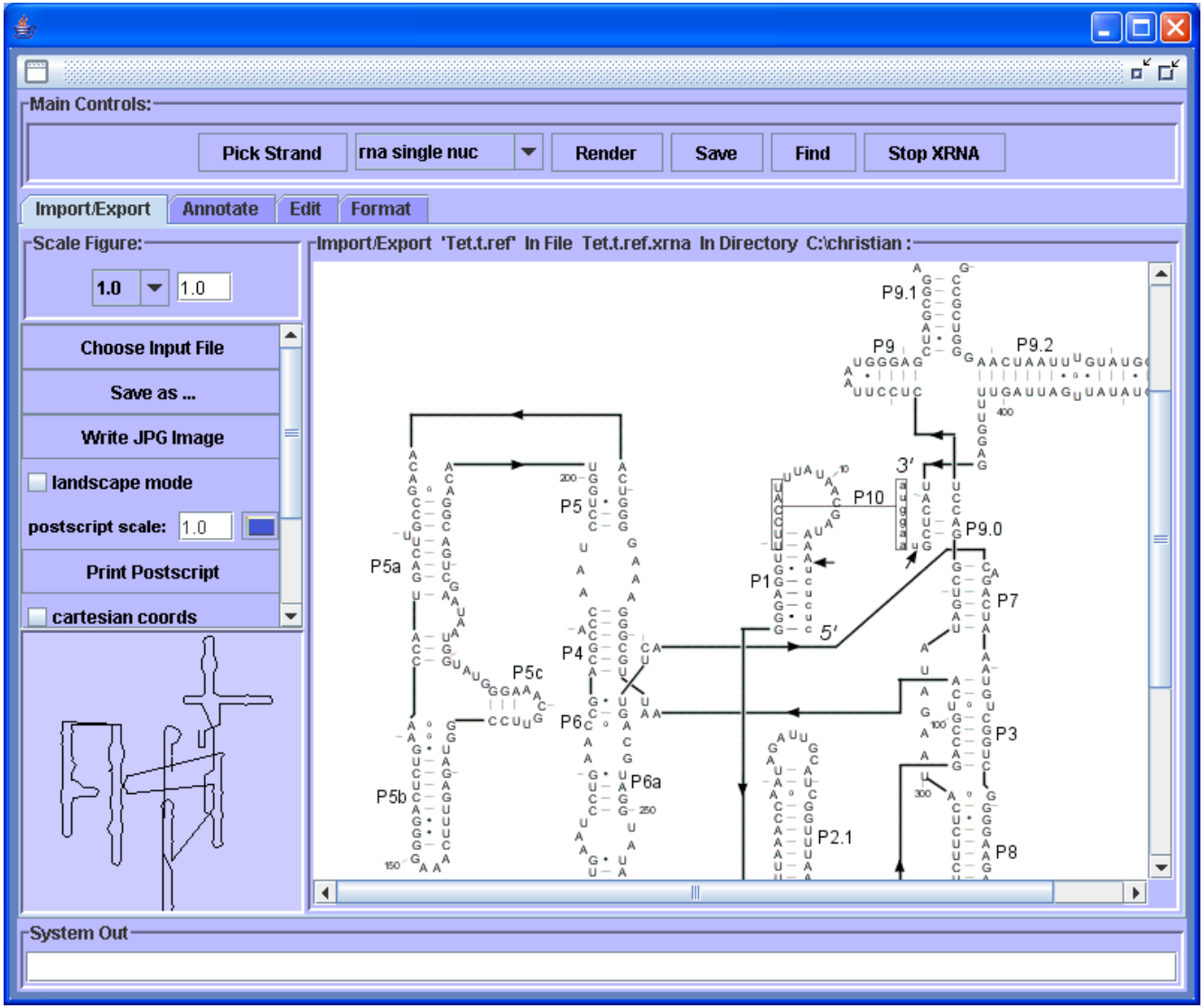




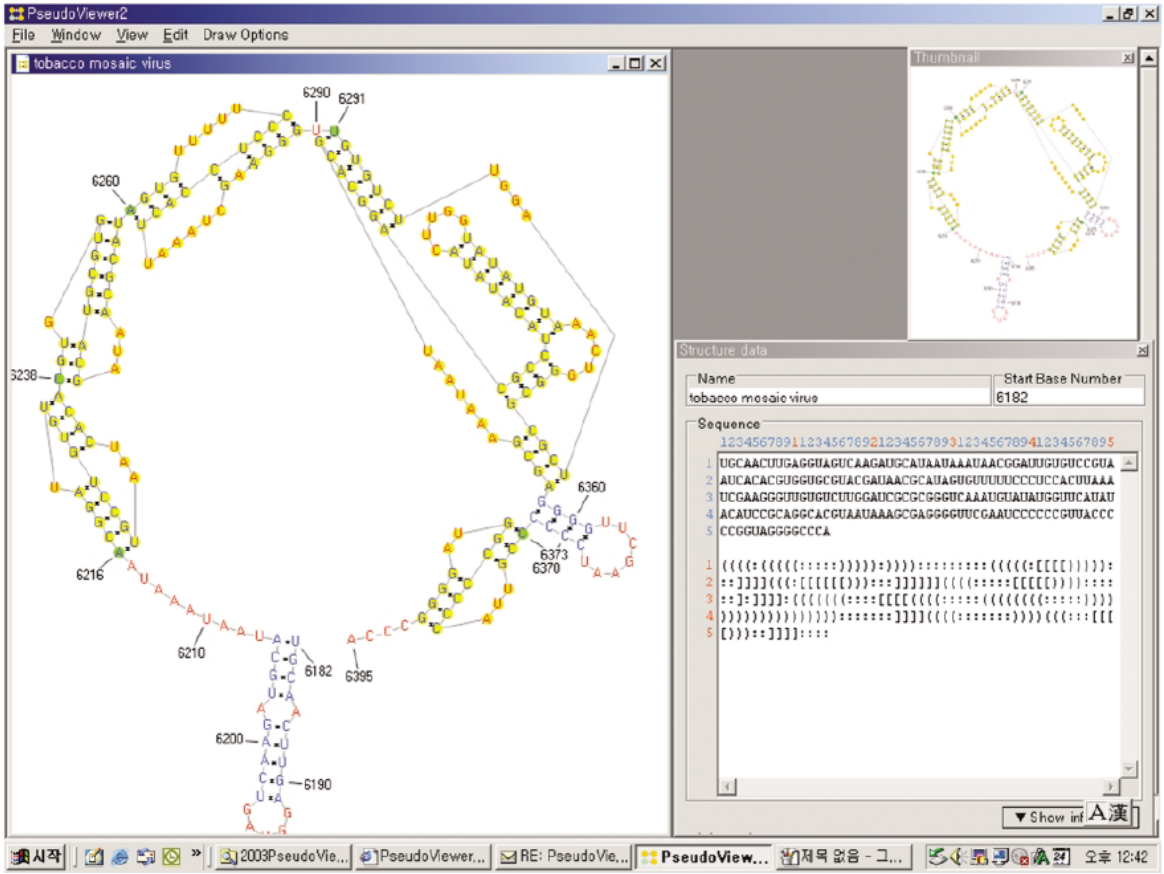
P. de Rijk and R. de Wachter, “RnaViz, a program for the visualization of RNA secondary structure,” *Nucleic Acids Research* 25(22), pp 4679-4684, 1997.

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

xRna

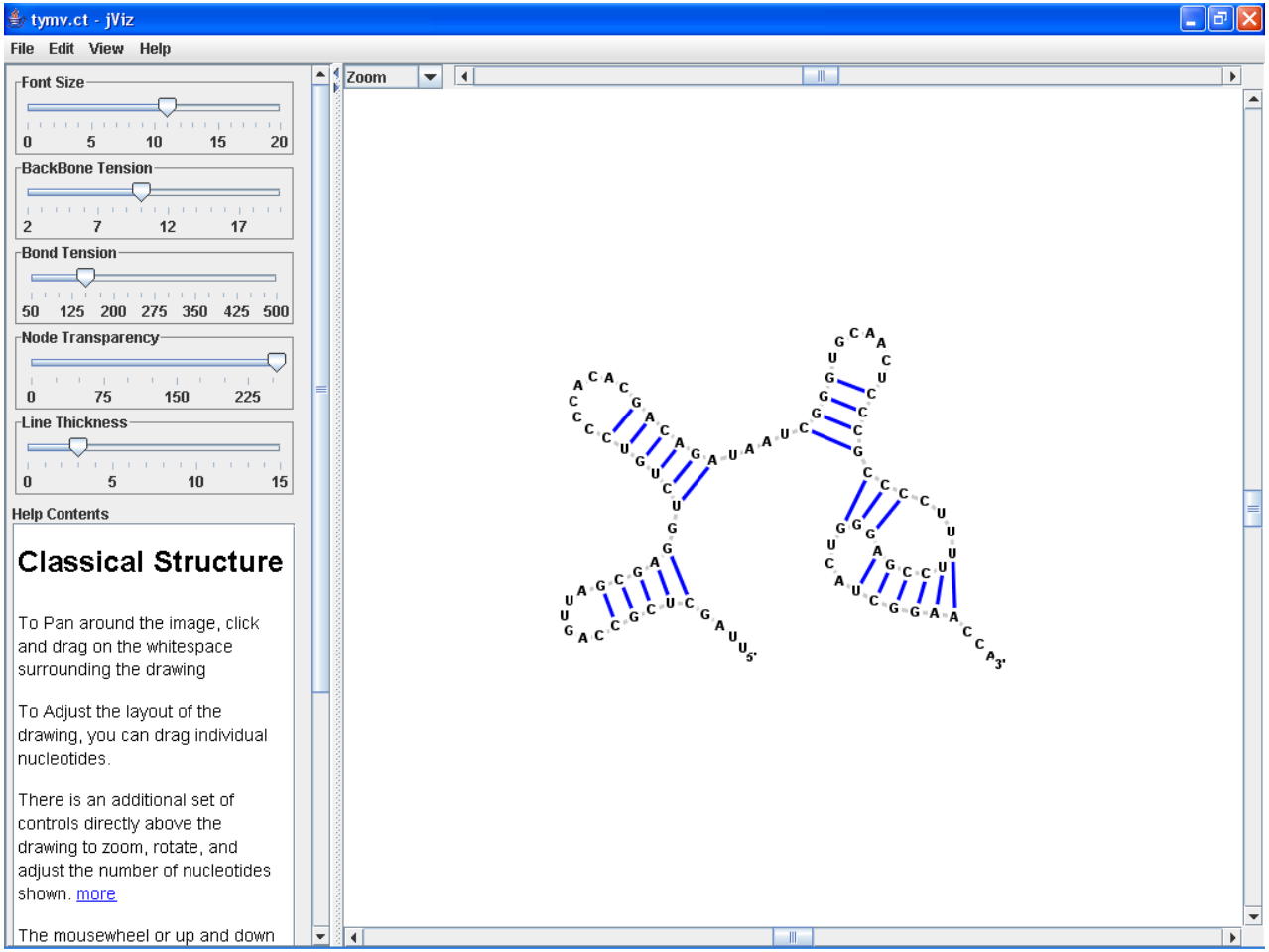

Main Controls:
Pick Strand
rna single nuc
Render
Save
Find
Stop XRNA
Import/Export
Annotate
Edit
Format
Scale Figure:
1.0
1.0
Choose Input File
Save as ...
Write JPG Image
landscape mode
postscript scale:
1.0
Print Postscript
cartesian coords
Import/Export 'Tet.t.ref' In File Tet.t.ref.xrna In Directory C:\christian :
P9.1
P9
P9.2
P10
P9.0
P7
P5
P5a
P5c
P4
P6
P6a
P5b
P1
P2.1
P3
P8
3′
5′
System Out

PseudoViewer



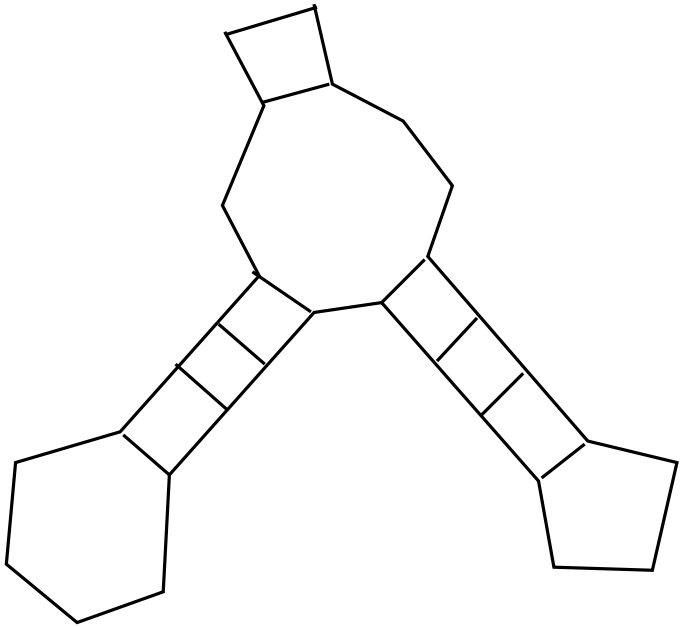
K. Han, Y. Lee, and W. Kim, "PseudoViewer: automatic visualization of RNA pseudoknots," *Bioinfomatics* Vol. 18 Suppl. 1, pp S321-328, 2002

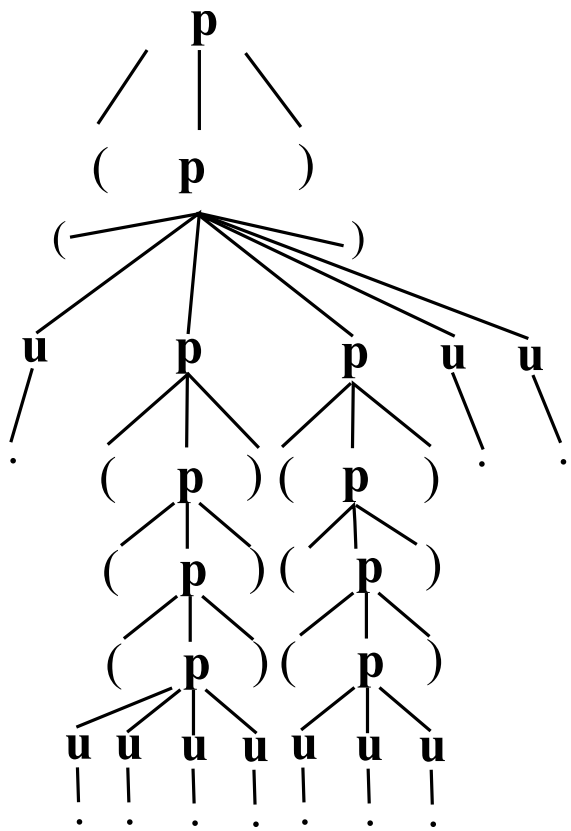
K. Han and Y. Byun, "PsuedoViewer2: visualization of RNA pseudoknots of any type," *Nucleic Acids Research* 31(13), pp3432-3440, 2003



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.

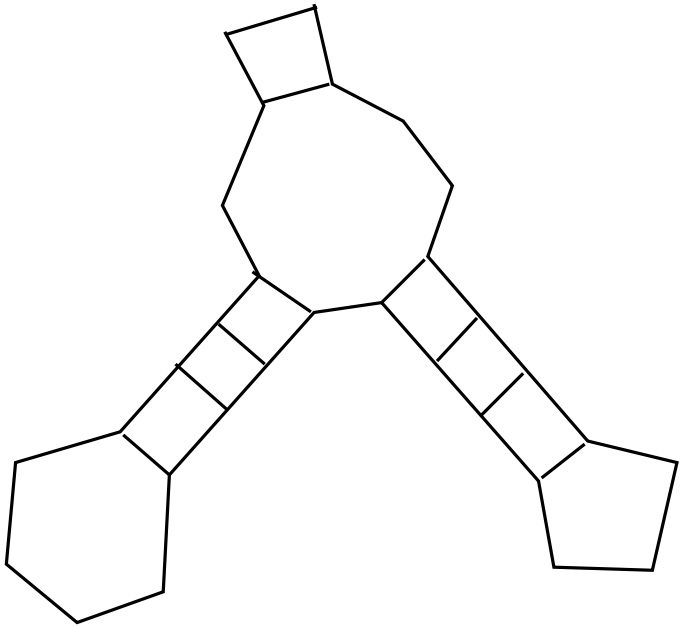
Structure Tree

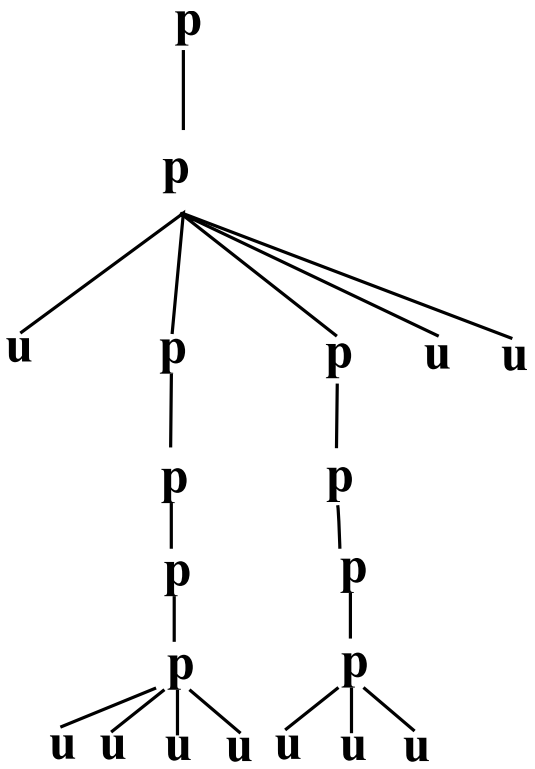




- Simple to compute the structural tree
- Simple to lay out this tree
- Problems: occlusions of subtrees and pseudoknots

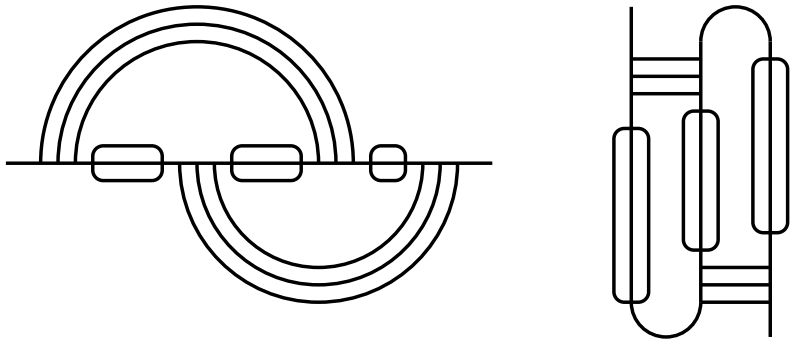
Structure Tree



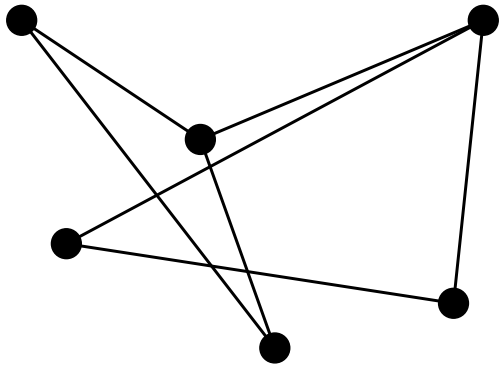


- Simple to compute the structural tree
- Simple to lay out this tree
- Problems: occlusions of subtrees and pseudoknots

Psudoknots

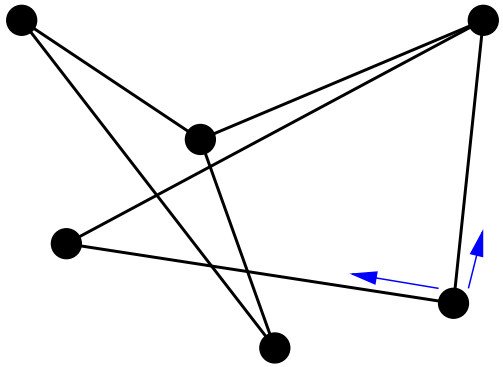


Force- Directed Methods



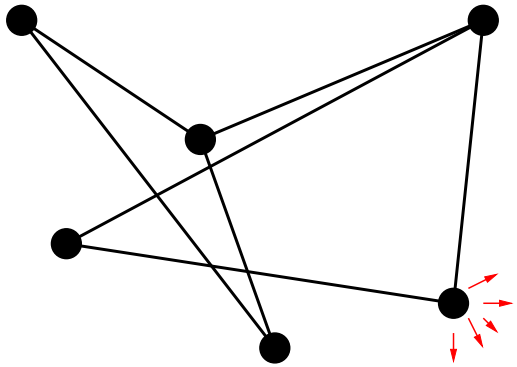
- general (undirected) graph layout
- computationally expensive ($O(n^3)$)
- non-deterministic
- low probability of crossings and occlusions

Force- Directed Methods



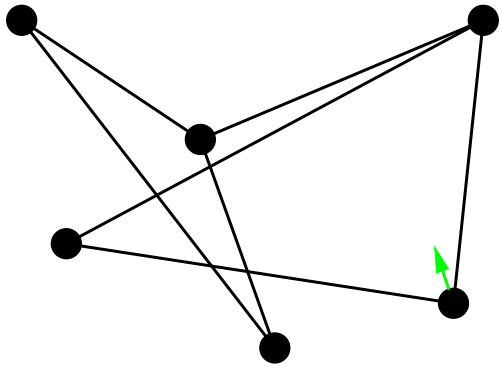
- general (undirected) graph layout
- computationally expensive ($O(n^3)$)
- non-deterministic
- low probability of crossings and occlusions

Force- Directed Methods



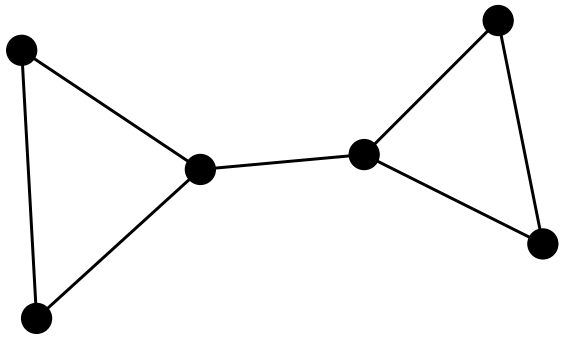
- general (undirected) graph layout
- computationally expensive ($O(n^3)$)
- non-deterministic
- low probability of crossings and occlusions

Force- Directed Methods



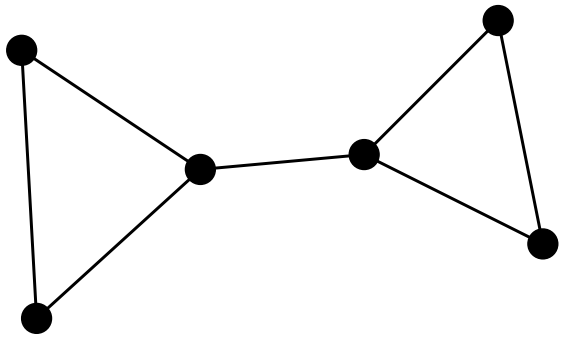
- general (undirected) graph layout
- computationally expensive ($O(n^3)$)
- non-deterministic
- low probability of crossings and occlusions

Force- Directed Methods



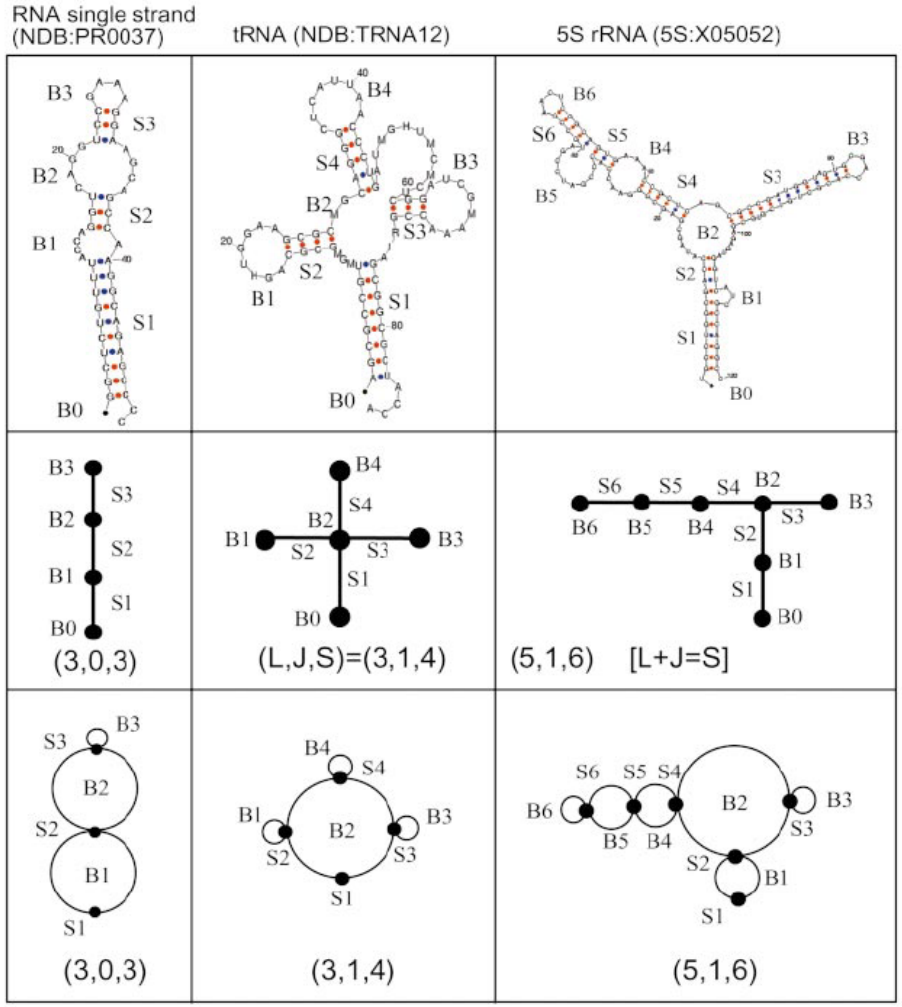
- general (undirected) graph layout
- computationally expensive ($O(n^3)$)
- non-deterministic
- low probability of crossings and occlusions

Force- Directed Methods



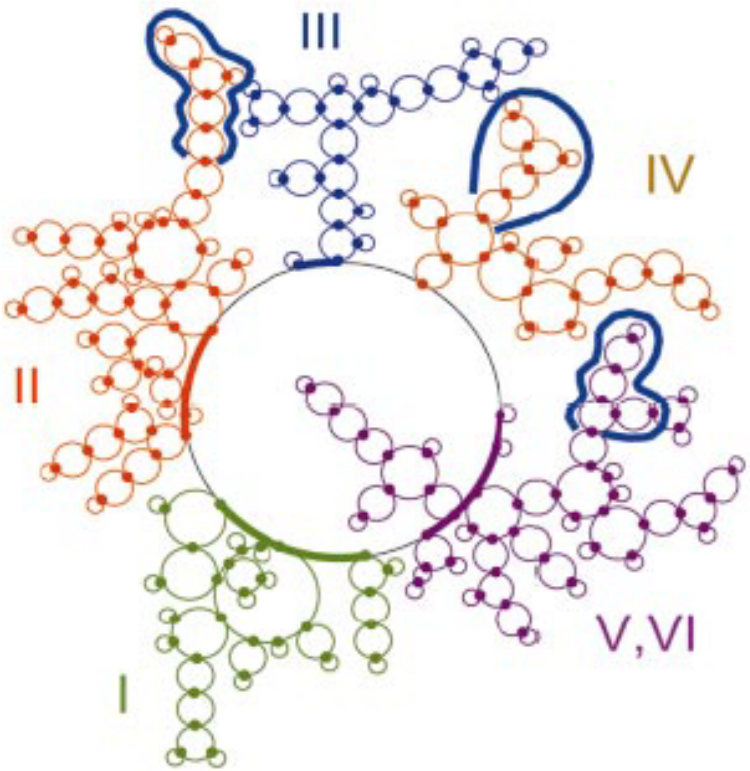
- general (undirected) graph layout
- computationally expensive ($O(n^3)$)
- non-deterministic
- low probability of crossings and occlusions

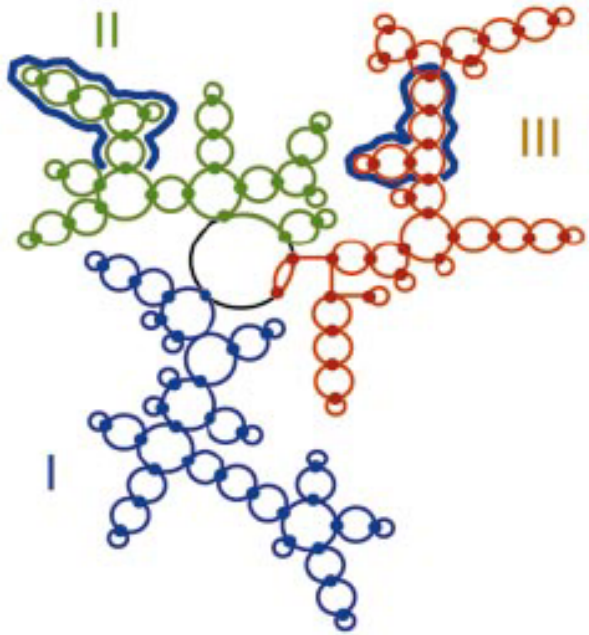
Dual Graph



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.

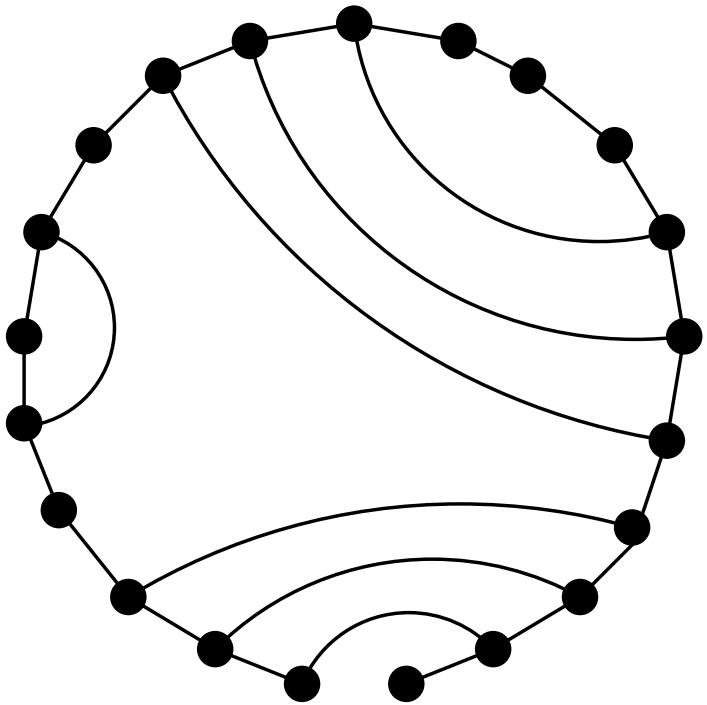
Dual Graph



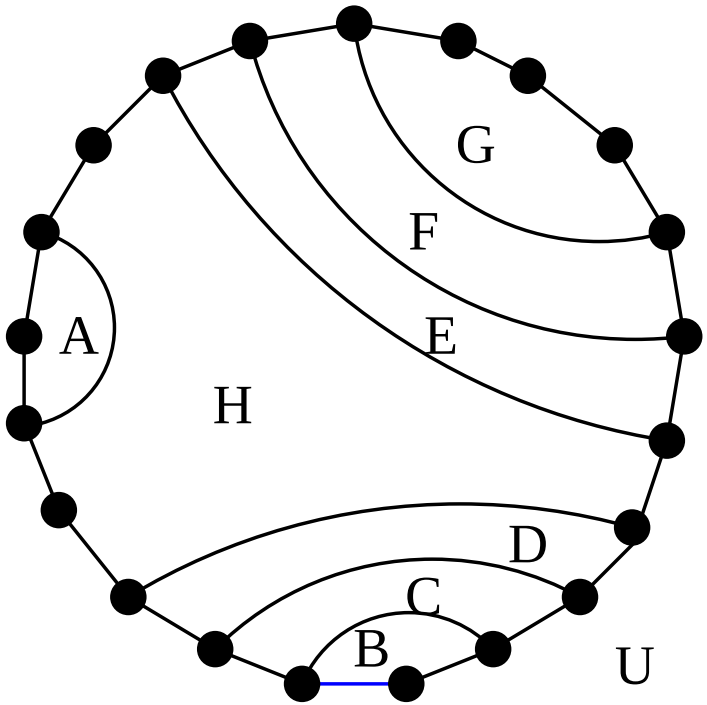


H.H. Gan, S. Pasquali, and T. Schlick, “Exploring the repertoire of RNA secondary motifs using graph theory; implications for RNA design,” *Nucleic Acids Research* 31(11), pp 2926-2943, 2003.

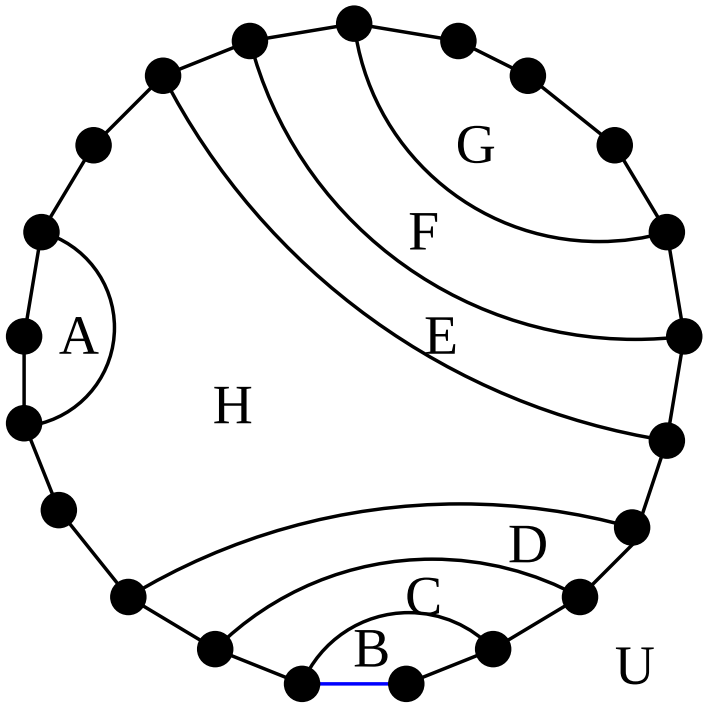
Dual Graph Minus U

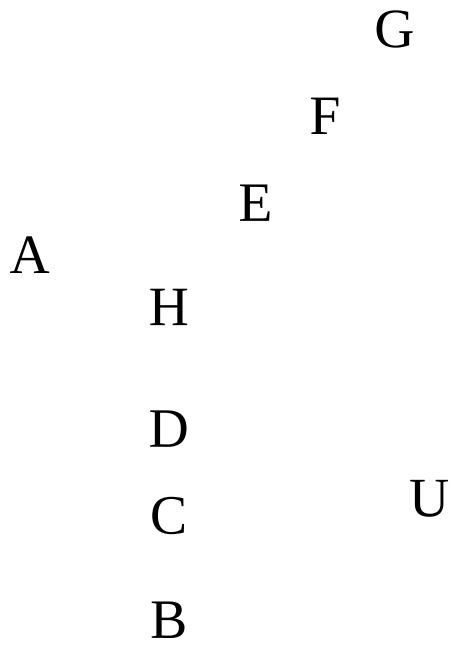


Dual Graph Minus U

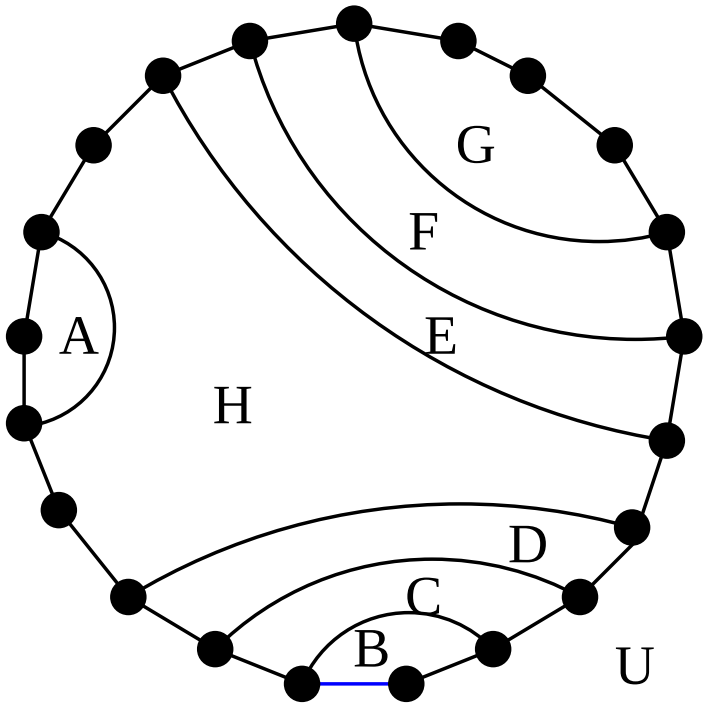


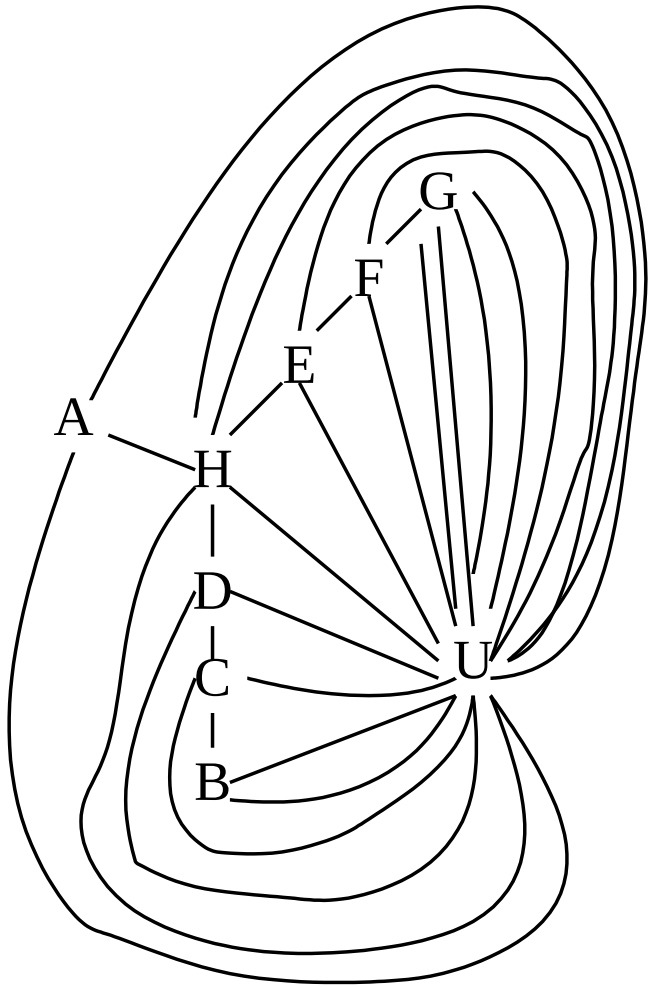
Dual Graph Minus U



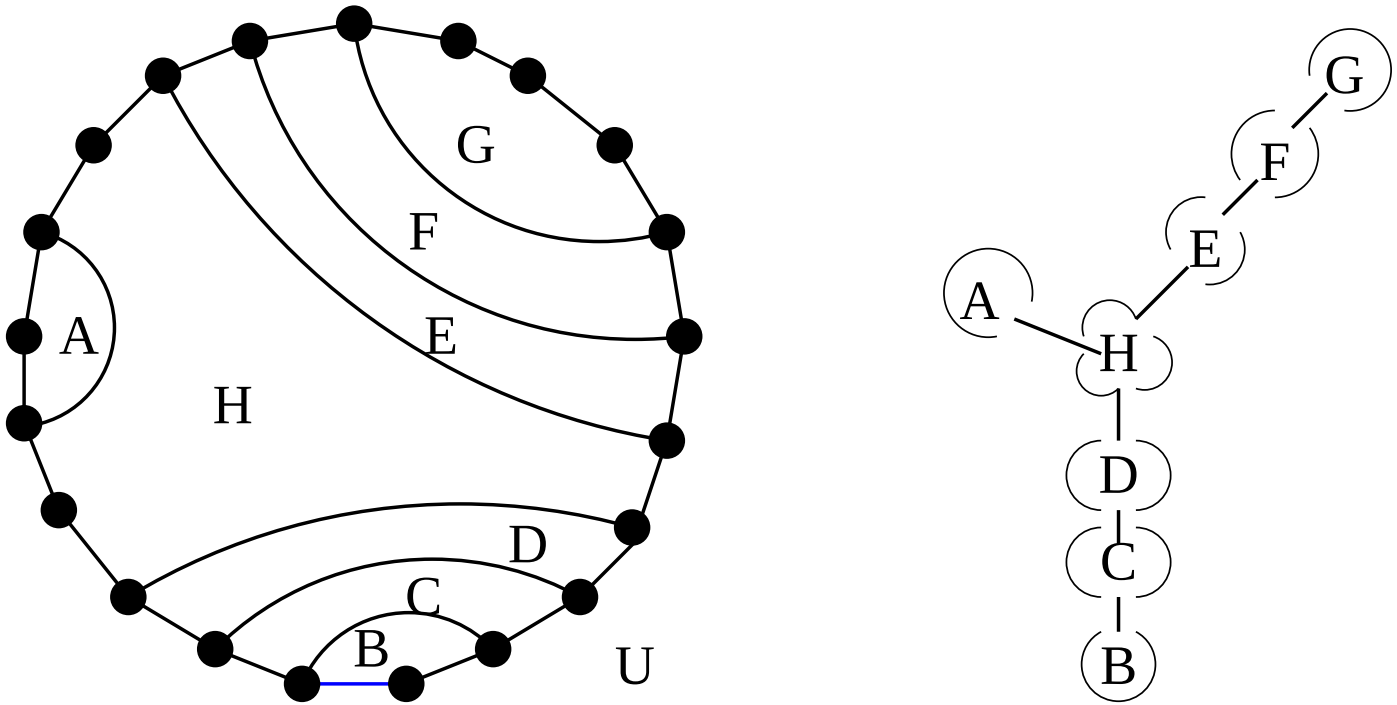


Dual Graph Minus U

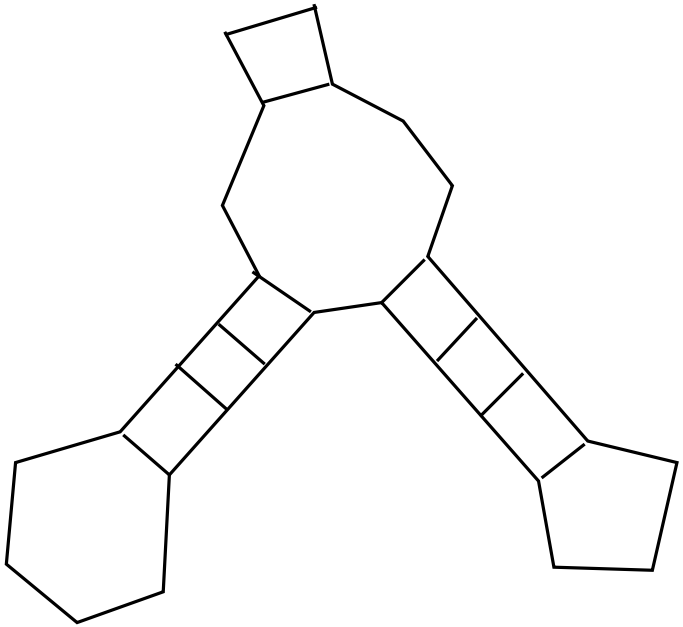


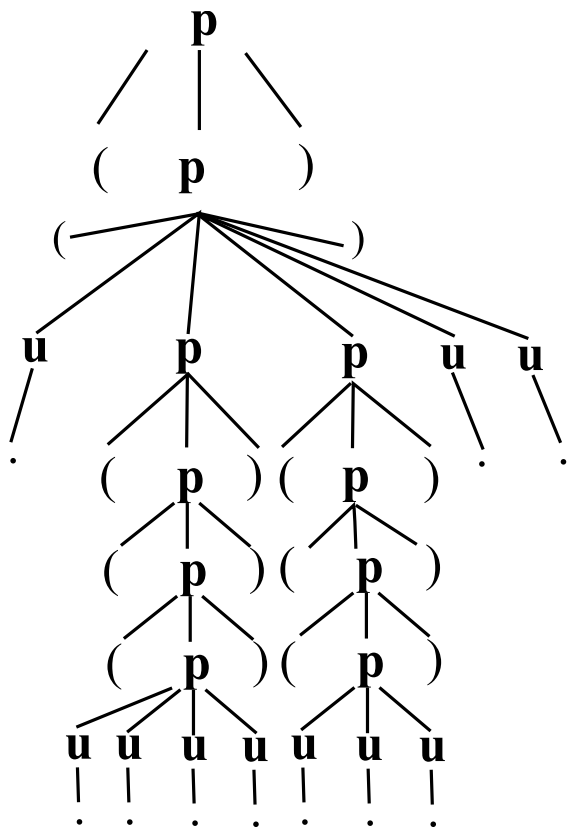


Dual Graph Minus U



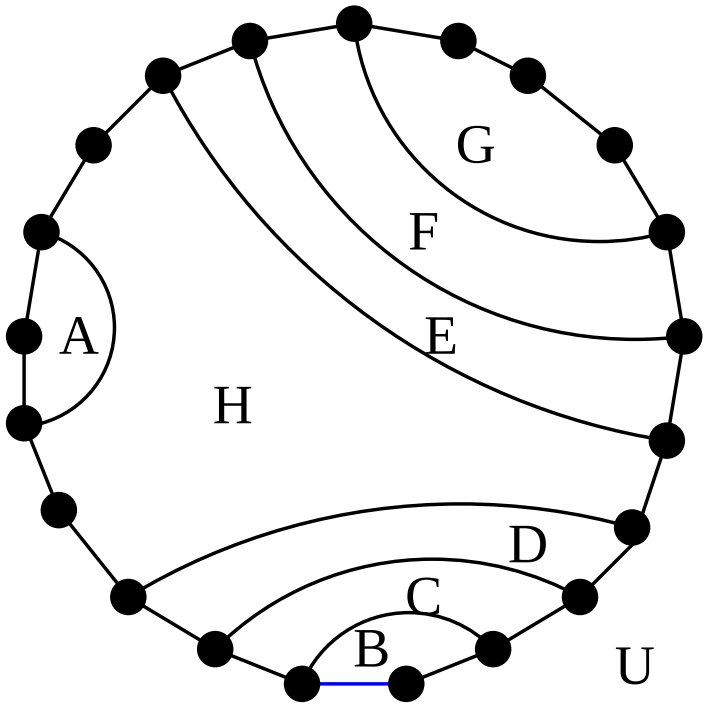
Structure Tree

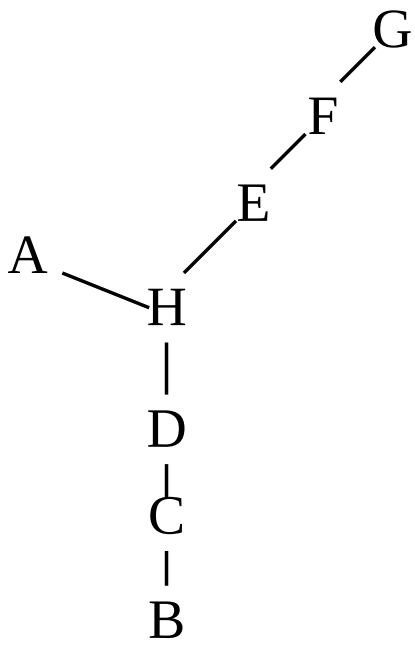




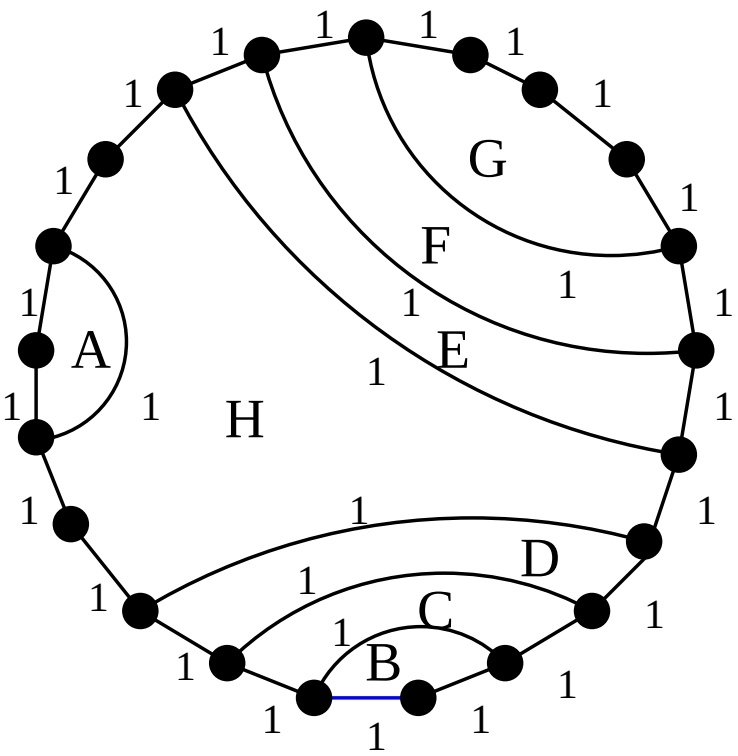
- Simple to compute the structural tree
- Simple to lay out this tree
- Problems: occlusions of subtrees and pseudoknots

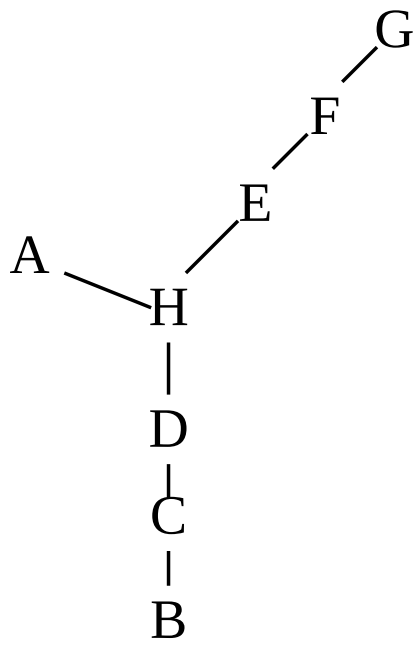
Reduced Dual Graph Minus U



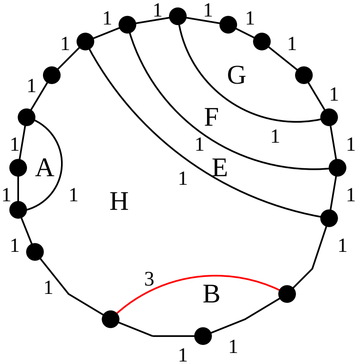


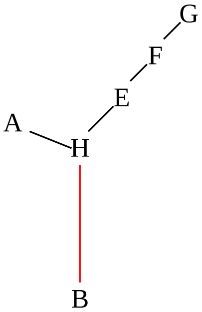
Reduced Dual Graph Minus U



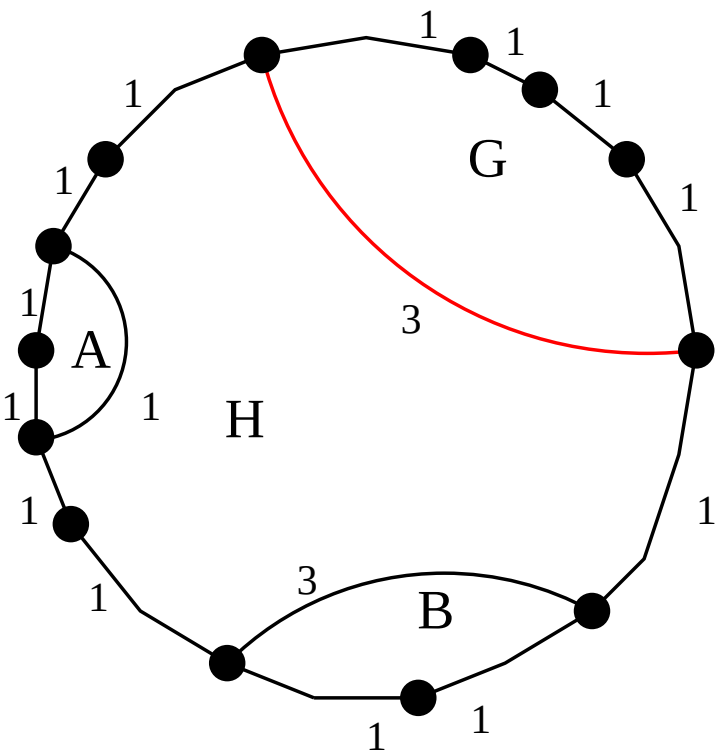


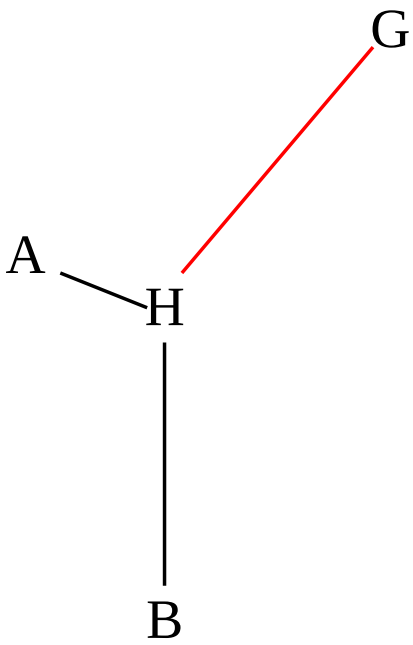
Reduced Dual Graph Minus U



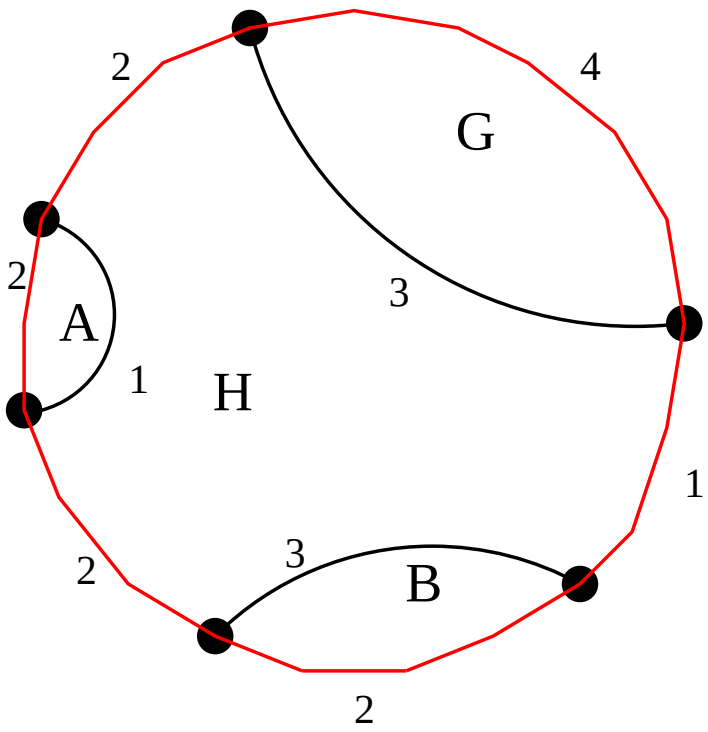


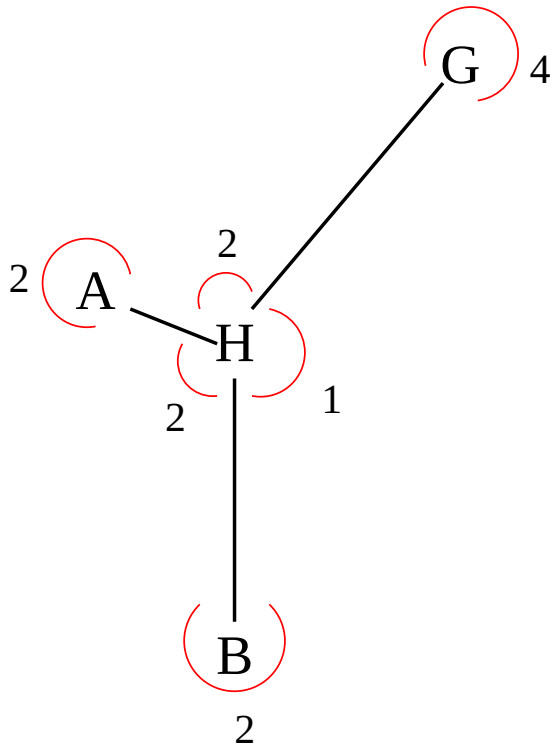
Reduced Dual Graph Minus U



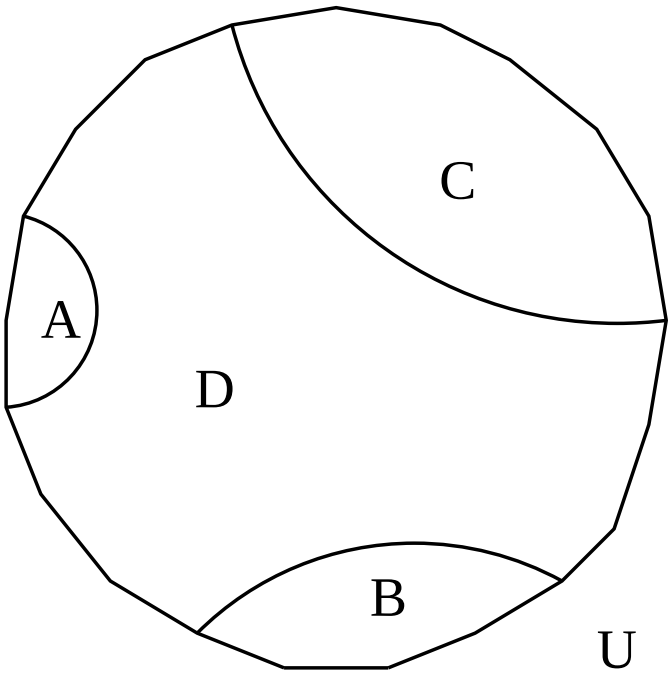


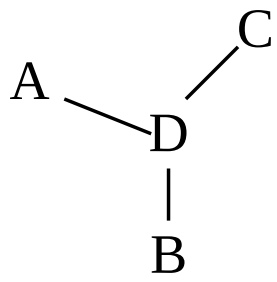
Reduced Dual Graph Minus U



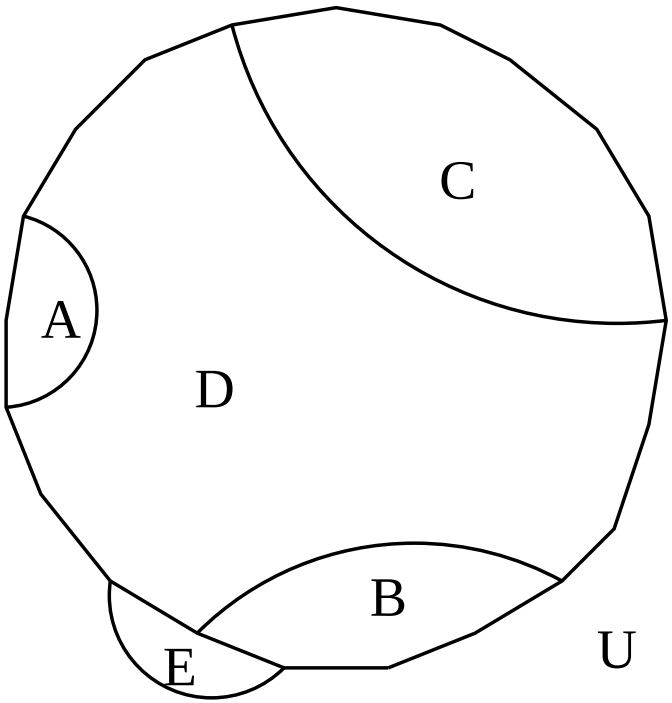


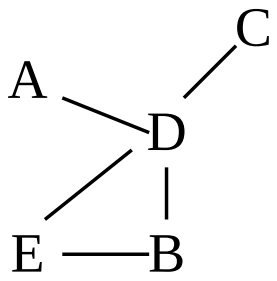
Dual Graphs with Pseudoknots

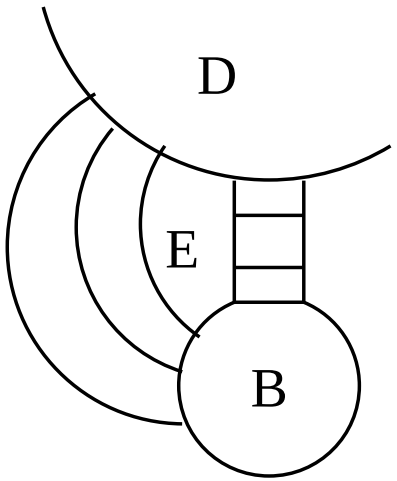




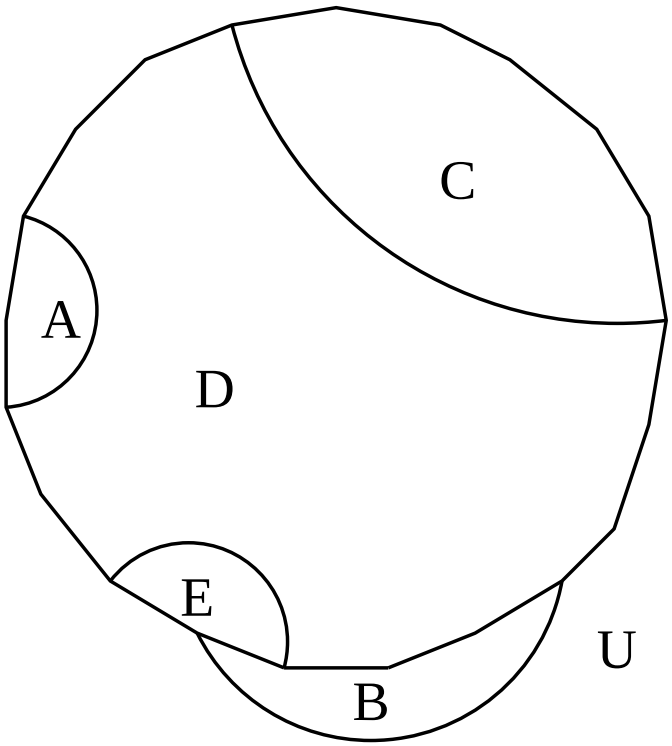
Dual Graphs with Pseudoknots

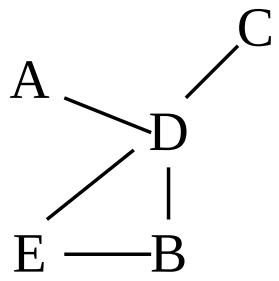


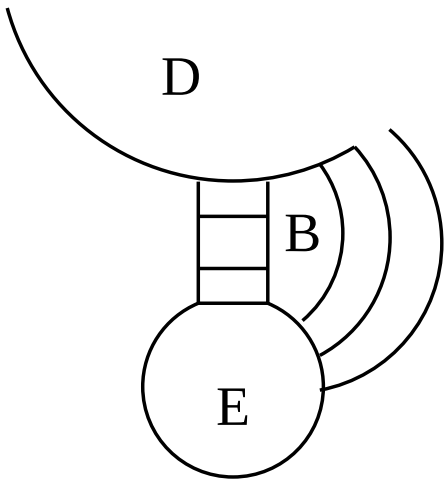




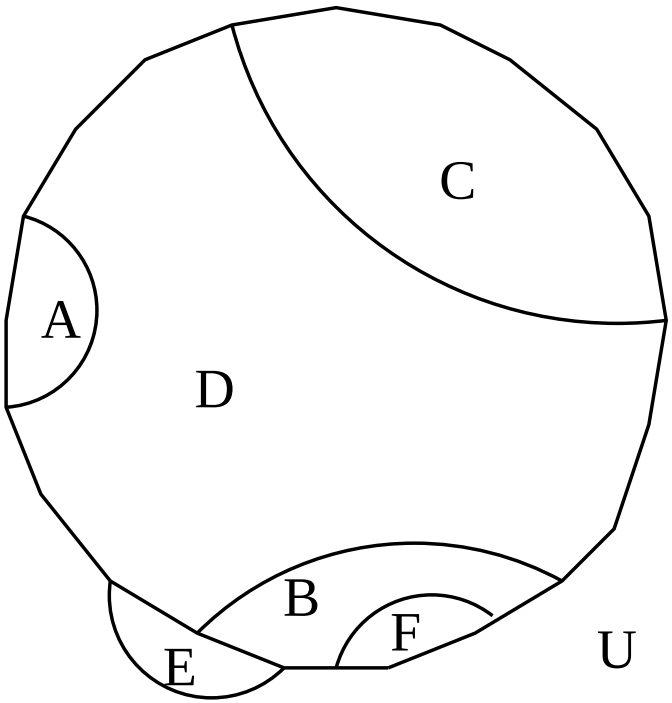
Dual Graphs with Pseudoknots

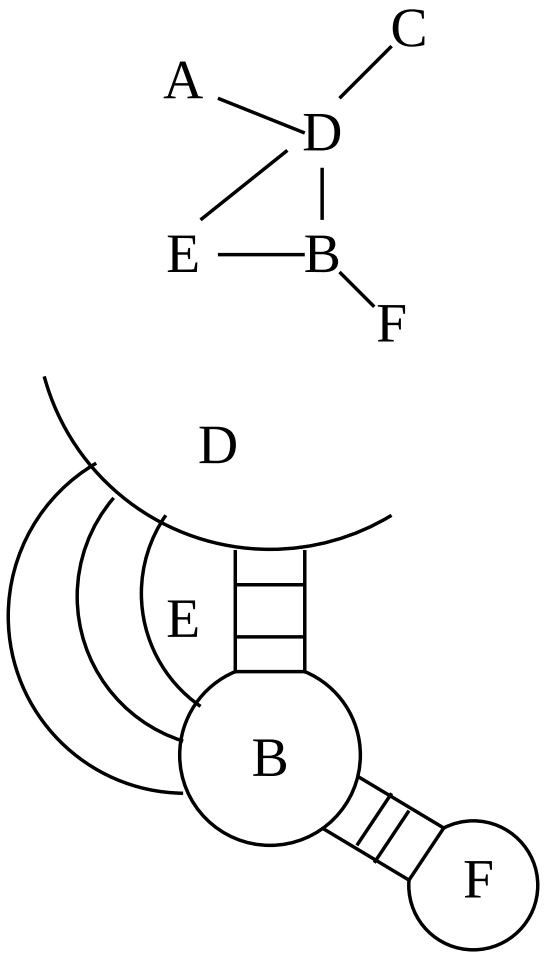




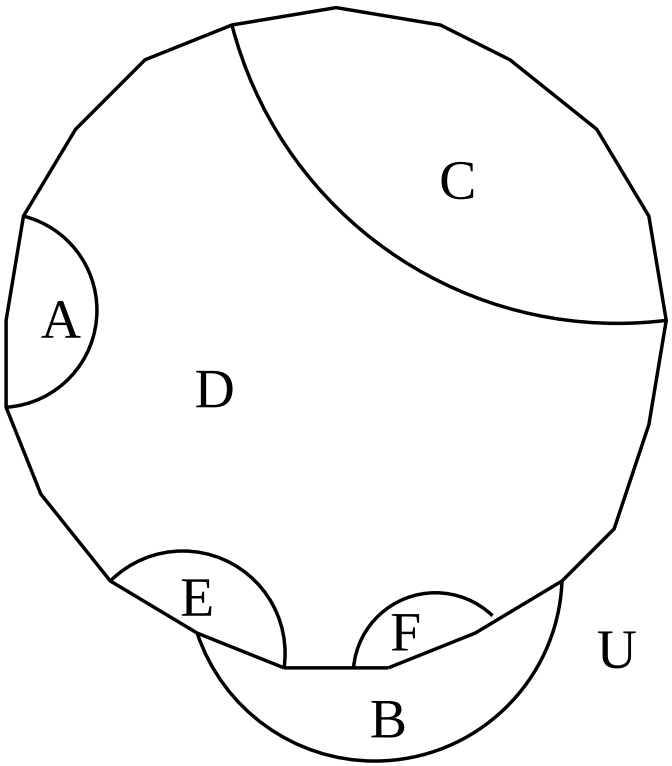


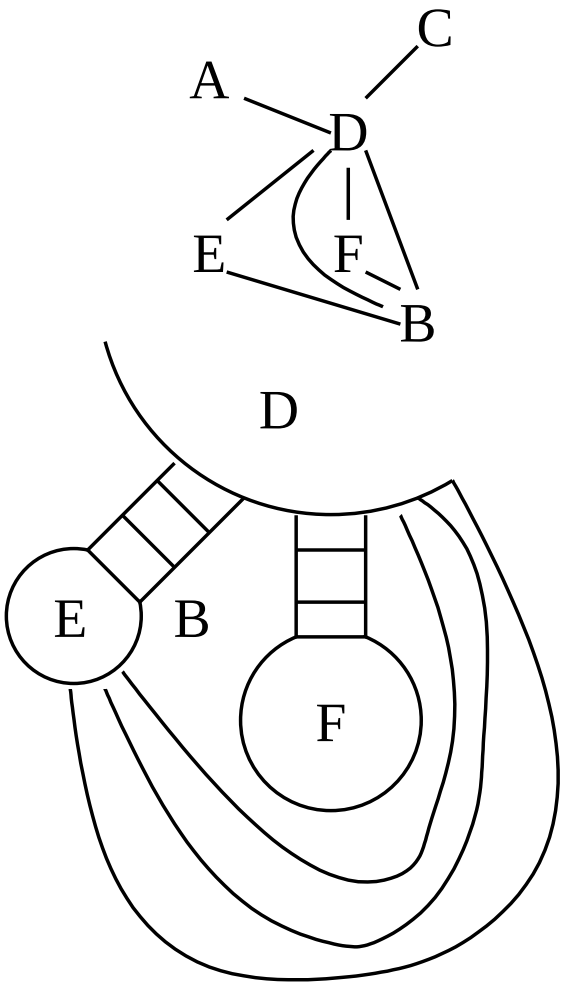
Dual Graphs with Pseudoknots



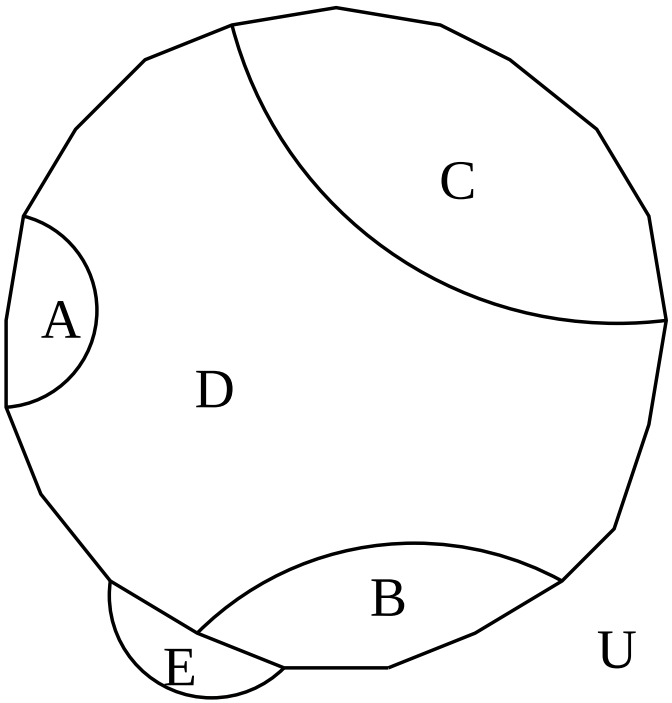


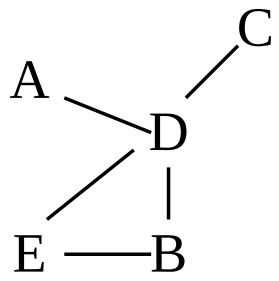
Dual Graphs with Pseudoknots

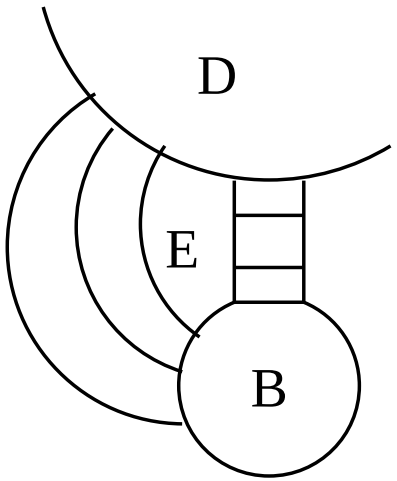




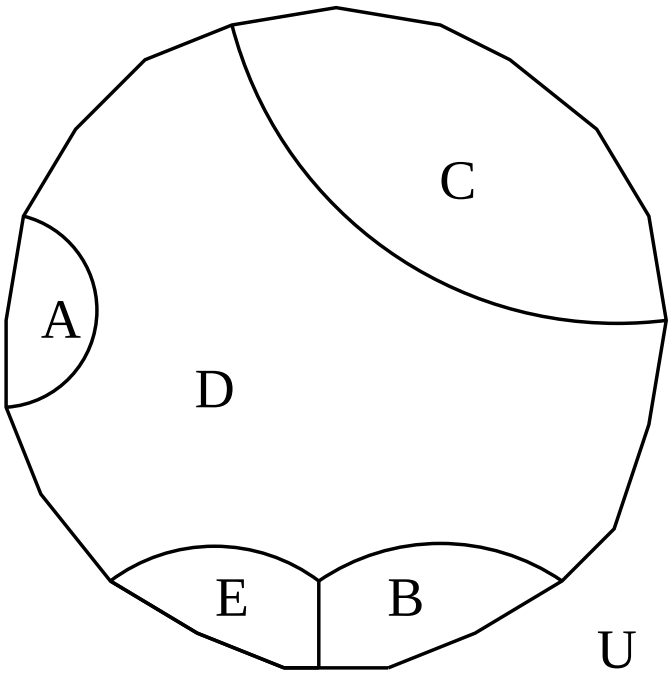
Dual Graphs with Pseudoknots



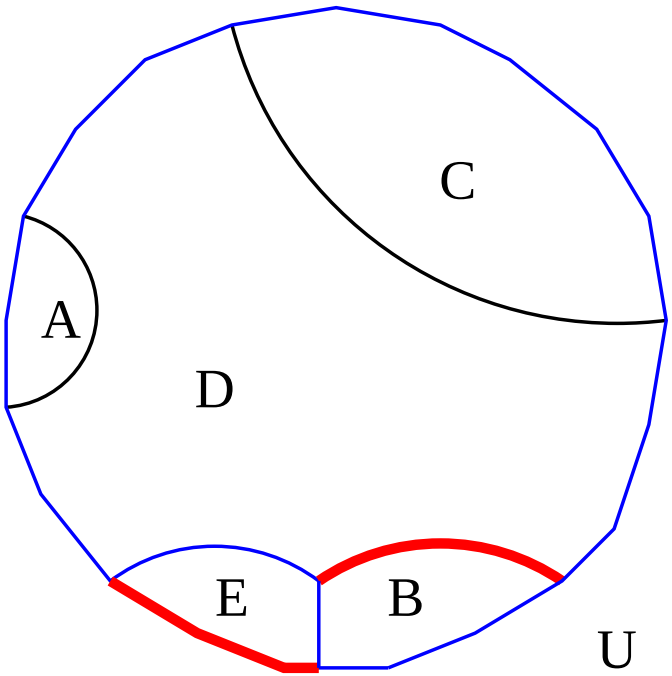




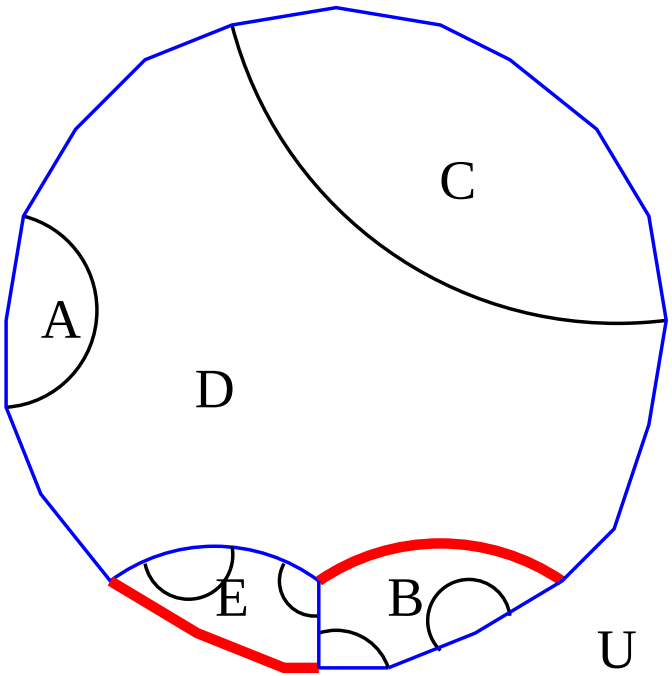
Dual Graphs with Pseudoknots



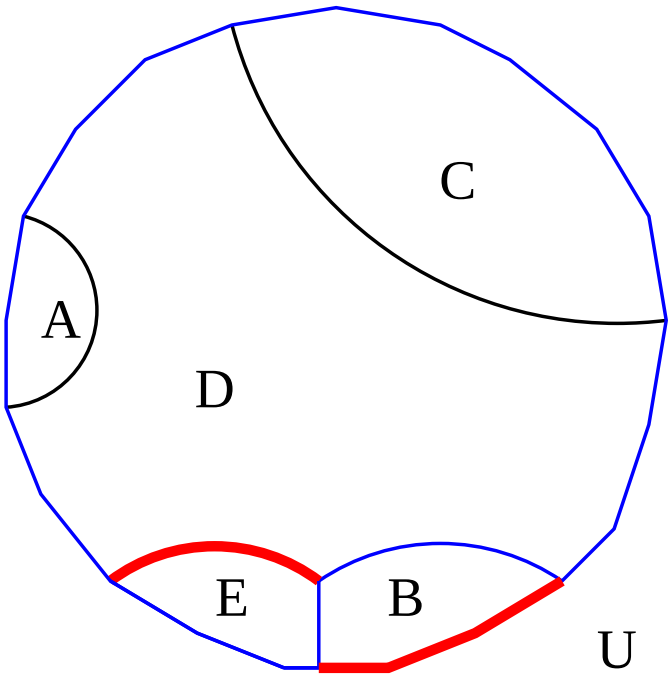
Dual Graphs with Pseudoknots



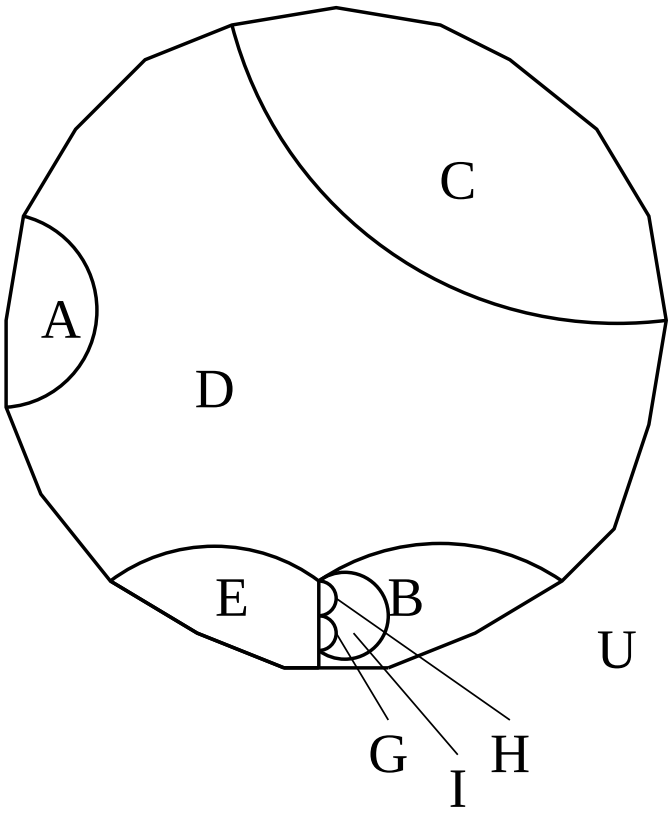
Dual Graphs with Pseudoknots



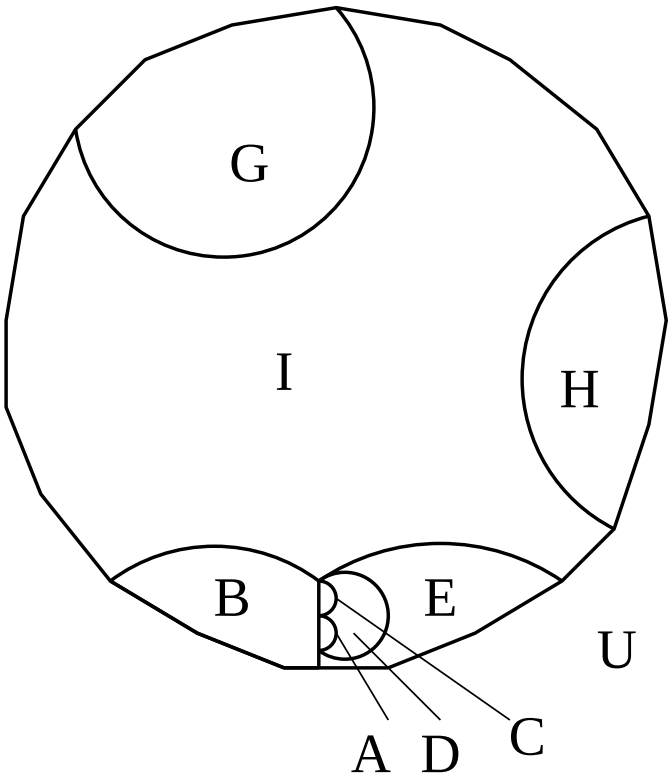
Dual Graphs with Pseudoknots



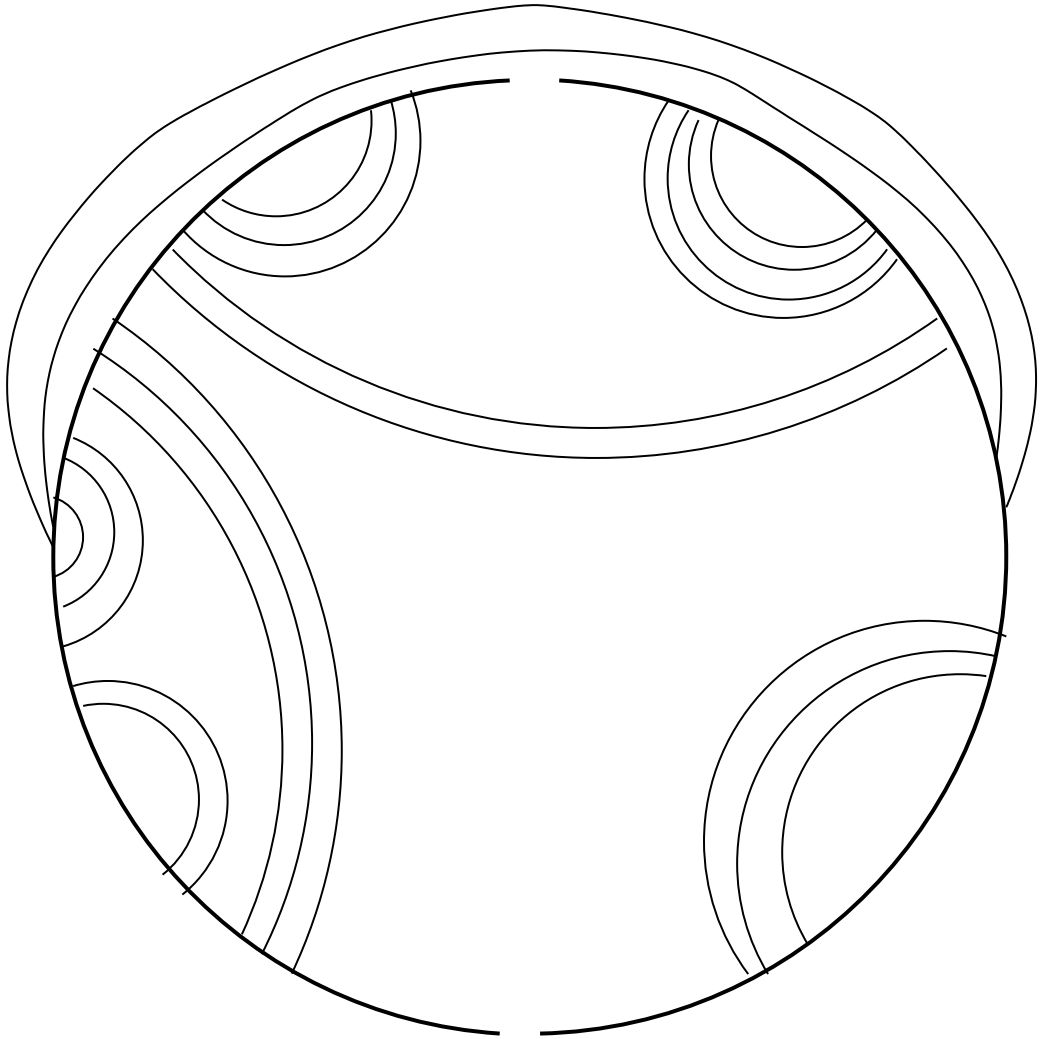
Dual Graphs with Pseudoknots



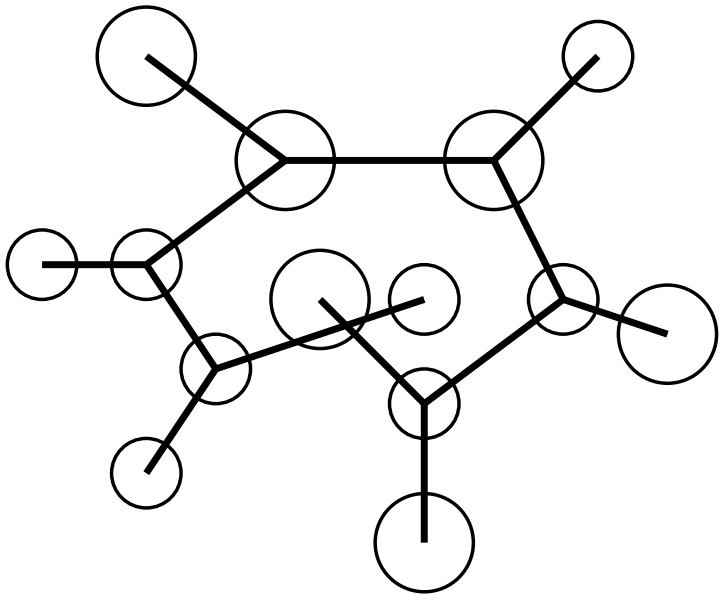
Dual Graphs with Pseudoknots



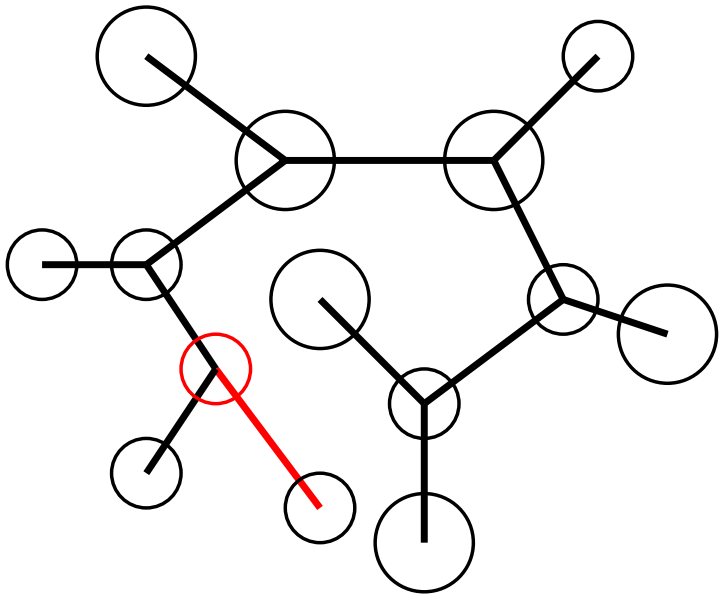
Is the visualization of pseudoknots a solution to visualization cotranscriptional folding?



The Occlusion Problem

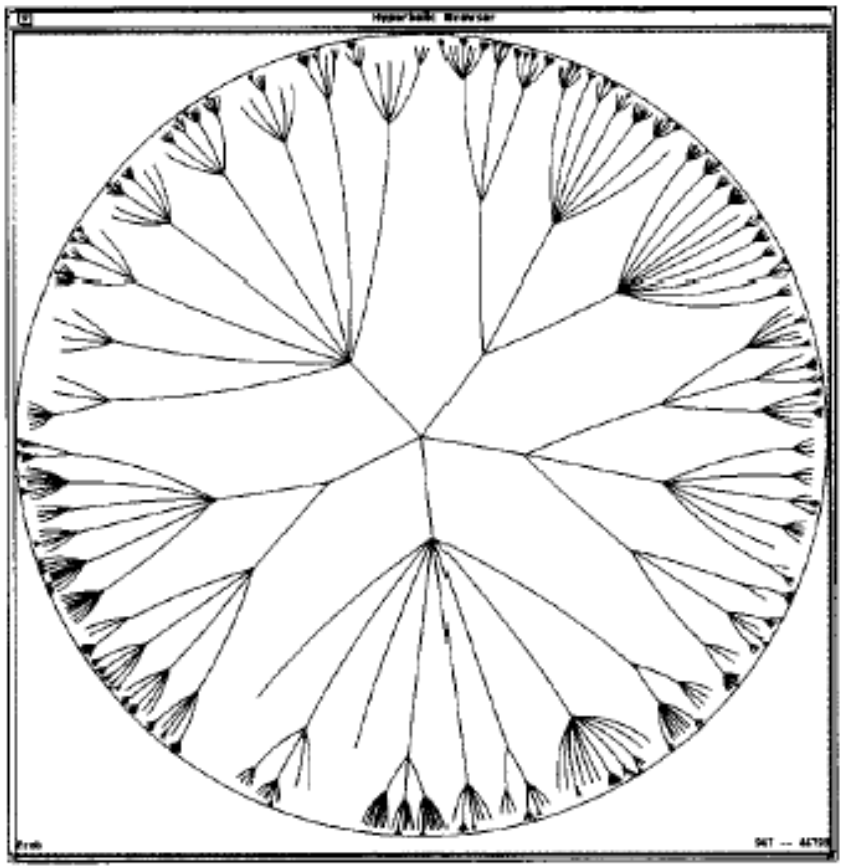


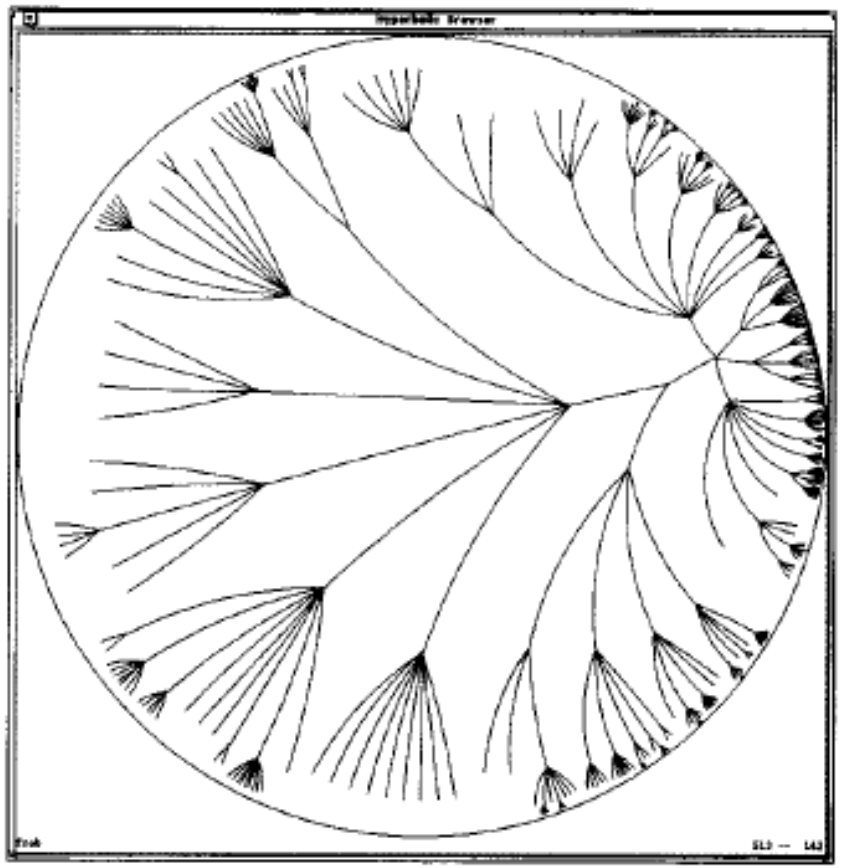
The Occlusion Problem



The Occlusion Problem

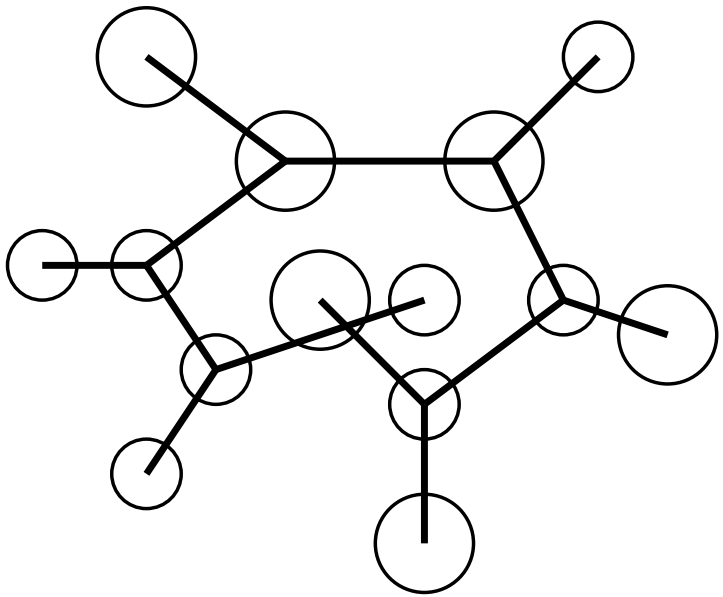
- Layout in hyperbolic space





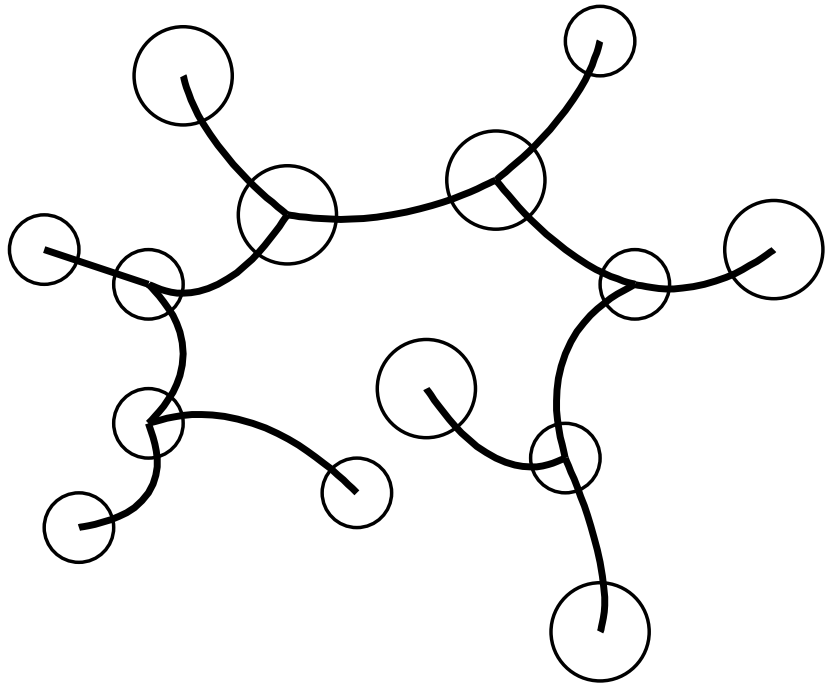
The Occlusion Problem

- Force- Directed- “Flexible Stack” Layout.



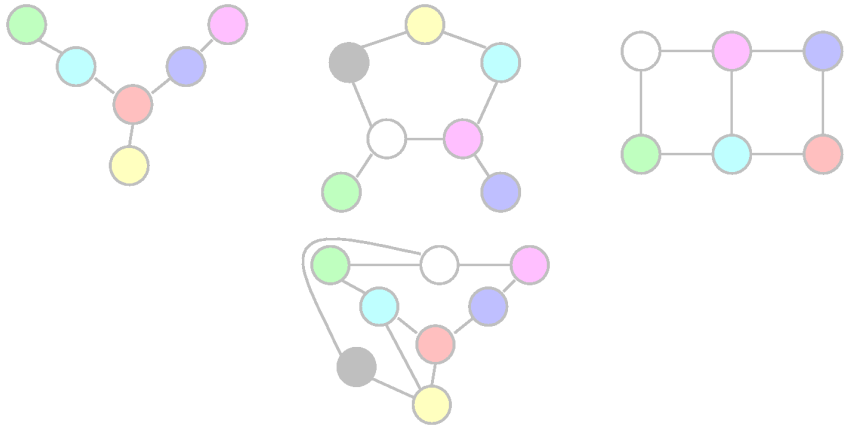
The Occlusion Problem

- Force- Directed- “Flexible Stack” Layout.



Visualization of Consensus Structures

$G_i = (V_i, E_i)$



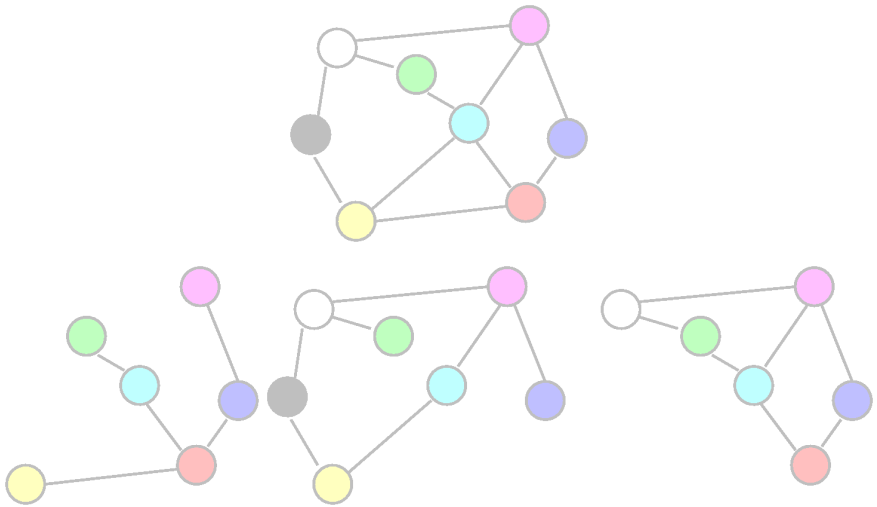
$G = (\bigcup V_i, \bigcup E_i)$

no layout

---

with layout

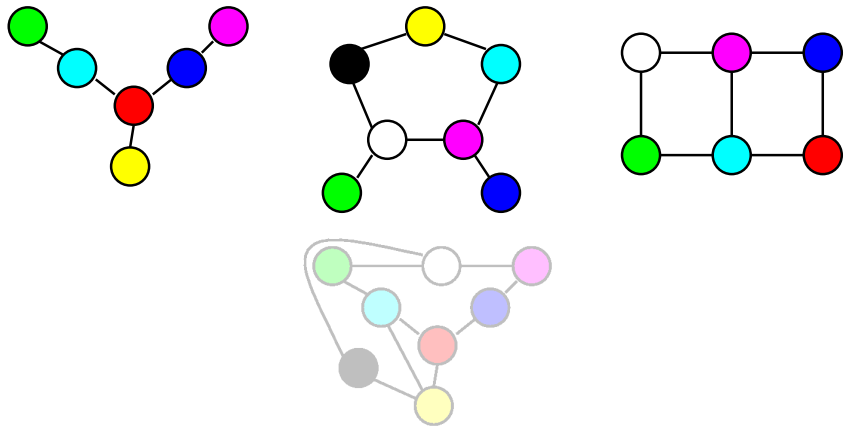
$L = L(G)$



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

Visualization of Consensus Structures

$G_i = (V_i, E_i)$



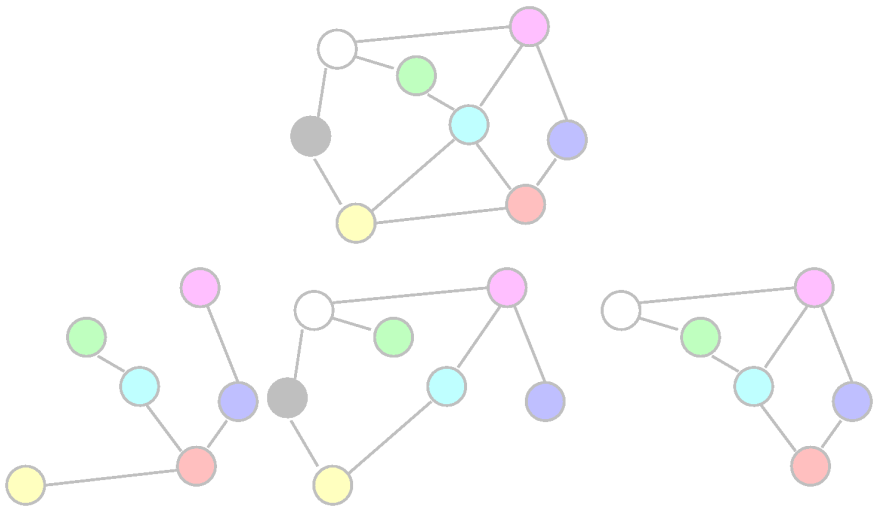
$G = (\bigcup V_i, \bigcup E_i)$

no layout

---

with layout

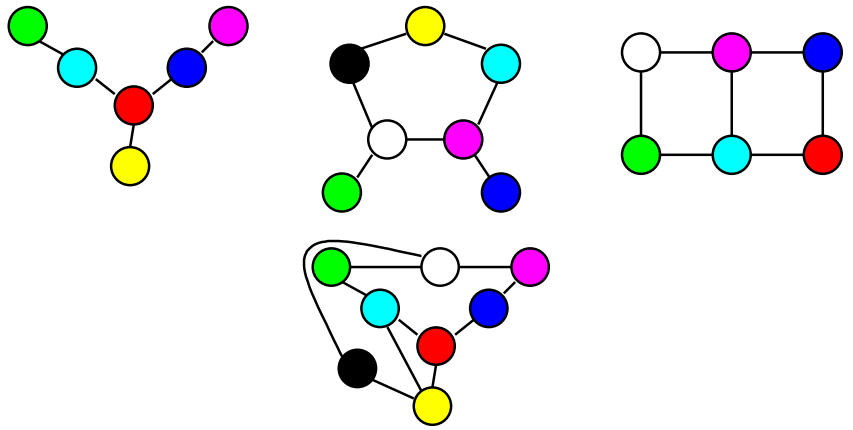
$L = L(G)$



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

Visualization of Consensus Structures

$G_i = (V_i, E_i)$



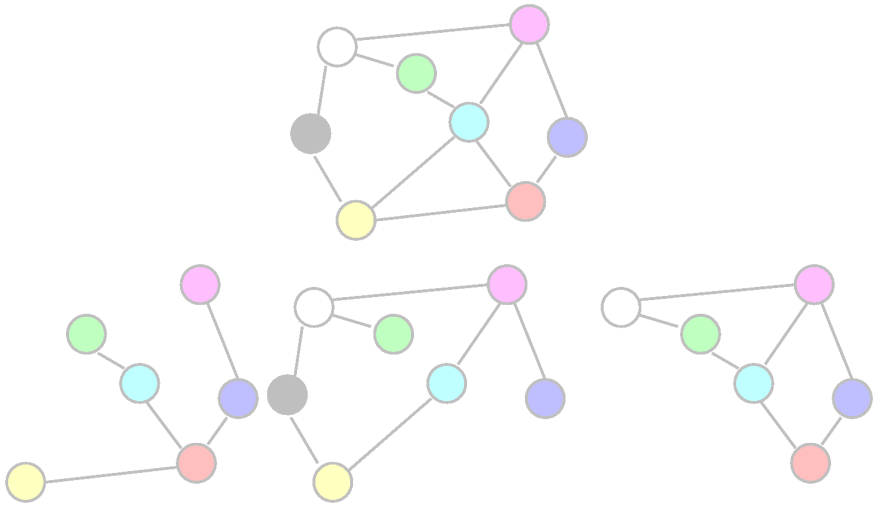
$G = (\bigcup V_i, \bigcup E_i)$

no layout

---

with layout

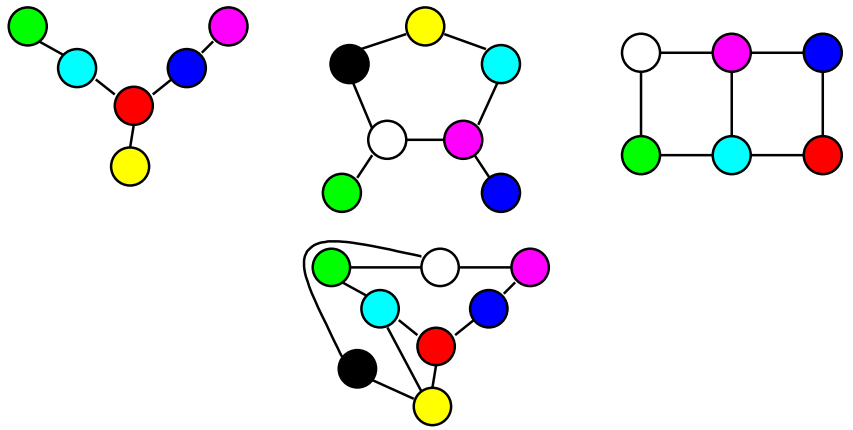
$L = L(G)$



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

Visualization of Consensus Structures

$G_i = (V_i, E_i)$



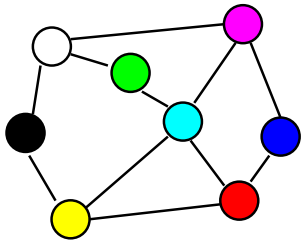
$G = (\bigcup V_i, \bigcup E_i)$

no layout

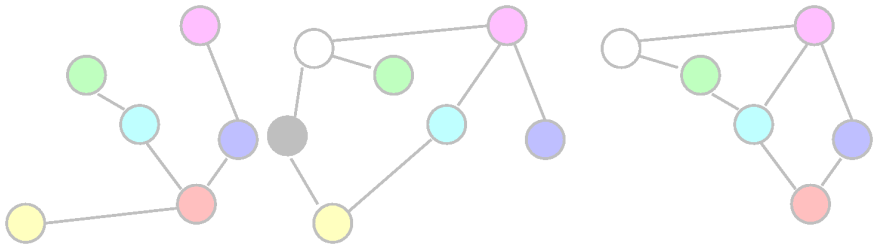
---

with layout

$L = L(G)$

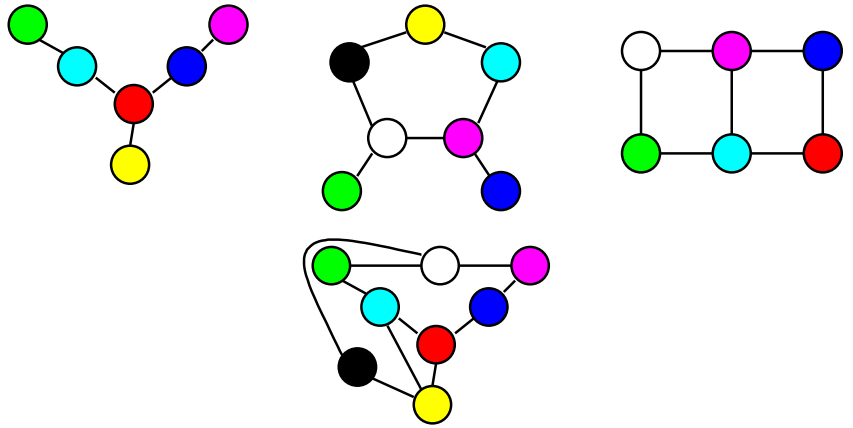


$L(G_i) = L(G)|_{G_i}$



Visualization of Consensus Structures

$G_i = (V_i, E_i)$



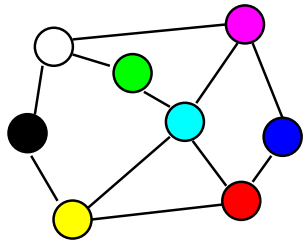
$G = (\bigcup V_i, \bigcup E_i)$

no layout

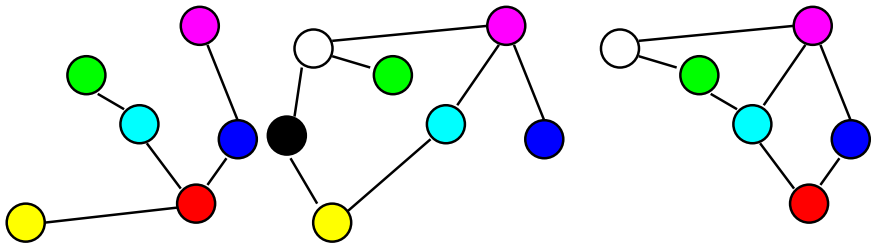
---

with layout

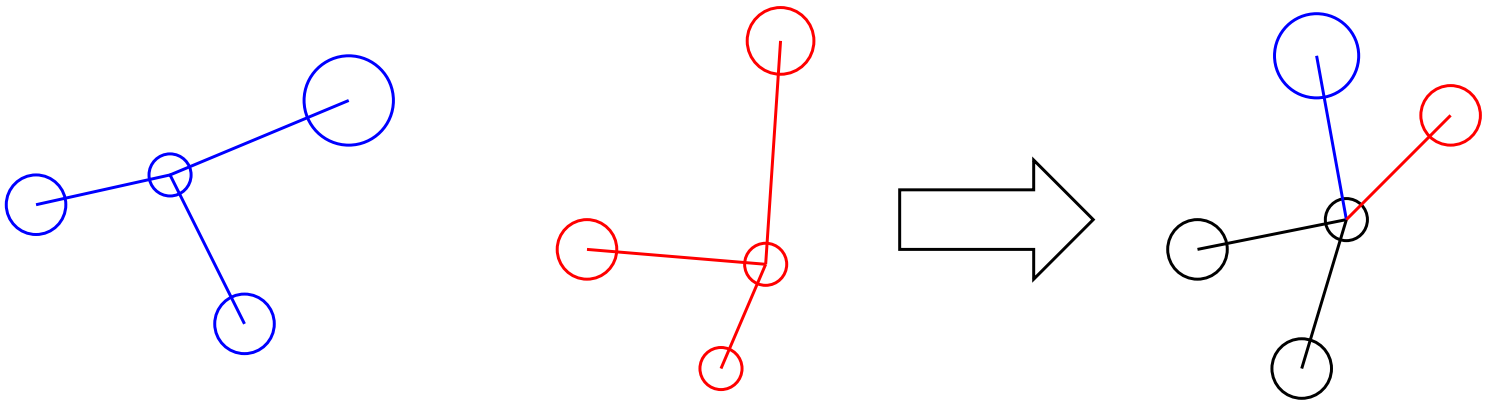
$L = L(G)$



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



Consensus Structure



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

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!