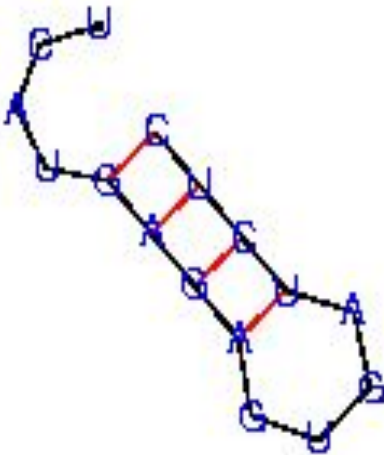
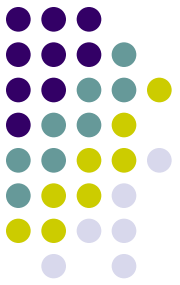


# Dynamic Programming Approaches





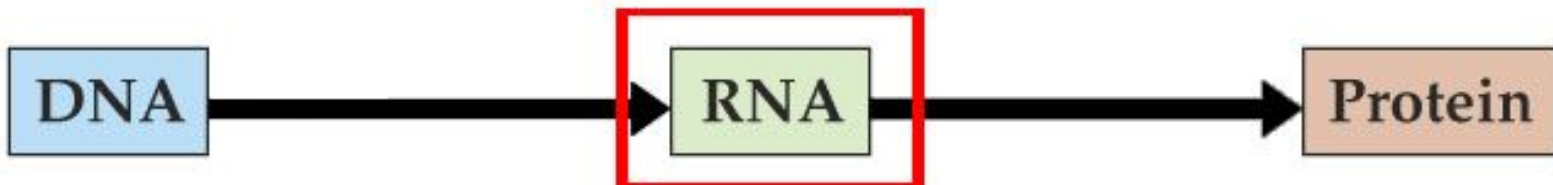
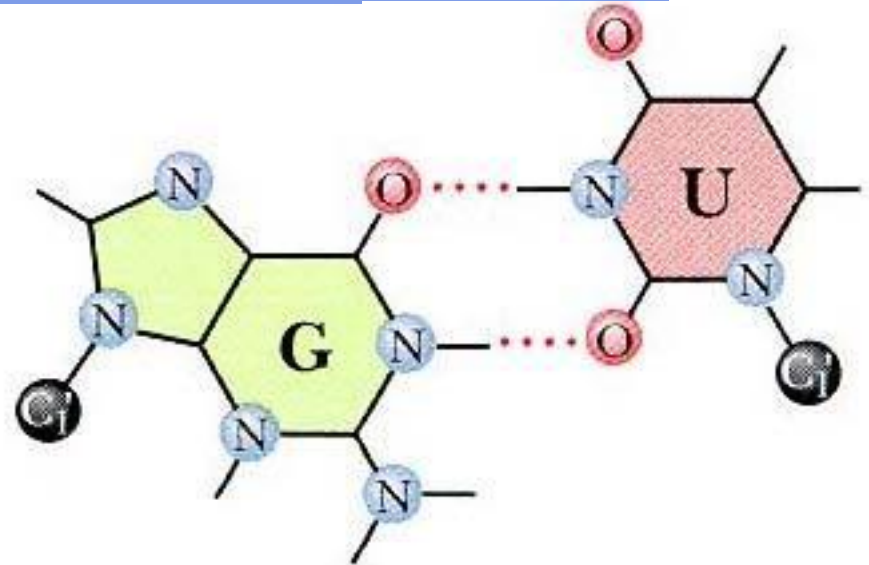
# Outline

- RNA folding
- Dynamic programming for RNA secondary structure prediction
- Covariance model for RNA structure prediction

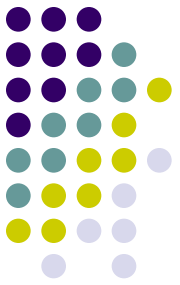
# RNA Basics

- RNA bases A,C,G,U
- Canonical Base Pairs
  - A-U
  - G-C
  - G-U
- “wobble” pairing
- Bases can only pair with **one** other base.

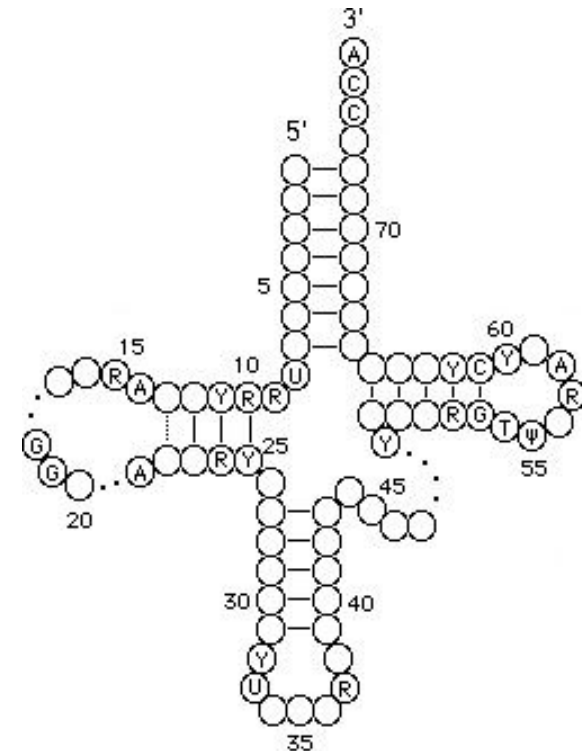
3 Hydrogen Bonds – more stable



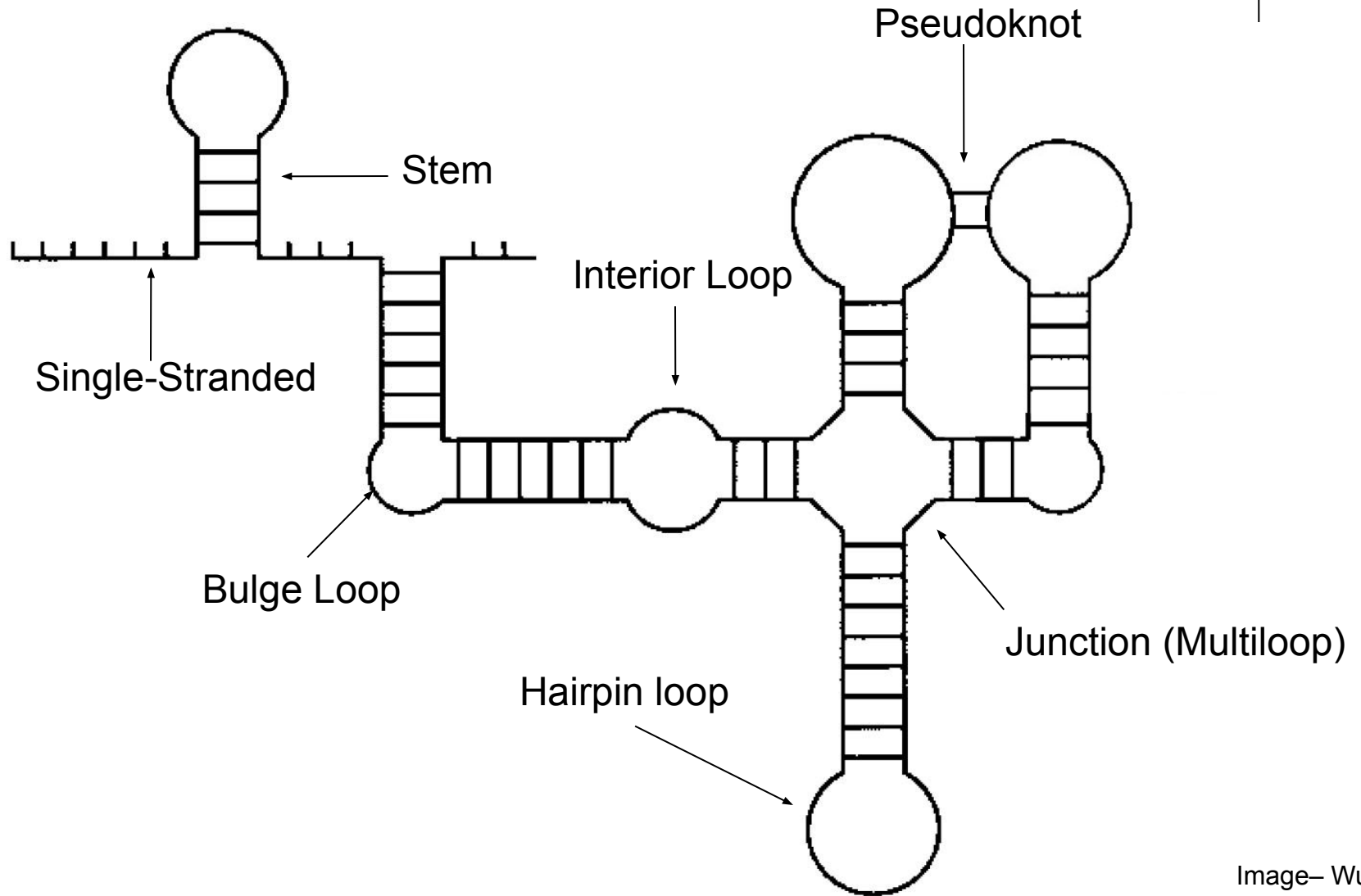
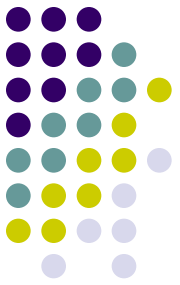
# RNA Basics



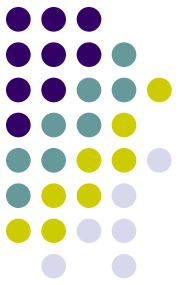
- transfer RNA (tRNA)
- messenger RNA (mRNA)
- ribosomal RNA (rRNA)
- small interfering RNA (siRNA)
- micro RNA (miRNA)
- small nucleolar RNA (snoRNA)



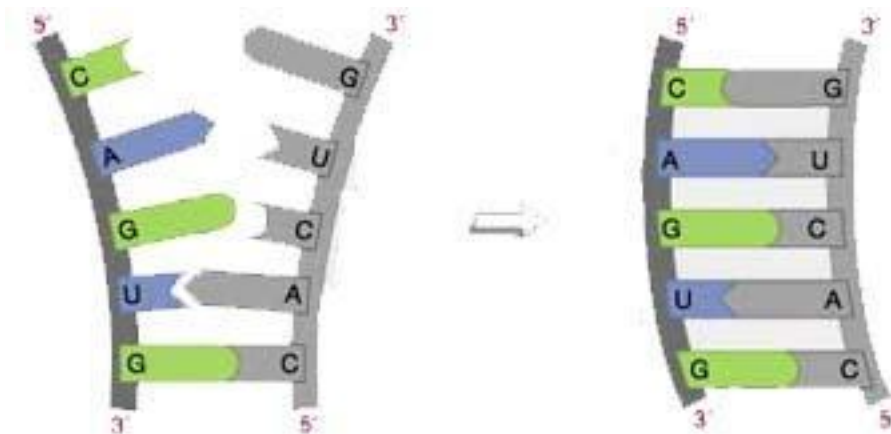
# RNA Secondary Structure



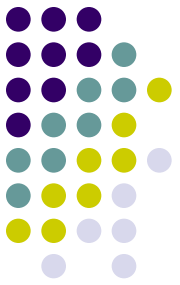
# Sequence Alignment as a method to determine structure



- Bases pair in order to form backbones and determine the secondary structure
- Aligning bases based on their ability to pair with each other gives an algorithmic approach to determining the optimal structure



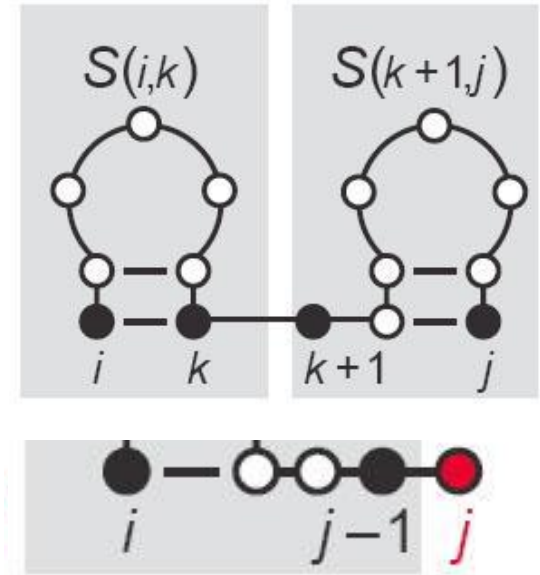
# Base Pair Maximization – Dynamic Programming Algorithm



*$S(i,j)$  is the folding of the subsequence of the RNA strand from index  $i$  to index  $j$  which results in the highest number of base pairs*

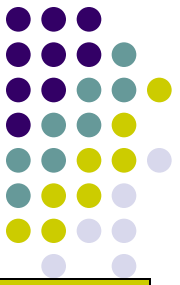
*Maximizing Base Pairs*

$$S(i,j) = \max \begin{cases} S(i+1, j-1) + 1 & [\text{if } i,j \text{ base pair}] \\ S(i+1, j) \\ S(i, j-1) \\ \max_{i < k < j} S(i, k) + S(k+1, j) \end{cases}$$

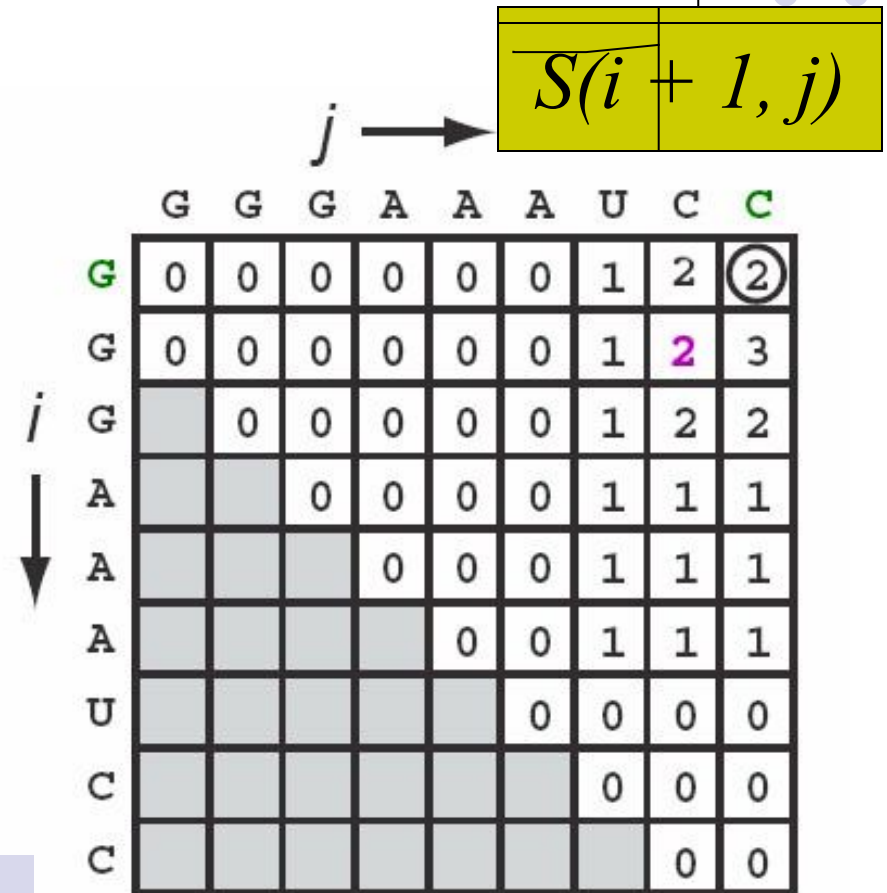


*If no base pair at  $i$  and  $j$*

# Base Pair Maximization – Dynamic Programming Algorithm



- Alignment Method
  - Align RNA strand to itself
  - Score increases for feasible base pairs
- Each score independent of overall structure
- Bifurcation adds extra dimension

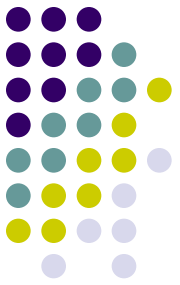


Dynamic Programming – possible paths

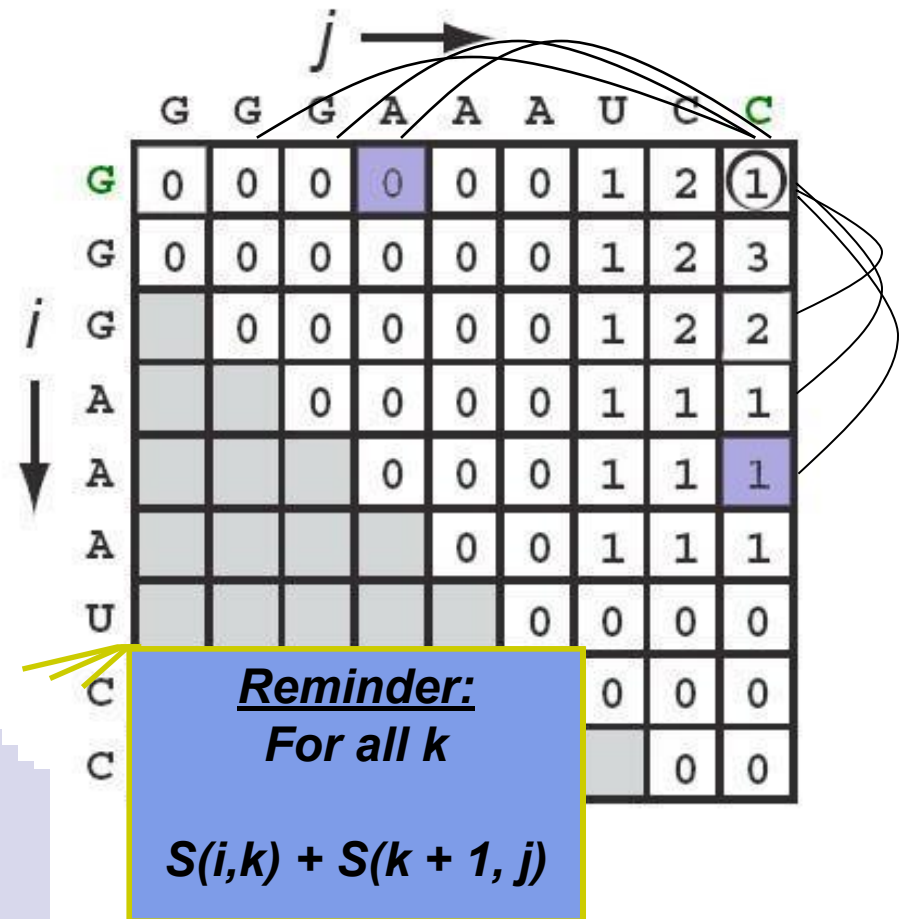
$$S(i + 1, j - 1) + 1$$



# Base Pair Maximization – Dynamic Programming Algorithm

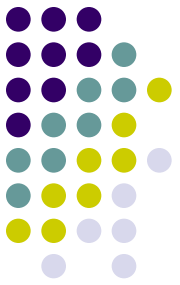


- Alignment Method
  - Align RNA strand to itself
  - Score increases for feasible base pairs
- Each score independent of overall structure
- Bifurcation adds extra dimension



Bifurcation – add values  
for all  $k$

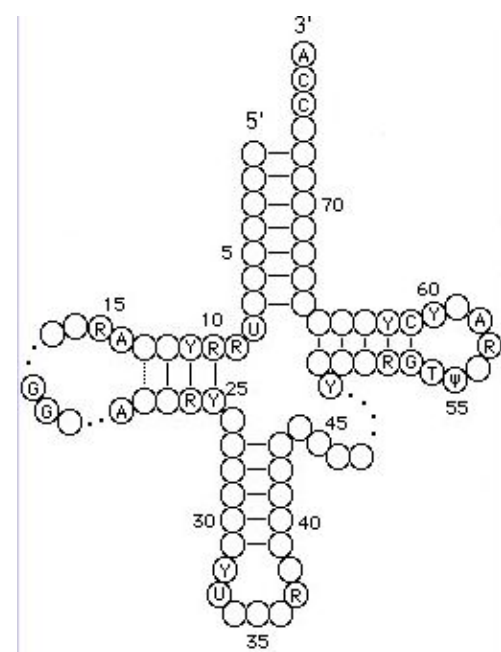
# Base Pair Maximization - Drawbacks



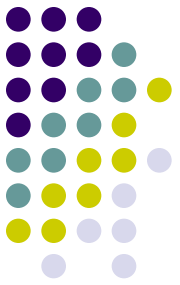
- Base pair maximization will not necessarily lead to the most stable structure
  - May create structure with many interior loops or hairpins which are energetically unfavorable
- Comparable to aligning sequences with scattered matches – not biologically reasonable

# Energy Minimization

- Thermodynamic Stability
  - Estimated using experimental techniques
  - Theory : Most Stable is the Most likely
- No Pseudoknots due to algorithm limitations
- Uses Dynamic Programming alignment technique
- Attempts to maximize the score taking into account thermodynamics
- MFOLD and ViennaRNA



# References



- [How Do RNA Folding Algorithms Work?](#). S.R. Eddy. [Nature Biotechnology](#), 22:1457-1458, 2004.