# BioE 131, HW4 — by Gautam Machiraju

1. **What does a high mutual information imply about the nucleotides at positions *i* and *j*?**

   In general terms, the mutual information of two random variables measures the mutual dependence of said random variables. In particular, a higher mutual information between two nucleotides therefore implies that nucleotides *i* and *j* are highly dependent on each other to maintain the structure and function of the tRNA structure they conform to.

   For example, if positions *i* and *j* have higher mutual information, they may correspond to a base-pairing relationship, such as in a tRNA stem structure. Additionally, a higher mutual information may also correspond to positions *i* and *j* in the anticodon loop, which requires the correct anticodon to base pair with the tRNA's corresponding codon.

2. **With the aid of a sketch (may be hand drawn or computer generated), summarize what the "top 20" and "top 50" results tell you about the secondary structure of transfer RNA.**

Entry 2   Entry 3   Entry 1
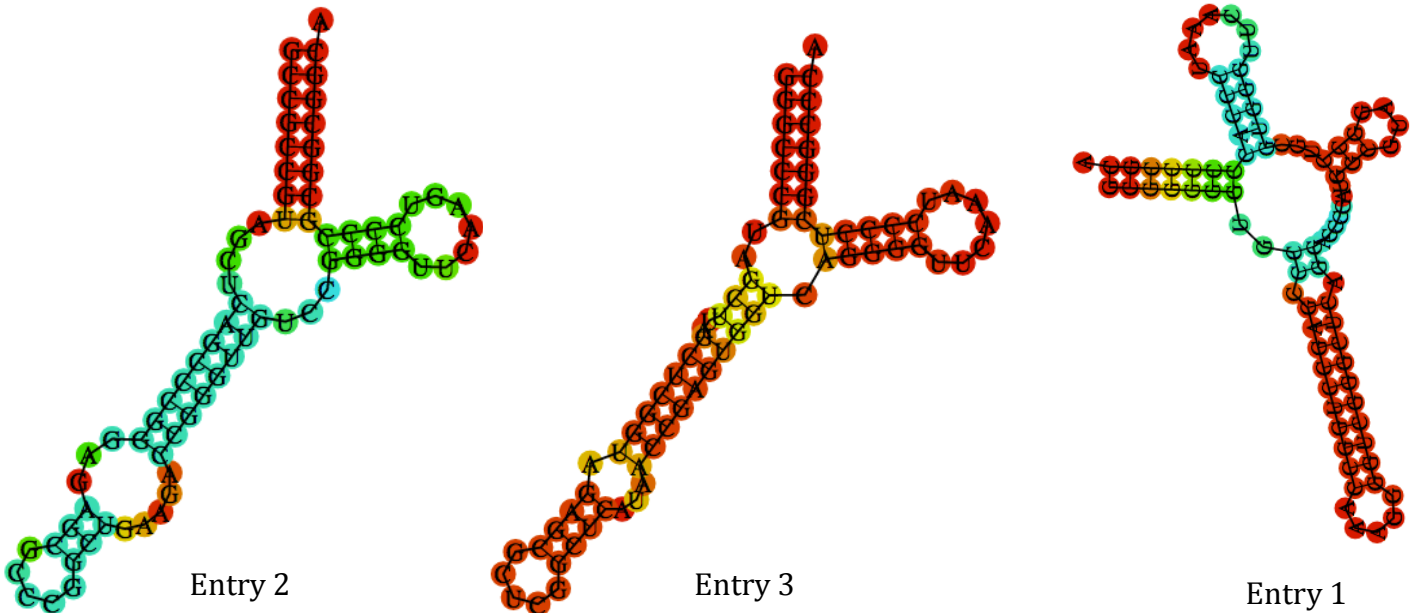
   As previously discussed, high mutual information denotes that the sequences have high dependence on each other. The "top 20" and "top 50" results identify the 20 and 50 most prevalent pairings of nucleotide positions in the sample population. This implies that these are the 20 and 50 most significant interactions between nucleotide positions in the sequence, which can hint at the structure of the physical molecule. For example, in the figure below, one would expect many of the paired nucleotides that are boxed to be in the top 20 - top 50 since a given nucleotide on one side of the double stranded sections should correlate very strongly with its complementary nucleotide in the position directly across from it.

Above are RNA folding structures for entries 1, 2, and 3 from the RNAfold web server — the structures were determined for each individual sequence entry, and not for the entire "tRNA.stock" Stockholm file, due to the large computational load. Nonetheless, these individual structures glean insight into what the complete structure would resemble (i.e. number of loops and stems) when factoring all mutual information scores — in which we could represent the fifty column pairings with the highest mutual information as base pairings.
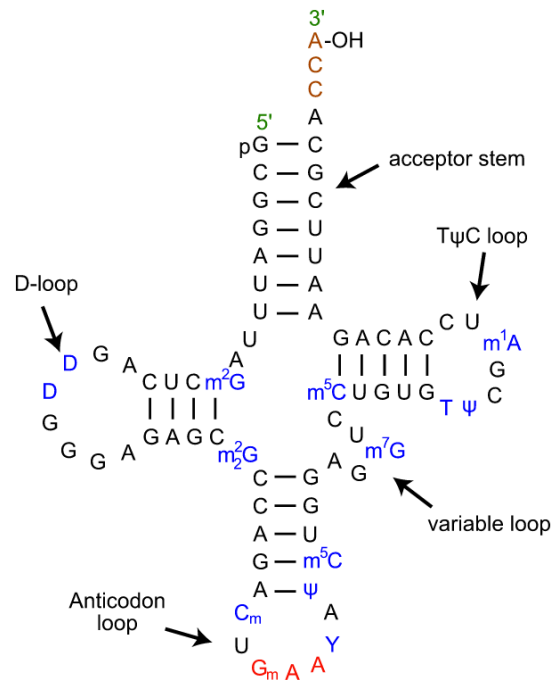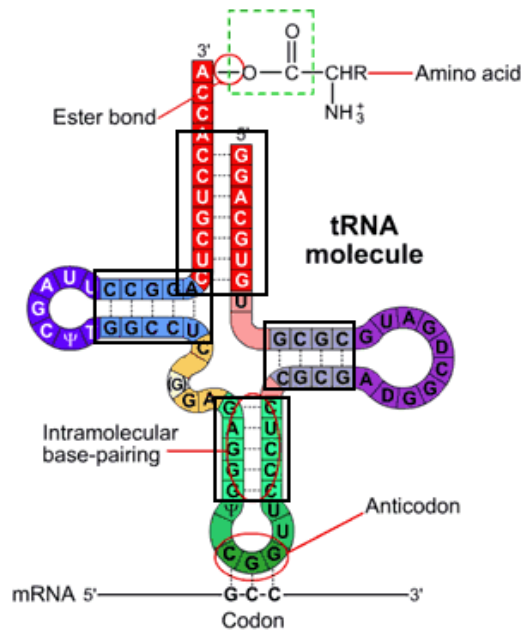

Image Sources: Wiley and Wikipedia (respectively)