William Mitchell

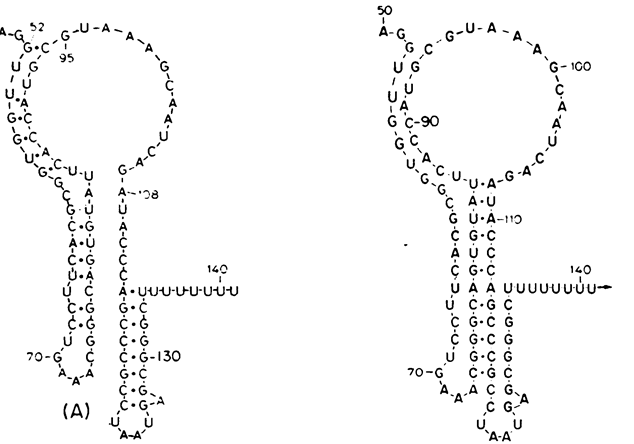
BNFO301

**Tryptophan Regulation by the Formation of Structures in Bacteria**

**Introduction:**

Tryptophan is one of the 20 amino acids that are essential for life. Regulation of the gene that is responsible for the synthesis of Tryptophan is key for living organisms. Over, under, or absence of this amino acid could cause the death of the organism. Bacteria have an interesting way of regulating this particular gene. They do this by forming a structure in the non-coding region upstream and adjacent to the Tryptophan gene. This structure is known as a stem loop.

Stem loop structures form in the RNA during transcription by the nucleotides base pairing with their compliment in the same RNA. The stem loop structure can form more than one structure, depending on what structure forms will allow or terminate transcription. This process is known as attenuation. The structure that forms is based on the concentration of tRNA that encodes for tryptophan in low concentrations one structure is preferred and in high concentrations the other structure is preferred. In the figures below the Figure marked (A) does not allow for transcription of the Tryptophan gene. The other figure will allow the tryptophan gene to be transcribed and translated.



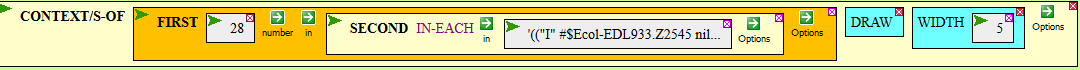
In E.coli the stem loop structure has been well studied and sequenced. This project s purpose is to determine if this structure is conserved in other Enterobacteriacane.

**Methods:**

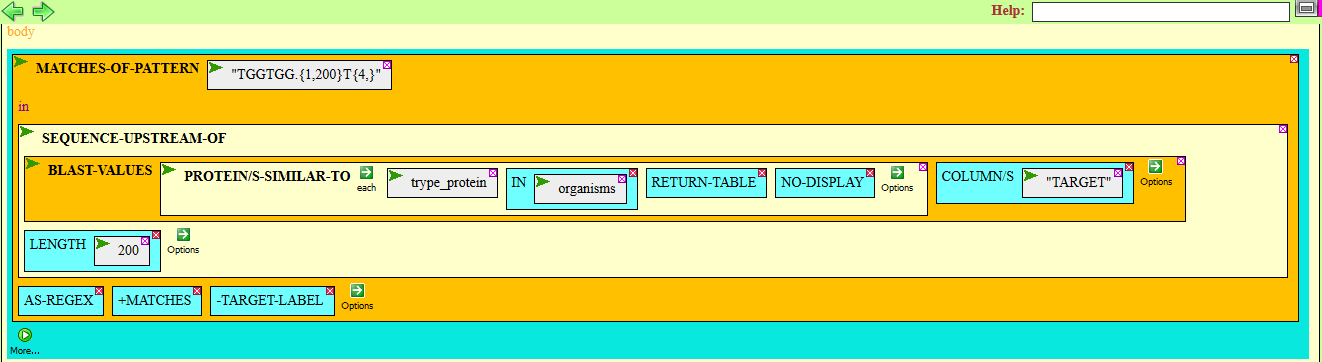
The known sequence AAGTTCACGTAAAAAGGGTATCGACAATGAAAGCAATTTTCGTACTGAAAGGTTGGTGGCGCACTTCCTGAAACGGGCAGTGTATTCACCATGCGTAAAGCAATCAGATACCCAGCCCGCCTAATGAGCGGGCTTTTTTTTGAACAAAATTAGAGAATAACA found in E.coli k 12 the red section indicating the section where the stem loop structure form was blasted in Genbank.



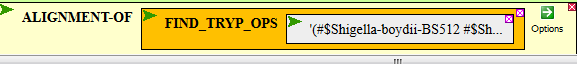
. Then I used Context-of function to find the sequence of the trypE gene which is the first gene in the synthesis of the tryptophan.



Then took that sequence and converted it to a protein sequence. Then I wrote this function to look for the upstream sequences of the trypE gene.



Then I used this function to align the data.

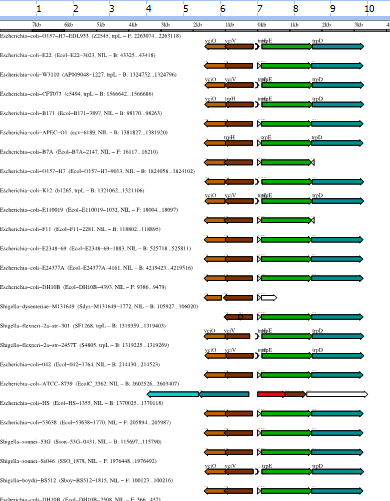


Then I used the Mfold (http://mfold.rna.albany.edu/results/1/15Apr27-01-58-45-fcc91c6cba/)program to fold the sequence to compare structures .

**Results**

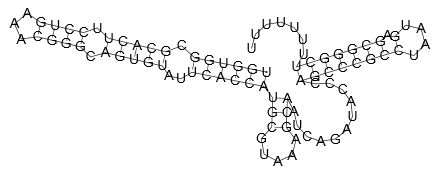
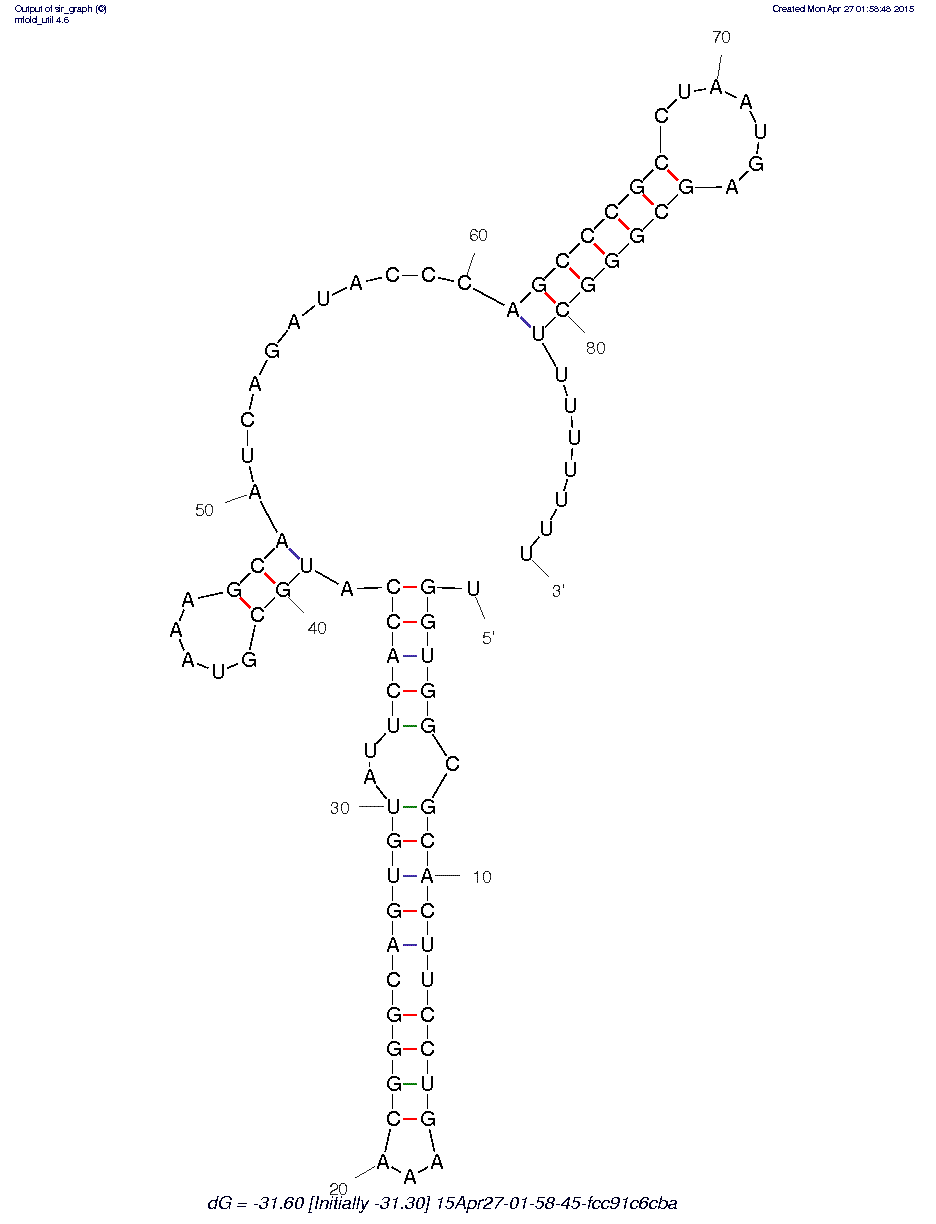
After blasting the known operon sequence n GenBank there were 187 hits that had 100% match. All of the hits were in E.coli and Shigella. There were other hits but the e-value was much to high to consider them.

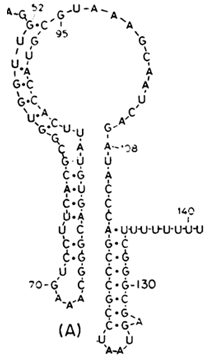
The results of the Context –of function indicate that the trypE gene is adjacent to the upstream sequence of the tryptophan gene in E.coli and Shigella.



By finding the trypE gene was next to the upstream sequence of the tryptophan gene it was then changed to a protein sequence to locate potential upstream sequences in other bacteria.

The result of the Mfold is the structure below on the left. The other structure on the right was done in CyanoBike. The structures are very similar, as you can see both of these programs allow for G-U binding to obtain the lowest energy state of the structures. The structure at the bottom does not allow for G-U binding .





**Works Cited**

Nunvar, Jaroslav, Tereza Huckova, and Irena Licha. “Identification and Characterization of Repetitive Extragenic Palindromes (REP)-Associated Tyrosine Transposases: Implications for REP Evolution and Dynamics in Bacterial Genomes.” BMC Genomics 11 (2010): 44. PMC. Web. 15 Apr. 2015.

Petrillo, Mauro et al. “Stem-Loop Structures in Prokaryotic Genomes.” BMC Genomics 7 (2006): 170. PMC. Web. 15 Apr. 2015.

Di Nocera, Pier Paolo, Eliana De Gregorio, and Francesco Rocco. “GTAG- and CGTC-Tagged Palindromic DNA Repeats in Prokaryotes.” BMC Genomics 14 (2013): 522. PMC. Web. 15 Apr. 2015.

Roberto Kolter, Charles Yanofsky.” Genetic analysis of the tryptophan operon regulatory region using site-directed mutagenesis “ Journal of Molecular Biology Volume 175, Issue 3, 25 May 1984, Pages 299–312

Yanofsky, C et al. “The Complete Nucleotide Sequence of the Tryptophan Operon of Escherichia Coli.” Nucleic Acids Research 9.24 (1981): 6647–6668. Print.