

CISC 471 HW 4 Part 2

Hershil Devnani (20001045)

March 12, 2021

1 Programming

Please refer to the submitted python files and README.md for the solutions to Part 1.

2 Theory

1. Lesson 4.2: How could a bacterium produce a peptide that is not encoded by the bacterium's genome?

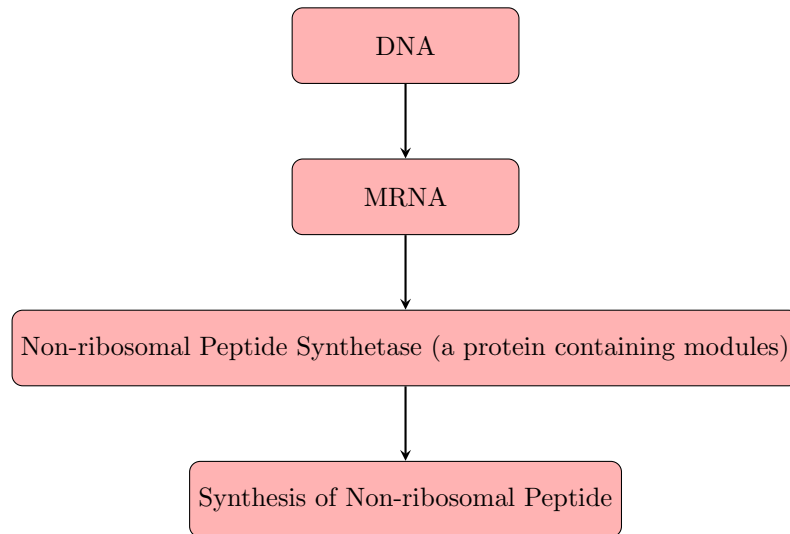
One way that a bacterium produces a peptide that is not directly encoded by the bacterium's genome is by using non-ribosomal peptide synthetase (NRPS). NRPS encode for non-ribosomal peptides (NRP) and an example of some non-ribosomal peptides given by the text is of tyrocidine and gramicidine.

Peptides, such as tyrocidine and gramicidine, are synthesized by means of a large protein known as NRP synthetase (NRPS). This is contrary to both the conventionally known method of RNA translation to peptides via ribosomes and what many scientist originally hypothesised when trying to reverse engineer and find a matching potential 30-mer in the *Bacillus brevis* genome that would encode the Tyrocidine B1 antibiotic.

NRPS are proteins that are transcribed the by MRNA and have an affinity for a particular amino acid. These NRPS contain what are known as modules and they assemble peptides by growing them one amino acid at a time. Essentially, an NRPS is made up of different modules and each module is responsible for adding a single amino acid to the peptide or NRP. The NRPS is responsible for the amino acids that are used to synthesize the resulting peptide, and as a result, the peptides produced via the NRPS are not directly encoded in a given bacterium's genome.

NRP molecules are very versatile since their construction is not limited by the 20 amino acids that are encoded by RNA. Rather, the modules, or NRPS, that code for an NRP can have an affinity for 500 different amino acids. In this way, NRP molecules can be used as antibiotics, used by the bacteria that produces them to communicate with cells and for recycling applications.

As such, the result of the production of an NRP via NRPS renders the reverse sequencing of a peptide to mRNA to DNA to be ineffective since the modular nature of the peptide synthesis would simply not be possible if: the peptide contains amino acids that are not coded by mRNA; or a coding sequence that will not be found in the genome. NRP sequences are also known as template independent, since they follow the sequence of production below:



2. Lesson 4.4: How many subpeptides does a cyclic peptide of length n have?

The number of subpeptides that a cyclic peptide of length n has can be represented by the following:

$$n * (n - 1)$$