Please use Python3 for the following questions involving programming.

1. Analyze the code given in provide\_documentation.py and please add appropriate documentation.
2. Write the contents of function Site\_finder(full\_seq, subseq) in add\_fxn\_contents.py along with a set of unit tests.
3. Write a search function to satisfy a given set of criteria (described below). Please include documentation and test code.

DNA sequences are comprised of A, T, G, and C letters only. DNA letters are called ‘nucleotides’. The letter ‘N’ refers to any nucleotide, and the letter ‘R’ refers to G or A only. (How many unique sequences are described by the ‘GNRA’ motif?) The term ‘complementary’ refers to the pairings A:T and G:C such that the following sequences are complementary:

ATCG

||||

TAGC

A single strand of DNA can fold. If folding occurs, a stable hairpin-like structure can form if enough complementary nucleotides exist in the appropriate locations. For example, the string ‘ATTCGGTAACGCAT’, representing a single strand of DNA, can fold up to form the following hairpin structure:

**T A**

**G A**

G-C

C-G

T C

T-A

seq\_start - A-T - seq\_end

The “loop region” is shown in bold font and the “stem region” is colored grey. The stem region is also referred to as a ‘complementarity region’. Complementary nucleotides in a stem region are said to be ‘paired’ with each other. Pairing interactions between complementary nucleotides are indicated by the ‘-’.

In this task, you are to find and output all substrings of a DNA sequence given the following set of rules. Note that the example 14 NT sequence shown above matches the below criteria.

Rule 1: substring length >= 12 and <=100

Rule 2: A ‘GNRA’ motif is at the center of the subsequence

Rule 3: The subsequence contains at least 4 pairs of complementary nucleotides.

Rule 4: Each half of each complementary pair must be located equidistant from the center of the ‘GNRA’ motif as shown in the diagram of the hairpin structure above.

Rule 5: Both nucleotides of one of the complementary pairs must be a distance of <= 2 from either end of the ‘GNRA’ motif

Rule 6: At least 70% of nucleotides in a complementarity region must be participating in a pairing interaction.

The input sequence for this question is in “question3\_input.txt”.