Abrar Abidi
Fall 2015

Computational Biology Independent Project

Data:

One of the reasons I chose this project was that I would be able to explore unfamiliar territory in my own knowledge computational biology, while largely being able to check the validity and accuracy of my code every step of the way with the help of Rosalind datasets. Thus, though I only leave you with a single data file from Rosalind in the form of “data.txt,” I in fact inputted and tested every previous dataset supplied by Rosalind before the spectral convolution part of the problem. I only moved on from one part of the problem to another when I saw that my code returned the expected outputs for given inputs from Rosalind.

I designed my code to be able to read the data in the format given by Rosalind directly, so that I wouldn’t have to tamper with it manually. There is also a data file called “masses.txt,” which simply supplies my master code with all the masses corresponding to each amino acid. This file is again read directly, and unlike “data.txt” should be left as is, given that it contains standard values.

W 0.2 52 (n) -0.2 (s) 0.2 (st) 0.2 (a)-0.2 (n) -0.2 (d) -0.2 (a) -0.2 (rd) -0.24-0.2 (l) -0.2(u) -0.2 (e) -

left