DDS Negative Selection Prediction

## **Functional Description:**

The Software will allow users to determine the percentage of self-recognizing T-Cells will survive negative selection.

**User Interface**

Command line interface. Simply need to input desired files and maybe a few parameters by the user. Otherwise a basic GUI might be worth designing although likely unnecessary.

**Prioritization**

1. Read Data (tests + code)
2. Read Fasta files about the genes that have been mutated
3. Use prior training data to check for genetic variations and determine new binding coefficients
4. Compute strength of PPI of new genetic variation and current
5. Determine whether the lymphocyte underwent positive or negative selection
6. Export result to output file

**Modules**

1. Detecting variation in the genetic sequence
   1. Fasta file input from user will be analyzed. Files will be of AIRE, CD4, CD8, and RANKFX
   2. Will be comparing to known genetic sequence provided by a database
   3. Blosum Matrix will be utilized to determine if variation is a sign of a completely different protein altogether or if it’s the presence of a new variation.
   4. This will speed up the process of determining KBinding as it permits us to make relative assumptions of Kbinding based on the likelihood of an amino acid substitution and if that AA substituted shares similar properties to the original.
2. Generate randomized set of Lymphocytes with various different combinations of CD8 and CD4 expressions.
3. Calculating Kbinding from AA sequence
   1. Use formulas from online sources to determine strength of binding based on Amino Acid Sequence
      1. Could also find developed packages that perform calculation for me
   2. Strength of binding to AIRE will be calculated based on the residues present on binding domain
4. Express output as various excel files describing combinations of AA sequences between the input files that were first given.

**Goals and Milestones**

1. Write High level Docs
2. Setup Github Repository
3. Research
4. Refine Documentation accordingly
5. Start coding each module in order listed
6. Prototype 1
7. Peer Review
8. Prototype 2

**Modules and Dependencies**

Unsure for the moment