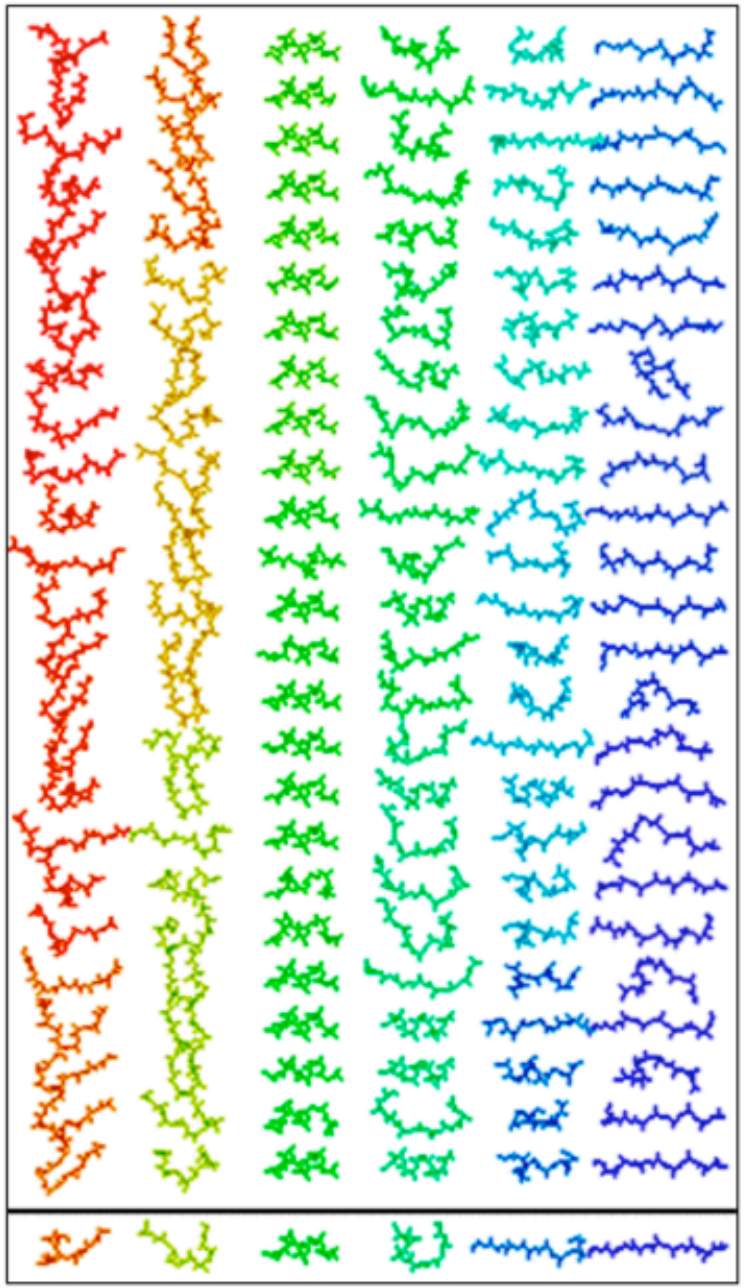
How to Pick Fragments

James Thompson
Baker Lab

Fragment Assembly



Basic Overview

1. Use PSI-BLAST to create an amino acid profile for each position in the query protein.
2. Predict secondary structure of sequence.
3. Pick fragments with similar profiles and secondary structure to the query sequence.

Configuring PSI-BLAST

- PSI-BLAST setup relies on several programs and data files available from the NCBI:
  - blastpgp
  - makemat
  - NR database

Configuring PSI-PRED

- PSI-PRED relies on the following files available from David Jones:
  - psipred
  - psipass2
  - psipred data files

Compiling PSI-PRED

- Compiling PSI-PRED is simple:

  cc -O sspred_avpred.c -lm -o psipred

  cc -O sspred_hmulti.c -lm -o psipass2

  cc -O seq2mtx.c -lm -o seq2mtx

  cc -O pfilt.c -lm -o pfilt

- README and LICENSE files are helpful.

Getting Rosetta Stuff

- Necessary files:
  - make_fragments.pl
  - pNNMAKE.gnu
  - vall.dat.2006-05-05

Configuring make_fragments.pl

- Lines 1-75 contain configuration information.
- Simply edit make_fragments.pl to point to the appropriate locations for all of the files.

```
my $PSIPRED = "/Users/tex/src/psipred/bin/psipred";
my $PSIPASS2 = "/Users/tex/src/psipred/bin/psipass2";
my $PSIPRED_DATA = "/Users/tex/src/psipred/data";
```

Running make_fragments.pl

./make_fragments.pl

-verbose

-nosam

-nojufo

-noprof

1ubi_.fasta

Links

- BLAST - ftp://ftp.ncbi.nih.gov/blast/executables/
- PSIPRED - http://bioinf.cs.ucl.ac.uk/downloads/psipred/
- fragments-support@rosettacommons.org