

**Restoring order in protein structure:
Applications in Tuberculosis**
PRIMES Conference 2013

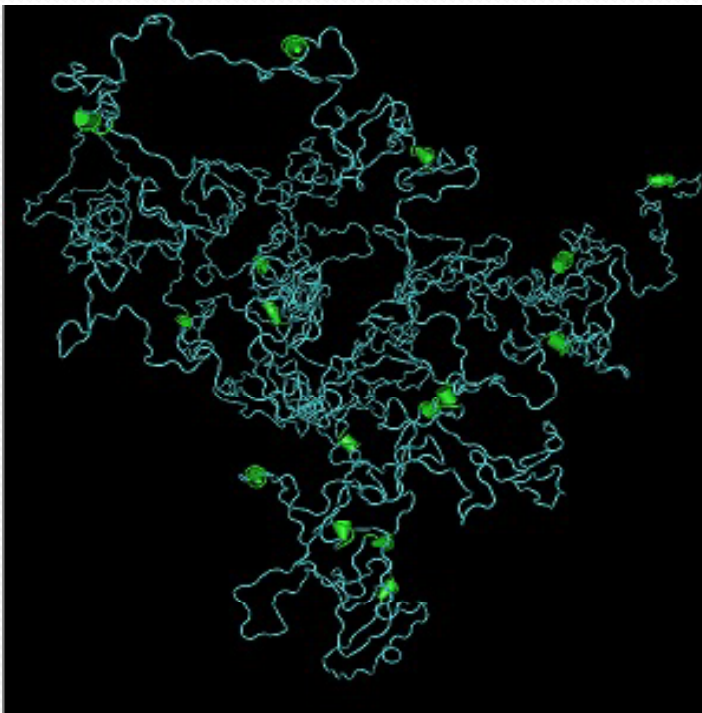
Ben Zheng
Mentor: Dr. Gil Alterovitz

May 19, 2013

What is an intrinsically disordered protein?

Definition

An intrinsically disordered protein (IDP) is a protein that lacks a stable tertiary structure when isolated in vitro.



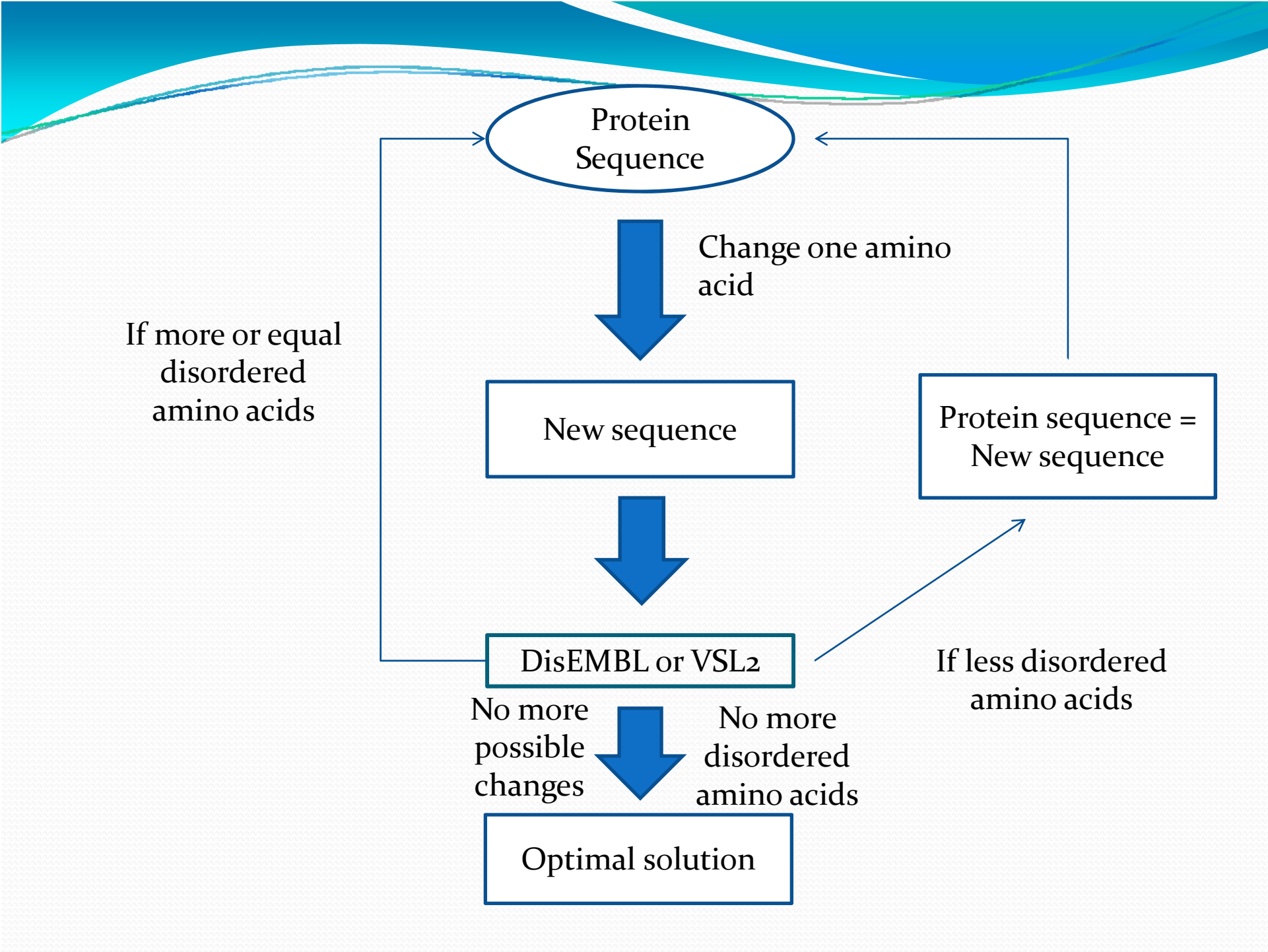
brca1

- Correlation between IDPs and drug resistance in tuberculosis
- Restoring order = crystallization of protein



Hill-Climbing Algorithms

- Optimization technique
- Start with arbitrary solution
- Find single change that will improve the solution most
- Commit and repeat until best solution is found



Protein Sequence

Change one amino acid

New sequence

Protein sequence = New sequence

If more or equal disordered amino acids

DisEMBL or VSL2

If less disordered amino acids

No more possible changes

No more disordered amino acids

Optimal solution



Patterns in changes made

- Changing an E (Glutamic Acid) to an F (Phenylalanine)
- Changing various proteins (L, N, P, R, S, etc.) to a V (Valine)
- Changing a K (Lysine) to an L (Leucine)
- Changing a S (Serine) to a T (Threonine)
- Changing a V (Valine) to a W (Tryptophan)
- Changes concurrent with “order promoting” amino acids found in Campen et al., Prot. Pept. Lett. (2008)

Back to TB



Rv0682 (before processing)

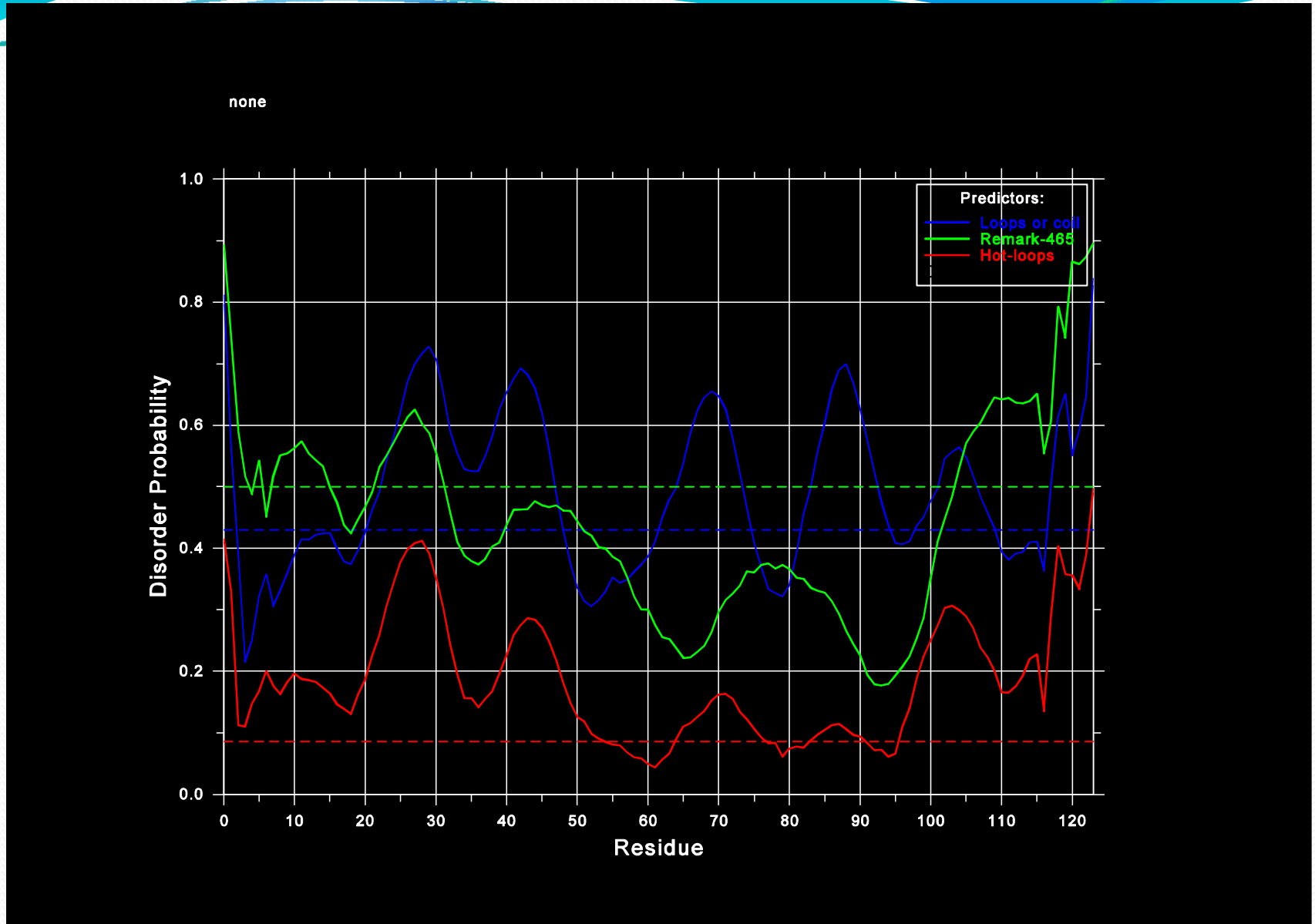
- Classifications we would like to cure: MDR-TB and XDR-TB
- Protein: associated with drug-resistance in TB, not found in PDB
- TBDreaM Database: Rv0682

Collecting Results

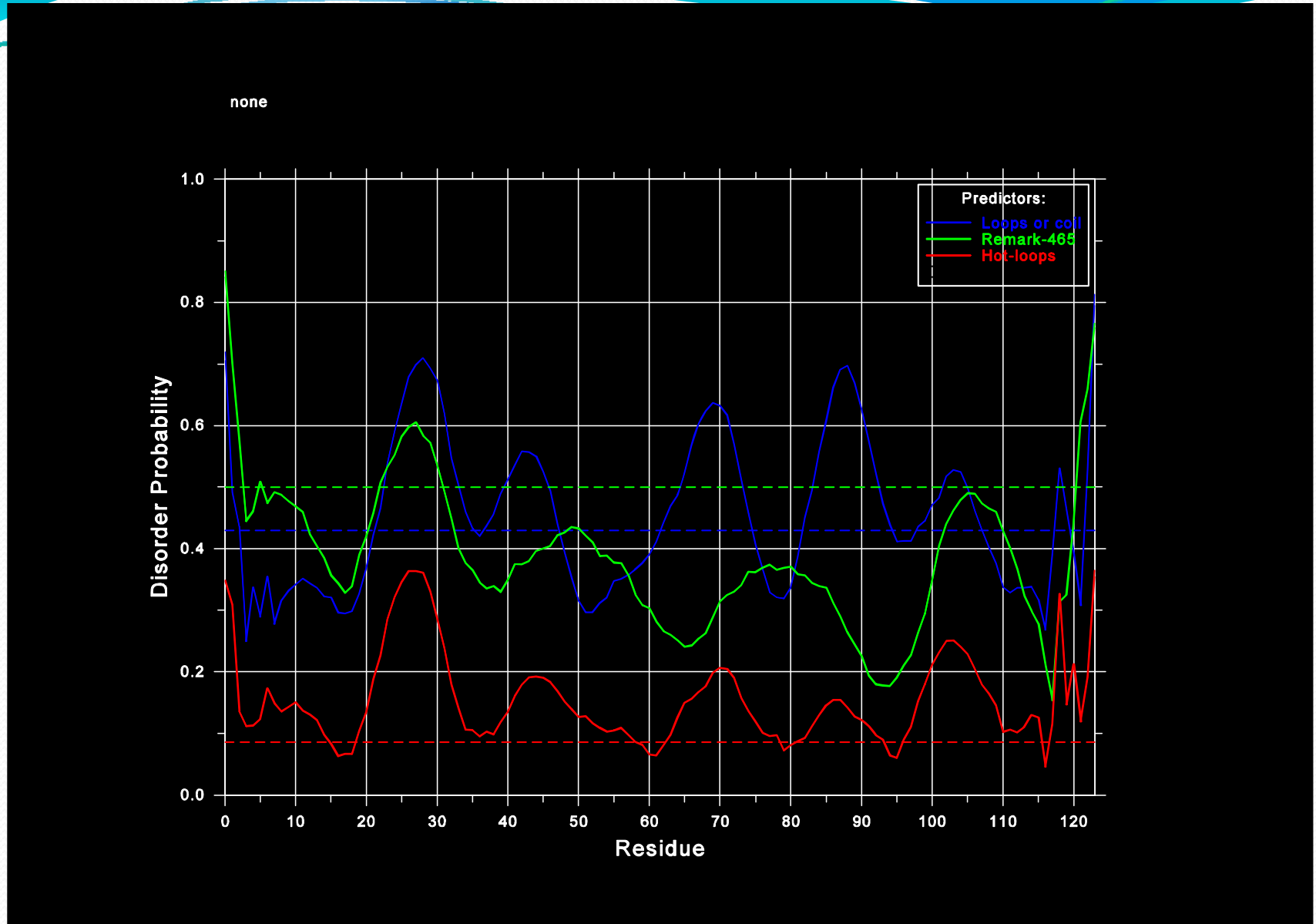


Rv0682 (After processing)

- Trajectory Directed Ensemble Sampling (TraDES) creates and analyzes “unfolding movies”
- Movies can be viewed using Cn3D structure viewer
- Disorder can be measured in terms of probability and radius of gyration

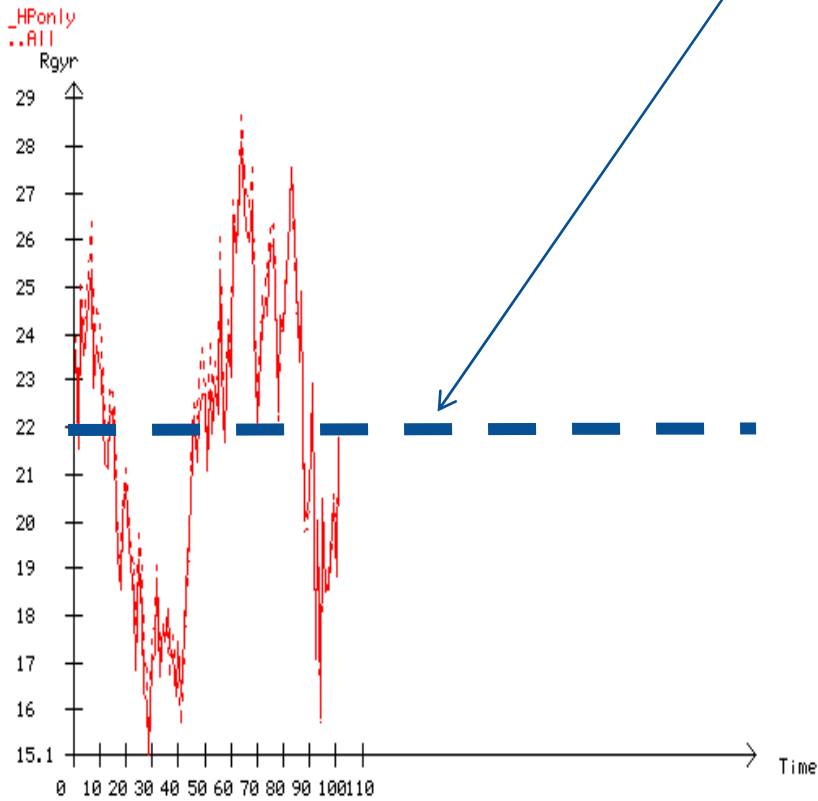


Disorder Probability in DisEMBL for Rv0682 (before processing)

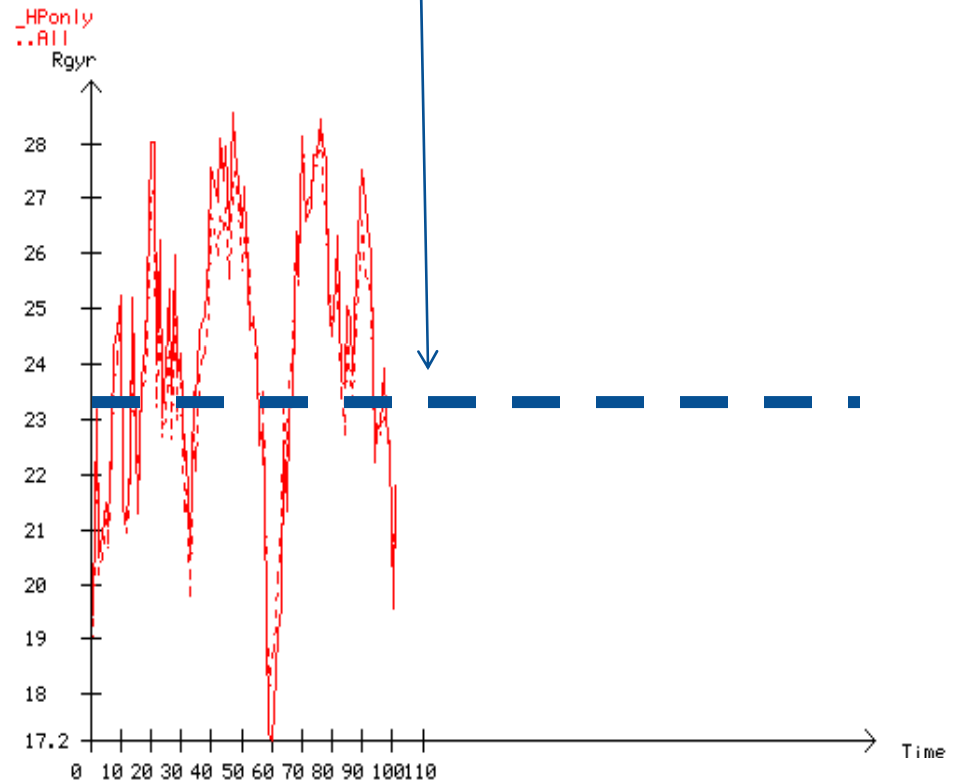


Disorder Probability in DisEMBL for Rv0682 (after processing)

Approximation of average



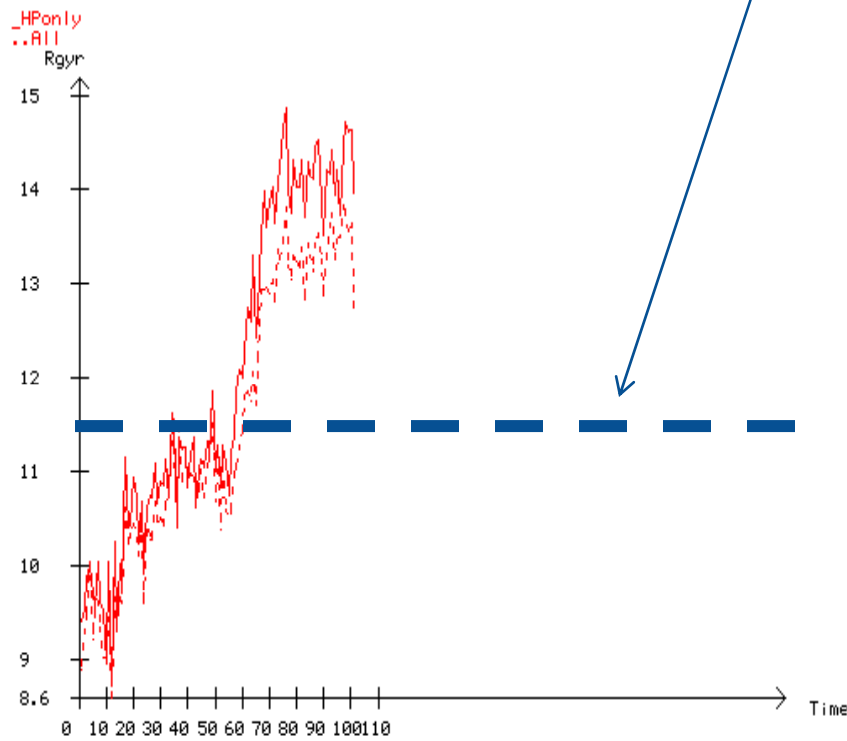
Radius of gyration over time of first 50 amino acids (before processing)



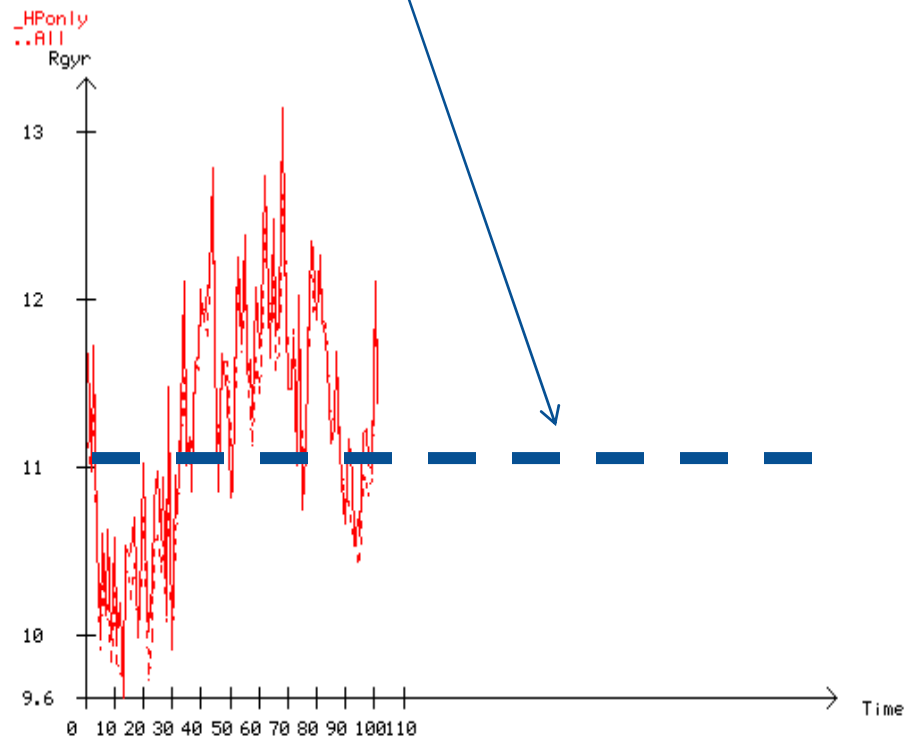
Radius of gyration over time of first 50 amino acids (after processing)



Approximation of Averages



Radius of gyration over time of last 23 amino acids (before processing)



Radius of gyration over time of last 23 amino acids (after processing)

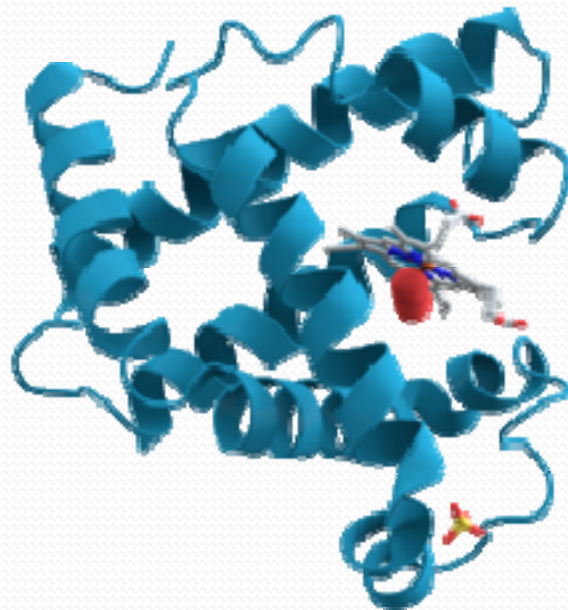


Conclusions

- Probability predictor: amino acids at beginning of protein less affected by the algorithm than those at end of the protein
- Average radii of gyration: increased after processing for first 50, decreased after processing for last 23
- TraDES analysis more accurate for shorter sequences due to less atomic collisions
- Conclusion: predictor reflective of results, disorder decreases around significant changes in probability

Future Work

- Tests with more proteins, including brca1
- Different algorithms
- Real life testing





Acknowledgements

- Thanks to Dr. Gil Alterovitz for mentoring me
- Thanks to Peijin Zhang for explaining concepts to me
- Thanks to my parents for helping me with programming
- Thanks to PRIMES for presenting me with this research opportunity