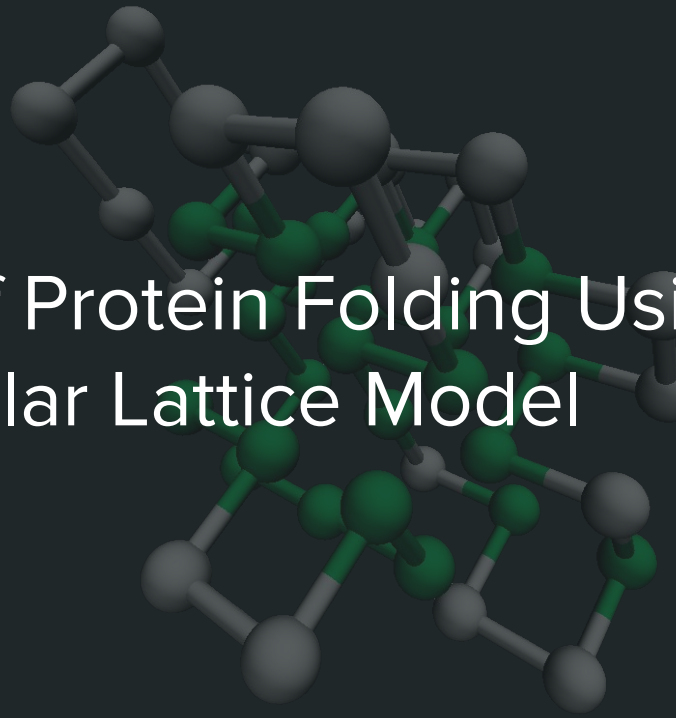


Computer Simulation of Protein Folding Using the 3D Hydrophobic-Polar Lattice Model

Richard Wang

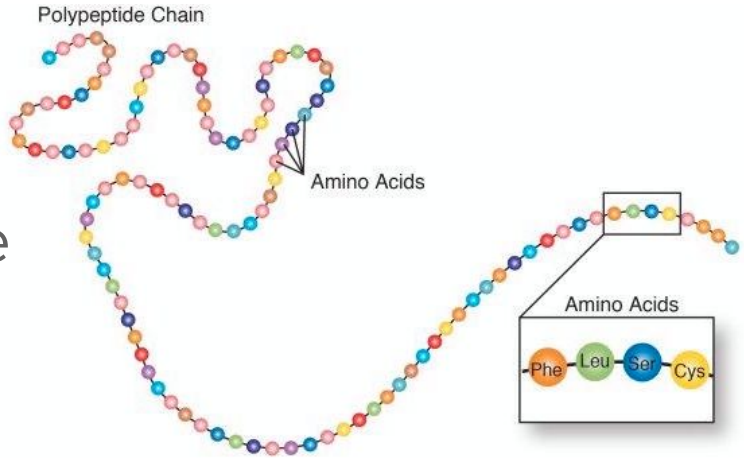
U.S. National Library of Medicine
NCBI

Background

- Proteins
- Protein Folding
- Hydrophobic Interactions
- Goals

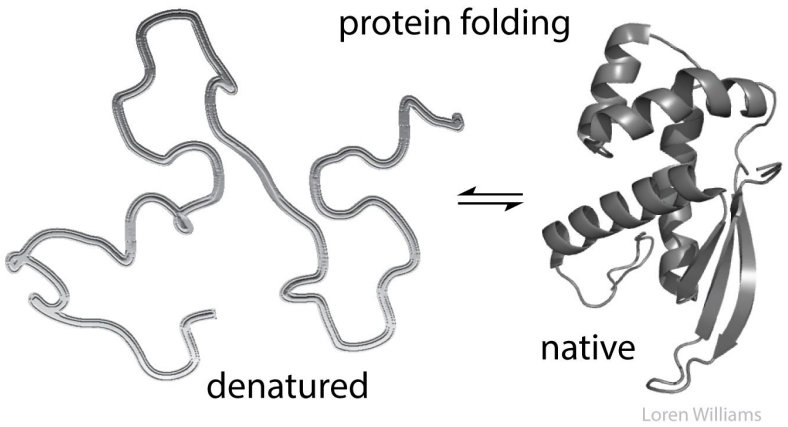
Proteins

- Complex polymer structures
  - Chain of amino acids
  - Folds spontaneously into a unique, stable, functional shape
- Participate in nearly all cellular functions
- Structure = Function



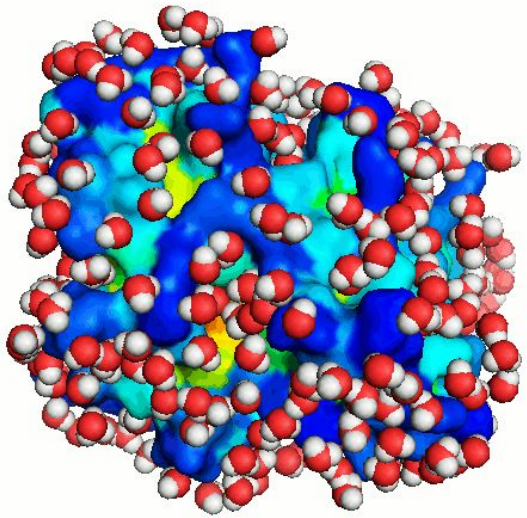
Protein Folding

- Unsolved puzzle
- Many factors involved
  - Most important: hydrophobic interactions
- Computationally expensive to simulate detailed models
  - NP-complete



Hydrophobic Interactions

- Protein enveloped by jiggling water molecules
- Creating hydrophobic cores



Goals

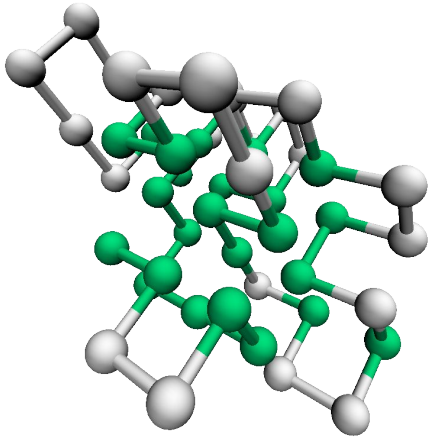
- To develop a protein folding simulation using the HP lattice model and Monte Carlo dynamics
- Compare different types of energy functions

The Program

- Software
- Model
- Features
- Algorithms

Software

- Programmed in Python 2.7.2
- 3D interactive output through Visual Molecular Dynamics 1.9.2



Model

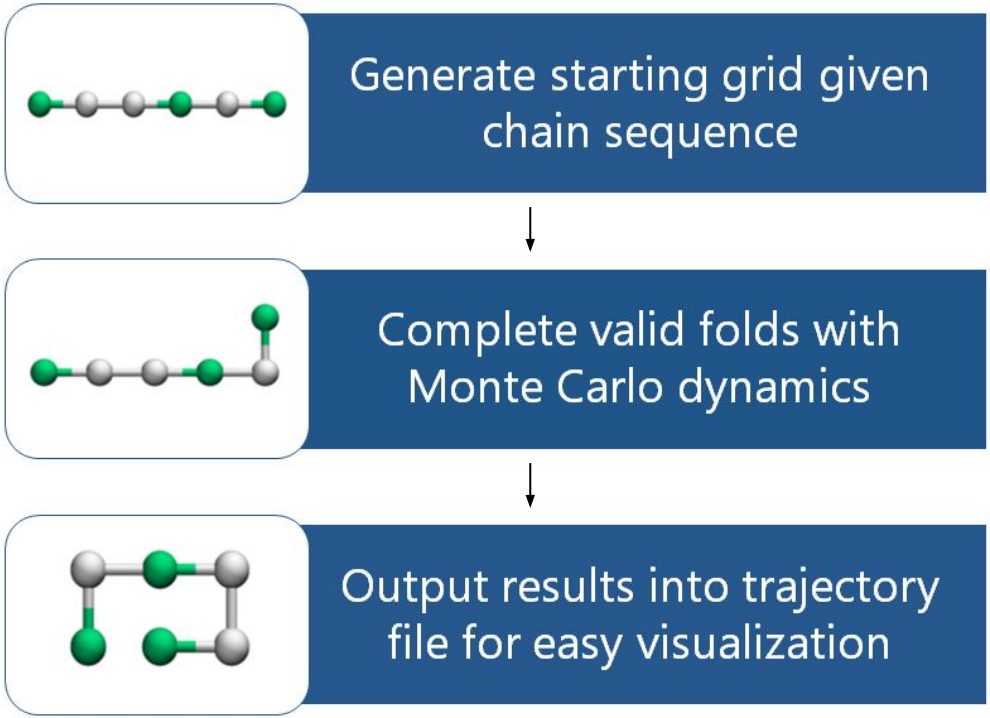
- 3-D Hydrophobic-Polar Lattice Model
- Considers only water interactions
- Amino acids represented by bead of “H” or “P”
  - Hydrophobic: water repelling
  - Polar: water attracting
- 3-D Coordinate grid for placement

Features

- Interactive graphical output
- Multiple energy functions
- Monte Carlo calculations
- Simulated annealing/anti-annealing

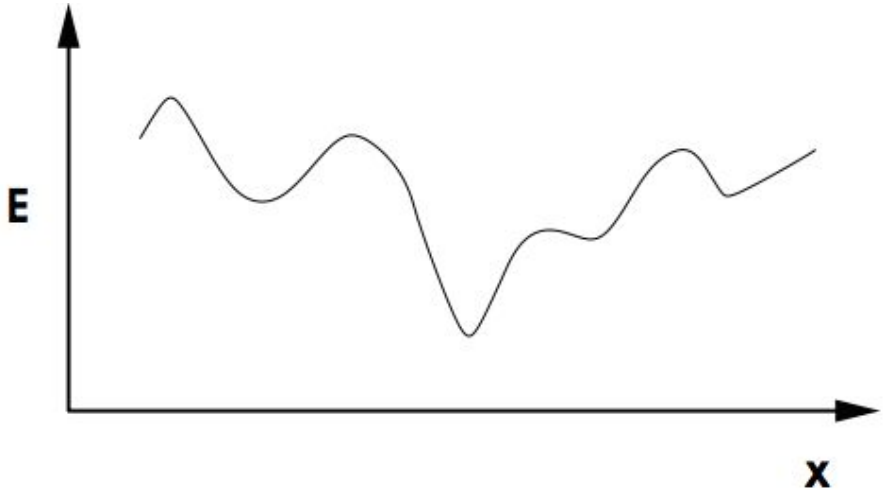
Algorithm

- Find lowest energy structure given amino acid sequence



Monte Carlo Probability Calculations

- Used to simulate water molecule bombardment of proteins



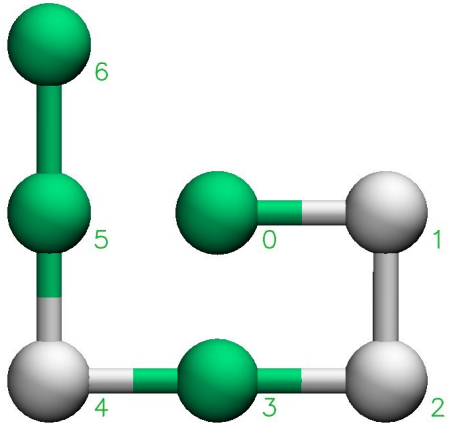
$$p = e^{-\frac{\Delta E}{kT}}$$

$$r = random(0,1) \quad r < p\,?$$

Contact-based Energy Function

- Uses only adjacent H-H contacts when determining energy value
- Inherent flaw

Considered: 5-0, 3-0
Not Considered: 6-0, 6-3, 5-3



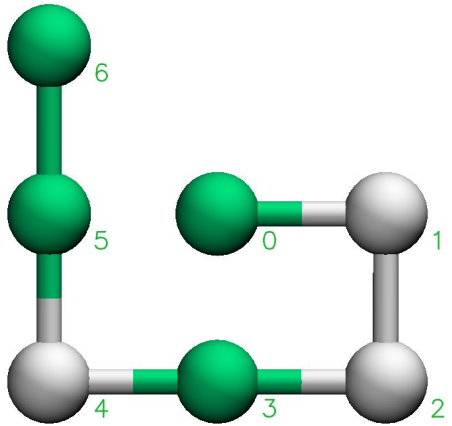
Green: Hydrophobic, White: Polar

Distance-based Energy Function

- Inverse summation of distances between all H beads
- Guaranteed energy change when folding

$$E = -\sum \frac{1}{\sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}}$$

Considered: 5-0, 3-0, 6-0, 5-3, 6-3



Green: Hydrophobic, White: Polar

Simulated Annealing

- Continuously decrease temperature throughout simulation
- Encourages risky folds near beginning of simulation, safe folds near end

$$p = e^{-\frac{\Delta E}{kT}}$$

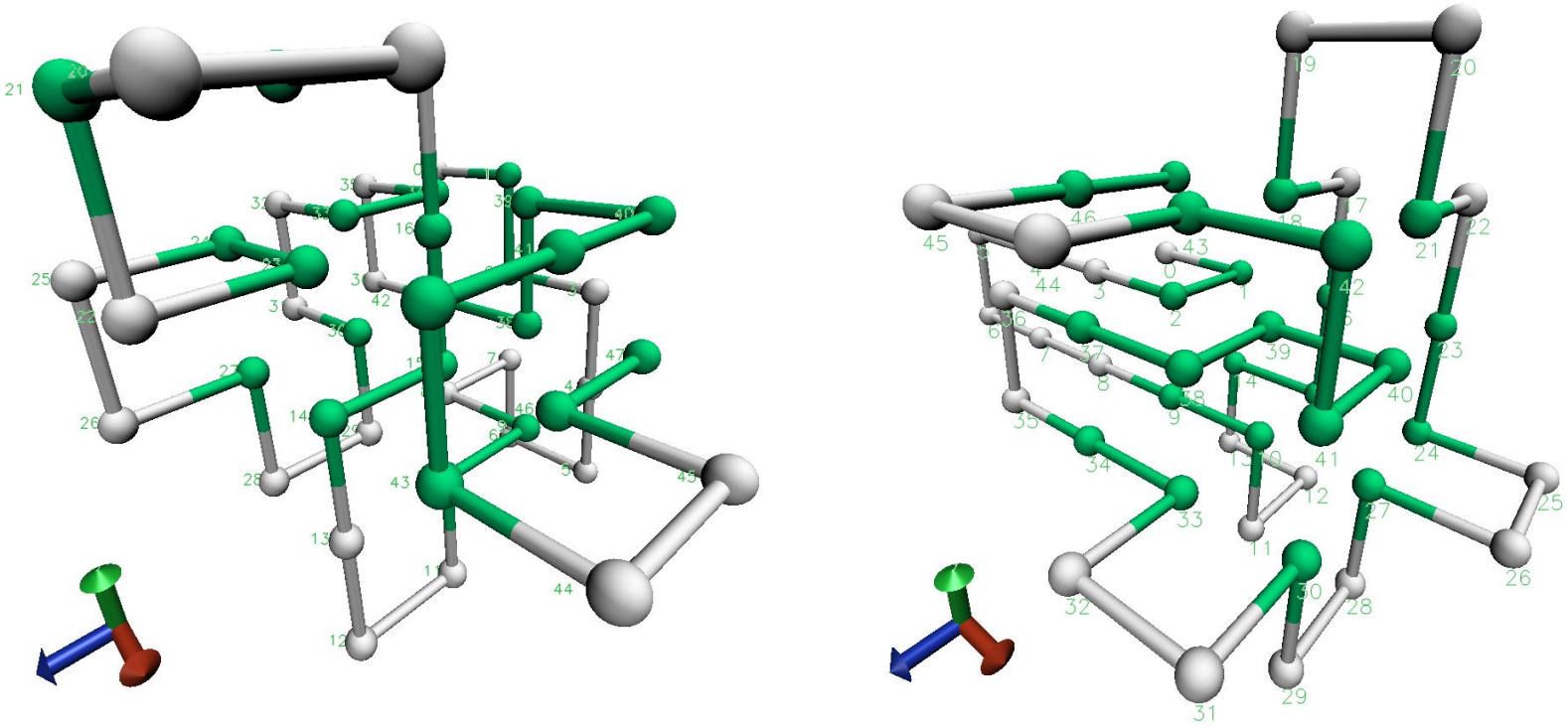
Anti-Annealing

- Used to examine funnel-like structure
- Slowly raise temperature starting from absolute zero and native state
  - Wait until protein breaks out of native structure, then fold with decreasing temperatures

$$p = e^{-\frac{\Delta E}{kT}}$$

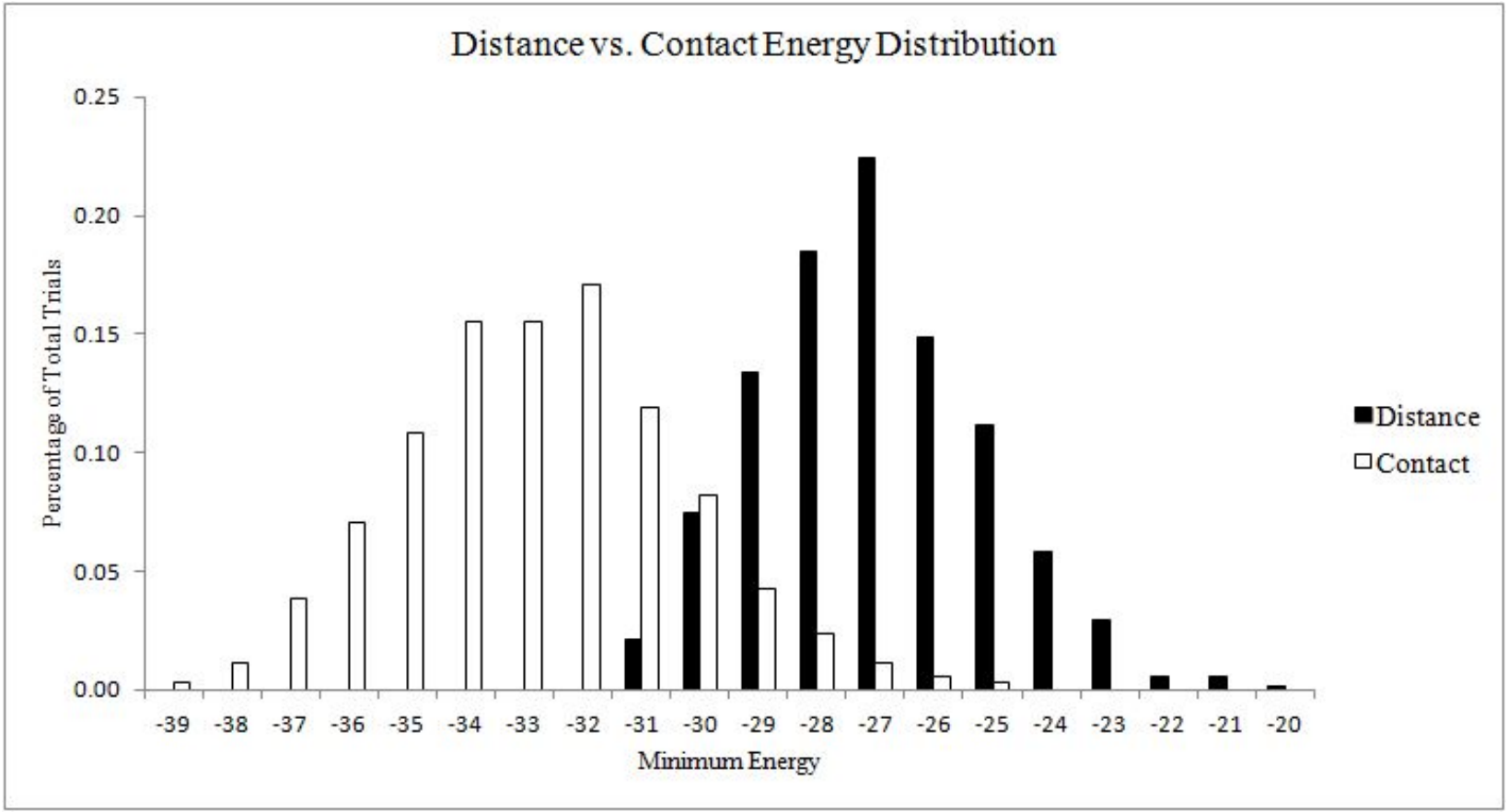
Results

- Data
- Conclusions



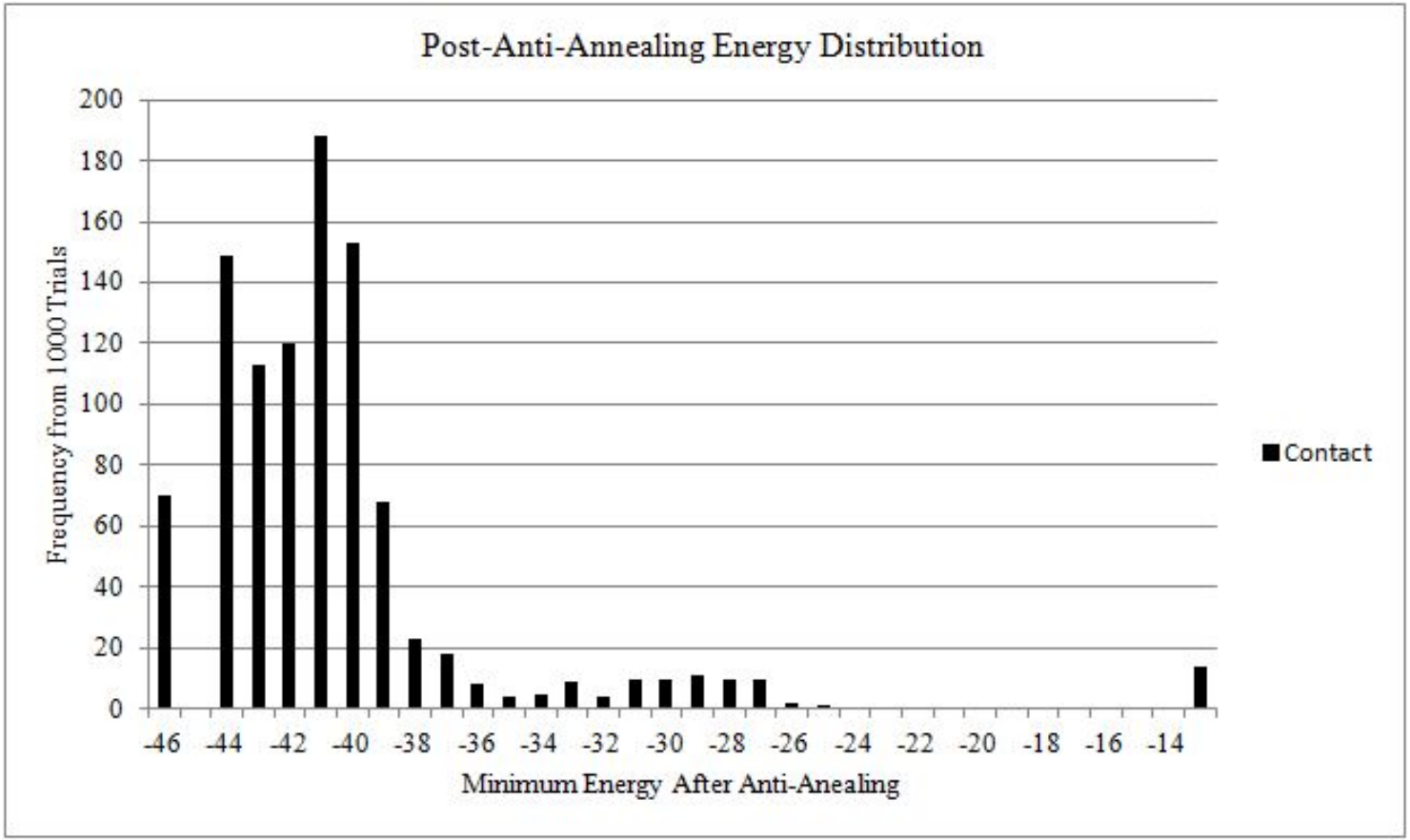
Lowest energy fold found (-41) versus known native state (-46)

PHHPPPPPPHHPPPHHHPHPPHPHHPPHPPHPPHHPPHHHHHHHPPHH



Findings (708 trials):

PHHPPPPPPHHPPPHHHPHPPHPHHPPHPPHPPHHPPHHHHHHHPPHH



Anti-Annealing: Contact

PHHPPPPPPHHPPPHHHPHPPHPHHPPHPPHPPHHPPHHHHHHHPPHH

Conclusions

- Program folds protein chain to near-optimal state in computationally efficient manner
- Contact-based energy function superior
- Jagged energy landscape, with “golf hole” like funnel

The Future

- Future work
- Applications

Future Work

- Improved algorithms
  - Heuristic folding patterns
  - Parallel threading
  - Optimization
- Off-lattice simulations

Applications

- Structure-based drug design
  - Developing proteins with specific functions
- Precursor to development of new algorithms

Acknowledgements

- Dr. Siqian He
- SMCS Teachers
- Friends and Family

Questions?