GIVEN A PROTEIN STRUCTURE GENERATE PROTEIN ENSEMBLE USING BACKRUB METHOD

Presented By:-

Ashish Malgawa 17CS60R81

Suyog Chadawar 17CS60R57

Shubham Yadav 17CS60R47

Guided By:-

Prof. Pralay Mitra

Ms. Abantika Pal

MOTIVATION

- Proteins undergo conformational fluctuations in response to thermal energy, binding events, and mutation.
- Understanding and predicting such excursions around the native state of a protein is a key challenge in computational molecular biology.

OUR ALGORITHM

- Input
  - PDB FILE, Segment Size, number of steps
- Output:
  - Backbone ensemble of the input structure

CONT..

- Randomly select an angle alpha from the range of permitted angles.
- Randomly select a segment from all allowed segments with the given segment size.
- Apply BackRub move.
- Use SCWRL to fit side chains.
- Use FOLDX to calculate energy of the newly generated protein conformation.
- Select the newly generated protein conformation with the probability given by Metropolis Criterion.
- Repeat all these steps until the terminating condition is satisfied.

BACKRUB MOVE

- Randomly select a segment from all allowed segments with the given segment size.

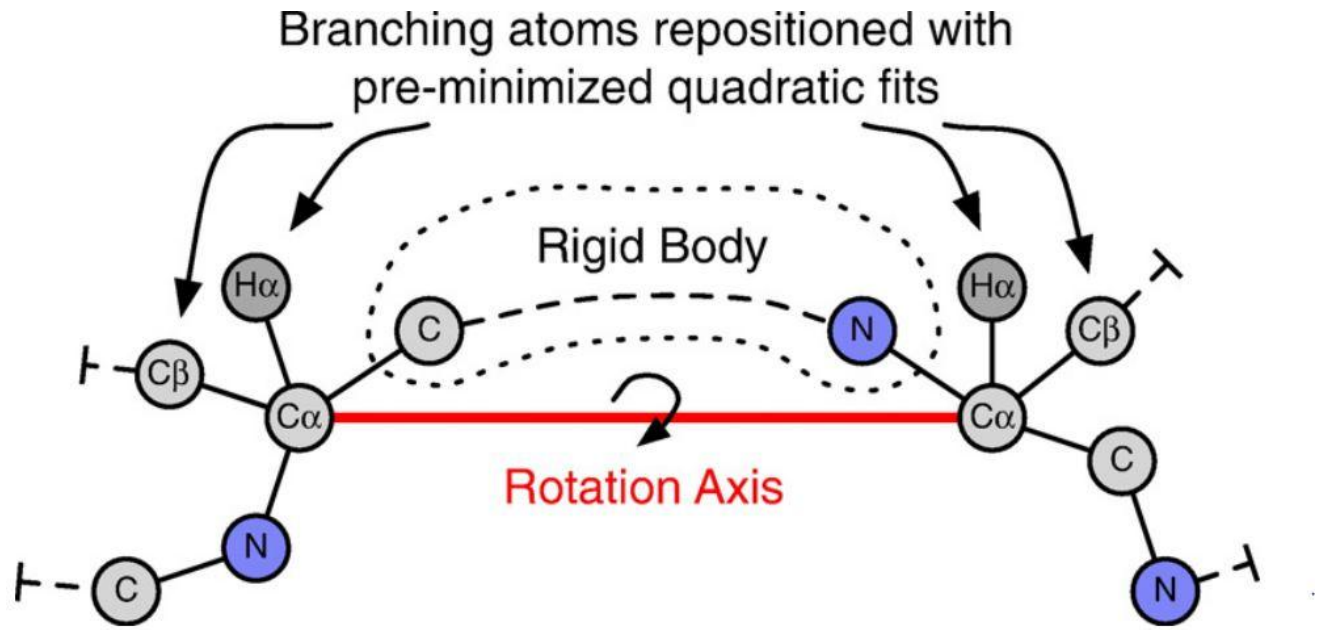
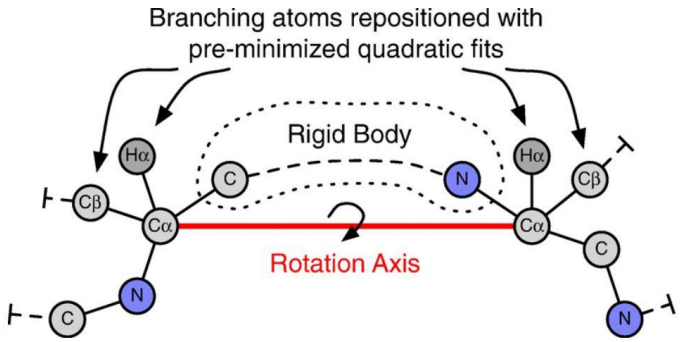


Fig from Smith, Colin A., and Tanja Kortemme. "Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction." Journal of molecular biology 380.4 (2008): 742-756.

BACKRUB MOVE

- Delete coordinate information of all side chains in the selected segment.
- Randomly select angle within the specified range.
- Axis vector is calculated by considering the boundary Cα atoms.
- All the atoms between boundary Cα atoms are rotated about this axis by selected angle.
- Using selected angle, calculated axis and coordinates of boundary Cα atoms, we can calculate rotation matrix.



BACKRUB MOVE

$$\begin{bmatrix} \frac{u^2+(v^2+w^2)\cos\theta}{L} & \frac{uv(1-\cos\theta)-w\sqrt{L}\sin\theta}{L} & \frac{uw(1-\cos\theta)+v\sqrt{L}\sin\theta}{L} & \frac{\Big(a(v^2+w^2)-u(bv+cw)\Big)(1-\cos\theta)+(bw-cv)\sqrt{L}\sin\theta}{L} \\ \frac{uv(1-\cos\theta)+w\sqrt{L}\sin\theta}{L} & \frac{v^2+(u^2+w^2)\cos\theta}{L} & \frac{vw(1-\cos\theta)-u\sqrt{L}\sin\theta}{L} & \frac{\Big(b(u^2+w^2)-v(au+cw)\Big)(1-\cos\theta)+(cu-aw)\sqrt{L}\sin\theta}{L} \\ \frac{uw(1-\cos\theta)-v\sqrt{L}\sin\theta}{L} & \frac{vw(1-\cos\theta)+u\sqrt{L}\sin\theta}{L} & \frac{w^2+(u^2+v^2)\cos\theta}{L} & \frac{\Big(c(u^2+v^2)-w(au+bv)\Big)(1-\cos\theta)+(av-bu)\sqrt{L}\sin\theta}{L} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

BACKRUB MOVE

$$\begin{bmatrix} \frac{\Big(a(v^2+w^2)-u(bv+cw-ux-vy-wz)\Big)(1-\cos\theta)+Lx\cos\theta+\sqrt{L}(-cv+bw-wy+vz)\sin\theta}{L} \\ \frac{\Big(b(u^2+w^2)-v(au+cw-ux-vy-wz)\Big)(1-\cos\theta)+Ly\cos\theta+\sqrt{L}(cu-aw+wx-uz)\sin\theta}{L} \\ \frac{\Big(c(u^2+v^2)-w(au+bv-ux-vy-wz)\Big)(1-\cos\theta)+Lz\cos\theta+\sqrt{L}(-bu+av-vx+uy)\sin\theta}{L} \end{bmatrix}$$

BACKRUB MOVE

- Using this rotation matrix we calculate new coordinates of backbone atoms in the selected segment.
- These new coordinates are reflected back in a new PDB file.
- This PDB file is given as input to SCWRL.

SIDE CHAIN FITTING

```
Original PDB
ATOM    573  N   THR A  76      13.900  49.836  47.615  1.00 21.77           N
ATOM    574  CA  THR A  76      12.588  49.229  47.853  1.00 21.77           C
ATOM    575  C   THR A  76      12.644  47.789  48.354  1.00 21.77           C
ATOM    576  O   THR A  76      11.639  47.333  48.980  1.00 21.77           O
ATOM    577  CB  THR A  76      11.689  49.261  46.557  1.00 22.27           C
ATOM    578  OG1 THR A  76      12.474  48.582  45.540  1.00 22.27           O
ATOM    579  CG2 THR A  76      11.266  50.631  46.003  1.00 22.27           C
ATOM    580  N   GLY A  77      13.749  47.127  48.065  1.00 19.48           N
ATOM    581  CA  GLY A  77      13.830  45.709  48.502  1.00 19.48           C
ATOM    582  C   GLY A  77      13.131  44.862  47.457  1.00 19.48           C
ATOM    583  O   GLY A  77      13.013  43.629  47.598  1.00 19.48           O
ATOM    584  N   ALA A  78      12.659  45.451  46.353  1.00 18.08           N
ATOM    585  CA  ALA A  78      11.985  44.712  45.281  1.00 18.08           C
ATOM    586  C   ALA A  78      13.056  44.023  44.418  1.00 18.08           C
ATOM    587  O   ALA A  78      14.245  44.474  44.446  1.00 18.08           O
ATOM    588  CB  ALA A  78      11.073  45.628  44.448  1.00 18.40           C
```

SIDE CHAIN FITTING

```
After Backrub

ATOM    573  N   THR A  76      13.968  49.437   48.3  1.00 21.77           N
ATOM    574  CA  THR A  76      12.588  49.229  47.853  1.00 21.77           C
ATOM    575  C   THR A  76      11.974  47.907  48.304  1.00 21.77           C
ATOM    576  O   THR A  76      10.709  47.818  48.347  1.00 21.77           O
ATOM    580  N   GLY A  77      12.834  46.948  48.594  1.00 19.48           N
ATOM    581  CA  GLY A  77      12.277  45.632  49.001  1.00 19.48           C
ATOM    582  C   GLY A  77      11.946  44.859   47.74  1.00 19.48           C
ATOM    583  O   GLY A  77      11.419   43.73  47.794  1.00 19.48           O
ATOM    584  N   ALA A  78      12.252  45.392  46.552  1.00 18.08           N
ATOM    585  CA  ALA A  78      11.985  44.712  45.281  1.00 18.08           C
ATOM    586  C   ALA A  78      13.081  43.657  45.055  1.00 18.08           C
ATOM    587  O   ALA A  78      14.187  43.784  45.671  1.00 18.08           O
```

SIDE CHAIN FITTING

After Side Chain Fitting

```
ATOM    701  N   THR A  76      13.968  49.437  48.300  1.00           N
ATOM    702  CA  THR A  76      12.588  49.229  47.853  1.00           C
ATOM    703  C   THR A  76      11.974  47.907  48.304  1.00           C
ATOM    704  O   THR A  76      10.709  47.818  48.347  1.00           O
ATOM    705  H   THR A  76      14.514  48.731  48.418  1.00           H
ATOM    706  CB  THR A  76      12.476  49.295  46.318  1.00           C
ATOM    707  HG1 THR A  76      12.404  51.169  46.199  1.00           H
ATOM    708  OG1 THR A  76      12.891  50.588  45.861  1.00           O
ATOM    709  CG2 THR A  76      11.038  49.061  45.879  1.00           C
ATOM    710  N   GLY A  77      12.834  46.948  48.594  1.00           N
ATOM    711  CA  GLY A  77      12.277  45.632  49.001  1.00           C
ATOM    712  C   GLY A  77      11.946  44.859  47.740  1.00           C
ATOM    713  O   GLY A  77      11.419  43.730  47.794  1.00           O
ATOM    714  H   GLY A  77      13.725  47.066  48.553  1.00           H
ATOM    715  N   ALA A  78      12.252  45.392  46.552  1.00           N
ATOM    716  CA  ALA A  78      11.985  44.712  45.281  1.00           C
ATOM    717  C   ALA A  78      13.081  43.657  45.055  1.00           C
ATOM    718  O   ALA A  78      14.187  43.784  45.671  1.00           O
ATOM    719  H   ALA A  78      12.638  46.205  46.555  1.00           H
ATOM    720  CB  ALA A  78      11.934  45.718  44.142  1.00           C
```

ENERGY CALCULATION

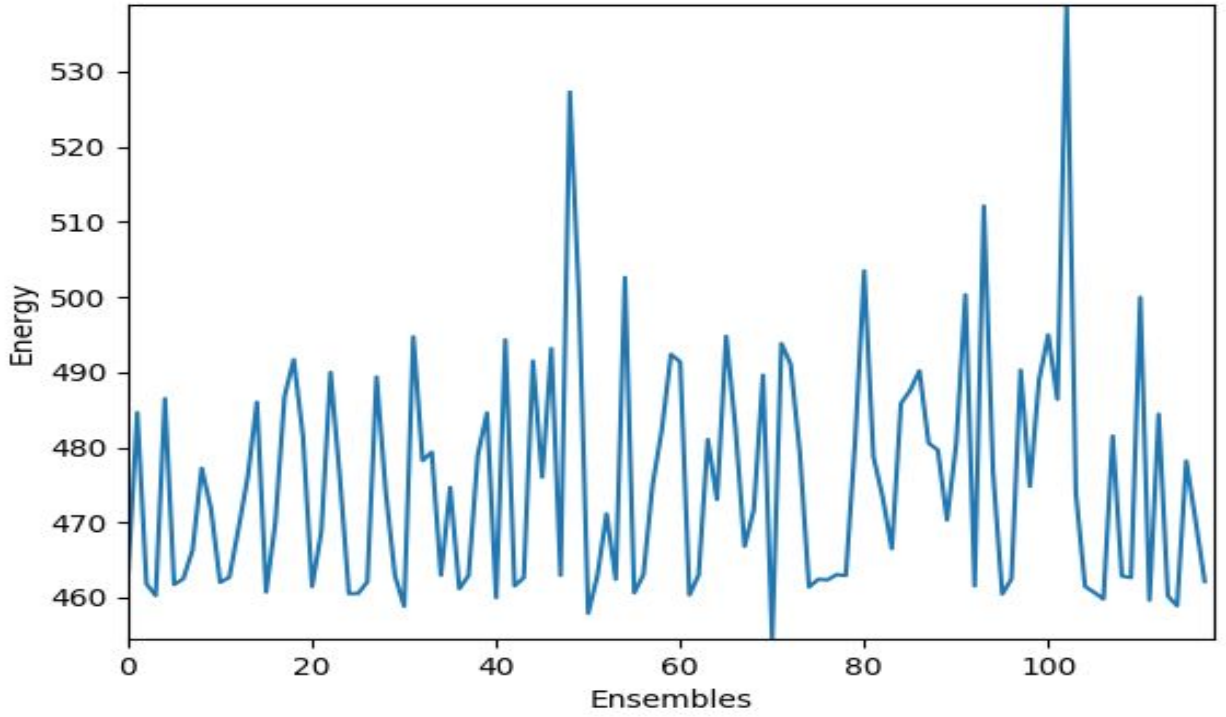
```
   ***                                     ***
   ***     Jesper Borg, Frederic Rousseau  ***
   ***    Joost Schymkowitz, Luis Serrano  ***
   ***    Peter Vanhee, Erik Verschueren   ***
   ***     Lies Baeten, Javier Delgado     ***
   ***       and Francois Stricher         ***
   *** and any other of the 9! permutations ***
   ***   based on an original concept by   ***
   ***   Raphael Guerois and Luis Serrano  ***
   *****************************************

Stability >>>

1 models read: 6apu.pdb
./6apu_0_ST.fxout
Output Directory: ./
Output File: 6apu_0_ST.fxout
Configuration File: config_6apu_0_ST.cfg
number of incomplete residues = 1
completing sidechain atoms of LYSA197

BackHbond       =              -248.52
SideHbond       =              -53.18
Energy_VdW      =              -439.18
Electro         =              -26.69
Energy_SolvP    =              613.50
Energy_SolvH    =              -592.03
Energy_vdwclash =              41.08
energy_torsion  =              17.22
backbone_vdwclash=             223.06
Entropy_sidec   =              212.41
Entropy_mainc   =              526.01
water bonds     =              0.00
helix dipole    =              1.41
loop_entropy    =              0.00
cis_bond        =              2.46
disulfide       =              0.00
kn electrostatic=              -0.54
partial covalent interactions = -3.38
Energy_Ionisation =            0.38
Entropy Complex =              0.00
-----------------------------------------------------------
Total          =                                  50.95

FINISHING STABILITY ANALYSIS OPTION
```

METROPOLIS CRITERION

$$Pr[c \to c'] := \begin{cases} 1 & \text{if } \Delta E \leq 0, \\ e^{\frac{-\Delta E}{T}} & \text{otherwise.} \end{cases}$$

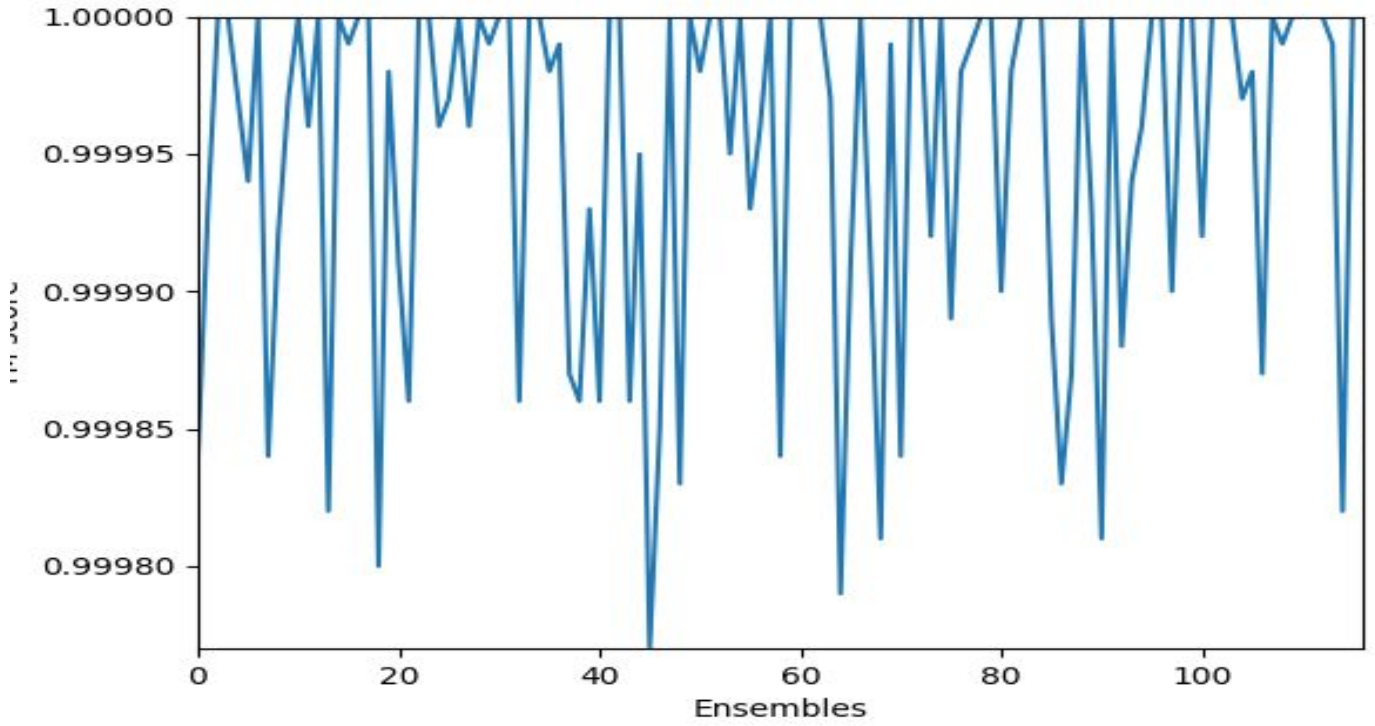
RESULTS for 1cwu (Version 1.0)

Total of 117 ensembles were generated after running the program for 1500 iterations.



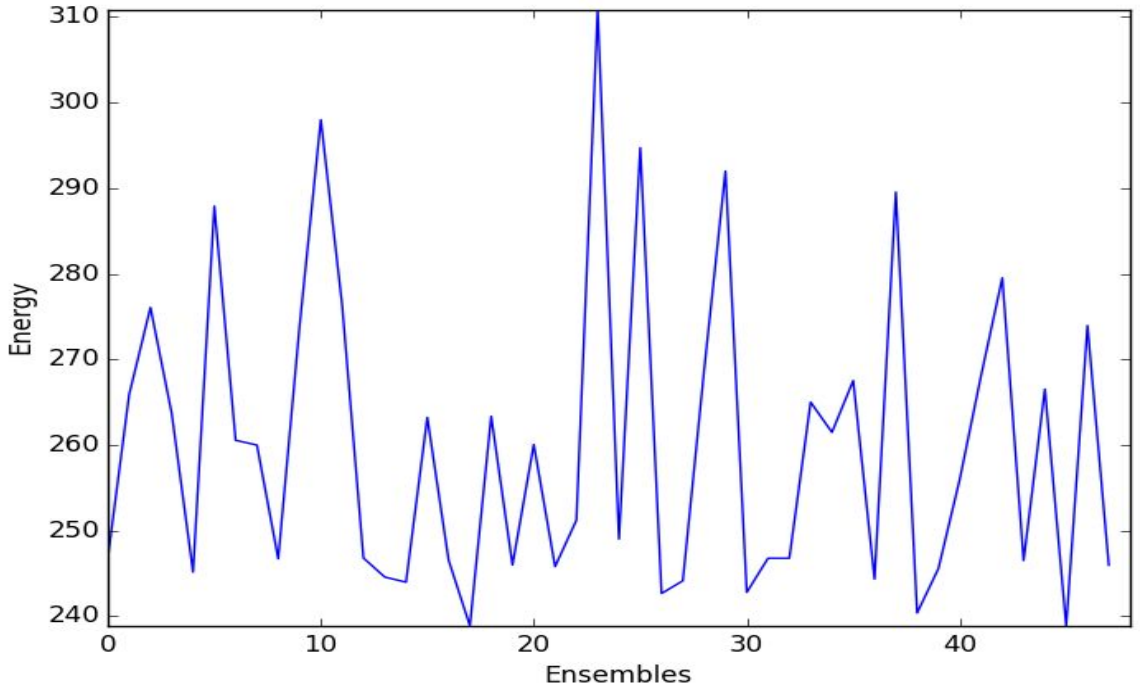
TM-score RESULT (Version 1.0)

TM-score results of the ensembles (in previous slide) when compared with original protein.



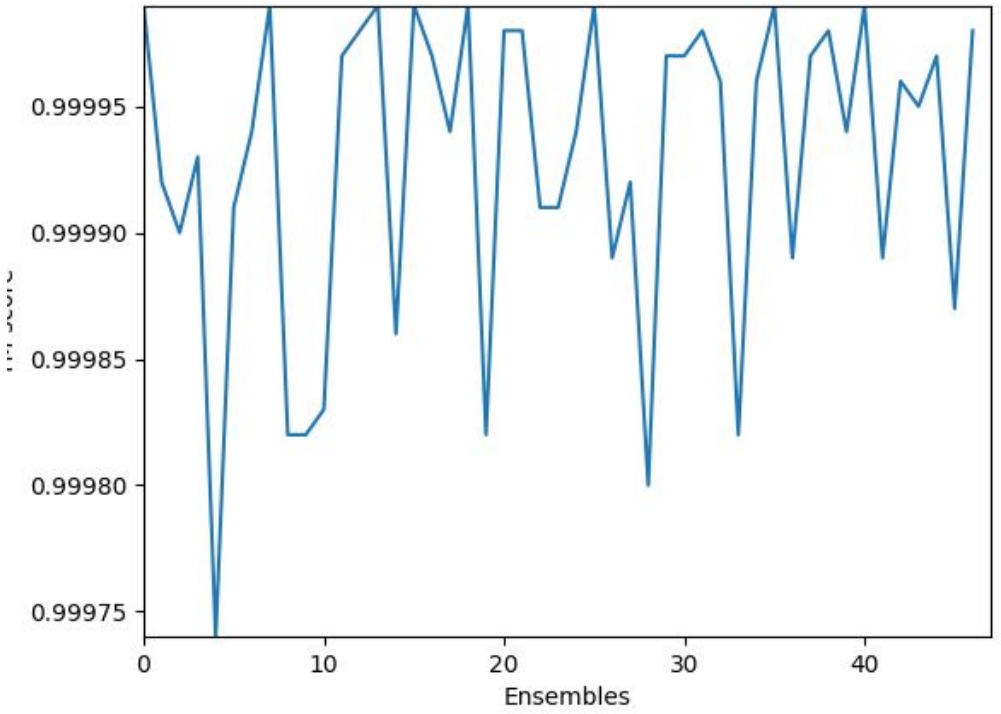
RESULTS for 1cvl (Version 1.0)

Total of 47 ensembles were generated after running the program for 500 iterations.



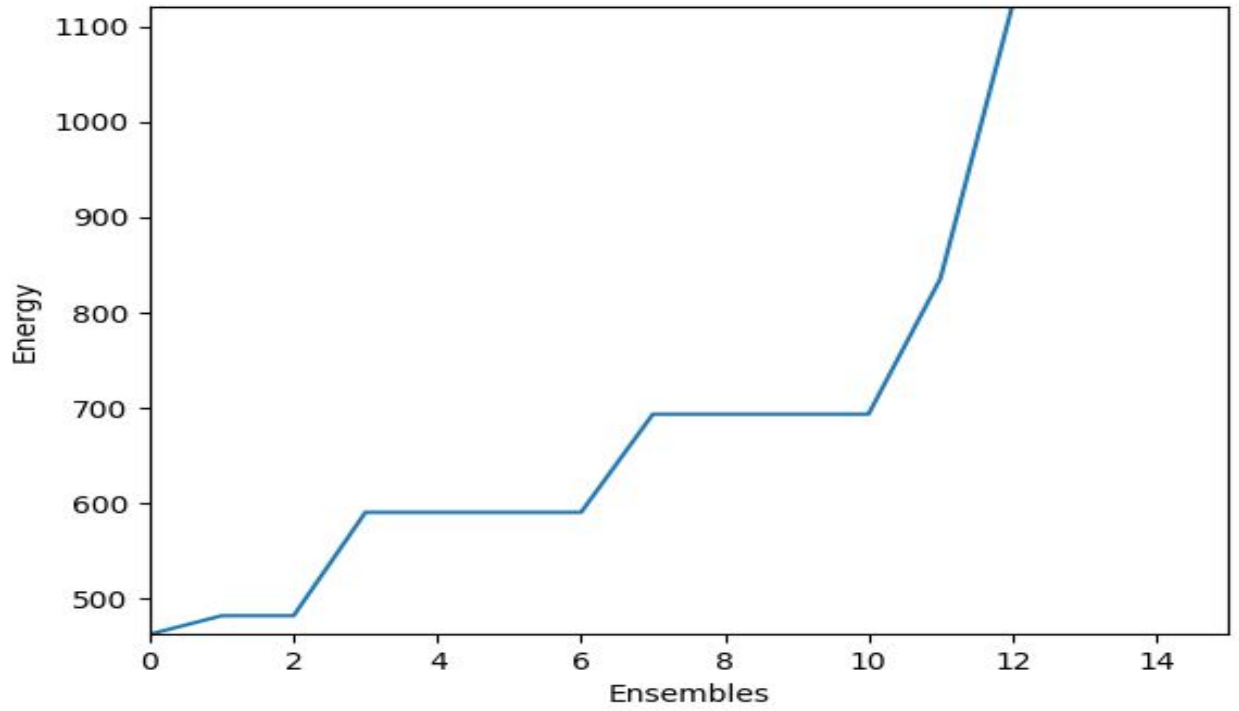
TM-score RESULT (Version 1.0)

TM-score results of the ensembles (in previous slide) when compared with original protein.



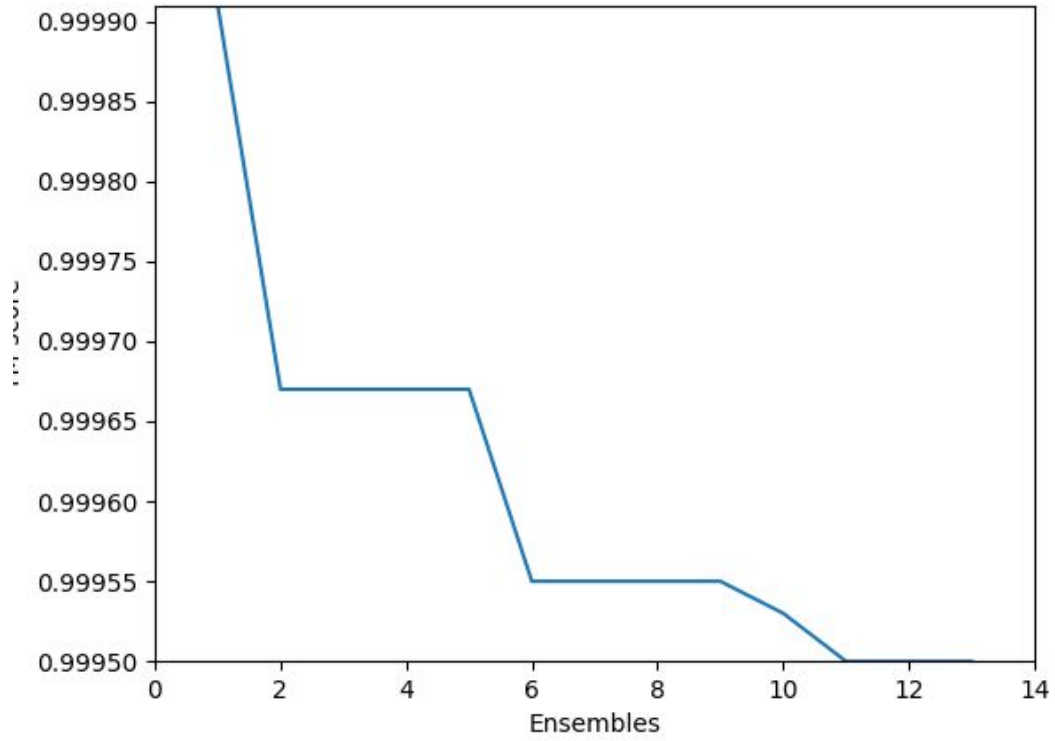
RESULTS for 1cwu (Version 2.0)

Total of 13 ensemble were generated after running for 50 iterations.



TM-score RESULT (Version 2.0)

TM-score results of ensembles when compared with original protein.



REFERENCES

- Thachuk, Chris, Alena Shmygelska, and Holger H. Hoos. "A replica exchange Monte Carlo algorithm for protein folding in the HP model." *BMC bioinformatics* 8.1 (2007): 342.
- Smith, Colin A., and Tanja Kortemme. "Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction." Journal of molecular biology 380.4 (2008): 742-756.
- Lauck, Florian, et al. "RosettaBackrub—a web server for flexible backbone protein structure modeling and design." Nucleic acids research 38.suppl_2 (2010): W569-W575.

THANK YOU