Protein Folding Simulations

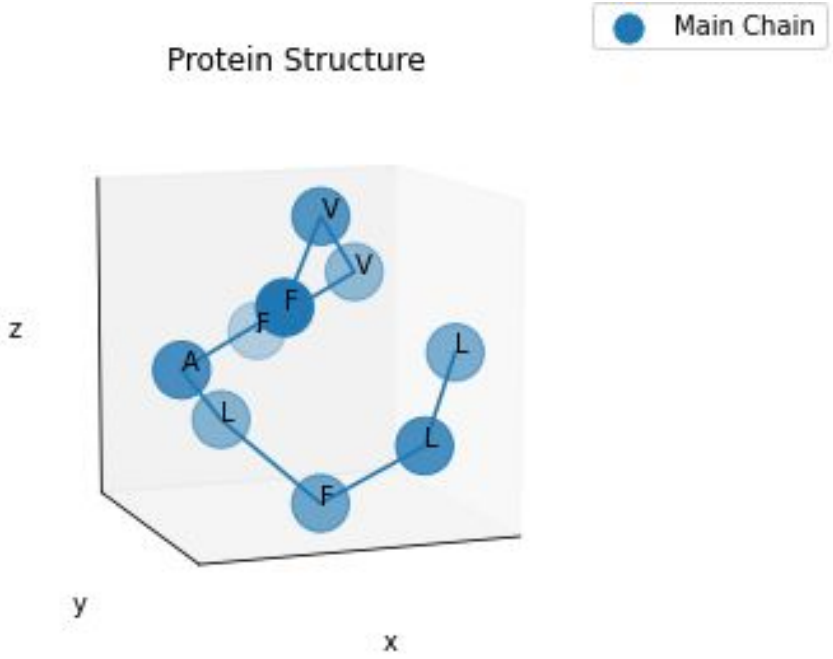
Results

QKRISHI

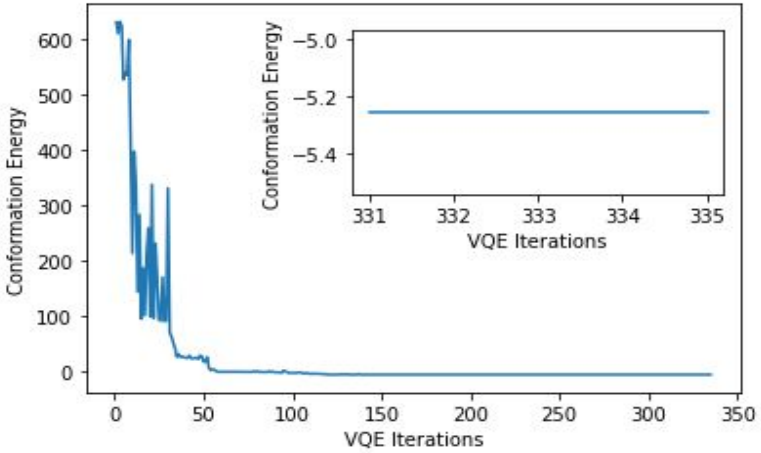
Introduction

- Peptide based vaccines and drugs are increasingly explored for treating SARS-CoV-2. Some of the potential epitopes are LLFLAFVVF, FLLVTLAIL, TLACFVLAA, GDAALALLL, AQFAPSASA, FAMQMAYRF and FVFLVLLPL.
- Simulations are run using Qiskit to obtain the minimum energy fold (active protein) for these sequences.
- Implementation:
  - Based on "**Resource-Efficient Quantum Algorithm for Protein Folding**" by Anton Robert et al.
  - Coarse Grained Model with a Tetrahedral Lattice
    - Monomers encoded as beads
    - Turns encoded as qubits
    - Results present as coordinates of the beads
  - Algorithm used: CVar-VQE
  - No side chains considered for this run

Sequence: LLFLAFVVF



- Minimum Conformational Energy: -5.255439 Hartree
- The folded protein's main sequence of turns is: [1, 0, 3, 2, 0, 1, 2, 0]

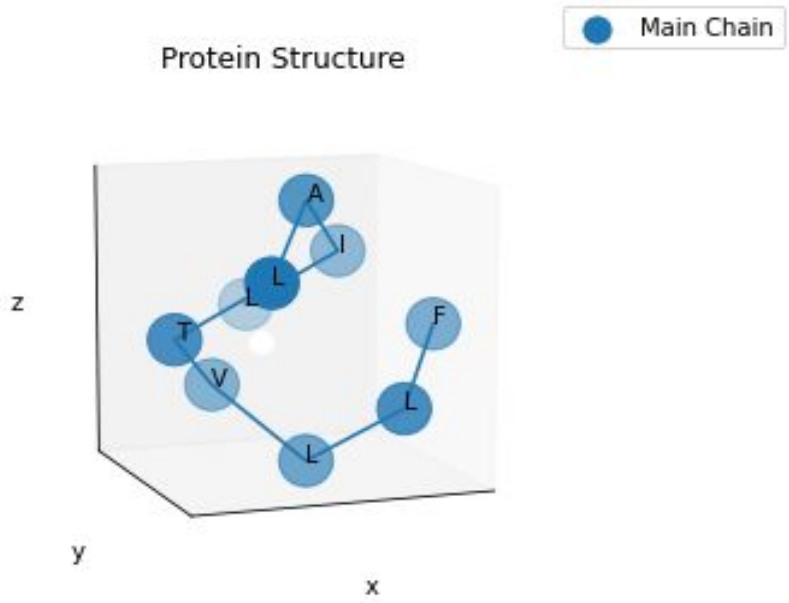


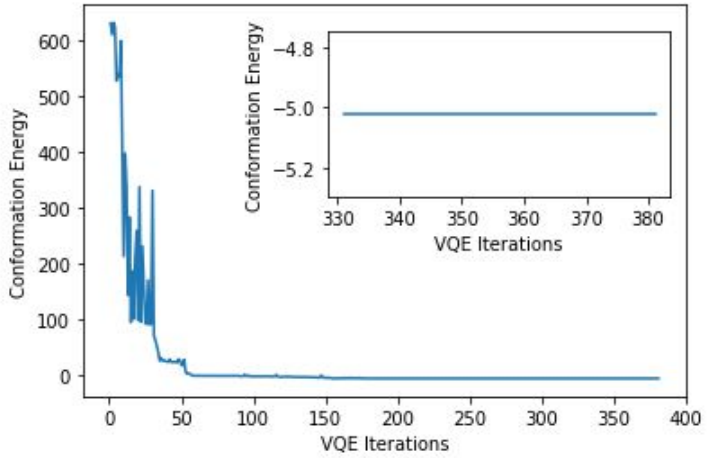
Coordinates

- L = (0.00 ,0.00 ,0.00)
- L = (0.57735, 0.57735, -0.57735)
- F = (1.15470, 0.00, -1.15470)
- L = (1.73205, -0.57735, -0.57735)
- A = (2.30940, 0.00, 0.00)
- F = (1.73205, 0.57735, 0.57735)
- V = (1.15470, 0.00, 1.15470)
- V = ('0.57735, -0.57735, 0.57735)
- F = (1.15470, -1.15470, 0.00)

Total No. of Runs: 335

Sequence: FLLVTLAIL





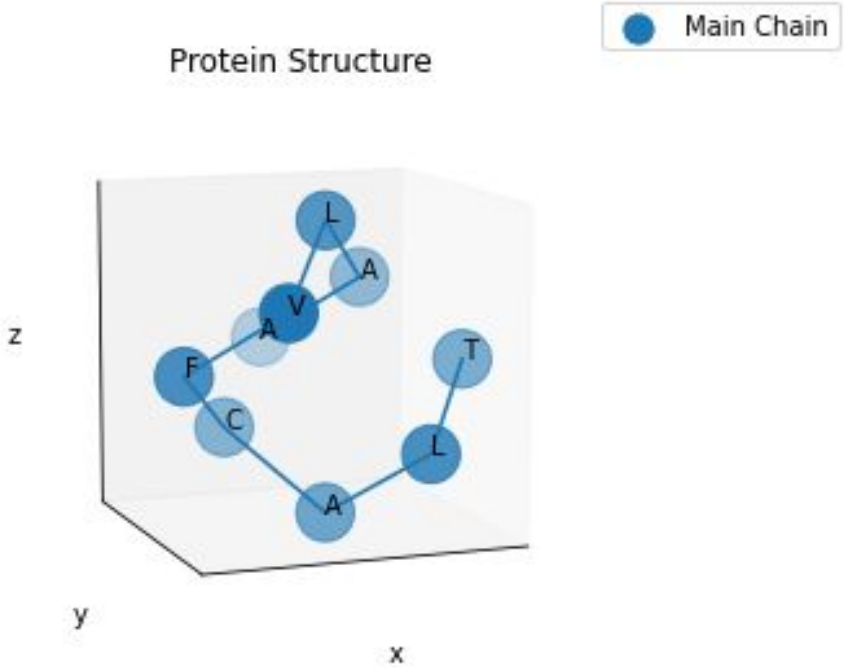
Coordinates:

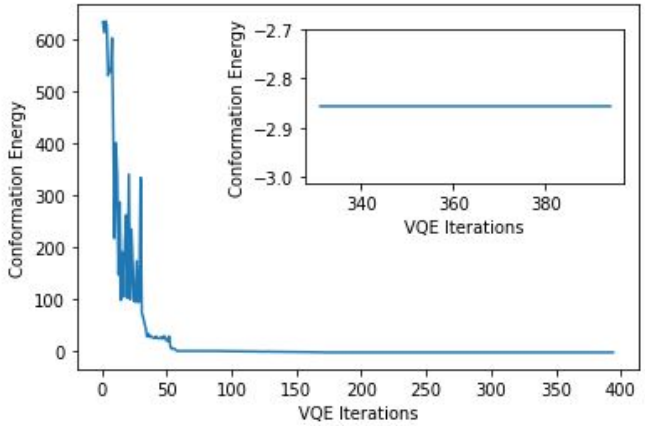
- F = (0.00 ,0.00 ,0.00)
- L = (0.57735, 0.57735, -0.57735)
- L = (1.15470, 0.00, -1.15470)
- V = (1.73205, -0.57735, -0.57735)
- T = (2.30940, 0.00, 0.00)
- L = (1.73205, 0.57735, 0.57735)
- A = (1.15470, 0.00, 1.15470)
- I = (0.57735, -0.57735, 0.57735)
- L = (1.15470, -1.15470, 0.00)

- Minimum Conformational Energy: -5.020990 Hartree
- The folded protein's main sequence of turns is: [1, 0, 3, 2, 0, 1, 2, 0]

Total No. of Runs: 381

Sequence: TLACFVLAA





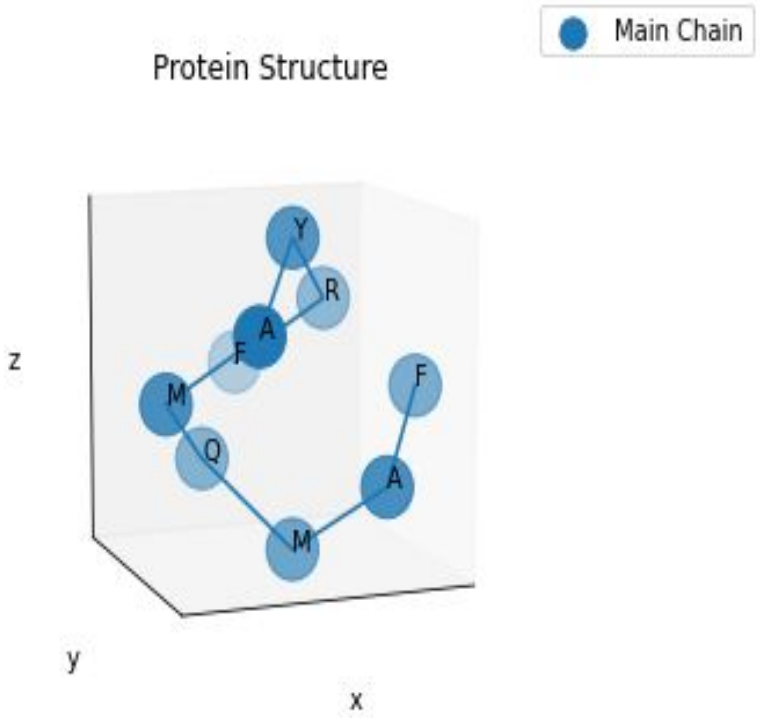
Coordinates:

- T = (0.00 ,0.00 ,0.00)
- L = (0.57735, 0.57735, -0.57735)
- A = (1.15470, 0.00, -1.15470)
- C = (1.73205, -0.57735, -0.57735)
- F = (2.30940, 0.00, 0.00)
- V = (1.73205, 0.57735, 0.57735)
- L = (1.15470, 0.00, 1.15470)
- A = (0.57735, -0.57735, 0.57735)
- A = (1.15470, -1.15470, 0.00)

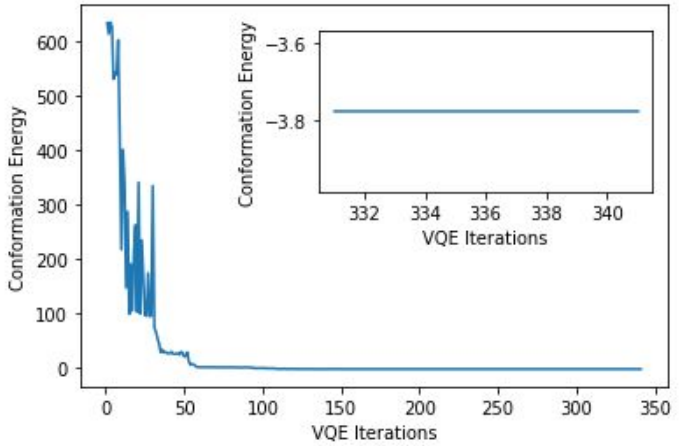
- Minimum Conformational Energy: -2.856981 Hartree
- The folded protein's main sequence of turns is: [1, 0, 3, 2, 0, 1, 2, 0]

Total No. of Runs: 394

Sequence: FAMQMAYRF



- Minimum Conformational Energy: -3.775386 Hartree
- The folded protein's main sequence of turns is: [1, 0, 3, 2, 0, 1, 2, 0]

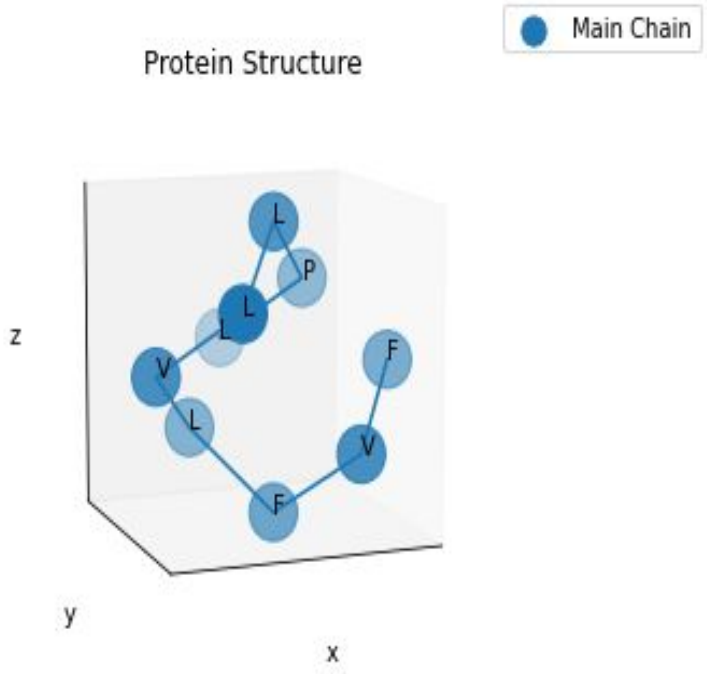


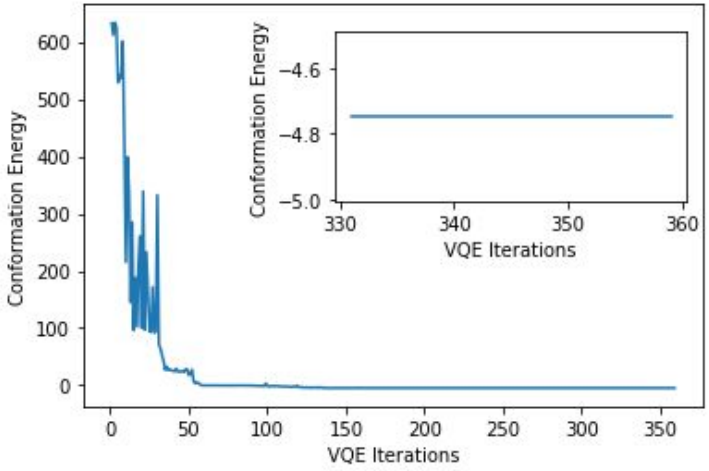
Coordinates:

- F = (0.00 ,0.00 ,0.00)
- A = (0.57735, 0.57735, -0.57735)
- M = (1.15470, 0.00, -1.15470)
- Q = (1.73205, -0.57735, -0.57735)
- M = (2.30940, 0.00, 0.00)
- A = (1.73205, 0.57735, 0.57735)
- Y = (1.15470, 0.0, 1.15470)
- R = (0.57735, -0.57735, 0.57735)
- F = (1.15470, -1.15470, 0.0)

Total No. of Runs: 341

Sequence: FVFLVLLPL





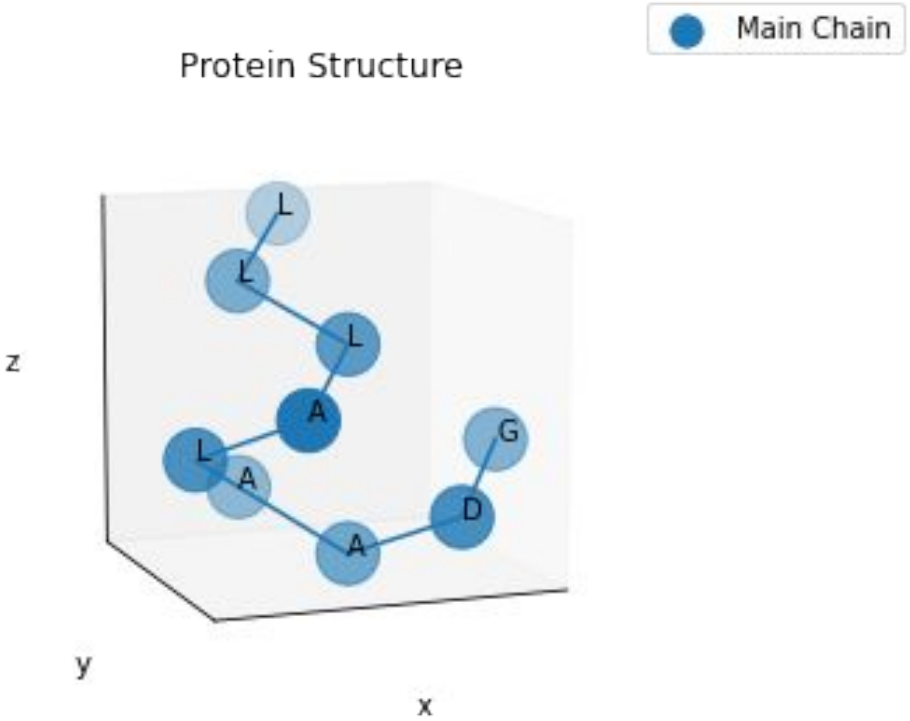
- Minimum Conformational Energy: -4.745997 Hartree
- The folded protein's main sequence of turns is: [1, 0, 3, 2, 0, 1, 2, 0]

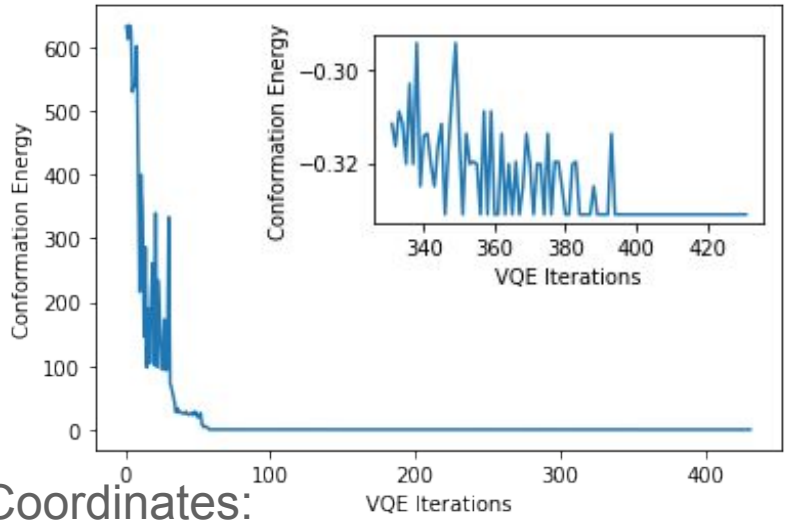
Coordinates:

- F = (0.00 ,0.00 ,0.00)
- V = (0.57735, 0.57735, -0.57735)
- F = (1.15470, 0.00, -1.15470)
- L = (1.73205, -0.57735, -0.57735)
- V = (2.30940, 0.00, 0.00)
- L = (1.73205, 0.57735, 0.57735)
- L = (1.15470, 0.00, 1.15470)
- P = (0.57735, -0.57735, 0.57735)
- L = (1.15470, -1.15470, 0.00)

Total No of Runs: 359

Sequence: GDAALALLL





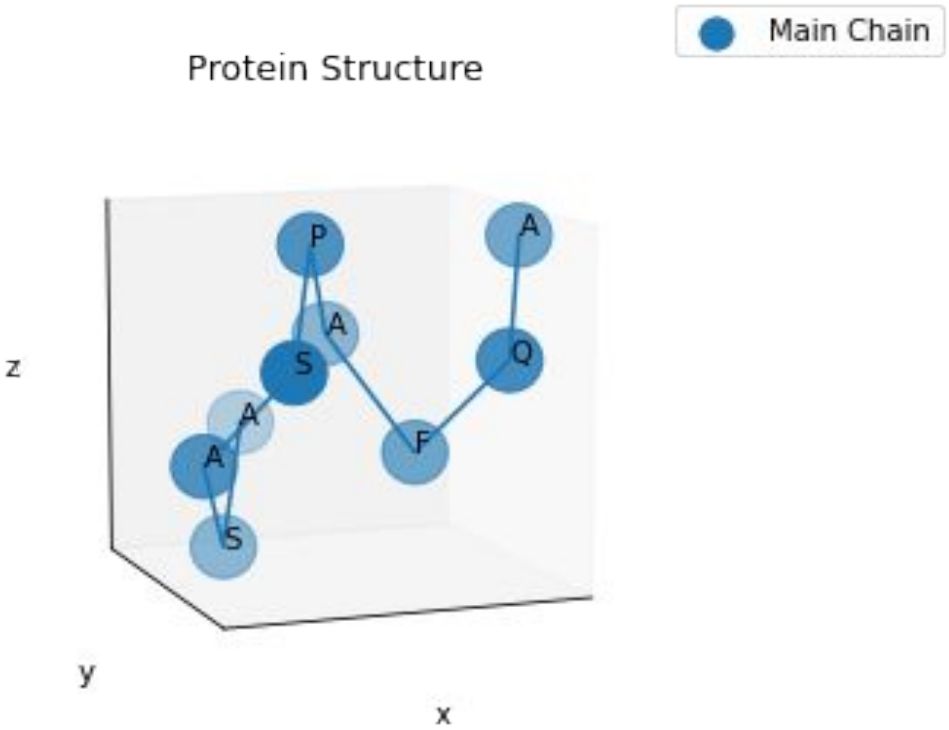
Coordinates:

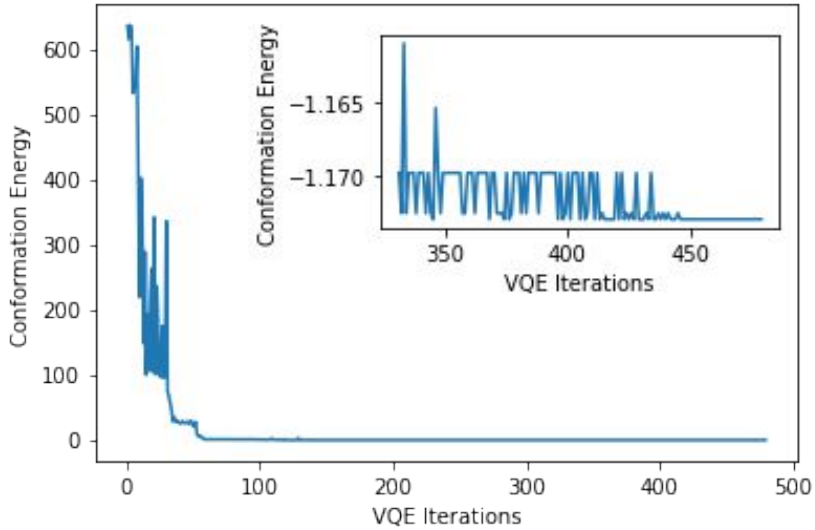
- G = (0.00 ,0.00 ,0.00)
- D = (0.57735, 0.57735, -0.57735)
- A = (1.15470, 0.00, -1.15470)
- A = (1.73205, -0.57735, -0.57735)
- L = (2.30940, 0.0, 0.0)
- A = (1.73205, 0.57735, 0.57735)
- L = (1.15470, 0.0, 1.15470)
- L = (1.73205, -0.57735, 1.73205)
- L = (1.15470, -1.15470, 2.30940)

- Minimum Conformational Energy: -0.331020 Hartree
- The folded protein's main sequence of turns is: [1, 0, 3, 2, 0, 1, 3, 1]

Total No. of Runs : 431

Sequence: AQFAPSASA



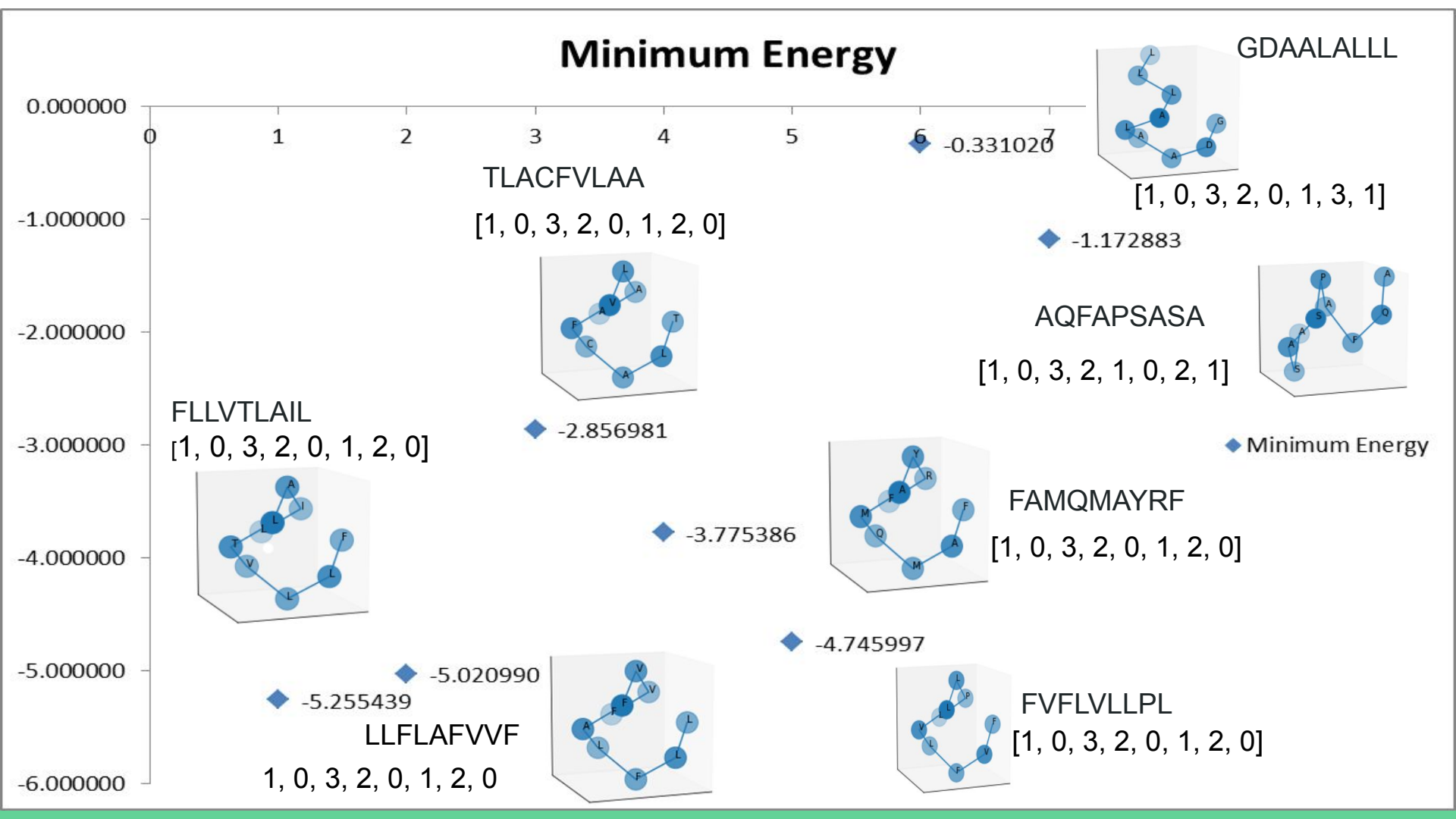


Coordinates:

- A = (0.00 ,0.00 ,0.00)
- Q = (0.57735, 0.57735, -0.57735)
- F = (1.15470, 0.00, -1.15470)
- A = (1.73205, -0.57735, -0.57735)
- P = (2.30940, 0.00, 0.00)
- S = (2.88675, 0.57735, -0.57735)
- A = (3.46410, 0.00, -1.15470)
- S = (2.88675, -0.57735, -1.732050)
- A = (2.30940, -1.15470, -1.15470)

- Minimum Conformational Energy: -1.172883 Hartree
- The folded protein's main sequence of turns is: [1, 0, 3, 2, 1, 0, 2, 1]

Total No. of Runs: 479


Minimum Energy
0.000000
-1.000000
-2.000000
-3.000000
-4.000000
-5.000000
-6.000000
0
1
2
3
4
5
6
7
-0.331020
GDAALALLL
[1, 0, 3, 2, 0, 1, 3, 1]
-1.172883
TLACFVLAA
[1, 0, 3, 2, 0, 1, 2, 0]
-2.856981
AQFAPSASA
[1, 0, 3, 2, 1, 0, 2, 1]
Minimum Energy
FLLVTLAIL
[1, 0, 3, 2, 0, 1, 2, 0]
-3.775386
FAMQMAYRF
[1, 0, 3, 2, 0, 1, 2, 0]
-4.745997
-5.020990
-5.255439
LLFLAFVVF
1, 0, 3, 2, 0, 1, 2, 0
FVFLVLLPL
[1, 0, 3, 2, 0, 1, 2, 0]

Next Steps

- Convert coarse grained coordinates into all atoms and vice versa.
- Run classical all atoms and coarse grained molecular dynamics simulations of individual ligands for comparing the quantum results.
- Perform docking of these ligands with corona virus protein and calculate the binding energy.
- Run classical all atoms and coarse grained molecular dynamics simulations of the (complex)docked system and evaluate the structural change.
- Perform electronic structure theory (EST) or QM/MM calculations if necessary.

Coordinate Conversion: VMD, PyMol

MD : GROMACS, Alphafold

EST : Gaussian, GAMESS, PySCF, Psi4

QM/MM : Gaussian, CP2K, NWChem

THANK YOU