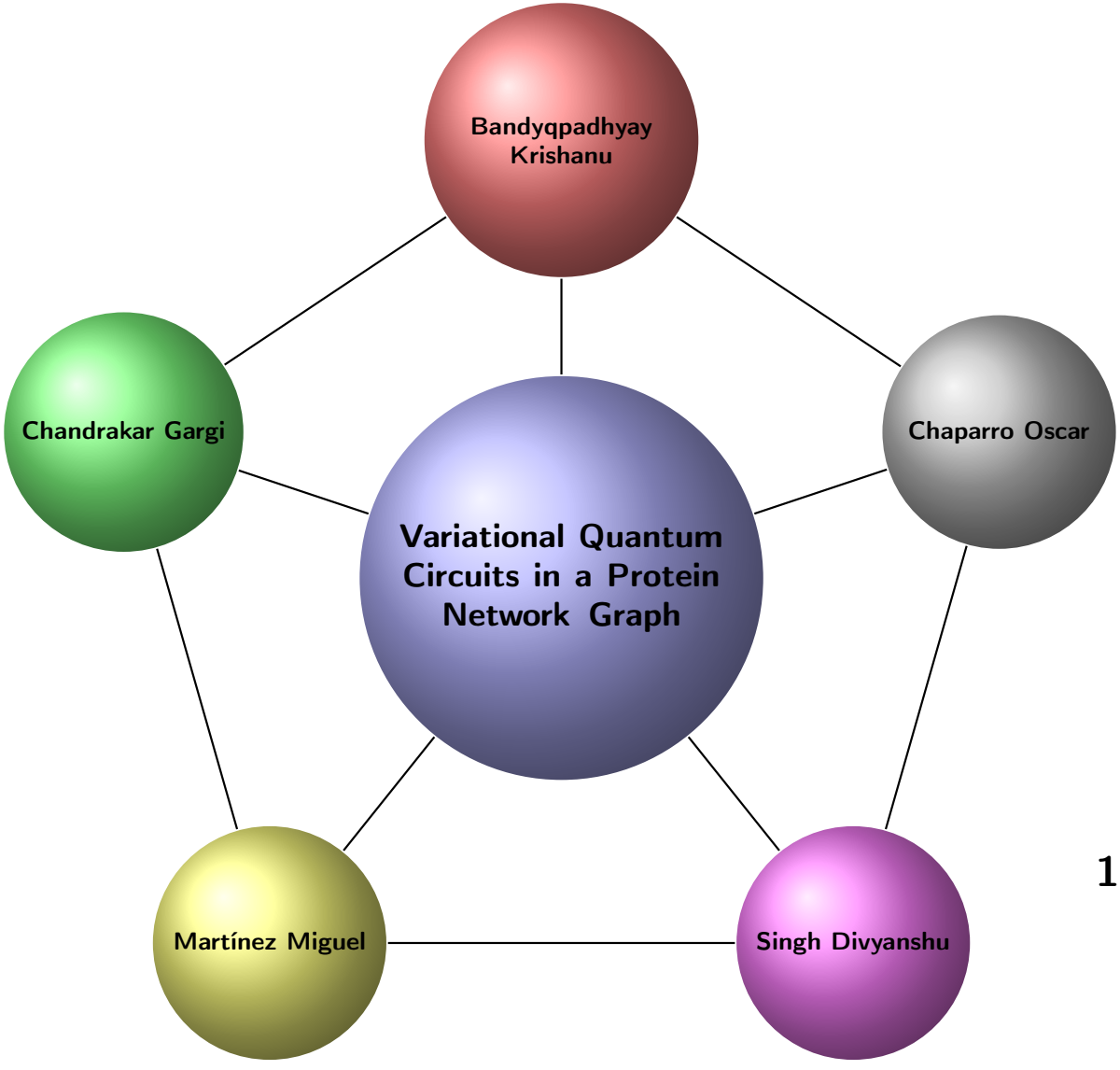
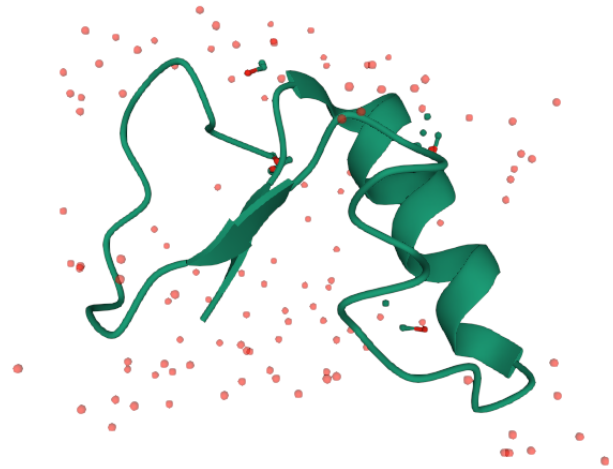
Qiskit Global Hackaton 2021

Qiskit





19-November-2021

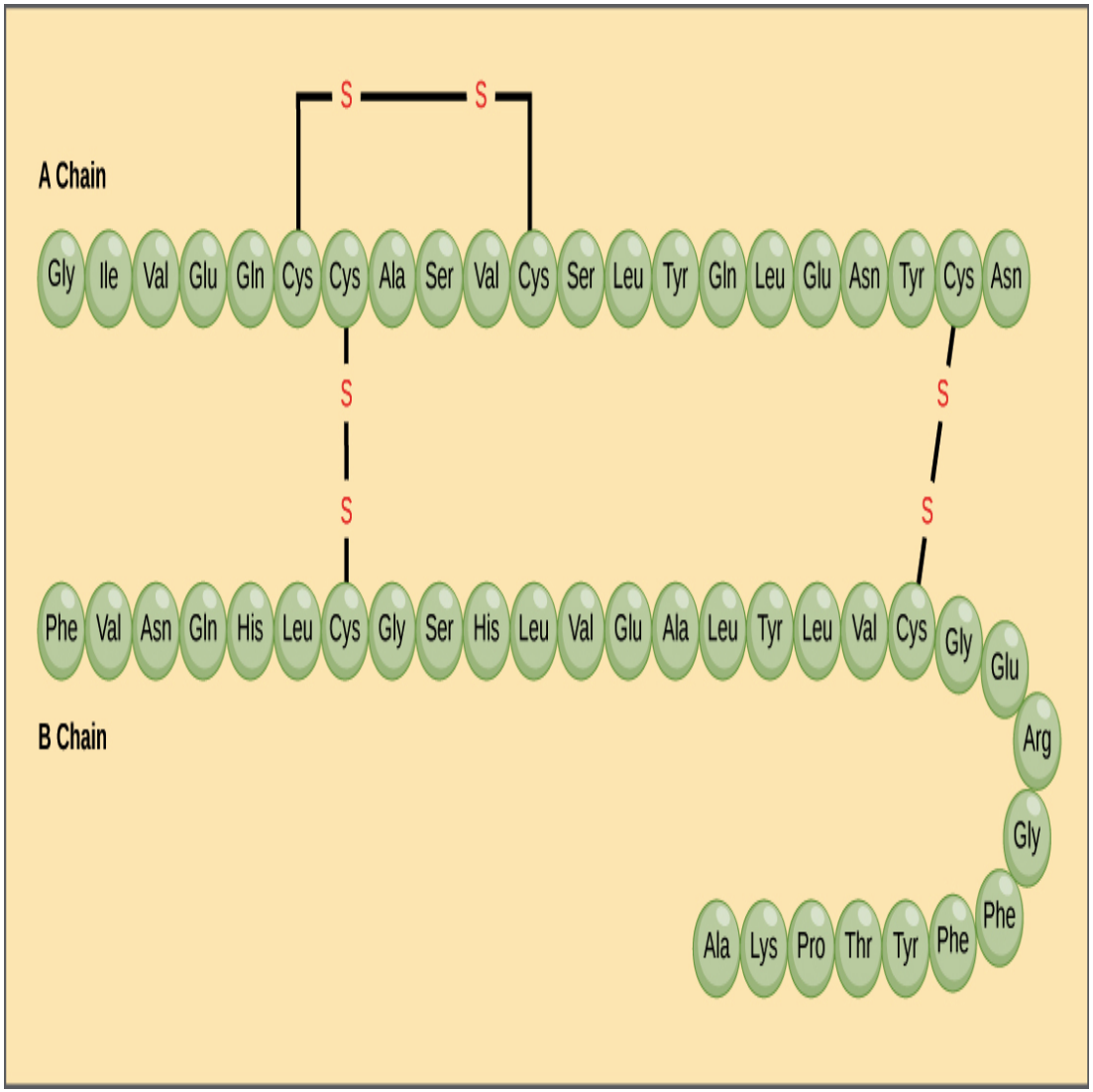
Contents

Objectives

- To find bounding amino acids function on a delimited network graph approach that characterized a chemical structure using a variational quantum circuit.
- To apply quantum circuits used in graph theory and optimization problems to map 3D-protein structure into chemical interaction network.
- Using the QAOA variational form to find an approximate solution for the protein folding quadratic program.
- To encourage people to push into quantum computing and solve real molecular biology problems faster than classical methods.

Introduction to the protein-protein interactions

- Protein are made of amino acids chain, which are biomolecules that have several biochemical interaction between them.
- 3D-protein structure depends on the the biochemical bondings between the structure to reach their functional form.
- Detect and found the portion of the chemical interaction that stabilize the molecule have several application as farmacology, biology and bioinformatics.



Bonding energy function

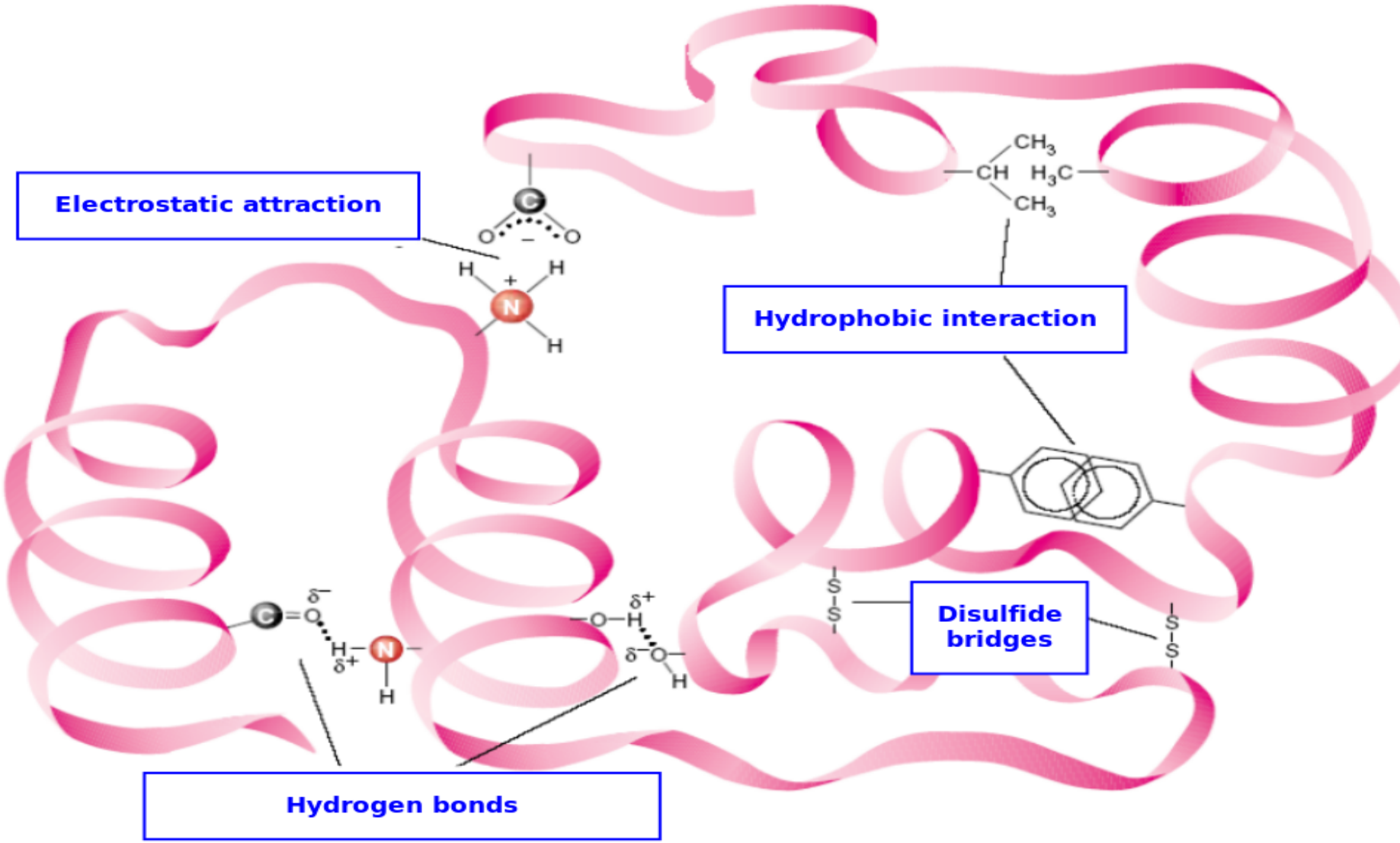
$$E_{total\ potential} = E_{structure} + E_{bonding} \tag{1}$$

$$E_{bonding} = \sum_{i=0}^{n} \sum_{j>i}^{n} \left(-\frac{A_{ij}}{r_{ij}^{6}} + \frac{B_{ij} e^{-C_{ij} r_{ij}}}{k_1} \right) + \sum_{i=0}^{n} \sum_{j>i}^{n} \frac{332 q_i q_j}{k_2 \epsilon r_{ij}} \tag{2}$$

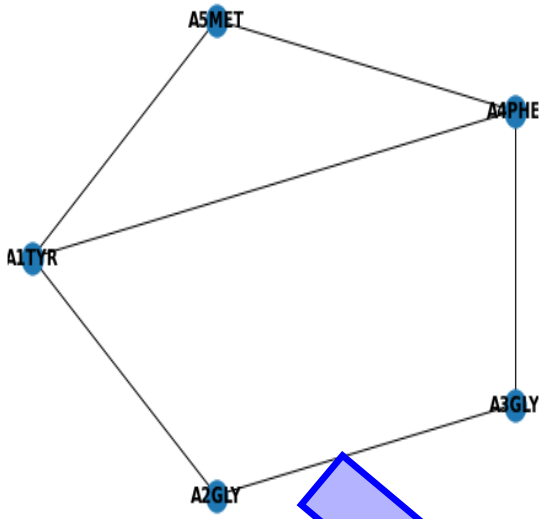
With n number of atoms in the molecule (YA., A., and HA, 2006).

- Hydrophobic.
- Disulfide (bridnge).
- Hydrogen bonds.
- Ionic and aromatic (partial).
- Cation-π.
- Peptide bond.

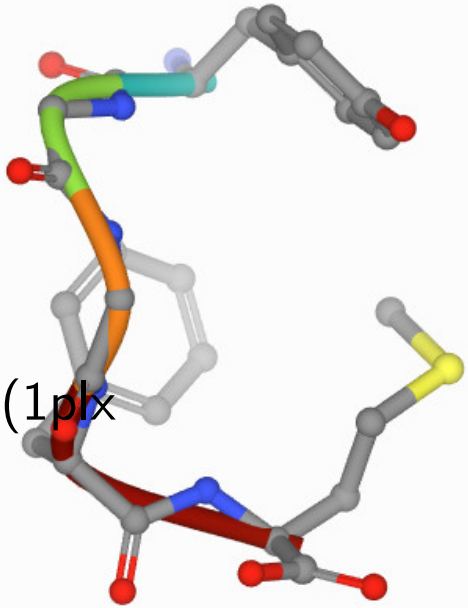
Main chemical bonding inside a protein-protein 3D-structure



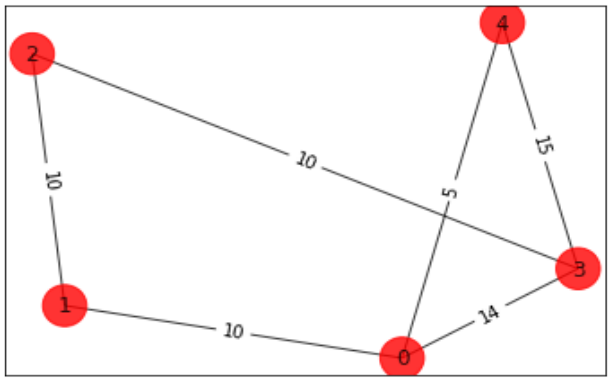
Graph network methodology



Amino Acids as nodes of graph network.

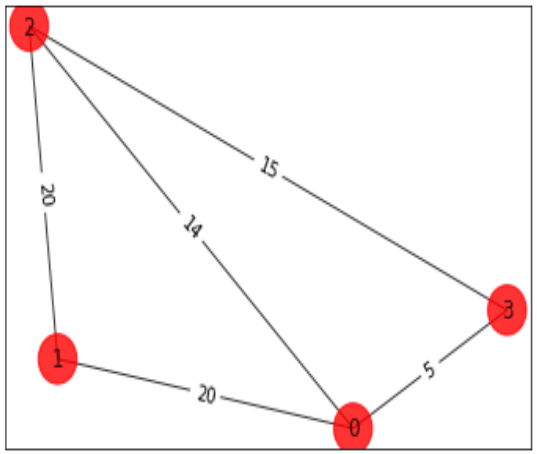


3D-protein structure (1plx (Harano, 2012)).



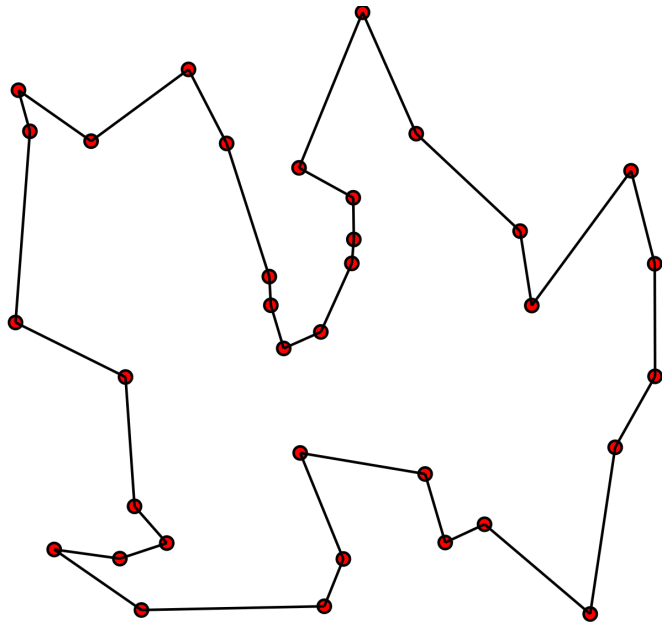
Weighted by the interacting potential function.

Graph reduction.



Classic solution

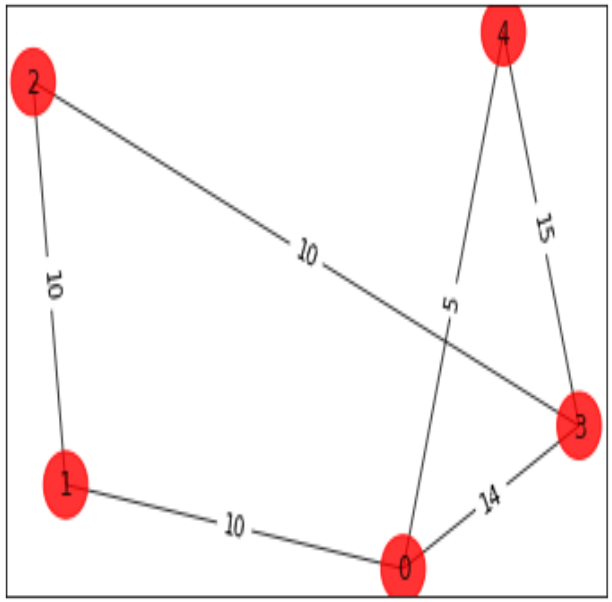
We implement a classical solution using the algorithm traveling salesman as an NP-complete problem. for protein folding problem based on a protein network graph to find bounding amino acids, the main objective is to find the nodes with more interactions or weights (proteins with more connections).



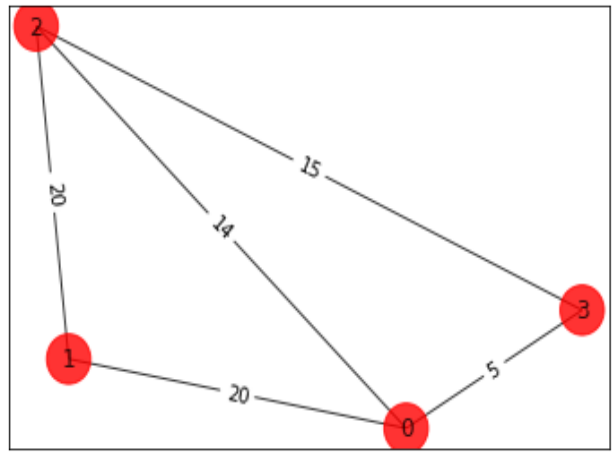
Protein network graph representation

Grouped amino acids = node value:

- A1TYR = 0
- A2GLY = 1
- A3GLY = 2
- A4PHE = 3
- A5MET = 4



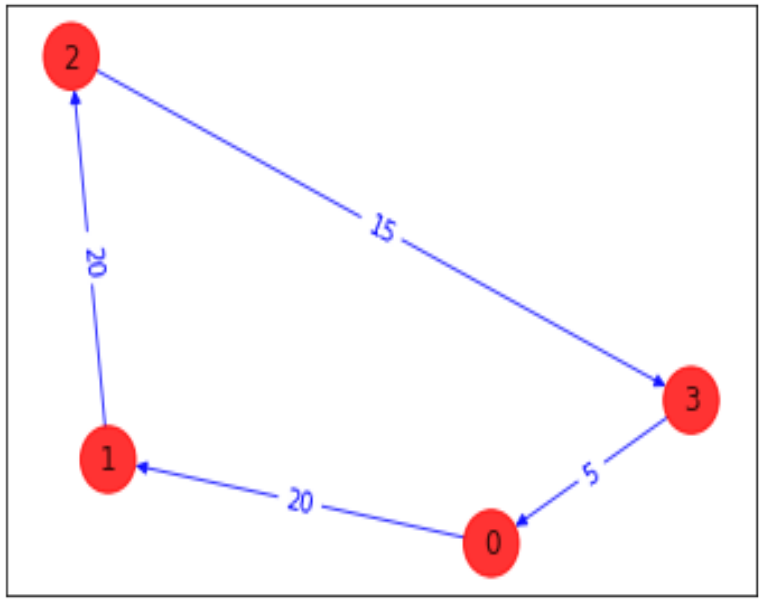
The weighted protein graph can be reduced and grouped with Bounding energy function (ANIS et al., 2021) [2]



Brute Force Approach

```
a=list(permutations(range(1,N)))
last_best_distance = 0
for i in a:
distance = 0
pre_j = 0
for j in i:
distance = distance + w[j,pre_j]
pre_j=j
distance = distance + w[pre_j, 0]
order = (0,) + i
if distance > last_best_distance :
best_order=order
last_best_distance=distance
print('order = ' + str(order)
+ ' Weight of protein folding = ' +
str(distance))
```

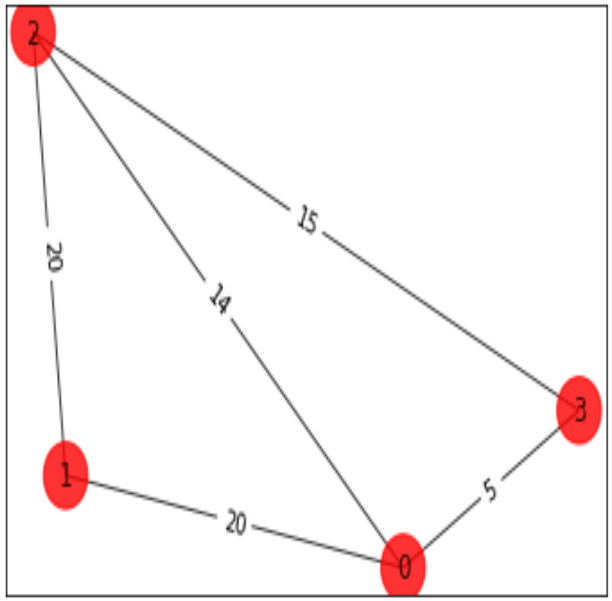
order = (0, 1, 2, 3) Weight of protein folding = 60
Best order from brute force = (0, 1, 2, 3) with total weight = 60



Quantum computing approach

Protein folding problem base on the travelling salesman.

$$Weight_{matrix} = \begin{bmatrix} 0 & 20 & 14 & 5 \\ 20 & 0 & 20 & 0 \\ 14 & 20 & 0 & 15 \\ 5 & 0 & 15 & 0 \end{bmatrix}$$



Cost function:

$$C(\mathbf{x}) = \sum_{i,j} w_{ij} \sum_{p} x_{i,p} x_{j,p+1} + A \sum_{p} \left(1 - \sum_{i} x_{i,p}\right)^2 + A \sum_{i} \left(1 - \sum_{p} x_{i,p}\right)^2, \quad (3)$$

Quadratic program to QUBO

Qubo matrix and vector:

$Q_{ij} = W_{ij} \qquad c_i = \sum_{i=1}^{n} i.$

In this case we need to maximize the cost function:

$$\sum_{i,j} w_{ij} \sum_{p} x_{i,p} x_{j,p+1} + A \sum_{p} \left(1 - \sum_{i} x_{i,p}\right)^2 + A \sum_{i} \left(1 - \sum_{p} x_{i,p}\right)^2$$

$$= \sum_{i,j} \sum_{p} x_{i,p} w_{ij} x_{j,p+1} + A \sum_{p} \left(1 - \sum_{i} x_{i,p}\right)^2 + A \sum_{i} \left(1 - \sum_{p} x_{i,p}\right)^2$$

$$= A \sum_{p} \left(1 - \sum_{i} x_{i,p}\right)^2 + A \sum_{i} \left(1 - \sum_{p} x_{i,p}\right)^2 + \sum_{i,j} \sum_{p} x_{i,p} w_{ij} x_{j,p+1}$$

$$= c^T x + x^T Q x,$$

Running on a Quantum Computer

Qiskit provides automatic conversion from a suitable *QuadraticProgram* to an Insing Hamiltonian, which then allows to leverage all the *MinimumEigenSolver*.

Solving a QUBO with the minimum eigen optimizer based on VQE

```
spsa = SPSA(maxiter=300)
ry = TwoLocal(qubitOp.num_qubits, 'ry', 'cz', reps=5,
entanglement='linear')
vqe = VQE(ry, optimizer=spsa, quantum_instance=quantum_instance)

result = vqe.compute_minimum_eigenvalue(qubitOp)
```

```
energy: -2417.052734375
time: 454.48226857185364
feasible: True
solution: [2, 3, 0, 1]
solution objective: 60.0
```

References I

- ANIS, MD SAJID et al. (2021). *Qiskit: An Open-source Framework for Quantum Computing*. DOI: 10.5281/zenodo.2573505.
- Harano, Yuichi (2012). "Application of Hydration Thermodynamics to the Evaluation of Protein Structures and Protein-Ligand Binding". In: *Entropy* 14.8, pp. 1443–1468. ISSN: 1099-4300. DOI: 10.3390/e14081443. URL: https://www.mdpi.com/1099-4300/14/8/1443.
- YA., Arnautova, Jagielska A., and Scheraga HA (2006). "A new force field (ECEPP-05) for peptides, proteins, and organic molecules". In: *The Journal of Physical Chemistry B* 10.110, pp. 5025–5044. DOI: 10.1021/jp054994x. URL: https://pubs.acs.org/doi/10.1021/jp054994x.