Protein Structures

Primary Structure

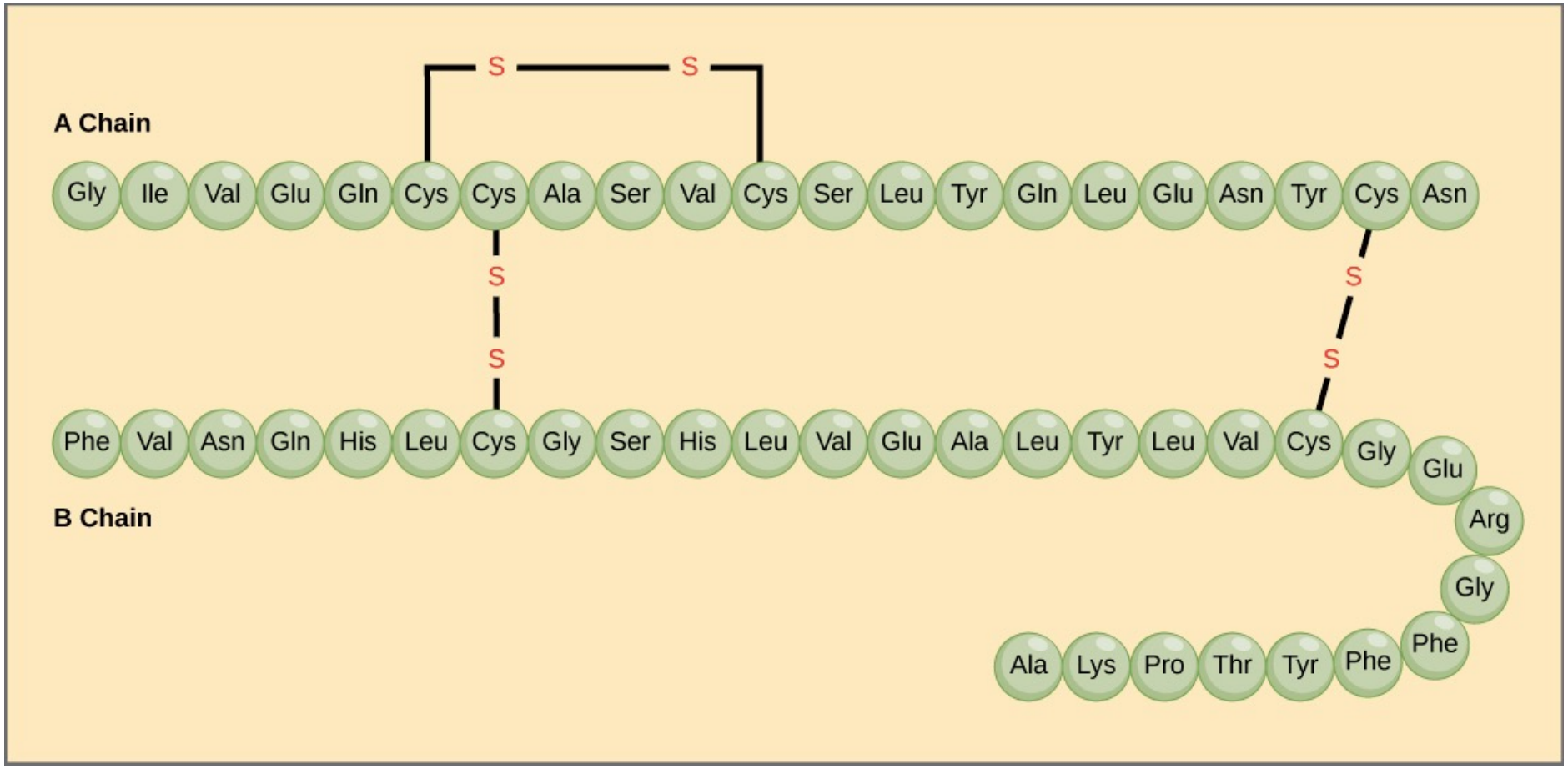
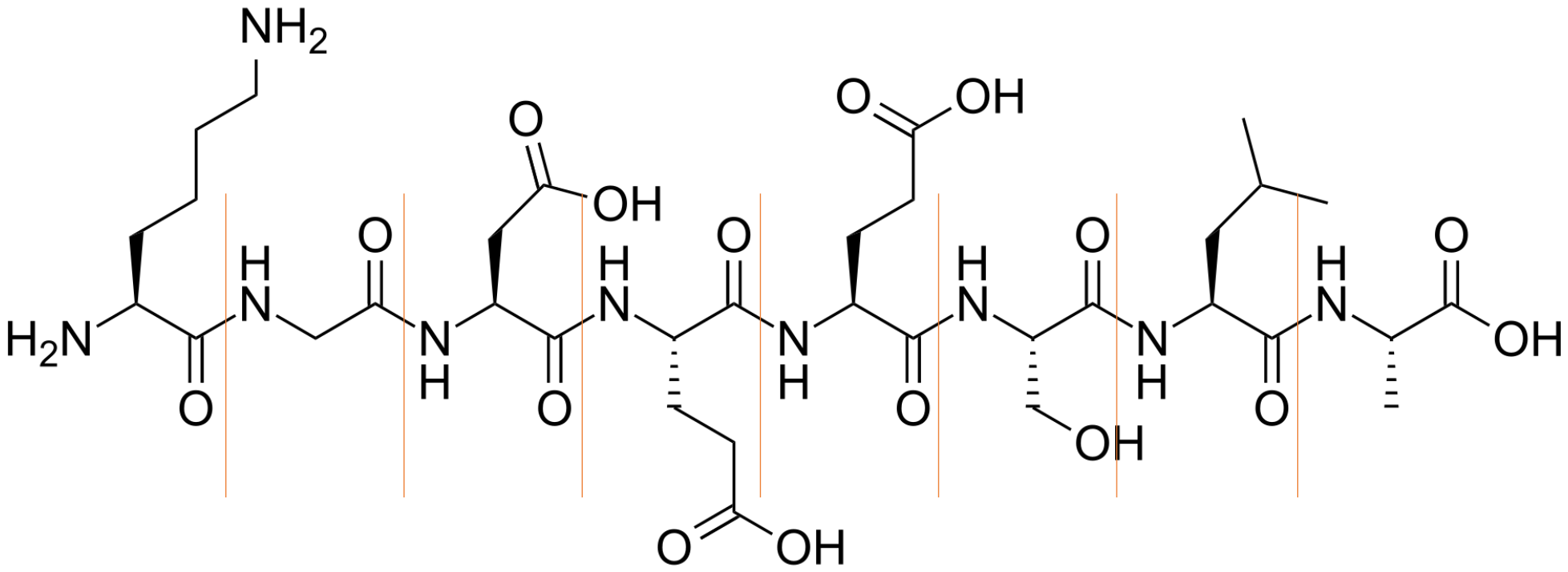


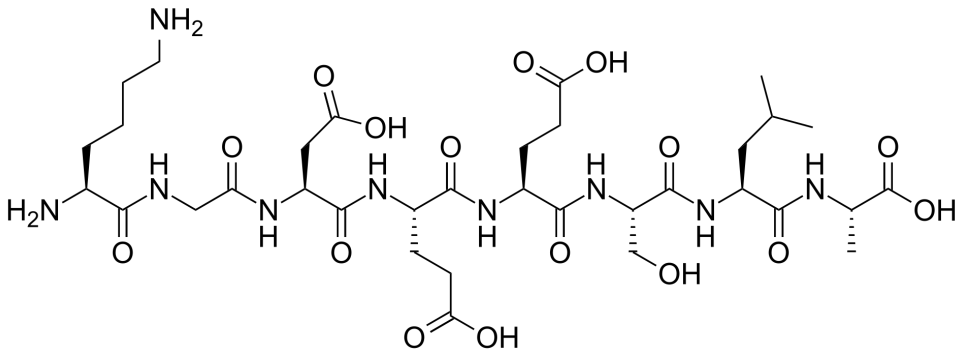
Image Credit: Khanacademy

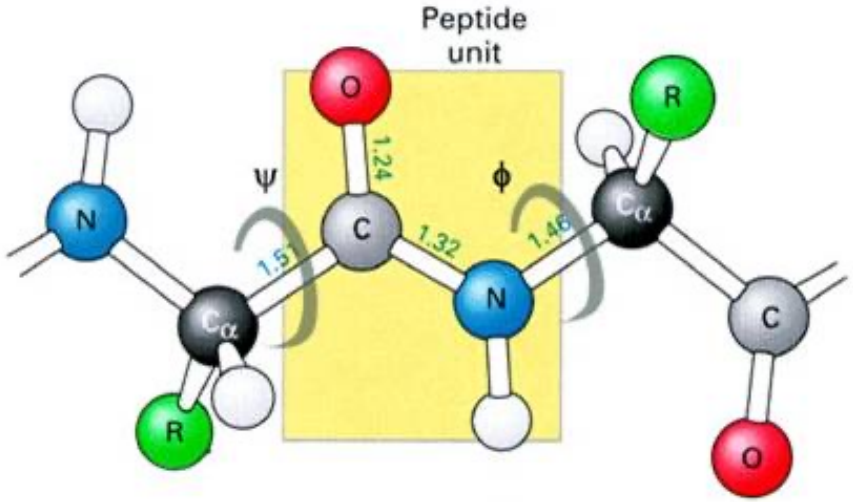
Beefy meaty peptide (delicious peptide)



KGDEESLA

Torsion Angles





Secondary Structure

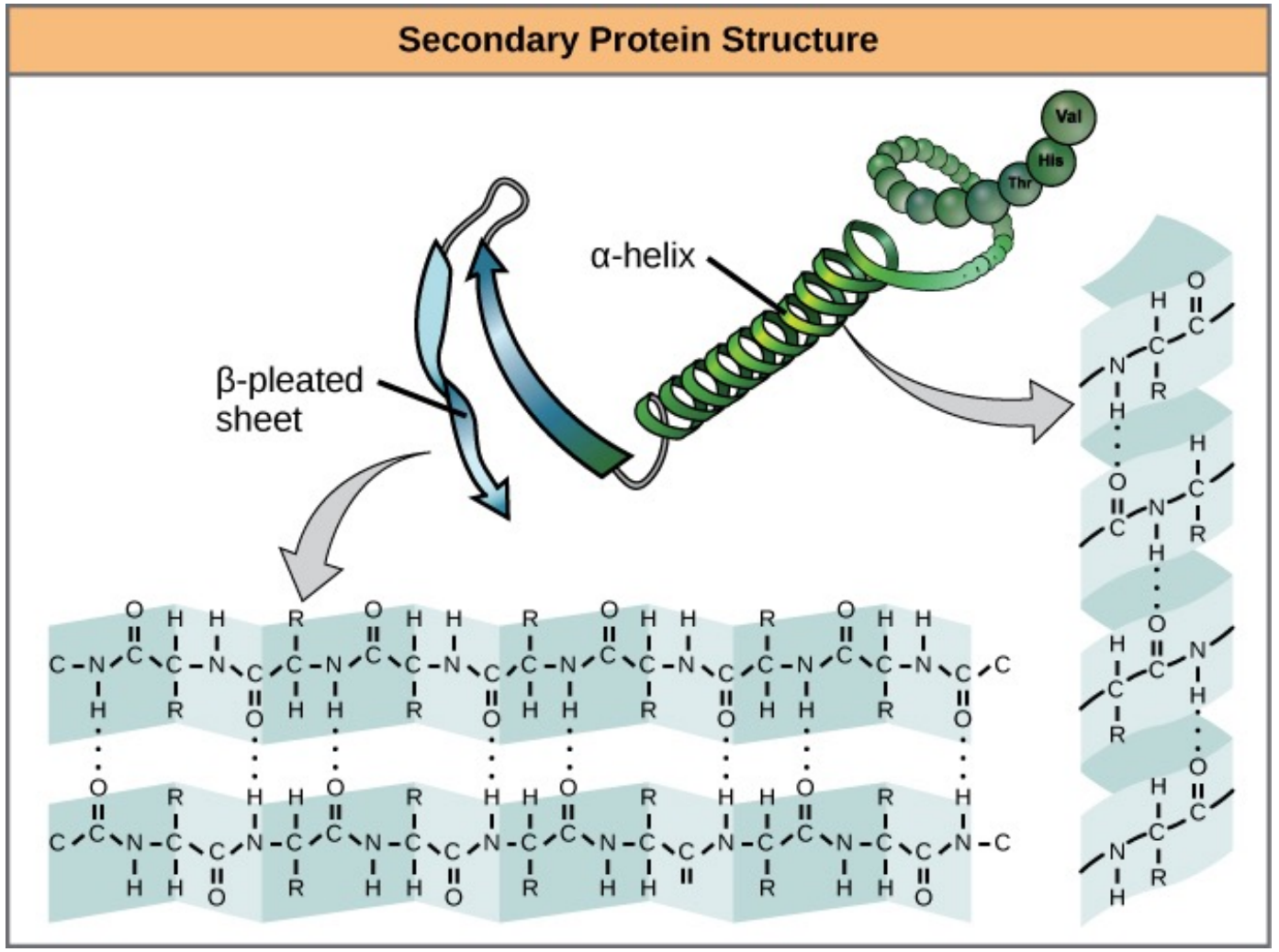
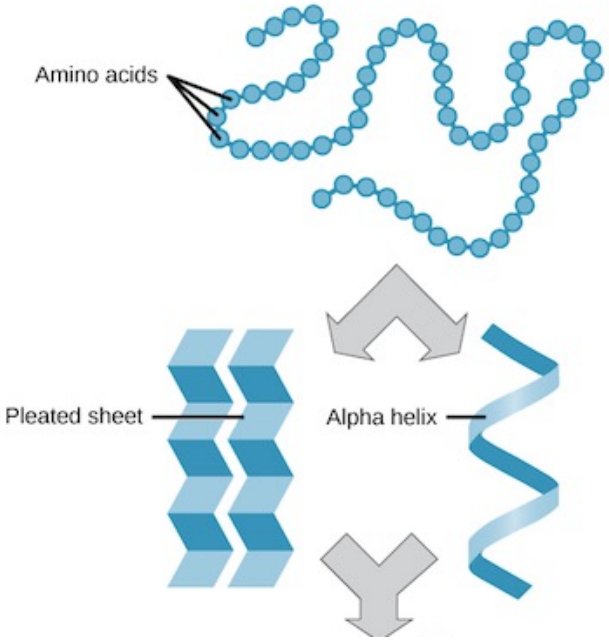


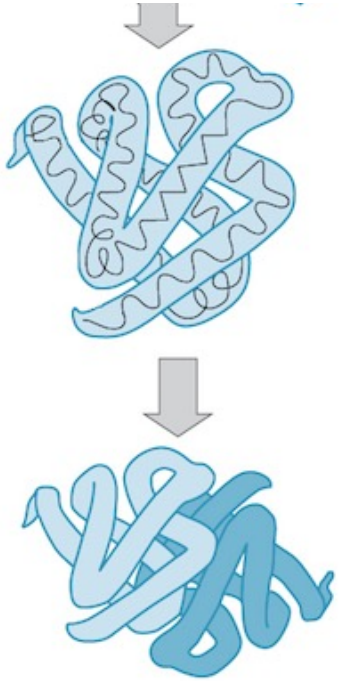
Image credit: OpenStax Biology.

Tertiary and Quaternary Structures



Primary protein structure
sequence of a chain of animo acids

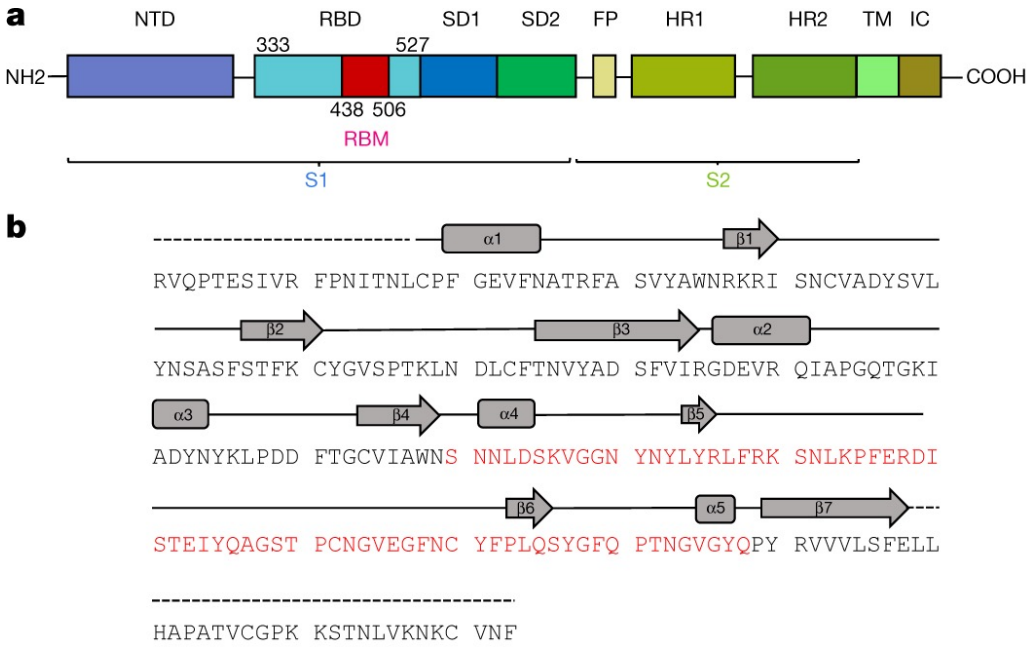
Secondary protein structure
hydrogen bonding of the peptide backbone causes the amino acids to fold into a repeating pattern

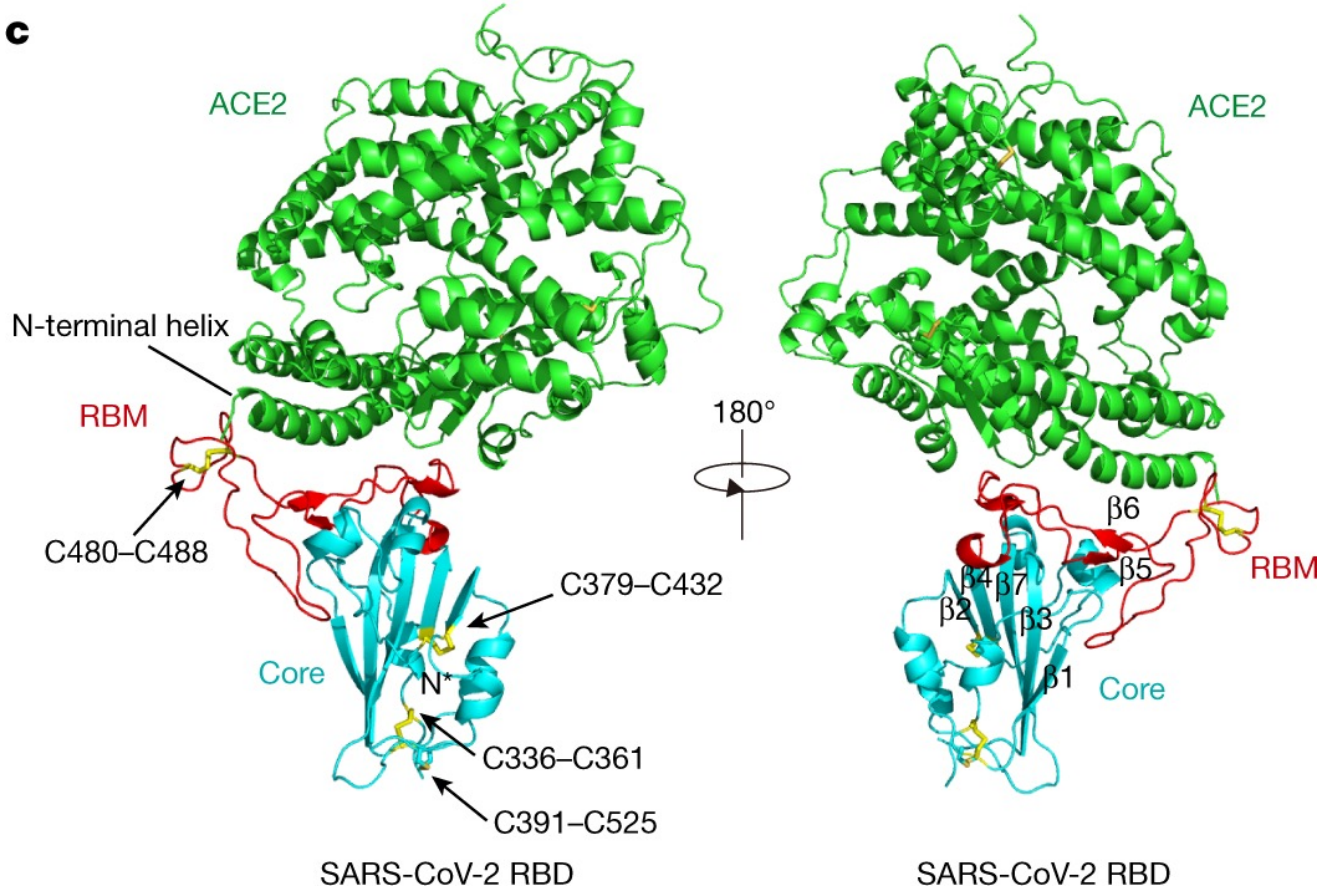


Tertiary protein structure
three-dimensional folding pattern of a protein due to side chain interactions

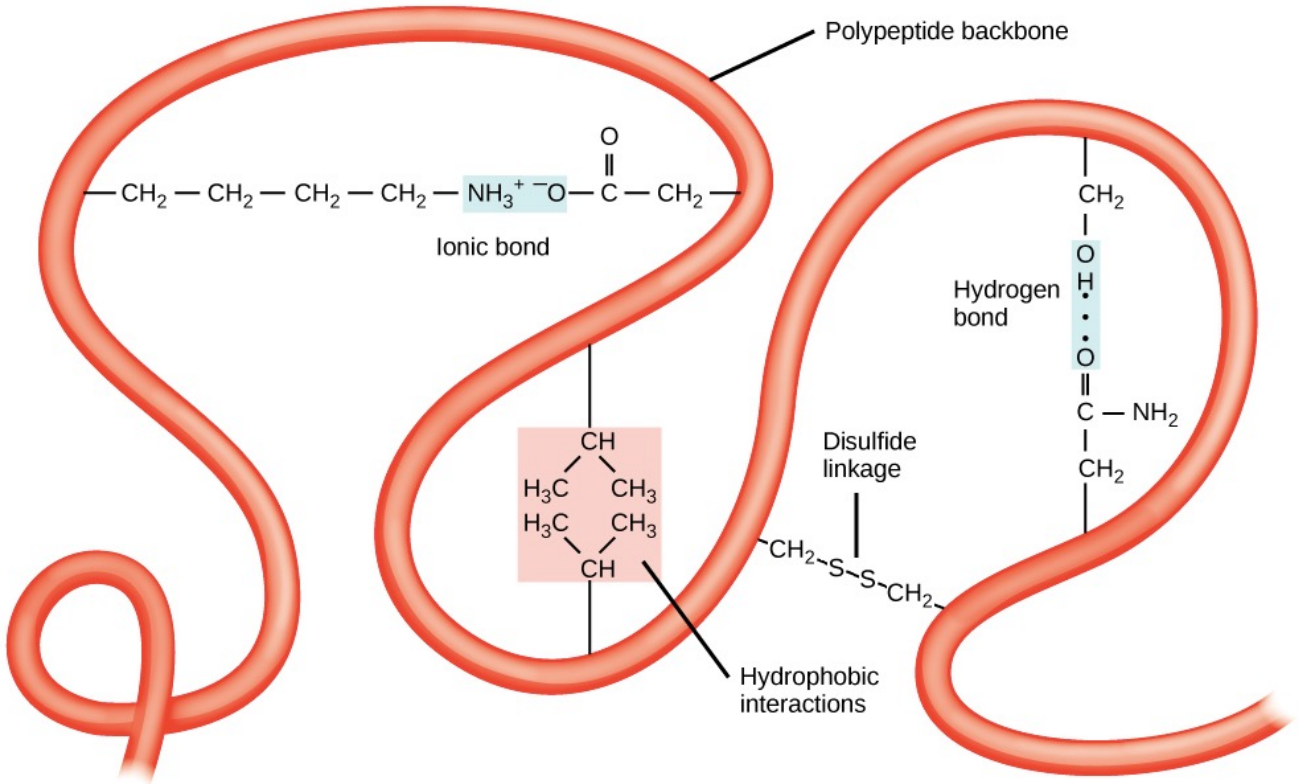
Quaternary protein structure
protein consisting of more than one amino acid chain

SARS-COV2 Spike protein and ACE



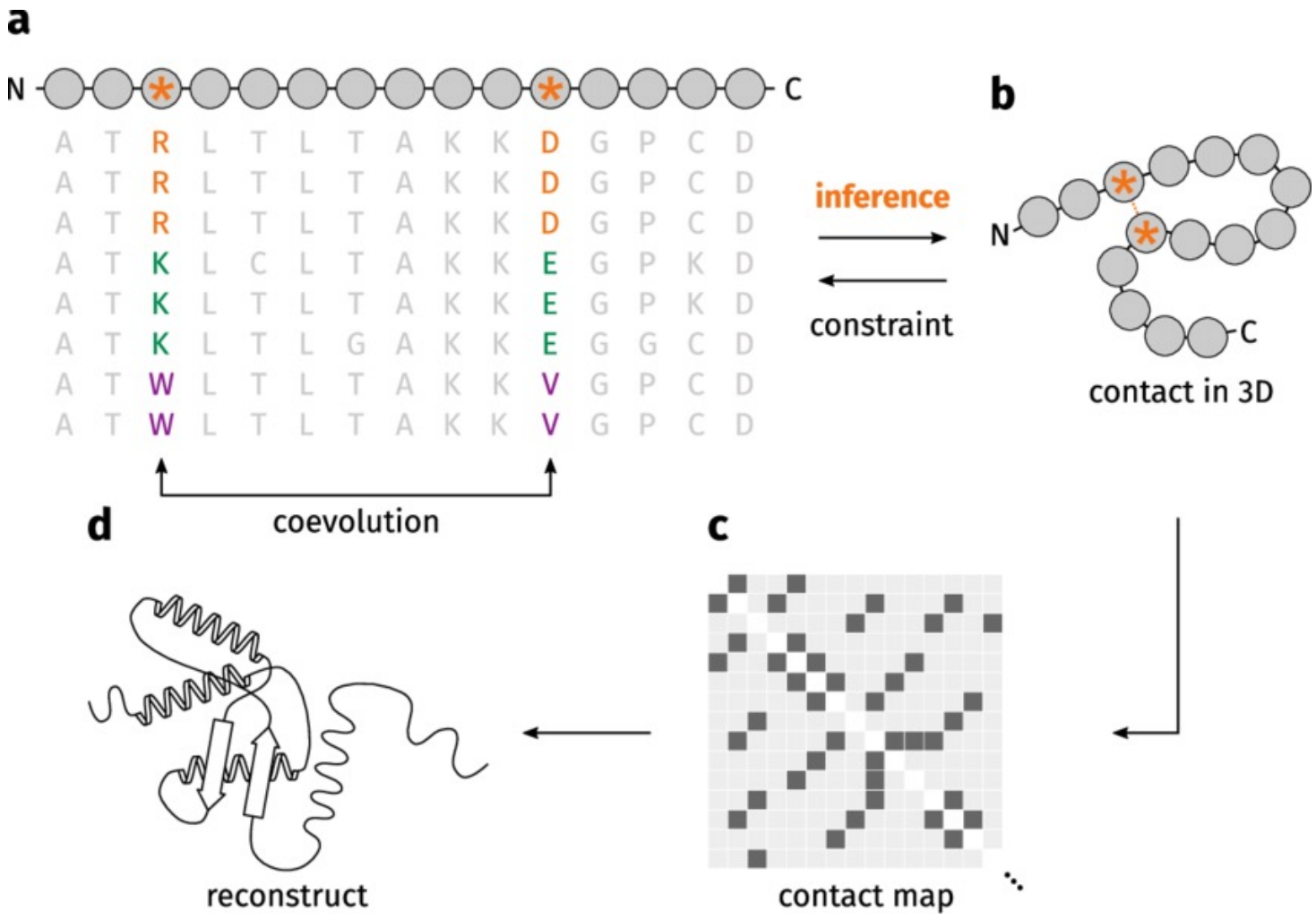


Free Energy Minimization



- Certain interactions reduces the “free energy”.
- It is believed that the natural structure minimizes the free energy.

Additional Information (e.g. Coevlolution)



- Additional information can be incorporated in the “energy” function.
- It’s therefore just a scoring function.

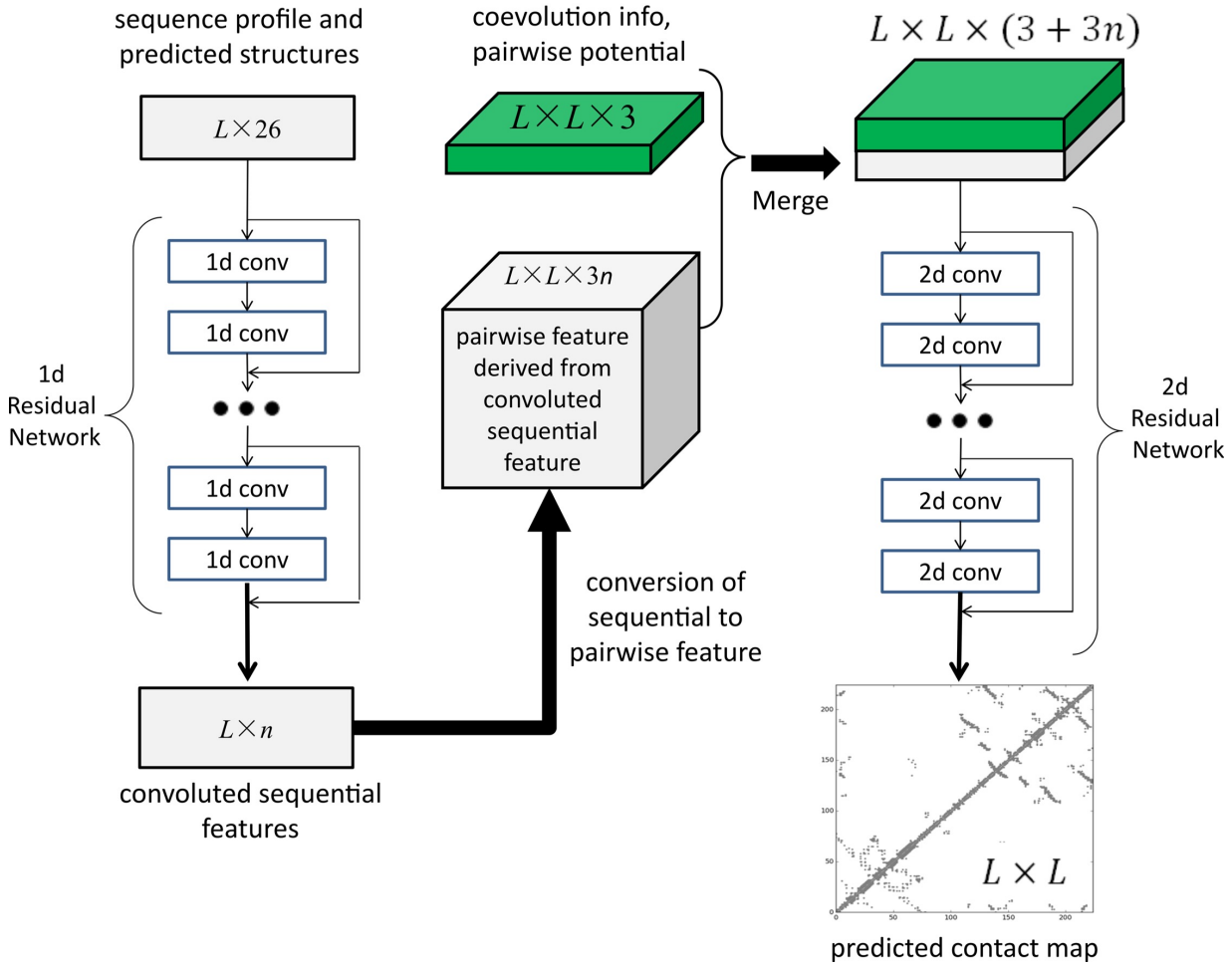
Image credit: DOI:10.1038/s41598-019-55047-4

Search Approach

- Monte Carlo
  - Sampling of the confirmational landspace and minimize energy
- Threading
  - Use existing protein structures as models

Raptor X Contact Map

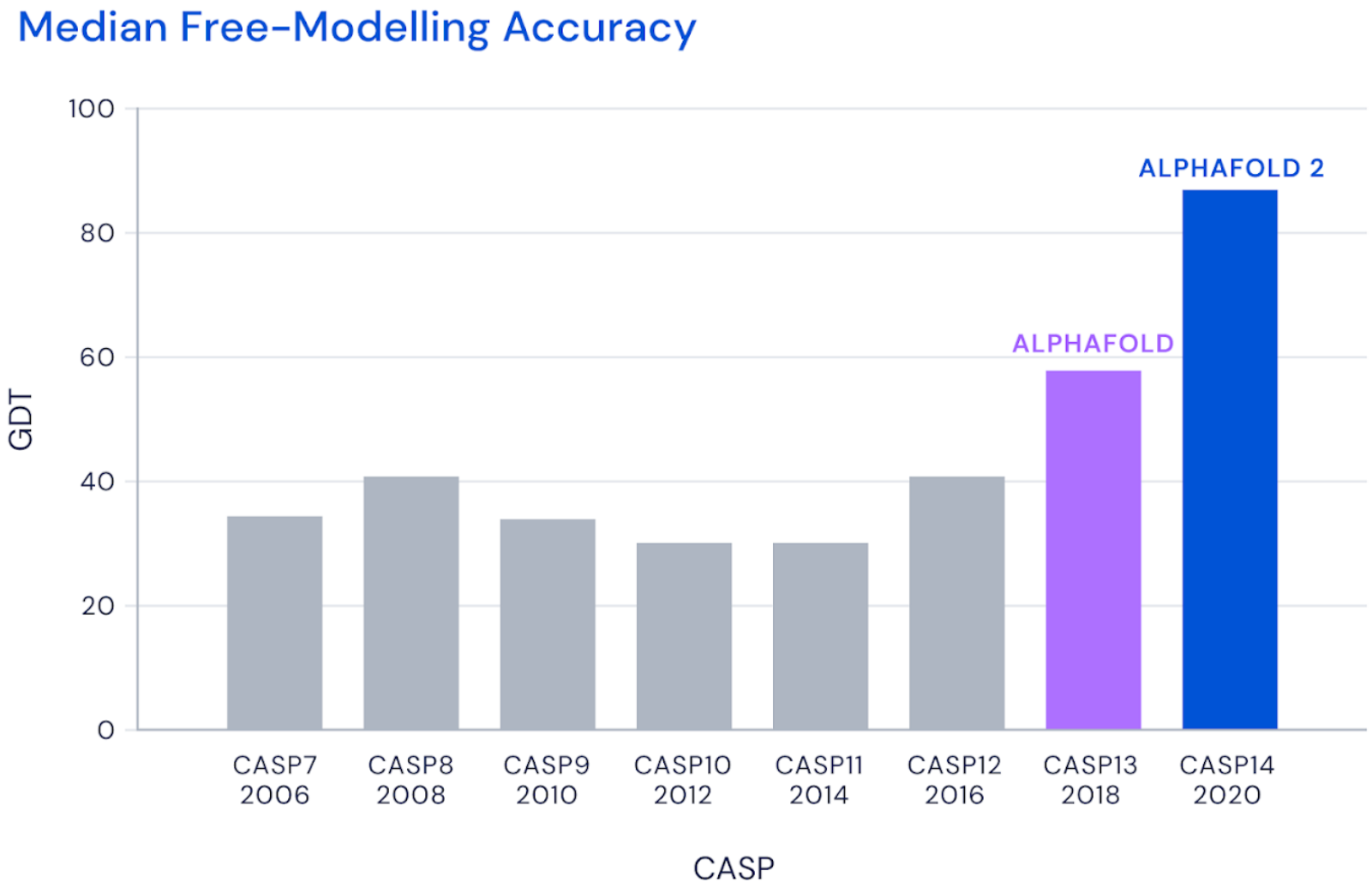
- Use ResNet to predict the contact map.
- Use some traditional approach to predict structure assisted by contact map.
- Wang S, Sun S, Li Z, Zhang R, Xu J (2017) Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. PLoS Comput Biol 13(1): e1005324. doi:10.1371/journal.pcbi.1005324



Alpha Fold

- AlphaFold
  - Takes Jinbo Xu's approach + a gradient descent search strategy to minimize the energy.
- AlphaFold2
  - Predict the contact map (2D representation) with neural network.
  - Followed by torsion angel and structure prediction with neural network.
  - New network structures.

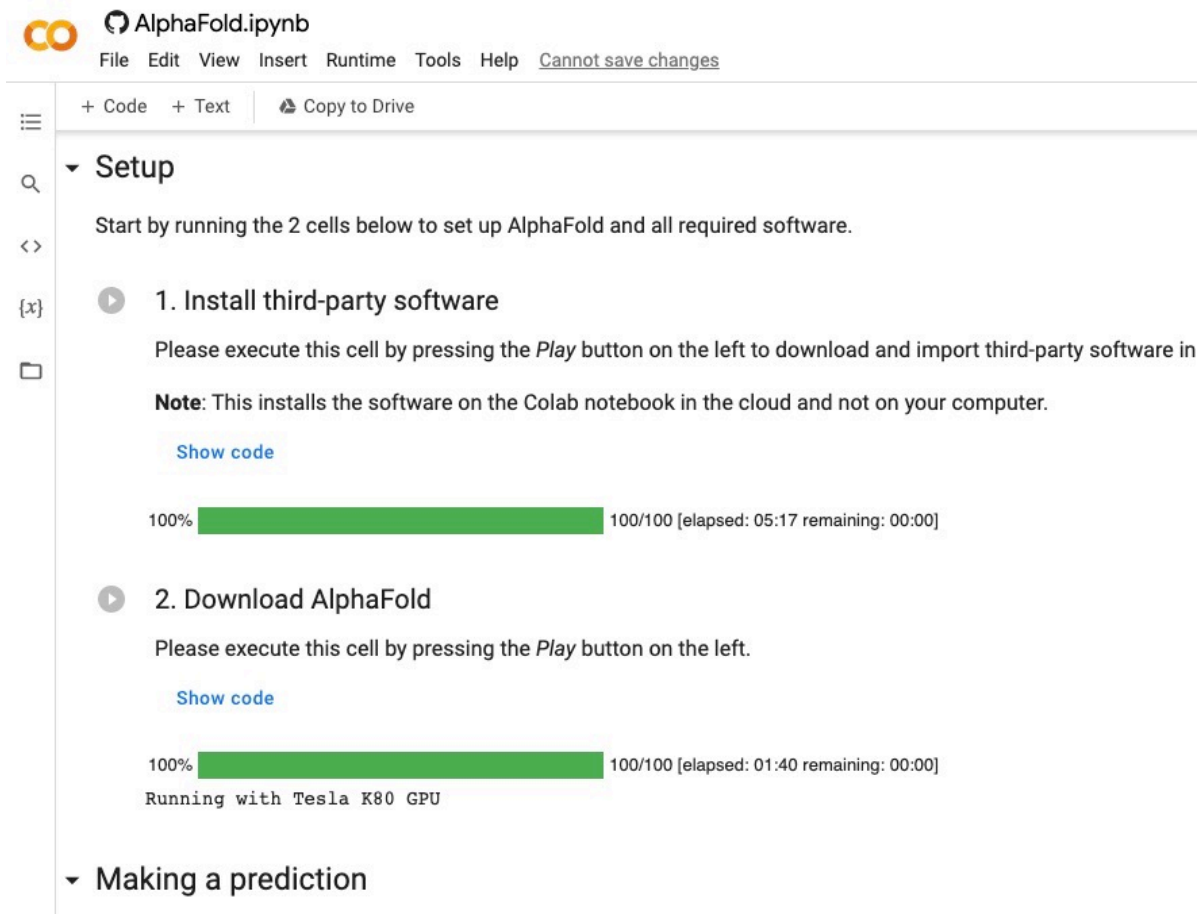
CASP: Critical Assessment of Protein Structure Prediction



Structure Prediction Software

- Rosetta: David Baker, University of Washington
- RaptorX: Jinbo Xu, Toyota Technological Institute at Chicago
- Alpha Fold: DeepMind
- Many others..

Use Alpha Fold online



https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb

Use Alpha Fold Online

3. Enter the amino acid sequence(s) to fold ⬇️

Enter the amino acid sequence(s) to fold:

- If you enter only a single sequence, the monomer model will be used.
- If you enter multiple sequences, the multimer model will be used.

sequence_1: " MAAHKGAEHHHKAAEHHEQAAKHHHAAAEHHEKGEHEQAAHHADTAYAHHKHAEEHAAQAAKHDAEHHAPKPH

sequence_2: " Insert text here

sequence_3: " Insert text here

sequence_4: " Insert text here

sequence_5: " Insert text here

sequence_6: " Insert text here

sequence_7: " Insert text here

sequence_8: " Insert text here

Show code

```
Using the single-chain model.
```

4. Search against genetic databases

Once this cell has been executed, you will see statistics about the multiple sequence alignment (MSA) that will be used by AlphaF

Show code

```
Getting MSA for sequence 1
Searching mgnify: 100% 147/147 [elapsed: 29:00 remaining: 00:00]
58 unique sequences found in uniref90 for sequence 1
110 unique sequences found in smallbfd for sequence 1
9 unique sequences found in mgnify for sequence 1
```

5. Run AlphaFold and download prediction

Once this cell has been executed, a zip-archive with the obtained prediction will be automatically downloaded to your computer.

In case you are having issues with the relaxation stage, you can disable it below. Warning: This means that the prediction might have distracting sr

run_relax: ☑

Show code

AMBER relaxation: 100% 7/7 [elapsed: 31:30 remaining: 00:00]

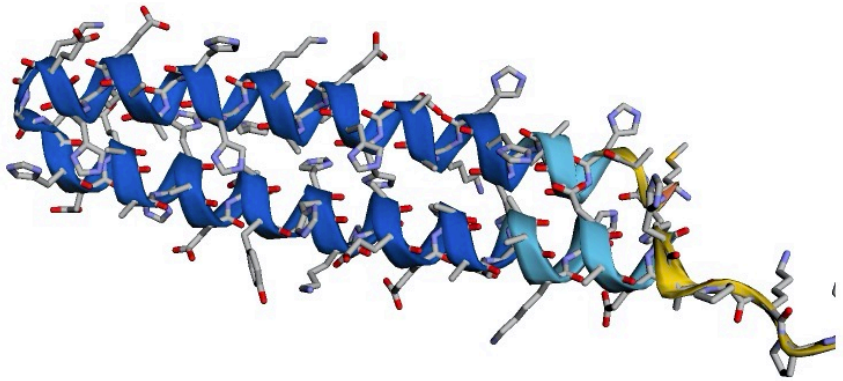
Mod

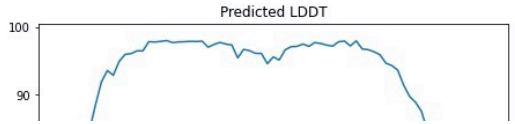
Very lo

Low (7(

Confide

Very hig





Protein Structure Resources

- Structure databases
  - PDB (Protein Data Bank)
    - https://www.rcsb.org/
  - Alpha Fold Predicted
    - https://alphafold.ebi.ac.uk/
- Structure visualization software
  - Pymol