

Highly accurate protein structure prediction with AlphaFold

Jumper, J., Evans, R., Pritzel, A. *et al.*

*Presentation by
Yannick Wattenberg*

What are Proteins?

- Proteins are long chains of amino acids
- They perform many functions within organisms

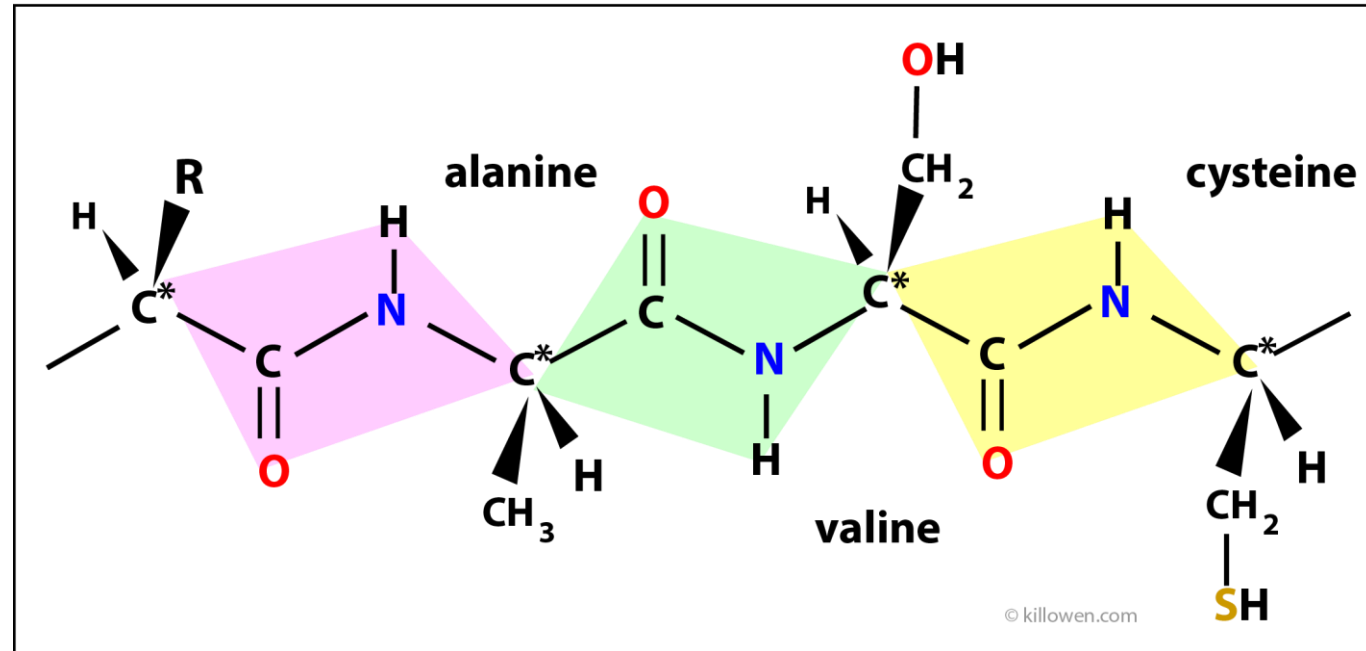


Fig. 1: Primary structure of a protein

Protein Folding

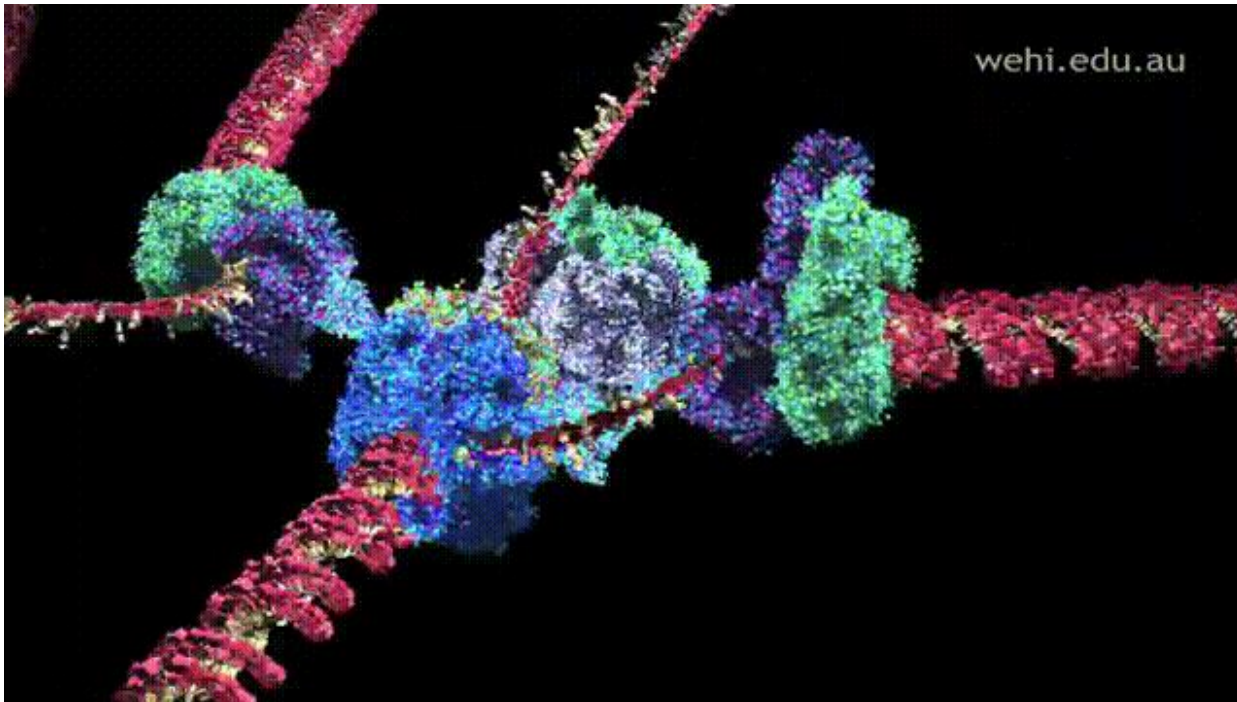


Fig. 25: DNA replication by DNA polymerase

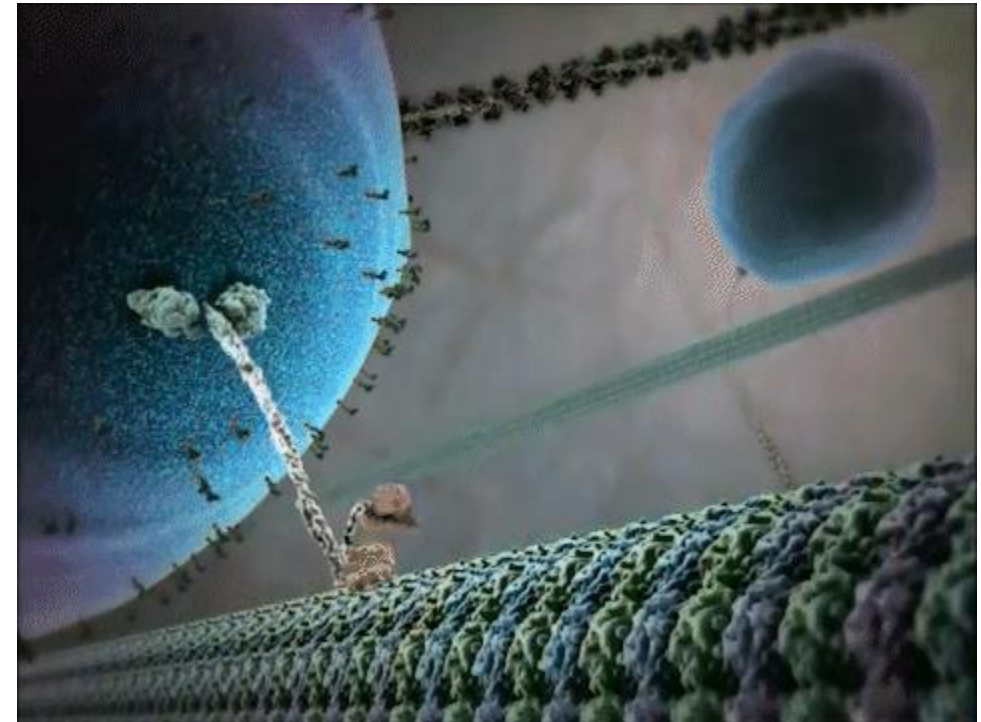


Fig. 2: Kinesin protein walking on microtubule

Protein Folding

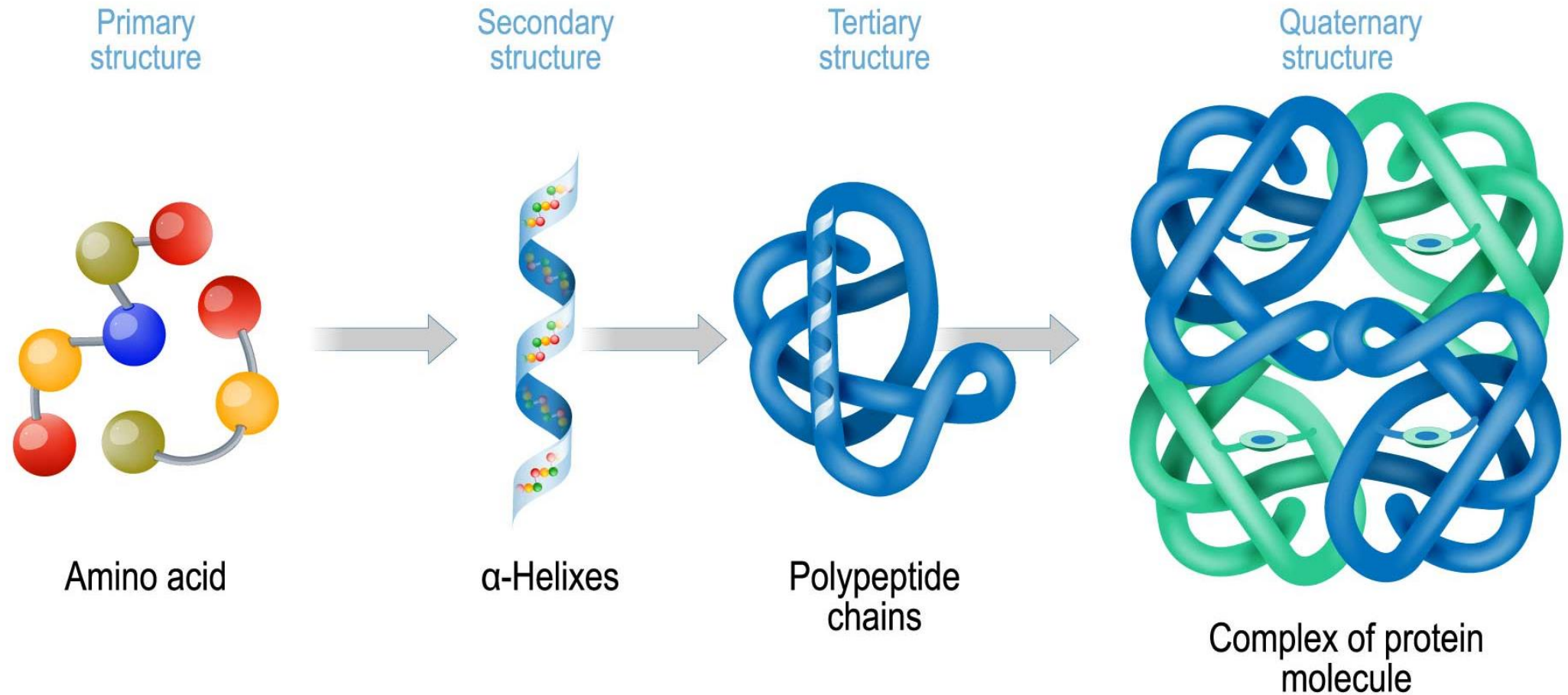


Fig. 3: The different folding structures of proteins

Why prediction?

- Finding the structure of proteins is time intensive
- The process can span years for complex proteins

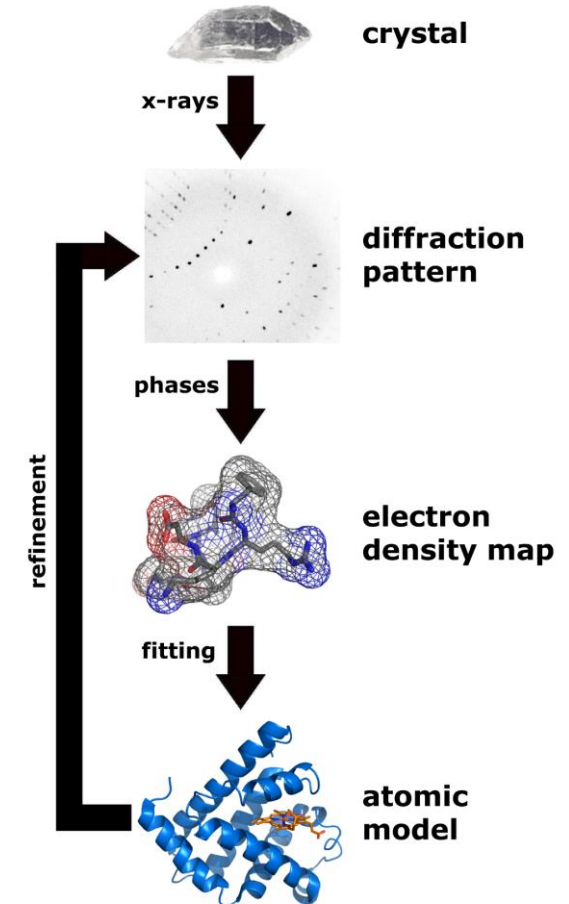


Fig. 4: Experimental process of finding a protein structure

CASP

- Participants are asked to predict the structure of Proteins
- Predictions are made on Proteins with previously unknown structure
- Problem of structure prediction considered solved at a GDT-TD of 90

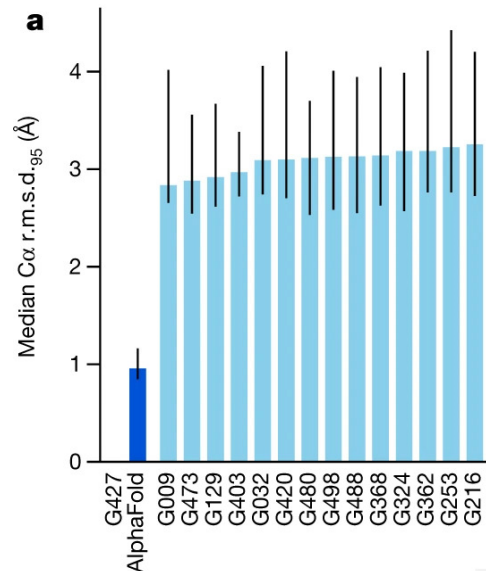


Fig. 5: Performance of AlphaFold2 in CASP14

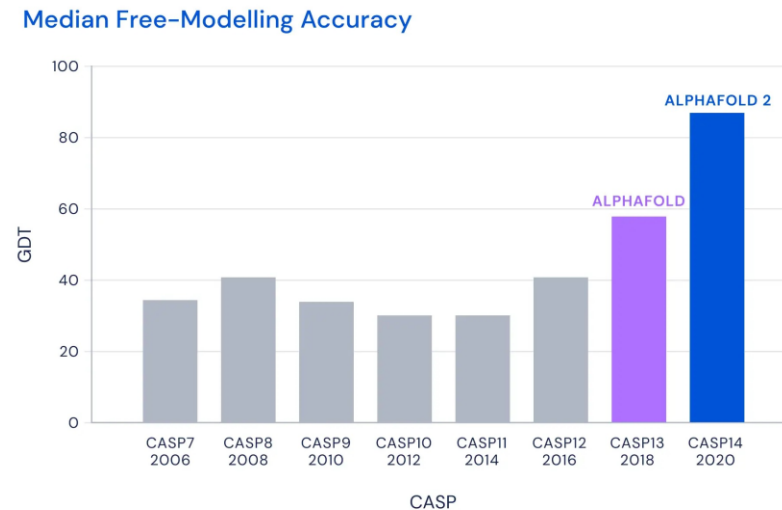


Fig. 6: Performance of AlphaFold2 compared to previous competitions

The Network

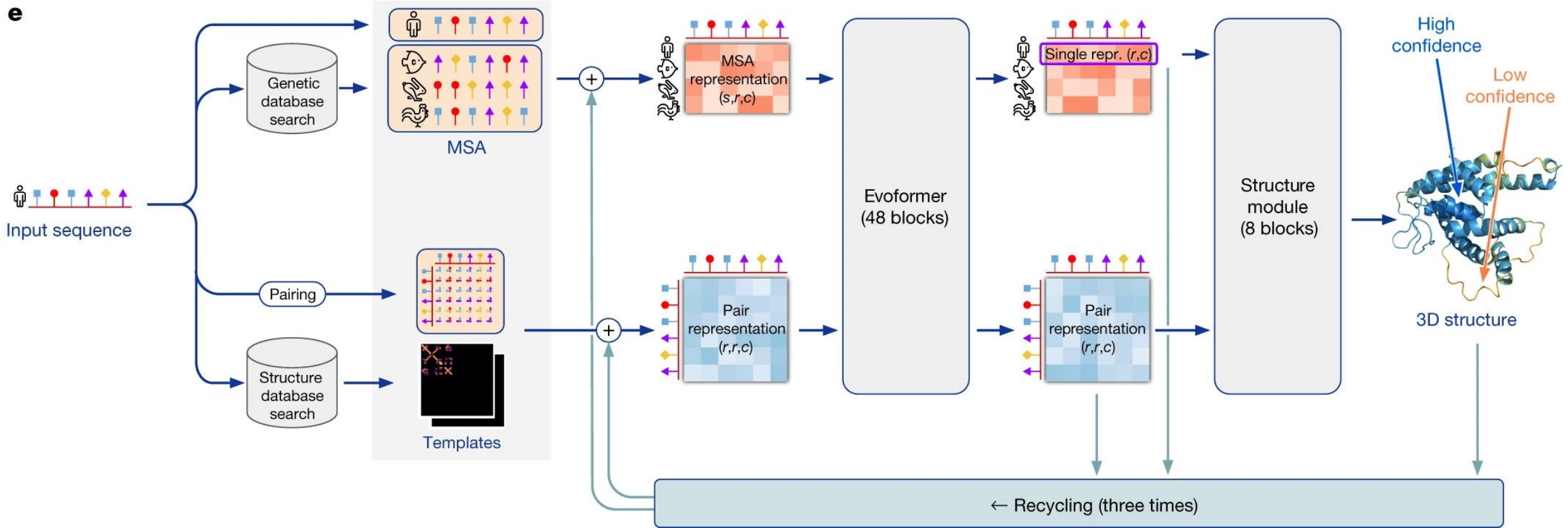


Fig. 7: The Structure of the AlphaFold Network

The Network

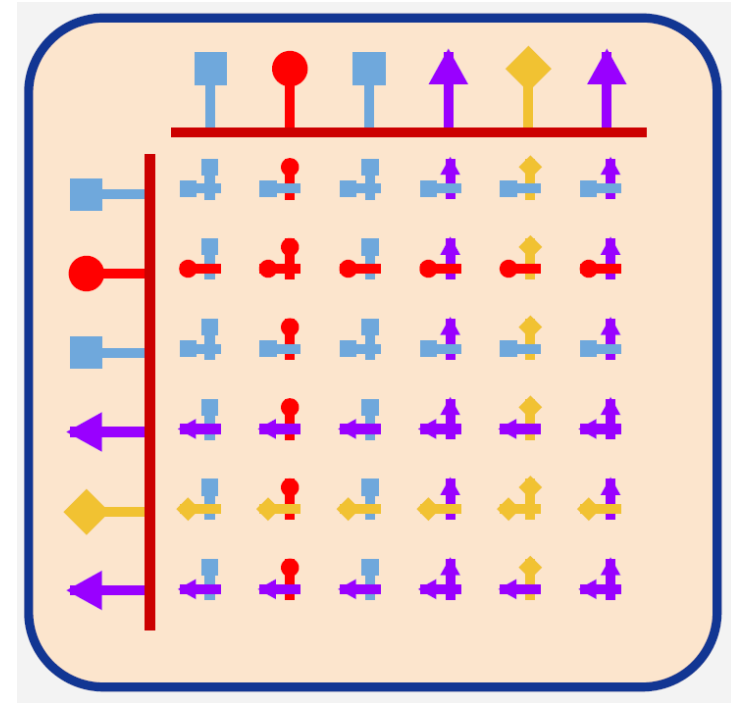
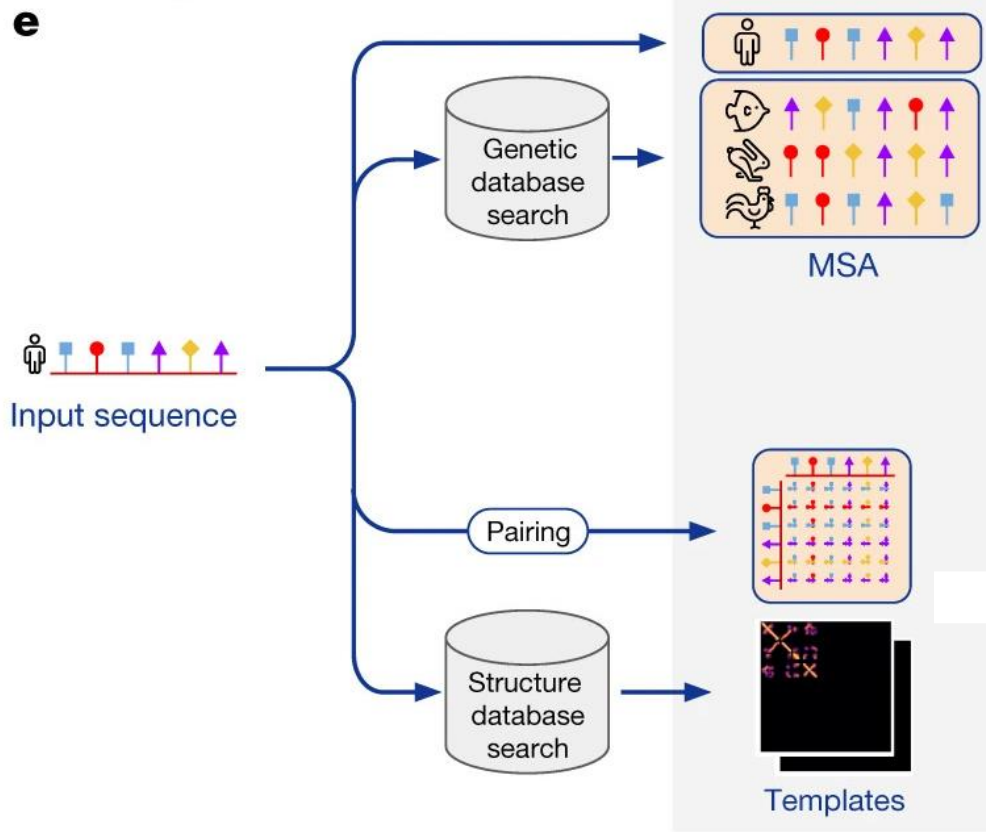


Fig. 7: The Structure of the AlphaFold Network

Fig. 23: The Structure of the AlphaFold Network

Input

- MSA maps evolutionary relationships between residues
- Conserved regions are likely to be functionally important

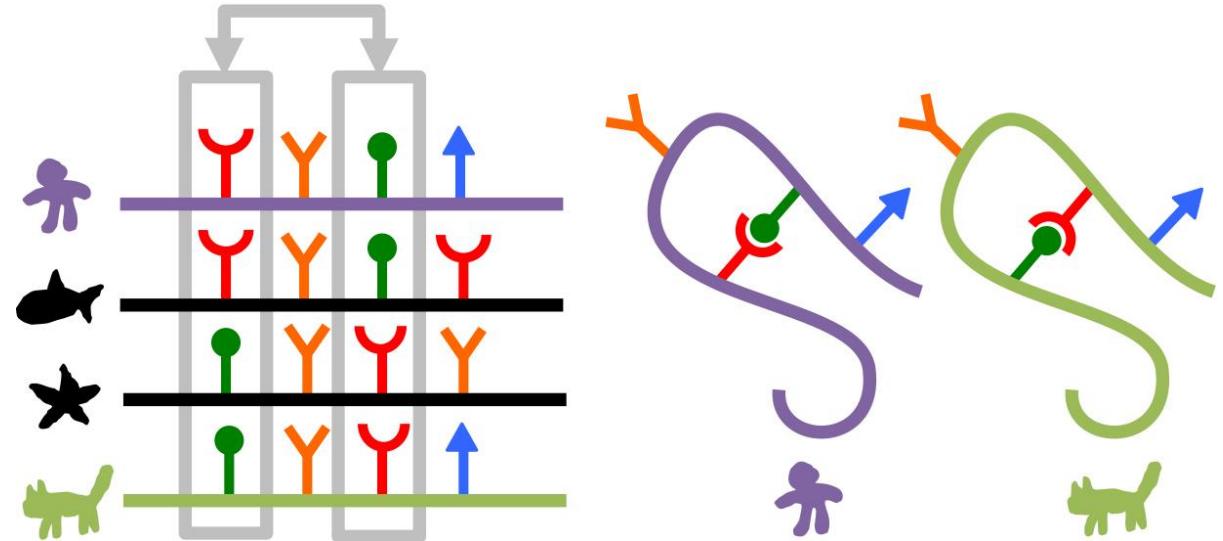


Fig. 8: Illustration of MSA

Input

- The model uses the structures of similar proteins as a basis
- These templates are used to bias the pair representation

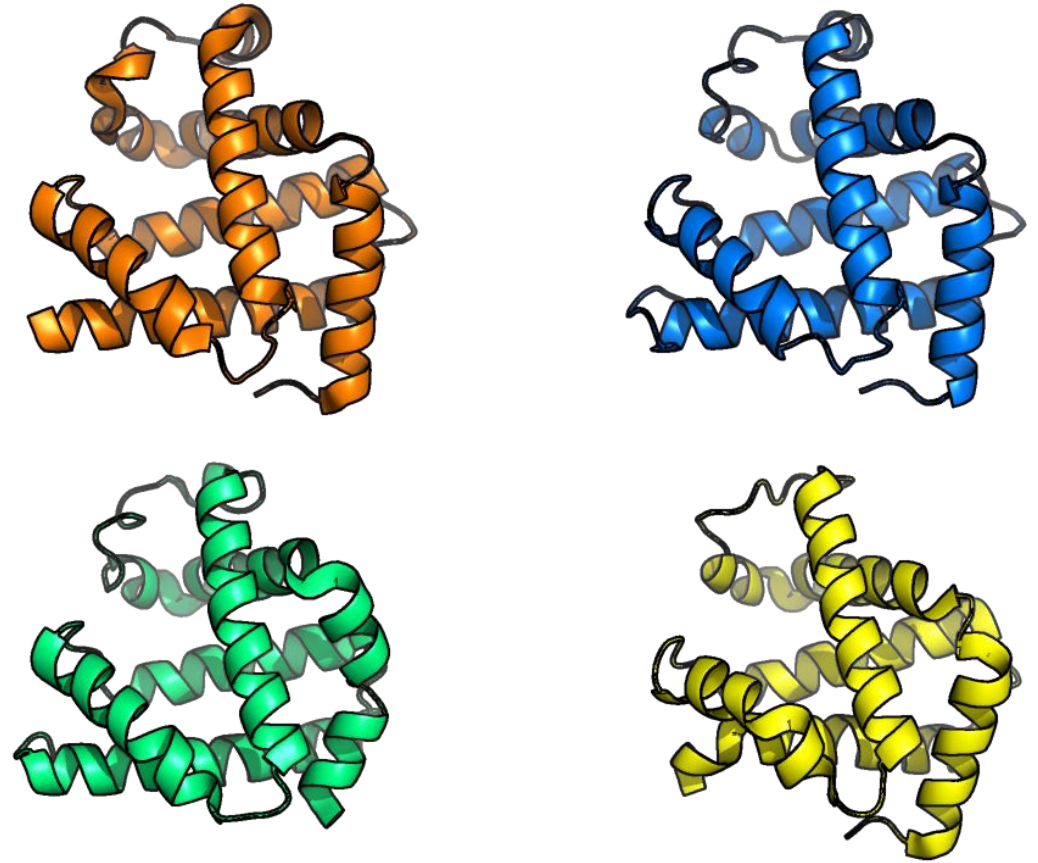


Fig. 9: Examples of myoglobin from different animals

Evoformer

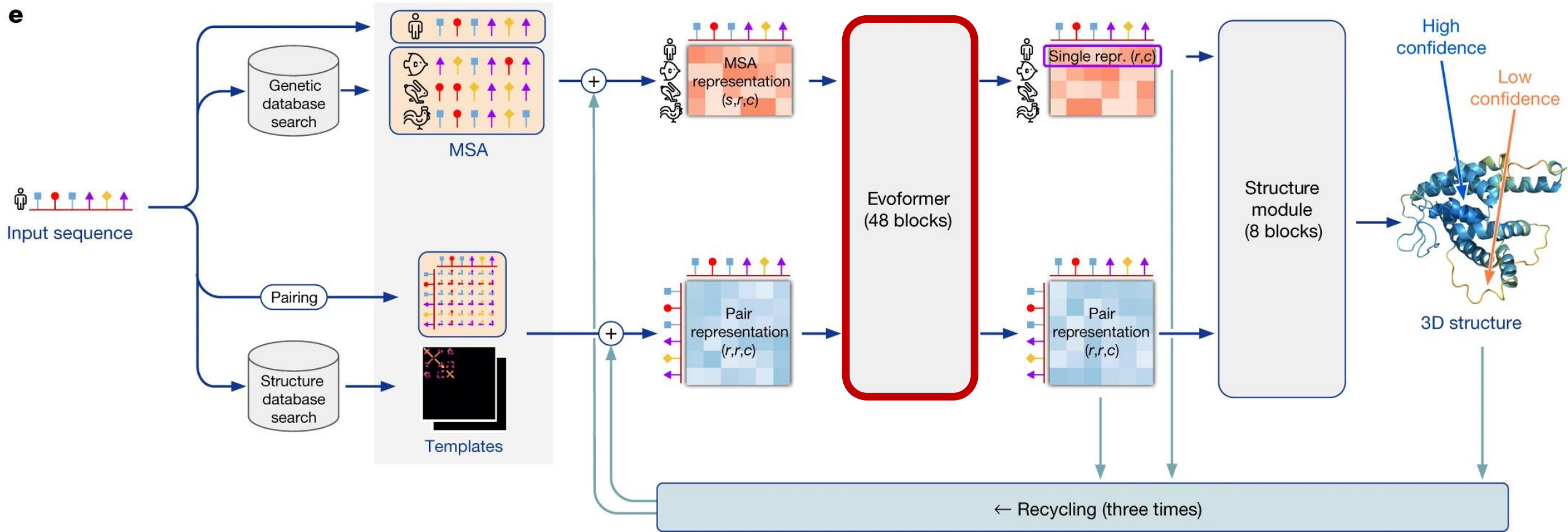


Fig. 7: The Structure of the AlphaFold Network

Evoformer

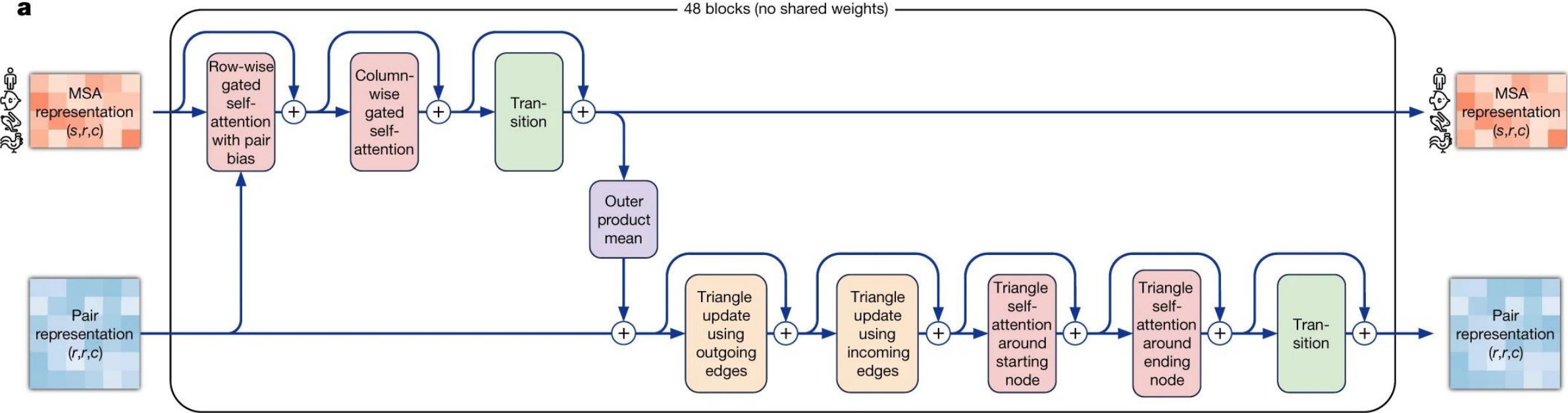


Fig. 10: The Structure of the Evoformer block

Row/Column-wise Attention

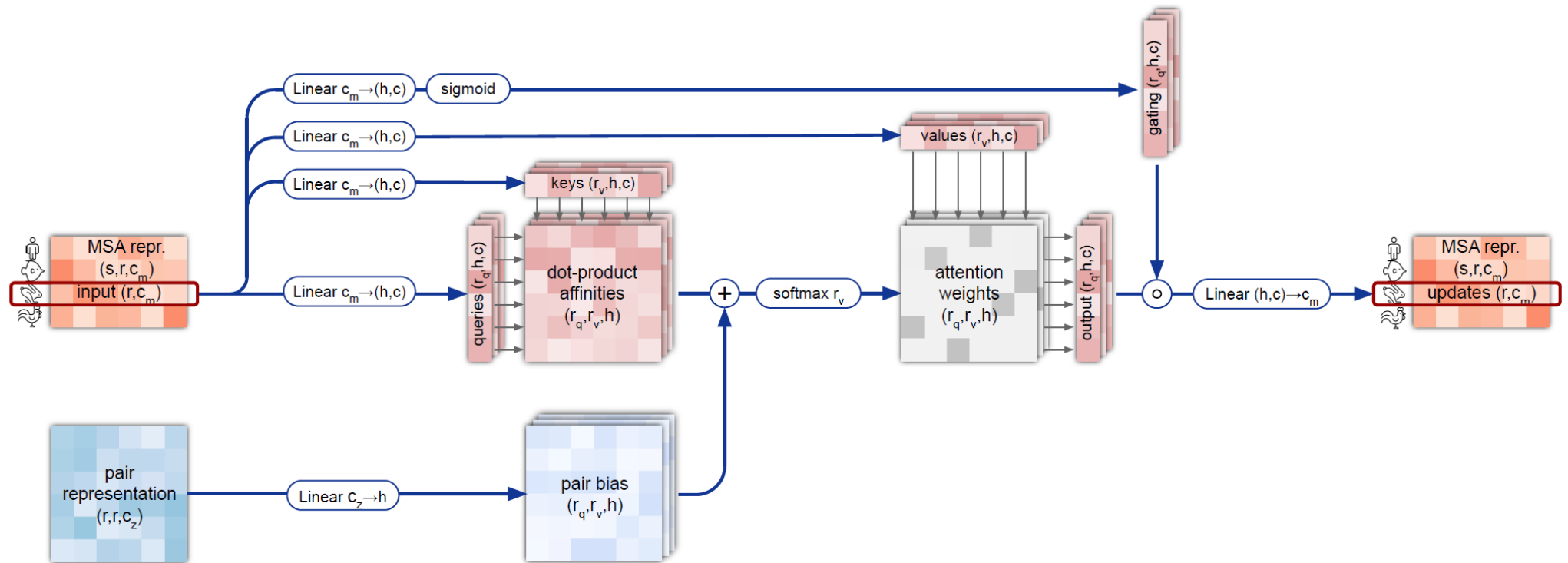


Fig. 11: Row attention block

Evoformer

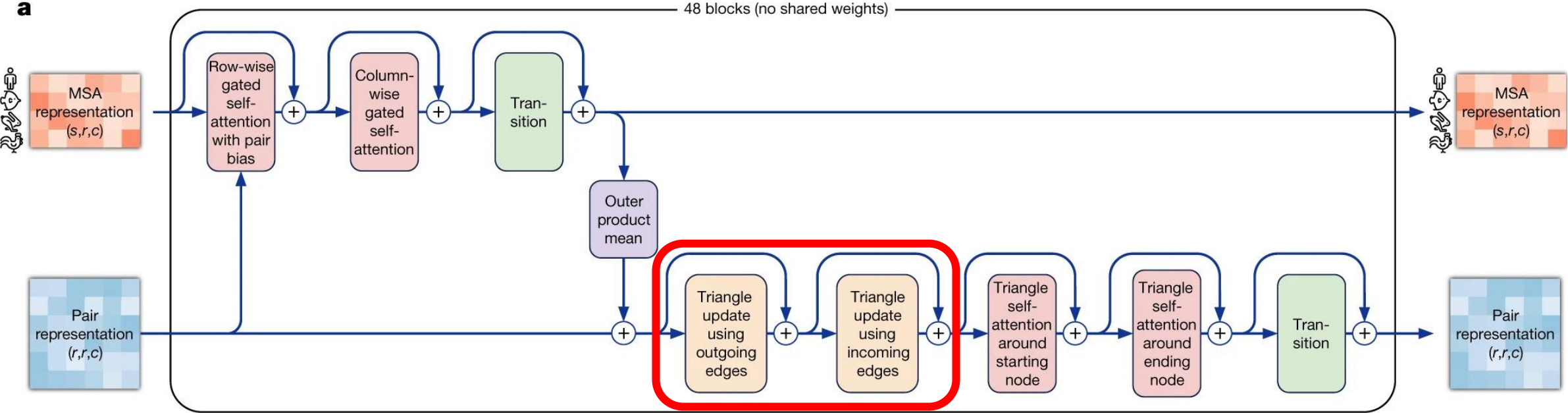


Fig. 10: The Structure of the Evoformer block

Triangle updates

- Basic idea: adhere to triangle inequality
- Update an edge based on all the triangles it is involved in

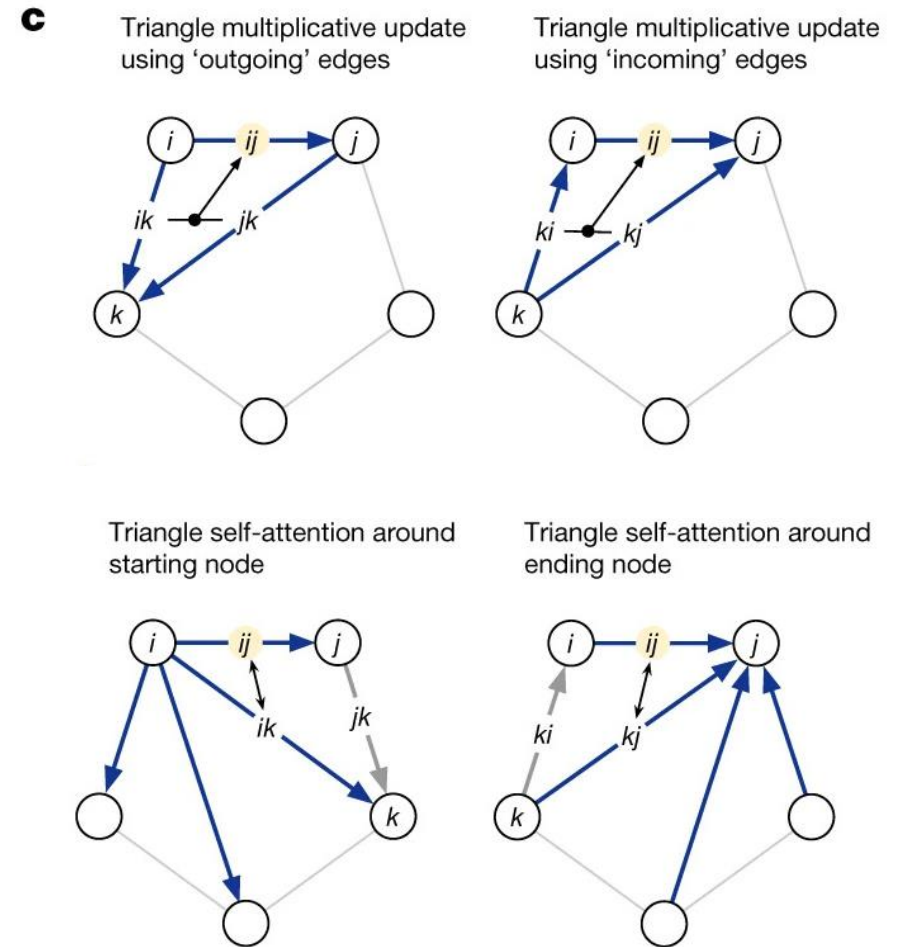


Fig. 12: Illustration of the Triangle Blocks

Structure module

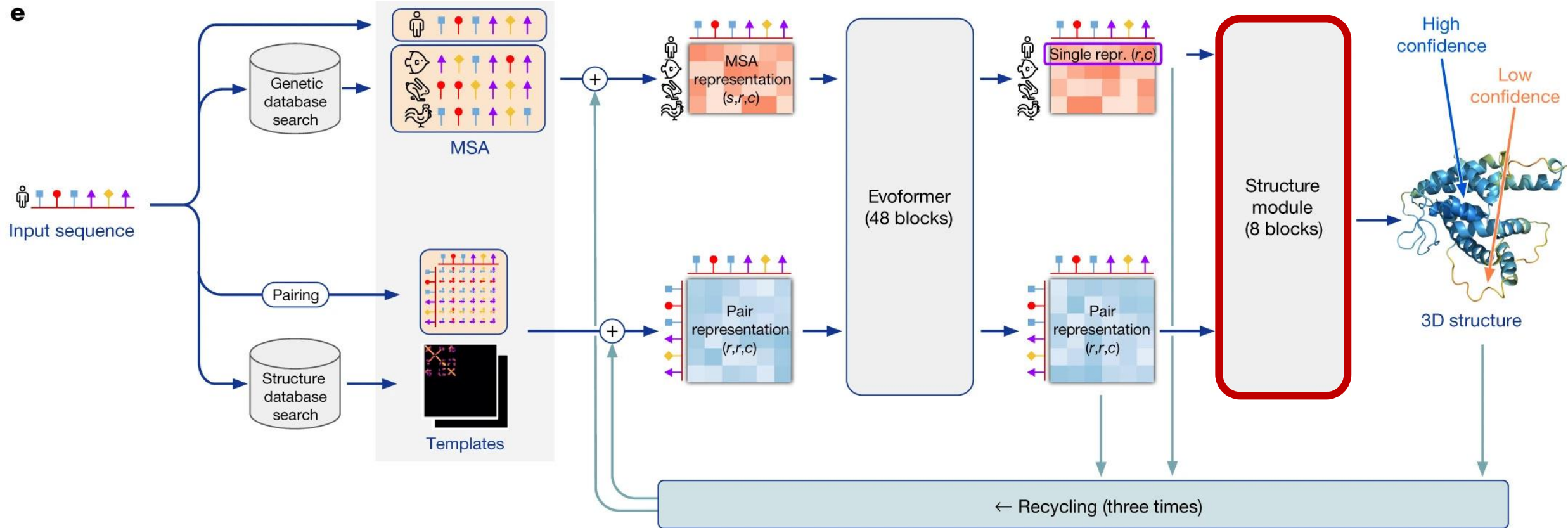


Fig. 7: The Structure of the AlphaFold Network

Structure module

- Protein backbone modeled as 3D gas of triangles
- No enforcement of chain

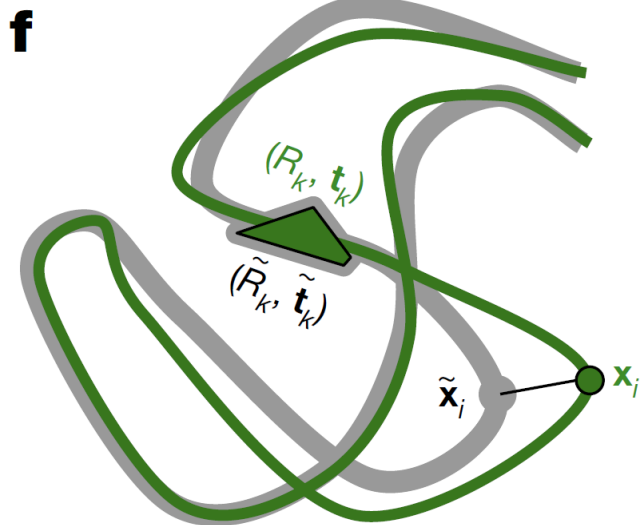


Fig. 24: Impact of small changes on the backbone

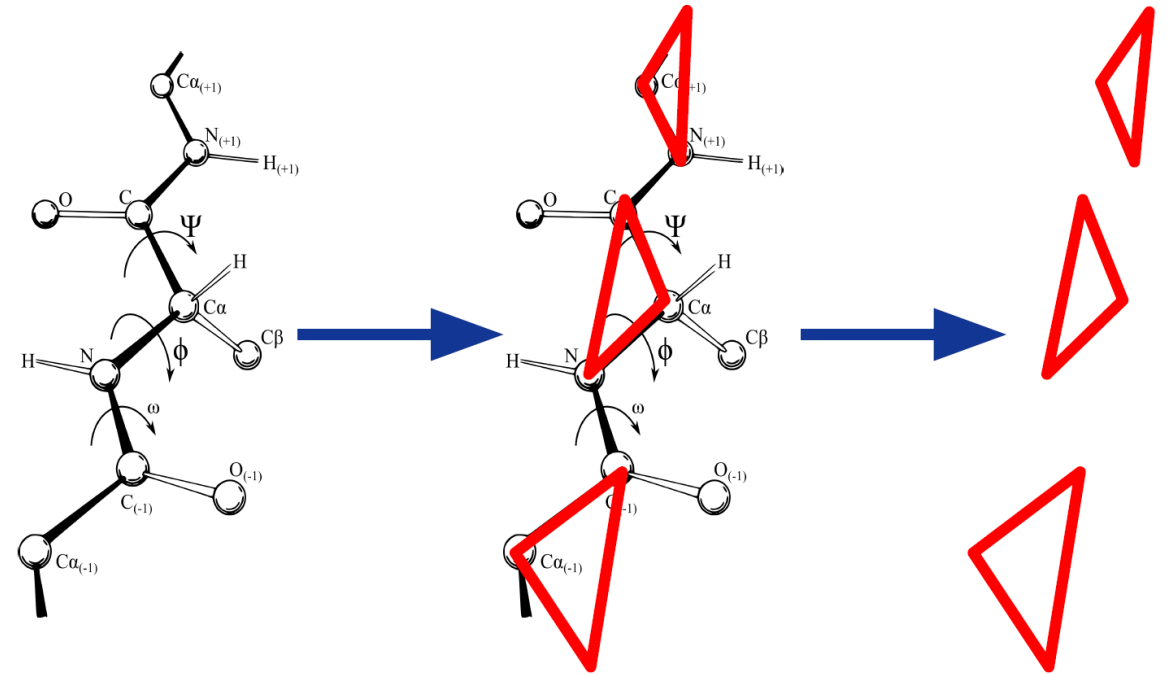


Fig. 13: Illustration of the ridged body gas

Structure module

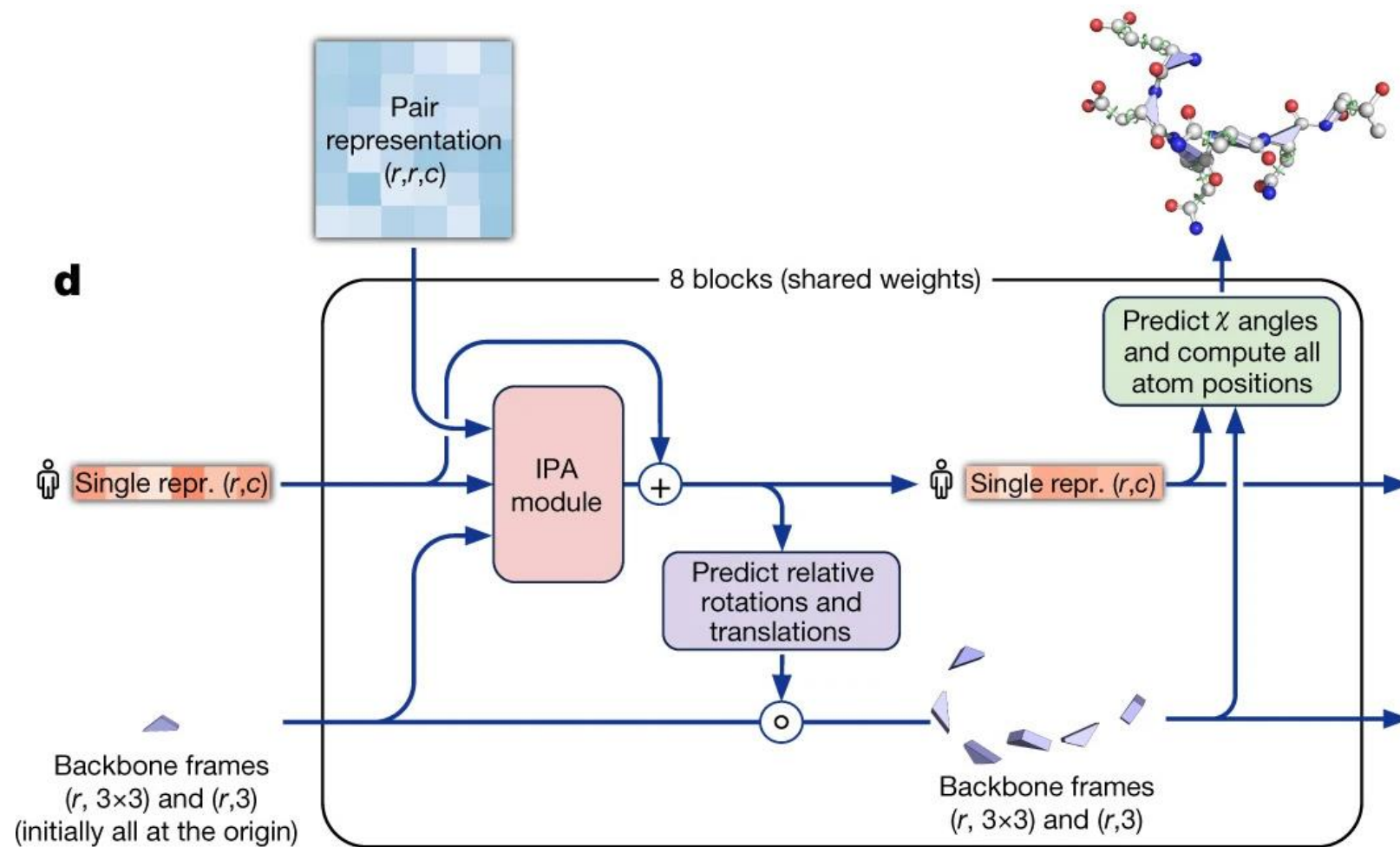


Fig. 14: Structure module

Invariant Point Attention

- 3D equivariant transformer updates triangles (position/rotation)
- Coordinates in local frame are invariant in respect to global frame

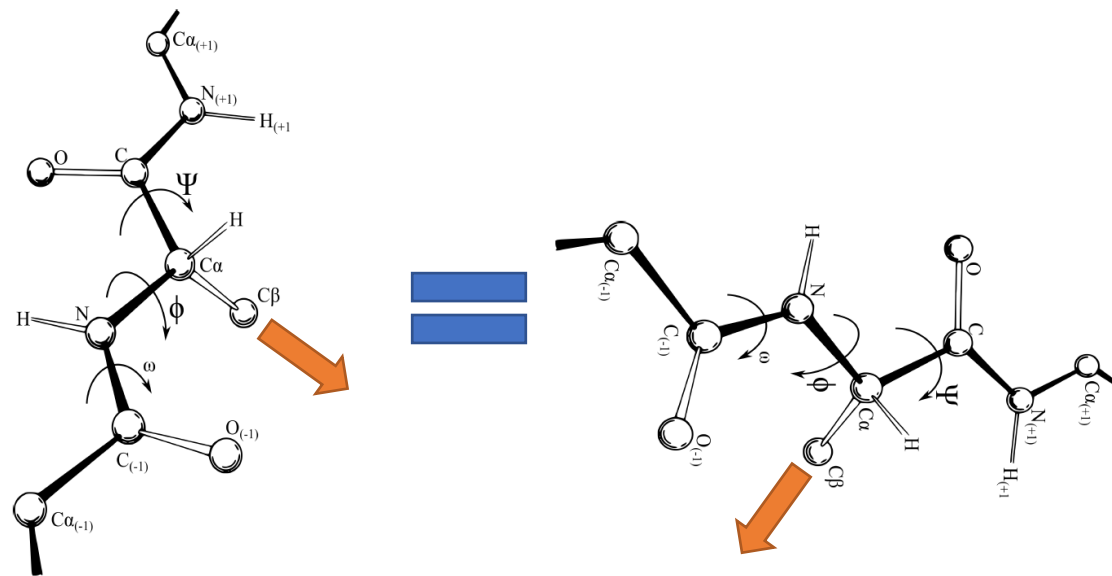


Fig. 13: Illustration of equivariance w.r.t global location

Residue prediction

- The atom location of the side chains can be parametrized using only four additional angles

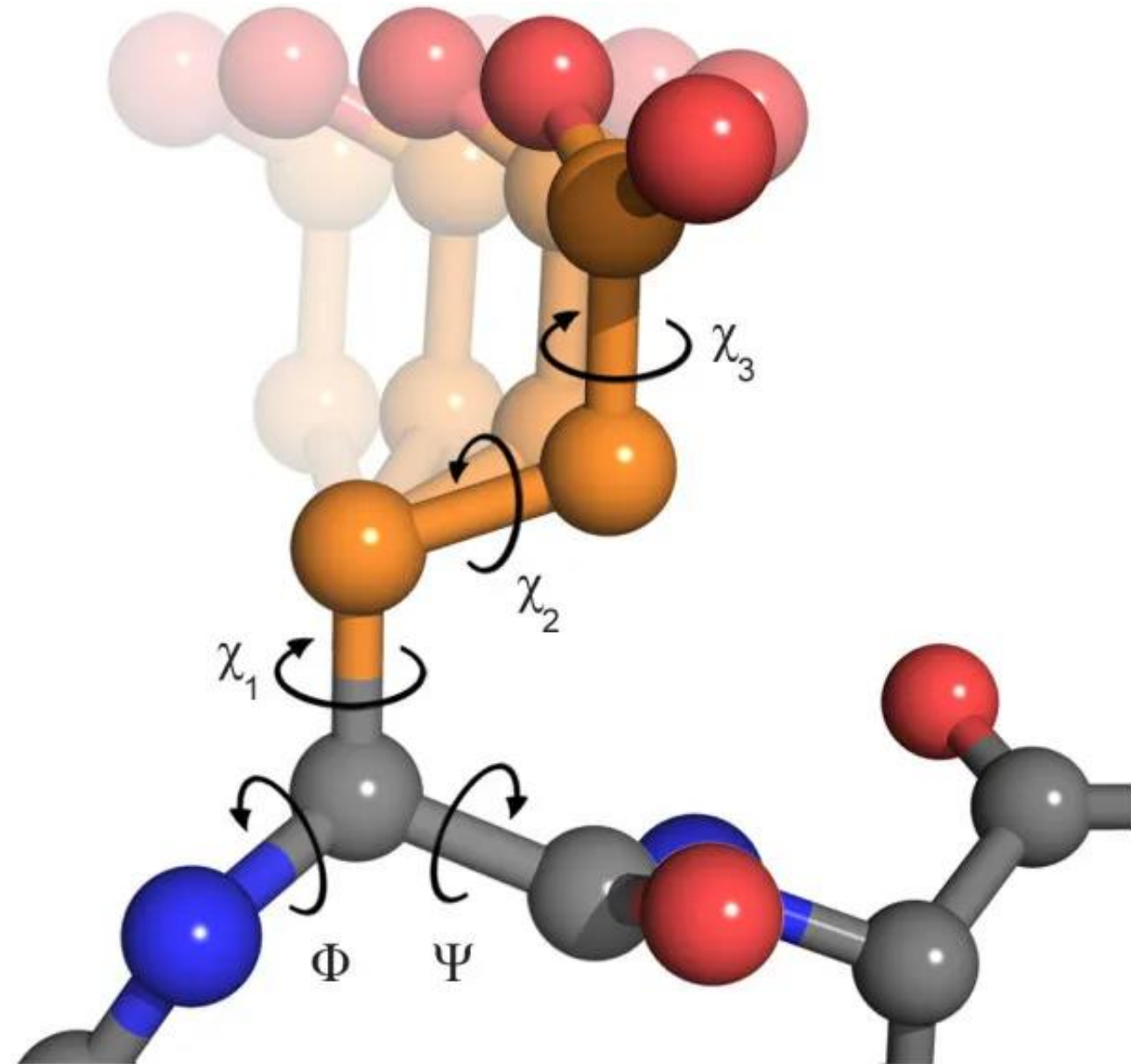


Fig. 15: Defining rotation angles for protein structure

The Model

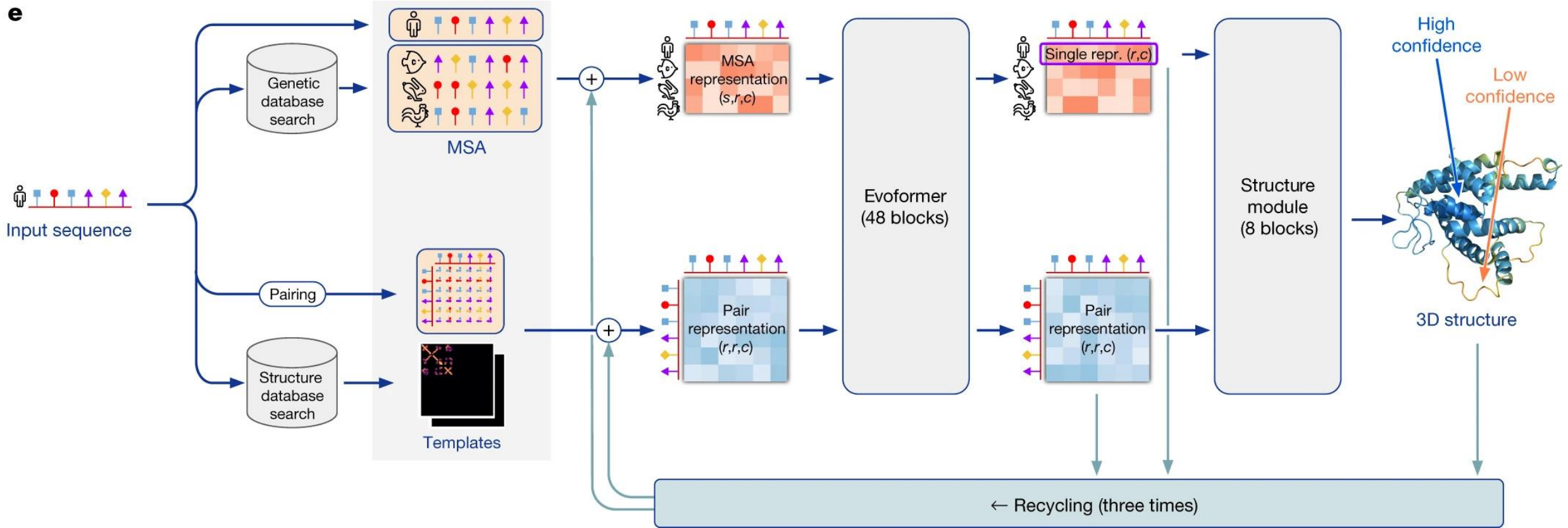


Fig. 7: The Structure of the AlphaFold Network

Physical correction



Recycling iteration 0, block 01
Secondary structure assigned from the final prediction

Model Confidence

- Alpha fold produces confidence measures for each residue
- This is essential for the interpretability of the produced structures

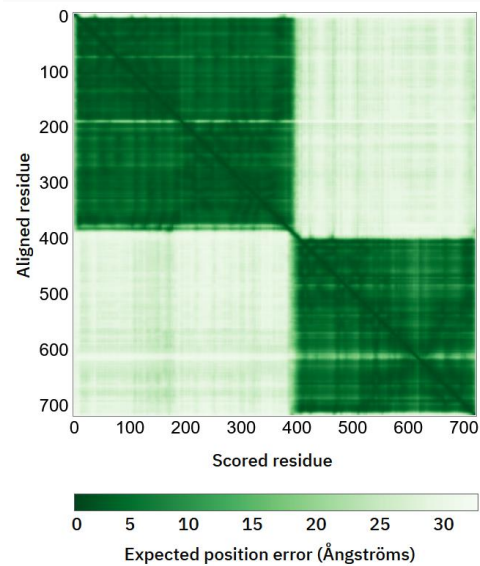
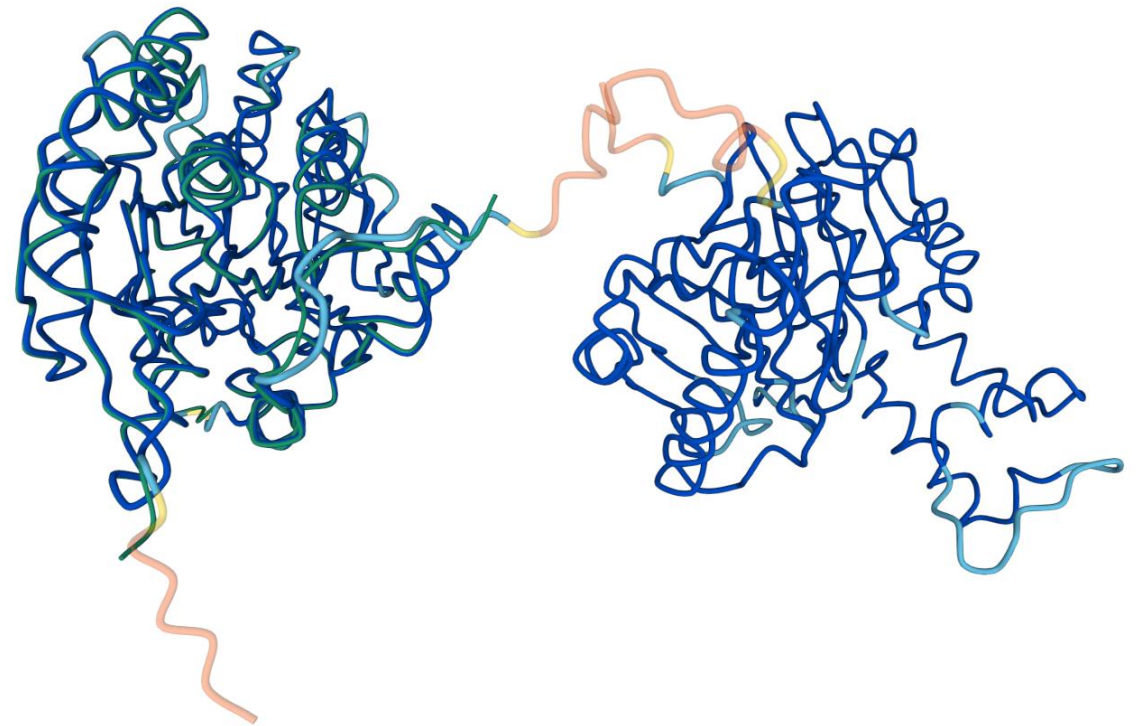


Fig. 17: Confidence Matrix



<https://alphafold.com/entry/Q9Y223>

Fig. 18: Alpha Fold prediction overlaid with experimental structure (green)

Results

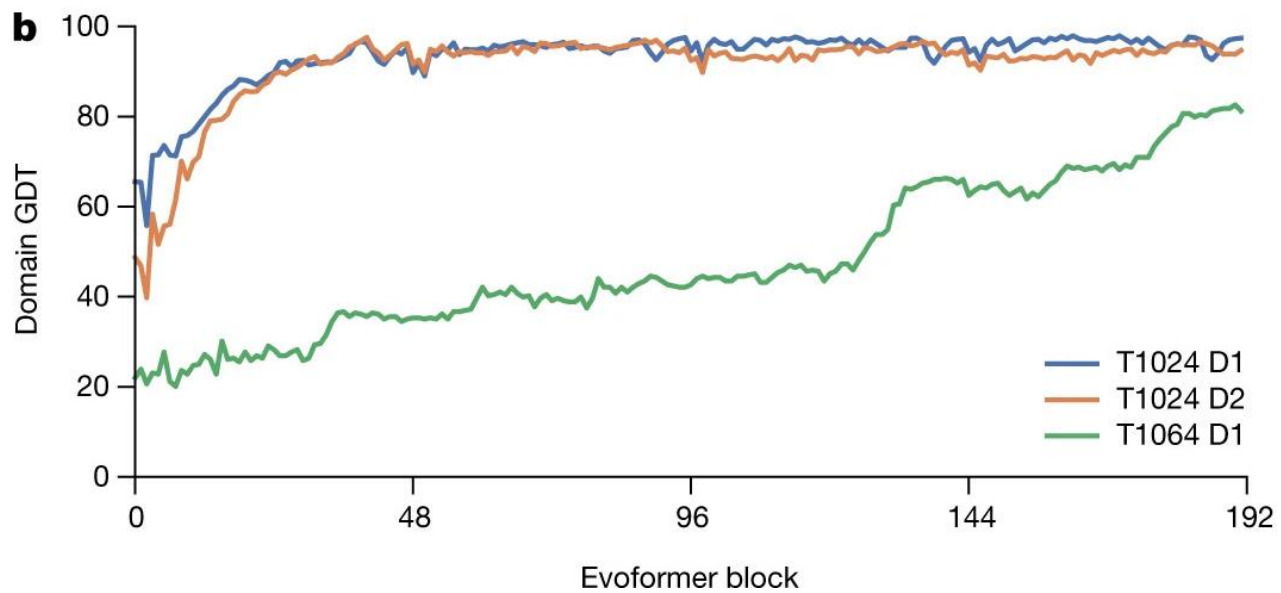
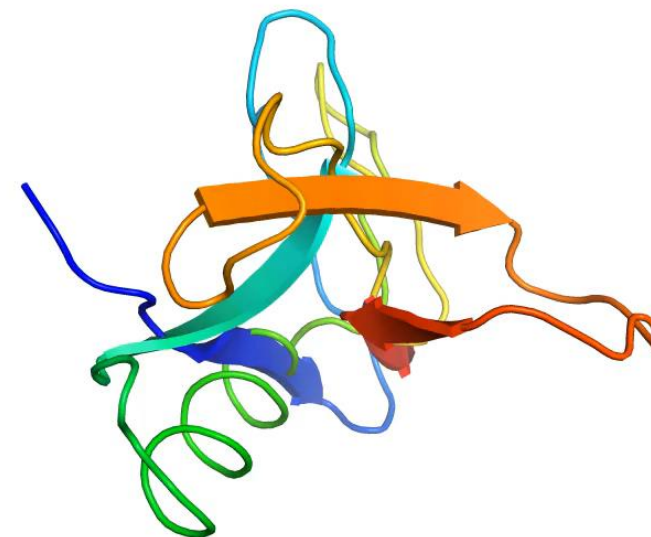


Fig. 20: GDT against number of Evoformer blocks



Recycling iteration 0, block 01
Secondary structure assigned from the final prediction

Fig. 21: Folding of T1064

Results

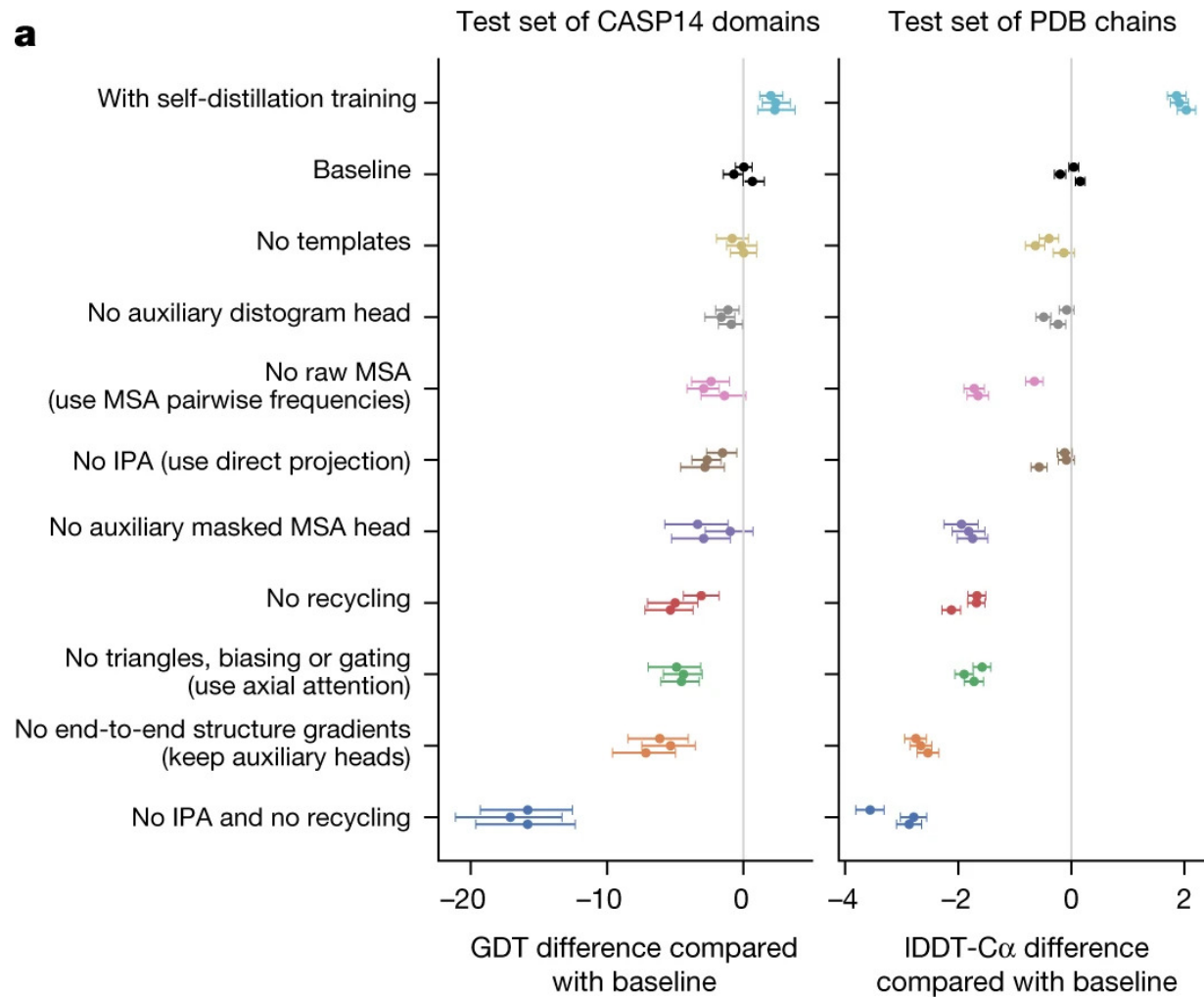


Fig. 22: Structure of SARS-CoV-2 ORF3a

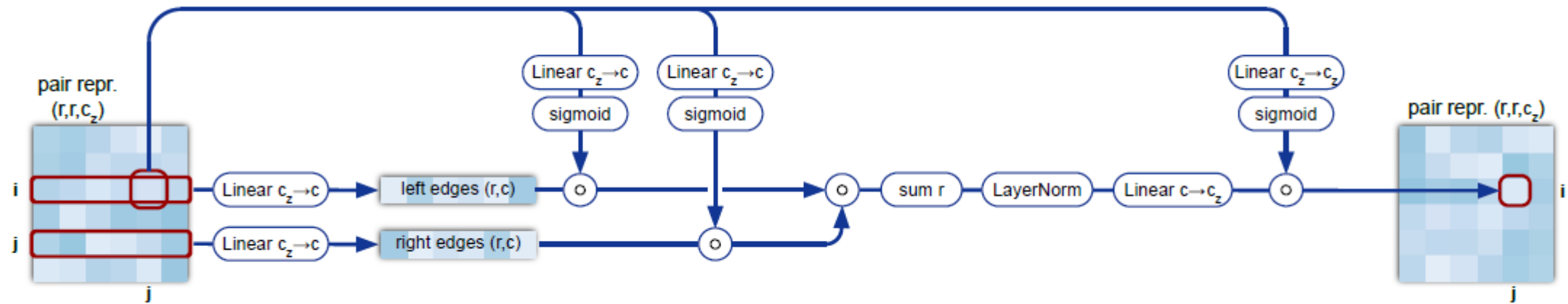
Sources

- <https://www.britannica.com/science/protein>
- [Tim Green: Highly accurate protein structure prediction with AlphaFold – YouTube](#)
- [AlphaFold Protein Structure Database](#)
- [Putting the power of AlphaFold into the world’s hands](#)
- [AlphaFold: a solution to a 50-year-old grand challenge in biology](#)
- [AlphaFold reveals the structure of the protein universe](#)
- [Timeline of a breakthrough](#)
- [Highly accurate protein structure prediction with AlphaFold | Nature](#)
- [AI@MIT Alphafold 2 Reading Group Presentation: Paper Deep Dive – YouTube](#)
- [bayesgroup.github.io/Figurnov_Alphafold.pdf at master · bayesgroup/bayesgroup.github.io · GitHub](https://github.com/bayesgroup/bayesgroup.github.io)
- [OpenFold2](#)
- [Multiple sequence alignment modeling: methods and applications | Briefings in Bioinformatics | Oxford Academic](#)
- [AlphaFold 2 is here: what’s behind the structure prediction miracle | Oxford Protein Informatics Group AlphaFold 2: Attention Mechanism for Predicting 3D Protein Structures | Artificial Intelligence & Software | PI IP LAW <https://alphafold.com/entry/Q9Y223>](#)
- [Invariant Point Attention in Alphafold 2 | by Jude Wells | Medium](#)
- https://www.nature.com/articles/s41586-019-1923-7.epdf?author_access_token=Z_KaZKDgtKzbE7Wd5HtwI9RgN0jAjWel9jnR3ZoTv0MCcgAwHMgRx9mvLjNQdB2TIQqaa7l420UctGo8vYQ39gg8IFWR9mAZtvsN_1PrccXflbc6e-tGSgazNL_XdtQzn1PHty21qdcxV7Pw-k3htw%3D%3D

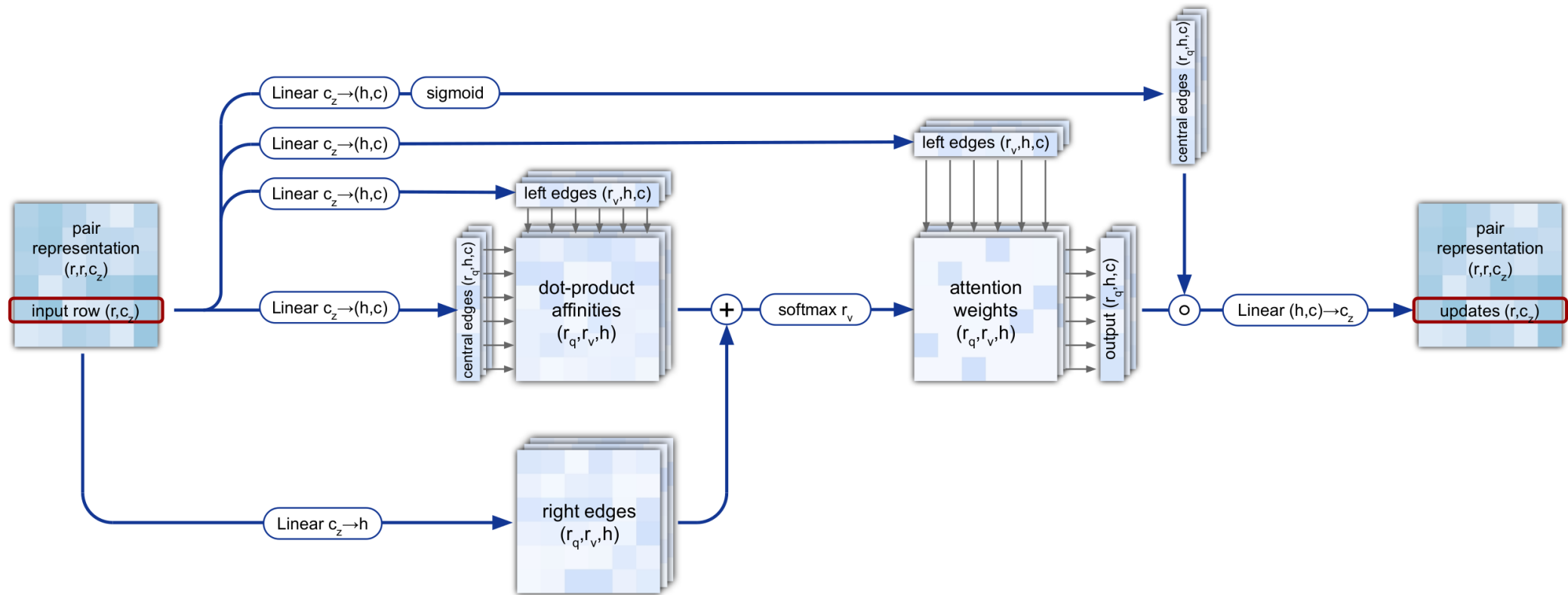
Image sources

- Fig. 1: <https://www.killowen.com/assets/amino%20acid8.png> 18.03
- Fig. 2: “Cellular Visions: The Inner Life of a Cell” By Beth Marchant / July 20, 2006
- Fig. 3: <https://martinswellness.com/media/wysiwyg/blog/protein-structure-1920x1200px.jpg> 11.03
- Fig. 4: https://en.wikipedia.org/wiki/X-ray_crystallography#/media/File:X_ray_diffraction.png 11.03
- Fig. 5: Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold (fig. 1)
- Fig. 6: <https://www.deepmind.com/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology> 11.03
- Fig. 7: Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold (fig. 1)
- Fig. 8: <https://jgi.doe.gov/seeking-structure-metagenome-sequences/cartoon-coevolution-sergey-o/>
- Fig. 9: <https://i0.wp.com/www.blopig.com/blog/wp-content/uploads/2021/07/myoglobin-1.png> 14.03
- Fig. 10: Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold (fig. 3)
- Fig. 11: Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold (sup. fig. 3 (Additional Material))
- Fig. 12: Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold (fig. 3)
- Fig. 13: https://github.com/bayesgroup/bayesgroup.github.io/blob/master/bmml_sem/2021/Figurnov_Alphafold.pdf 11.03
- Fig. 14: Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold (fig. 3)
- Fig. 15: <https://i0.wp.com/www.researchgate.net/profile/Jes-Frellsen/publication/44651362/figure/fig4/AS:267569261183037@1440804891902/Dihedral-angles-in-glutamate-Dihedral-angles-are-the-main-degrees-of-freedom-for-the.png?w=625&ssl=1> 11.03
- Vid. 16: [Supplementary Video 4](#) (Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold)
- Fig. 17: <https://alphafold.ebi.ac.uk/entry/Q9Y223> 11.03
- Fig. 18: <https://www.ebi.ac.uk/pdbe/pdbe-kb/proteins/Q9Y223> 11.03
- Fig. 19: <https://www.rcsb.org/3d-view/6XDC/1> 11.03
- Fig. 20: Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold (fig. 4)
- Fig. 21: https://static-content.springer.com/esm/art%3A10.1038%2Fs41586-021-03819-2/MediaObjects/41586_2021_3819_MOESM5_ESM.mp4 (Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold)
- Fig. 22: Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold (fig. 4)
- Fig. 23: Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold (fig. 1)
- Fig. 24: Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold (fig. 3)
- Fig. 25: <https://www.youtube.com/watch?v=OjPcT1uUZIE> 19.03

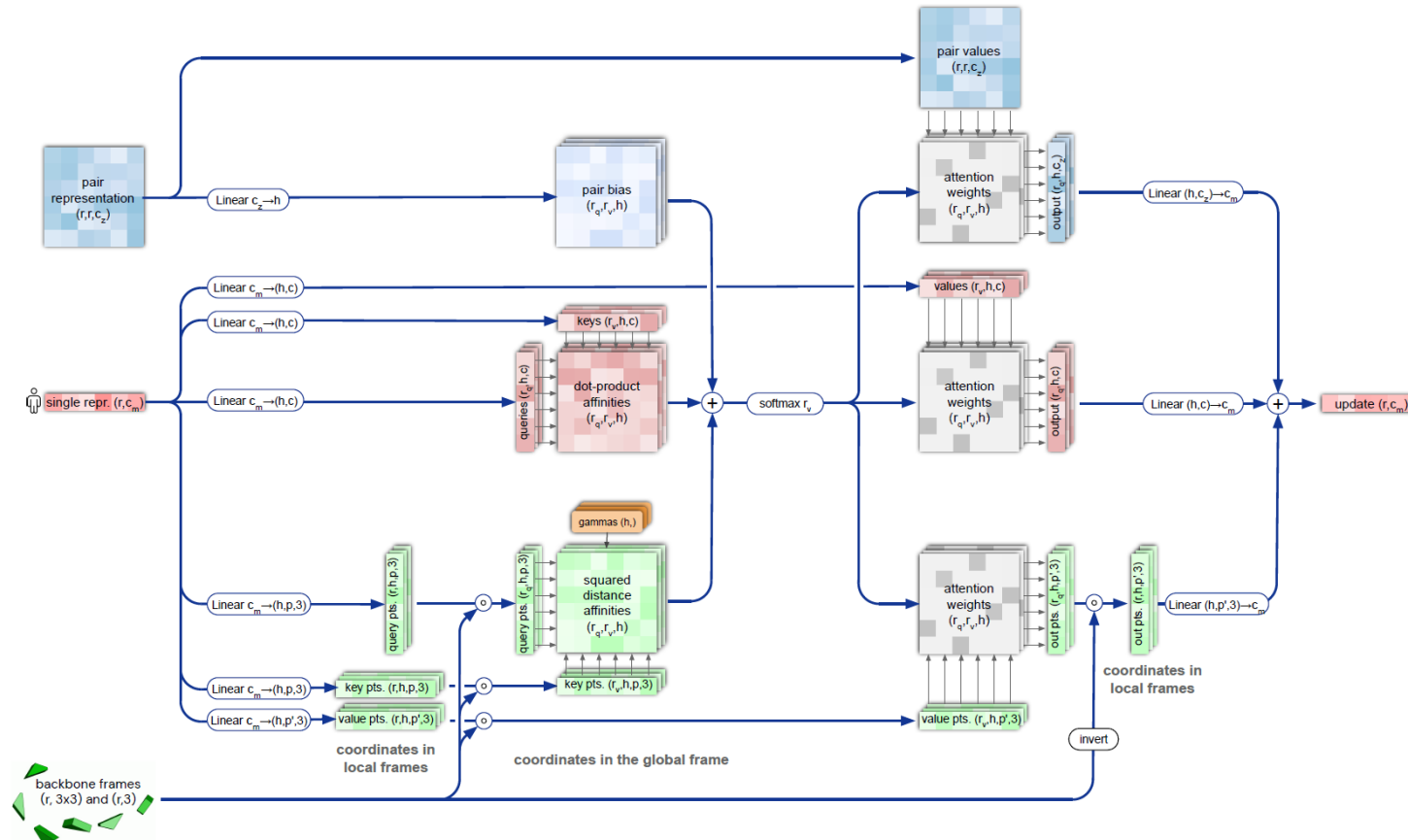
Triangle update



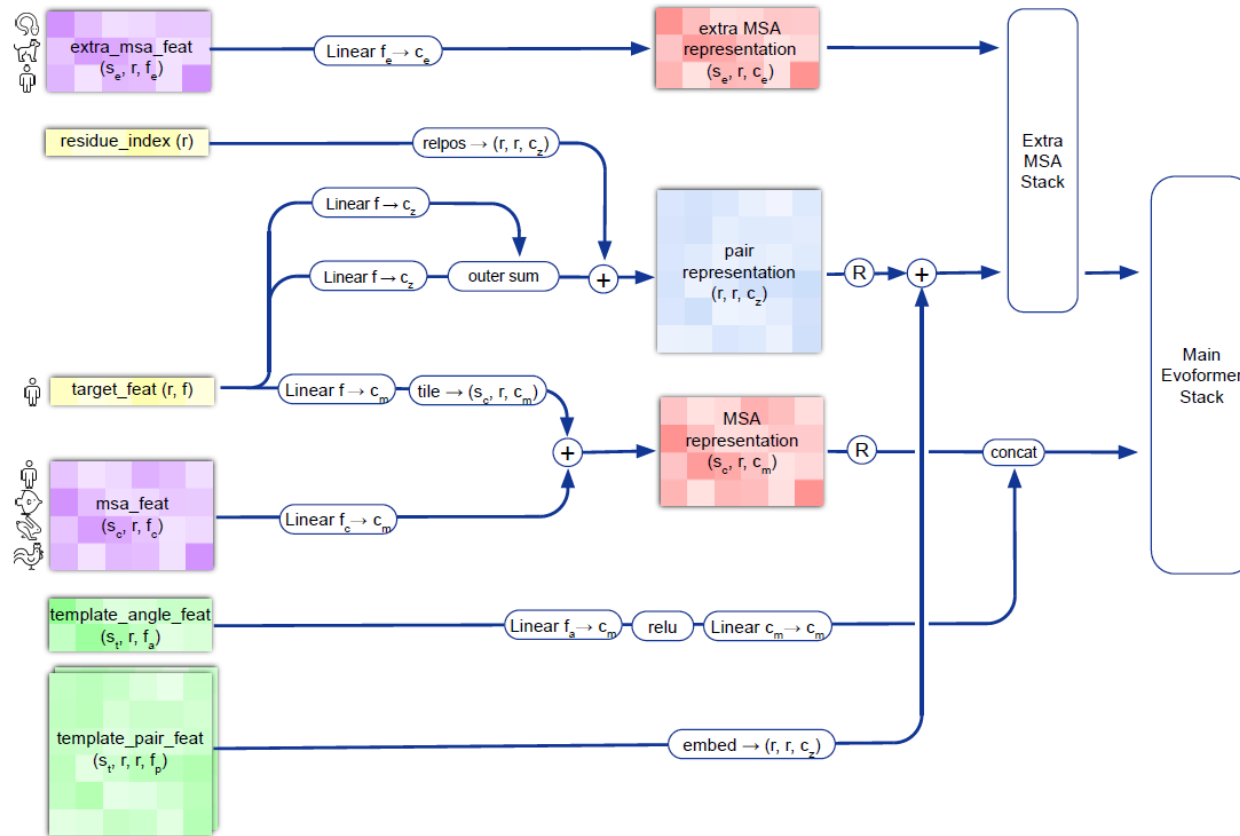
Triangle attention



IPA



Feature embedding



Alpha Fold 1

