



AlphaFold & Protein

Why in News?

Recently, DeepMind, a company based in London, announced that it had predicted the **three-dimensional** structures of more than 200 million **proteins using AlphaFold**.

What do we need to know about AlphaFold?

▪ About:

- **AlphaFold** is an [Artificial Intelligence-based protein structure](#) prediction tool.
- It is based on a computer system called **Deep Neural Network**.
 - **Neural networks** use a large amount of **input data** and provide the desired output exactly like how a human brain would.
 - The real work is done by the **black box** between the input and the output layers, called the **Hidden Networks**.
- AlphaFold is **fed with protein sequences** as input.
 - When protein sequences enter through one end, the **predicted three-dimensional structures** come out through the other.

▪ Mechanism:

- **The first step** uses the available structures of 1,70,000 proteins in the **Protein Data Bank (PDB)** to train the computer model.
 - Then, it uses the results of that training to **learn the structural predictions of proteins not in the PDB**.
 - Then, it uses the **high-accuracy predictions** from the first step to **retrain and relearn** to gain higher accuracy of the earlier predictions.
- By using this method, AlphaFold has now predicted the structures of the entire **214 million unique protein sequences** deposited in the **Universal Protein Resource (UniProt)** database.

▪ Implication:

- Knowing protein structure and function is essential to **understand human diseases**.
- Proteins are generally structured using **x-ray crystallography, nuclear magnetic resonance spectroscopy, or cryogenic electron microscopy**.
 - These techniques are not just time-consuming, **they often take years** and are based mainly on **trial-and-error methods**.
 - AlphaFold makes the process of **protein structuring easier**.
 - It is a watershed movement in science and **structural biology** in particular.
- AlphaFold has already helped hundreds of scientists accelerate their discoveries in **vaccine and drug development** since the first public release of the database nearly a year back.
- **Available Substitutes:**
 - AlphaFold is **neither flawless nor the only AI-based protein structure** prediction tool.
 - **RoseTTaFold**, developed by University of Washington in Seattle, U.S., is another tool.
 - Although **less accurate than AlphaFold**, it can predict the structure of protein complexes.

How should India Benefit from it?

- India needs to quickly take advantage of the AlphaFold database and **learn how to use the structures to design better vaccines and drugs**.
 - Understanding the accurate structures of **Covid-19** virus proteins in days rather than years will **accelerate vaccine and drug development against the virus**.
- India should also speed up implementation of **public-private partnerships** in the sciences.
 - It should facilitate **joint collaborations with the prevalent hardware muscle and data science talent** in the private sector and specialists in academic institutions to pave the way for data science innovations.

UPSC Civil Services Examination, Previous Year Questions (PYQs)

Q. What is the Cas9 protein that is often mentioned in the news? (2019)

- (a) A molecular scissors used in targeted gene editing
- (b) A biosensor used in the accurate detection of pathogens in patients
- (c) A gene that makes plants pest-resistant
- (d) A herbicidal substance synthesized in genetically modified crops

Ans: (a)

Explanation:

- CRISPR-Cas9 is a unique technology that enables geneticists and medical researchers to edit parts of the genome by removing, adding or altering sections of the DNA sequence.
- CRISPR is an acronym for “Clustered Regularly Interspaced Short Palindromic Repeats.”
- Cas9 is basically an enzyme that is used like a pair of scissors to cut two strands of DNA at a specific location to add, remove or repair bits of DNA.
- Therefore, option (a) is the correct answer.

[Source: TH](#)

PDF Refernece URL: <https://www.drishtias.com/printpdf/alphafold-protein>