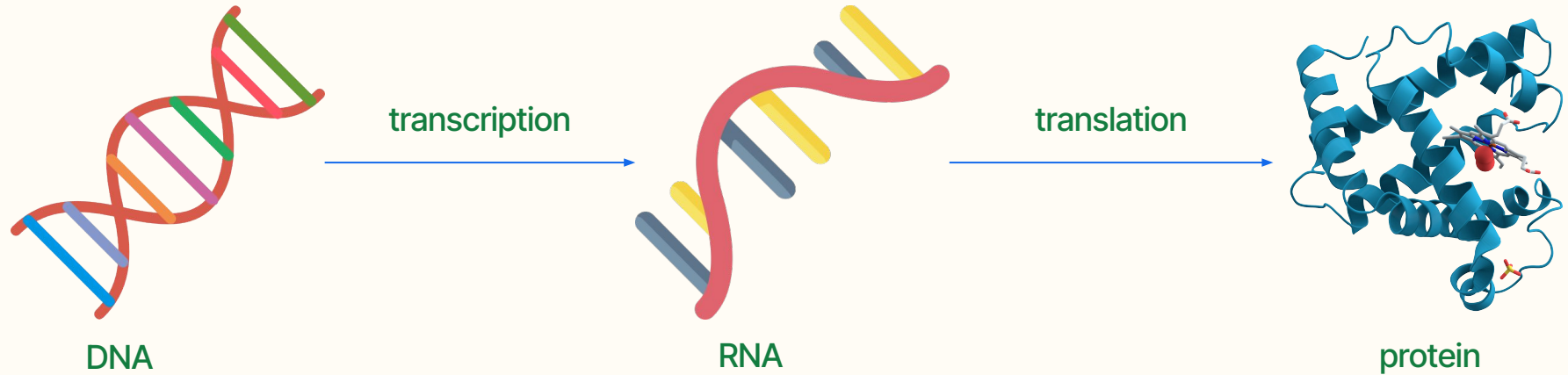# Protein Structure Prediction (AlphaFold)

David, Seth, Grace, Pallavi, Aryan
Group 7
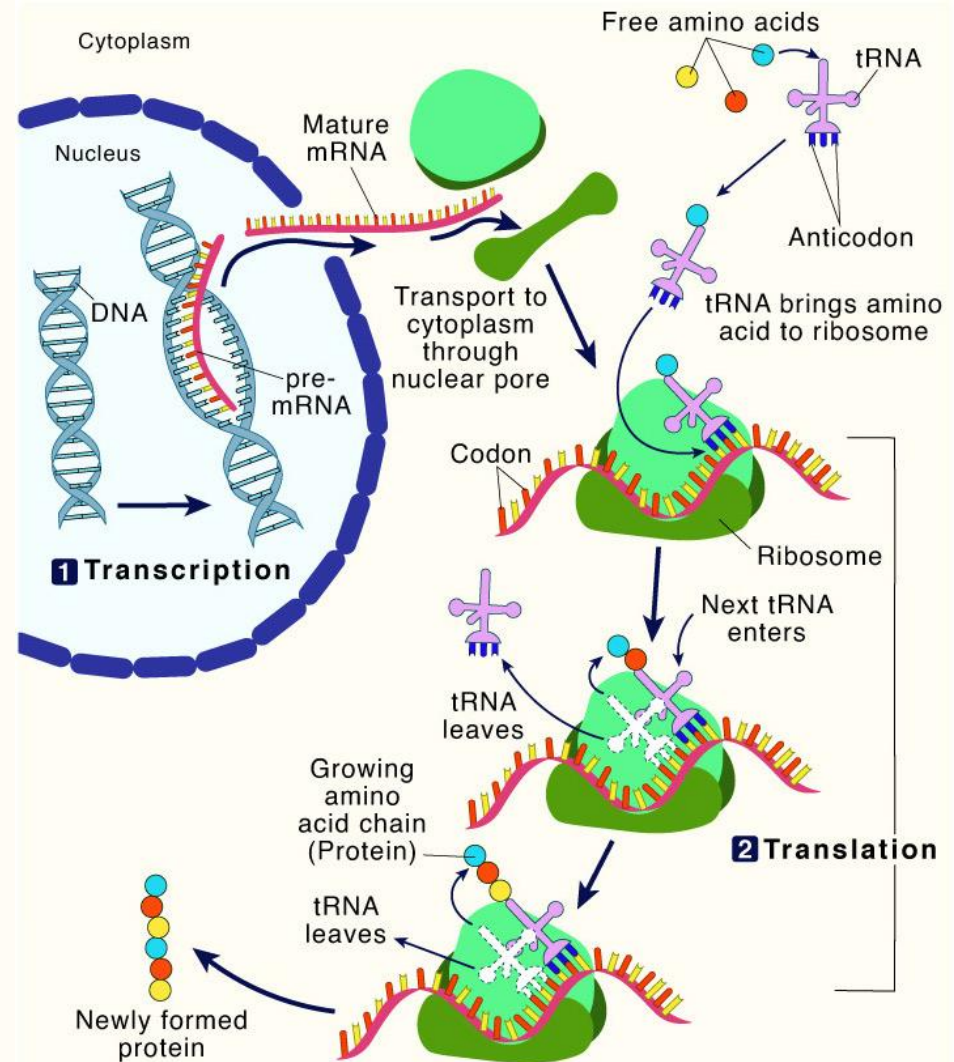
# Poll Question: When was the last time you took a biology class?

# The Central Dogma

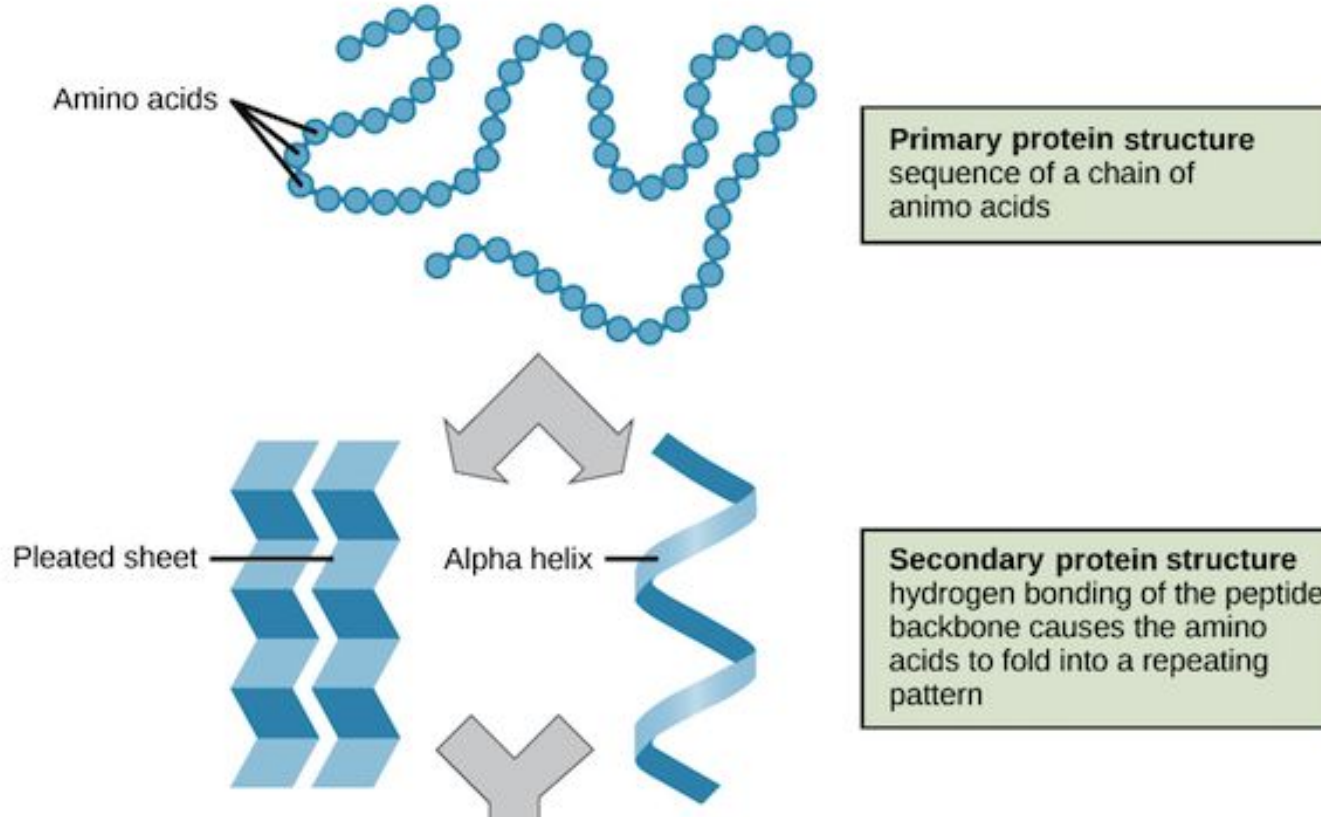DNA → transcription → RNA → translation → protein
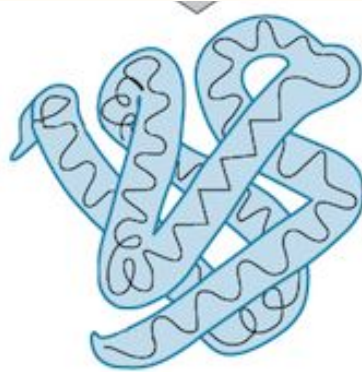
# Protein Synthesis

- Transcription: make mRNA from DNA
  - Uses RNA polymerase to unwind and bind to DNA
- Translation: make protein from mRNA
  - 3 steps:
    - Initiation
    - Elongation
    - Termination
- Protein folding occurs

# Protein Structures



Amino acids

**Primary protein structure**
sequence of a chain of
animo acids

Pleated sheet — Alpha helix —

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

# Protein Structures



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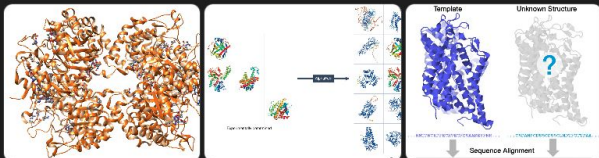three-dimensional folding pattern of a protein due to side chain interactions

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

# Real Impacts of Alpha Fold pt 1

- Usually, the impacts can be a little abstract or hard to understand for people who don't know biology
- For example, this was ChatGPTs first response to "give me examples of discoveries that have a made an impact on the world that were made possible by alphafold":



While this sounds cool, it means nothing to me and I imagine to most non-biology people here.
The next two slides have real life examples of impacts of AlphaFold driven research

# Real Impacts of Alpha Fold pt 2

- Honeybees are dying off at an alarmingly rapid rate
  - "In 2025, US honeybees suffered perhaps the largest colony die-off in the country's history, close to 60% for managed honeybee colonies."
- Breeding of more resilient honeybees is **essential to the survival of honeybees as a species**
- Current breeding cycles takes up to a year; identification of the Vg protein could help shorten breeding cycles to **mere weeks**
- This protein was predicted by AlphaFold 2 in 2021 and confirmed in 2025
  - The researcher claimed that: "It took me two days to do something that could have taken me years."
- Summary: AlphaFold 2 helps researcher predict protein which is crucial in honeybee preservation

# Real Impacts of Alpha Fold pt 3

- Global warming is having an impact on crop harvests of temperature sensitive staple crops
- GLYK is an enzyme in plants crucial to photosynthesis that is hypothesized to stop working when it overheats
- Through AlphaFold predictions, researchers were able to model the protein and realize why it struggled under heat
- Researchers were able to create similar enzymes without the same temperature weakness

# Implications of AI on Scientific Research

Does AI accelerate discovery?

Does it change what scientists study?

Does it actually translate into innovation?

# AlphaFold expanded Scientific Exploration

The consistent association of AlphaFold 2 with more novel protein structures suggests that the tool is used to characterise proteins in less chartered parts of the protein universe. These findings suggest the capacity of AI to open up additional avenues for research within a problem space, and is in line with existing work on the use of AlphaFold 2 (Yu, 2024).

- [Supports the "streetlight effect" discussion] The authors argue AlphaFold may actually push science *into darker territory*, not just easier problems.

- They interpret this as AI enabling exploration of previously inaccessible protein space, rather than reinforcing only well-known structures

# AlphaFold measurably increased productivity and impact

Academic productivity and quality: Links to AlphaFold 2 lead to a modest increase in publication rates for researchers (2.5%) and laboratories (5.1%), similar to other frontier developments. Citation counts for research papers building on AlphaFold 2 and other frontier developments exhibit increases between 25% and 30%. Normalising citation counts by field and year still yields positive associations, with AlphaFold 2 performing strongly for researchers and laboratories.

Key result:

- Labs and researchers using AlphaFold showed:
    - modest productivity gains
    - consistently higher citation impact
- AlphaFold's impact was **slightly stronger and more consistent than other frontier AI tools**

→ AlphaFold doesn't just help individuals; it shifts the whole field's output upward.

# The Nuance: Translational impact exists, but it isn't simple

Idea: AlphaFold shows a "varied and nuanced impact on translational outputs," meaning innovation pathways depend on deeper mechanisms, not just raw prediction accuracy

- Papers, researchers, and labs using AlphaFold all show **higher patent citation rates**
- But disease-focused research doesn't rise uniformly
- Clinical impact varies depending on whether you measure papers vs labs

→ AI helps invention, but doesn't automatically solve applied problems.

# AlphaFold as engineering infrastructure

The authors explicitly frame AlphaFold as engineering infrastructure, not automated discovery:

AlphaFold "does not generate new scientific knowledge," but by removing a bottleneck, it "qualitatively shifts the directionality of research."

AlphaFold:

- lowers cost barriers
- reduces risk of exploring new ideas
- enables efficient screening of massive hypothesis spaces

→ AI becomes a **new method of invention** not merely a faster calculator

# Discuss

## Does AI change what it means to "do science"?

I argue, from the paper, that it does!

AI changes what it means to do science; not by replacing scientists, but by reshaping which questions feel feasible.

Do you agree? Why or Why not?

# Question 2:

## Does AI narrow what "good science" looks like?

"We speculate that AlphaFold 2 lowers the cost barrier and reduces the risks associated with research portfolio diversification"

"There is concern that the uptake of AI will lead to scientists overly focusing on problems that are suited to AI"

"Debates cente[r] on whether AI will ultimately unlock substantial productivity gains and enrich science or, conversely, create "streetlight effects" that narrow research focus toward data-rich, low-hanging fruit"
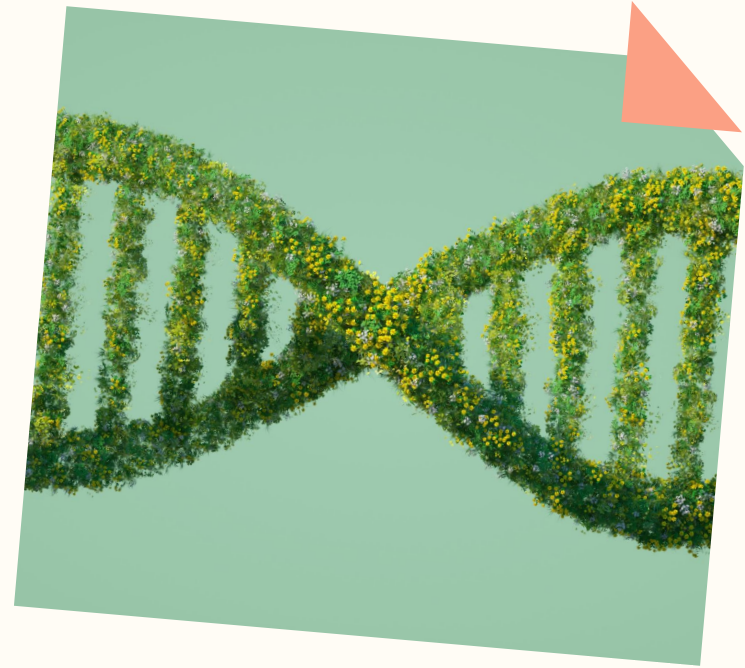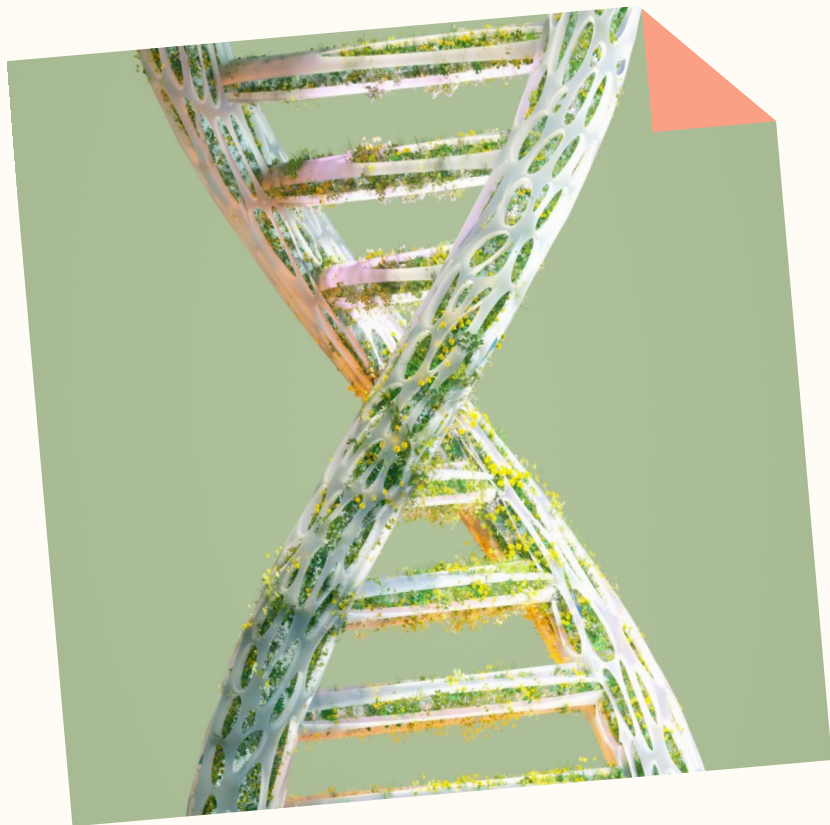
Question 3:

**If AlphaFold is getting protein structures right most of the time, does it still make sense to run expensive lab experiments just to confirm them?**
**And is it okay if we don't fully understand how the model works, as long as the output is correct?**

"AlphaFold 2 should be understood as a complement to experimental and domain-specific work rather than a substitute."

"High predictive accuracy does not necessarily imply mechanistic understanding."