- **Resources**
  - Pairwise Structure Alignment
  - CollabFold
- **What's Our Goal?**
  - We want to assess the the viability of adding a base-editor to the HiSCRIBE reverse transcriptase
  - AlphaFold can tell us:
    - Does adding this domain result in the misfolding of the base-editor or polymerase?
    - How long of a linker should we use?
- **Steps on Our Adventure**
  - *Finding a Gene*
    - The Original SCRIBE Paper
    - We're looking for the RT protein (from the Ec86 retron cassette in BL21)
    - Search the Gene on KEGG
    - Look for "Ec86" and find the one from *E. coli* BL21
    - Find the Sequence
  - *Predicting a Structure*
    - CollabFold Paper
      - 40-60x faster search than the original AlphaFold (thanks to a sped-up MSA step)
      - Can run on a Google Collab notebook in the cloud — doesn't require the researcher to have a powerful computer of 2TB of storage space for the databases!
      - They maintain a number of Collab notebooks which can be found on their GitHub
    - CollabFold GitHub
      - Here is where all of the code for CollabFold (and AlphaFold) is shared!
      - For basic folding simulations, we can use AlphaFold2 with MMseqs2
    - CollabFold Notebook
      - We can start by predicting the sequence of our Ec86, copying the amino acid sequence into the notebook
      - It can also be worth checking the `use_amber` box which enables model "relaxation", taking the predicted structure and wiggling residues around to relieve steric clashes
        - This typically improves the quality of a structure, but can slow down the folding process and sometimes lead to crashes
        - *We'll run without relaxation for now, to save some time!*
      - Here's One I Made Earlier!
  - *Finding a Base-Editor*
    - A MutaT7 Source Paper
    - AD is named TadA* -> Supplementary Materials -> Find TadA* in table -> Follow link to gene
    - The TadA* Fusion Protein

        - We can tell where the linker is from the long series of glycine and serines — TadA* is just before this!
        - We can add the TadA* and linker to the N-terminal of our polymerase protein
        - Here's One I Made Earlier
    - *Aligning the Sequences*
        - Shout-out to PyMol
        - RCSB Pairwise Structural Alignment
        - There are a few different ways you can align proteins — some in a rigid way (not changing the supplied structures) and others in a more flexible way (warping proteins into some alternative conformations to match)
        - *All of these methods are aligning structures — not sequences (like AlphaFold does)!*
        - RCSB Alignment Tool
            - Upload the two (or more) PDB files that you'd like to compare and select the chain of interest
                - For single-chain structures like ours, this will almost always be `A`
            - Align everything but the TadA* to the Ec86
            - Align both fusions to the TadA*
    - *Results*
        - It looks like the reverse transcriptase structure is relatively unaffected by the base editor when a linker is used!