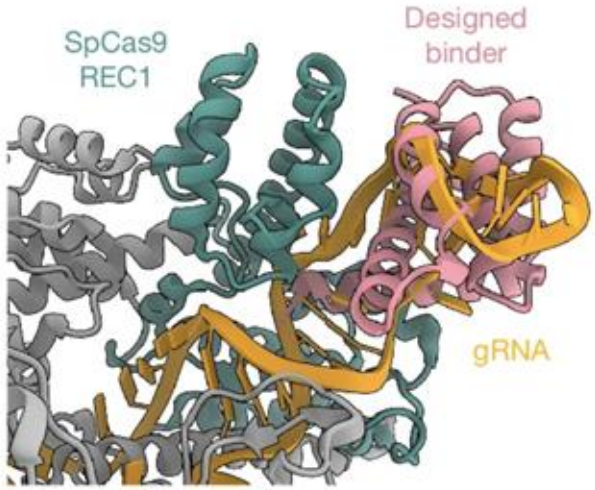
One-shot design of functional protein binders with BindCraft

Presented by Jack McMahon

Pacesa, M., Nickel, L., Schellhaas, C. *et al.* One-shot design of functional protein binders with BindCraft. *Nature* 646, 483–492 (2025). https://doi.org/10.1038/s41586-025-09429-6

Overview

- Background
  - De Novo Protein Design
  - Previous Work (RFdiffusion)
- Paper Aims
- Protein Design Pipeline
  - Binder Hallucination and Optimization
  - MPNNsol Sequence Optimization
  - Monomer Modelling and Filtering
- Applications
- Limitations
- Concluding Takeaways

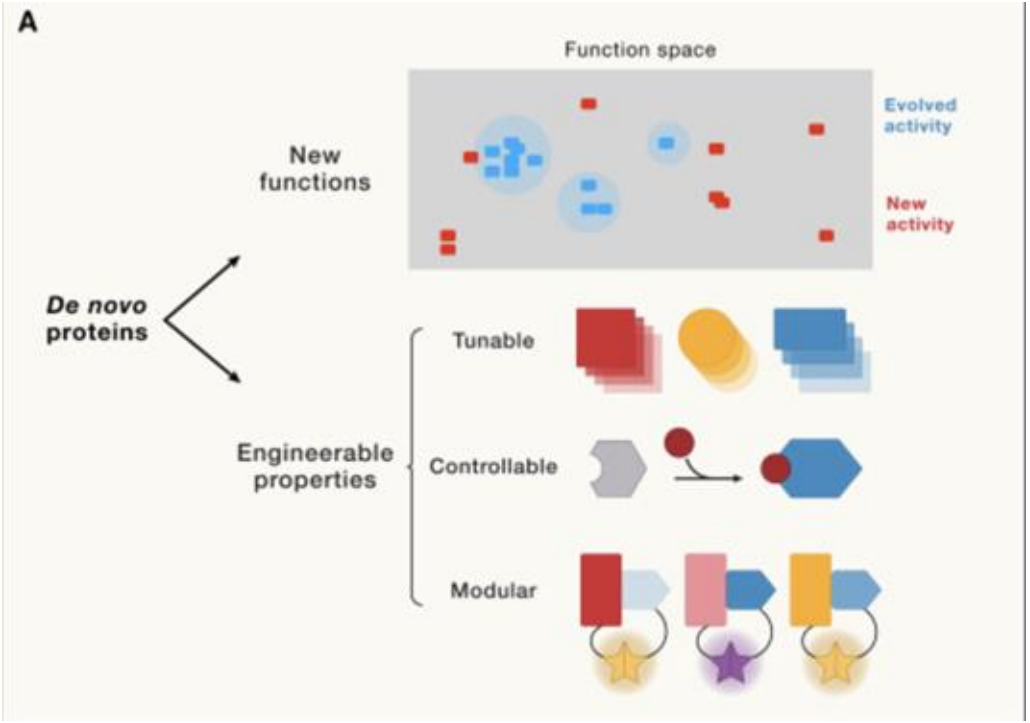


BindCraft applied to SpCas9

Significance of De Novo Protein Design

- Broad biological applications across drug discovery, diagnostics, industrial chemistry etc.
- Precise design potentially much more powerful than reengineering existing proteins

Goal: Given a desired function, predict associated structure and sequence of amino acids



Kortemme, T. (2024). De novo protein design—From new structures to programmable functions. *Cell*, *187*(20), 5522-5541. https://doi.org/10.1016/j.cell.2024.08.014

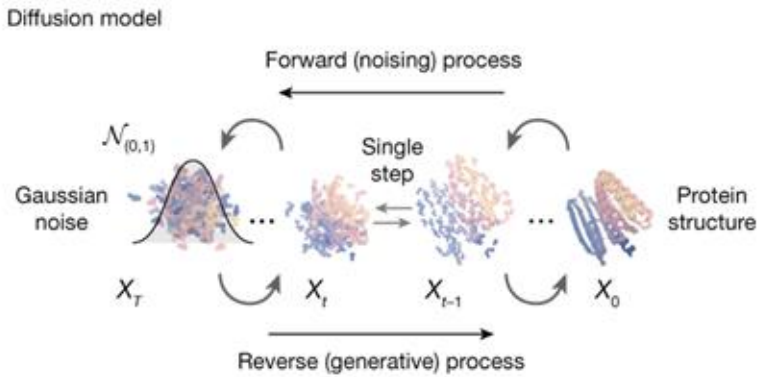
Pre-ML Methods

- Pre-ML methods to generate binders were time consuming and included very limited design control
  - Immunization, antibody library screenings, directed evolution etc
- Physics-based methods like Rosetta were computationally limited and had very low success rates
  - Required screening up to tens of thousands of candidates in vitro
  - Not well suited for binders

RFdiffusion

- Significant advancement in de novo protein design from the Baker lab in 2023
- Gaussian denoising to yield protein backbone for rigid target
- Scores structures using AlphaFold
- 0.1-10% in vitro success rates

Limitations: Low accuracy and disconnected generation steps



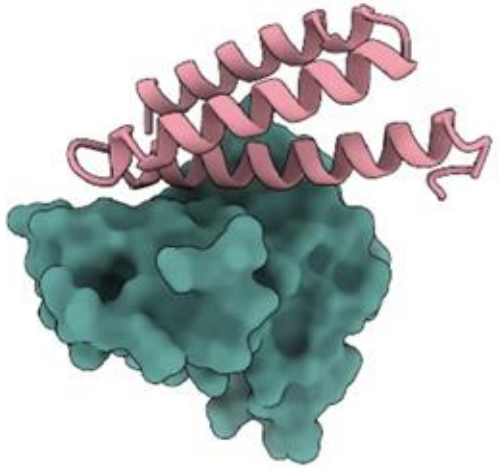
Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borst, A. J., Ragotte, R. J., Milles, L. F., Wicky, B. I. M., Hanikel, N., Pellock, S. J., Courbet, A., Sheffler, W., Wang, J., Venkatesh, P., Sappington, I., Torres, S. V., ... Baker, D. (2023). De novo design of protein structure and function with RFdiffusion. *Nature*, *620*(7976), 1089-1100. https://doi.org/10.1038/s41586-023-06415-8

Key Aims of BindCraft

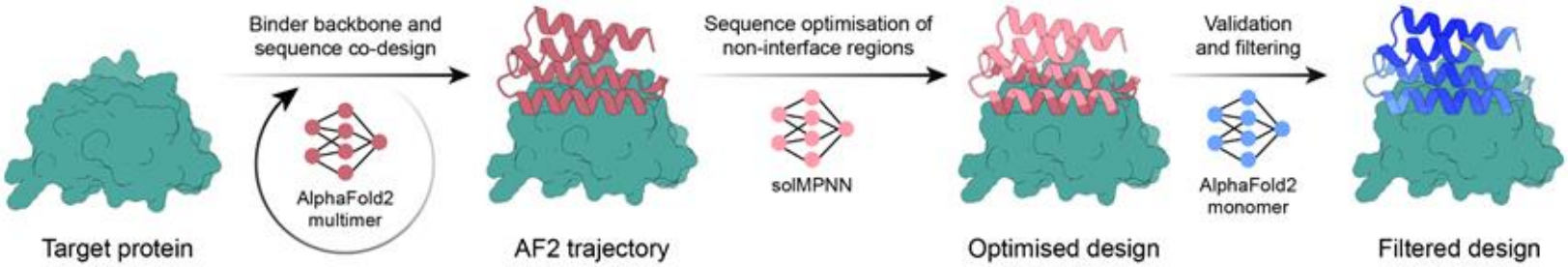
Rather than using Alphafold to filter predictions after generation, is there a way to incorporate AlphaFold directly into structure hallucination?

“One-Shot Design”

- Iteratively optimize structure hallucination using AF2 energy function
- Co-fold with target structure
- Build backbone and sidechains together

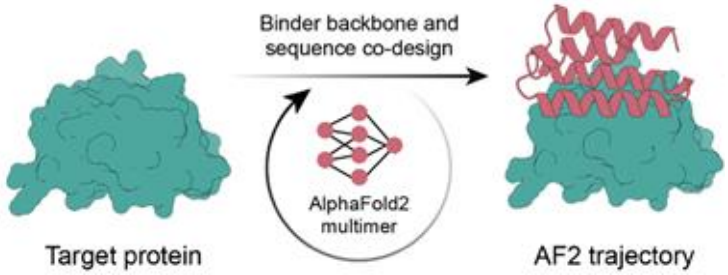


Protein Design Pipeline



1: Binder Hallucination and Optimization via AlphaFold

Input: Target PDB structure and binder specifications



Hallucination: Starts with random binder sequence, AF2 provides complex structure prediction

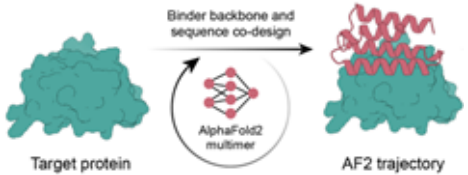
Loss Function:

- Prediction Confidence: pLDDT, i_pTM, predicted alignment error
- Packing: Residue contact loss (internal and external)
- Topology: Radius of gyration and helicity

Optimizes: Binder sequence

Parallelization: Runs process in parallel for hundreds of candidates

1: Binder Hallucination and Optimization via AlphaFold



Four stages of sequence optimization:

1: Continuous optimization
- -sequence identity is a blend of possible amino acids per position
- -75 iterations

2: Funnel towards realistic sequence
- -learning rate decays and normalizes logits to sequence probabilities
- -45 iterations

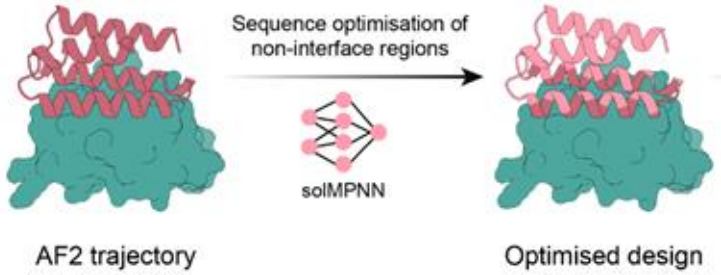
3: Move to discrete amino acids
- -Incorporate one-hot encoding of top amino acid per position
- -5 iterations

4: Stochastically sample different mutations
- -15 iterations

2: MPNNsol Sequence Optimization

MPNNsol refines AlphaFold predicted structure away from contacts to improve stability

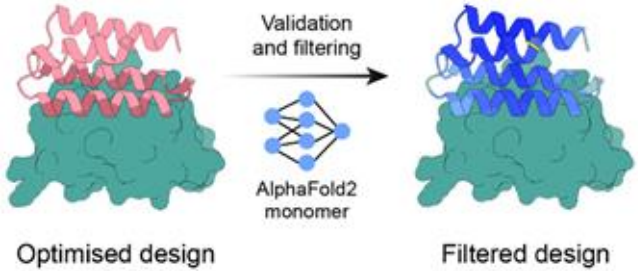
- Inverse folding model
- Binder residues at target interfaces are preserved
- Design 20 new sequences for the remaining binder core and surface residues
- Uses solubility-trained model to prevent hydrophobicity



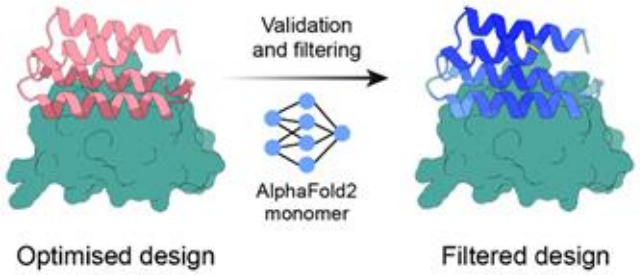
3: Monomer Modelling and Filtering

Quality control step to insure prediction viability

- Optimized sequences are predicted in monomer form using AlphaFold2
- Monte-carlo based energy relaxation via Rosetta FastRelax checks for physical constraints
- Aggregate quality metrics to choose top candidates (next slide)



3: Monomer Modelling and Filtering



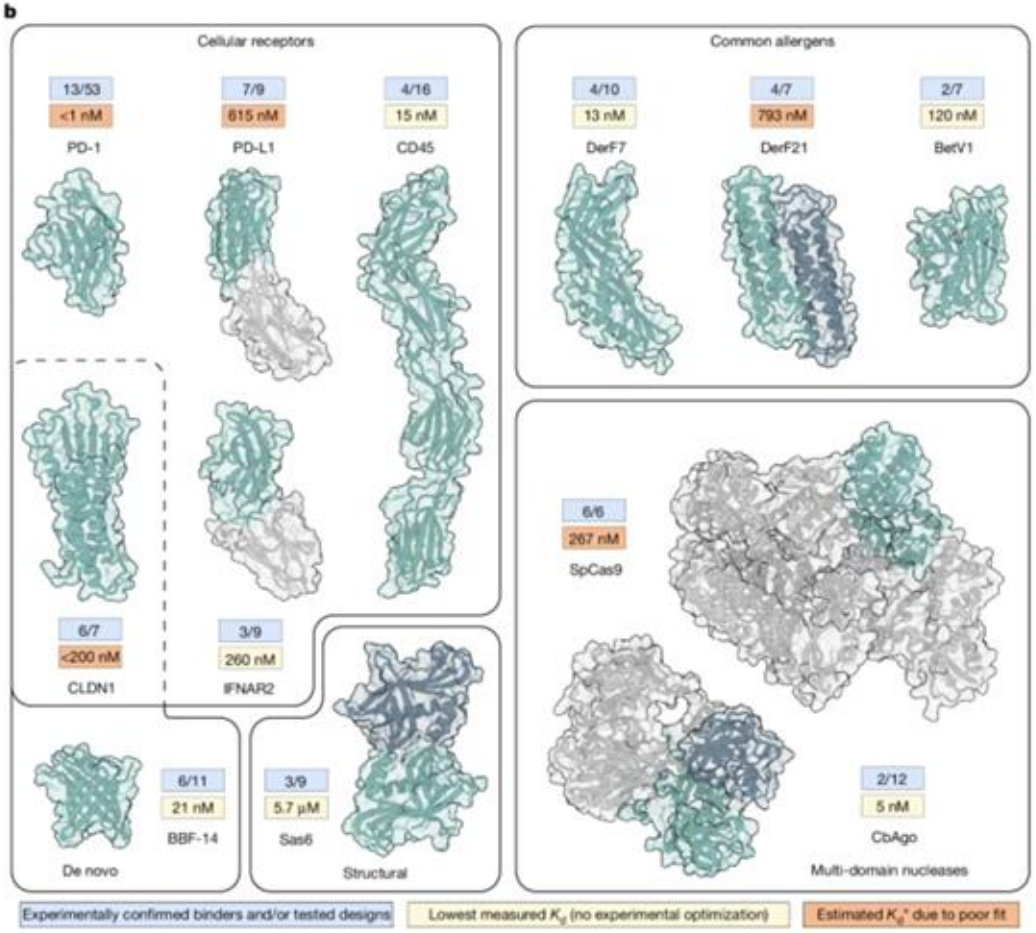
Filtering metrics:

- AF2: > 0.8 complex pLDDT
- AF2: > 0.5 interface confidence (i_pTM)
- AF2: > 0.35 predicted alignment error
- Rosetta Interface Shape: Above 0.6 complementarity
- > 3 hydrogen bonds at interface
- < 4 saturated hydrogen bonds at interface
- < 35% binder surface hydrophobicity
- < 0.35 angstrom rmsd bound and unbound structure
- <3 lysines and methionines at binder interface

Pipeline:
1: 100s of hallucinations -> 1,000s of MPNNsol structures -> top 10-20 designs

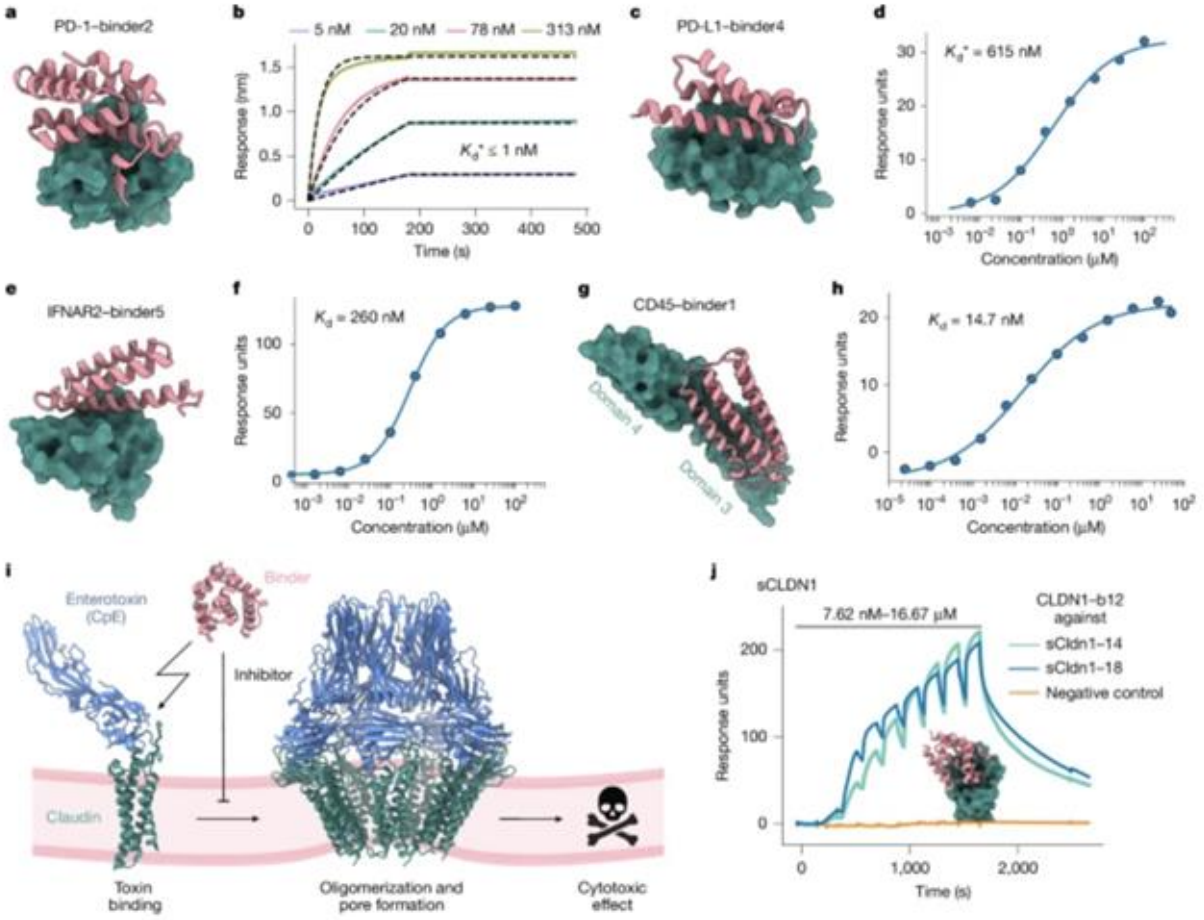
De Novo Binder Design

- Huge leap forward in accuracy -> 46% success in vitro
  - Far above previous 0.1%-10% rates
- Very high design affinities achieved
- Tested on cell receptors, allergens, and multi-domain nucleases



Cell-Surface Receptor Binders

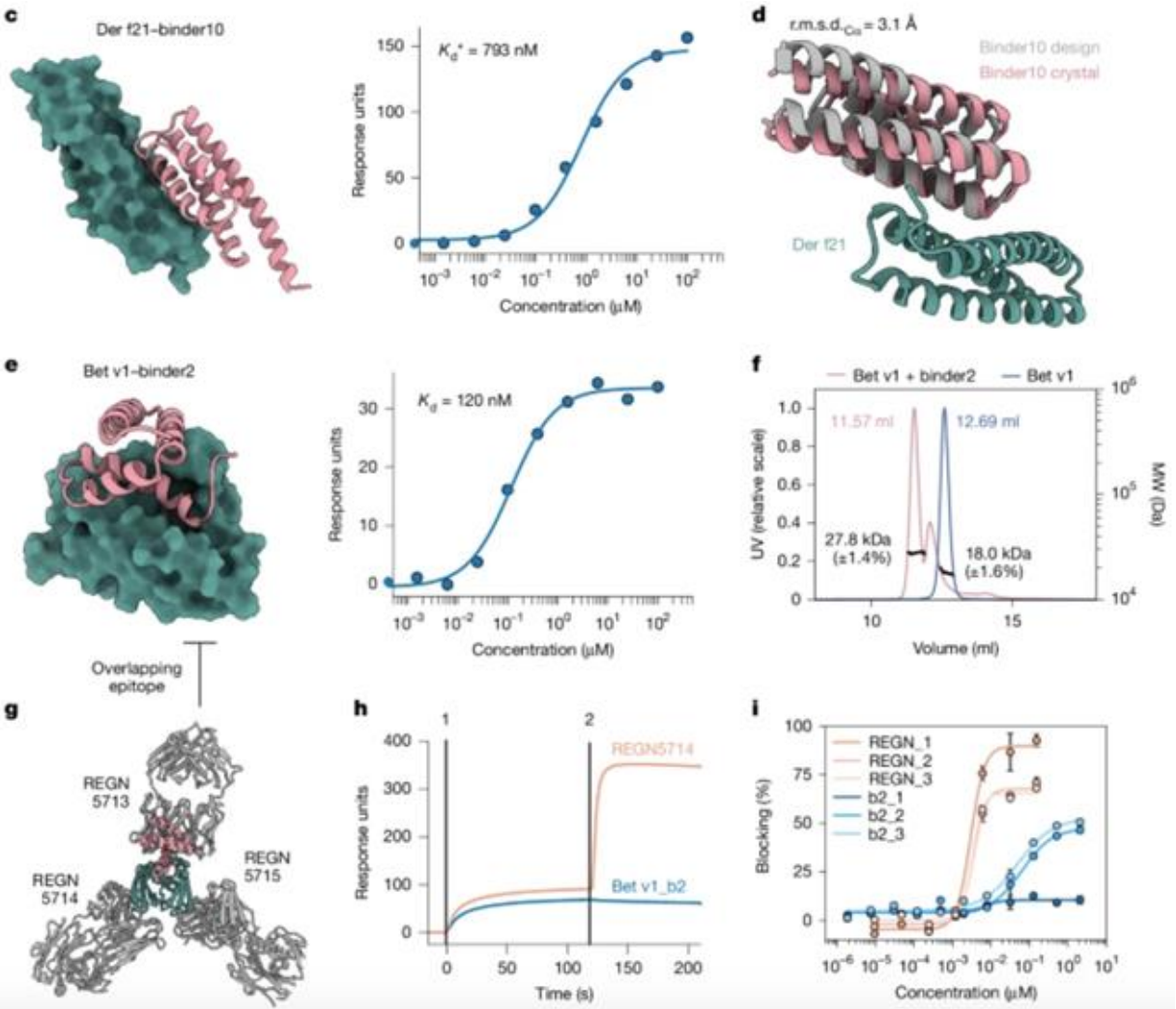
- PD-1 is a human immune checkpoint receptor
- Achieved up to 1nM dissociation constant
- Competition assay with pembrolizumab
- Extended to CD45 as previously uncharacterized binding site
- Measured surface plasmon resonance



Allergen Binders

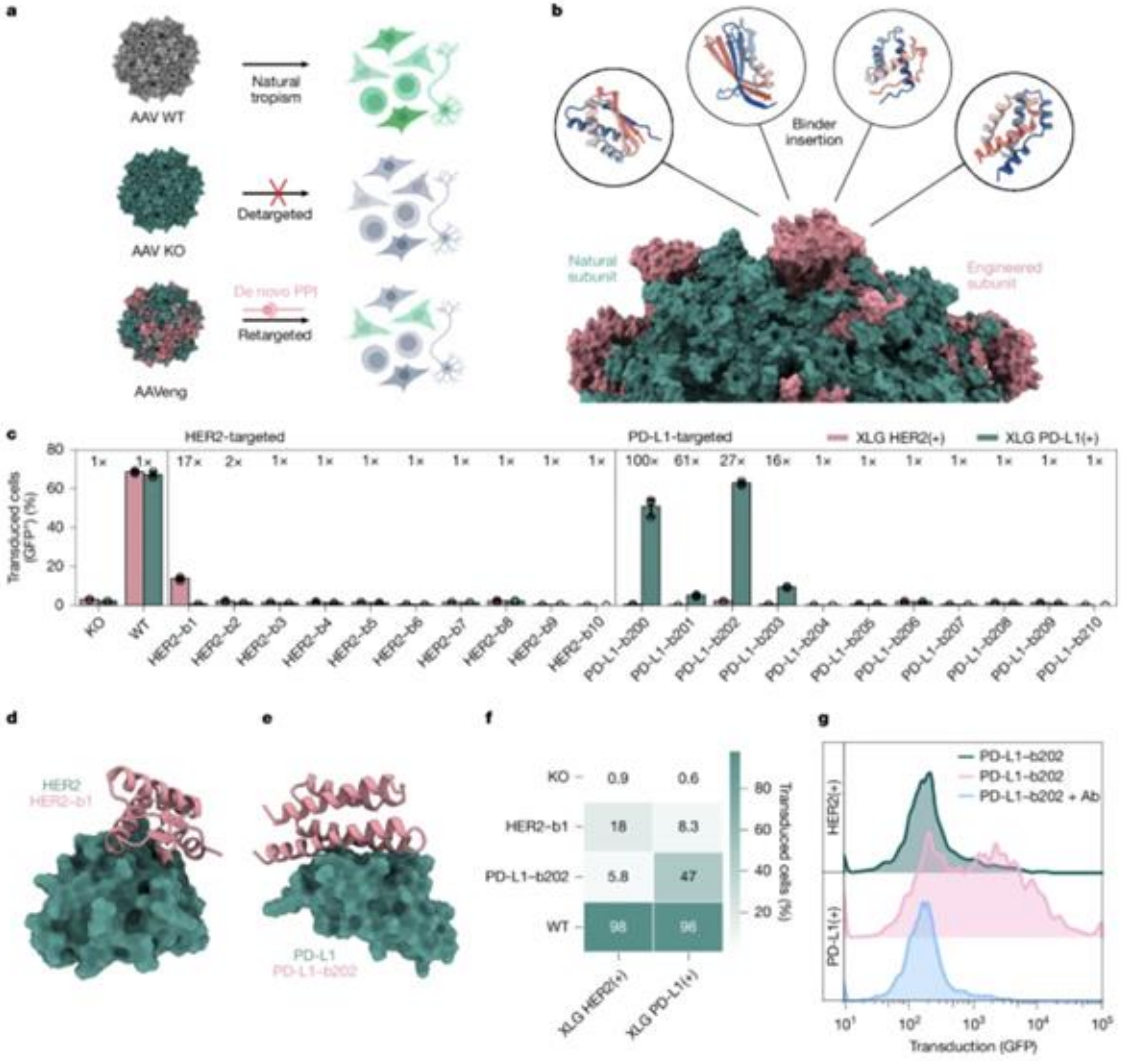
Goal: Ability to neutralize allergic reactions

- Dust mite (Der) and birch (Bet) allergens tested
- Designed binders achieved comparable performance to single antibodies
- Competition assay and blocking ELISA assay



Binders for Targeting Gene Delivery

- Adeno-associated viruses can deliver genes but have poor cell specificity
- Miniprotein binders can precisely target cell-type specific receptors
- Designed binder against HER2, measured with flow cytometry (GFP)



Limitations

- Computationally intensive
  - Running BindCraft requires access to GPUs
  - Creating a set of viable binders requires hundreds of iterations through AlphaFold for each sequence -> hours/days of GPU time
  - High memory constraints
- Inherits AlphaFold limitations
  - Doesn’t extend as well to proteins with intrinsically-disordered regions or multiple conformations
  - Expected to perform worse on less well-characterized folds

Questions

- How could future work improve the efficiency of binder hallucination and backpropagation?
- What types of design tasks would be potentially challenging for BindCraft, and what would accuracy rates look like on those tasks?
- Future work: head-to-head design benchmarks?

Concluding Takeaways

- With an accuracy rate of 46% (10%-100% range on tasks), BindCraft is a huge step forward for de novo binder design
- Despite higher computational costs, binder design is much more efficient overall without the need for large in vitro screening assays
- Enables a wide array of new protein design tasks, lowers barriers to experimentation due to high success rates
- Demonstrated success on unexplored binding sites is an exciting advancement, validates engineering potential