UT Southwestern Medical Center

Applied Protein Design – BindCraft

Dominika Borek, Ph.D.

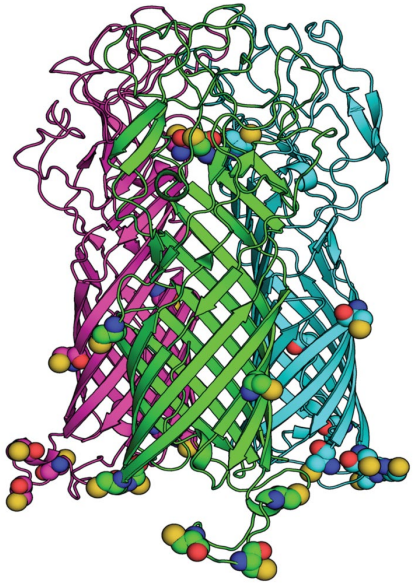
Disclosures:

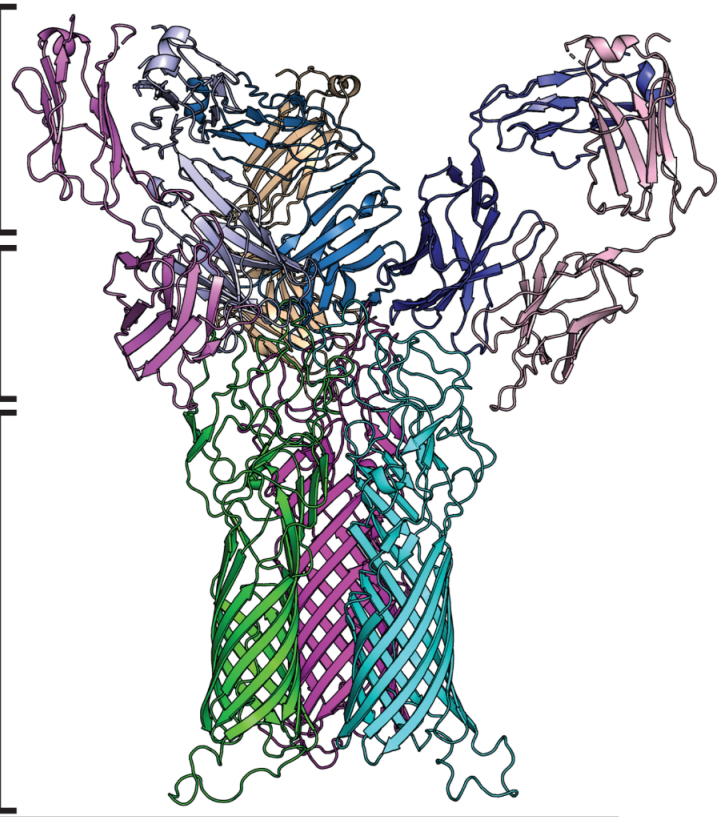
DB is a co-founder of Ligo Analytics, a company that commercializes technology related to concepts presented in this talk.

Macromolecules work through interactions

These interactions control:

- Signaling
- Immunity
- Infection
- Transport
- Many other biological processes

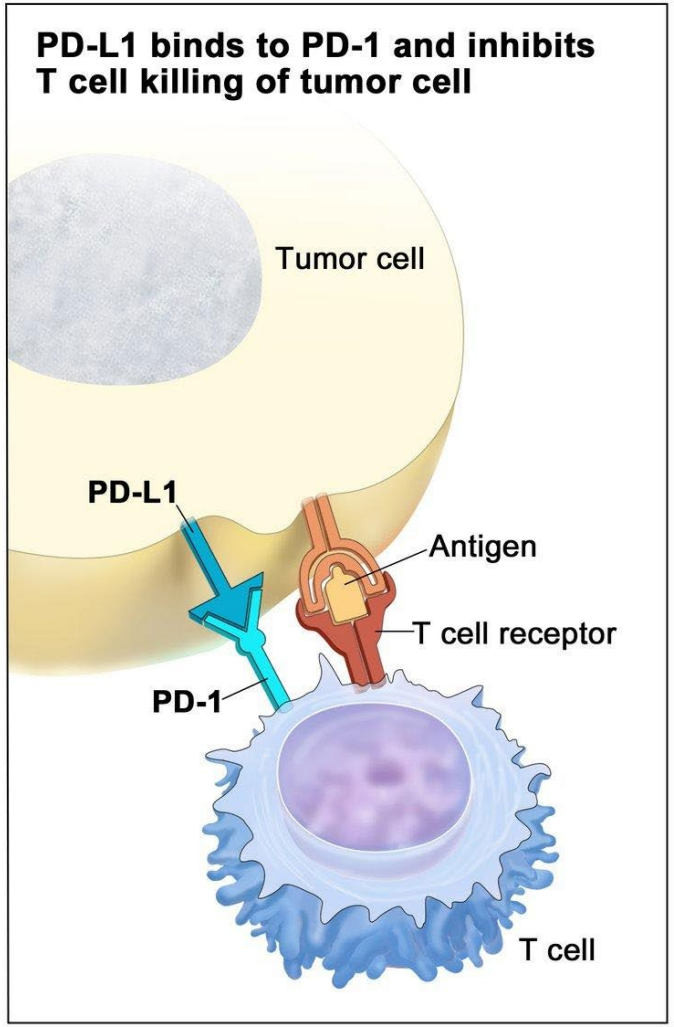


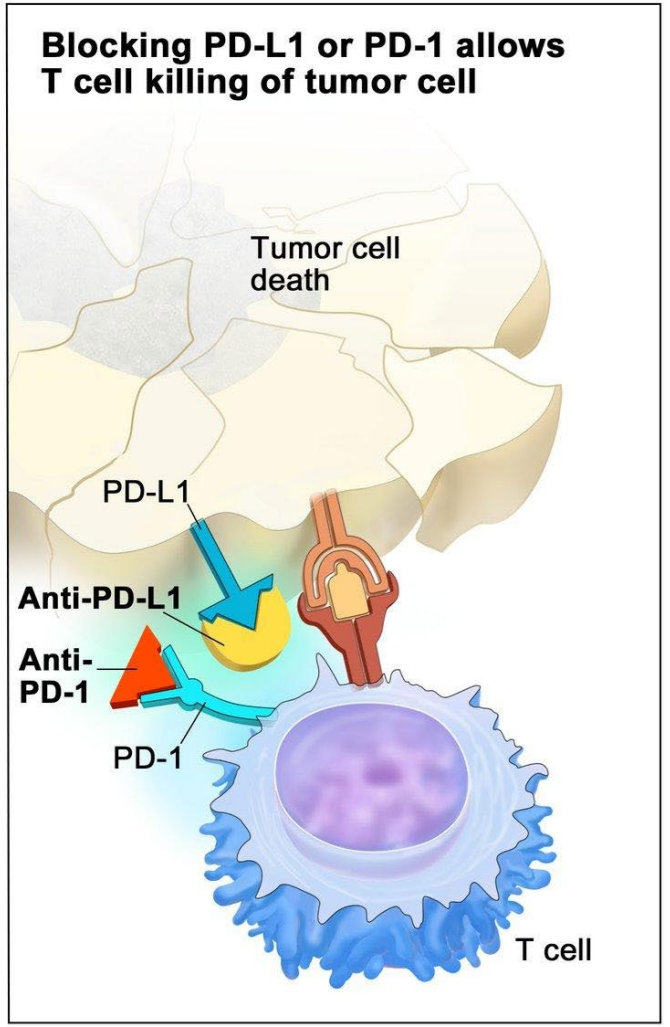


A binder is a molecule designed to recognize a specific surface on another molecule.

What can a binder help us do?

- **Block** an unwanted interaction
- **Activate or recruit** a biological process
- **Detect** a target in a diagnostic assay
- **Deliver** a payload to a selected cell or protein





Why design binders computationally?

Traditional discovery

- Immunization
- Library screening
- Directed evolution
- Experimental screening

Computational design

- Begin from a target structure
- Aim at a selected surface
- Generate new sequences and structures
- Experimentally test a smaller, prioritized set

Can we design a protein that fits a target surface before making it in the laboratory?

UT Southwestern Medical Center

What is BindCraft?

BindCraft is an automated, user-friendly pipeline that uses a target protein structure to generate new protein sequences predicted to fold and bind that target.

Article | Open access | Published: 27 August 2025

One-shot design of functional protein binders with BindCraft

Martin Pacesa, Lennart Nickel, Christian Schellhaas, Joseph Schmidt, Ekaterina Pyatova, Lucas Kissling, Patrick Barendse, Jagrity Choudhury, Srajan Kapoor, Ana Alcaraz-Serna, Yehlin Cho, Kourosh H. Ghamary, Laura Vinué, Brahm J. Yachnin, Andrew M. Wollacott, Stephen Buckley, Adrie H. Westphal, Simon Lindhoud, Sandrine Georgeon, Casper A. Goverde, Georgios N. Hatzopoulos, Pierre Gönczy, Yannick D. Muller, Gerald Schwank, ... Bruno E. Correia + Show authors

Nature **646**, 483–492 (2025) | Cite this article

Save article

187k Accesses | **257** Citations | **306** Altmetric | Metrics

Input

- Target structure in PDB format
- Target chain or chains
- Optional hotspot residues
- Binder-length range

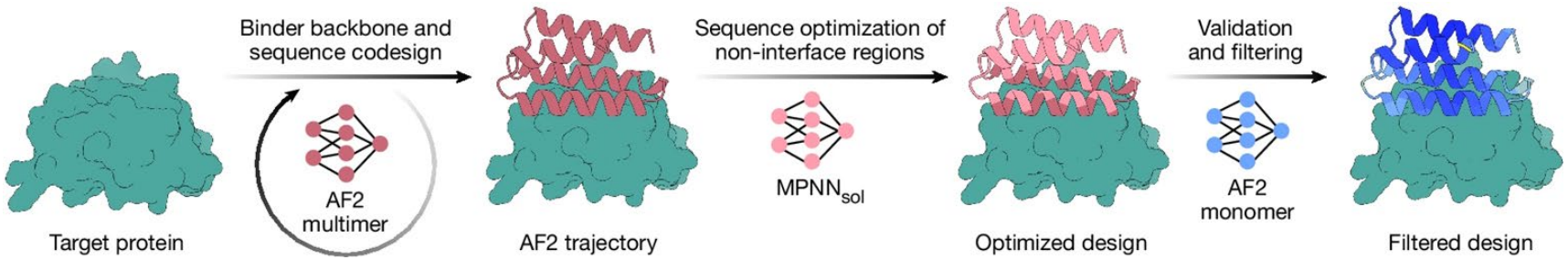
Output

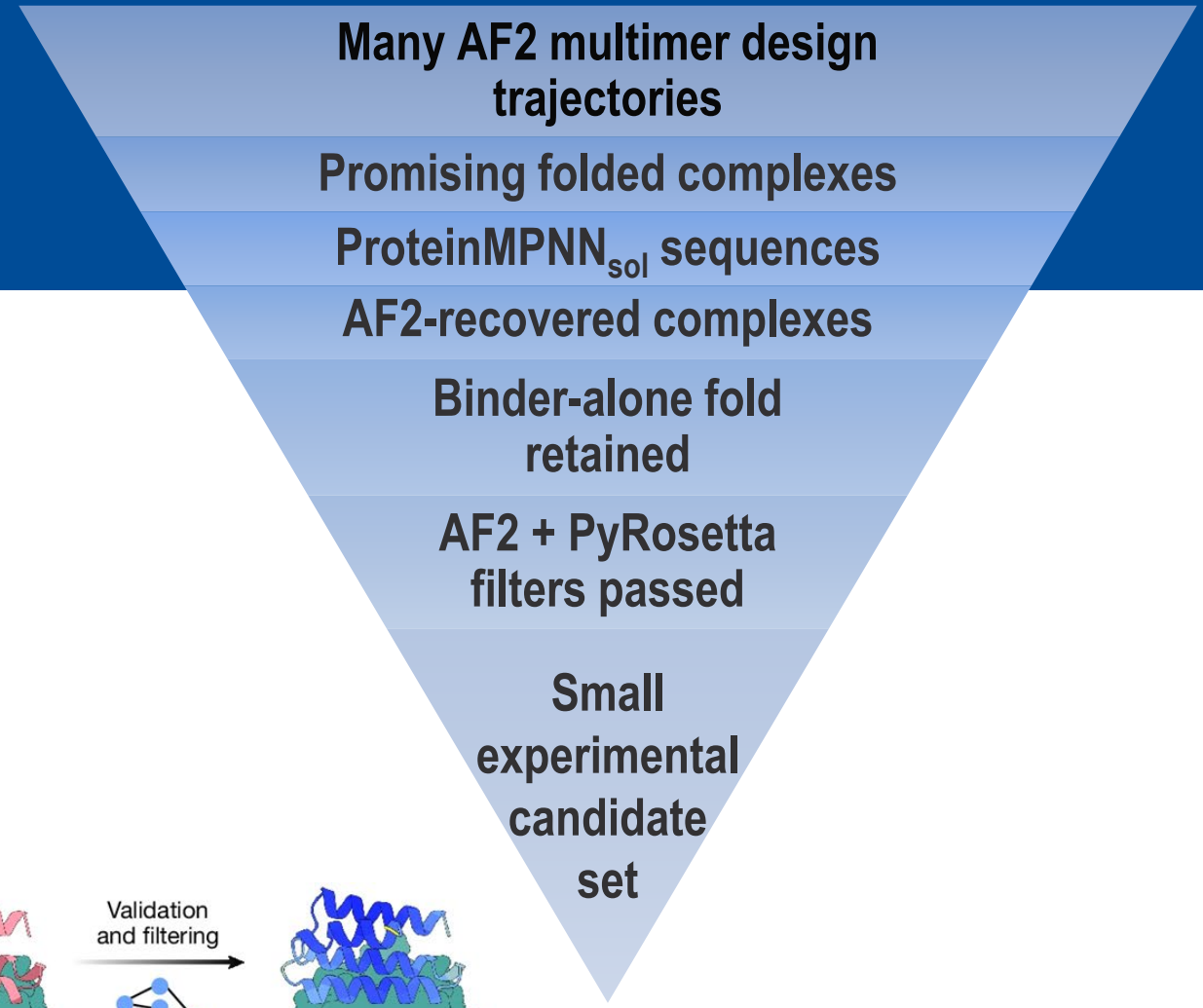
- Candidate binder sequences
- Predicted binder-target complexes
- Confidence and interface metrics
- Accepted and rejected design sets

BindCraft

BindCraft is a self-consistency funnel

- **BindCraft is not one prediction.**
- **It is a loop of generate → redesign → re-predict → filter.**
- **A trajectory is one complete design attempt.**

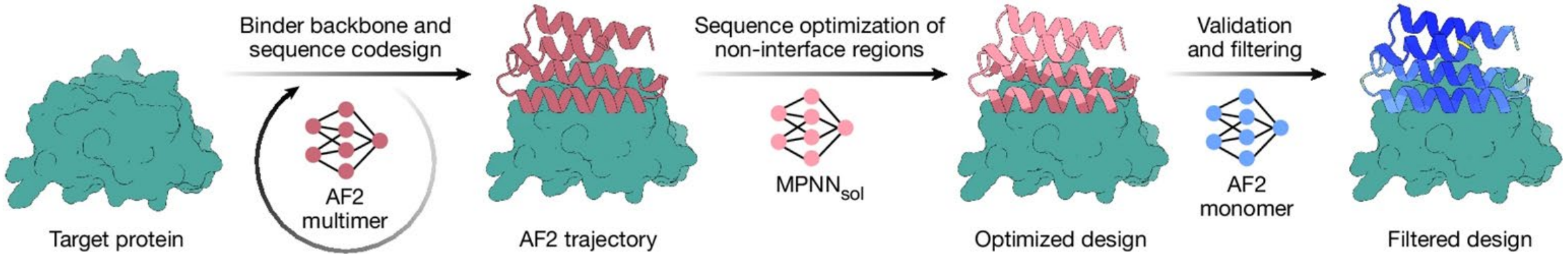




https://github.com/martinpacesa/BindCraft/blob/main/README.md

UT Southwestern Medical Center

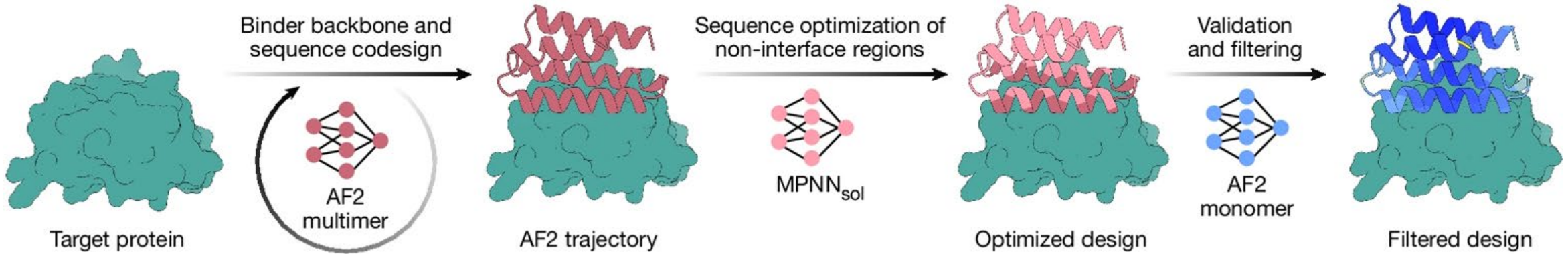
Input to AF2: How does BindCraft start a trajectory?



- **Each trajectory begins with a new random seed.**
- **One binder length is randomly selected from the user-defined range and remains fixed for that trajectory. An L × 20 matrix of nearly uniform amino-acid logits is randomly initialized for the binder.**
- **No starting binder backbone or poly-alanine sequence is provided.**
- **The variable binder representation is combined with the fixed target sequence and target structure (+ any optional hotspot information).**
- **AF2 Multimer predicts the complex in single-sequence, MSA-free mode.**

UT Southwestern Medical Center

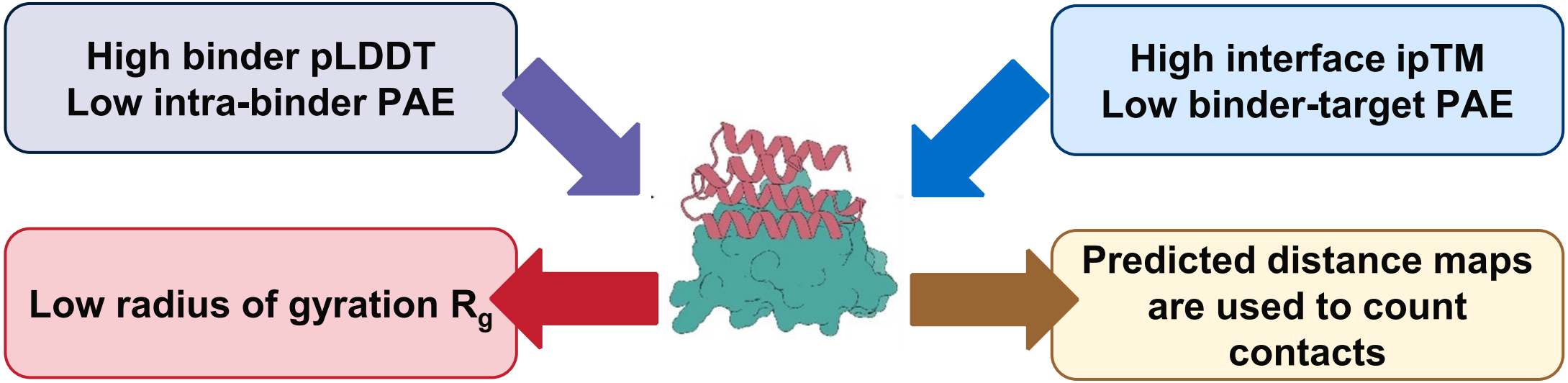
Turning AF2 Multimer from a predictor into a designer



- **AF2 Multimer predicts the structure of the target-binder complex.**
- **BindCraft evaluates the prediction using several design objectives (loss functions).**
- **The loss gradient is propagated back to the binder sequence logits (not to the AF2 model parameters!).**
- **The binder sequence is updated and the complex is predicted again.**
- **This cycle is repeated until a discrete binder sequence and structure emerge.**
- **ColabDesign (Ovchinnikov et al.) makes the AF2 prediction process differentiable, allowing the design loss to guide sequence updates.**

What is BindCraft asking AF2 to optimize?

BindCraft minimizes a weighted loss that favors a confidently folded, compact binder with a well-defined interface and sufficient contacts to the target.

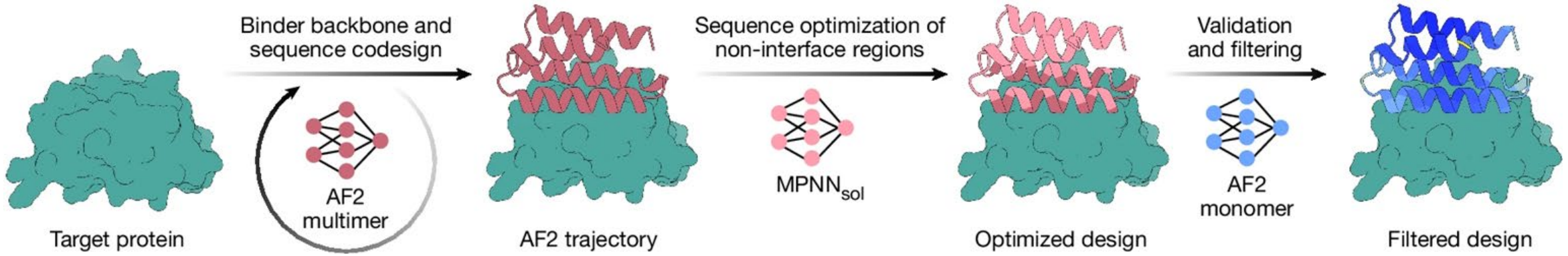


pLDDT (Predicted Local Distance Difference Test) → *How confident am I that this specific piece is built correctly?*

PAE (Predicted Aligned Error) → *If I place this domain, how wrong might the other domain be in relation to it?*

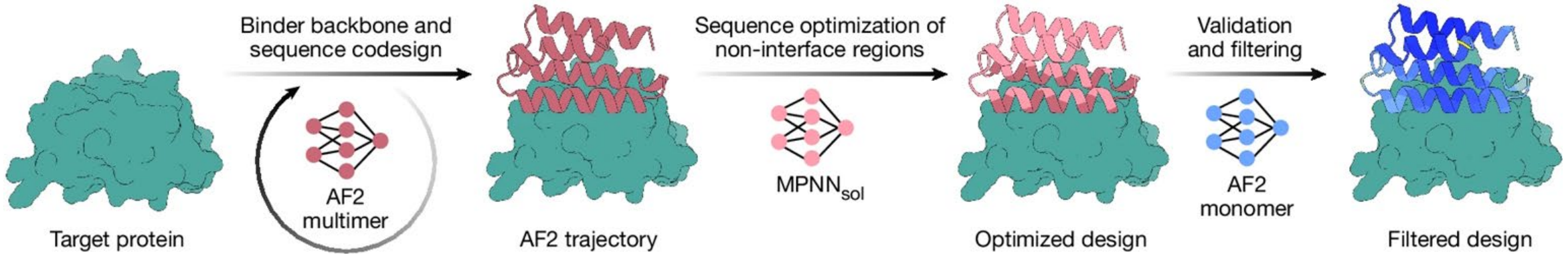
ipTM (interface predicted Template Modeling score) → *Are these proteins touching each other in the correct way?*

Turning AF2 Multimer from a predictor into a designer

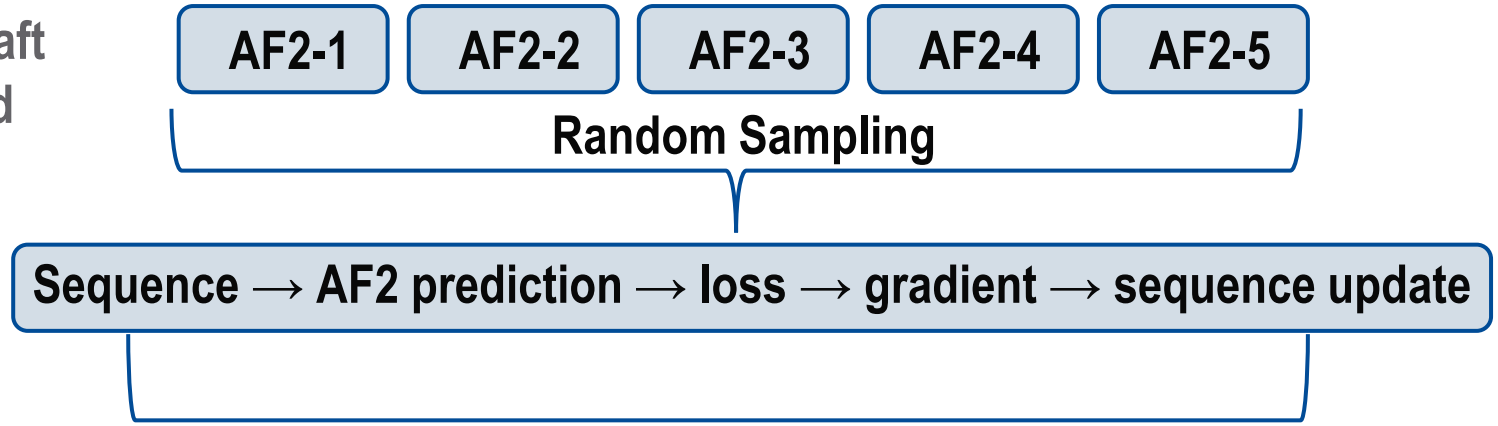


- AF2 Multimer predicts the structure of the target-binder complex.
- BindCraft evaluates the prediction using several design objectives (loss functions).
- The loss gradient is propagated back to the binder sequence logits (not to the AF2 model parameters!).
- The binder sequence is updated and the complex is predicted again.
- This cycle is repeated until a discrete binder sequence and structure emerge.
- ColabDesign (Ovchinnikov et al.) makes the AF2 prediction process differentiable, allowing the design loss to guide sequence updates.

Self-consistency check 1: AF2 model swapping

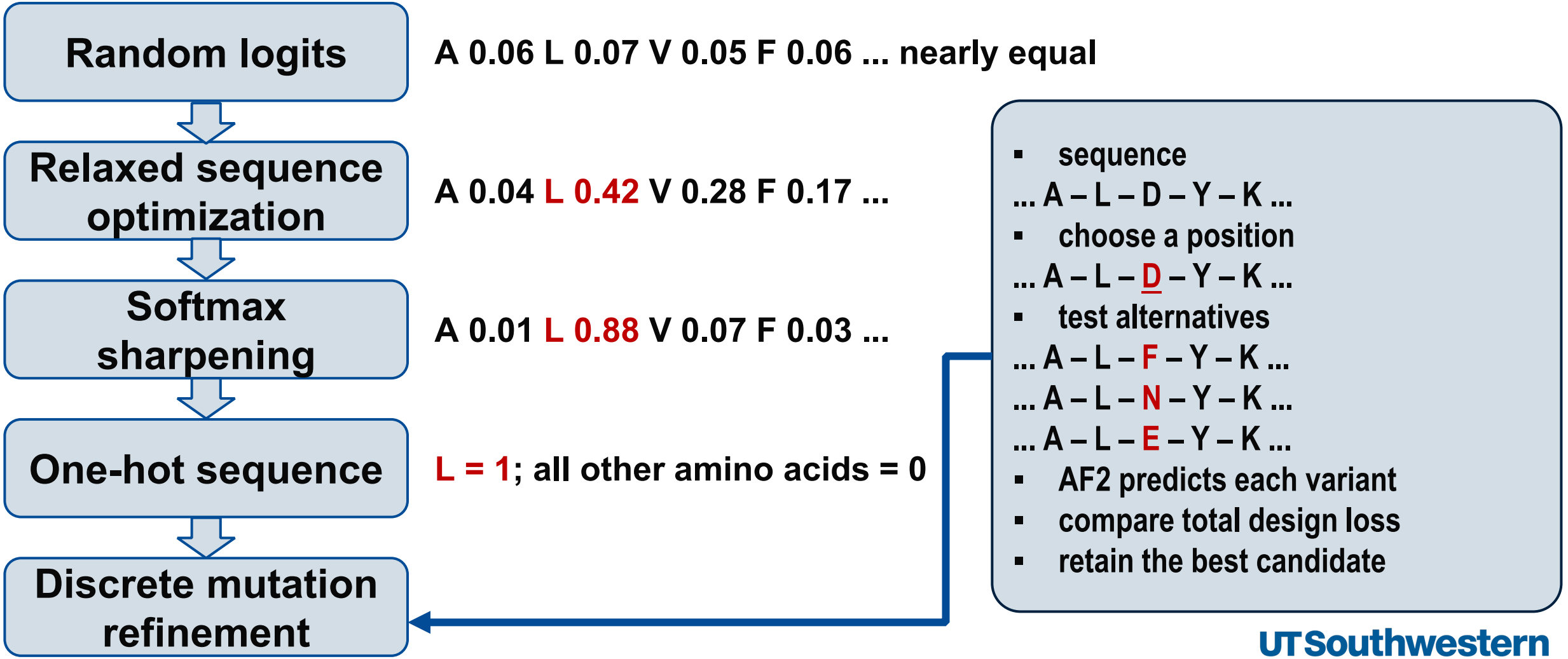


- During sequence optimization, BindCraft randomly alternates among five trained AF2 Multimer models.
- Each iteration uses one model to calculate the prediction, loss, and sequence gradient.
- This reduces overfitting to one AF2 model and favors sequences that remain plausible under changing model parameters.



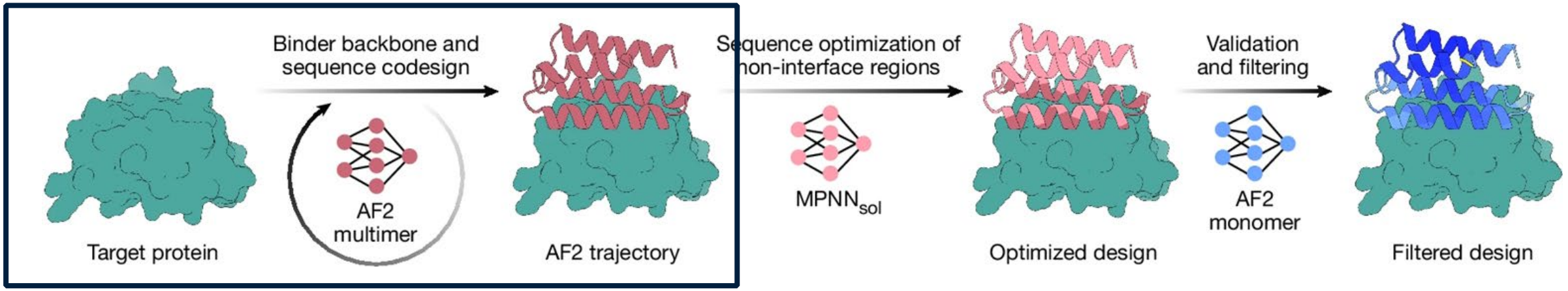
UT Southwestern Medical Center

How does BindCraft turn a soft sequence into a real protein sequence?



UT Southwestern Medical Center

Why do some AF2 trajectories fail?



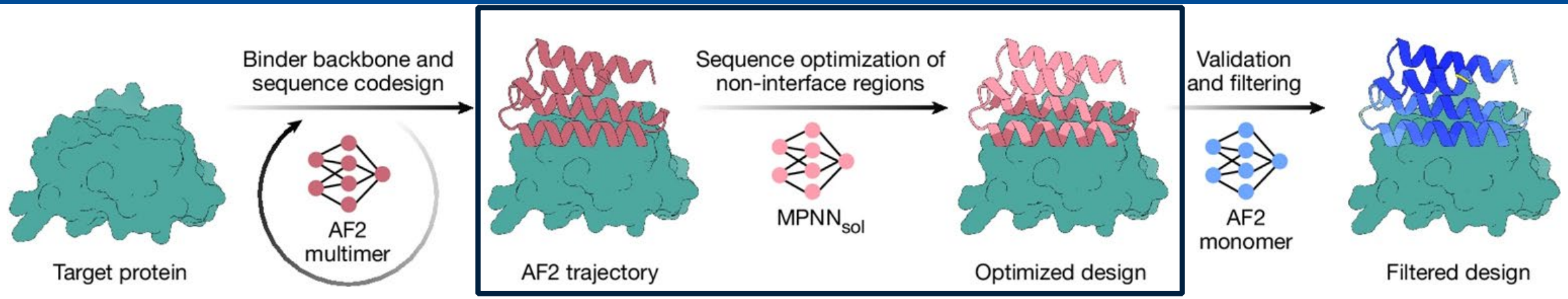
One AF2 trajectory produces:

- **One discrete binder sequence**
- **One predicted binder–target complex**
- **A record of the loss and confidence during optimization**

Why can it fail?

- **Confidence drops when the soft sequence becomes a real sequence**
- **The final binder remains poorly defined, low-confidence or makes too few contacts and "floats" above the target**
- **The predicted complex contains severe backbone clashes**

Protein MPNN redesigns the sequence while preserving the interface



Preserve:

- Binder backbone and binding mode from AF2 trajectory
- Interface residues identified and held fixed

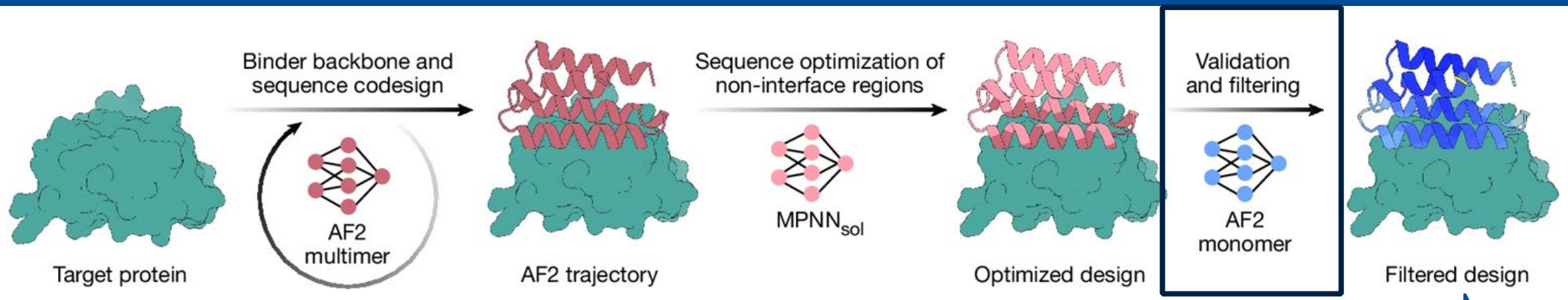
Redesign with MPNN$_{sol}$:

- Optimize core and exposed surface positions
- MPNN is a different neural network.
- 20 different redesigned sequences by default.

We have changed much of the sequence—but will the new sequence still produce the same fold and binding mode?

"Preserved" means fixed during sequence redesign. It does not guarantee that the new sequence will recover the interface. That is tested in the next AF2 step.

AF2 Monomer validation: can the redesigned sequence recover the complex?



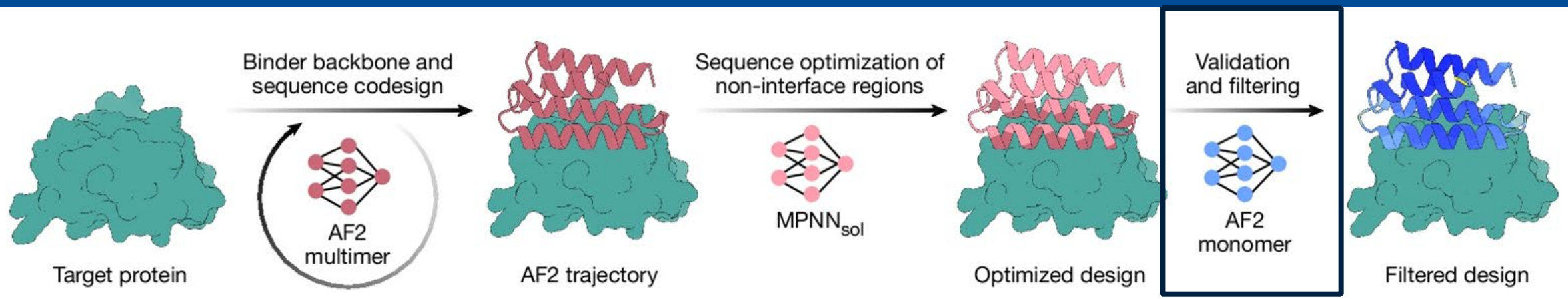
Input:

- MPNN-redesigned binder sequence
- Fixed target sequence and structural template
- Single-sequence, MSA-free prediction
- The original binder backbone is not provided as a template in the default workflow

Re-predict with different AF2 weights:

- **Use AF2 Monomer rather than the AF2 Multimer weights used for design**
- **Predict each MPNN sequence without backpropagation**
- **Published default: two template-capable AF2 Monomer models and three recycles**
- **The sequence is fixed (this is validation, not optimization)**

Can AF2 Monomer recover the designed complex?



Does the binder refold?

- Confident binder structure
- Similar fold to the original trajectory

Does it return to the same site?

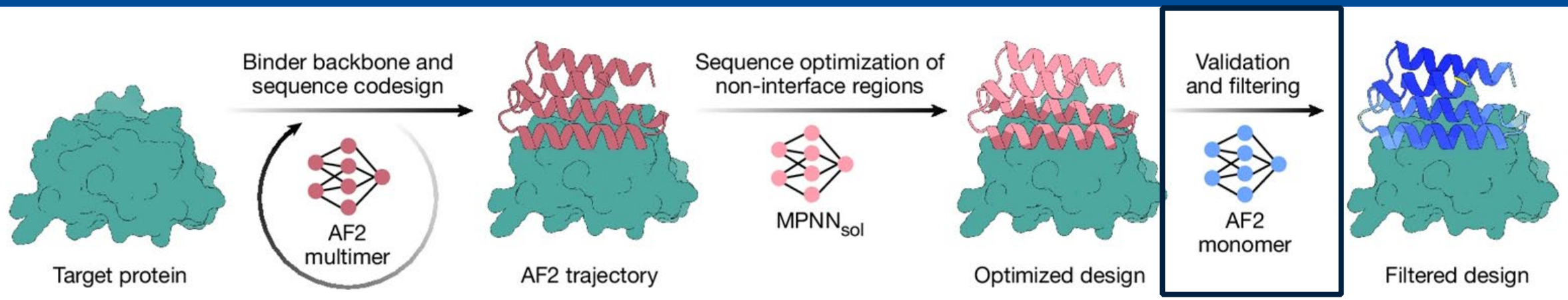
- Same target surface
- Similar orientation and binding mode

Is the interface confidently defined?

- High pLDDT and pTM
- High i-pTM
- Low i-pAE

Failure may mean a different binder fold, loss of docking, movement to another site, or low prediction confidence. The complex can be recovered in the presence of the target, but does the binder retain its fold on its own?

Binder-alone prediction: does the binder fold without the target?



BindCraft predicts the redesigned binder by itself and compares its structure with the binder conformation in the target-bound complex.

- **Low RMSD + high confidence: fold is likely encoded by the binder sequence.**
- **High RMSD or low confidence: binder may be flexible, unstable, or dependent on the target for folding.**

A low binder RMSD supports structural self-consistency, but it does not prove experimental stability or binding.

PyRosetta evaluation: does the interface pass a physical sanity check?

- AF2 asks whether the complex is predicted confidently
- PyRosetta asks whether its atomic geometry, packing, and interface chemistry are physically reasonable.

Relax:

- Apply a restrained FastRelax calculation
- Allow small backbone and side-chain adjustments
- Relieve local clashes without redocking the binder
- Whole-chain rigid-body movement is disabled

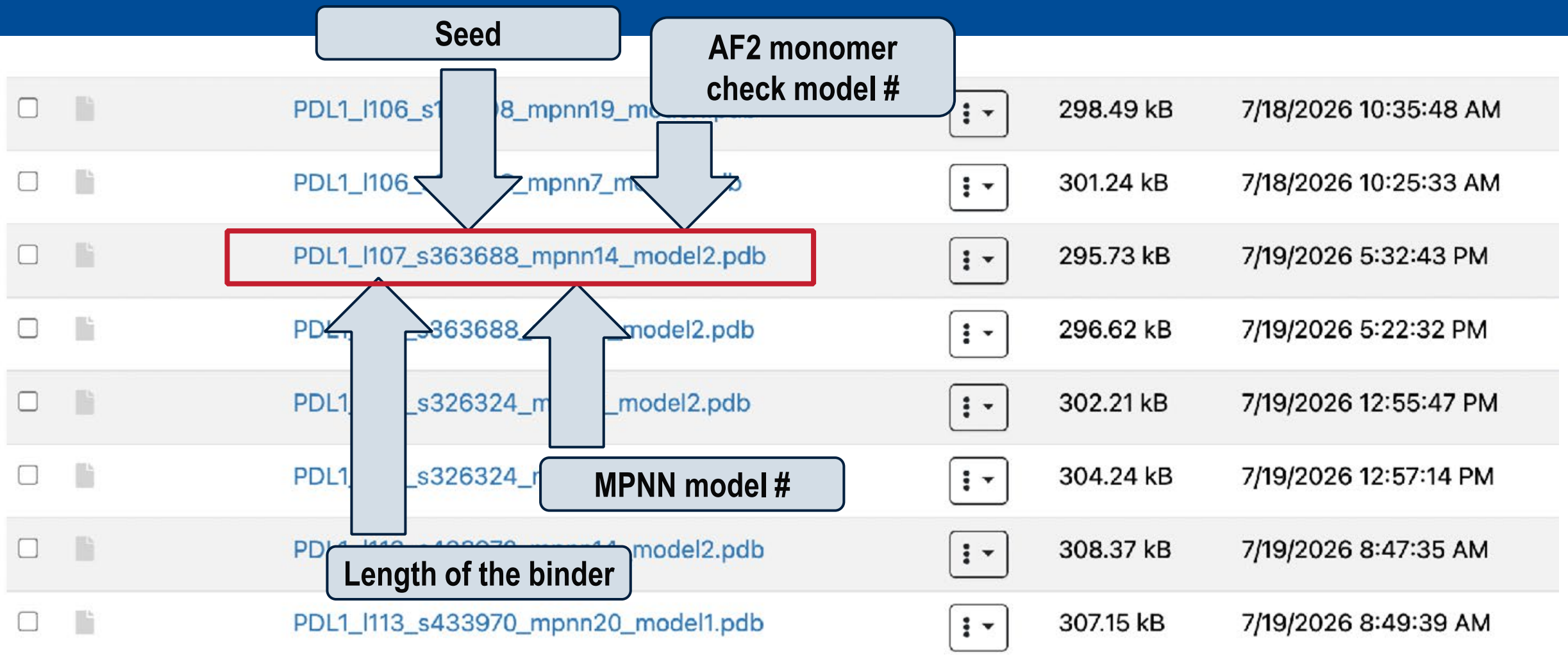
Analyze:

- Is the interface geometrically well packed? (# of clashes, SC, interface packing score)
- Is the interface large and energetically favorable? (buried surface area, Rosetta interface energy, dG)
- Is the interface chemistry reasonable? (H-bonds, buried polar atoms, hydrophobicity)

Filter:

- Combine PyRosetta measurements with AF2 confidence and structural recovery
- Advance only designs supported by several independent criteria

Final filtering: which designs enter the accepted set?



UT Southwestern Medical Center

Final filtering: which designs enter the accepted set?

Seed

BindCraft applies configurable thresholds to AF2 confidence, structural recovery, interface geometry, and Rosetta metrics.

Example:

AF2:

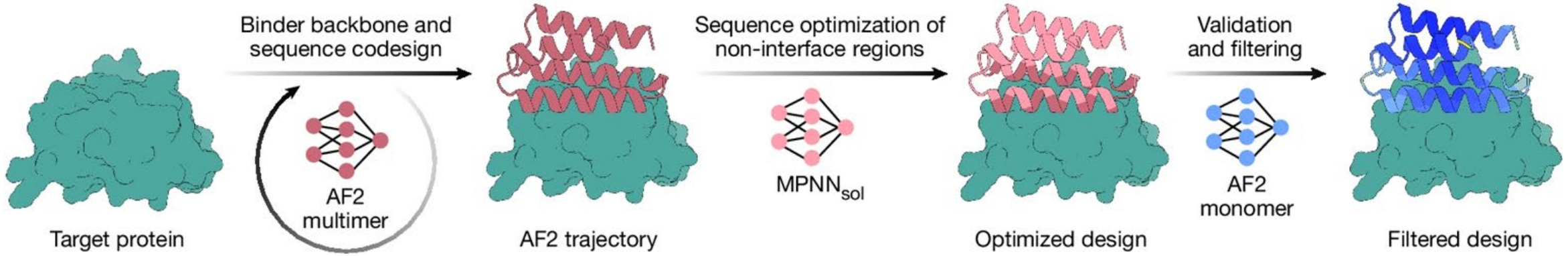
- complex pLDDT better than 0.8
- interface confidence (ipTM) higher than 0.5

Rosetta Interface Shape:

- shape complementarity higher than 0.6
- more than 3 H-bonds at interface
- 0.35 Å RMSD between bound and unbound structure

A design must pass all required filters.

What happens after BindCraft?

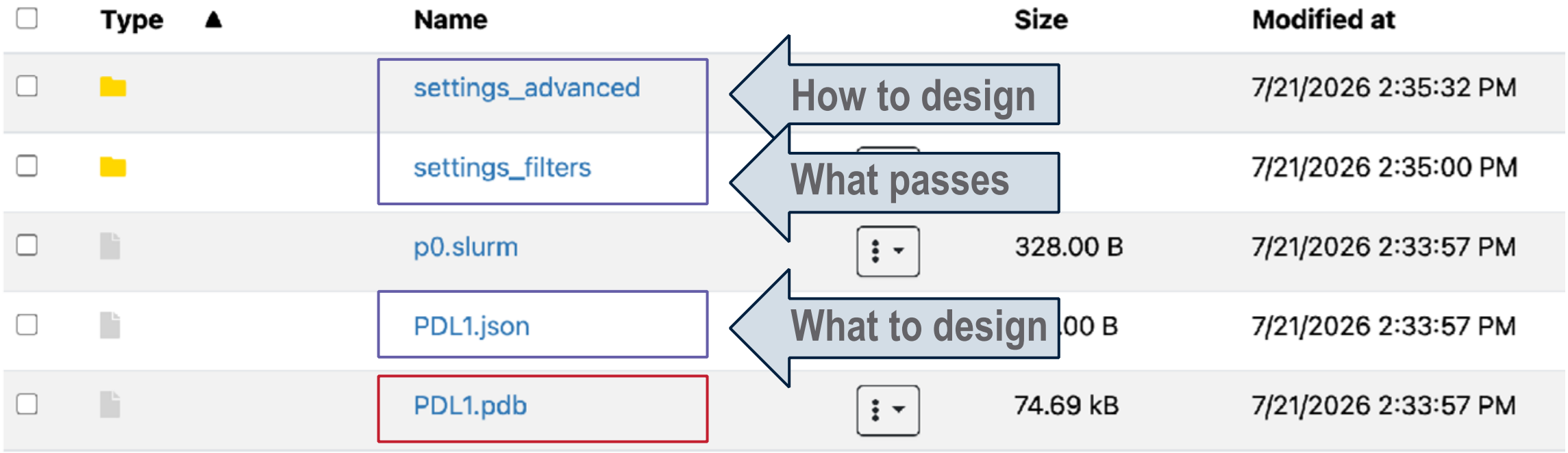


The next stage is to make the protein and test, in sequence, whether it can be produced, folds correctly, binds specifically, and performs the intended biological function.

UT Southwestern Medical Center

What does a BindCraft run require?

A BindCraft campaign requires a prepared target structure, three configuration files, a Linux software environment, and one CUDA-capable NVIDIA GPU per running process.

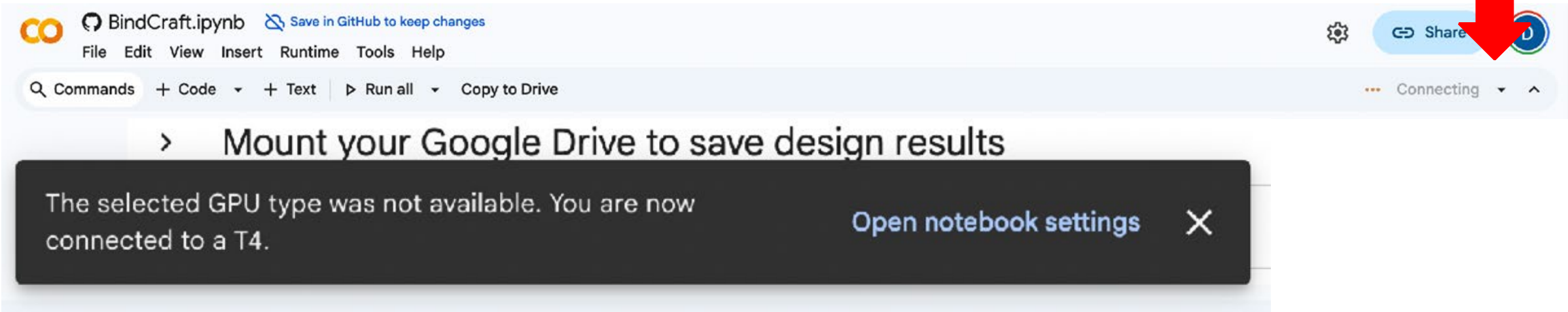


UT Southwestern Medical Center

Notebooks on Google Colab

- Web-form with coding interface
- Runs on Google servers

https://colab.research.google.com/github/martinpacesa/BindCraft/blob/main/notebooks/BindCraft.ipynb



- Protein modelling tools
  - ColabFold
  - ColabDesign
  - Follow @sokrypton.org on bsky

UT Southwestern Medical Center

Launching BindCraft

The same BindCraft pipeline can be launched:

- On a SLURM cluster
- On a local GPU workstation
- Tamarind Bio has BindCraft

Once launched, the pipeline continues automatically until it reaches the requested number of accepted designs or a stopping condition is triggered.

GPU memory determines the maximum design size

Plan GPU memory using the total modeled length: the retained target residues plus the largest possible binder (total amino acids).

1 GPU ⇒ 1 trajectory ⇒ 10 minutes on

GPU: model and performance determine how quickly the trajectory runs.

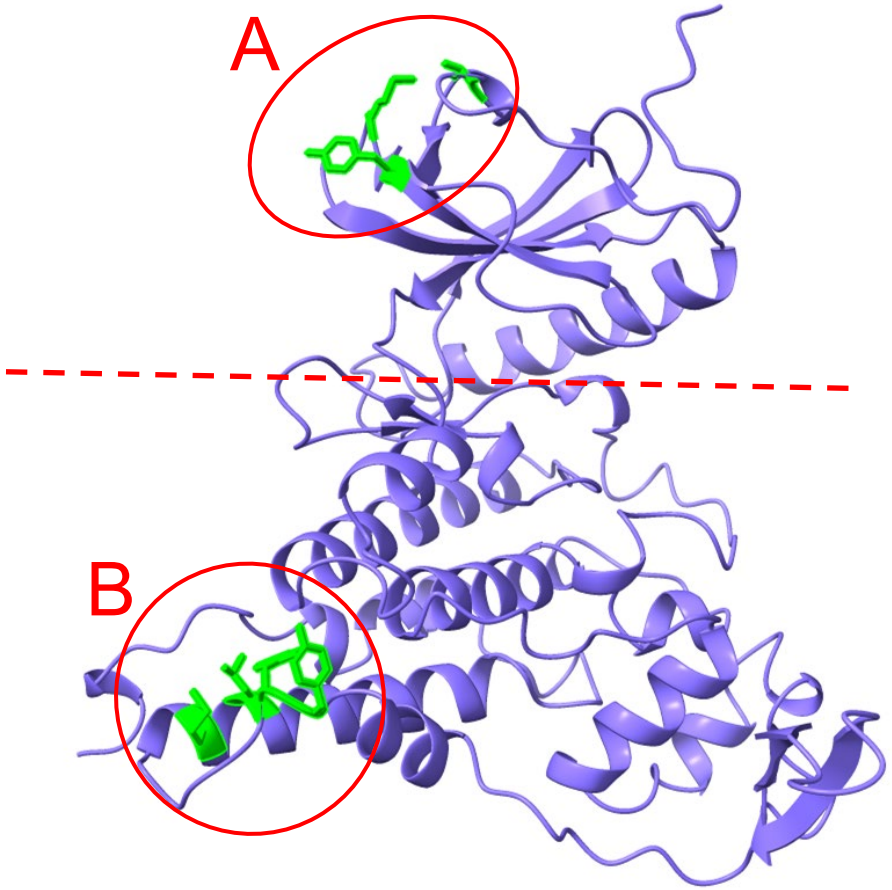
GPU VRAM: determines whether the trajectory fits; used in the AF2 calculation; usually the critical limit.

- 32 GB VRAM: Fits ~ 550 total amino acids (e.g., a ~ 300-residue target plus a ~ 250-residue binder).
- 48 GB VRAM: Fits ~ 750 total amino acids.
- 80 GB VRAM: Fits ~ 950 total amino acids.

System RAM: used by Python, structure processing and PyRosetta.

BindCraft cannot be parallelized across multiple GPUs. However, multiple instances of BindCraft jobs can be run on multiple GPUs, with job outputs into the same output folder to speed up design sampling.

Binder design with BindCraft – What to do?



- **Webform input:**
  - Target: 320 aa
  - Binder: 100 aa, site A

Out of memory!

- **Truncate C-term:**
  - N-term subdomain: 100 aa

Binder sticks to truncation site!

- **Mutate exposed res:**
  - Hydrophobics > A

Binder overlaps with truncated C-term domain

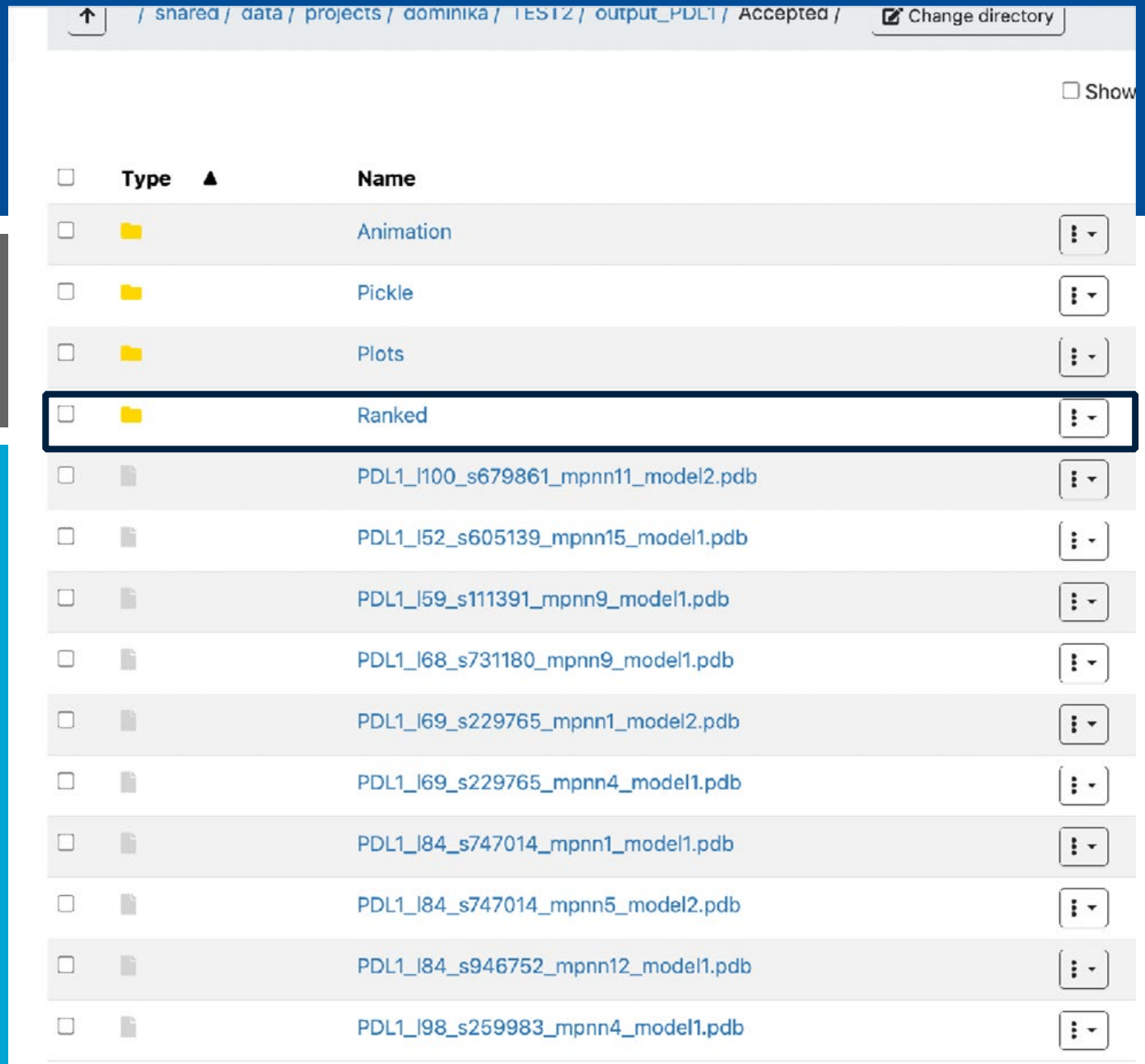
- **Modify site A**
  - Select residues away from C-term subdomain

Slide courtesy of Isabelle Phan

UT Southwestern Medical Center

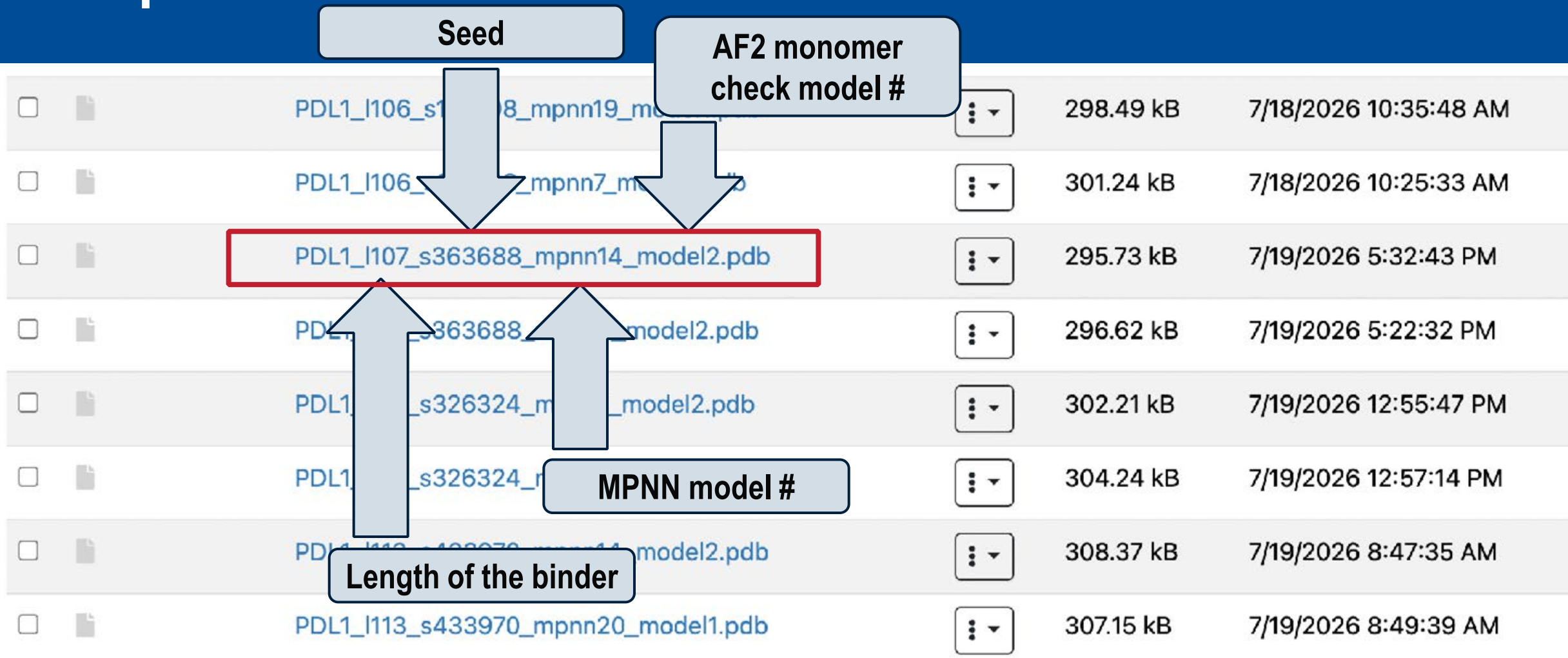
Where are my results?

Type	Name	Size	Modified at
folder	output_PDL1	-	7/16/2026 8:44:45 PM
folder	settings_advanced	-	7/16/2026 8:43:03 PM
folder	settings_filters	-	7/16/2026 8:44:10 PM
file	p0-a02.slurm	395.00 B	7/16/2026 8:27:05 PM
file	p0_bindcraft_7910.err	0.00 B	7/16/2026 8:44:31 PM
file	p0_bindcraft_7910.log	527.28 kB	7/17/2026 4:19:04 AM
file	PDL1.json	221.00 B	7/16/2026 8:25:53 PM
file	PDL1.pdb	74.69 kB	7/16/2026 8:25:44 PM



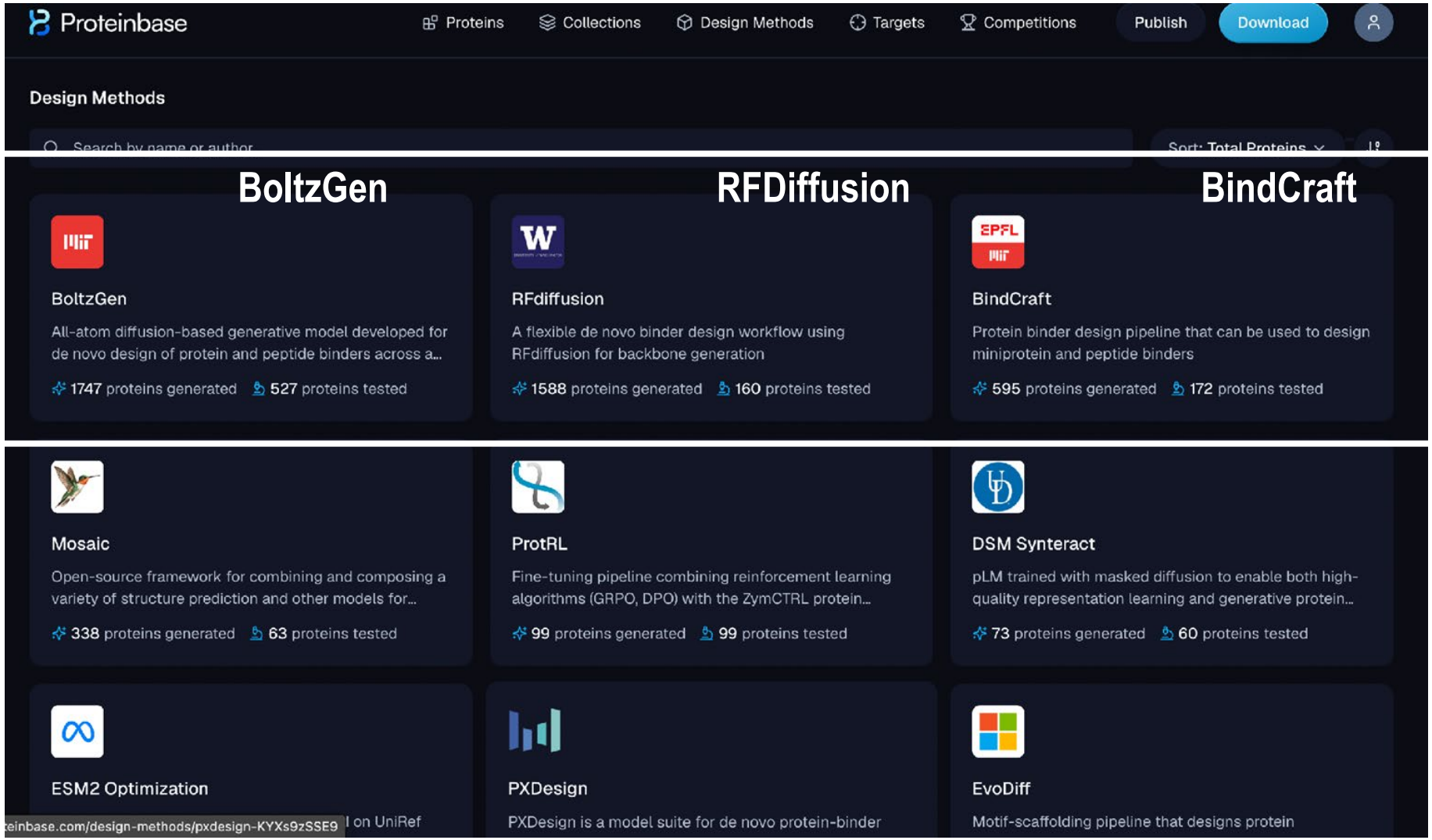
UT Southwestern
Medical Center

Output files



UT Southwestern Medical Center

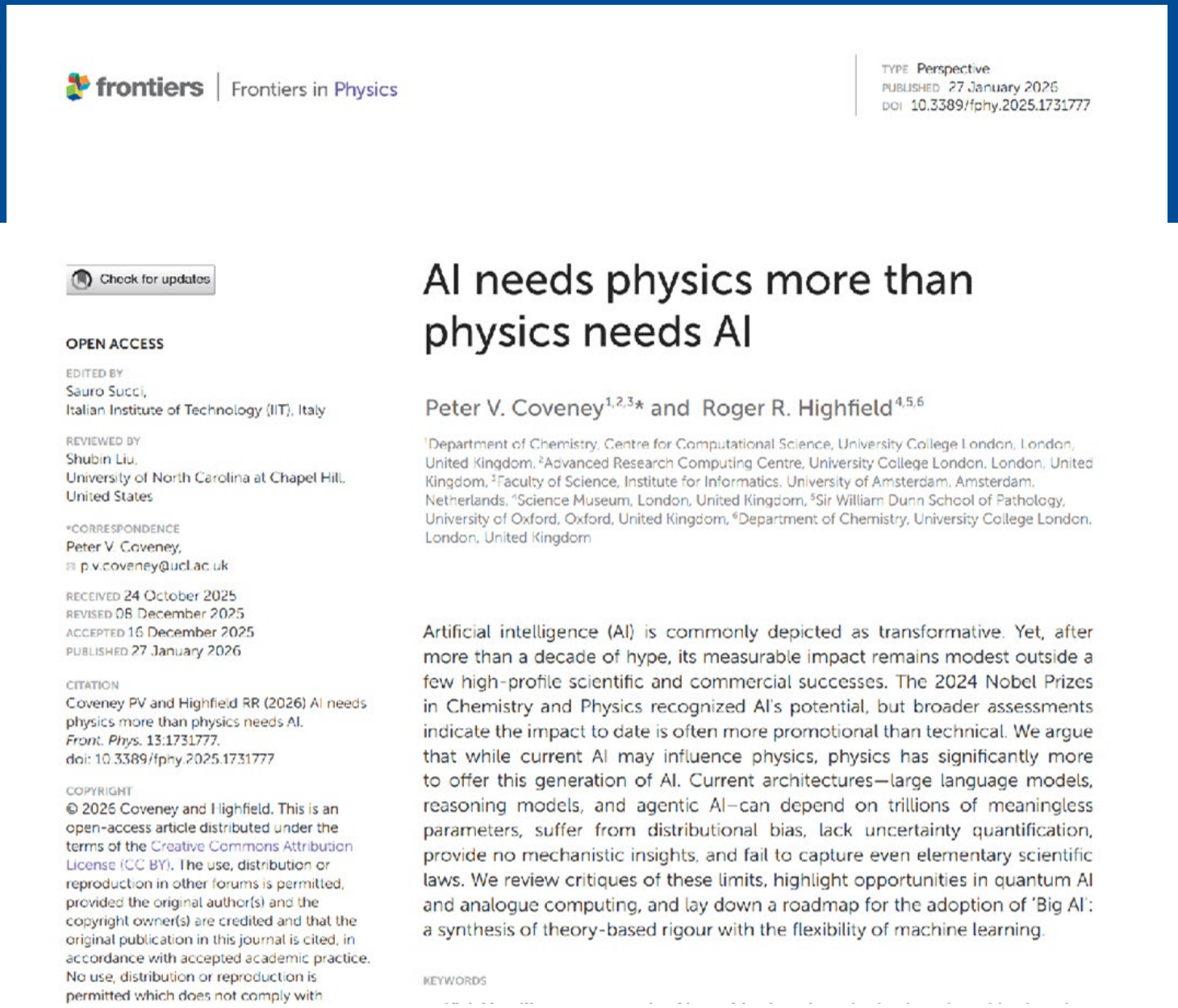
Other tools? – Many! Check frequently!



UT Southwestern Medical Center

Which one is the best???

- The bottleneck is not computational…
- We don't have a good measure of success other than expression, purification, and binding affinity measurements for each structural hypothesis separately;
- Plausible binders frequently do not bind = there is a problem…



frontiers | Frontiers in Physics

TYPE Perspective
PUBLISHED 27 January 2026
DOI 10.3389/fphy.2025.1731777

OPEN ACCESS

EDITED BY
Sauro Succi,
Italian Institute of Technology (IIT), Italy

REVIEWED BY
Shubin Liu,
University of North Carolina at Chapel Hill, United States

*CORRESPONDENCE
Peter V. Coveney,
p.v.coveney@ucl.ac.uk

RECEIVED 24 October 2025
REVISED 08 December 2025
ACCEPTED 16 December 2025
PUBLISHED 27 January 2026

CITATION
Coveney PV and Highfield RR (2026) AI needs physics more than physics needs AI.
Front. Phys. 13:1731777.
doi: 10.3389/fphy.2025.1731777

AI needs physics more than physics needs AI

Peter V. Coveney[1,2,3]* and Roger R. Highfield[4,5,6]

[1]Department of Chemistry, Centre for Computational Science, University College London, London, United Kingdom, [2]Advanced Research Computing Centre, University College London, London, United Kingdom, [3]Faculty of Science, Institute for Informatics, University of Amsterdam, Amsterdam, Netherlands, [4]Science Museum, London, United Kingdom, [5]Sir William Dunn School of Pathology, University of Oxford, Oxford, United Kingdom, [6]Department of Chemistry, University College London, London, United Kingdom

Artificial intelligence (AI) is commonly depicted as transformative. Yet, after more than a decade of hype, its measurable impact remains modest outside a few high-profile scientific and commercial successes. The 2024 Nobel Prizes in Chemistry and Physics recognized AI's potential, but broader assessments indicate the impact to date is often more promotional than technical. We argue that while current AI may influence physics, physics has significantly more to offer this generation of AI. Current architectures—large language models, reasoning models, and agentic AI—can depend on trillions of meaningless parameters, suffer from distributional bias, lack uncertainty quantification, provide no mechanistic insights, and fail to capture even elementary scientific laws. We review critiques of these limits, highlight opportunities in quantum AI and analogue computing, and lay down a roadmap for the adoption of 'Big AI': a synthesis of theory-based rigour with the flexibility of machine learning.

KEYWORDS

Proteinbase runs competitions and validates experimentally

BoltzGen

1. 1747proteins generated
2. 527 / 1747proteins validated
3. 93% expression rate
4. Binding Strength tested for 22%

Out of tested:

~82% does not bind

~9% strong binders

~9% medium binders

Less than 1% weak binders

RFDiffusion

1. 1588 proteins generated
2. 160 / 1588 proteins validated
3. 97% expression rate
4. Binding Strength tested for 7%

Out of tested:

~70% does not bind

~14% strong binders

~14% medium binders

Less than 1% weak binders

BindCraft

1. 595 proteins generated
2. 172 / 595 proteins validated
3. 97% expression rate
4. Binding Strength tested for 18%

Out of tested:

~78% does not bind

~17% medium binders

~5% weak binders

Less than 1% strong binders

UT Southwestern Medical Center

Acknowledgments

- All organizers
- Zbyszek Otwinowski's lab
- Yirui Guo (Ligo)
- All students (shoutout to Yingying He for letting me know about BoltzGen)
- All instructors (thank you Phil for checking SBGrid)
- SBGrid for stepping in in the last moment:
  - Ben Eisenbraum
  - Pete Meyer
  - Piotr Sliz