Article

https://doi.org/10.1038/s41589-025-01929-w

Accurate de novo design of high-affinity protein-binding macrocycles using deep learning

Received: 31 December 2024

Accepted: 2 May 2025

Published online: 20 June 2025

Check for updates

Stephen A. Rettie[1,2,3,15], David Juergens[2,4,15], Victor Adebomi[1,2,15], Yensi Flores Bueso[1,2,5,6], Qinqin Zhao[7], Alexandria N. Leveille[8], Andi Liu[7], Asim K. Bera[2], Joana A. Wilms[9,10], Alina Üffing[9,10,14], Alex Kang[2], Evans Brackenbrough[2], Mila Lamb[2], Stacey R. Gerben[2], Analisa Murray[2], Paul M. Levine[2], Maika Schneider[1,2,11], Vibha Vasireddy[1,2], Sergey Ovchinnikov[12], Oliver H. Weiergräber[10], Dieter Willbold[9,10], Joshua A. Kritzer[8], Joseph D. Mougous[7,13], David Baker[2,5,13] ✉, Frank DiMaio[2,5] ✉ & Gaurav Bhardwaj[1,2] ✉

A bit of background

Sorry in case you already know it

Sequence → Structure

The protein folding revolution: AlphaFold2

- General solution to protein structure prediction
- Predicts accurate protein structures in minutes
  - As opposed to months - years for experimental determination

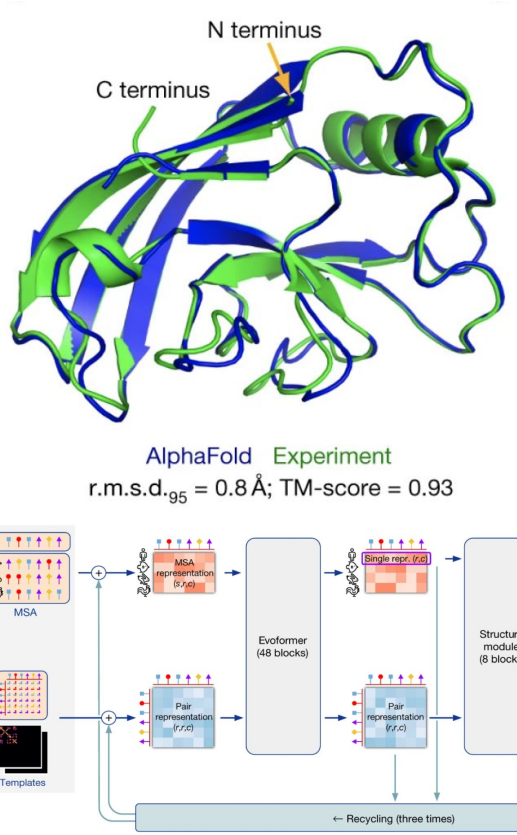


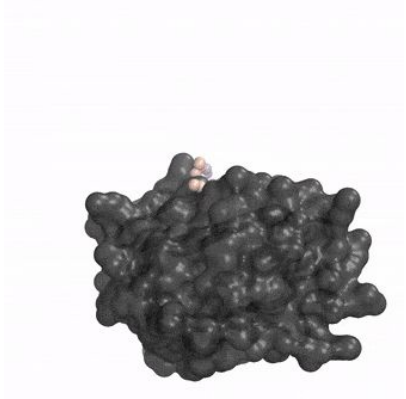
Image from: Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

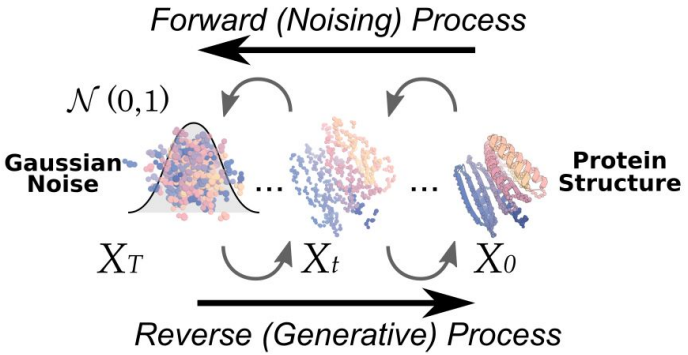
Very short intro to RFDiffusion

- DDPM on top of a structure prediction network

- conditioned on the target structure

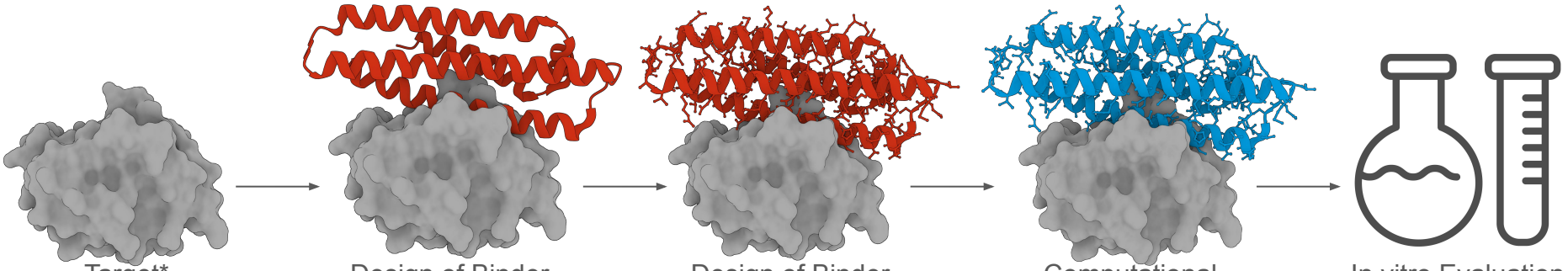
- will yield backbone of binding protein





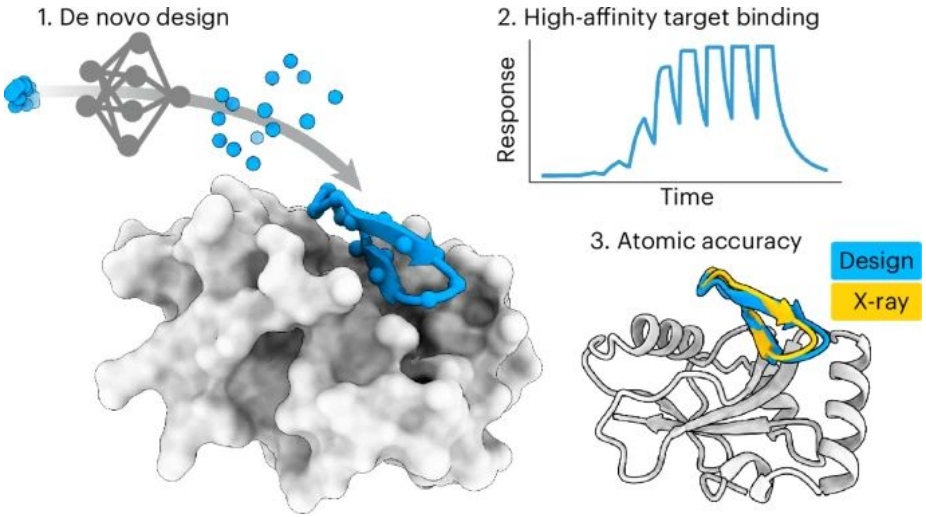
Figures from: Ian Haydon, IPD @ UW;
10.1126/science.ade2574

What is Protein Design - all together



* Target can also be of small molecule or nucleic acid type

The paper


1. De novo design
2. High-affinity target binding
Response
Time
3. Atomic accuracy
Design
X-ray

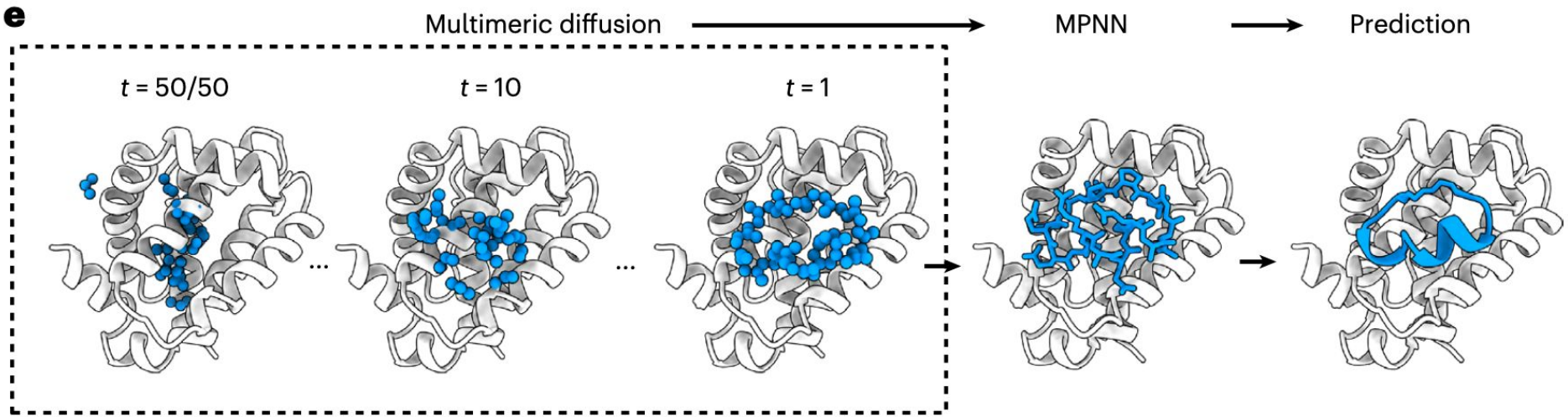
Showcase

Like in a cooking show

Challenges for current Macrocycle design

- time and cost-intensive HTS for discovery
- low mutational tolerance for hits
- multi-objective optimisation not possible
  - target binding
  - selectivity and membrane permeability

The RFpeptides pipeline

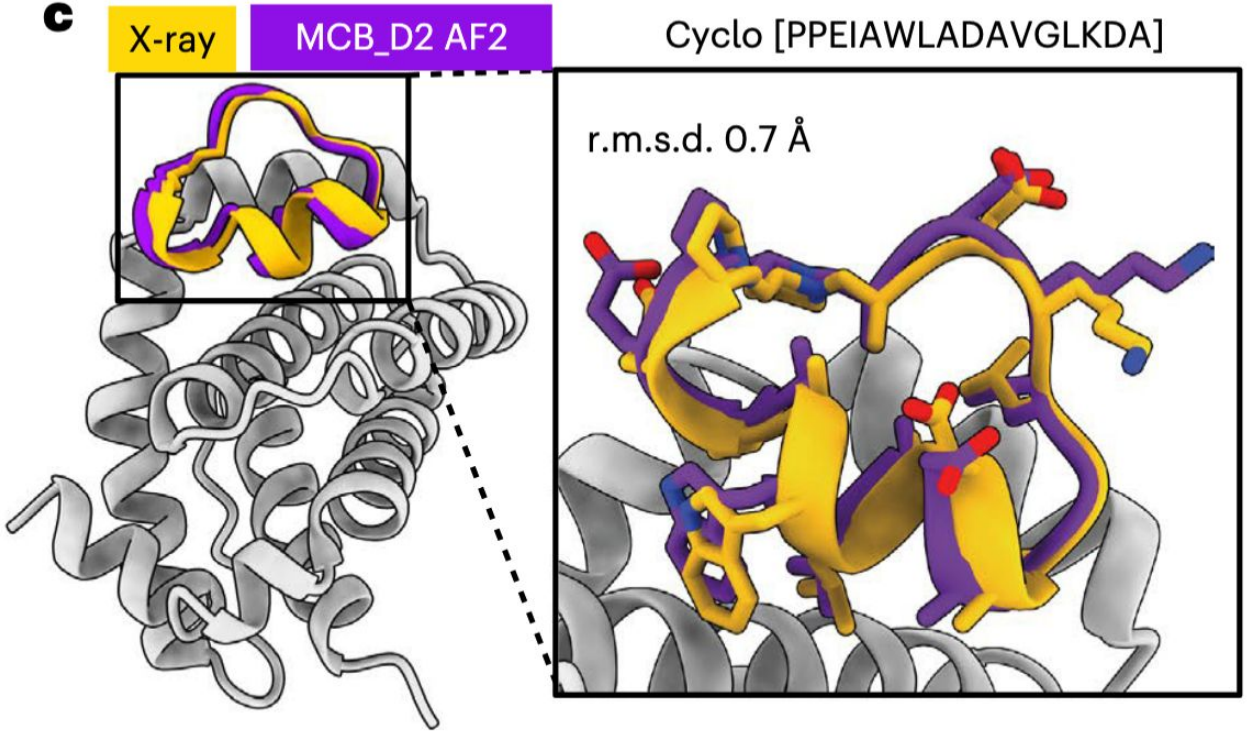


The design campaign

- myeloid cell leukemia 1 (MCL1) protein
  - part of the Bcl2 protein family (blocking cell death)
  - regulatory role in autophagy, cell survival, DNA repair and cellular proliferation
  - Aim: Inhibiting MCL1 to bind other proteins

- 9,965 structures → ProteinMPNN and AfCycDesign →
- 1,984 structures → physics-based in silico filtering →
- 27 designs for Fmoc-based solid-phase synthesis →
- 14 designs with acceptable yields→
- 3 designs binding →
- 1 best binder: 2 μM (SPR)

The design campaign

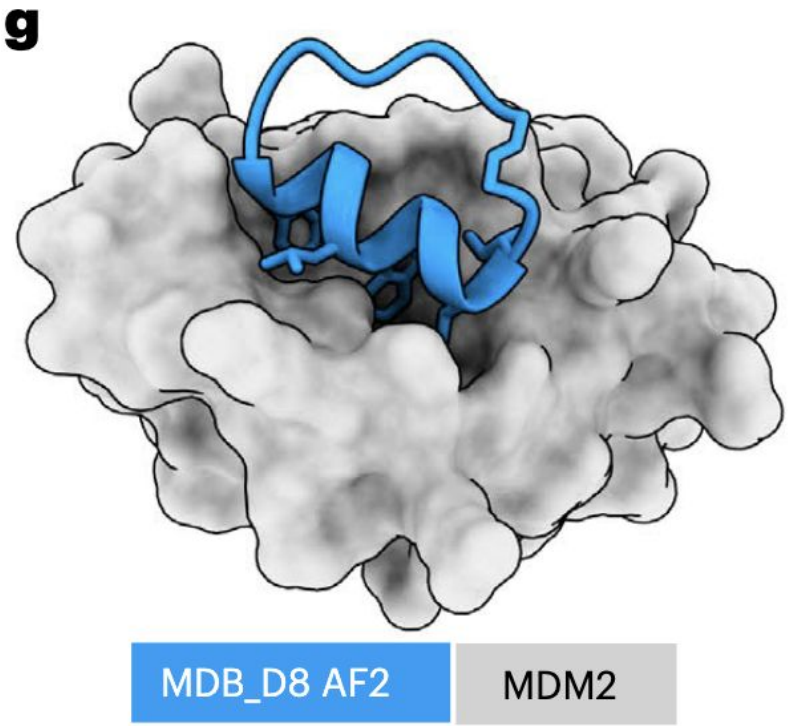


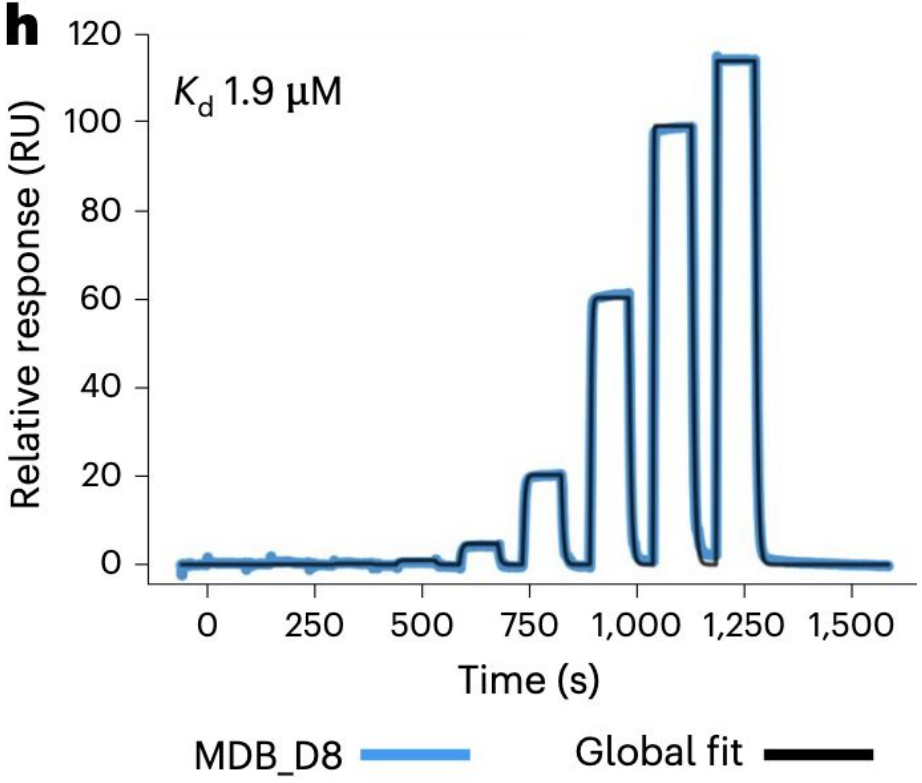
The design campaign

- MDM2
  - E3 ligase that interacts with tumor suppressor protein p53
  - critical roles in tumor growth and survival
  - Aim: Improve binding of ligase

- 40,000 structures → ProteinMPNN and AfCycDesign →
- 7,495 structures → physics-based in silico filtering →
- 11 designs for Fmoc-based solid-phase synthesis →
- 8 designs with acceptable yields→
- 3 designs binding →
- 1 best binder: 1.9 μM (SPR)

The design campaign





The design campaign

- γ-aminobutyric acid type A receptor-associated protein (GABARAP)
  - mediating autophagy through its role in autophagosome biogenesis and recruitment of cargo
  - resulting in lysosomal degradation of damaged or surplus proteins and organelles
  - therapeutic applications: late stage cancer, protein degradation

- 80,000 structures → ProteinMPNN and AfCycDesign + clustering →
- 335 structures → physics-based in silico filtering →
- 13 designs for Fmoc-based solid-phase synthesis →
- 6 designs with acceptable yields→
- 2 designs binding →
- 1 best binder: 6 nM (SPR) → capable of inhibiting the interface with K1

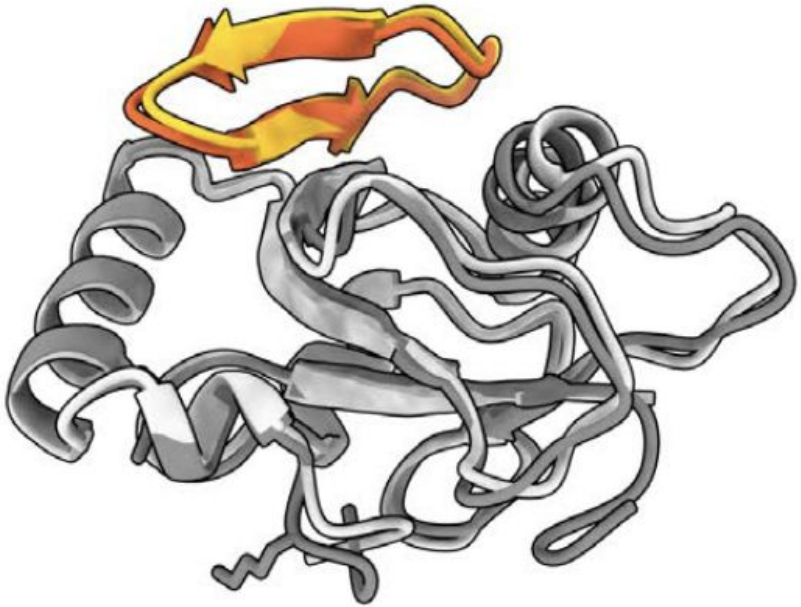
c

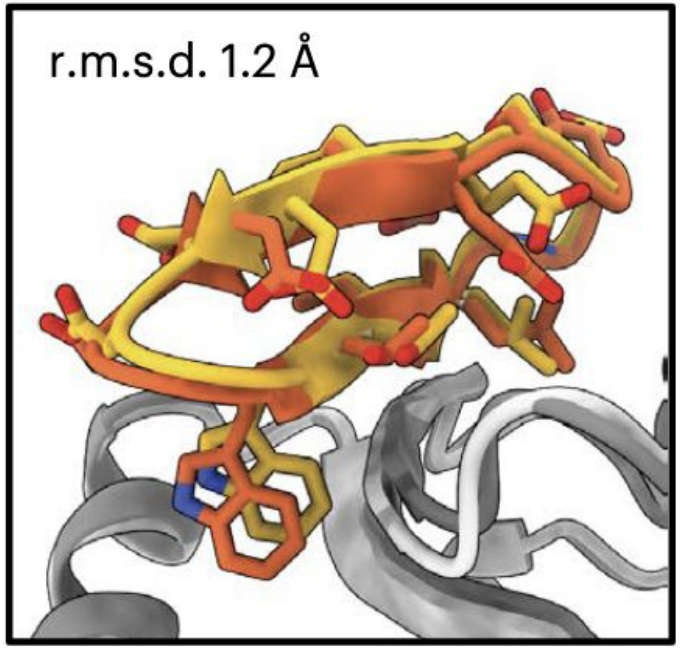
X-ray

GABARAPL1 | GAB_D8

AF2

GABARAP | GAB_D8





In short

- Achieves high-affinity binders with fewer than 20 candidates per target (vs. trillions in library-based approaches)
- Enables targeted design for specific protein patches and sites
- Provides atomically accurate models for structure-guided optimization
- Allows optimization for multiple properties simultaneously (binding affinity, membrane permeability)
- Bypasses complex structure determination bottlenecks

They do not address

- Oral availability
- Expensive synthesis
- Lack of broader applicability (only 3 targets)
- Off-target interactions
- Metabolic stability
- Pharmakokinetics, Immunogenicity, ADMET, Toxicity etc. (*)

Showcase

Present the results