Empowering Next-Generation Drug Discovery via Agentic AI

www.sgo.or.kr



Woo Youn Kim
(KAIST / HITS Inc.)

DECLARATION OF INTERESTS

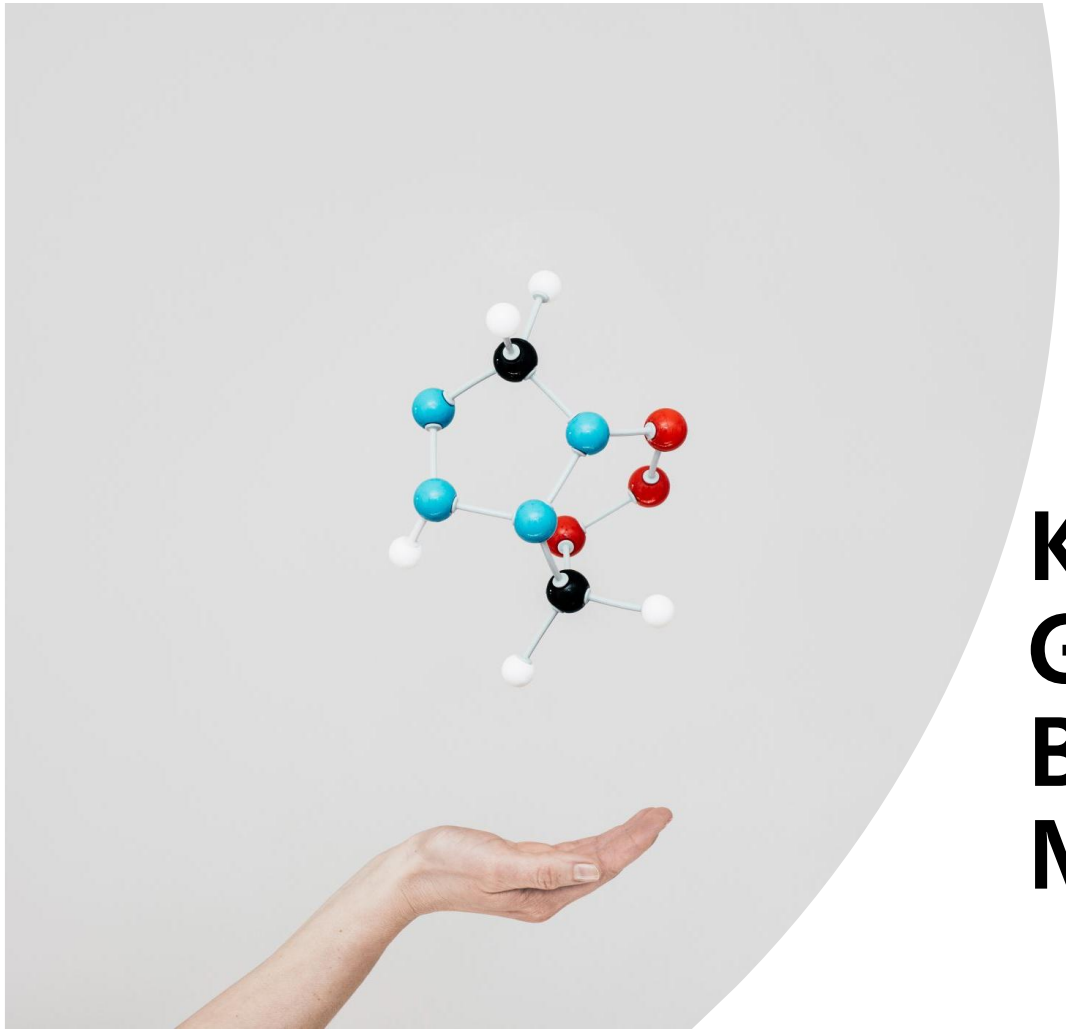
- **Employment / Leadership**
  CEO & Co-founder, HITS Inc.
- **Stock ownership**
  HITS Inc.
- **Products discussed in this talk**
  K-Fold and HyperLab are developed by KAIST and HITS Inc., respectively.

Contents

1. K-Fold: Next Generation of Bio-Foundation Model

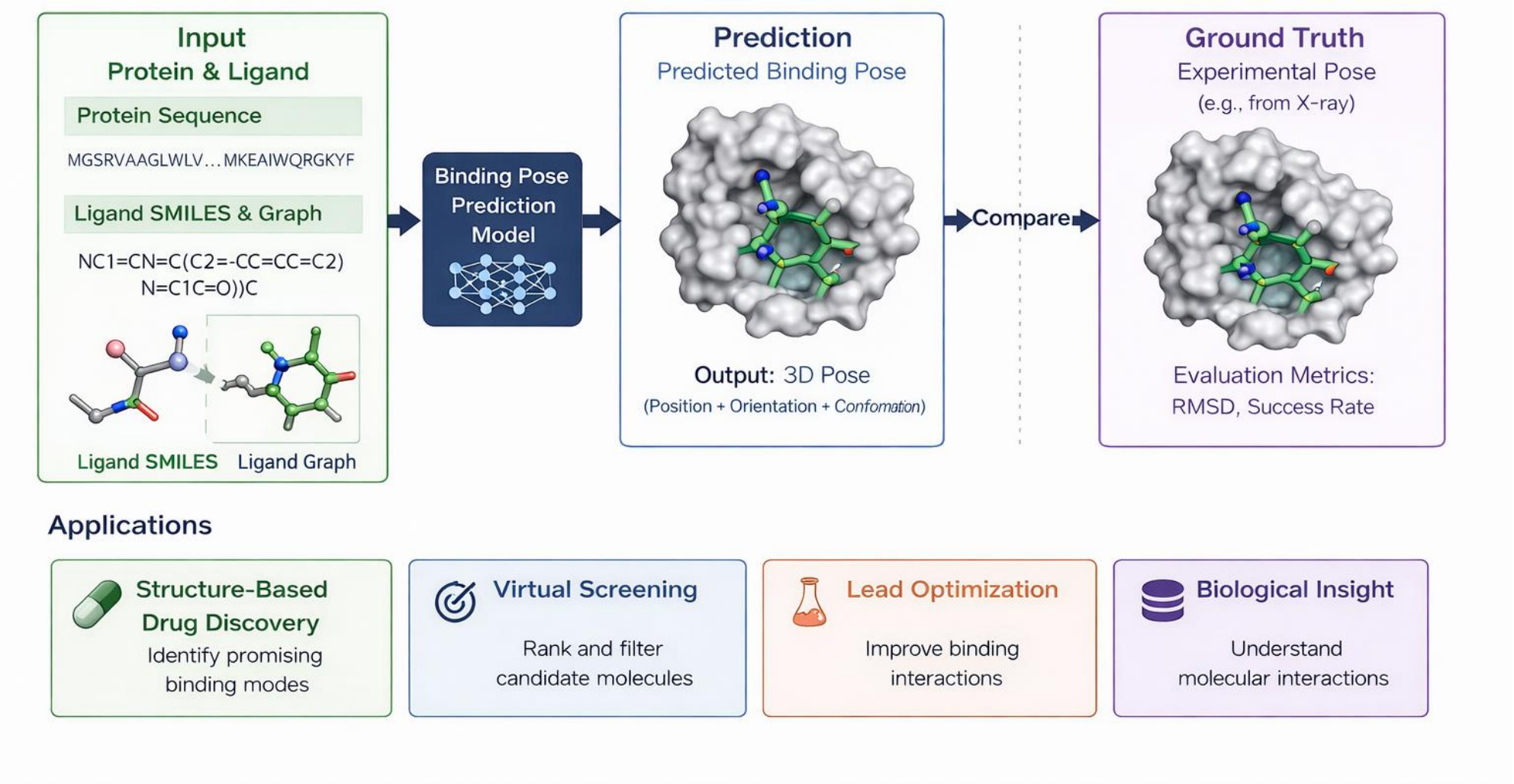
2. HyperLab: Bio Co-Scientist

3. Conclusion



K-Fold: Next Generation of Bio-Foundation Model

Importance of Biomolecular Structure Prediction



2024 Nobel Prize in Chemistry



Ill. Niklas Elmehed © Nobel Prize Outreach

David Baker

Prize share: 1/2



Ill. Niklas Elmehed © Nobel Prize Outreach

Demis Hassabis

Prize share: 1/4



Ill. Niklas Elmehed © Nobel Prize Outreach

John Jumper

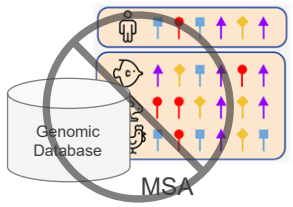
Prize share: 1/4

- David Baker: "**for computational protein design**“ ➔ 불가능 하다고 믿었던 일.
- Demis Hassabis and John Jumper: "**for protein structure prediction**“ ➔ 50년 난제

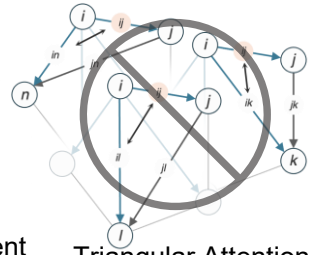
Next-Generation Bio Foundation Model: K-Fold

We re-define the generative modeling in bio as *"**Structure Prediction** ➔ **Structure Change Prediction**"*

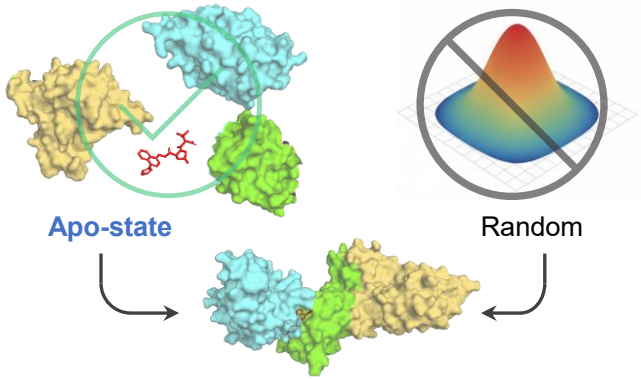
Efficient/Faster Modeling



Multiple Sequence Alignment



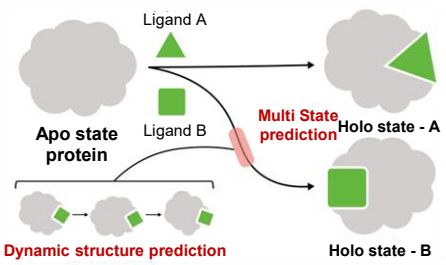
Triangular Attention



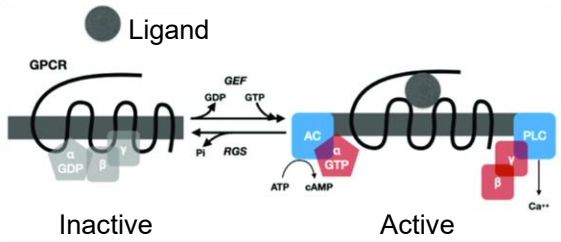
Efficient/fast structure prediction pipeline

Implementation of a protein-complex data structure prediction framework

Modeling of Multiple States



Multi-state modeling of proteins **induced** by ligand binding

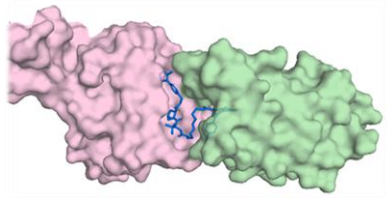


Modeling of **Signaling Pathways**
(G-Protein Coupled Receptor)

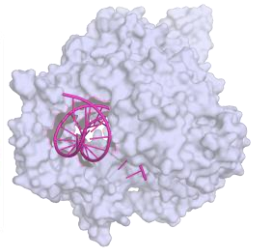
Clarify biological mechanisms

Model dynamic biological pathways

Generalization on Multiple Modalities



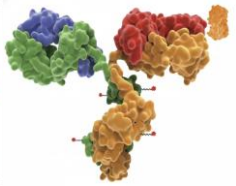
Small Molecule Compounds
(Yuhan Corporation - Leclaza)



SARS-CoV-2 Virus
mRNA-Protein Complex



GLP-1-like **Synthetic Protein**
(Novo Nordisk - Wegovy)



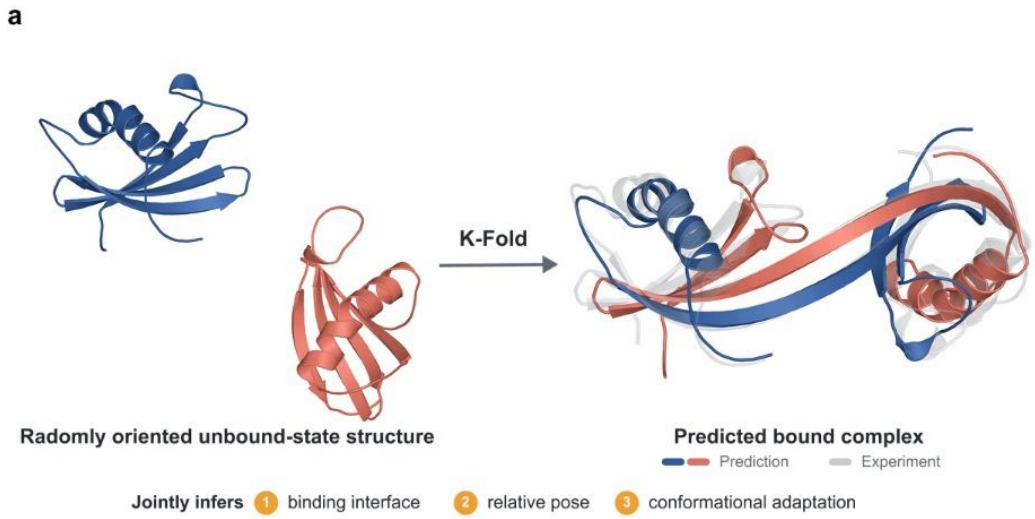
Antibody-drug Conjugate
(molecular glue)

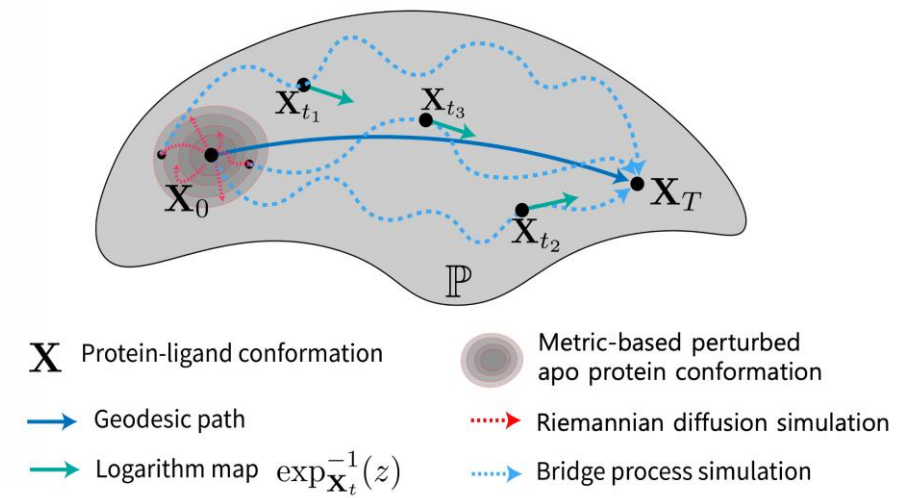
Change in the drug design paradigm

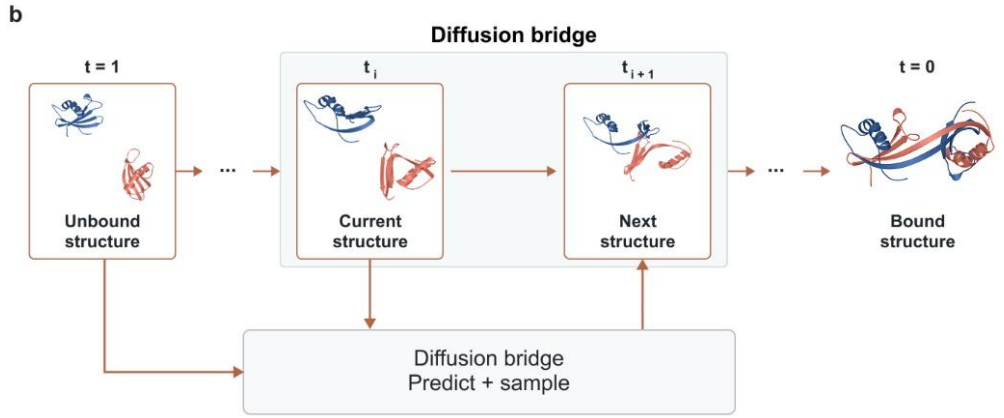
Combined understanding and design of multiple molecular modalities

Conformational Change Prediction with Diffusion Bridge

- **Goal: Learning Apo-to-Holo structure changes**
- A **distribution shift problem** with **diffusion bridge modeling**





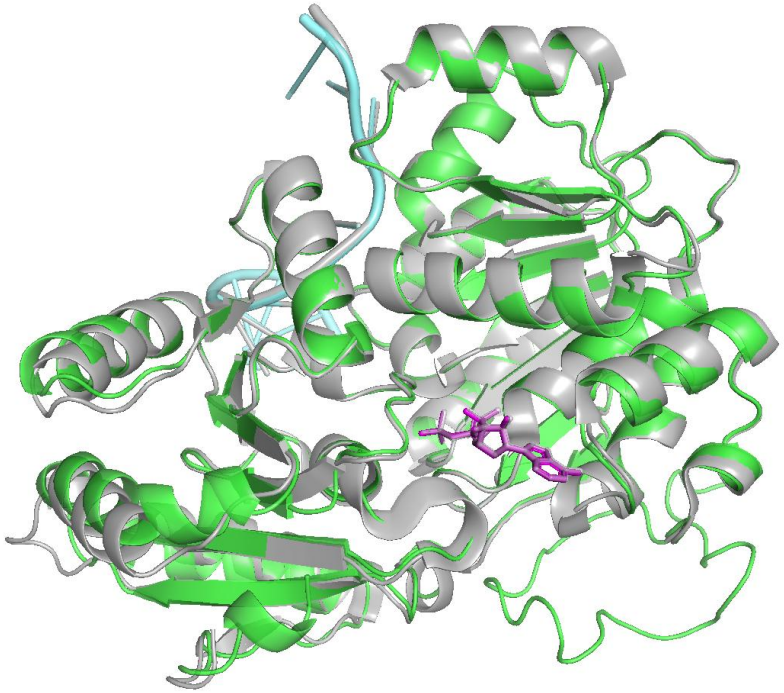


1. Start with apo states
2. Simulate structural change
3. No MSA required

Prediction Example: Multi-modality Systems

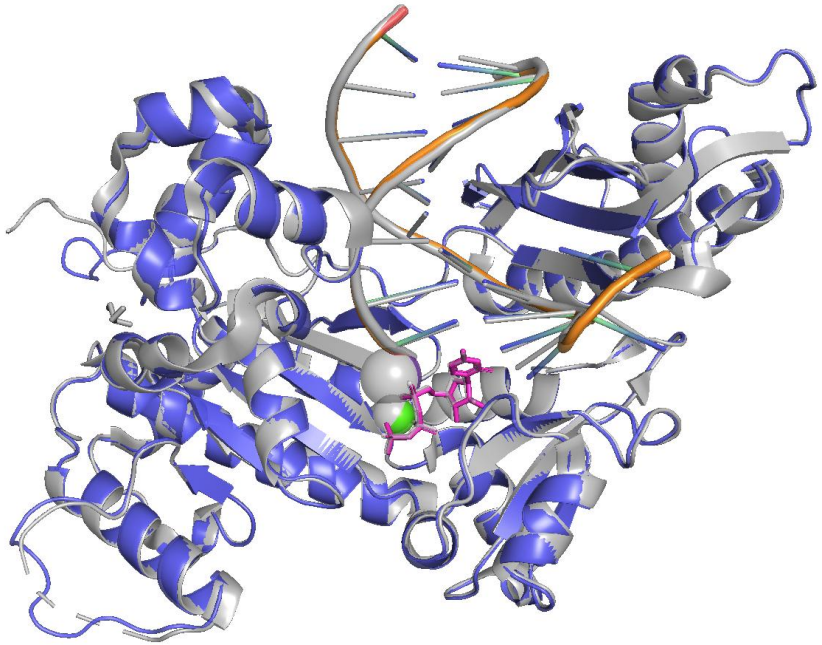
K-Fold predicts three-dimensional structures of **multi-modal, multi-molecular** systems.

(a) 9AVR (Protein, RNA, molecule)



K-Fold predicted (best of 5)

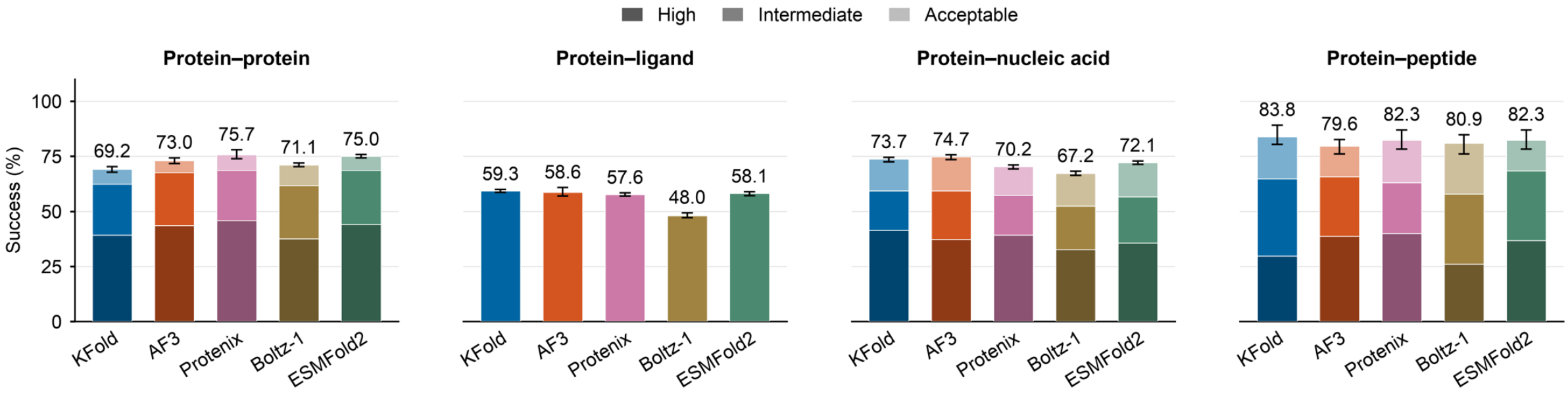
(b) 8V7B (Protein, DNA, molecule, ions)



K-Fold predicted (best of 5)

Ground truth

K-Fold Benchmark Results: General Set



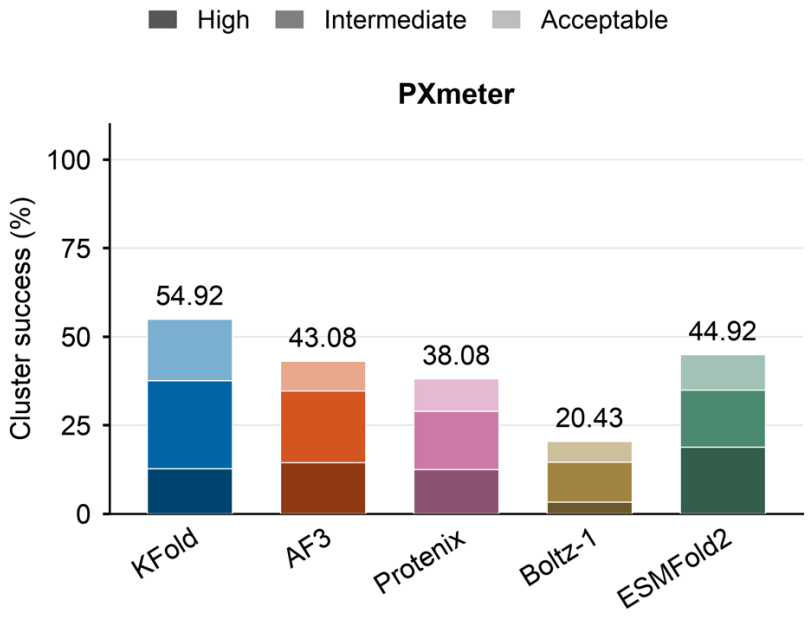
10 seeds x 5 samples ➔ Top1 of $_{10}C_5$ combinations

K-Fold: ~10 s per complex, against ~25 min for AlphaFold3 (NVIDIA B200).

FoldBench

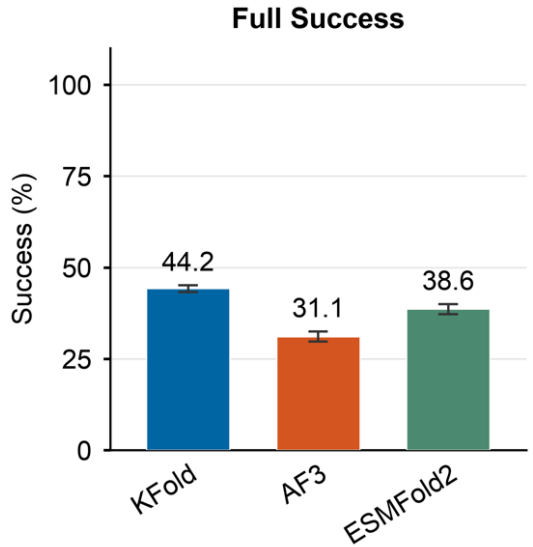
K-Fold Benchmark Results: Challengeable Sets

Antibody-Antigen



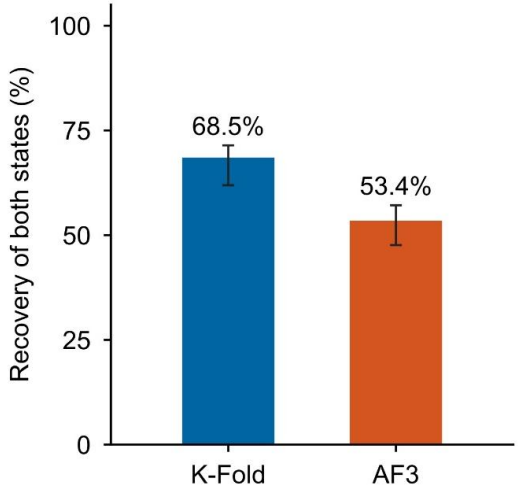
PXM-2025

Molecular Glue



MGBench

GPCR (agonist vs antagonist)



GPCRBench

10 seeds x 5 samples ➔ Top1 of $_{10}C_5$ combinations

K-Fold: ~10 s per complex, against ~25 min for AlphaFold3 (NVIDIA B200).



HyperLab:
Bio Co-Scientist

HyperLab: AI Co-Scientist powered by Multi-Agent System

AI Tools
예측 및 생성 모델 라이브러리

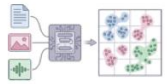
K-Fold
- 차세대 Co-folding 모델
- TPD, mRNA 모달리티 지원
- 빠른 예측 속도



물성 예측 모델
- FM 기반 AI 모델
- 결합 에너지 예측
- 구조 안정성 예측



Foundation Model
- 서열&구조 표현학습



툴 라이브러리 제공

모델 호출

Agentic AI
가상 실험실 운영 체계



PI 에이전트
총괄 조율



실험 에이전트
분자 생성 · 결합 구조 예측
Mutagenesis · 코돈 최적화

생분자



모델 호출

평가 에이전트
물성 예측 기반 평가
결합력 · 안정성 · 독성



학습 에이전트
결과 기록 · 가설 업데이트
Knowledge Base 갱신
wet-lab 결과 우선 학습

○ 총괄 Agent ○ 전문 Agent

고신뢰 데이터 환류

wet-lab 실험 결과
고신뢰 데이터 → 학습 에이전트로 환류

HyperLab Demo



한눈에 보는 스펙

HyperLab 3.0의 핵심 규모를 여섯 개 지표로 정리

1M tokens

최대 컨텍스트

7 종 + 이후 더 추가 예정

탑재 AI 모델

RFdiffusion · Protein/LigandMPNN · K-Fold

120 + α

분석 툴

샌드박스 내 on-the-fly 설치 가능

193 개 + 이후 더 추가 예정

전문 스킬셋

17개 분야

200 + α

연결 데이터베이스

web query 지원

13M nodes

보유하고 있는 바이오 지식그래프

엣지 약 2 억, 출처 보유 100%

무엇을 할 수 있는가

HyperLab 3.0으로 수행 가능한 핵심 작업

1	단백질 바인더·항체 설계	RFdiffusion·ProteinMPNN·K-Fold 기반 de novo 설계
2	복합체 결합 구조 예측	단백질-저분자, 단백질-단백질 등 복합체 구조 예측
3	오믹스 데이터 분석	유전체, 전사체, 단백체 등의 데이터로부터 멀티오믹스 분석
4	문헌 조사 및 정리	관련 문헌 및 지식 그래프 탐색·요약·근거 정리
5	가상 실험 워크플로우	실험 워크플로우 구축 및 자동 실행
6	그 외 에이전트 작업	보고서 작성, 슬라이드, 이미지 생성 등 다양한 데일리 기능들

Case 1: Design of Antibody Targeting AHSP

Workflow of antibody design in HyperLab.

01 Preflight research

Target structure & epitope
by web search and structural analysis.

Design options
Antibody format, CDR, etc.

Campaign scale
Number of generations and predictions.

02 Design pipeline

Backbone generation
RFdiffusion (RFantibody variation)

Sequence design
ProteinMPNN (CDR only)

Co-folding
Structural validation with K-Fold.

03 Candidate selection

Filtering & ranking
Confidence scores, scRMSD, epitope contact, etc.

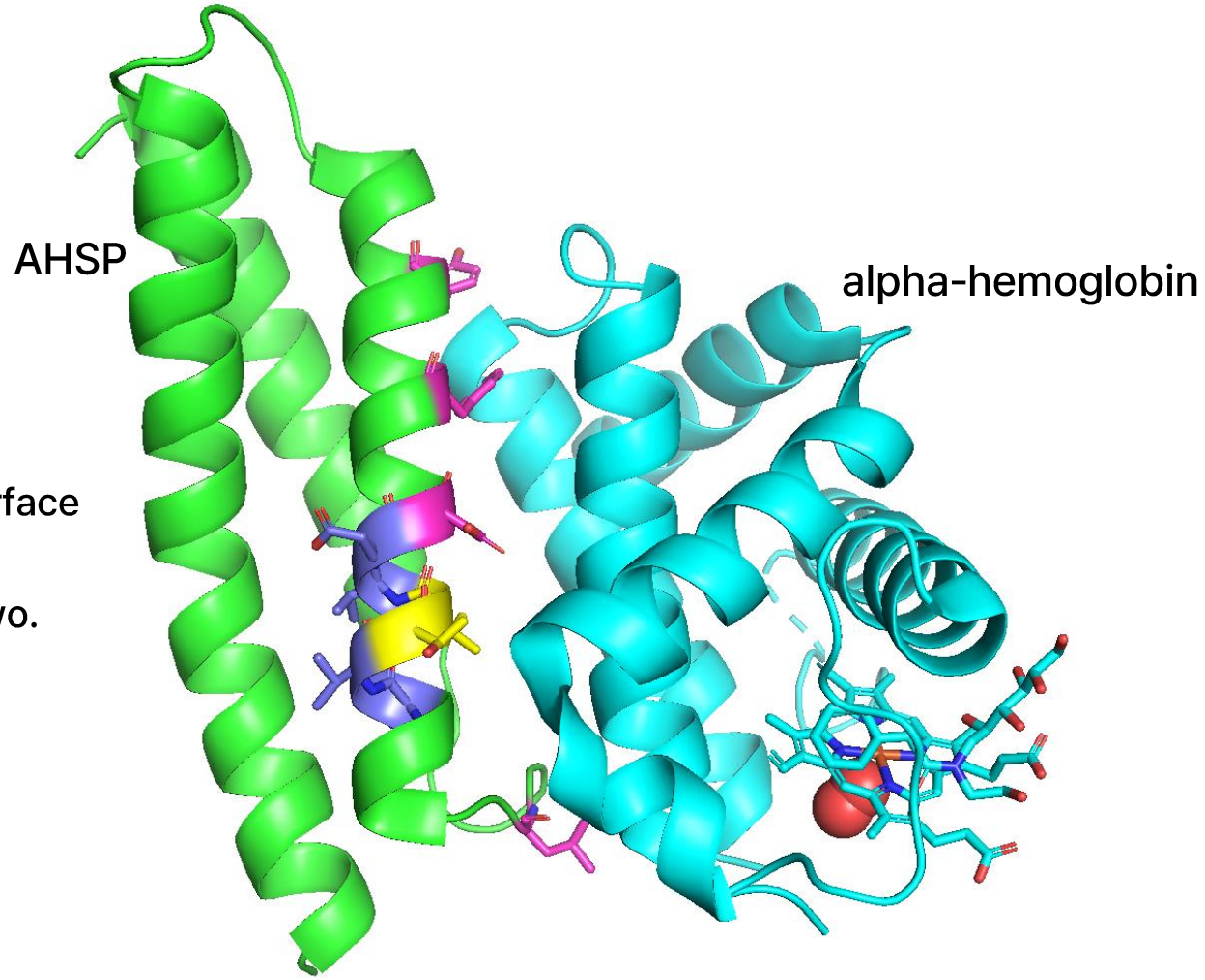
Top recommendation with reasoning

Bennett, N.R. et al. Nature 649, 183–193 (2026).

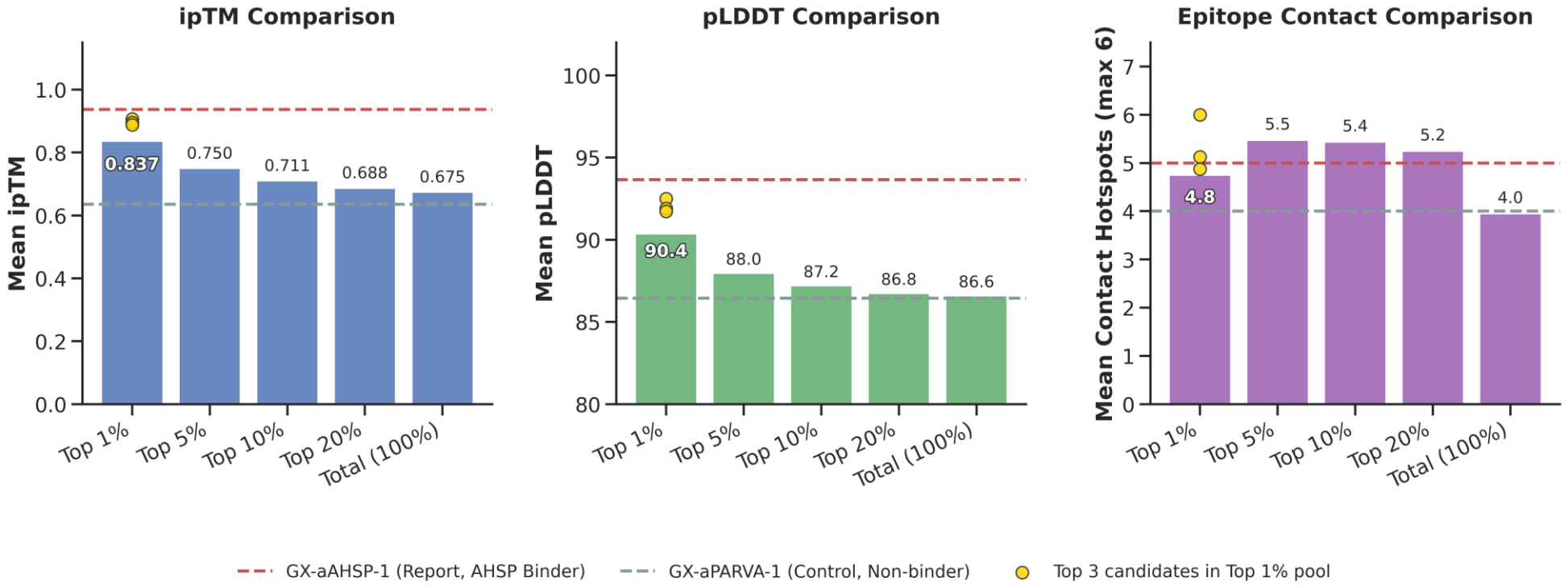
Case 1: Automated Identification of Epitope

- Violet: Epitope (MVTVVE) of a known antibody PCE 5.
- Pink: Residues from native interface & SASA analysis.
- Yellow: Overlap of the above two.

Final suggestion: Pink ∪ yellow.



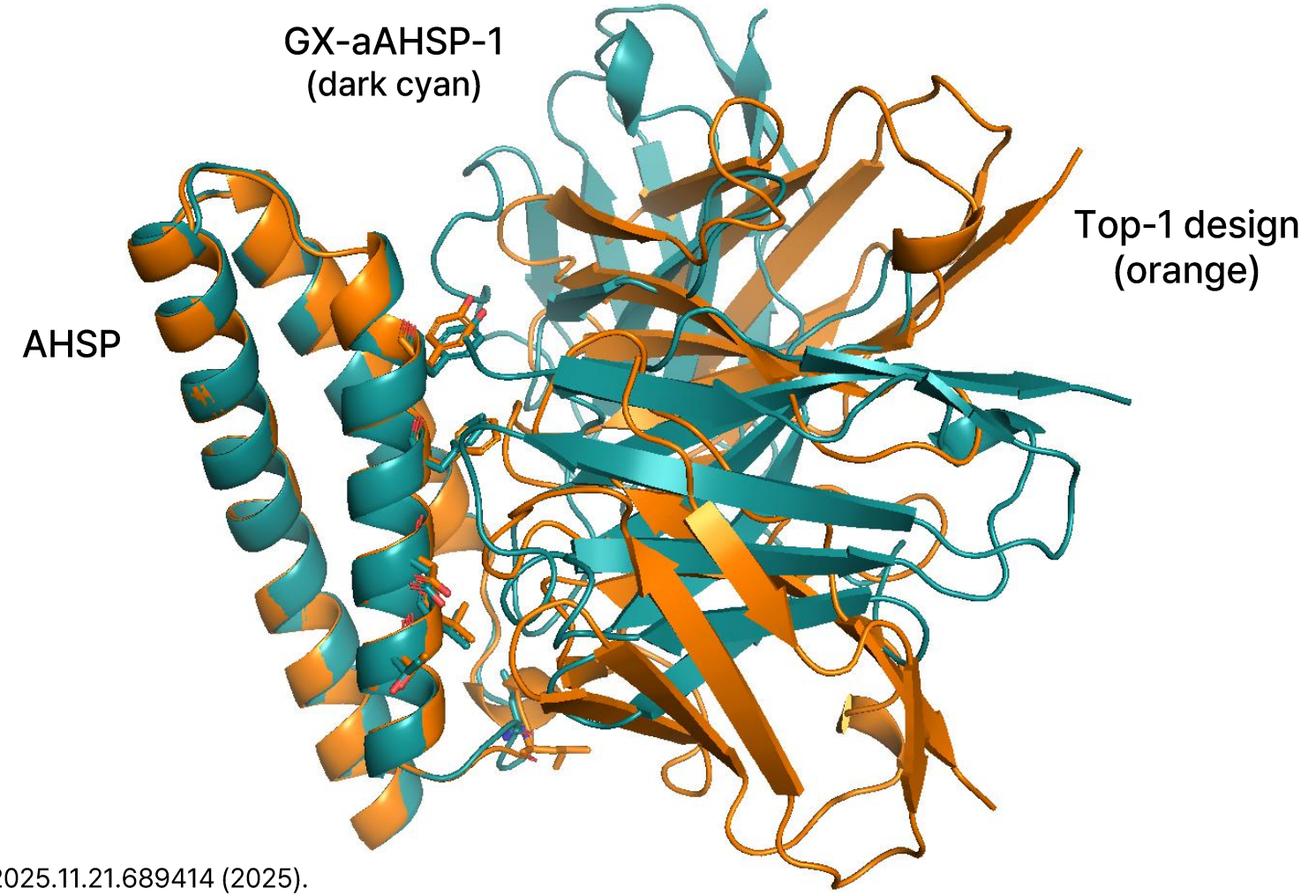
Case 1: Reference Comparison



- Total pool = 800 designs
- Rank by agent := (Sum of normalized ipTM, pLDDT, contact) – 15·scRMSD.

Cha, H. et al. bioRxiv 2025.11.21.689414 (2025).

Case 1: Predicted Structure of a Top-Design Example



Cha, H. et al. bioRxiv 2025.11.21.689414 (2025).

Case 1: Design Campaign Scale & Time Summary

Campaign scale (Quick test)

Backbone	RFdiffusion	100 backbones
Sequence	ProteinMPNN	8 sequences / backbone
Co-folding	복합체 구조 샘플링	5 samples / sequence

Total 4,000 co-folding inferences

Design speed (on 1 GPU)

Backbone	2 min / backbone
Sequence	~0.1 sec / sequence
K-Fold	30 sec / 5 samples

Zero MSA wait time

MSA computation time ~ 10 min per complex on avg.

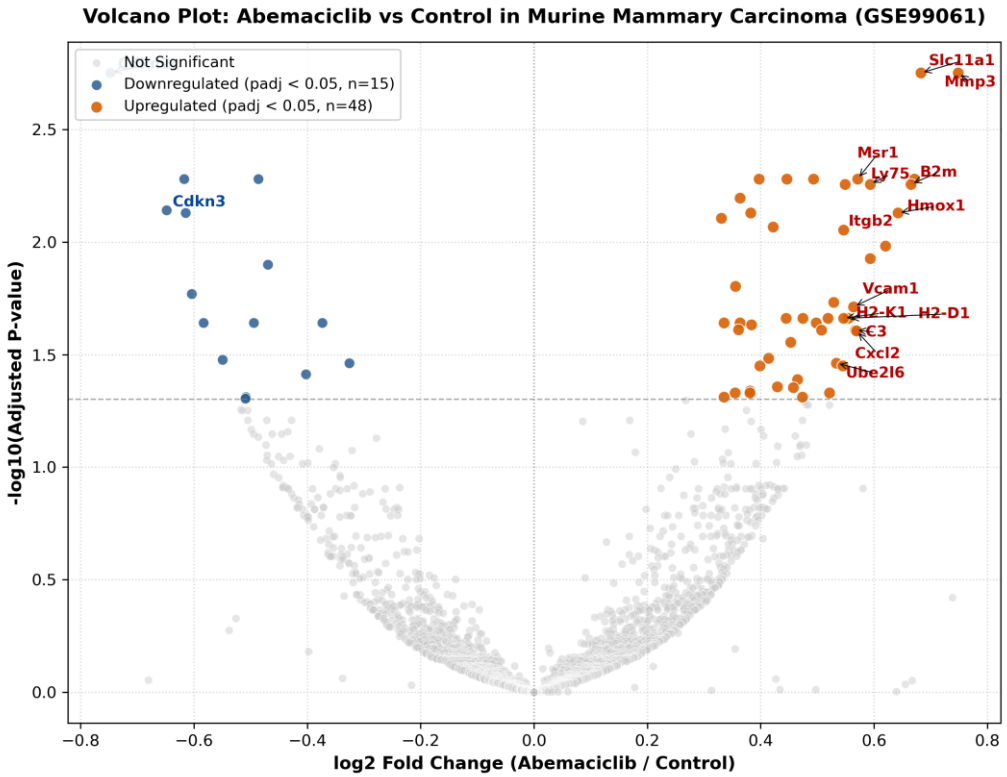
Total campaign wall time ~ 2 hours @ 5 GPUs.

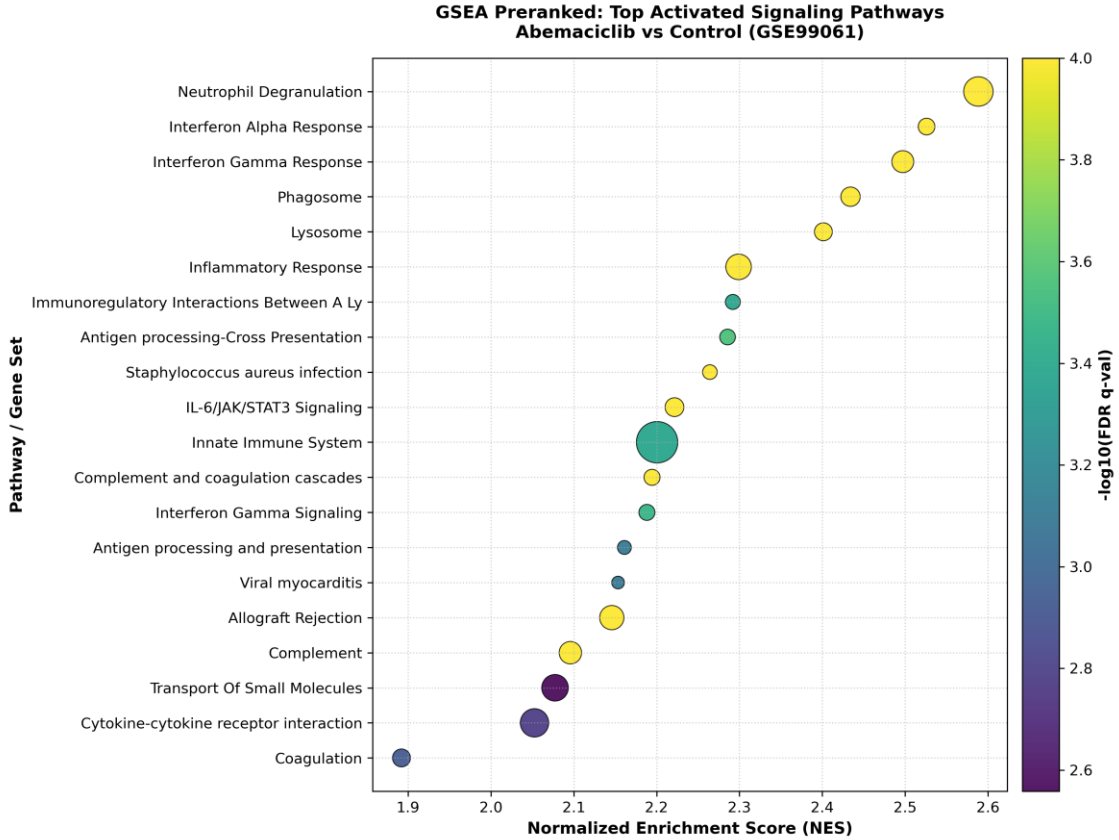
(Total wall time with MSA ~ 75 hours.)

Case 2: RNA-seq Analysis in HyperLab

- 3,826개 유전자에 대한 RNA-Seq raw counts데이터로부터 DEG 분석을 요청
- 분석 파이프라인 설계 → 분석 진행(PCA, DEG, ORA, GSEA) → 결론 도출

GSE99061 데이터셋으로 DEG 분석을 진행해서, in vivo 마우스 유방암 모델에 아베마시클립(abemaciclib)을 처리했을 때 발현이 증가하는 신호 전달 경로들을 찾아주세요.

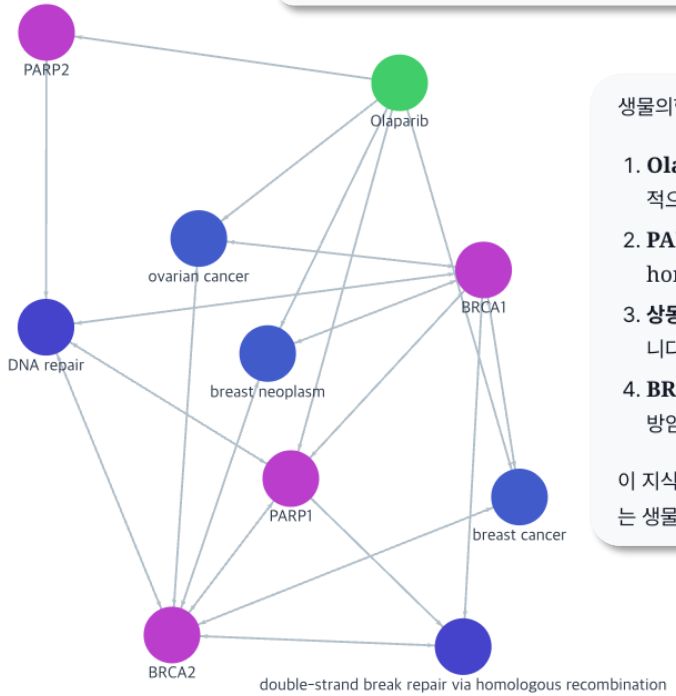




Case 3: Literature Search in HyperLab

- 문헌 조사가 필요한 생물학적 내용을 질문
- 생물 의학 지식 그래프에 기반한 답변 제공

Olaparib이 일부 난소암 및 유방암 환자에게서만 뚜렷한 항종양 효과를 보이는 분자 수준의 선택적 작용 기전에 대한 내용들을 관련 문헌에서 찾아 근거와 함께 정리해 주세요.



생물의학 지식그래프 분석 결과, Olaparib과 질환 간의 분자적 경로는 다음과 같이 검증된 상호작용으로 연결됩니다:

1. **Olaparib → PARP1 / PARP2 저해**: Olaparib(`CHEBI:83766`)은 PARP1(`NCBIGene:142`) 및 PARP2(`NCBIGene:10038`)의 촉매 활성을 억제하고 물리적으로 결합합니다 (DrugBank, BindingDB, Pharos, OpenTargets).
2. **PARP1/2 → DNA 복구 / 상동재조합 경로 관여**: PARP1/2는 유전자 온톨로지(GO)의 DNA repair (`GO:0006281`) 및 double-strand break repair via homologous recombination (`GO:0000724`) 생물학적 프로세스에 직접 관여합니다.
3. **상동재조합 → BRCA1 / BRCA2**: BRCA1(`NCBIGene:672`)과 BRCA2(`NCBIGene:675`)는 상동재조합 이중가닥 절단 복구(`GO:0000724`)의 핵심 실행 인자입니다.
4. **BRCA1/2 → 난소암 및 유방암 인과관계**: UniProt, DisGeNET, ClinVar에 큐레이션된 바와 같이, BRCA1/2의 병원성 변이는 난소암(`MONDO:0008170`)과 유방암(`MONDO:0007254` / `MONDO:0021100`) 발생의 직접적인 유전적 원인으로 작용합니다.

이 지식그래프 경로는 단일 질환 타깃 억제가 아닌, 두 독립적인 DNA 수리 경로(PARP 매개 경로와 BRCA 매개 상동재조합 경로)의 교차 지점에서 치료 효과가 발생하는 생물학적 네트워크 구조를 명확히 보여줍니다.



Conclusion

오믹스 데이터에서 신약 후보까지, AI로 효율화 하는 Co-Scientist



Acknowledgements

KAIST Team

[김우연 교수 연구실]
김형우, 서성환, 문석현, 이중원,
김준형, 박민하, 김신우

[황성주 교수 연구실]
김준형, 김동기, 백진헌, 조재형

[안성수 교수 연구실]
장효순, 김태원, 서현진

[오병하 교수 연구실]
배영남, 한현규, 최유진, 최정훈

[김호민 교수 연구실]
전우성, 이준엽

[이규리 교수 연구실]
신윤수, Calvin

HITS

[AI연구본부]
배성한, 임재창, 정원호, 황상연

[플랫폼사업본부]
김규범, 김명섭, 김아람, 박하진,
정승연, 조성상, 최범권, 최윤서

AI-CRED Institute

과학기술정보통신부 Ministry of Science and ICT

nipa

AI 특화 파운데이션 모델 프로젝트

HyperLab

HyperLab 가입 후 베타 체험 크레딧 제공

- https://wooyoun.kaist.ac.kr/
- https://hits.ai/
- https://hyperlab.ai/
- https://omicshorizon.ai/



2026년
대한부인종양학회
제31차 추계 심포지엄

2026.10. 23. (금) - 24. (토)

정부세종컨벤션센터 / 세종 코트야드 바이 메리어트

대한부인종양학회
Korean Society of Gynecologic Oncology