Agent-Guided De Novo Design of Nanobody Binders Against a Novel Cancer Target

Yue Zhao[1,†], Melih Yilmaz[1,†], Edward Lee[1,†], Chuanyui Teh[1,†], Lan Guo[1], Kemal Sonmez[1], Luca Giancardo[1], Gordon Trang[1], Fangda Xu[1], Madelyn Espinosa-Cotton[2], Nai-Kong V Cheung[2,*], Jiwon Kim[1,*], Xinyun (Nina) Cheng[1,*]

[1]Amazon Web Services, Applied AI Solutions, Life Sciences | [2]Memorial Sloan Kettering Cancer Center

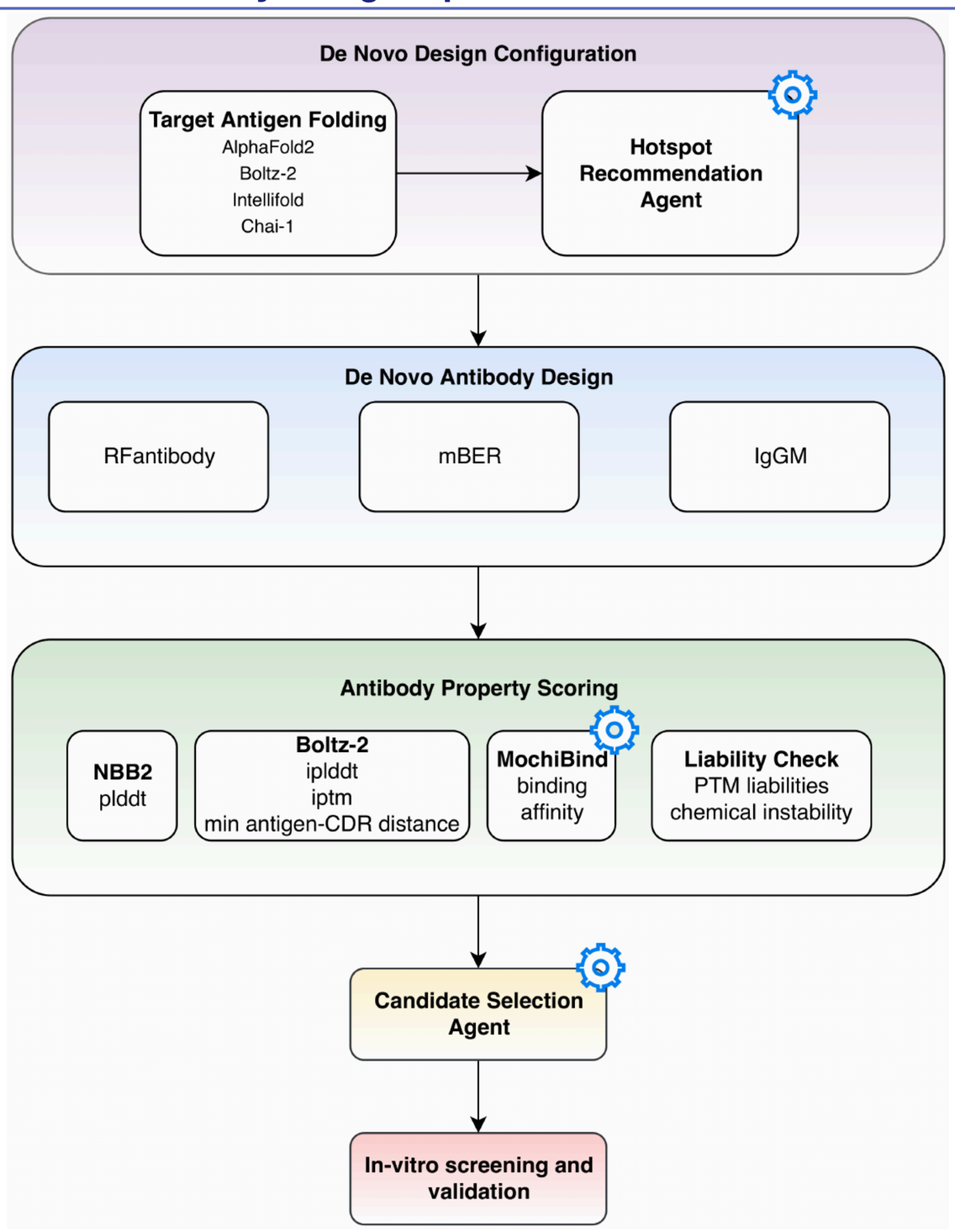
[†]Equal contribution *Corresponding authors

bioRxiv preprint

Motivation & Approach

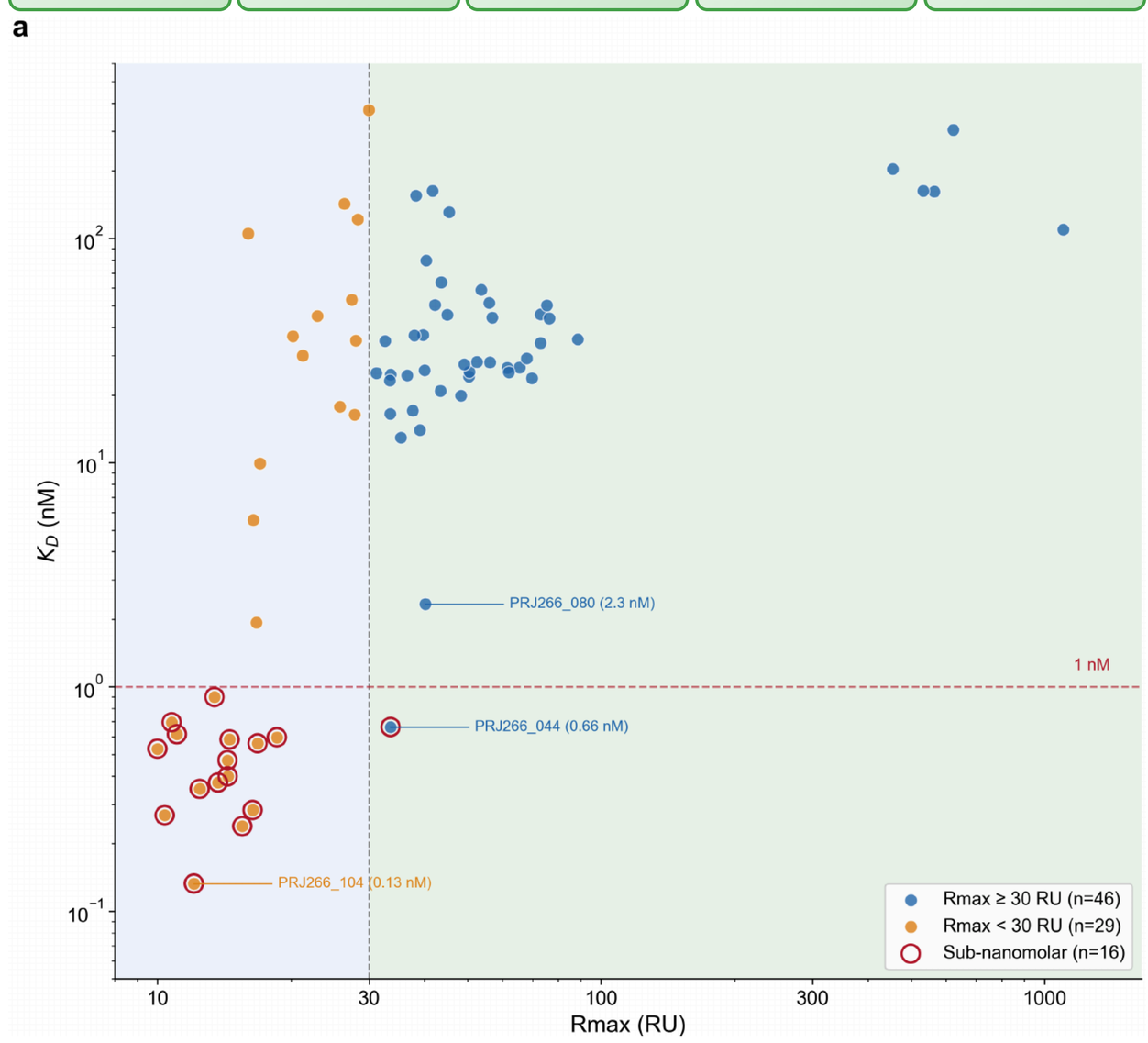
- Traditional antibody discovery: **6–12 months**, limited epitope control
- Novel DSRCT cancer target (identified at MSK) — **no structure, no prior antibodies**
- We test whether **generalized priors from PDB complexes** can design functional binders against an unseen surface
- Fully computational, agent-guided workflow with **evidence-grounded epitope control**

De novo Nanobody Design Pipeline



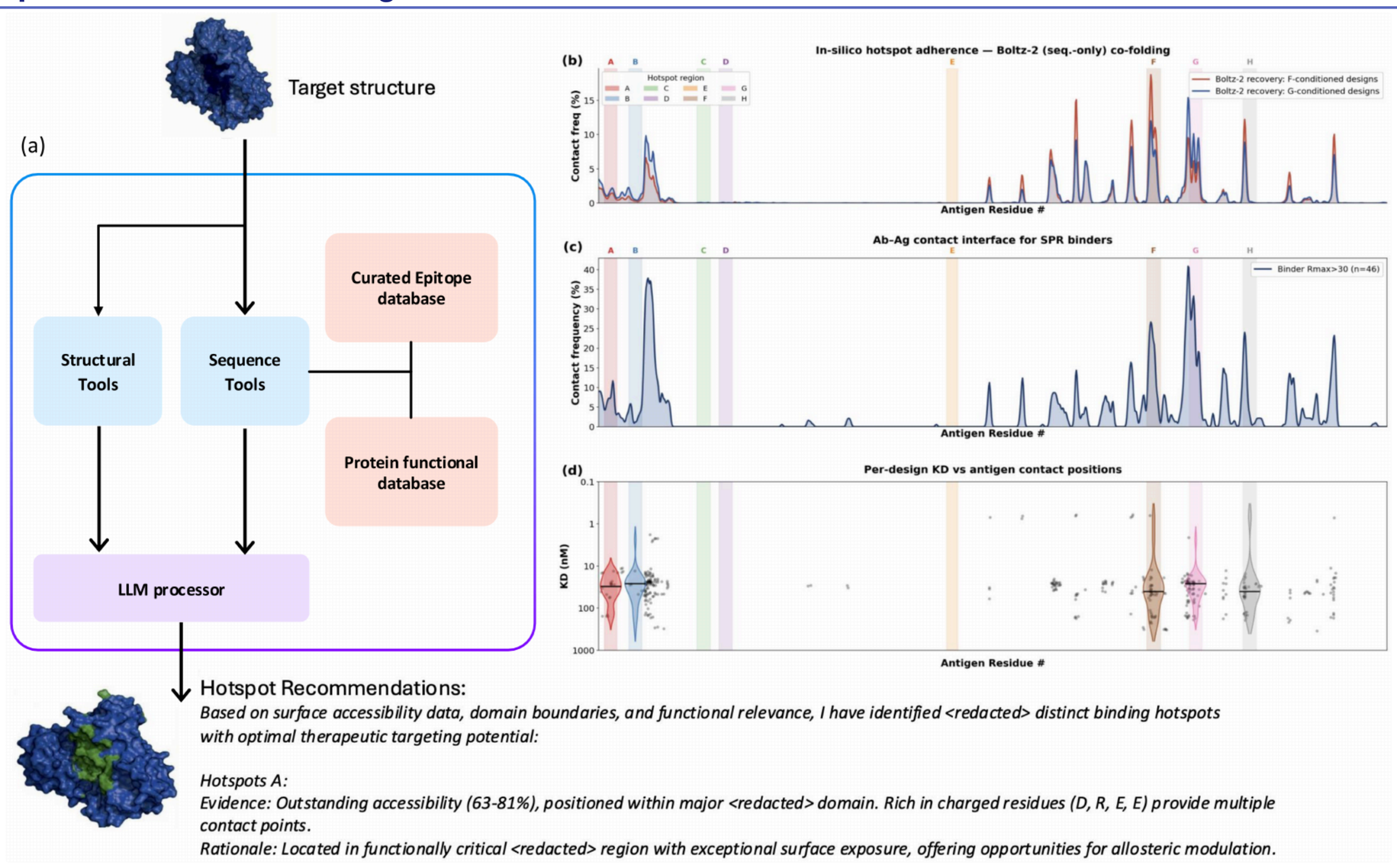
Key Experimental Results

288K	100K	46/116	0.66 nM	31.7 nM
Designed	Screened	SPR (39.7%)	Best K_D	Median K_D



SPR (Surface Plasmon Resonance): measures antibody–antigen binding strength in real time.
K_D (Dissociation Constant): quantifies binding affinity (lower = tighter binding).
R_{max}: maximum signal amplitude (higher = more active binders and more reliable kinetic fits).

Hotspot Recommendation Agent



- LLM-orchestrated agent integrates **7 bioinformatics tools and 2 databases**
- Proposes **8 distinct hotspot regions** with physical & functional rationale
- **~80% top-5 accuracy** on curated holdout (n=76)
- Grounded in verifiable tool outputs — no hallucinated epitopes

Takeaways & Next Steps

- Fully computational, agent-guided workflow yields **nanomolar–sub-nM binders** against a novel target with no structure or prior antibody data.
- Grounding the LLM in **deterministic tool outputs** constrains epitopes to evidence.
- Bottleneck shifts from experimental throughput to **model & scoring quality**; in-silico folding metrics correlate weakly with binding.
- **Next:** DBTL cycle 1 complete (116 SPR-characterized) → developability profiling → target-specific ML for directed evolution.

Contact: cheungn@mskcc.org | kimaaj@amazon.com | cxinyun@amazon.com