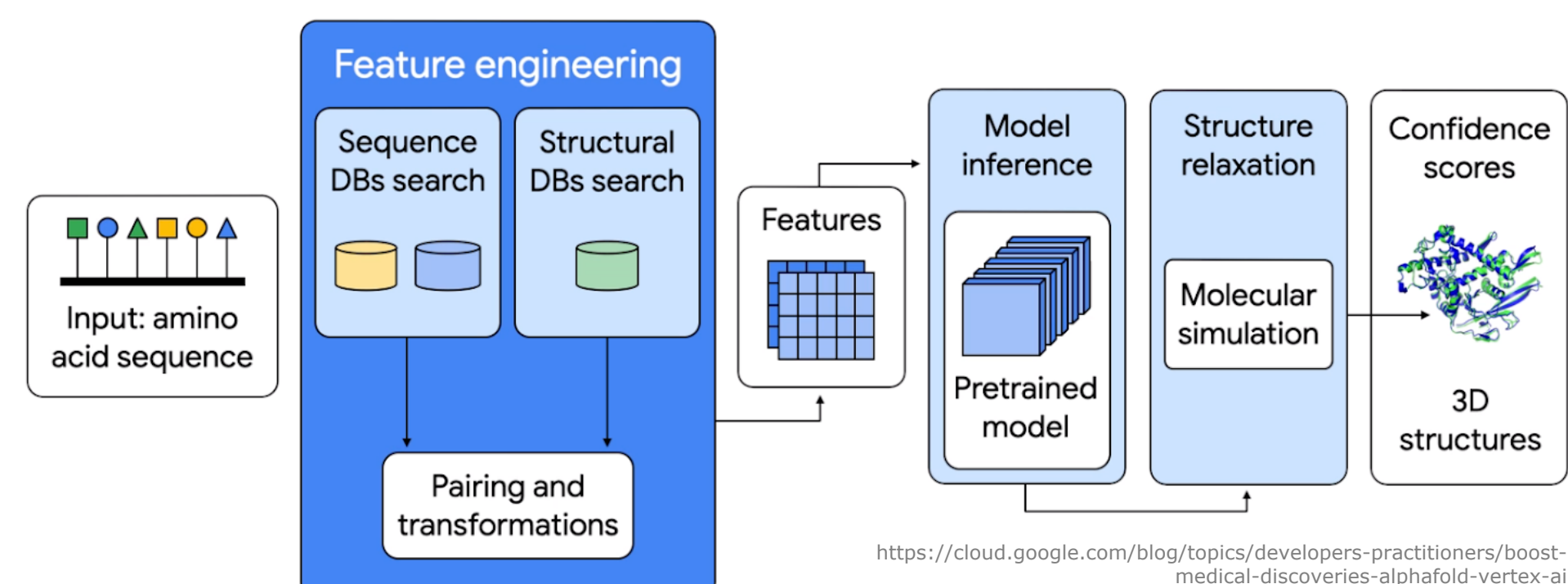


Protein Structure Prediction

AlphaFold-style models are revolutionizing drug-discovery!



Where is the Current Data Sourced

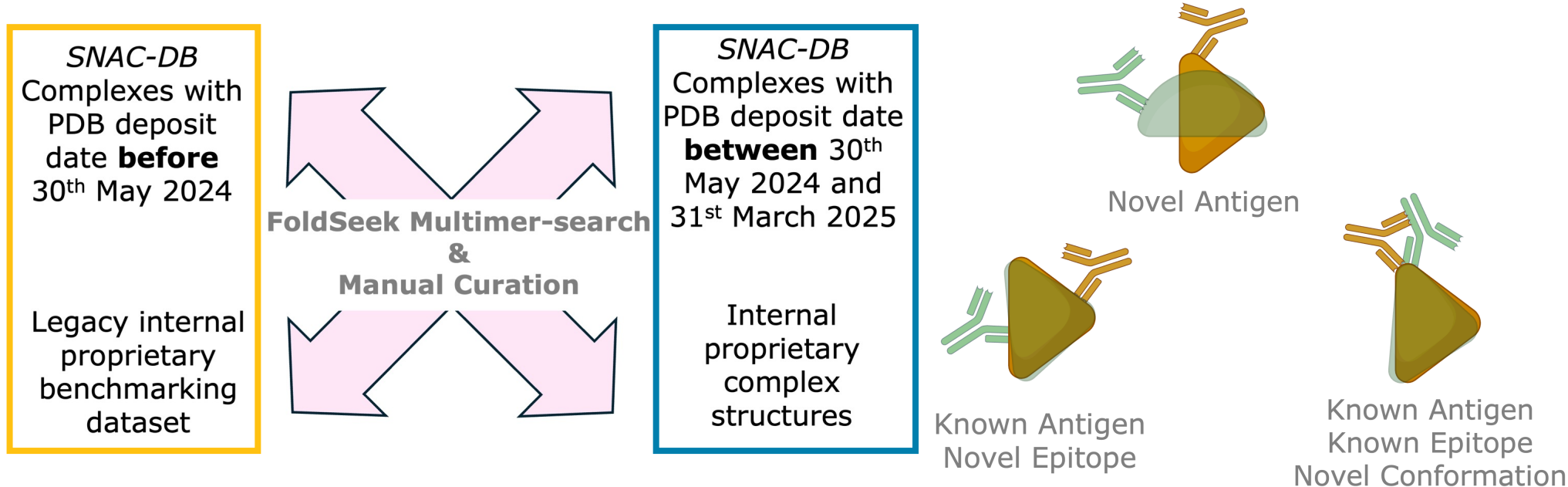
RCSB PDB Protein Data Bank¹

- PDB – the world's central archive of experimentally determined 3D macromolecular structures (X-ray crystallography, NMR, cryo-EM).
- Pros:** Comprehensive coverage.
- Cons:** Heterogeneous naming, missing residues, crystal-packing artifacts.

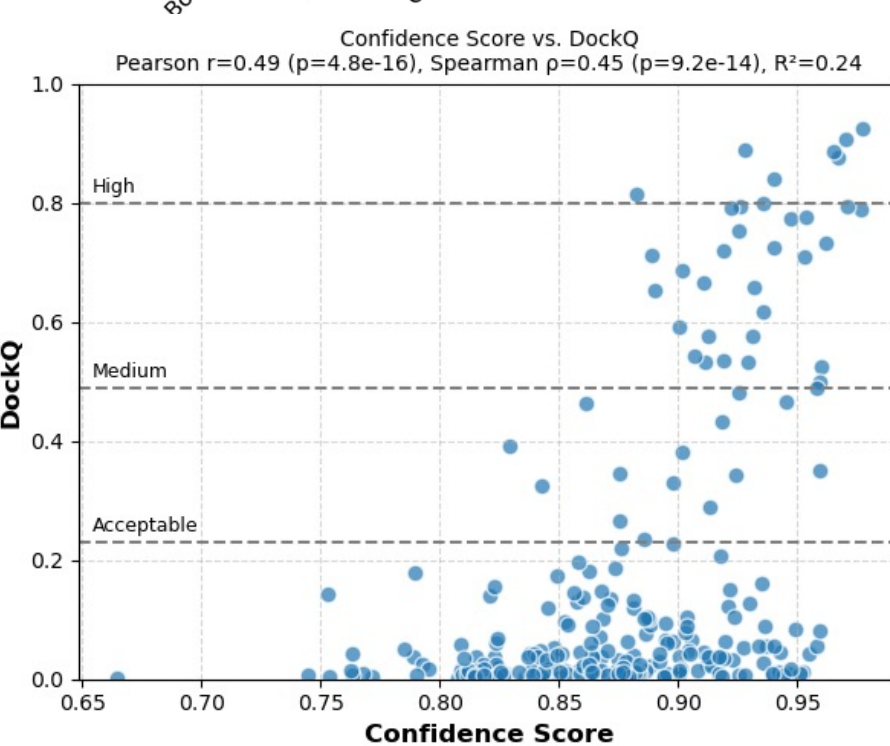
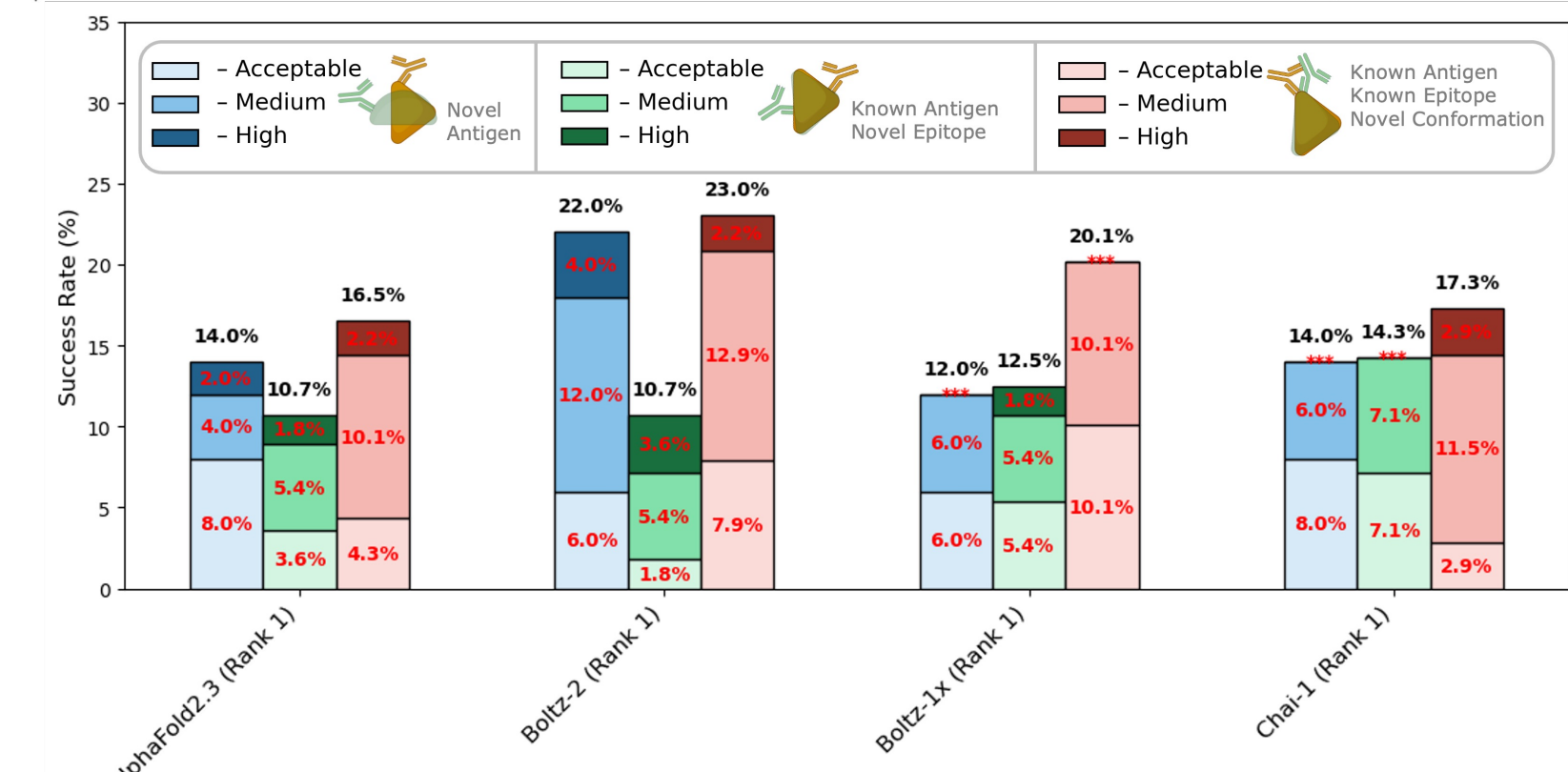
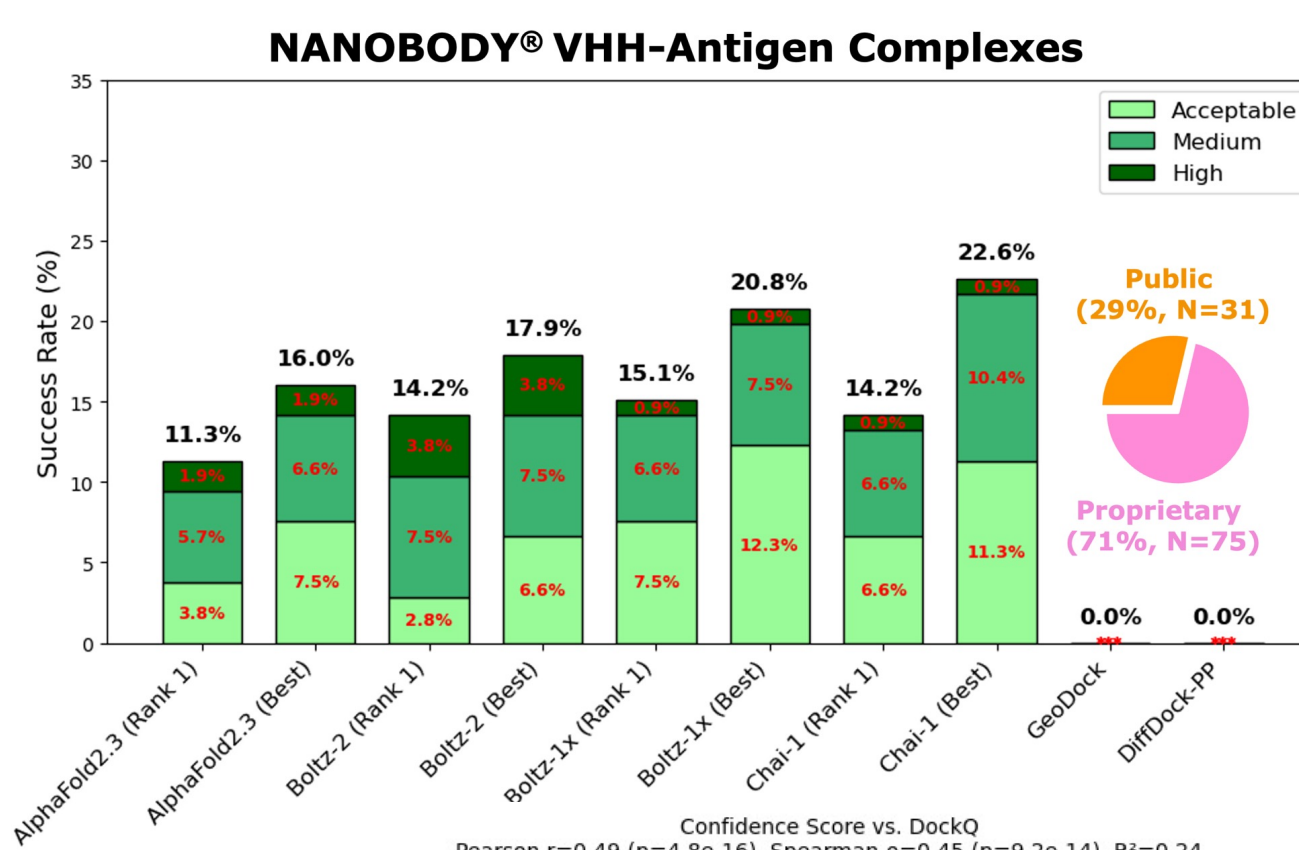
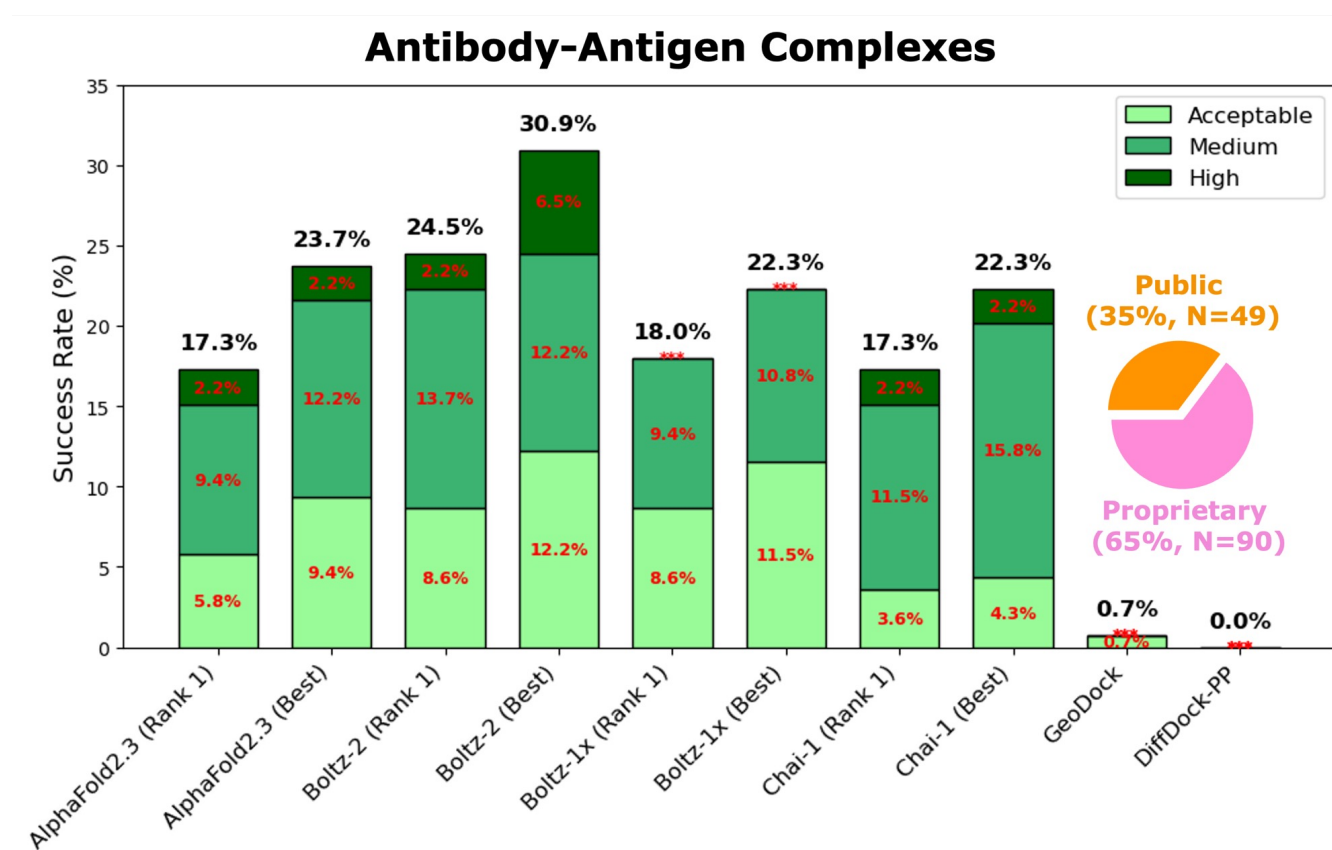
SAbDab Structural Antibody Database²

- SAbDab – a curated PDB subset of Ab/Nb-Ag complexes from asymmetric units.
- Pros:** VH/VL/VHH annotation, standardized variable domains.
- Cons:** Issues with the accuracy of multi-chain antigens, excludes TCR and Ig-Ig interactions.

Towards Rigorous Benchmarking



- Genuine Novelty:** Filters out any overlap with prior training sets using TM-score to ensure evaluation on truly unseen targets, epitopes, and conformations.
- Industry-relevant benchmark:** Mix of public and proprietary structures.



References and Data Availability:

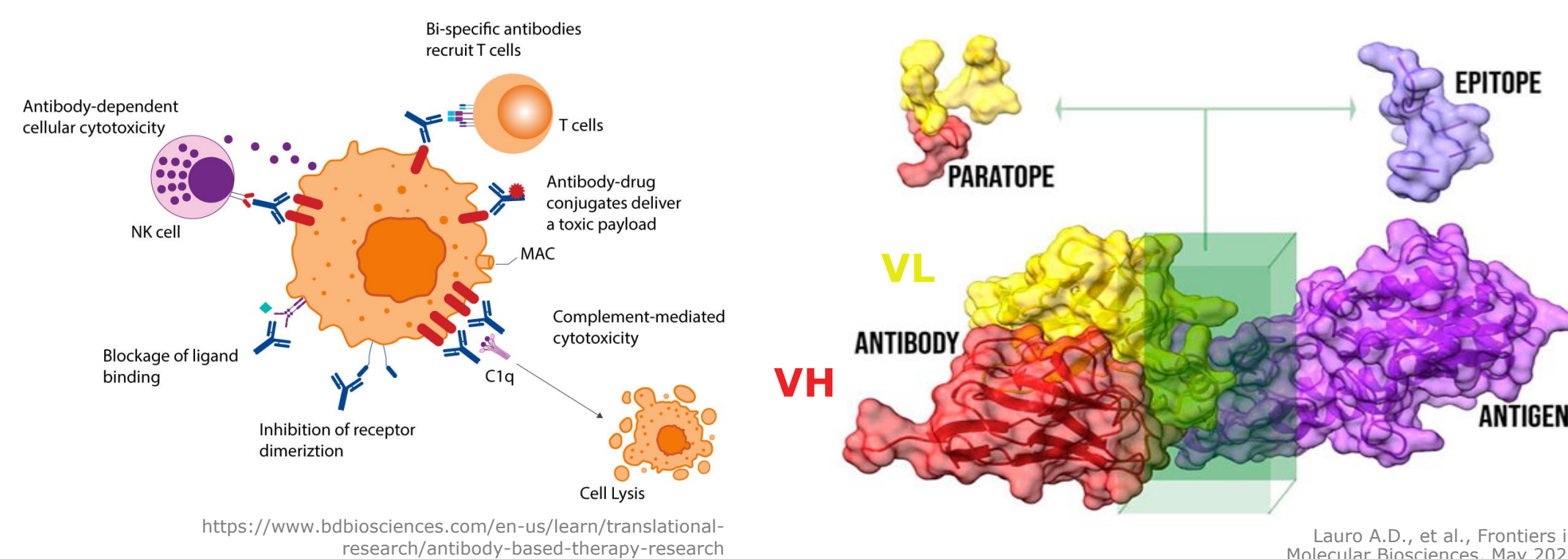
¹Berman, H.M., et al (2000). Nucleic Acids Res. 42. 28:235-242
<https://github.com/Sanofi-Public/SNAC-DB>



²Dunbar, J., Krawczyk, K. et al (2014). Nucleic Acids Res. 42. D1140-D1146
<https://zenodo.org/uploads/15870003>

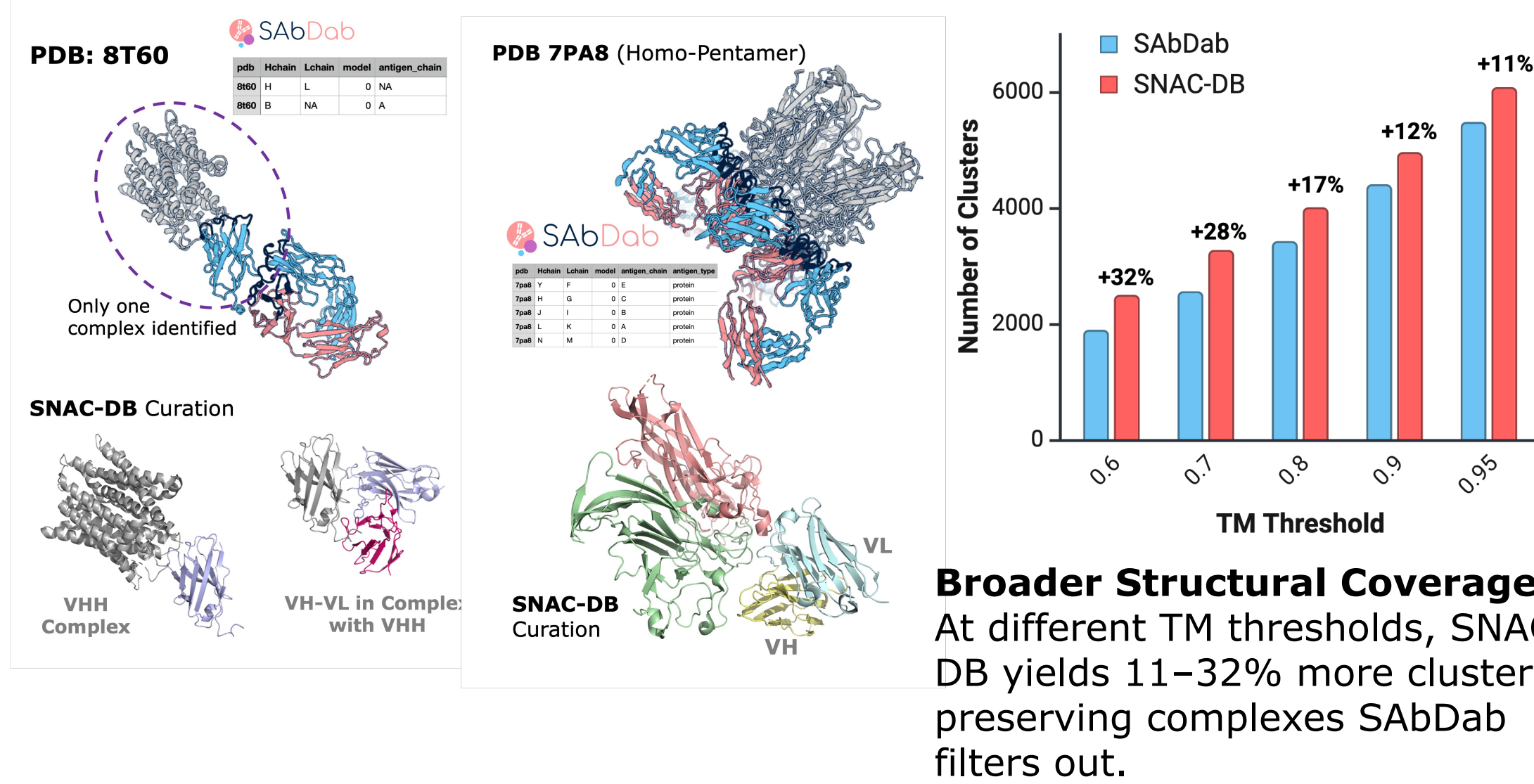
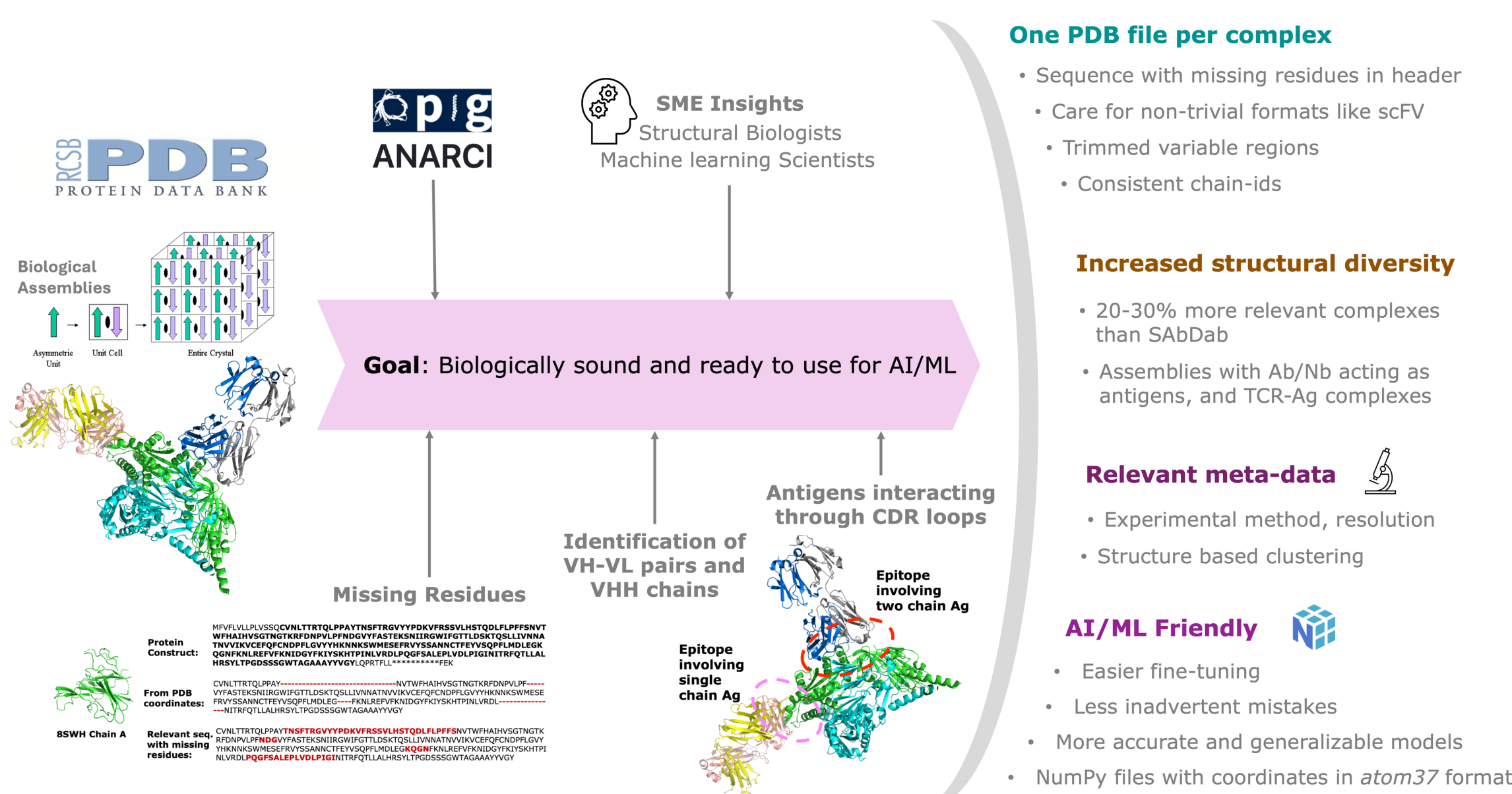


Importance of Antibody and NANOBODY® VHH



- Therapeutic antibodies and nanobodies are the **fastest growing** drug class.
- Highly variable** and flexible nature of CDR loops make them a **blind spot** for AI/ML.
- Poor predictions risk wasted time and money** in early-stage discovery.

Pipeline Overview



Conclusions & Future Directions

- Data Matters:** Boltz-2's extra training on newer PDB entries yields noticeably better Ab-Ag and Nb-Ag predictions than purely architectural improvements.
- Poor generalization:** All models struggle on novel epitopes, revealing a tendency to "remember" familiar binding sites instead of truly extrapolating.
- Unreliable confidence metrics:** Even when the correct pose is sampled, internal scoring rarely ranks it first, limiting trust in top-ranked predictions. Boltz-2's internal confidence scores correlate weakly with actual DockQ quality (Spearman $\rho \approx 0.45$).
- SNAC-DB closes critical gaps.** Immediate lift to structural coverage, better multi-chain antigens, missed complexes: TCR-Ag, Ab-Nb, Nb-Ab, weak-cognate interactions.
- Next steps:** fine-tune models on SNAC-DB to evaluate how the performance is impacted.