ORIGIN-1: EXPERIMENTALLY VALIDATED GENERATIVE AI PLATFORM FOR *DE NOVO* ANTIBODY DESIGN AGAINST "ZERO-PRIOR" EPITOPES

Simon Levine*, Jonathan Edward King*, Jacob Stern*, David Grayson*, Raymond Wang*, Rui Yin*, Umberto Lupo*, Paulina Kulytė*, Ryan Matthew Brand*, Tristan Bertin*, Robert Pfingsten*, Jovan Cejovic*, Chelsea Chung, Breanna K. Luton, Andrew Hagemann, Robel Haile, Elliot Medina, Pankaj Panwar, Oleksii Dubrovskyi, Chase LaCombe, Zahra Anderson, Derrik Mildh, Scott Benjamin, Joe Kaiser, Joseph Ferron, Marta Sarrico, Alexandria Kershner, Apurva Mishra, Kai R. Ejan, Emily K. Marsh, Paul Bringas, Phetsamay Vilaychack, Kyra Chapman, Jacob Ripley, Muttappa Gowda, Kathryn M. Collins, Cailen M. McCloskey, Jeremiah S. Joseph, Rylee Ripley, Shaheed A. Abdulhaqq, Audree Feltner, Michael Guerin, Jeffrey Goby, Jesse Hendricks, Danielle Castillo, Sean McClain, Douglas Ganini, Derek Shpiel, James Mategko, Eder Cruz Garcia, Masoud Zabet-Moghaddam, John M. Sutton, Zheyuan Guo, Sean M. West, Janani S. Iyer†, Amir Shanehsazzadeh†

*Equal contribution

†Correspondence to ashanehsazzadeh@absci.com, jiyer@absci.com

ABSTRACT

Generative artificial intelligence has advanced antibody discovery, yet *de novo* design against epitopes lacking structural precedent remains a fundamental challenge. Here we present Origin-1, a generative AI platform that integrates epitope-conditioned all-atom structure generation (AbsciGen), paired CDR sequence design (IgDesign2), and a co-folding-based scoring protocol (AbsciBind) to select high-confidence binders. We test Origin-1's ability to design antibodies against "zero-prior" epitopes, or target sites with no available antibody-antigen or protein-protein complex structures. Across ten human protein targets, Origin-1 identified antibodies for four targets in fewer than 100 design attempts each. Cryogenic electron microscopy confirmed atomic accuracy (3.0–3.1 Å resolution; DockQ 0.73–0.83) and AI-guided affinity maturation produced an IL36RA antagonist with 104 nM potency. These results establish a framework for antibody design without structural precedent.

1 INTRODUCTION

Antibodies are essential therapeutics due to their ability to bind diverse antigens with high affinity and specificity (Klein et al., 2024; Pantaleo et al., 2022). However, traditional discovery methods remain resource-intensive and provide limited control over epitope targeting, motivating computational approaches for *de novo* design with precise epitope specification.

Antibody–antigen recognition is mediated by six complementarity-determining regions (CDRs), whose sequence and conformational variability make CDR design the core computational challenge (James et al., 2003). Recent advances have enabled progress in protein structure prediction (Abramson et al., 2024; Jumper et al., 2021) and *de novo* design (Cao et al., 2022; Watson et al., 2023), yet existing approaches exhibit significant limitations and varying levels of experimental validation. Structure generation methods lack atomic resolution (Watson et al., 2023; Bennett et al., 2026), are not antibody-specialized (Qu et al., 2024; Geffner et al., 2025), or incur substantial computational expense (Pacesa et al., 2025; Cho et al., 2025). Sequence design models treat heavy and light chains independently (Dauparas et al., 2022; Dreyer et al., 2023) or lack deep integration of paired language representations with geometric encoding (Kenlay et al., 2024).

Beyond design, stringent filtering is critical to advance high-confidence candidates to experimentation (Zambaldi et al., 2024). Although folding model confidence metrics can evaluate binding like-

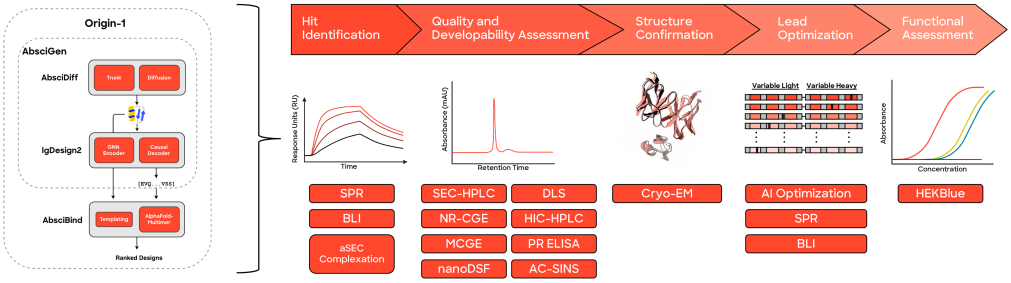


Figure 1: **Origin-1 platform and experimental validation cascade.** *Left:* AbsciGen generates antibody designs via AbsciDiff (epitope-conditioned all-atom diffusion) and IgDesign2 (paired CDR sequence design). AbsciBind evaluates designs using a modified co-folding protocol to select high-confidence candidates. *Right:* Experimental cascade: designs are expressed as mAbs and validated through Surface Plasmon Resonance (SPR) screening, Biolayer Interferometry (BLI) confirmation, analytical Size Exclusion Chromatography (SEC) complexation, developability assessment, Cryogenic Electron Microscopy (cryo-EM) structure determination, and functional assays.

lihood (Evans et al., 2021b; Roney & Ovchinnikov, 2022), they perform poorly for antibody–antigen complexes (Yin & Pierce, 2024; Hitawala & Gray, 2025).

We introduce Origin-1, a design-and-score platform that addresses these limitations. AbsciGen enables epitope-conditioned all-atom structure generation via AbsciDiff (diffusion-based generation) and IgDesign2 (paired sequence design). AbsciBind scores candidates for experimental prioritization. We demonstrate that Origin-1 generates antibodies against "zero-prior" epitopes—sites lacking any reported complex structure involving the target—from four human protein targets.

2 Methods

Origin-1 performs two tasks: (1) generate antibody structures and sequences likely to bind to targets of interest, and (2) score candidates to select high-confidence designs (Figure 1). AbsciGen comprises AbsciDiff (structure generation) and IgDesign2 (sequence design). AbsciBind provides scoring and filtering.

2.1 Structure Generation

AbsciDiff is a diffusion-based all-atom model fine-tuned from Boltz-1 (Wohlwend et al., 2024) for epitope-conditioned antibody design. We model the forward process as $dX_t = \sqrt{2t}\, dB_t$ and train the model to approximate $\mathbb{E}[X_0|X_t]$ given conditioning information: antigen structure/sequence, framework structure/sequence, and epitope residues.

Key modifications include: (1) masking CDR positions and inter-chain features while providing epitope conditioning via a binary residue vector; (2) an intermediate sequence hypothesis module that predicts design region sequences and recycles them through the trunk; (3) template featurization from user-provided structures; and (4) Atom14 representation for design regions with virtual atoms at Cα positions, enabling open-ended sidechain generation. For more details on structure generation, refer to Appendix B.1.

2.2 Sequence Design

IgDesign2 models CDR sequences autoregressively: $p_\theta(R|X, C) = \prod_{i=1}^{N} p_\theta(r_i|r_{<i}, X, C)$. The architecture combines a Graph Neural Network (GNN) encoder (PiFold-based (Gao et al., 2023)), a causal transformer decoder, and a paired antibody language model (IgBert (Kenlay et al., 2024)) for refinement. At each pLM layer, structure embeddings are fused with language embeddings via trainable projections. The model is pre-trained on protein–protein interactions, then fine-tuned on antibody–antigen complexes. For more details on sequence design, refer to Appendix B.2.

2.3 Scoring

Standard AlphaFold-Multimer (AFM) (Evans et al., 2021a) approaches for binder discrimination yield high false-negative rates for antibody–antigen complexes due to template masking. AbsciBind addresses this by: (1) providing designed structures as multimer templates with retained amino acid tokens; (2) masking sidechains except Cβ; and (3) disabling inter-chain template distance masking.

We compute an AbsciBind Score as the mean of standard Interface Predicted Template Modeling (ipTM) and Antibody-Aligned ipTM (which merges antibody/antigen chains to focus on the interface) scores. Designs with ligand Root-Mean-Square Deviation (RMSD) >5 Å between designed and predicted structures are filtered. For more details on AbsciBind score, refer to Appendix B.3.

3 Results

3.1 Experimental Setup

Targets. We selected ten targets and identified zero-prior epitopes per target. 9/10 targets did not have PDB complex structures. One target (AZGP1) had one PDB entry in complex with a non-antibody protein binder (PDB:3ES6), but the associated epitope was not used for design. All targets also had $\leq$60% sequence similarity to proteins with known complexes. IL36RA was selected as a target of particular therapeutic interest. Given IL36RA's role in dampening proinflammatory responses that may promote anti-tumor immunity (Finucane et al., 2025), antibody-mediated inhibition of IL36RA could enhance immune infiltration into "immune-cold" tumors.

Design. For each target, we sampled 3,360 design specifications varying framework, CDR lengths, and epitope subsampling. Through iterative search, we selected 95 designs per target.

Validation. Designs were expressed as mAbs and screened by SPR against targets and off-targets. Hits were validated by BLI and analytical SEC. Developability was assessed via thermal stability, polyreactivity, and hydrophobicity assays. Select complexes were characterized by cryo-EM. IL36RA binders were tested in HEKBlue functional assays.

3.2 In Silico Results

We benchmarked AbsciGen against RFantibody on a subset of the previously-selected targets (COL6A3, AZGP1, CHI3L2, IL36RA). AbsciGen substantially outperformed RFantibody (Figure 2), with 28.4% of designs achieving an AbsciBind Score $\geq$0.5 compared to just 1.5%—a 19-fold improvement.

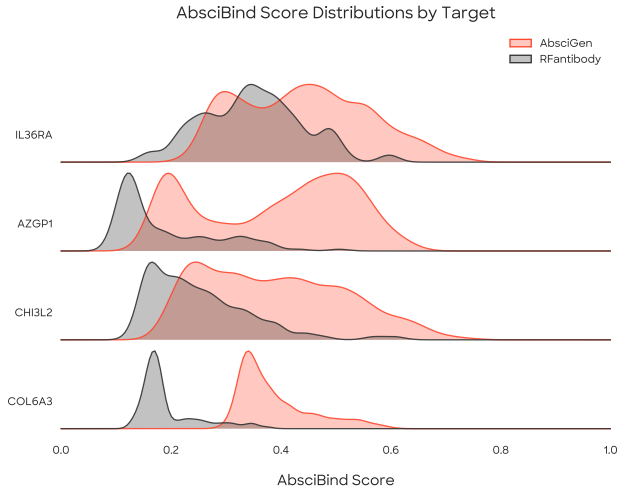
We also evaluated AbsciBind's ability to discriminate binders from non-binders across eight antibody-antigen systems. AbsciBind achieved better classification performance (AUROC) compared to six reference methods including GeoFlow-V2, GeoFlow-V3, AFM-IG, Protenix, Boltz-2, and Chai-1 (Figure 3).

For additional *in silico* results, including supervised benchmarking on targets with known structures (Appendix Figures 11–12), humanness assessment (Appendix Figure 14), and design diversity analysis (Appendix Figures 16–17), refer to Appendix F.

3.3 In Vitro Results

Experimental Validation. SPR identified nine hits across four targets from $\sim$400 total designs: three against COL6A3, four against AZGP1, one against CHI3L2, and one against IL36RA. All hits bound their designed antigen without binding off-targets (Figure 4C; Appendix Figures 18–21). BLI confirmed COL6A3 and AZGP1 hits; CHI3L2 and IL36RA hits showed weak affinity ($K_D > 2\,\mu$M). Analytical SEC confirmed complexation for all four targets (Appendix Figure 23).

AI-Guided Affinity Maturation. We applied lead optimization using single-mutant libraries scored by AbsciBind and protein language models. This yielded improvements across all targets: 68-fold for IL36RA (from $\sim$7 μM to 104 nM; Figure 4C,D) and $\sim$4-fold for CHI3L2 (Appendix Figure 22).



(A)

Method	Mean Absci Bind Score	% Designs $\geq$ 0.5
RFantibody	0.270 ± 0.112	1.5%
AbsciGen	$\mathbf{0.422 \pm 0.116}$*	**28.4%**

*$p < 0.001$, Mann-Whitney U test.

(B)

Figure 2: **AbsciGen outperforms RFantibody on four targets.** (A) AbsciBind Score distributions by target. (B) Summary statistics.

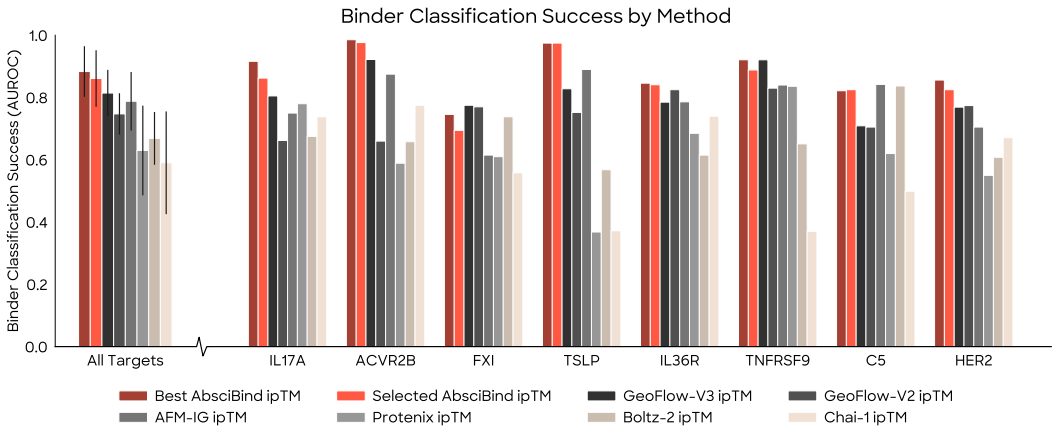
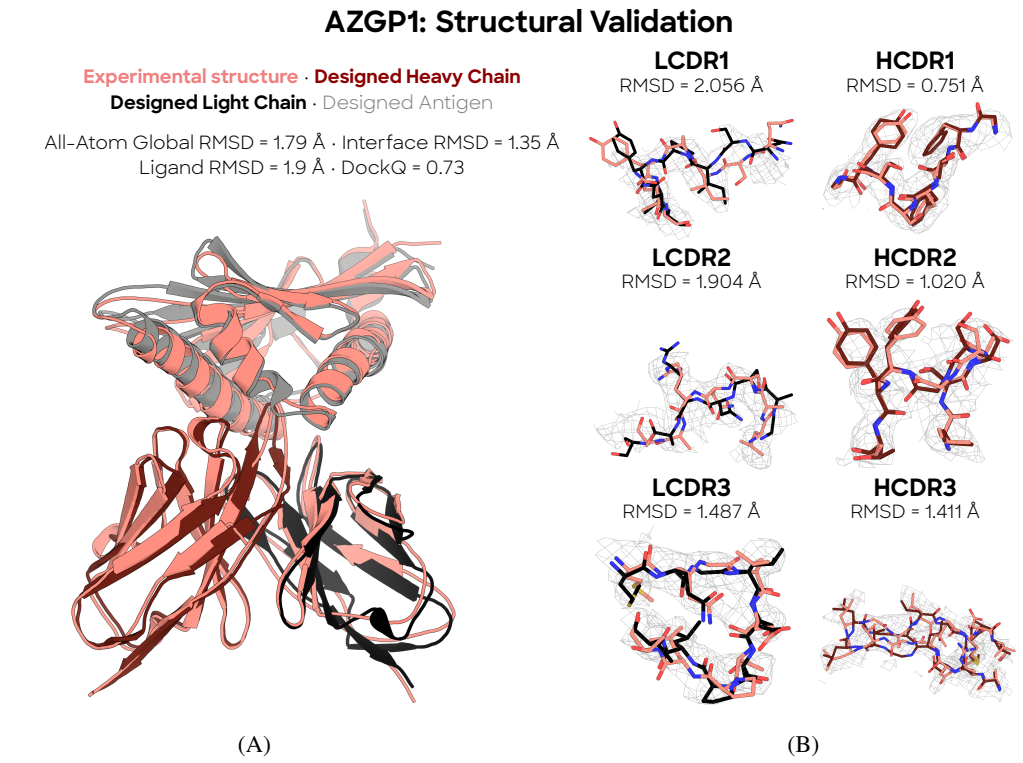


Figure 3: **AbsciBind discriminates binders from non-binders.** Classification performance (AUROC) across eight antibody-antigen systems compared to six reference methods.

Structural Validation. Cryo-EM structures of COL6A3 and AZGP1 complexes (3.0 Å and 3.1 Å resolution) showed high fidelity with designs. COL6A3: interface RMSD 0.96 Å, DockQ 0.83; CDR RMSDs 0.66–1.49 Å (Appendix Figure 18). AZGP1: interface RMSD 1.35 Å, DockQ 0.73; CDR RMSDs 0.75–2.06 Å (Figure 4A,B; Appendix Figure 19). Processing workflows are provided in Appendix Figures 24 and 25.

Functional Validation. Affinity-matured IL36RA variants were tested in HEKBlue assays measuring IL36 receptor signaling. While the parent showed no activity, optimized variants displayed antagonism correlating with affinity, with the best achieving EC_{50} = 104 nM (Figure 4E; Appendix Figure 21).

Developability. All binders met therapeutic criteria for polyreactivity, polydispersity, and thermal stability ($T_m > 68$°C), with minor flags for hydrophobicity (IL36RA) and self-association (AZGP1).



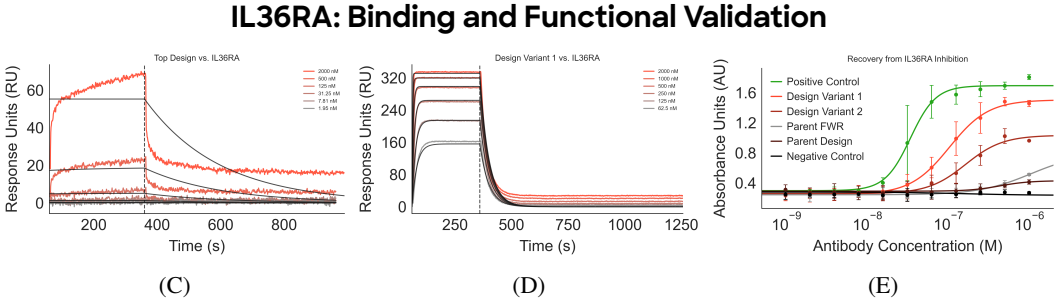


Figure 4: **Structural, binding, and functional validation.** Representative data from four validated targets (COL6A3, AZGP1, CHI3L2, IL36RA). *Top (AZGP1):* (A) Cryo-EM structure at 3.1 Å resolution showing overlay of designed and experimental structures. (B) Per-CDR analysis confirms all six CDRs match design with RMSD <2.1 Å. *Bottom (IL36RA):* (C) SPR confirms parent design binds IL36RA ($K_D > 2$ μM). (D) AI-guided affinity maturation improves binding to 104 nM (68-fold improvement). (E) HEKBlue functional assay demonstrates antagonism with EC_{50} = 104 nM. Cryo-EM validation for COL6A3 and binding data for all targets are provided in the Appendix.

4 Discussion

Origin-1 achieves *de novo* antibody design against "zero-prior" epitopes. In fewer than 100 attempts per target, we identified validated binders for 4/10 targets, with cryo-EM confirming atomic accuracy and functional assays demonstrating therapeutic potential.

Our hit rates are ostensibly lower than reports designing against known interfaces (Pacesa et al., 2025; Chai Discovery Team et al., 2025). In interpreting this, one should consider both the design challenge undertaken and the extent of experimental validation conducted to support a reported hit rate. Origin-1 must identify viable epitopes without guidance from solved structures, and we require binding confirmation across multiple orthogonal assays.

The integration of generative design with specialized scoring proved essential—AbsciBind's discrimination enabled efficient prioritization despite the zero-prior setting. The 68-fold affinity im-

provement for IL36RA demonstrates that weak initial binders can be rescued through AI-guided optimization.

Limitations. Success on 4/10 targets indicates room for improvement. Improving epitope prediction, sampling diversity, and scoring calibration are priorities for future work.

Impact. Origin-1 enables targeting epitopes without structural precedent, potentially unlocking difficult-to-drug targets. AbsciGen substantially outperforms RFantibody while generating more human-like sequences (Appendix Figure 14). Our methods are open-source and we release experimental data to support community benchmarking.

REFERENCES

Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016): 493–500, 2024. doi: 10.1038/s41586-024-07487-w.

Nathaniel R. Bennett, Joseph L. Watson, Robert J. Ragotte, Andrew J. Borst, DéJenaé L. See, Connor Weidle, et al. Atomically accurate de novo design of antibodies with RFdiffusion. *Nature*, 649: 183–193, 2026. doi: 10.1038/s41586-025-09721-5.

Longxing Cao, Brian Coventry, Inna Goreshnik, Buwei Huang, William Sheffler, Joon Sung Park, et al. Design of protein-binding proteins from the target structure alone. *Nature*, 605(7910): 551–560, 2022. doi: 10.1038/s41586-022-04654-9.

Chai Discovery Team, Jacques Boitreaud, Jeremy Dent, Daniel Geisz, Matt McPartlon, Joshua Meier, et al. Zero-shot antibody design in a 24-well plate. *bioRxiv*, 2025. doi: 10.1101/2025.07.05.663018.

Yehlin Cho, Martin Pacesa, Zuobai Zhang, Bruno E. Correia, and Sergey Ovchinnikov. BoltzDesign1: Inverting all-atom structure prediction model for generalized biomolecular binder design. *bioRxiv*, 2025. doi: 10.1101/2025.04.06.647261.

Justas Dauparas, Ivan Anishchenko, Nathaniel Bennett, Hao Bai, Robert J. Ragotte, Lukas F. Milles, et al. Robust deep learning-based protein sequence design using ProteinMPNN. *Science*, 378 (6615):49–56, 2022. doi: 10.1126/science.add2187.

Frédéric A. Dreyer, Daniel Cutting, Constantin Schneider, Henry Kenlay, and Charlotte M. Deane. Inverse folding for antibody sequence design using deep learning. *arXiv*, 2023. doi: 10.48550/arXiv.2310.19513. 2023 ICML Workshop on Computational Biology.

James Dunbar, Konrad Krawczyk, Jinwoo Leem, Terry Baker, Angelika Fuchs, Guy Georges, et al. SAbDab: The structural antibody database. *Nucleic Acids Research*, 42(D1):D1140–D1146, 2014. doi: 10.1093/nar/gkt1043.

Richard Evans, Michael O'Neill, Alexander Pritzel, Natasha Antropova, Andrew Senior, Tim Green, Augustin Žídek, Russ Bates, Sam Blackwell, Jason Yim, Olaf Ronneberger, Sebastian Bodenstein, Michal Zielinski, Alex Bridgland, Anna Potapenko, Andrew Cowie, Kathryn Tunyasuvunakool, Rishub Jain, Ellen Clancy, Pushmeet Kohli, John Jumper, and Demis Hassabis. Protein complex prediction with alphafold-multimer. *bioRxiv*, 2021a. doi: 10.1101/2021.10.04.463034. URL `https://www.biorxiv.org/content/early/2021/10/04/2021.10.04.463034`.

Richard Evans, Michael O'Neill, Alexander Pritzel, Natasha Antropova, Andrew Senior, Tim Green, et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021b. doi: 10.1101/2021.10.04.463034.

Méabh Finucane, Elizabeth Brint, and Aileen Houston. The complex roles of IL-36 and IL-38 in cancer: friends or foes? *Oncogene*, 44(13):851–861, 2025. doi: 10.1038/s41388-025-03293-4.

Zhangyang Gao, Cheng Tan, and Stan Z. Li. PiFold: Toward effective and efficient protein inverse folding. In *International Conference on Learning Representations*, 2023. URL `https://openreview.net/forum?id=oMsN9TYwJ0j`.

Tomas Geffner, Kieran Didi, Zhonglin Cao, Danny Reidenbach, Zuobai Zhang, Christian Dallago, Emine Kucukbenli, Karsten Kreis, and Arash Vahdat. La-proteina: Atomistic protein generation via partially latent flow matching. *arXiv*, 2025. doi: 10.48550/arXiv.2507.09466.

Mario Geiger, Emine Kucukbenli, Becca Zandstein, and Kyle Tretina. Accelerate drug and material discovery with new math library NVIDIA cuEquivariance, 2024. URL `https://developer.nvidia.com/blog/accelerate-drug-and-material-discovery-with-new-math-library-nvidia-cuequivariance/`.

Fatima N. Hitawala and Jeffrey J. Gray. What does AlphaFold3 learn about antibody and nanobody docking, and what remains unsolved? *mAbs*, 17(1):2545601, 2025. doi: 10.1080/19420862.2025.2545601.

Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin De Laroussilhe, Andrea Gesmundo, et al. Parameter-efficient transfer learning for NLP. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97, pp. 2790–2799, 2019. URL http://proceedings.mlr.press/v97/houlsby19a.html.

Leo C. James, Pietro Roversi, and Dan S. Tawfik. Antibody multispecificity mediated by conformational diversity. *Science*, 299(5611):1362–1367, 2003. doi: 10.1126/science.1079731.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, 2021. doi: 10.1038/s41586-021-03819-2.

Henry Kenlay, Frédéric A. Dreyer, Aleksandr Kovaltsuk, Dom Miketa, Douglas Pires, and Charlotte M. Deane. Large scale paired antibody language models. *PLOS Computational Biology*, 20 (12):e1012646, 2024. doi: 10.1371/journal.pcbi.1012646.

Christian Klein, Ulrich Brinkmann, Janice M. Reichert, and Roland E. Kontermann. The present and future of bispecific antibodies for cancer therapy. *Nature Reviews Drug Discovery*, 2024. doi: 10.1038/s41573-024-00896-6.

Luis S. Mille-Fragoso, John N. Wang, Claudia L. Driscoll, Haoyu Dai, Talal Widatalla, Xiaowei Zhang, et al. Efficient generation of epitope-targeted de novo antibodies with germinal. *bioRxiv*, 2025. doi: 10.1101/2025.09.19.677421.

Martin Pacesa, Lennart Nickel, Christian Schellhaas, Joseph Schmidt, Ekaterina Pyatova, Lucas Kissling, et al. One-shot design of functional protein binders with BindCraft. *Nature*, 646(8084): 483–492, 2025. doi: 10.1038/s41586-025-09429-6.

Giuseppe Pantaleo, Bruno Correia, Craig Fenwick, Victor S. Joo, and Laurent Perez. Antibodies to combat viral infections: development strategies and progress. *Nature Reviews Drug Discovery*, 21:676–696, 2022. doi: 10.1038/s41573-022-00495-3.

Wei Qu, Jiawei Guan, Rui Ma, Ke Zhai, Weikun Wu, and Haobo Wang. P(all-atom) is unlocking new path for protein design. *bioRxiv*, 2024. doi: 10.1101/2024.08.16.608235.

James P. Roney and Sergey Ovchinnikov. State-of-the-art estimation of protein model accuracy using AlphaFold. *Physical Review Letters*, 129(23):238101, 2022. doi: 10.1103/PhysRevLett.129.238101.

Jeffrey A. Ruffolo, Aadyot Bhatnagar, Joel Beazer, Stephen Nayfach, Jordan Russ, Emily Hill, et al. Adapting protein language models for structure-conditioned design. *bioRxiv*, 2024. doi: 10.1101/2024.08.03.606485.

Schrödinger, LLC. The PyMOL molecular graphics system, version 3.1.3.1. URL https://pymol.org.

David Sehnal, Sebastian Bittrich, Mandar Deshpande, Radka Svobodová, Karel Berka, Václav Bazgier, Sameer Velankar, Stephen K. Burley, Jaroslav Koča, and Alexander S. Rose. Mol* viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Research*, 49(W1):W431–W437, 2021. doi: 10.1093/nar/gkab314.

Joseph L. Watson, David Juergens, Nathaniel R. Bennett, Brian L. Trippe, Jason Yim, Helen E. Eisenach, et al. De novo design of protein structure and function with RFdiffusion. *Nature*, 620 (7976):1089–1100, 2023. doi: 10.1038/s41586-023-06415-8.

Jeremy Wohlwend, Gabriele Corso, Saro Passaro, Mateo Reveiz, Ken Leidal, Wojtek Swiderski, et al. Boltz-1: Democratizing biomolecular interaction modeling. *bioRxiv*, 2024. doi: 10.1101/2024.11.19.624167.

Rui Yin and Brian G. Pierce. Evaluation of AlphaFold antibody-antigen modeling with implications for improving predictive accuracy. *Protein Science*, 33(1):e4865, 2024. doi: 10.1002/pro.4865.

Vinicius Zambaldi, David La, Alexander E. Chu, Harshnira Patani, Amy E. Danson, Tristan O. C. Kwan, et al. De novo design of high-affinity protein binders with AlphaProteo. *arXiv*, 2024. doi: 10.48550/arXiv.2409.08022.

Zaixiang Zheng, Yifan Deng, Dongyu Xue, Yi Zhou, Fei Ye, and Quanquan Gu. Structure-informed language models are protein designers. In *Proceedings of the 40th International Conference on Machine Learning*, ICML'23. JMLR.org, 2023.

A Related Work

Generative All-Atom Models for Structure Design. Recent advances in generative structure modeling have demonstrated that operating directly in all-atom space can yield more accurate and physically realistic protein designs. RFdiffusion (Watson et al., 2023) first extended diffusion-based generative modeling to protein backbone design. While effective on certain targets, the method is limited to backbone coordinates and relies on a computationally heavy architecture that scales poorly. More recently, P(all-atom) (Qu et al., 2024) adapted components of AlphaFold3 (Abramson et al., 2024) for fully atomistic, unconditional protein design. This framework improves output granularity and sample efficiency relative to earlier equivariant backbone-only models, but the published model is unconditional and trained on a broad, non-targeted dataset, limiting its immediate usefulness for controlled conditional design. La-Proteina (Geffner et al., 2025) introduced a partially latent flow-matching approach that explicitly models backbones and represents side chains and sequences using low-dimensional latent variables per residue. This improves efficiency and supports joint sequence–structure generation. However, it does not directly support conditional, all-atom design of epitope-specific antibodies.

Generative Models for Antibody Design. RFantibody (Bennett et al., 2026) addresses some of these limitations by fine-tuning RFdiffusion (Watson et al., 2023) on antibody-antigen complexes, enabling conditional backbone generation with target specificity. The pipeline couples this backbone generator with ProteinMPNN (Dauparas et al., 2022) for subsequent sequence design. However, scalability of training and inference remains constrained by the computational demands of the underlying RFdiffusion (Watson et al., 2023) architecture. More recently, Germinal (Mille-Fragoso et al., 2025) demonstrated efficient *de novo* epitope-targeted antibody generation by unifying AlphaFold-Multimer (Evans et al., 2021b) with an antibody-specific language model under a shared gradient-based optimization framework. In contrast, our approach directly trains an all-atom diffusion model on a large-scale dataset of antibody–antigen complexes, simultaneously leveraging rich learned representations and enabling efficient, fully atomistic conditional structure generation.

Protein Inverse Folding. Machine learning-based inverse folding models have rapidly advanced protein sequence design conditioned on desired structures. ProteinMPNN (Dauparas et al., 2022) introduced a message-passing neural network architecture that achieved high sequence recovery rates across diverse protein folds, demonstrating strong generalization across protein families. A follow-up study demonstrated that combining ProteinMPNN with AlphaFold2 (Jumper et al., 2021) for self-consistency assessment enables more efficient design of protein binder sequences than traditional energy-based methods alone (Dauparas et al., 2022). ProteinMPNN is widely used in protein sequence design systems, including for antigen-specific antibody design. AbMPNN (Dreyer et al., 2023) adapted inverse folding specifically for antibodies, incorporating variable region constraints to improve CDR sequence recovery and design. PiFold (Gao et al., 2023) demonstrated superior performance on the general protein inverse folding problem through an enhanced featurization scheme and a more efficient, non-autoregressive Graph Neural Network (GNN) architecture. Benchmarking against ProteinMPNN and other inverse folding models showed substantially lower sequence perplexity and higher amino acid recovery. However, both ProteinMPNN and PiFold are limited by the availability of solved protein structures, which is significantly smaller than the space of known protein sequences. To address this limitation, LM-Design (Zheng et al., 2023) integrated structure-based models (ProteinMPNN and PiFold) with protein language models trained on large sequence datasets, boosting sequence recovery performance by 4–12 percent. However, all of these models have limitations in terms of how tightly they integrate structural and sequence conditioning information.

Antibody Sequence Design. Our primary objective is to generate antibody sequences that bind specific target antigens without known CDR sequence context and thus models trained for this data-constrained task are of particular interest. AbMPNN (Dreyer et al., 2023) and other antibody-specific inverse folding models were developed by fine-tuning general-purpose models on antibody structural data from SAbDAb (Dunbar et al., 2014). These models relied on training datasets augmented with synthetic structures for sequences from the Observed Antibody Space (OAS). Such antibody-specific models demonstrated enhanced CDR re-design capability in terms of sequence recovery, loop self-consistency RMSD, and Ab-Ag affinity ranking over general-purpose inverse folding models when

evaluated *in silico* with full backbone structure and sequence context. More recently, ProseLM (Ruffolo et al., 2024) offered a comprehensive framework for structure-informed protein design by integrating a causal encoder with the ProGen2 pLM, combining novel architecture with extensive experimental validation. Their antibody-specific results include a novel PD-1 therapeutic antibody binder with 2.2 nM affinity. In combination with these efforts, the advent of pre-trained paired-antibody language models such as IgBert and IgT5 (Kenlay et al., 2024) has presented an opportunity for further innovation in the antibody-specific, structure-conditioned sequence design space.

B Detailed Architectures

B.1 Structure design via AbsciDiff

AbsciDiff (Figure 5) is a diffusion-based all-atom generative model fine-tuned from Boltz-1 (Wohlwend et al., 2024) for epitope-conditioned antibody design. Key modifications include antibody- and docking-specific feature masking and conditioning strategies, an intermediate sequence hypothesis module with recycling, integration of optimized CuEquivariance kernels (Geiger et al., 2024), and support for structural templates. The sampling procedure follows Boltz-1, preserving the sample efficiency and stability of the parent model. In the following sections, we describe the diffusion formulation, feature generation strategy, architectural changes, and training protocol.

B.1.1 AbsciDiff Diffusion Formulation

Following Boltz-1, we model the noise added to the native structure of an antibody-antigen complex $X_0 \in \mathbb{R}^{N\times 3}$ through a forward diffusion process,

$$dX_t = \sqrt{2t}\, dB_t, \tag{1}$$

where dB_t is $(N \times 3)$-dimensional Brownian motion.

To generate new complexes, we reverse this diffusion process over T timesteps (Figure 7). Starting from $X_T \sim \mathcal{N}(0, I)$, we iteratively denoise using a model P_θ parametrized by a neural network and conditioned on antigen and framework structure templates X_a, X_f, sequences S_a, S_f, and epitope residues E_a. We denote the full conditioning information as $C = \{X_a, X_f, S_a, S_f, E_a\}$.

We train this model to approximate the expected coordinates of the original complex structure:

$$P_\theta(X_t, t; C) \approx \mathbb{E}_{X_0|X_t}[X_0], \tag{2}$$

with the standard denoising loss

$$\mathcal{L}(\theta) = w_t \cdot \|P_\theta(X_t, t; C) - X_0\|^2, \tag{3}$$

where w_t is a weight proportional to the variance of the noising process at timestep t.

B.1.2 AbsciDiff Featurization

Antigen, Framework, and Epitope Conditioning. AbsciDiff's most significant departure from Boltz-1 concerns feature masking and conditioning. For our intended use, we assume antigen structures are known, and we aim to design or redesign only the CDR regions while keeping therapeutic framework sequences and structures intact. We additionally condition on an epitope defined during training as the set of antigen residues within 6 Å of the antibody, down-sampled via a geometric distribution ($p = 0.3$) for robustness. We mask sequence and structure features for all CDR positions, as well as all inter-chain pair features. Docking is thus guided only by the token-wise binary epitope vector, implemented analogously to Boltz-1's pocket-conditioning feature.

Design Region Representation. Amino acid identities in the design region are initially set to the unknown (UNK) amino acid token for both sequence input and reference conformer lookup. During structure prediction, the model is trained to perform atom superposition, as in P(all-atom) (Qu et al., 2024). For this task, all amino acids in the design region are represented in Atom14 notation, with excess "virtual" atoms placed at the location of their residue's Cα (Figure 6). Amino acids outside the design region have no virtual atoms since their identities are known *a priori*. Other than the sequence hypothesis head described below, AbsciDiff performs no sequence decoding on designed amino acids.

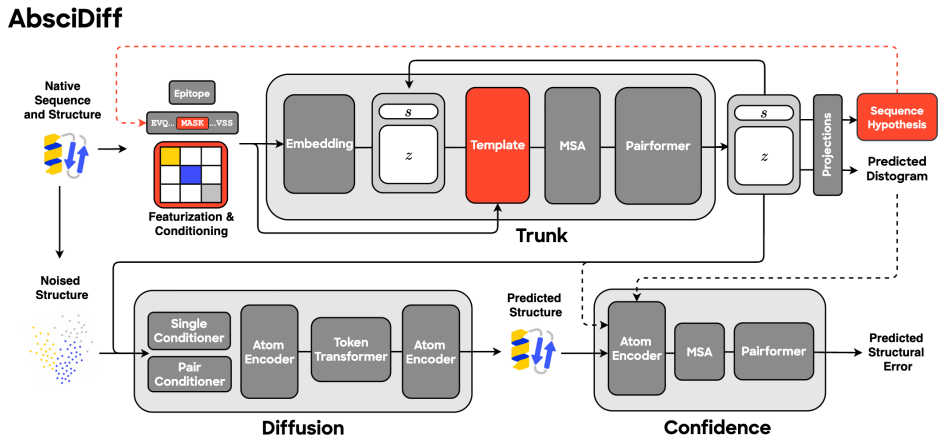


Figure 5: **AbsciDiff architecture.** Building on the architecture of Boltz-1, AbsciDiff uses a pairformer-based trunk to provide conditioning information to a diffusion module. Modifications include invariant feature conditioning via an AlphaFold3-like template module (Abramson et al., 2024), as well as a module to predict a sequence hypothesis. The sequence hypothesis is used to update the model's input features for a second forward pass through the model. Dashed lines indicate a stopped gradient. Red indicates modifications to Boltz-1.

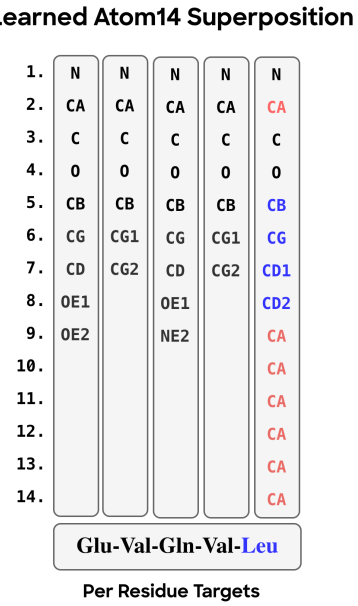


Figure 6: **Atom14 representation.** An example of AbsciDiff's atom superposition technique. Non-design region amino acids are represented with their true atomic structure since their identities are fixed a priori. Design region residues (see Leu in blue) use the Atom14 representation, where excess virtual atoms beyond the residue's actual atom count are placed at the Cα position.

Template Structures. AbsciDiff also adds support for structural template featurization in two ways: endogenous templating (in which pairwise residue distance information is derived from a user-provided structure) and exogenous templating (where the same information is sourced from external template databases). In both cases, structural information is encoded and embedded as is performed in AlphaFold3. During development, we found that endogenous templating was the more practical and effective strategy for providing structural guidance without introducing bias. Thus, the final training runs included only endogenous template information.

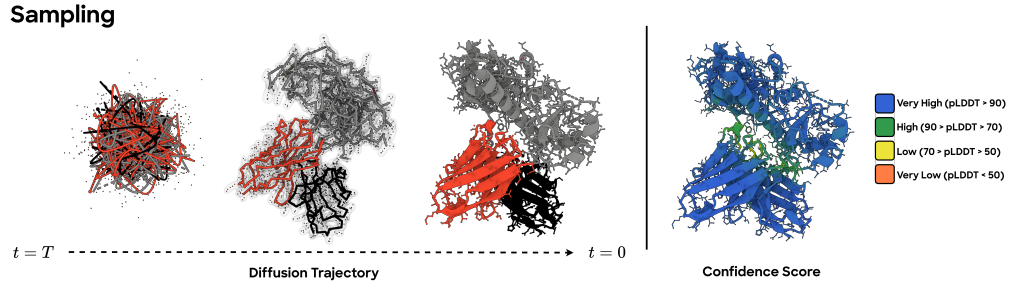


Figure 7: **Reverse diffusion trajectory for joint antibody–antigen structure generation.** Starting from a fully corrupted state X_T (*left*), AbsciDiff performs iterative denoising transitions $p(X_{t-1}|X_t, t)$ (*middle*), progressively recovering global geometry and local structure until producing a final denoised sample X_0 (*center right*). Subsequently predicted confidence computed on the final structure is shown on the far right (color scaled by residue pLDDT, where blue represents high confidence and red represents low confidence).

Multiple Sequence Alignments (MSAs). AbsciDiff supports MSA featurization as in Boltz-1, though it remains disabled by default. We found that full MSA feature sets performed no better than a one-hot encoding of the input sequence when conditioned on endogenous templates.

B.1.3 AbsciDiff Trunk Module and Intermediate Sequence-Informed Design

An auxiliary prediction head in AbsciDiff produces a design region sequence hypothesis from the final single (s) representation of the trunk. The predicted logits are decoded into the most likely amino acid tokens and, along with corresponding atom-level reference conformer features, are recycled through the trunk with gradient flow stopped. Sequence prediction is trained using cross-entropy loss against the native sequence. We hypothesize that intermediate sequence prediction and conformer recycling assists the design process by focusing the search space of potential residue identities and local atomic conformations prior to the diffusion model performing final design.

B.1.4 AbsciDiff Confidence Module and Design Pre-Ranking

Though the underlying confidence module architecture is the same as Boltz-1, AbsciDiff introduces a new on-model ranking and filtering strategy. Like Boltz-1, the diffusion module emits M samples per sampling pass, which are subsequently scored using the confidence module to generate predicted Template Modeling scores (pTM), predicted Local Distance Difference Test scores (pLDDT), predicted Docking Error (pDE), and predicted Aligned Error (pAE). We normalize these scores (noting that pDE and pAE are unbounded and must be transformed and inverted) before averaging to produce a composite ranking score by which the top $k \leq M$ candidates are selected. Although the confidence module is efficient—and its scores well-correlated with supervised metrics—we do not rely on it alone for evaluating real-world binding affinity. As such, we employ this on-model scoring strategy for efficient pre-ranking, thereby reducing the downstream computational cost of re-folding and scoring designs with AbsciBind.

B.1.5 AbsciDiff Training

Fine-tuning. We initialize AbsciDiff using the weights from the provided Boltz-1 checkpoint and finetune the model on an internally curated variant of the SAbDab dataset (Dunbar et al., 2014). As described earlier, redundant dataset entries were filtered based on sequence similarity thresholds, and data were split using a combination of temporal and sequence similarity-based criteria. Training follows Boltz-1 defaults with the number of learning rate steps reduced to 20 and an effective batch size of 64 for 10 epochs.

Cropping. Input structures are spatially cropped to a maximum of 512 residues, with larger crop sizes showing minimal improvement on sampled antibody-antigen interface quality. The cropping strategy centers the representation on the antibody-antigen interface and requires that: (1) the variable

antibody region (Fv) is always included; (2) antibody constant domains are always excluded; (3) the remaining budget is allocated to antigen residues in order of the smallest distance to the epitope.

B.2 Sequence Design via IgDesign2

IgDesign2 (Figure 8) combines: 10-layer PiFold GNN encoder (dim 128, 4 heads, 30 neighbors), 10-layer causal transformer decoder (4 heads, FFN dim 512), and IgBert pLM with layer-wise fusion via trainable MLPs. In the following sections, we describe the problem formulation, model architecture, as well as pre-training and fine-tuning protocols.

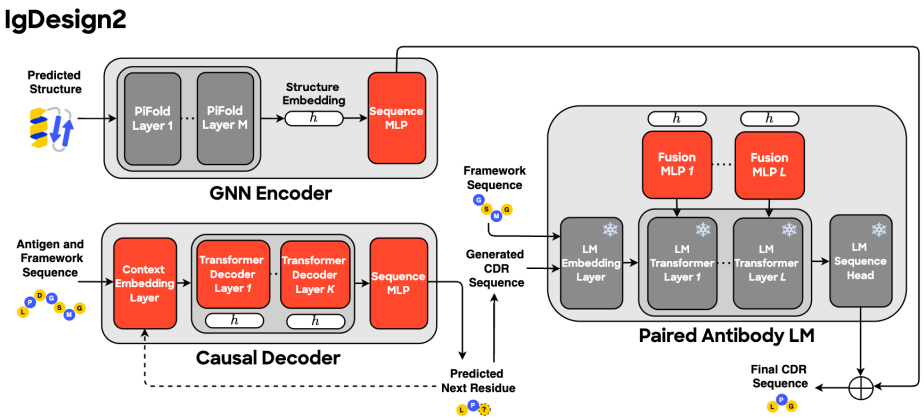


Figure 8: **IgDesign2 architecture.** A GNN encoder processes antibody-antigen structure features, followed by a causal transformer decoder for autoregressive sequence generation. The decoder supports both temperature-weighted sampling (for diversity) and beam search (for high-likelihood sampling). Generated sequences are refined by a structure-aware paired antibody language model. Red indicates model components developed specifically for this work. Snowflakes indicate frozen pretrained modules. Solid arrows show forward information flow. Dotted arrows show recycled information flow.

B.2.1 IgDesign2 Problem Formulation

We model the joint probability of the unknown sequence region R using an autoregressive factorization:

$$R = (r_1, r_2, \ldots, r_N), \tag{4}$$

$$p_\theta(R|X, C) = \prod_{i=1}^{N} p_\theta(r_i|r_{<i}, X, C), \tag{5}$$

where each amino acid r_i is conditioned on previously generated positions $r_{<i}$, backbone structure X, and known sequence context C. Our goal is to find the neural network parameterization θ that maximizes the likelihood of the training data under this scheme.

B.2.2 IgDesign2 Model Architecture

IgDesign2 combines a Graph Neural Network (GNN) encoder, a causal transformer decoder, and a protein language model (pLM) refinement module into a structure-conditioned sequence design system. The input to the encoder consists of the four-atom representation of the protein backbone [N, C_α, C, O] residues in which each residue constitutes a node in a graph with k edges connecting to its nearest neighbors in Euclidean space. We leverage PiFold's (Gao et al., 2023) node and edge featurization scheme and mirror their GNN message passing architecture for encoding the three-dimensional geometry of each residue based on its atomic coordinates. The decoder ingests the structure embeddings and any antigen and antibody framework sequence context for conditioning the autoregressive CDR sequence generation process, which is applied in a randomly shuffled decoding order via temperature-weighted sampling or beam search.

The heavy and light chain sequences are then provided to the IgBert paired antibody language model (Kenlay et al., 2024) for structure-aware refinement. At every layer of the pLM's transformer, the pLM latent embeddings are fused with the final structure embedding in a shared low-dimensional space before being projected back to the pLM dimension for processing by the subsequent transformer layer (Ruffolo et al., 2024; Houlsby et al., 2019; Gao et al., 2023). After the transformer has fully processed the fused representation, the final latent representations of both the GNN and pLM are decoded by their respective sequence prediction heads and integrated with a residual connection to produce the final sequence output. This jointly optimized "generate-and-refine" approach combines the strengths of both sequence modeling paradigms to enhance the quality of our conditional CDR designs.

B.2.3 IgDesign2 Pre-Training

We pre-train the encoder and decoder on protein-protein interaction examples from an internally curated variant of SAbDab (Dunbar et al., 2014), spatially cropped in three dimensions around the interfacial region of the two chains to a maximum of 500 total residues. We use the Adam optimizer with a learning rate of 10^{-3}, $\beta_1 = 0.9$, and $\beta_2 = 0.999$, along with an effective batch size of 32 and a standard cross-entropy loss over 20 possible amino acids. Early stopping is applied when the loss has not improved for 10 consecutive epochs on the held-out validation set. We retain the published hyperparameters of the original PiFold model for the encoder (dimension 128, 10 layers, 4 heads, 30 nearest neighbors in the structure graph), while the causal transformer decoder consists of 10 standard transformer decoder layers with four attention heads and a feedforward dimension of 512. We also apply Gaussian noise with a standard deviation of 0.1 Å to the coordinates of each input structure during training to encourage robustness.

B.2.4 IgDesign2 Fine-Tuning

We fine-tune the model with the same hyperparameters, aside from two adjustments: we reduce the effective batch size to 8 and terminate training when the cross-entropy loss over the masked region has not improved for 5 consecutive epochs on the held-out validation set. Antibody-antigen complexes are spatially cropped to a maximum of 600 residues around the interacting residues of the antigen while retaining the full antibody Fv. The pLM weights remain frozen throughout fine-tuning, while a trainable multi-layer perceptron (MLP) is introduced per pLM layer to fuse the incoming projections from PiFold and the pLM. Each MLP consists of three standard linear layers of dimension 128, separated by ReLU activation functions. The incoming projections from PiFold and the pLM reduce their respective embeddings to 64 dimensions prior to concatenation, while the final projection maps the 128-dimensional MLP output back to the pLM embedding dimension. We compute the cross-entropy loss over both the sequence produced by the causal decoder and the sequence refined by the pLM, summing them to compute the total loss.

B.3 Ranking via AbsciBind Score

The interface predicted TM-score (ipTM) is a confidence metric from AlphaFold-Multimer that estimates the accuracy of predicted inter-chain interfaces. For a predicted structure with chains A and B, ipTM measures the expected TM-score of interface residues:

$$\text{ipTM} = \frac{1}{L_{\text{interface}}} \sum_{i \in \text{interface}} \frac{1}{1 + (d_i/d_0)^2} \tag{6}$$

where d_i is the predicted error for residue i and d_0 is a length-dependent normalization factor. Higher ipTM indicates greater confidence in the predicted interface geometry.

Antibody-Aligned ipTM. Standard ipTM treats heavy and light chains separately, diluting the antibody–antigen signal. We introduce Antibody-Aligned ipTM, which merges the heavy and light chains into a single "antibody" entity before computing ipTM:

$$\text{ipTM}_{\text{Ab-aligned}} = \text{ipTM}([\text{H} \oplus \text{L}], \text{Ag}) \tag{7}$$

This focuses the metric on the therapeutically relevant antibody–antigen interface rather than intra-antibody chain contacts.

AbsciBind Score. The final AbsciBind Score is the arithmetic mean of standard ipTM and Antibody-Aligned ipTM:

$$\text{AbsciBind Score} = \frac{1}{2}(\text{ipTM} + \text{ipTM}_{\text{Ab-aligned}}) \tag{8}$$

Designs with ligand RMSD >5 Å between designed and predicted structures are filtered as likely non-binders.

Protocol. Our implementation modifies AF_Unmasked as follows: (1) full sequence input; (2) designed structure as template; (3) retain amino acid tokens; (4) mask sidechains except Cβ; (5) disable inter-chain template distance masking; (6) single-sequence mode; (7) disable dropout/masking. We use Model 2 (model_2_multimer_v3) for $\sim$80% runtime reduction with comparable discrimination.

C Dataset Curation

We curated protein structures for training via: (1) PDB extraction with missing residue/atom resolution; (2) interface metadata extraction; (3) structure-based splitting.

Filtering. Fvs require: heavy chain present, antigen $\geq$15 residues, epitope $\geq$5 residues, no missing CDR residues, BSA $\geq$500 Å^2, resolution <9 Å (train) or <3.5 Å (val/test).

Splitting. Temporal: <2024-01-01 (train), 2024-01-01–2024-09-30 (val), >2024-09-30 (test). Homology: >40% antigen similarity removed. Final: 3,242/63/84 antibody–antigen; 29,835/356/322 protein–protein.

D Library Design Pipeline

Wide Structure Search. 3,360 specs/target $\rightarrow$ 24 samples/spec $\rightarrow$ top 3 by confidence $\rightarrow$ 1 sequence each $\rightarrow$ top 5% by max/median ipTM.

Deep Structure Search. 10$\times$ oversample $\rightarrow$ filter L-RMSD <5 Å $\rightarrow$ top 500 (25% by ipTM, 75% by intersection score) with clustering for diversity.

Sequence Search. 320 seqs/structure $\rightarrow$ filter liabilities $\rightarrow$ top 20 by perplexity $\rightarrow$ AbsciBind $\rightarrow$ final 95 designs.

E Data Availability

SPR screening data for *de novo* designs and lead optimization variants, including sequences, target annotations, and binding classifications, are available at `https://github.com/AbSciBio/origin-1`. Computational models generated by AbsciGen and AbsciBind for all reported designs are also provided.

F Additional In Silico Results and Methods

F.1 AbsciBind Validation

We validated the AbsciBind protocol against standard AlphaFold-Multimer approaches (Figure 9).

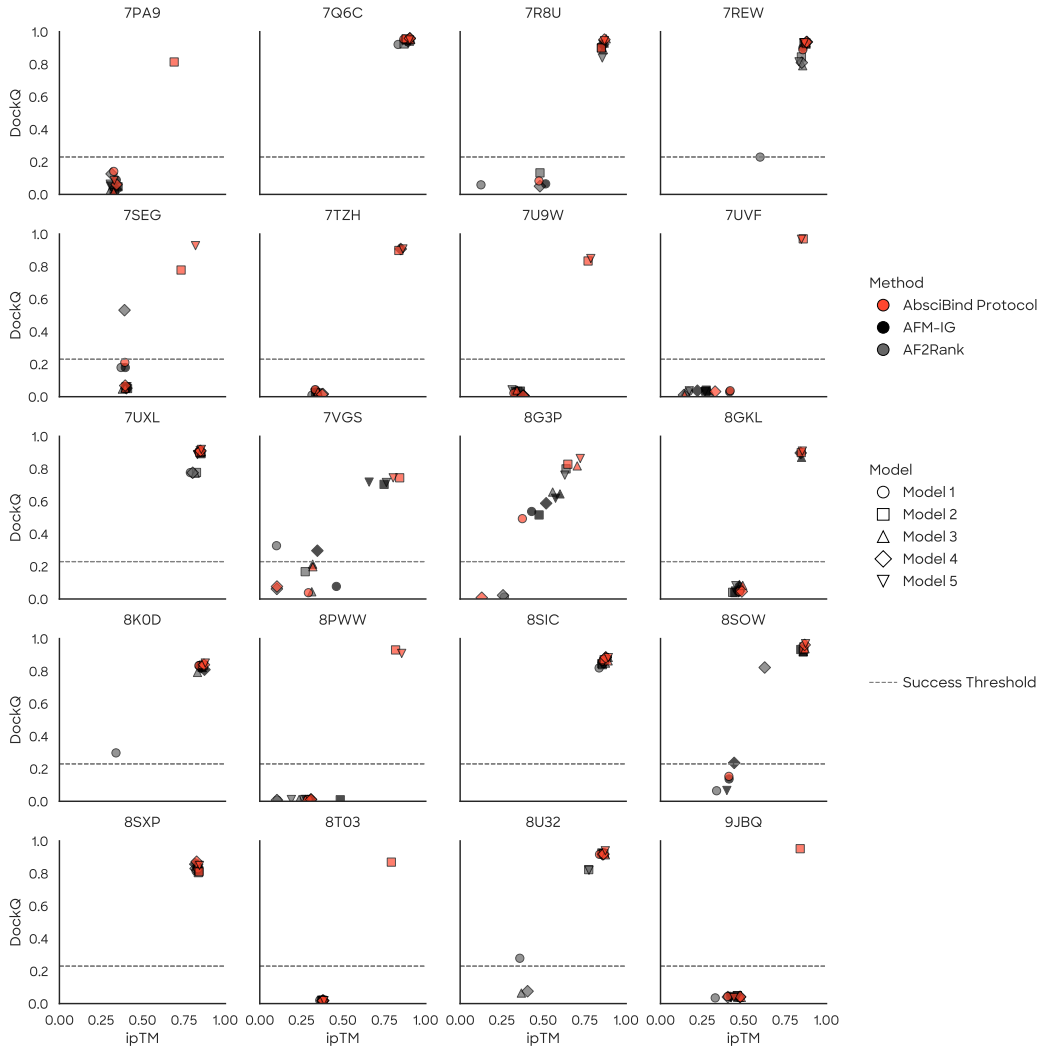


Figure 9: **AbsciBind protocol validation.** Distribution of DockQ and ipTM scores for AlphaFold-Multimer predictions generated using the initial-guess (AFM-IG), AF2Rank, and AbsciBind protocols. The gray dashed line at $y = 0.23$ marks the CAPRI threshold separating incorrect from acceptable predictions. DockQ scores are computed by comparing the merged antibody heavy–light chains against the antigen chain.

F.2 Coordinate Initialization for RFantibody Benchmarking

De novo antibody design using RFantibody required careful attention to coordinate initialization to achieve reasonable outputs. Without modification, we observed reduced performance compared to the supervised benchmarking setting, including chain breaks in predicted loop structures and glycine-dominated sequences from ProteinMPNN.

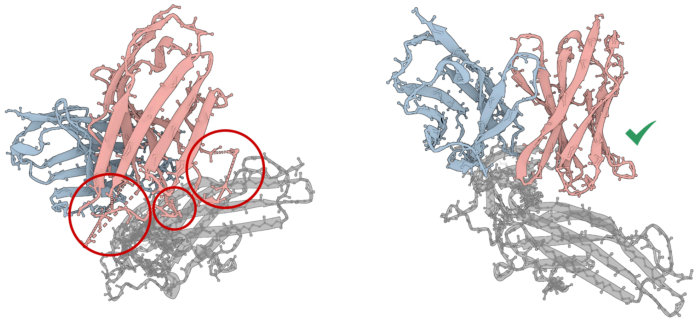


Figure 10: **RFantibody loop quality depends on coordinate initialization.** Chain breaks appear when loop coordinates are placed at the origin (*left*), but not when idealized coordinates are provided (*right*). Provided framework PDB files all include ground-truth loop coordinates, motivating our preprocessing changes for a fair *de novo* benchmark.

To avoid potential information leakage under the *de novo* setting, our pipeline preprocesses structures such that CDR loop coordinates are placed at the origin for the designed number of residues, and CDR sequences are populated with glycines. While RFantibody fully masks features for designed loops, we found that initial loop coordinates influence the diffusion process. Specifically, RFantibody's reverse diffusion implementation during inference parameterizes the initial noise distribution on the input Cα coordinates at timestep $t = T$ (rather than sampling from an unconditional prior), causing spatial biases in input coordinates to persist through noising.

By default, when a designed loop length differs from the input "framework" PDB, RFantibody initializes coordinates using idealized backbones with random noise (via `adjust_loop_lengths`). However, when the loop length already matches—as in our pipeline, where we prepare PDBs for each desired length with origin-initialized CDR coordinates—the model uses the provided coordinates directly. Furthermore, provided examples in the public RFantibody release include actual loop coordinates from the RCSB PDB entry.

To ensure a fair benchmark on examples lacking provided loop coordinates while also limiting changes to the RFantibody codebase, we applied an idealized initialization based on loop "dilation" logic to all CDR coordinates:

$$\mathbf{x}_{\text{CDR}}(0) \leftarrow \mathbf{x}_{\text{ideal}} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{U}(-5, 5)^3, \tag{9}$$

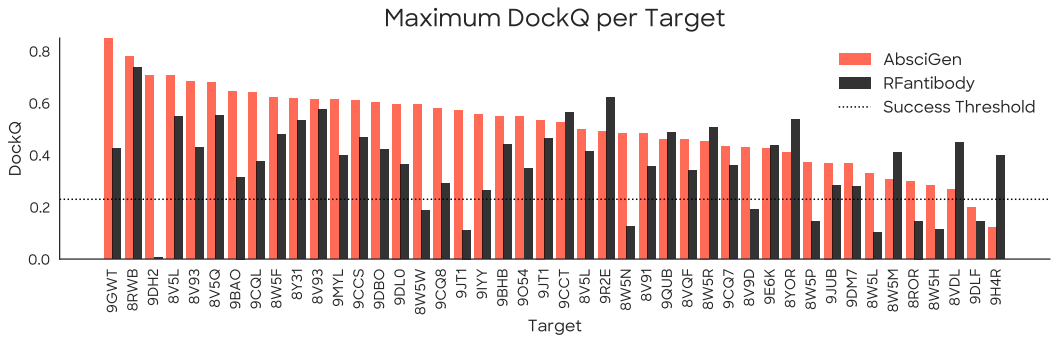
where the idealized backbone coordinates are defined as:

$$\mathbf{x}_{\text{ideal}} = \{N(-0.527, 1.359, 0),\ C\alpha(0, 0, 0),\ C(1.523, 0, 0)\}. \tag{10}$$

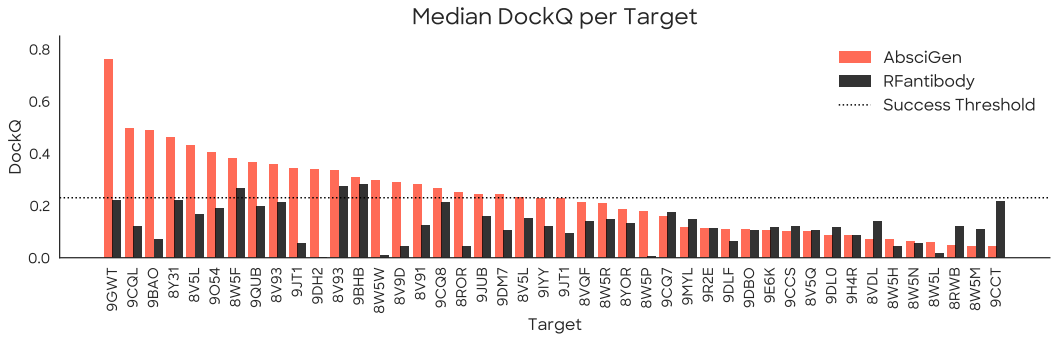
This approach does not fully decouple designs from initialization—loop length and random seed still influence the result—but it respects the intention of the existing loop length dilation methodology without introducing more substantial modifications. Unless otherwise noted, all benchmarking results use this modified initialization.

F.3 Supervised Benchmarking

We benchmarked AbsciGen against RFantibody on targets with known structures (Figures 11–12).



(A)



(B)

Figure 11: **Supervised benchmark: DockQ by target.** (A) Per-target comparison of maximum DockQ scores for AbsciGen and RFantibody. The acceptable threshold (DockQ = 0.23) is indicated by the dashed line. (B) Per-target comparison of median DockQ scores. AbsciGen achieves acceptable DockQ scores ($\geq$0.23) on 95.5% of targets, compared to 77.3% for RFantibody.

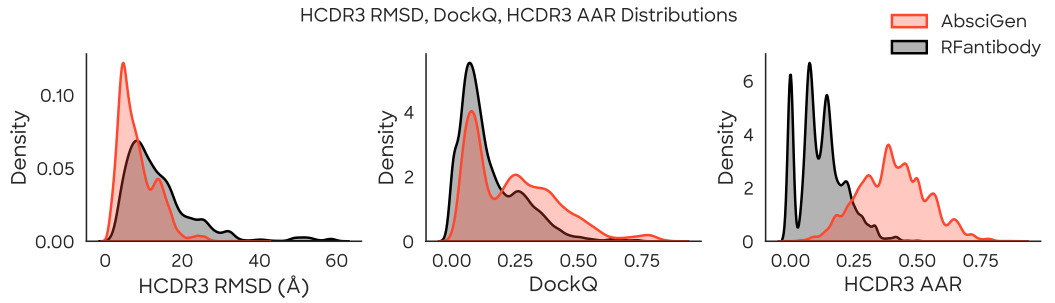
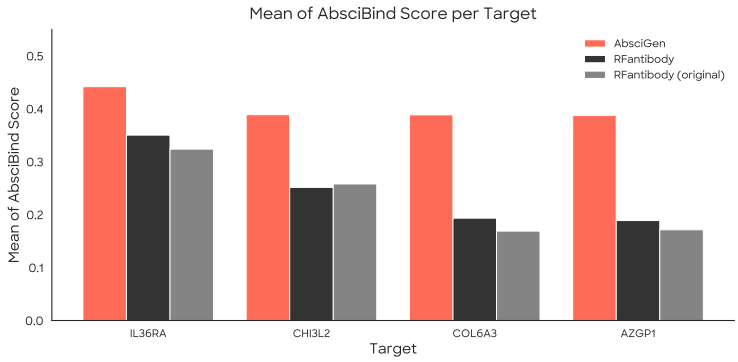


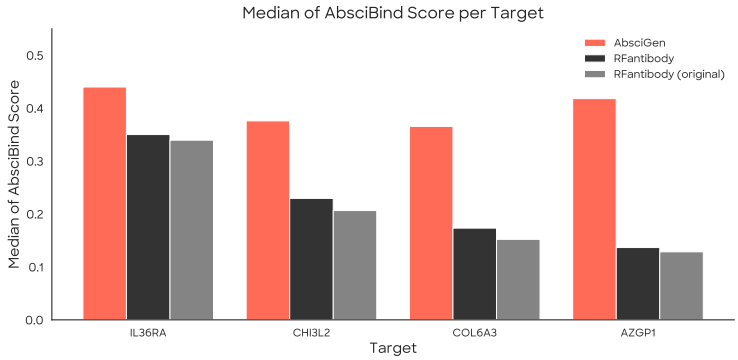
Figure 12: **Distribution of supervised benchmark metrics.** Distribution of HCDR3 RMSD, DockQ, and amino acid recovery (AAR) for AbsciGen and RFantibody across all generated samples. The dashed line indicates the DockQ success threshold (0.23). Lower HCDR3 RMSD indicates better structural agreement with the native binding pose.

F.4 Unsupervised Benchmarking

We evaluated AbsciGen on zero-prior targets where no ground-truth structures exist (Figure 13).



(A)



(B)

Figure 13: **Unsupervised benchmark: AbsciBind Scores by target.** (A) Per-target comparison of mean AbsciBind Scores for AbsciGen and RFantibody. (B) Per-target comparison of median AbsciBind Scores. The success threshold of 0.5 is indicated by the dashed horizontal line. The original RFantibody implementation without initialization modification is included for comparison.

F.5 Humanness Assessment

We assessed sequence humanness using OASis percentile scores (Figure 14).

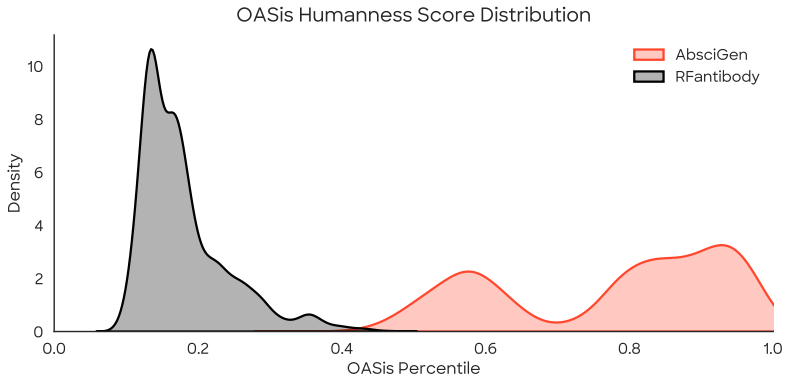


Figure 14: **OASis humanness scores.** Distribution of Observed Antibody Space identity (OASis) percentile scores for sequences generated by AbsciGen and RFantibody across four targets (COL6A3, AZGP1, CHI3L2, IL36RA). Higher percentiles suggest sequences are more likely to exhibit human-like antibody qualities. AbsciGen achieves significantly higher humanness scores, attributed to IgDesign2's fine-tuning on antibody-antigen complexes and integration of the IgBert paired language model.

F.6 Epitope Ablation

We evaluated how epitope contact conditioning affects design quality (Figure 15).

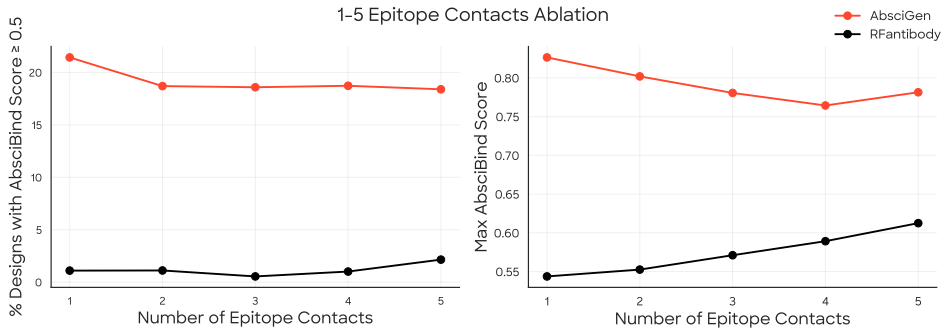


Figure 15: **Epitope contact conditioning study.** Evaluation of how varying the number of provided epitope contact residues impacts binding quality as measured by AbsciBind Score. No significant improvement in the fraction of designs achieving acceptable AbsciBind Score ($\geq$0.5) is observed with increasing contact information. *Left*: Fraction of designs with AbsciBind Score $\geq$0.5. *Right*: Maximum AbsciBind Score observed across generated samples. Adding contact constraints may restrict pose diversity for AbsciGen, while RFantibody benefits from enhanced conditioning.

F.7 Design Diversity

We also analyzed sequence and structural diversity of generated designs (Figures 16–17, Table 1).

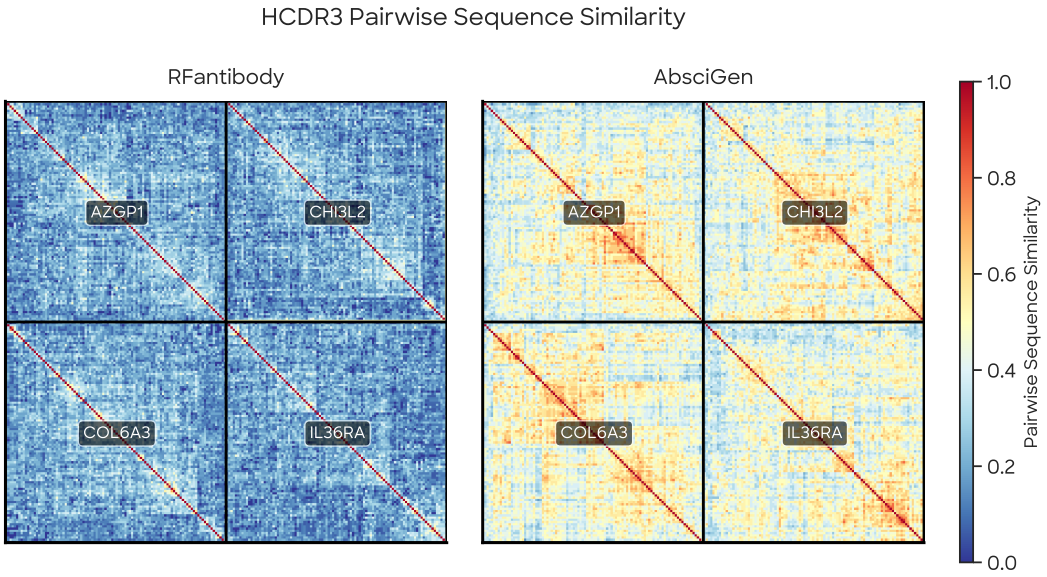


Figure 16: **HCDR3 pairwise sequence similarity matrices, grouped by target.** Each pixel shows the pairwise sequence similarity (PSS) score between two sequences generated for the same target (red = high similarity/low diversity, blue = low similarity/high diversity). Sequences are hierarchically clustered within each target block. RFantibody (*left*) shows more diverse outputs per structure, while AbsciGen (*right*) generates more similar sequences within each structure but achieves comparable target-level diversity through varied backbone generation.

Table 1: HCDR3 sequence diversity at structure and target levels.

	Structure-level		**Target-level**	
Method	Frac Unique ↑	PSS ↓	Frac Unique ↑	PSS ↓
RFantibody	$\mathbf{0.90 \pm 0.01}$*	$\mathbf{0.74 \pm 0.01}$*	1.00 ± 0.00	$\mathbf{0.20 \pm 0.01}$*
AbsciGen	0.49 ± 0.03	0.92 ± 0.00	$\mathbf{1.00 \pm 0.00}$	0.46 ± 0.01

PSS = Pairwise Sequence Similarity (lower = more diverse). *$p < 0.05$, Mann-Whitney U test. AbsciGen exhibits lower per-structure diversity but achieves full target-level diversity through varied backbone generation.

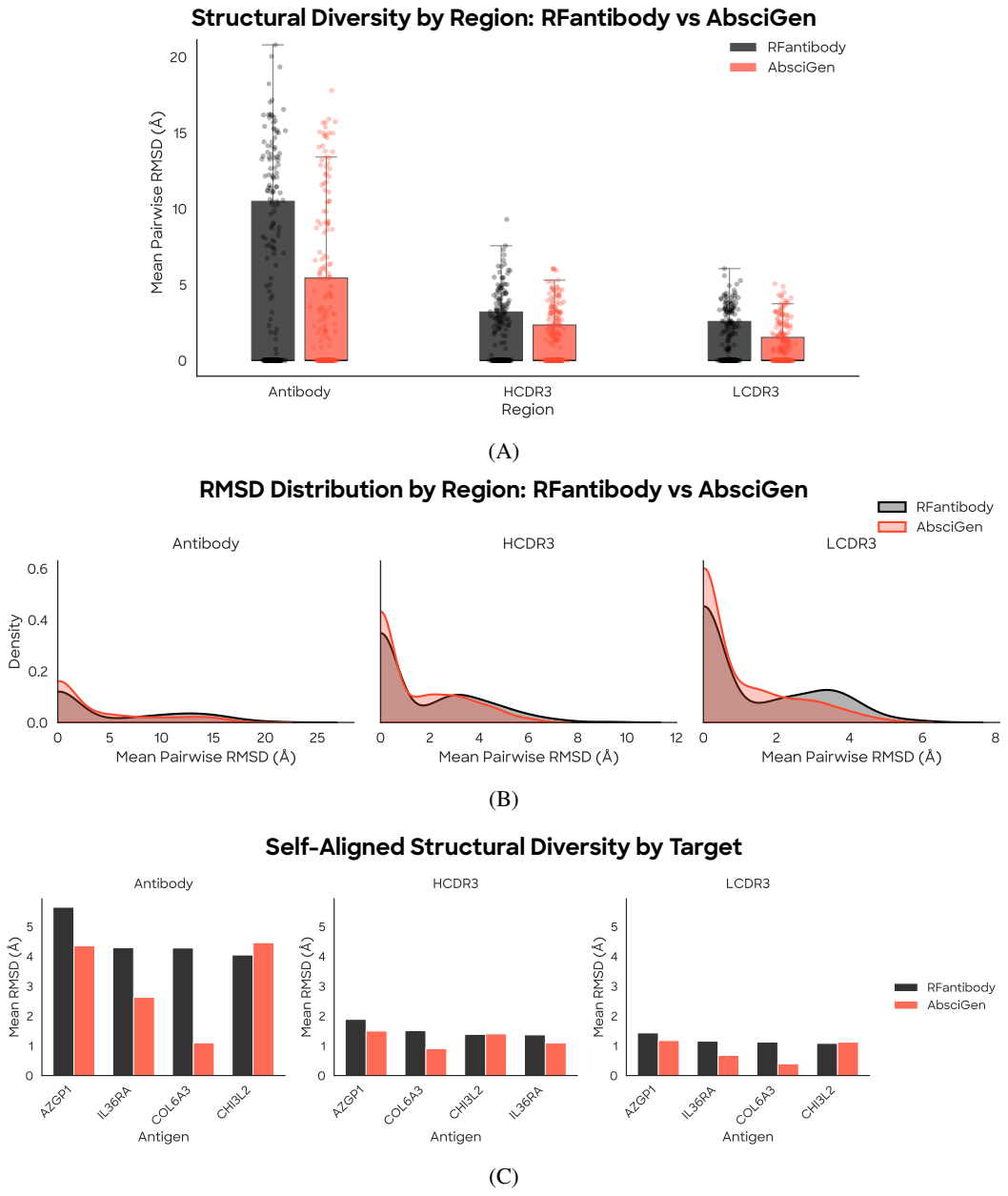


Figure 17: **Self-aligned structural diversity comparison.** Pairwise backbone RMSD (N, Cα, C, O atoms) was computed within groups of comparable antibody structures sharing the same antigen target, CDR lengths, and framework lengths. Each region was optimally superimposed onto itself using the Kabsch algorithm before computing RMSD, measuring conformational diversity independent of global orientation. (A) Boxplots show the distribution of mean pairwise RMSD across comparable groups for antibody Fv, HCDR3, and LCDR3 regions. (B) Kernel density estimates showing the full distribution of RMSD values. (C) Mean RMSD per antigen target, showing consistent trends across diverse antigens. RFantibody (black) exhibits greater structural diversity than AbsciGen (red) across all regions and targets.

G Additional In Vitro Results and Methods

G.1 Lead Optimization

Single-mutant libraries for CDRs + paratope (<5 Å from antigen). Scored by ESM ensemble, AbLang2, AbsciBind. Selection: alanine scan, germline reversions, positions improved by structure + sequence models. 94 variants/parent.

G.2 Experimental Protocols

SPR. Carterra LSAXT, SAHC30M chips, antigens 2 μM serially diluted 4-fold, 1:1 Langmuir fit. Hits: >10 RU, on-target only.

BLI. Gator Prime, Anti-Human Fab probes, 5 μg/mL loading, 31.25–2000 nM antigen, $R^2 > 0.95$.

SEC. 1.3:1 antigen:Fab, Superdex200.

Cryo-EM. Titan Krios 300 kV. AZGP1: Falcon4i, 0.932 Å/px. COL6A3: K3, 0.824 Å/px. RELION 4.0, CryoAtom, PHENIX. Structure figures were generated using PyMOL (Schrödinger, LLC) and Mol* (Sehnal et al., 2021).

HEKBlue. IL36γ (2 pM), IL36RA (50 nM), antibody 1 nM–1 μM, SEAP/QuantiBlue 620 nm.

G.3 Developability Results

Table 2: Developability assessment of top binders.

Target	PR	HIC	AC-SINS	T_m	PDI
COL6A3	3.5/3.8	1.03	1.6	68.0	0.00
AZGP1	15.8/8.5	N.D.	25.7	77.9	0.08
IL36RA	3.0/3.6	1.19	0.6	70.6	0.00
CHI3L2	3.6/4.0	0.91	-0.1	70.2	0.01

PR = Polyreactivity (RNS); HIC = Hydrophobic Interaction Chromatography (RRT); AC-SINS = self-association (nm); T_m = melting temperature (°C); PDI = polydispersity index.

G.4 Per-Target Binding Validation

Complete validation data for all four targets are shown below, including SPR binding, BLI kinetics, SEC complexation, cryo-EM structures, and functional assays.

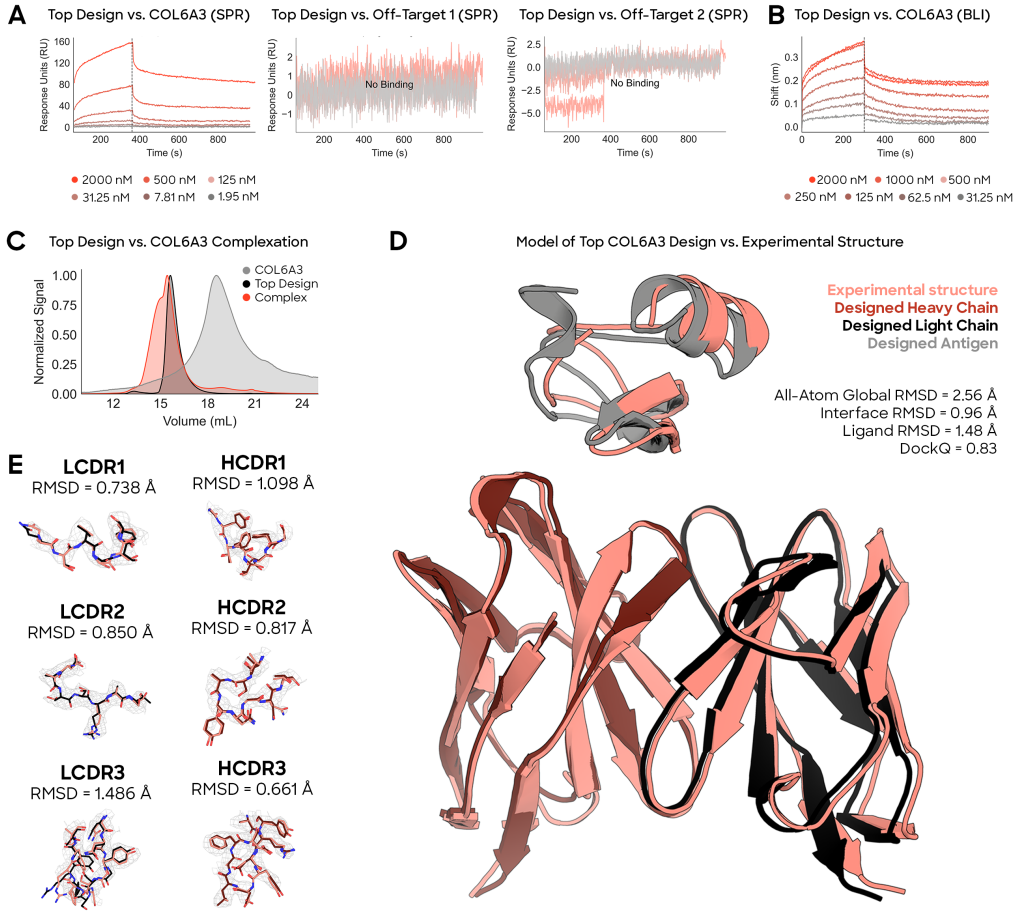


Figure 18: ***In vitro* experimentation confirms that Origin-1 generated a *de novo* binder against a novel epitope on COL6A3.** (A) SPR demonstrates that the top design binds to COL6A3 in mAb format and does not bind to either of two unintended targets ("Off-Target 1" and "Off-Target 2"). (B) BLI confirms that the top design binds to COL6A3 in Fab format. (C) Complexation experiment confirms that the top design, in Fab format, binds to COL6A3 in solution. (D) Cryo-EM of top design complexed with COL6A3 confirms epitope-specificity and atomic accuracy of the Origin-1 computational model. (E) CDR-specific analysis of model vs. solved complex structure confirms atomic accuracy.

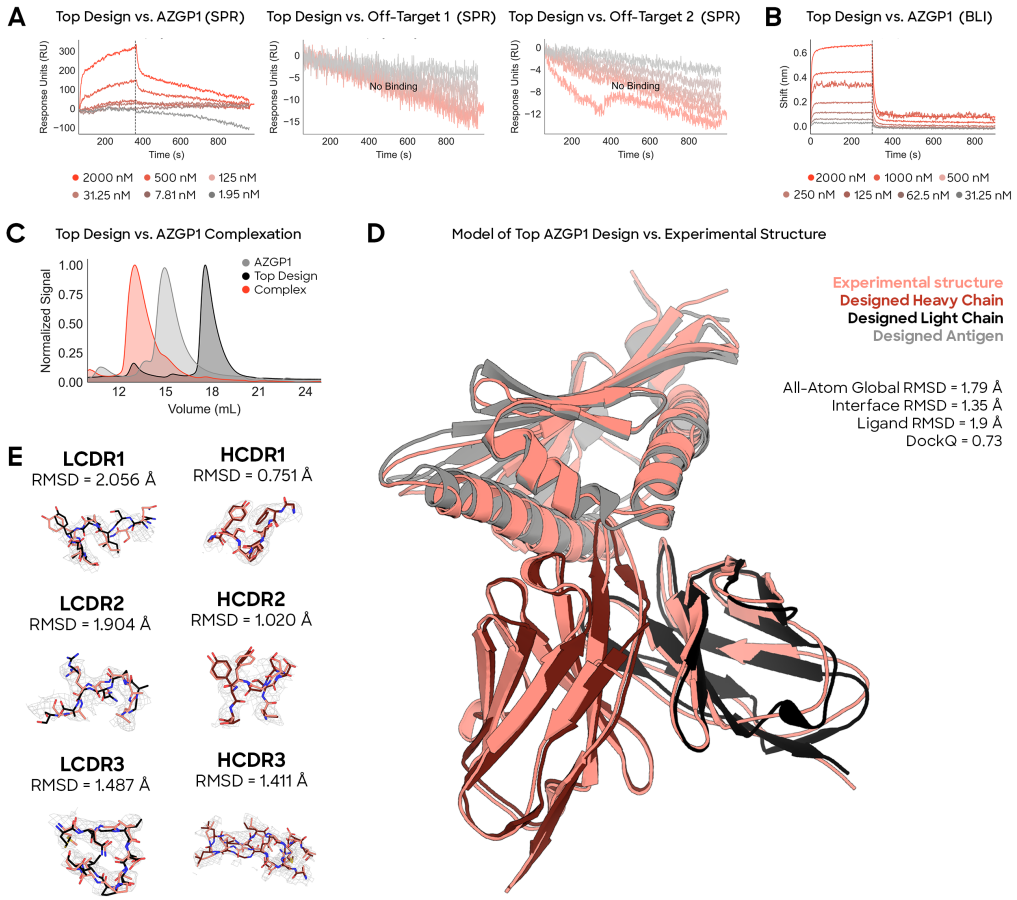


Figure 19: ***In vitro* experimentation confirms that Origin-1 generated a *de novo* binder against a novel epitope on AZGP1.** (A) SPR demonstrates that the top design binds to AZGP1 in mAb format and does not bind to either of two unintended targets ("Off-Target 1" and "Off-Target 2"). (B) BLI confirms that the top design binds to AZGP1 in Fab format. (C) Complexation experiment confirms that the top design, in Fab format, binds to AZGP1 in solution. (D) Cryo-EM of top design complexed with AZGP1 confirms epitope-specificity and atomic accuracy of the Origin-1 computational model. (E) CDR-specific analysis of model vs. solved complex structure confirms atomic accuracy.

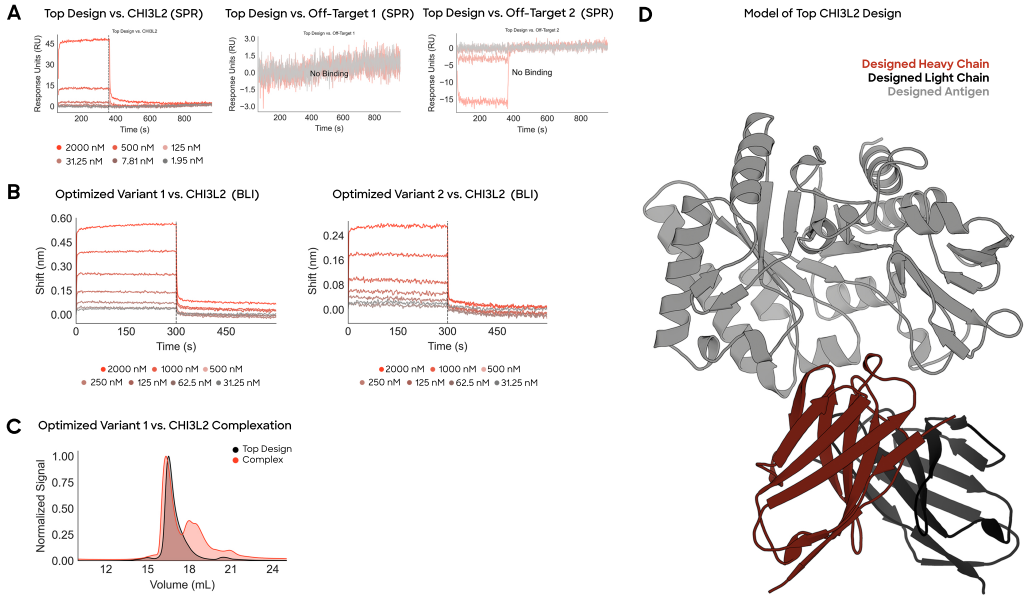


Figure 20: ***In vitro* experimentation confirms that Origin-1 generated a *de novo* binder against a novel epitope on CHI3L2.** (A) SPR demonstrates that the top design binds to CHI3L2 in mAb format and does not bind to either of two unintended targets ("Off-Target 1" and "Off-Target 2"). (B) BLI confirms that the two optimized variants bind to CHI3L2 in Fab format. (C) Complexation experiment confirms that optimized variant 1, in Fab format, binds to CHI3L2 in solution. (D) Origin-1 computational model of top design in complex with CHI3L2 is shown.

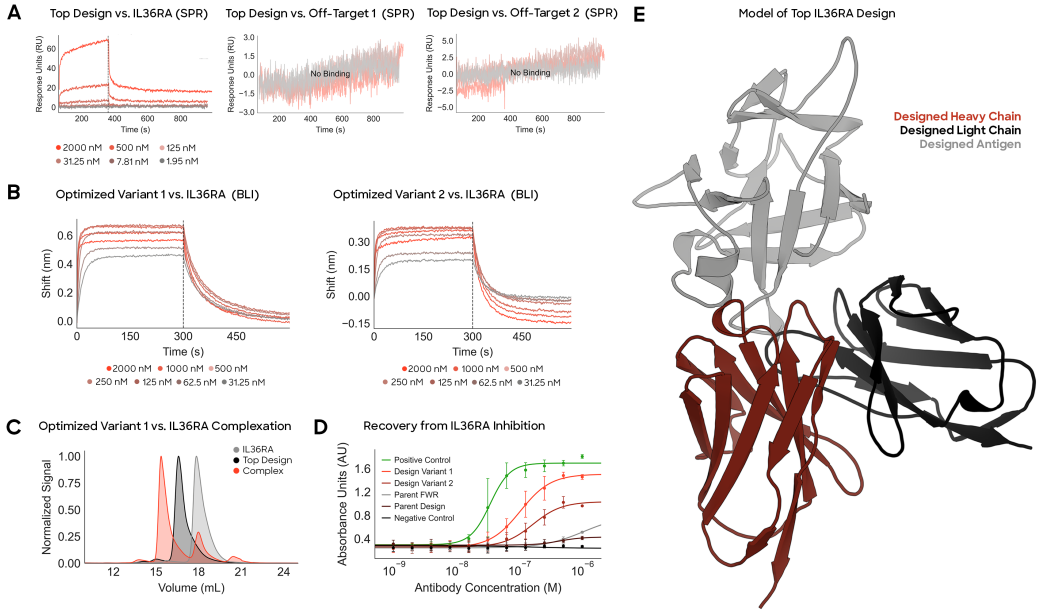


Figure 21: ***In vitro* experimentation confirms that Origin-1 generated a *de novo* binder against a novel epitope on IL36RA.** (A) SPR demonstrates that the top design binds to IL36RA in mAb format and does not bind to either of two unintended targets ("Off-Target 1" and "Off-Target 2"). (B) BLI confirms that the two optimized variants bind to IL36RA. (C) Complexation experiment confirms that optimized variant 1, in Fab format, binds to IL36RA in solution. (D) HEKBlue functional assays demonstrate that the top two optimized variants antagonize IL36RA inhibition. (E) Origin-1 computational model of top design in complex with IL36RA is shown.

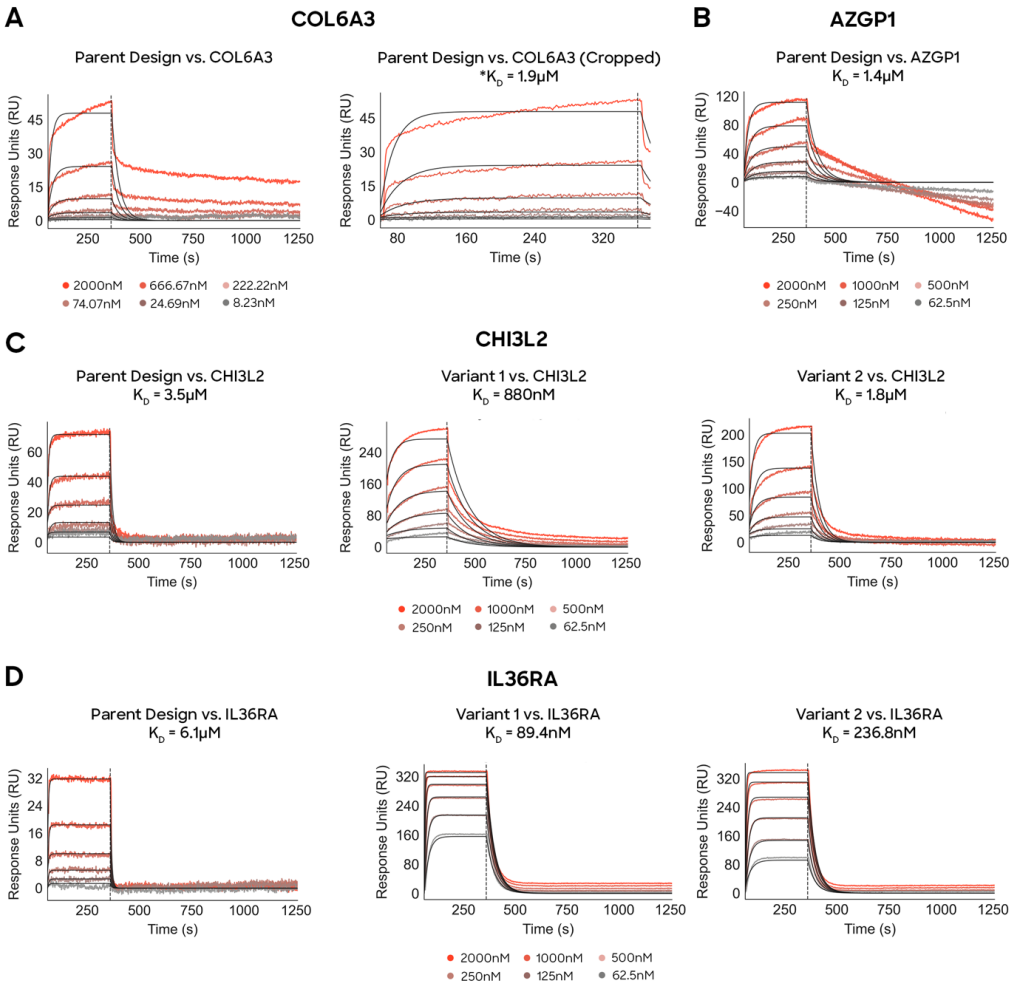


Figure 22: **Per-target SPR sensorgrams including measured binding affinities for Origin-1 parent and optimized variant designs.** (A) *Left*: Parent design against COL6A3 binds to COL6A3. *Right*: Binding affinity is measured from cropped SPR sensorgram reflecting binding of COL6A3 Parent Design against COL6A3. * indicates that binding affinity is computed from cropped sensorgram. (B) Parent design against AZGP1 binds to AZGP1. (C) *Left*: Parent design against CHI3L2 binds with micromolar affinity. One round of AI-based affinity maturation leveraging AbsciBind improves binding affinity by approximately 4X (*Middle*) and 2X (*Right*). (D) *Left*: Parent design against IL36RA binds to IL36RA with micromolar affinity. One round of AI-based affinity maturation leveraging AbsciBind improves binding affinity by approximately 68X (*Middle*) and 26X (*Right*).

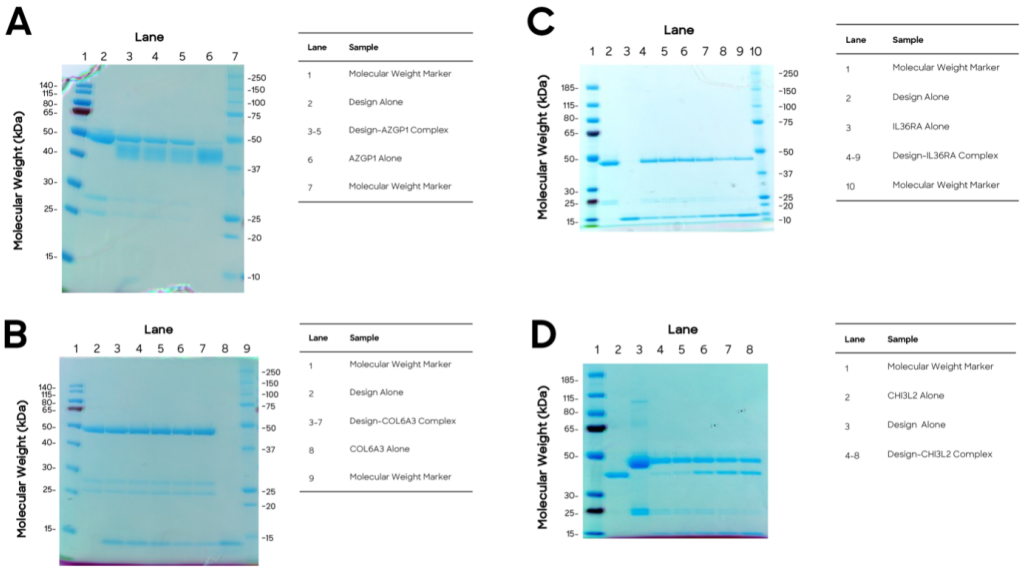


Figure 23: **SDS-PAGE gels for validated antigen-Fab combinations.** (A) SDS-PAGE gel with lanes corresponding to AZGP1 alone, Fab alone, and selected AZGP1-Fab complex fractions from chromatogram seen in Figure 19. (B) SDS-PAGE gel with lanes corresponding to COL6A3 alone, Fab alone, and selected COL6A3-Fab complex fractions from chromatogram seen in Figure 18. (C) SDS-PAGE gel with lanes corresponding to IL36RA alone, Fab alone, and selected IL36RA-Fab complex fractions from chromatogram seen in Figure 21. (D) SDS-PAGE gel with lanes corresponding to CHI3L2 alone, Fab alone, and selected CHI3L2-Fab complex fractions from chromatogram seen in Figure 20.

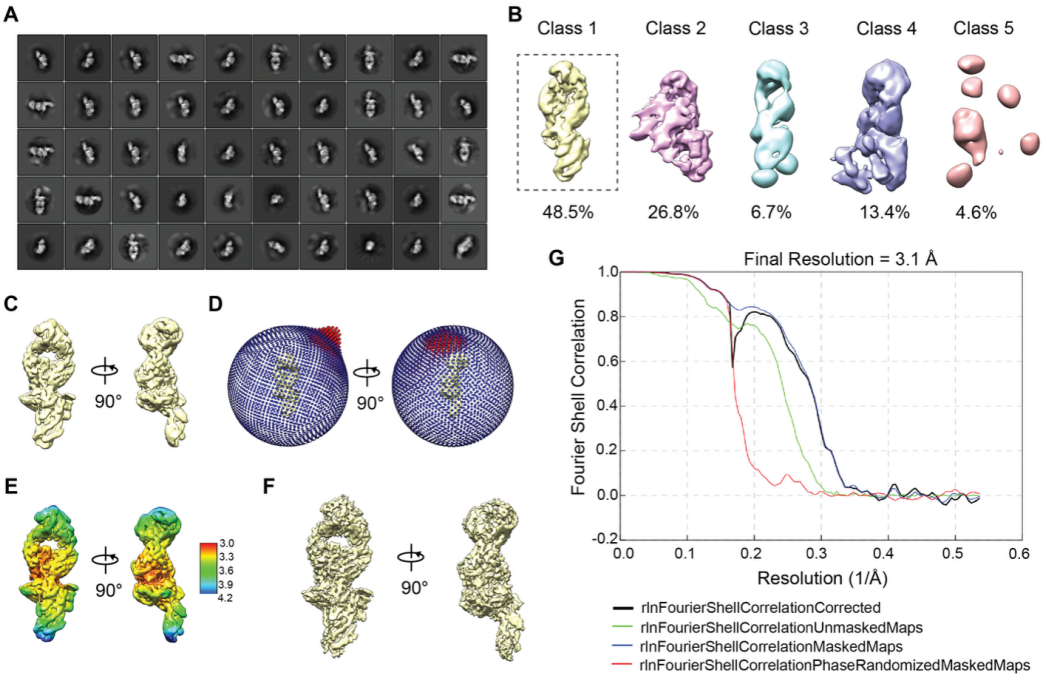


Figure 24: **Cryo-EM processing workflow of AZGP1-Design structure.** (A, B) Representative 2D class averages and 3D classes. (C–F) Global 3D reconstruction, angular distribution of particle orientations, local resolution estimations and sharpened map. (G) Gold-standard FSC plots for the AZGP1-Design complex.

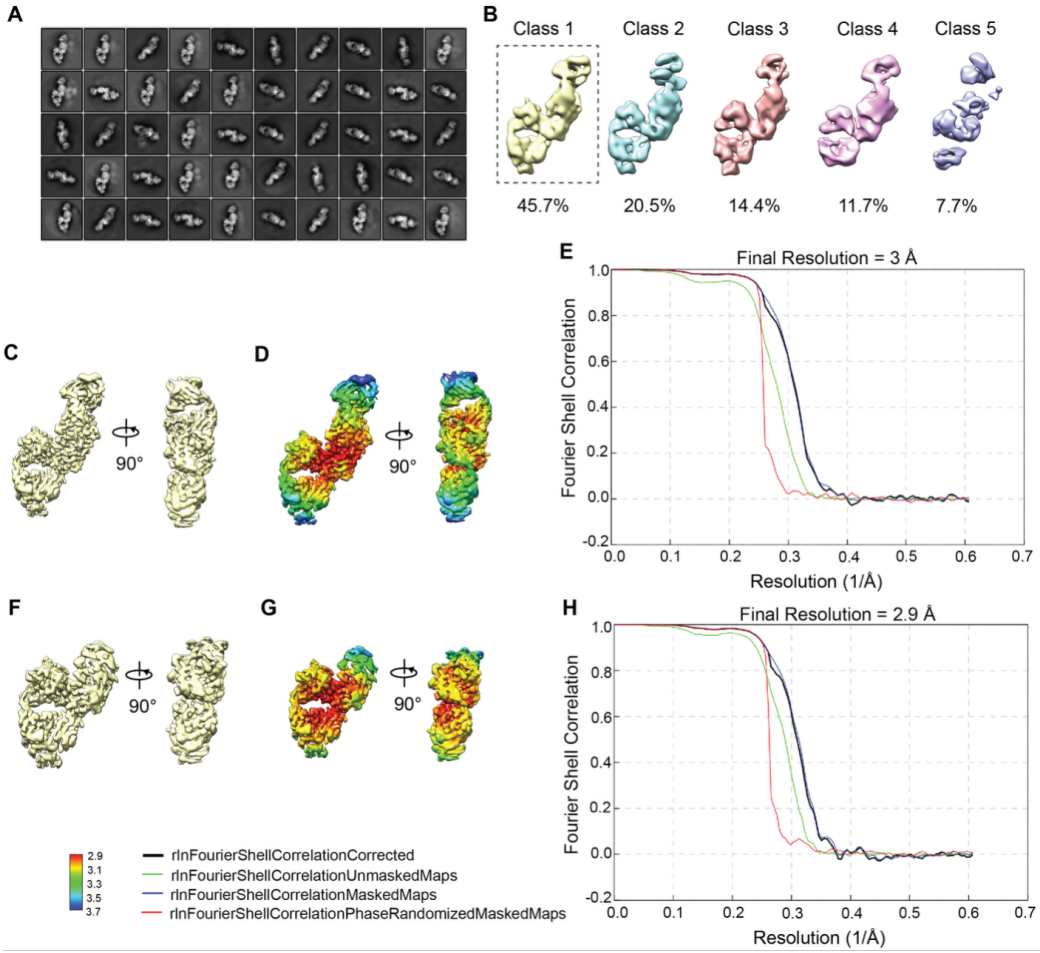


Figure 25: **Cryo-EM processing workflow of COL6A3-Design-Anti Kappa VHH-NabFab structure.** (A, B) Representative 2D class averages and 3D classes. (C–E) 3D reconstruction, local resolution estimations and gold-standard Fourier shell correlation (FSC) plots for the COL6A3 complex global map. (F–H) Locally refined 3D reconstruction map, local resolution estimations and gold-standard FSC plots for the COL6A3-Design region.