Accelerating Antibody Optimization with *AlphaBind*, a Domain-Specific Deep Learning Affinity Prediction Model

Kerry L. McGowan*, Aditya A. Agarwal, James Harrang, David Noble, Adrian W. Lange, Miranda C. Lahman, Jeffrey Adamo, Juliana Barrett, Paxton Reed, David A. Younger, Randolph M. Lopez, Ryan O. Emerson

* Poster presenter

A-Alpha Bio, Inc.
Seattle, WA

Abstract

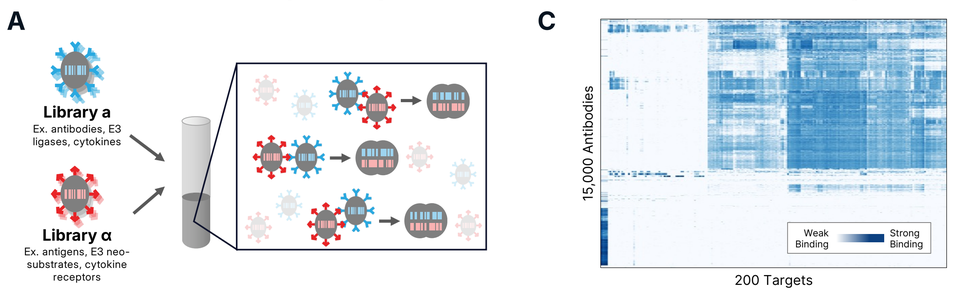
Optimizing antibodies across multiple dimensions, such as affinity, specificity, and developability, is essential for innovating new therapeutics. However, this is often an intensive, manual process that requires prolonged iteration on sequence design and significant *in vitro* testing.

Sequence selection for favorable developability characteristics can be successfully guided *in silico* using *AlphaBind,* a lightweight trained affinity regressor, to traverse a vast antibody sequence landscape while maintaining or improving binding affinity. Our *AlphaBind* model is trained on 7.5 million antibody-antigen binding affinities generated using our yeast display platform *AlphaSeq*. We also used *AlphaBind* to downstream engineer an anti-TIGIT humanoid scFv that was limited by liability motifs in CDRH3 and several non-human germline residues within a mutable window. This was validated *in vitro*, where the new lead showed a 3.4-fold greater affinity for human TIGIT compared to the prior affinity-optimized candidates.

Our results demonstrate that *AlphaBind* and *AlphaSeq* can produce thousands of antibody candidates with superior binding affinity relative to the parental sequence. Our state-of-the-art affinity prediction model can screen candidates at scale *in silico*, generating highly specific leads with improved biodevelopability, despite only selecting for on-target binding.

AlphaSeq Measures Binding Affinities At Scale

Millions of protein-protein interactions measured in one assay

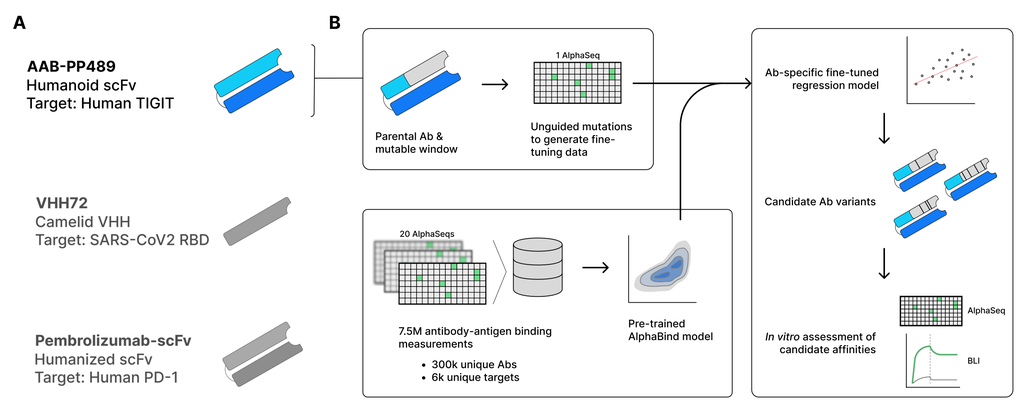


AlphaSeq applies synthetic biology and next-generation sequencing (NGS) to measure protein-protein interactions at a library-on-library scale.

(A) Two libraries containing barcoded protein sequences for display on the yeast cell surface are mixed in liquid culture. Interactions between surface-displayed proteins drive agglutination and cell fusion.
(B) The number of fused cells with a given protein pair depends on protein interaction strength. Counts are derived from NGS sequencing of barcodes.
(C) Example: *AlphaSeq* dataset measuring ~3 million interaction affinities in a single assay.
(D) *AlphaSeq*-measured affinities strongly correlate with BLI-measured affinities.

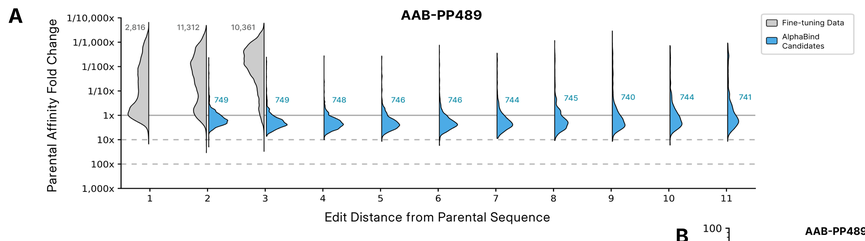
AlphaBind Generates Affinity-Optimized Antibodies

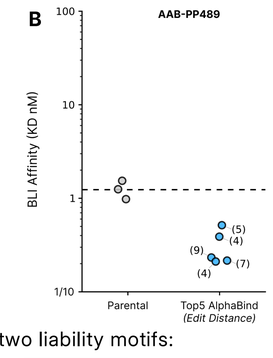
A lightweight affinity regressor to traverse a vast antibody sequence space



(A) A parental lead anti-TIGIT humanoid scFv, AAB-PP489.
(B) *AlphaSeq* was used to generate a training dataset of approximately 30,000 variants within a 99 amino acid mutable window of the parental. The *AlphaBind* regressor, pre-trained on unrelated antibody-antigen affinity data, was fine-tuned with the training dataset to generate an antibody-specific sequence-to-affinity regressor and then used to generate sequences optimized for binding affinity. *AlphaBind* predictions were validated *in vitro* using *AlphaSeq* for large-scale validation and biolayer interferometry (BLI) for a subset of top predictions.

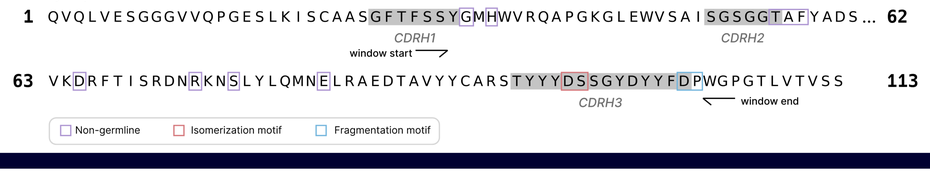
AlphaBind Produces Improved Candidates with Multiple Mutations





(A) The fine-tuned *AlphaBind* model was used to generate 750 candidates at each edit distance: 2 through 11 mutations from parental. *AlphaSeq* was used to prospectively assess the affinity of optimized candidates. Affinity is reported relative to parental binding affinity. Numerical annotations represent the number of candidates tested at each edit distance.
(B) BLI validation results for the top 5 *AlphaBind* predictions for each parental antibody. The geometric mean affinity for codon-shuffled copies of AAB-PP489 (parental antibody) are displayed as a dashed line. Five (100%) of the affinity-optimized candidate antibodies were successfully expressed, and all had improved affinity compared to parental.

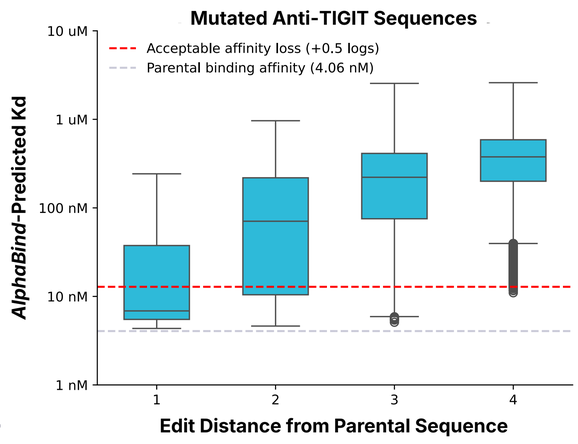
After affinity optimization, AAB-PP3115 (derived from AAB-PP489) contained two liability motifs:



AlphaBind-Guided Developability Engineering Refines Antibodies *In Silico*

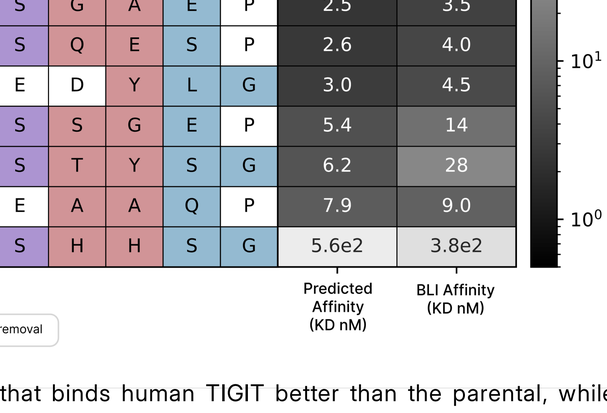
Break liability motifs with minimal affinity loss

- Exhaustively mutated 1-4 residues to remove liability motifs
- Predicted binding affinity using *AlphaBind* regressor for all possible sequences (19^4 = 130,321 predictions)
  - Majority of mutations predicted to ablate binding
- 736 sequences without liabilities (0.56%, or 1 in 177 sequences) were selected to continue
  - Accepted a minimal affinity loss (within 0.5 logs of parental)



Partially revert V and J genes to human germline sequence

- Exhaustively reverted the 736 sequences without liabilities up to 9 non-human germline residues identified with ANARCI[2]
  - 512 (2^9) possible germline variations per sequence
- Predicted binding affinity using *AlphaBind* regressor for all possible sequences (512 × 736 = 376,832)
- 1,974 germlined sequences without liabilities (0.52%, or 1 in 190 sequences) were selected for multi-objective optimization
  - Accepted a minimal affinity loss (within 0.5 logs of parental affinity)
  - Favorable predicted immunogenicity and biodevelopability (NaturalAntibody[3], TAP[4])



AlphaBind Delivers a De-Risked Antibody Lead Optimized Across Multiple Properties

Multi-objective optimization was used to select heavily germlined candidates with the best model-predicted affinities. The top 11 candidates were selected for *in vitro* validation by BLI. 10/11 (91%) candidates expressed, 7/11 (64%) maintained affinity within 2x of parental, and 3/11 (27%) had improved affinity for human TIGIT.

PP3117, the new lead antibody, is a derisked, germlined sequence that binds human TIGIT better than the parental, while maintaining or improving other properties:

	Immunogenicity			Expression	Agglutination	Binding Affinity (IgG)			Melting Temp.
Sequence	Edit Distance†	Germlined Positions†	Liability Motifs†	BLI Titer (μg/mL)	Analytical SEC (% MP)	BLI K_D x human TIGIT (nM)	BLI K_D x mouse TIGIT (nM)	KinExA K_D x human TIGIT (fM)	DSF Tm (°C)
parental	0	0	2	259.1	99.4	3.10	4.73	--	69.51
lead	7	4	0	334.1	98.4	1.97	4.77	309.81	69.38

***AlphaBind*, along with fine-tuning PPI data and an accompanying preprint, has been publicly released under an open-source license: github.com/A-Alpha-Bio/alphabind**

Key Takeaways

- Large scale PPI training data is key for building a fine-tuned antibody affinity oracle
- *AlphaBind* enables efficient exploration of a vast antibody sequence space to optimize affinity and developability properties
- *AlphaSeq* and *AlphaBind* enable rapid engineering of antibodies with femtomolar binding, reduced predicted immunogenicity, zero predicted liabilities, and mouse cross-reactivity

For more information on our platforms, pipeline, and partnerships, visit aalphabio.com or get in touch at contact@aalphabio.com

References

1. Agarwal AA, Harrang J, Noble D, McGowan KL, Lange AW, Engelhart E, Lahman MC, Adamo J, Yu X, Serang O, Minch KJ, Wellman KY, Younger DA, Lopez RM, Emerson RO. 2024. AlphaBind, a domain-specific model to predict and optimize antibody-antigen binding affinity. bioRxiv. https://doi.org/10.1101/2024.11.11.622872
2. Dunbar J, Deane CM. 2015. ANARCI: antigen receptor numbering and receptor classification. Bioinformatics. 32:(2)298-300. https://doi.org/10.1093/bioinformatics/btv552
3. NaturalAntibody. NaturalAntibody Research App, https://abstudio.naturalantibody.com (2024)
4. Raybould, MIJ, Marksa C, Krawczyka K, Taddeseb B, Nowaka J, Lewisc AP, Bujotzekd A, Shie J, Deane CM. 2019. Five computational developability guidelines for therapeutic antibody profiling. Proc Natl Acad Sci USA. 116:(10)4025-4030. https:// 10.1073/pnas.1810576116