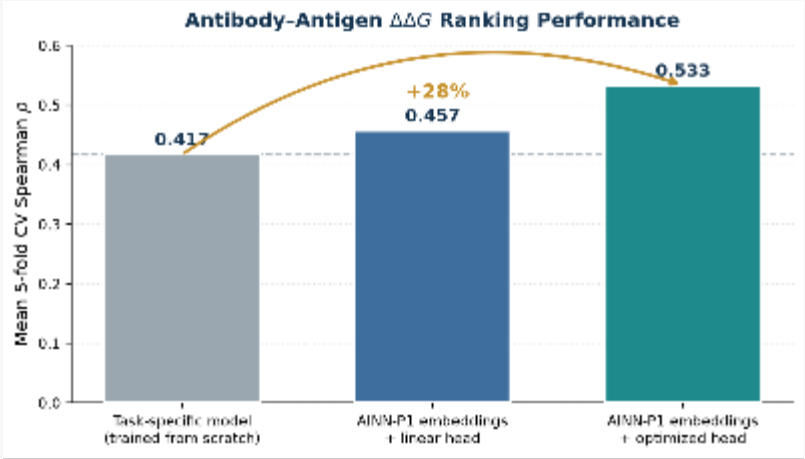
Ainnocence Shows Its Protein Foundation Model Improves Antibody-Antigen Affinity Ranking by 28% - From Sequence Alone

Ainnocence's foundation model trained from scratch, requiring no structure, no multiple sequence alignment, no fine-tuning, training in hours rather than weeks



Mean five-fold cross-validated Spearman ρ for antibody-antigen ΔΔG ranking. All configurations use identical data and folds. Source: (Wang et al., 2026)

SAN FRANCISCO , CA, UNITED STATES, August 3, 2026 /EINPresswire.com/ -- Ainnocence Inc., an AI-driven pharmaceutical & material IP-generating platform company, today released research showing that representations learned by its AINN-P1 protein foundation model substantially improve a core task in antibody engineering: ranking candidate antibody-antigen pairs by binding affinity. Using AINN-P1 as a frozen encoder, the company improved the mean Spearman rank correlation for predicted change in binding free energy (ΔΔG) from 0.42 to 0.53, a relative gain of approximately 28% over a task-specific model trained from scratch on the same data and evaluated under an identical five-fold cross-validation protocol.

> A plain linear probe on our frozen embeddings outperforms a full network trained end-to-end on the same labels. That tells you the biology is already in the representation."
>
> *Dr. Lurong Pan, Founder and CEO of Ainnocence*

The findings are described in a new preprint, "AINN-P1: A Compact Sequence-Only Protein Language Model Achieves Competitive Fitness Prediction on ProteinGym," by Roger Wang, Kevin Jin, and Lurong Pan.

Key findings

28% relative improvement in ΔΔG ranking quality (Spearman ρ: 0.417 → 0.533) under identical five-fold cross-validation.

A simple linear model on frozen AINN-P1 embeddings (ρ = 0.457) already beats a task-specific network trained end-to-end from scratch (ρ = 0.417) direct evidence that the gain comes from representation quality, not from added model capacity.
Sequence-only. No co-crystal structure, docked model, or multiple sequence alignment is required at any stage.
Orders-of-magnitude lower training cost: seconds per fold rather than hours, with the foundation model held frozen.

Target	Hits/Tested	Hit Rate (%)	Best KD
TSLP	13/9	69.2	1pM
IL36R	20/11	55	1pM
GD2 X CD3	44 / 85	51.8	129 pM
IL15	34 / 74	45.9	392 pM
CD19	50 / 124	40.3	28 pM
Activin A X GDF8	46 / 116	39.7	18 pM
TSHR	41 / 121	33.9	N/A
FcRn	33 / 98	33.7	49 pM
C1q	26 / 84	31	540 pM

Representative High-Yield Antibody Discovery Campaigns on SentinusAI®

Why it matters

Antibody affinity maturation requires ranking large panels of candidate variants under tight experimental-label budgets. Building a bespoke predictor for each new target is data-hungry, slow and hard to reproduce across campaigns.

Many of the strongest published affinity and fitness predictors depend on structural input, an experimentally solved or computationally docked three-dimensional model of the complex or on deep multiple sequence alignments. For novel antibody-antigen pairs, which is precisely the regime that matters in discovery, co-crystal structures usually do not exist, and docked models introduce error at the exact interface the prediction depends on.

Ainnocence's result is achieved without either. AINN-P1 encodes antibody and antigen sequences directly, and a lightweight supervised head ranks candidates from the resulting embeddings. Because the foundation model is never fine-tuned, the same embeddings can be reused across targets and objectives, and each downstream head trains in seconds enabling rapid, low-data iteration inside a live discovery campaign.

"The most informative result in this study is the one that looks least impressive on paper," said Dr. Lurong Pan, Founder and CEO of Ainnocence. "A plain linear probe on our frozen embeddings outperforms a full network trained end-to-end on the same labels. That tells you the biology is already in the representation. We are not winning by building a bigger model on top, we are winning because AINN-P1 has learned something real about how proteins bind, from sequence alone."

How it was measured

The team framed affinity maturation as a learning-to-rank problem: because downstream decisions depend on which candidates advance to the wet lab, the relative ordering of candidates matters more than absolute ΔΔG values. Spearman rank correlation served as the primary metric, with NDCG and AUC tracked as secondary measures.

Three configurations were evaluated on a curated antibody-antigen ΔΔG dataset. All shared identical inputs, labels, and cross-validation folds, and feature normalization were fit only on training folds to prevent information leakage isolating the effect of the representation itself.

Mean five-fold cross-validated Spearman ρ for antibody-antigen ΔΔG ranking. All configurations use identical data and folds. Source: (Wang et al., 2026)

“Because we keep the foundation model frozen, these numbers are a floor, not a ceiling,” said Roger Wang, co-author of the study. “We have not yet spent a single gradient step specializing AINN-P1 to antibody-antigen data. That is the next experiment, and we expect the margin to widen.”

What comes next

Because the foundation model is held frozen throughout, the reported accuracy represents a conservative lower bound. Ainnocence plans a task-adaptive fine-tuning stage that specializes AINN-P1 to antibody-antigen affinity data, alongside higher-capacity model variants, multi-objective heads that optimize affinity jointly with developability and specificity and closed-loop integration of experimental feedback. Prospective wet-lab validation and broader benchmarking across additional targets are underway.

The capability is being integrated into SentinusAI®, Ainnocence’s biologics and antibody design platform.

Representative outcomes from recent antibody discovery campaigns run on SentinusAI® are summarized in Table 1, spanning cytokines, receptors, immune checkpoints, and bispecific formats.

Representative High-Yield Antibody Discovery Campaigns on SentinusAI®

Availability

The preprint is available at AINN-P1: A Compact Sequence-Only Protein Language Model Achieves Competitive Fitness Prediction on ProteinGym. The AINN-P1 model and the curated antibody-antigen ΔΔG dataset are proprietary to Ainnocence Inc. aggregate results are reported in full in the preprint. Inquiries regarding data or model access may be directed to the corresponding author.

About Ainnocence, Inc.

Founded in 2021 and headquartered in California, Ainnocence is a next-generation biotechnology company transforming drug discovery and synthetic biology through AI-based, sequence-first engineering. The company’s self-evolving platform evaluates up to 10 billion molecules spanning proteins, antibodies, small molecules, nucleic acids, and chemical formulations within hours to weeks, enabling rapid, multi-objective design across therapeutic, biological, and chemical systems. By reducing R&D timelines and costs while increasing success rates, Ainnocence empowers industry and academic partners to pursue complex biological innovation with greater precision and control.

Lurong Pan, PhD
Ainnocence
+1 205-249-7424
service@ainnocence.com
Visit us on social media:
LinkedIn
YouTube
X
