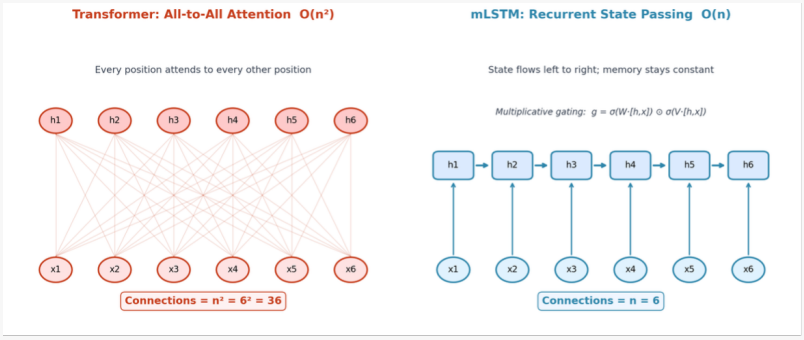
Ainnocence's AINN-P1 Demonstrates Efficient Protein Prediction and Antibody Engineering

How does a 167M-parameter protein foundation model outperform a 650M general-purpose model on the tasks that matter?



Transformer all-to-all attention (left) requires n² connections; mLSTM recurrent state passing (right) requires n steps, with a memory footprint that does not grow with sequence length.

SAN FRANCISCO , CA, UNITED STATES, August 12, 2026 /EINPresswire.com/ -- Ainnocence Inc., an AI-driven drug discovery company, today announced new results for AINN-P1, a purpose-built protein foundation model designed to deliver efficient and transferable predictions for protein engineering and biologics discovery. AINN-P1 takes a sequence-first approach, generating protein representations directly from amino acid sequences without requiring multiple sequence alignments, structure prediction, or external functional annotations. Conventional pipelines rely on homolog search, which is slow and requires large databases, while AlphaFold-class structure prediction still costs minutes per sequence. The encoder uses a multiplicative LSTM rather than a Transformer. Where all-to-all attention requires n² connections and a key-value cache that grows with sequence length, recurrent state passing requires n steps at a memory footprint that does not grow at all linear O(n) time complexity against the Transformer's quadratic O(n²). The model is trained autoregressively, predicting the next amino acid, rather than with the masked language modeling objective used by the ESM family.

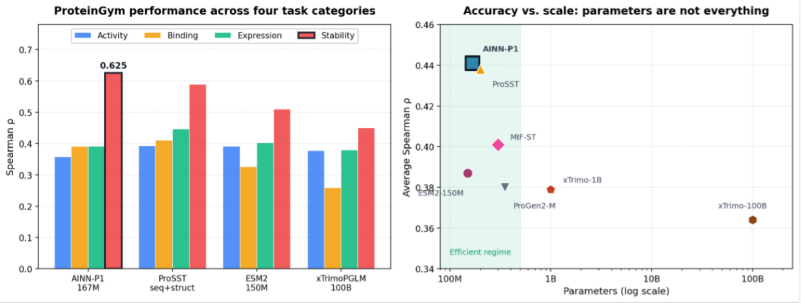
> Whether a model can support decisions on the next program and not just reproduce patterns from the last one, that distinction is critical for drug discovery”
>
> *Dr. Lurong Pan, Founder and CEO of Ainnocence*

Leading Stability Prediction on ProteinGym

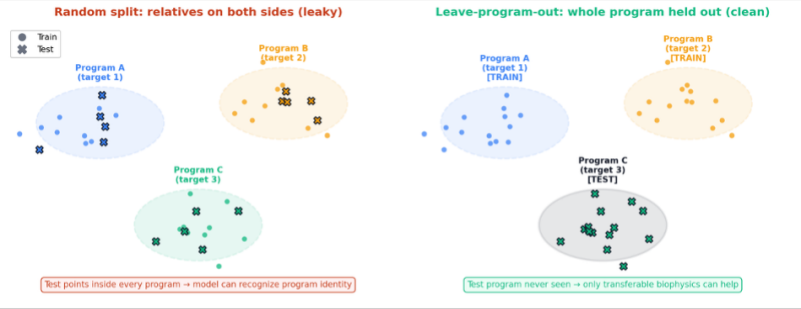
ProteinGym consolidates deep mutational scanning assays across four task categories: activity, binding, expression and stability. AINN-P1 recorded an average Spearman rho of 0.441 and a stability score of 0.625, 6% above the structure-aware ProSST and 39% above the 100B-parameter xTrimoPGLM. Stability is a gating property for biologics developability, since a protein

that will not fold or survive manufacturing cannot proceed regardless of its affinity.

Across activity, expression and stability the gap between AINN-P1 and the structure-aware ProSST is 0.03-0.06, and on binding it is 0.02, indicating that sequence alone encodes more of the relevant constraint than is commonly assumed.



Left: per-category performance, with AINN-P1 leading on stability. Right: accuracy vs. scale the 100B model sits below the 167M model. AINN-P1 values are few-shot frozen-embedding results; baselines are published leaderboard values.



Left: under a random split, every program contains test points, so a model can win by recognizing program identity. Right: under leave-program-out, an entire program is held out and only transferable biophysics can help.

Testing Generalization on New Antibody Programs

The evaluation problem

Antibody discovery data are organized into programs. Each program targets a specific antigen and candidates within it come from the same clonal families, sharing large stretches of framework sequence. Under a random train/test split, close relatives land on both sides. A model can then score well simply by learning that a sequence belongs to a given program and that the program expresses well without capturing any of the biophysics that governs expression.

Ainnocence evaluated VHH single-domain antibody expression under both protocols, using proprietary data from real discovery programs. Under the random split the three encoders are nearly indistinguishable. Under leave-program-out the gap opens immediately.

AINN-P1 exceeded the general-purpose 650M ESM2 by 0.15 AUC on new programs (0.810 vs. 0.660) with 3.9x fewer parameters and matched the finetuned ESM2 within 0.006 AUC without any task-specific finetuning. To rule out the downstream classifier as the driver, seven classifiers were evaluated across 21 encoder-by-classifier configurations; the ESM2-base column was weakest at every classifier position, indicating the effect is a property of the embeddings rather than the head fitted on top of them.

“The most important result for us is not simply that AINN-P1 performs well on a benchmark. It is that the model maintains strong performance when evaluated on antibody programs it has not seen before. That distinction is critical for drug discovery, where the real test is whether a model can support decisions on the next program, not just reproduce patterns from the last one,” said

Dr.Lurong Pan, Founder and CEO of Ainnocence.

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.

Dr. Lurong Pan
Ainnocence Inc.
+1 205-249-7424
lurong.pan@ainnocence.com
Visit us on social media:
LinkedIn
YouTube
X
