



White Paper

Built to Inherit the Curve

How LensAI turns each new generation of AMD EPYC into faster, lower-cost immunogenicity screening, with no code changes

Alignment-free, HYFT®-based humanness scoring on a sharded, stateless pipeline that rides the AMD EPYC performance curve, demonstrated at hpc8a for high-throughput antibody immunogenicity risk assessment.

MindWalk Holdings Corp. (NASDAQ: HYFT)

A Bio-Native AI company

Executive Summary

Immunogenicity is one of the most consequential and most underappreciated risks in biologics development. When a candidate triggers anti-drug antibodies (ADAs), the result can be reduced efficacy, altered pharmacokinetics, or safety findings that delay or terminate a program. Identifying these risks late in development is expensive; identifying them early is transformative.

Moving immunogenicity assessment upstream creates a scale problem. Modern AI-driven discovery can generate thousands of molecules, variants, and design hypotheses before a program reaches formal preclinical testing. A useful screening system therefore needs two capabilities at once: biological context rich enough to make the assessment meaningful, and compute architecture scalable enough to run that assessment across individual molecules, large antibody datasets, and mutation-scan workloads.

That is the role of the MindWalk platform stack described in this paper. HYFT® Technology provides the biological representation method: it turns sequence and structural biology into reusable, function-aware patterns that can be indexed and queried. ReefIQ™ is the biological context layer built from that representation: it organizes and enriches life-sciences data using HYFT-derived patterns and relationships, so candidate molecules can be evaluated against shared biological context rather than against isolated files or one-off models. LensAI™ is the reasoning and application layer that runs on ReefIQ. Its Immunogenicity Screening workflow combines HYFT-based humanness scoring with MHC II binding prediction and population allele weighting to generate a composite immunogenicity-risk score. AMD EPYC™-powered AWS instances provide the compute substrate for that workflow: they accelerate LensAI's sharded, stateless pipeline, but they are not the biology layer.

With that stack in view, LensAI brings immunogenicity screening to a speed and cost point that matches the pace of modern AI-driven discovery. The scalability is architectural rather than brute-force. HYFT patterns are indexed once and reused as biologically grounded subsequence fingerprints, so the humanness scoring step avoids repeated per-query alignment against the full reference corpus. LensAI pairs that indexed context with MHC II binding prediction and population allele weighting in a sharded, stateless workflow that can move from individual molecules to hundreds of thousands of molecules and mutation-scan workloads on the same software stack.

Running this architecture on 5th-generation AMD EPYC processors compounds the advantage. In the benchmark reported here, AMD EPYC C8a and Hpc8a instances delivered a 35–37% speed increase and a 29–43% cost reduction per sequence compared with Intel C8i instances. At full scale, LensAI screened the entire Patent and Literature Antibody Database (PLAbDab) [1] and the Fitness Landscape for Antibodies (Flab) Database [2], which contain 348,550 and 890,634 heavy and light chain sequences respectively, in 2 hours and 12 minutes for the PLaBdab dataset and 4 hours and 41 minutes for Flab on AMD EPYC hpc8a instances using AWS Batch.

This is the MindWalk layer thesis in concrete form. Value compounds in the layer, not in any single component that runs on top of it. On the data axis, ReefIQ becomes more useful as the HYFT-derived representation, curation, and connected evidence grows across programs. On the compute axis,

LensAI is designed to capture the gains from newer high-core-count CPU generations automatically, without an algorithmic rewrite. Hpc8a is the current measured point on that curve.

This white paper describes the characteristics that make LensAI immunogenicity screening scalable by design, the AMD EPYC generational curve that the architecture is designed to exploit, the benchmarks that quantify the gain at the current generation, and the decisions that scalable immunogenicity assessment can unlock for drug discovery teams.

Why Scalable Immunogenicity Screening Matters

Modern discovery generates more candidates than wet labs can test

Modern AI methods inspire qualitatively new drug development pipelines. They produce many potential drug candidates that need to pass a significant number of subsequent screenings and checks before finding their way into the clinic. With wet lab assays still expensive and comparatively slow, the focus has shifted to computational methods to perform these large-scale screenings.

The scale of the problem can become very large for several reasons:

- De novo design pipelines may generate large panels of candidate clones that need early immunogenicity triage before downstream characterization.
- Researchers may want to scan the effects of many single- and multiple-point mutations or other modifications. Even a single antibody heavy chain produces 2,420 variants under a single-point mutation scan, and the count grows exponentially with multi-point mutations.
- Whole antibody-library screens can contain hundreds of thousands of heavy- and light-chain sequences, depending on the dataset snapshot, filtering rules, and deduplication convention. MHC II risk assessment multiplies the workload because each candidate must be evaluated across an allele panel to estimate population-level presentation risk.

Why early ADA risk assessment is critical

Immune responses to therapeutic proteins can derail even the most carefully designed drug candidates. When the immune system identifies a treatment as foreign, it may trigger the production of anti-drug antibodies (ADAs). These responses can alter how a drug is distributed in the body, reduce its therapeutic effect, or create safety concerns that were not apparent during earlier studies. The consequences are often serious delays, added costs, program redesigns, or even full discontinuation.

This is not something to be considered only when a drug is close to clinical testing. It is a risk that needs to be addressed from the beginning. Regulatory agencies increasingly expect sponsors to demonstrate that immunogenicity has been evaluated in early discovery, not as a final check before filing. This shift reflects lessons learned from earlier products that failed late because they had not been properly screened.

Early-stage risk assessment lets developers ask the right questions at the right time. Are there T-cell epitopes likely to trigger immune recognition? Is the candidate similar enough to self-proteins to escape detection? Could minor sequence changes reduce the chances of immunogenicity without compromising function? Identifying epitope clustering or T-cell activation hotspots during discovery enables teams to make targeted modifications in regions such as the variable domain. These adjustments can reduce immunogenicity risk without compromising target binding, helping to streamline development and avoid costly rework later in the process.

Beyond candidate selection, immunogenicity screening improves resource allocation. If a molecule looks risky, there is no need to invest heavily in downstream testing until it has been optimized. Assessing immunogenicity is not just about risk avoidance. It is about building a faster path to clinical success, but only if the assessment can keep up with the volume of candidates generated upstream.

LensAI Immunogenicity Screening: How It Works

LensAI is the reasoning and application layer that runs on ReefIQ, the biological context layer for life sciences powered by HYFT® Technology. Immunogenicity screening is one of the AI Workflows LensAI delivers on top of that enriched context.

Modelling and predicting immunogenic response are a long-standing problem. The biological mechanisms are not fully understood, available quantitative data is limited, and the data that does exist is noisy: ADA incidence, the most common measure of immunogenicity, can vary significantly for the same drug depending on how it was measured.

Despite these challenges, the LensAI Immunogenicity Screening pipeline provides a useful predictive statistical model of ADA incidence and supports ADA incidence risk classification. In a study covering 217 therapeutic antibodies [2], LensAI classified 197 out of 217 (approximately 91%) antibodies as having an ADA incidence risk greater than 30%.

As a biologically faithful model, the LensAI Immunogenicity composite score is built from three features:

- **Peptide binding to MHC II proteins.** This determines whether the candidate drug contains peptides that could, in principle, be presented to T-cell receptors.

- **Distribution of MHC II protein alleles in the population of interest.** Allele frequencies in patients or in a target population are taken into account to quantify how often the immunogenic response is likely to occur.
- **HYFT-based humanness assessment of peptides.** Peptides recognized as “self” by human T-cells (and B-cells) are filtered out, since they are less likely to provoke an immunogenic response.

ARCHITECTURE

LensAI™ Immunogenicity Scoring Pipeline

How composite scores at the sequence, epitope, and amino-acid level are assembled from three independent signal sources.

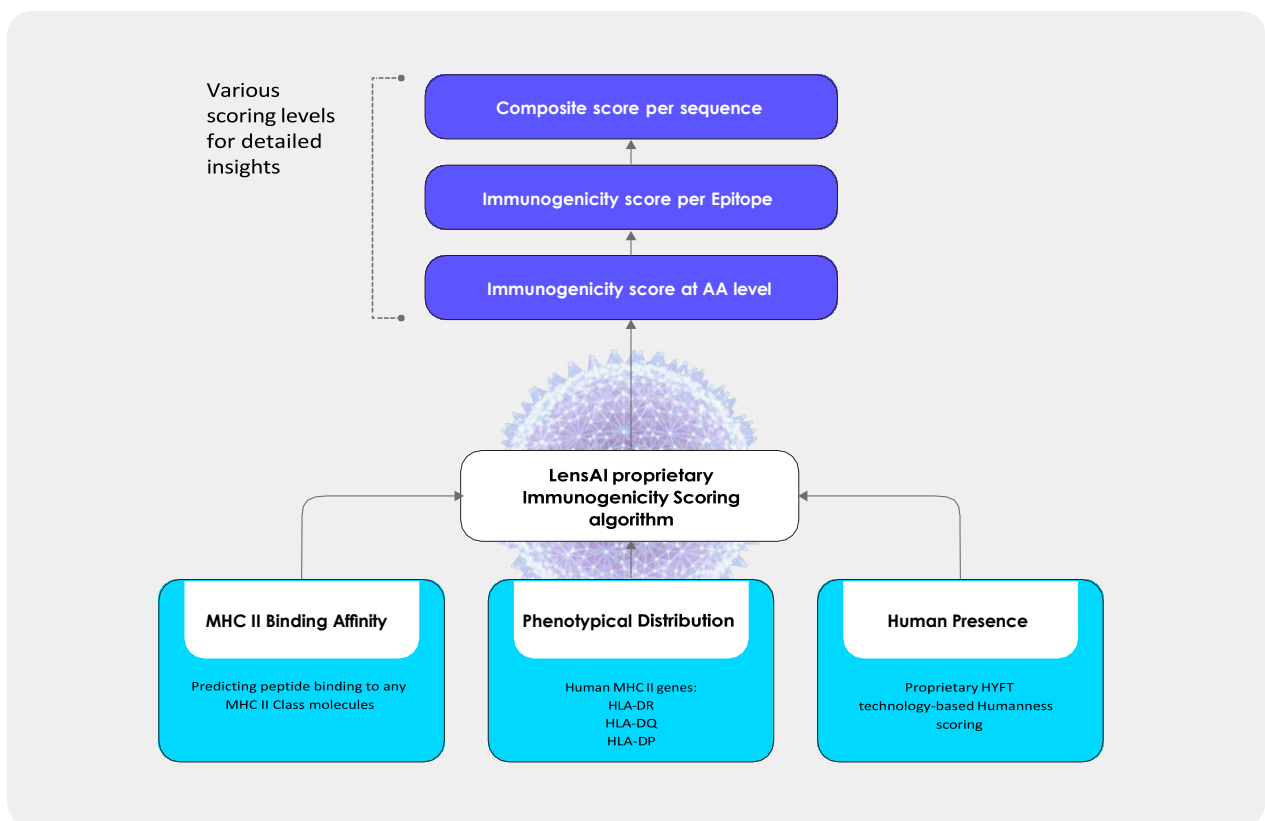


Figure 1. How the LensAI composite immunogenicity score is assembled from MHC II binding, population allele distribution, and HYFT-based humanness assessment.

Scalability by Design: HYFTs and Alignment-Free Humanness Scoring

HYFT patterns: indexed once, reused everywhere

The LensAI immunogenicity pipeline breaks into three independent components: HYFT-based humanness assessment, MHC II binding prediction, and scoring. Its humanness assessment scales not through brute force, per-query alignment but by operating on HYFT patterns: biologically grounded subsequence fingerprints that act as units of meaning across sequence, structure, and function. Because these patterns are indexed once, LensAI runs retrieval and screening over a compact, semantically structured search space rather than the far larger one implied by naive k-mer enumeration or full-alignment approaches. HYFT patterns capture conserved, biologically constrained sequence features and self-similarity relationships, and the LensAI pipeline pairs them with state-of-the-art MHC binding models to produce composite immunogenicity scores.

Alignment-free humanness scoring as an index lookup

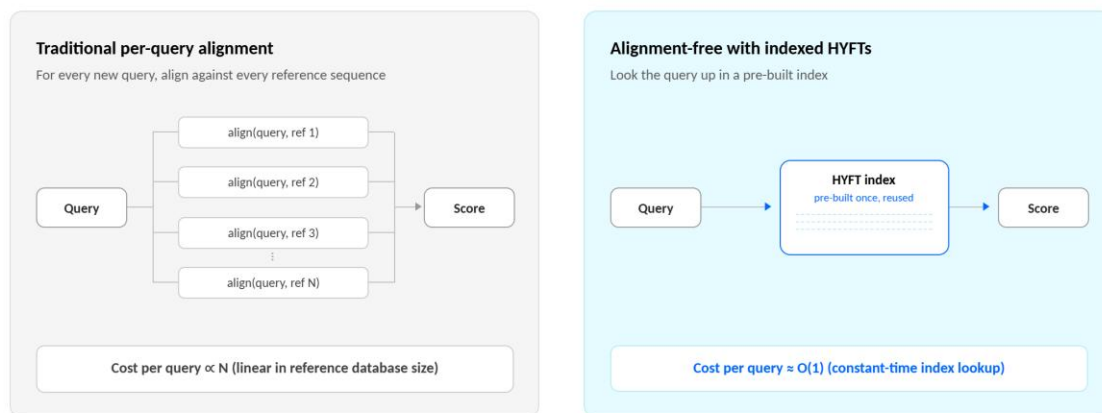
Humanness assessment is the lightest component of the immunogenicity pipeline by design. Because HYFT patterns are pre-indexed, evaluating whether a peptide looks human reduces to an index lookup: a highly optimized operation that does not require per-query alignment to a reference set. This is what makes humanness assessment alignment-free in LensAI, and it is the architectural reason the humanness step finishes in roughly one minute on a single 64-core machine for 6,220 sequences.

The practical consequence is that this step in any large-scale screen is also the cheapest.

SCALABILITY

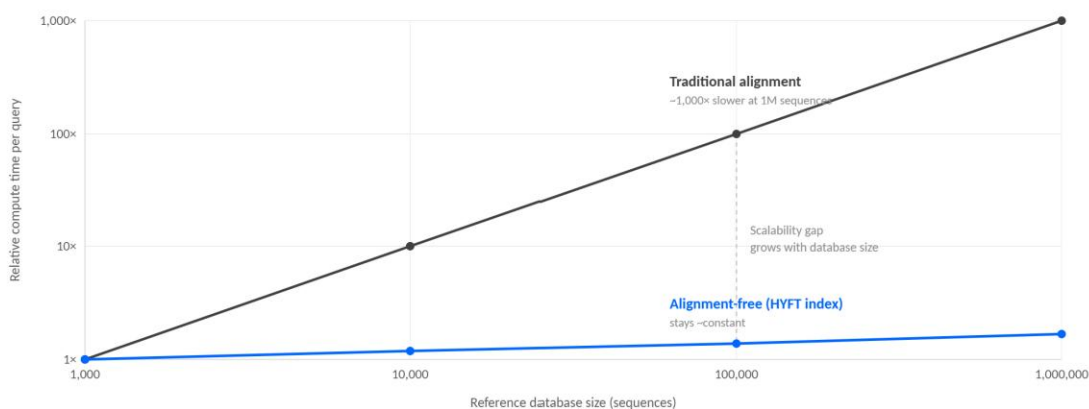
Alignment-free screening scales. Per-query alignment does not.

Why a pre-indexed HYFT layer turns humanness scoring into a constant-time lookup.



Compute cost per query as the reference database grows

relative time, log-log scale; lower is better



Conceptual illustration of asymptotic scaling. In LensAI, HYFT patterns are indexed once, so per-peptide humanness scoring reduces to an index lookup.

Figure 2. Alignment-free, indexed HYFT humanness scoring scales far more favorably than per-query alignment as the number of sequences grows.

A pipeline built for scalability

MHC II binding prediction is the most computationally demanding step, as LensAI screens against numerous MHC II protein alleles, each of which has a separate machine learning model that predicts the binding affinity.

Scalability is achieved because inputs are sharded into chunks that are processed in parallel; chunks share no state during scoring; and the same architecture trivially extends from one machine to dozens.

This combination, an indexed HYFT layer plus a stateless, sharded pipeline, is what allows immunogenicity screening to scale from a single antibody to hundreds of thousands of antibodies on the same software stack. It is also what makes the pipeline hardware-agnostic at the core: because no worker depends on the specific chip beneath it, faster silicon is inherited rather than integrated.

How LensAI Inherits the AMD EPYC Curve

LensAI is built to parallelize efficiently across heterogeneous compute, and high core count CPUs serve as the workhorse for large-scale immunogenicity screening. On AWS, this maps cleanly onto the EC2 C8a and EC2 Hpc8a instance family, powered by 5th Gen AMD EPYC processors. Because the pipeline is indifferent to the specific core it runs on, each new generation of EPYC silicon is inherited by the same code.

Amazon EC2 C8a instances

C8a instances are the latest generation compute-optimized instances powered by 5th Gen AMD EPYC processors, with a maximum frequency of 4.5 GHz. They deliver up to 30% better performance than C7a instances, with even higher gains for some workloads. C8a instances are ideal for high performance, compute-intensive workloads such as batch processing, distributed analytics, high performance computing (HPC), ad serving, highly scalable multiplayer gaming, and video encoding.

Amazon EC2 Hpc8a instances (NEW)

The HPC equivalent of the C8a compute instances, powered by 5th Gen AMD EPYC processors (Turin). Each Hpc8a instance delivers 192 physical cores across two sockets (96 cores per socket), 768 GiB of system memory, and up to 300 Gbps of Elastic Fabric Adapter (EFA) network bandwidth, with clock speeds up to 4.5 GHz and SMT disabled. AWS reports up to 40% higher performance, 42% greater memory bandwidth, and 25% better price-performance compared to Hpc7a.

The pattern to notice is generational. Each step, C7a to C8a and Hpc7a to Hpc8a, delivered double-digit gains on identical code. LensAI is built to consume the next step the same way.

Three properties of C8a and Hpc8a instances matter for this workload:

- **AVX-512, VNNI, and bfloat16 support.** Vectorized and reduced-precision instructions accelerate the MHC II binding prediction models that dominate runtime.
- **Improved memory bandwidth.** C8a instances deliver 33% higher memory bandwidth compared to C7a instances. High-speed access to data in memory matters when the same indexed HYFT structures are touched repeatedly across many parallel workers.
- **1 vCPU = 1 physical core.** Unlike earlier generations, C8a and Hpc8a instances do not use Simultaneous Multi-Threading. Every vCPU corresponds to a physical CPU core, which gives more predictable performance for compute-bound, parallel scoring.

These characteristics translate directly into measurable gains for LensAI immunogenicity screening, with no code changes required. That last phrase is the whole thesis: the improvement is inherited, not engineered per generation.

Where the Curve Sits Today: 5th-Gen AMD EPYC vs Intel

To locate the current position on that curve, we compared LensAI immunogenicity screening on a subset of 18,660 PLAbDab antibody sequences on AMD C8a instances, Hpc8a instances, and Intel C8i instances. All setups used 192-core machines on AWS in the eu-west-1 region for C8i and C8a, and the eu-north-1 region for the Hpc8a instances, with AWS Batch and on-demand instances. There were no code changes between the setups; all settings were kept identical. The runs were repeated five times, and results are averaged across runs. The results are displayed in Table 1.

Instance Type	Pricing per vCPU-h	Total vCPU Hours	Cost per 1,000 Sequences
Intel C8i	\$0.0503/hour	334 vCPU-h	\$0.9*
AMD C8a	\$0.0578/hour	208 vCPU-h	\$0.64*
AMD Hpc8a	\$0.044/hour	217 vCPU-h	\$0.51*

Table 1. Benchmarking results on 18,660 sequences. The averaged results over 5 runs are presented. *Experiments were conducted in July 2026; pricing reflects rates at that time.

On the same workload and the same code, AMD C8a delivered a **35% speed increase** and approximately **29% lower cost per sequence** compared to Intel C8i. For the AMD Hpc8a instances, there is a **37% speed increase** and a **43% lower cost per sequence**. We therefore proceeded to the full Flab screen on AMD Hpc8a instances.

The comparison is a snapshot, not the story. It marks where the inheritable curve sits at the current generation. The durable point is that the same architecture is built to occupy the next point on that curve when it ships, without a rewrite.

Large-Scale Demonstration: Full PLAbDab and Flab Screening

As a representative large-scale example, LensAI immunogenicity screening was performed on the entire Patent and Literature Antibody Database (PLAbDab) [1] and the Fitness Landscape for Antibodies (Flab) Database [2], which contain 348,550 and 890,634 heavy and light chain sequences respectively. Antibodies with a non-commercial license were removed from the Flab database.

PLAbDab [1] is a self-updating repository containing more than approximately 171,050 paired antibody sequences (as of August 2025) derived from patents and academic papers. FLAb is the largest publicly available therapeutic antibody dataset designed to train and benchmark protein AI models [2]. Each dataset is divided into chunks of approximately 18,660 sequences each. The LensAI immunogenicity pipeline processed each of these chunks in parallel on 192-core machines using AWS Batch. All computations were performed on AWS Hpc8a instances. The results are displayed in Table 2.

Dataset	Number of CPUs in parallel	Total vCPU Hours	Elapsed time
PLAbDab (348,550 seq)	1,890 vCPUs (Hpc8a)	4,164 vCPU·h	2 h 12 m
Flab (890,634 seq)	1,890 vCPUs (Hpc8a)	8,856 vCPU·h	4 h 41 m

Table 2. Benchmarking results on the PLaBdab and Flab datasets.

Using the same code deployment, the PLaBdab screening on AWS c7a instances required approximately twice as much time as on the HPC8a instances in a previous benchmark. This example illustrates how AMD EPYC CPU performance amplifies the “scalable by design” nature of HYFT-based screening, and how the same code turns raw generational hardware gains into finished results without modification.

Insights Only Made Possible by Scale

Scalable screening is not just a cost story. It is an insight story.

Large-scale immunogenicity screening also enables exhaustive in silico engineering, such as evaluating all single-amino-acid substitutions across an antibody’s heavy chain, where each input expands to thousands of variants and multi-point scans grow exponentially.

Figure 3 shows the effect of every possible single amino acid substitution on the LensAI immunogenicity composite score for the heavy chain of Adalimumab. Substitutions in the regions 55-65 substantially decrease the immunogenicity score. A composite score below 30 corresponds to low immunogenic risk (predicted ADA below 30%). A composite score above 54 corresponds to high immunogenic risk (predicted ADA above 30%). A single antibody scan generates 2,420 sequences; multi-point scans grow exponentially, illustrating why scalable screening is required for engineering work.

MUTATION SCANNING

Where a single substitution moves the needle

An illustrative single-point mutation scan across a 121-residue heavy chain — the same ~2,420-variant sweep referenced in the text.

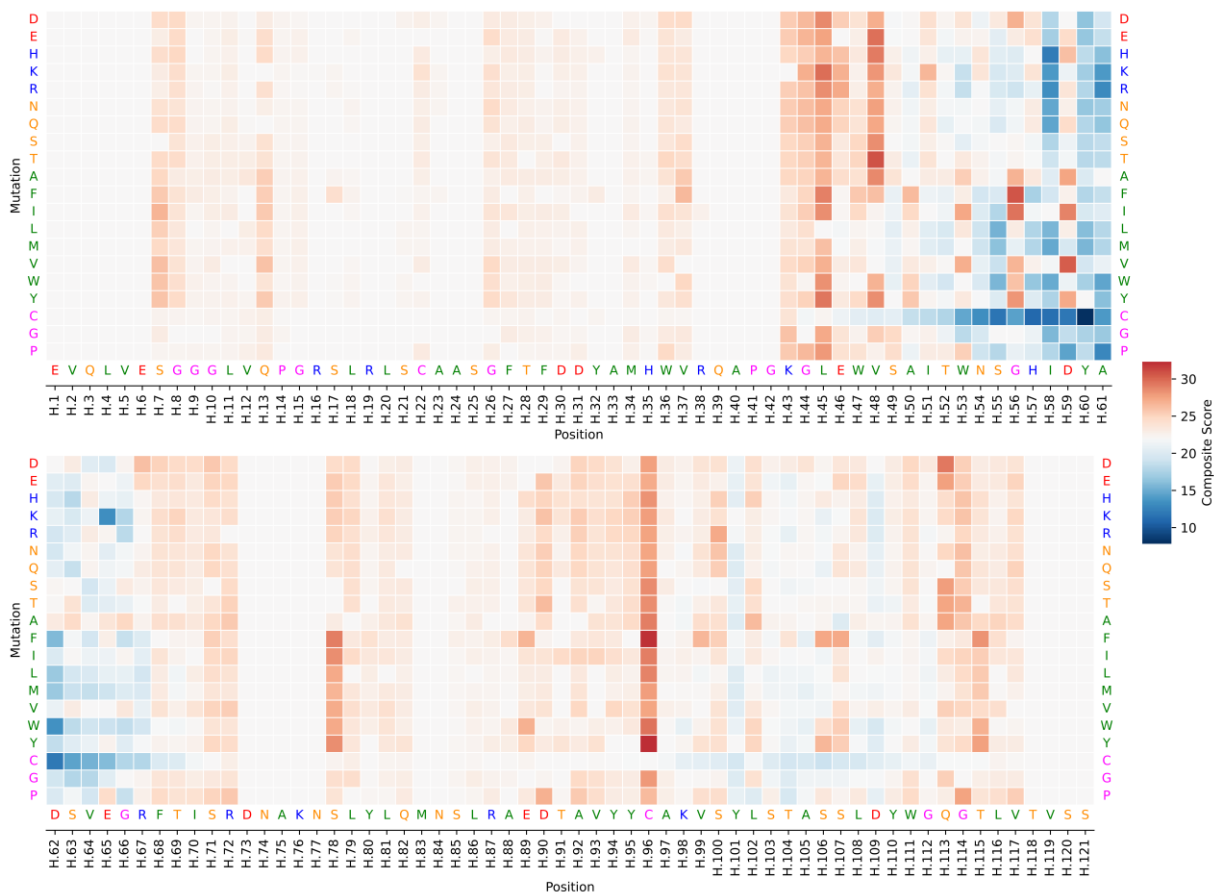


Figure 3. Mutation effect plot for the heavy chain of Adalimumab: every possible single-amino-acid substitution effect on the LensAI immunogenicity score is shown.

What Scalable Immunogenicity Screening Unlocks

Scalable, biologically faithful immunogenicity screening changes how a discovery program operates.

- **Earlier risk detection.** High-risk sequences can be flagged before synthesis, reducing the chance of late-stage failures driven by ADA findings that should have been visible months earlier.
- **Better candidate selection.** Teams can rank and triage thousands of de novo or library-derived candidates by population-level immunogenic risk in hours, not weeks.
- **Practical in silico engineering.** Exhaustive single- and multi-point mutation scans become tractable, enabling epitope removal.
- **Better budget efficiency.** By screening orders of magnitude more candidates within a fixed computational framework, teams can increase the number of viable lead candidates per

dollar spent and reduce wet-lab investment in candidates that were never going to clear immunogenicity gates.

The Compounding Layer: Why This Is the MindWalk Pattern

MindWalk is a Bio-Native AI company that built AI on top of more than 40 years of biology heritage, founded in 1983. MindWalk's thesis is that durable value in life-sciences AI compounds in shared infrastructure, not only in an individual model, dataset, or processor.

That principle appears twice in this paper, on two different axes.

On the data axis, HYFT® Technology enriches biological data at ingestion and connects it into the biological representation, 660 million patterns joined by 25 billion relationships, refined over 20 years of curation. Because the representation is shared infrastructure, every program that runs on it makes every other program more valuable, and that enrichment reaches back to data already on the platform without the customer doing anything. A corpus in ReefIQ is worth more in three years than it is today.

On the compute axis, the same architecture that makes humanness scoring an index lookup is what lets LensAI inherit the hardware curve. Each generation of AMD EPYC silicon makes the identical pipeline faster and cheaper, and LensAI consumes that improvement with no code changes. hpc8a is the current point on that curve; the architecture is built to ride it forward.

Both axes describe one company. MindWalk builds the layer where the gains accumulate: in the biological representation that gets richer with every program, and in a compute architecture built to inherit every generation of silicon. That is why immunogenicity screening at this speed and cost is not a benchmark that ages. It is a position on a curve the platform is built to keep climbing.

Conclusion

Immunogenicity screening at scale is feasible today, and it is feasible on AMD. The LensAI immunogenicity pipeline analyzes more than 890,000 antibody sequences in under four and a half hours on AMD EPYC-powered AWS Hpc8a instances, with no code changes relative to a baseline Intel run, while delivering a **37% speed gain** and approximately **43% lower cost per sequence** at the current generation.

What makes this durable is that none of it is pinned to a single chip. HYFT patterns are indexed once and reused, so humanness assessment is alignment-free and effectively constant-time per peptide. The pipeline is sharded and parallelizable from the first line of design. AMD EPYC's high core count, AVX-512, VNNI, bfloat16, and SMT-free vCPU mapping then translate that architecture into measurably faster, cheaper runs, and each new EPYC generation is inherited by the same code.

The result is not a race won once. It is a platform, and a deepening compute relationship with AMD, built to keep winning as the hardware improves.

By moving immunogenicity assessment to the earliest stages of discovery, where it can have the greatest impact, LensAI on AMD EPYC helps researchers reduce late-stage failures, shorten timelines, lower project costs, and improve the likelihood of clinical success. As biological datasets and design spaces continue to expand, and as AMD continues to advance EPYC, platforms that combine biological fidelity with an architecture built to inherit the compute curve are likely to become foundational for next-generation discovery and engineering workflows.

About MindWalk

MindWalk Holdings Corp. (NASDAQ: HYFT) is a Bio-Native AI company building the BioIntelligence infrastructure that life sciences AI and agentic AI require, integrating AI, data, and advanced wet lab capabilities into one connected discovery ecosystem. At its core is HYFT® Technology, a proprietary, function-aware representation of biology. Its HYFT fingerprints span sequence and structural biology and, refined over 20 years of curation, form a continuously evolving biological representation of 660 million patterns and 25 billion relationships. This enriched biological representation is the architecture behind ReefIQ™, the biological data substrate that provides context for life sciences, enriching data at ingestion and growing more valuable with each program run on it, and LensAI™, the reasoning and application layer for target discovery, candidate diligence, portfolio decision support, and the agentic AI workflows pharma is racing to deploy. By design, value compounds in this HYFT representation layer, not in any individual AI model that runs on top of it.

References

- [1] Abanades B., et al. “The Patent and Literature Antibody Database (PLAbDab): an evolving reference set of functionally diverse, literature-annotated antibody sequences and structures.” Nucleic Acids Research, Volume 52, Issue D1, 5 January 2024, Pages D545-D551.
- [2] Chungyoun, Michael, Jeffrey Ruffolo, and Jeffrey Gray. “FLAb: Benchmarking deep learning methods for antibody fitness prediction.” BioRxiv (2024): 2024-01.
- [3] Marks C., et al. “Humanization of antibodies using a machine learning approach on large-scale repertoire data.” Bioinformatics, 2021 Nov 18;37(22):4041-4047.

Trademarks and notices

HYFT® is a registered trademark of MindWalk Holdings Corp. LensAI™ and ReefIQ™ are trademarks of MindWalk Holdings Corp. or its subsidiaries; ReefIQ™ registration is pending. All other trademarks are the property of their respective owners.

AMD, the AMD Arrow logo, EPYC, and combinations thereof are trademarks of Advanced Micro Devices, Inc. Amazon Web Services, AWS, Amazon EC2, and related marks are trademarks of Amazon.com, Inc. or its affiliates. Intel is a trademark



of Intel Corporation. References to AMD, AWS, Intel, and other third-party products describe the compute environment used to run these benchmarks and do not imply any partnership, sponsorship, or endorsement.

For more information, contact info@mindwalkAI.com | mindwalkAI.com

©2026 MindWalk Holdings Corp. All rights reserved.