IntuitionLabs

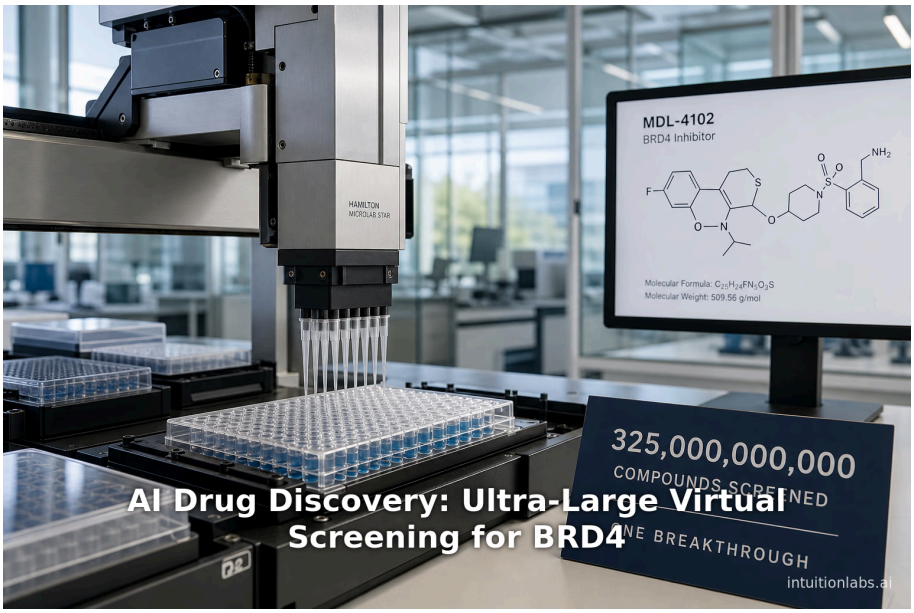
INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

AI Drug Discovery: Ultra-Large Virtual Screening for BRD4

> Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-05-01 · ai drug discovery · ultra-large virtual screening · brd4 inhibitors · bet proteins · computational chemistry · high-throughput screening · oncology therapeutics

Original article: https://intuitionlabs.ai/articles/ai-ultra-large-virtual-screening-brd4



In this article

1. Executive Summary
2. Introduction and Background
3. The Drug Discovery Challenge
4. Ultra-Large Virtual Screening (ULVS)
5. BRD4 and Oncology Targets
6. Advances in Virtual Screening and AI-Driven Discovery
7. Virtual Screening: from Millions to Billions
8. The Chemical Space and Screening Scale
9. Artificial Intelligence and ML in Drug Design
10. The 325-Billion Molecule Screen and MDL-4102 Discovery
11. Campaign Overview
12. MDL-4102: Properties and Novelty
13. Comparison to Known BRD4 Inhibitors
14. Data Analysis and Evidence
15. Case Study: The MDL-001 Antiviral Program
16. Discussion: Implications, Critiques, and Future Directions
17. Transforming Hit Discovery
18. Limitations and Critical Perspectives
19. Future Directions
20. Conclusion
21. References

Executive Summary

Drug discovery continues to struggle against the vastness of chemical space and the high failure rate of traditional pipelines. **Model Medicines**, an AI-driven biotechnology company, has aggressively tackled these challenges by deploying **ultra-large virtual screening (ULVS)**. In late 2025, using its *GALILEO*™ AI platform on Google Cloud, Model Medicines executed a virtual screening of **325 billion** candidate molecules (the company's October 2025 page cites a throughput of 325 billion compounds per day without giving a run time; its May 2026 release describes the screen as run "in a single 24-hour period," a company claim that is not independently verified) (www.biospace.com) (modelmedicines.com). According to the company, this throughput yielded **MDL-4102**, which it describes as a "first-in-class, selective BRD4 inhibitor with no observable BRD2/BRD3 activity"; no peer-reviewed data are available (modelmedicines.com). The company also reports chemical novelty versus clinical BET inhibitors based on ECFP4 Tanimoto similarity, without publishing the data (modelmedicines.com). Together, these results suggest that giga-scale AI-driven searches of chemical space can uncover entirely new classes of drug candidates once deemed unattainable.

This report provides an in-depth analysis of Model Medicines' BRD4 oncology effort. It contextualizes the MDL-4102 project within the history and current state of **ultra-large virtual screening**, details Model Medicines' methodologies and platform, and examines the biological target **BRD4** and its role in cancer therapeutic

development. The company's claims about the 325-billion-molecule screen are summarised (no screening data have been published), highlighting how the company says the GALILEO platform translated computational scale into a candidate molecule with novel chemistry. We compare MDL-4102 to existing BET inhibitors (Table 1), discuss Model Medicines' broader AI capabilities (e.g. MDL-001 antiviral program), and critically assess the potential and limitations of AI-driven drug discovery (modelmedicines.com) (www.scientific-computing.com). Future directions – including trillion-scale screening, multiparameter optimization with AI agents, and the eventual path toward clinical trials – are also explored.

Introduction and Background

The Drug Discovery Challenge

Traditional drug discovery is laborious and expensive, often requiring a decade or more and billions of dollars to bring a single new medicine to market (medicalxpress.com). In practice, nearly 90% of drug candidates fail during clinical development (medicalxpress.com). A persistent bottleneck is **lead identification** – finding small molecules that bind the therapeutic target with high potency and selectivity, while also possessing drug-like properties. Historically, lead discovery has often relied on **high-throughput screening (HTS)** of physical compounds. Yet even large-scale HTS facilities can test only millions of compounds per day (www.biospace.com), a vanishingly small fraction of possible molecules. Approximately **10^60** "drug-like" small molecules are believed to exist in principle (www.biospace.com) (pmc.ncbi.nlm.nih.gov). The company describes the scale screened to date, relative to the space of possible molecules, as "less than a single drop of water in all the world's oceans" (www.biospace.com).

This vast **chemical space** implies that most promising chemotypes remain undiscovered. Virtual screening (VS) – computationally evaluating large libraries of structures against a target – offers a way to explore more of this space at lower cost. Early docking and similarity-based approaches sifted millions of candidates in silico (medicalxpress.com), but their throughput has historically lagged far behind the theoretical size of chemical space. Recent advances in compute power, library enumeration, and especially **AI/ML** have begun to shift this balance. The key insight is that screening more compounds can yield **better** leads: as Harvard Medical School's Gorgulla et al. note, "the more compounds you can screen, the better your top candidates will be, and the lower your rate of false positives" (medicalxpress.com).

Ultra-Large Virtual Screening (ULVS)

The concept of **ultralarge virtual screening (ULVS)** has emerged as computing has scaled. Research teams have sought to expand the number of docked or predicted molecules from millions to billions or trillions (www.biospace.com) (www.sciencedirect.com). Early pioneering work includes **VirtualFlow** (2020), an open-source platform from Harvard researchers, which prepared a ready-to-screen library of **1.4 billion** compounds (medicalxpress.com). VirtualFlow was designed to run on large computer clusters or cloud grids, demonstrating that even academic groups could screen well beyond previous limits. For example, a 10,000-core cluster could dock 1 billion compounds in about two weeks (medicalxpress.com).

Other approaches include physics-based methods and advanced ML models. Physics-based free-energy perturbation (FEP) methods, such as Schrodinger's *FEP+*, can estimate binding energetics. Machine-learning co-folding and affinity models such as *Boltz-2* are reported to operate at about two million compounds per day on BioHive 2 (modelmedicines.com). By contrast, **predictive ML models** – neural networks or parametric scoring functions trained on bioactivity data – can achieve much higher raw throughput. For instance, Atomwise's AtomNet deep-learning model has reportedly reached **8 billion molecules per day** on its latest systems (modelmedicines.com).Yet prior to 2025, all such efforts remained below the **hundreds of billions** barrier (modelmedicines.com) (www.sciencedirect.com). The company's own authors state: "To our knowledge, however, no prior predictive screen has surpassed the hundred-billion scale" (modelmedicines.com).

Model Medicines' campaign therefore represents a watershed moment. In partnership with Google Cloud, the company reports a throughput of **325 billion molecules** per day (www.biospace.com) (modelmedicines.com) – roughly 40 times the 8 billion per day it cites for Atomwise's AtomNet. This ultra-large scale promises to find entirely new chemotypes ("a fundamentally broader exploration of chemical space" (modelmedicines.com)) that smaller screens cannot reach. Nevertheless, ULVS faces challenges such as scoring accuracy, computational cost, and data bias (www.sciencedirect.com). Combining ML with structure-based docking or hybrid methods may help mitigate these issues (www.sciencedirect.com). Ultimately, Model Medicines shows how pushing throughput can be a "precision amplifier" in discovery (www.biospace.com).

BRD4 and Oncology Targets

Bromodomain-containing protein 4 (**BRD4**) is a transcriptional regulator in the BET (bromodomain and extra-terminal) family (pmc.ncbi.nlm.nih.gov). It contains two conserved N-terminal bromodomains (BD1 and BD2) that bind acetylated lysines on histones and transcription factors (pmc.ncbi.nlm.nih.gov). By recruiting mediators of transcription elongation (e.g., P-TEFb complex), BRD4 helps control expression of critical genes, including oncogenes like *c-MYC* (pmc.ncbi.nlm.nih.gov). Abnormal BRD4 activity has been implicated in various cancers, inflammation, and viral processes (pmc.ncbi.nlm.nih.gov). Notably, the fusion of BRD4 with the **NUT** protein drives aggressive NUT-midline carcinoma (pmc.ncbi.nlm.nih.gov), and BRD4 overactivity has roles in leukemia, lymphoma, and other tumors (pmc.ncbi.nlm.nih.gov).

Because it sits at a regulatory nexus, BRD4 has been an attractive, albeit challenging, drug target. Dozens of small-molecule **BET inhibitors** have been developed and tested in trials (pmc.ncbi.nlm.nih.gov). Examples include JQ1 (research tool), OTX015 (Merck MK-8628), ABBV-075 and ABBV-744 (two distinct compounds). These compounds generally bind the acetyl-lysine pocket and typically inhibit both BD1 and BD2 domains of BET proteins (pmc.ncbi.nlm.nih.gov). However, because BRD2 and BRD3 are highly homologous to BRD4 (pmc.ncbi.nlm.nih.gov), most inhibitors block multiple BET family members.

As of 2025, several BET inhibitors have advanced in clinical trials for various hematologic and solid tumors (pmc.ncbi.nlm.nih.gov). Table 1 lists selected BET inhibitors that have been tested in clinical trials. Selective inhibition of BRD4 alone (sparing BRD2/3) has long been viewed as a "holy grail," since it might retain anti-tumor activity while reducing off-target toxicity (pmc.ncbi.nlm.nih.gov) (modelmedicines.com).

Table 1. Selected BET inhibitors that have been tested clinically. The list is illustrative and does not indicate current development status.

Compound (Developer)	Indications studied
Molibresib (GSK525762, GSK)	NUT midline carcinoma, AML*, other cancers
BAY 1238097 (Bayer)	Solid tumors
INCB054329 (Incyte)	Solid tumors, lymphoma
Pelabresib (CPI-0610, Constellation)	Myelofibrosis
OTX015 (MK-8628, Merck)	AML, lymphoma, myeloma

*AML = acute myeloid leukemia.

A **BRD4-specific inhibitor** with novel chemotypes could provide a breakthrough. This is the promise of MDL-4102, discovered through a massive AI-enabled screen.

Advances in Virtual Screening and AI-Driven Discovery

Virtual Screening: from Millions to Billions

Early computational screening typically handled libraries of millions to tens of millions of compounds (medicalxpress.com). Even by 2020, state-of-the-art had advanced: “the more compounds you can screen, the better your top candidates will be” (medicalxpress.com). The VirtualFlow platform demonstrated practical access to ultra-large libraries, integrating a prepared set of 1.4 billion commercial-like molecules (medicalxpress.com). As Haribabu Arthanari (Harvard Medical School) said, VirtualFlow “truly democratizes ultra-large-scale screening”; the same report notes that a 10,000-core cluster, as is available at HMS, could screen 1 billion compounds in two weeks (medicalxpress.com).

More recently, hybrid workflows have emerged. Structure-based docking remains “the most validated approach for prospective ULVS” according to a 2026 review (www.sciencedirect.com). Yet pure docking is compute-intensive. Consequently, many efforts now **pre-filter** or guide docking with machine learning. For example, “deep docking” strategies rapidly score trillions of molecules with lightweight models, then dock top candidates. Other approaches train predictive ML models (e.g. graph neural nets) directly on bioactivity data to rank candidates.

Critically, combining ML with docking can **accelerate** and enrich ULVS pipelines (www.sciencedirect.com). ML surrogates can rapidly eliminate low-priority scaffolds, focusing expensive docking on the most promising leads. Recent studies even explore reinforcement learning to generate novel libraries and fold them back into the screening process. Such advances suggest ULVS is evolving into multi-stage workflows: generative models expand libraries, ML filters and ranks, and docking verifies top poses (www.sciencedirect.com) (medicalxpress.com).

The Chemical Space and Screening Scale

The sheer potential chemical space is astronomically large (10^60 “drug-like” compounds (www.biospace.com)). By contrast, even screening **billions** of molecules is only “a drop in the ocean” (on the order of 10^-50 of total space (www.biospace.com)). Nonetheless, exploring larger fractions of accessible space has demonstrable benefits. Hits from ULVS campaigns tend to be more *novel* and potent. For instance, Gorgulla et al.’s VirtualFlow team, an international team led by Harvard Medical School researchers, prepared a database of 1.4 billion chemical compounds for screening (medicalxpress.com). Large-scale libraries allow the application of tighter filters (e.g. on novelty, patentability, ADMET predictions) without eliminating too many candidates (modelmedicines.com). Model Medicines emphasizes that screening at larger scale yields “higher novelty, improved hit rates, and dramatically reduced attrition” (modelmedicines.com).

To date, most ULVS efforts have topped out in the low billions. For example, Atomwise claimed ~8×10^9 molecules per day screening throughput with its AtomNet model (modelmedicines.com), and Gorgulla’s VirtualFlow demonstrated rapid screening of ~10^9 on academic clusters (medicalxpress.com). Model Medicines’ 325-billion-molecule campaign (Table 2) is, by the company’s account, larger than these earlier efforts. By distributing inference massively on CPU instances, it extends the chemical search far past prior limits (modelmedicines.com).

<table> <caption>Table 2. Comparison of Large-Scale Virtual Screening Initiatives</caption> <thead> <tr><th>Initiative / Organization</th><th>Year</th><th>Screening Scale (molecules/day)</th> <th>Method</th><th>Achievement</th></tr> </thead> <tbody> <tr> <td>VirtualFlow (Gorgulla et al., HMS) </td> <td>2020</td> <td>n/a (1.4×10^9-compound library prepared)</td> <td>Docking pipeline (AutoDock Vina, etc.)</td> <td>Prepared a database of 1.4 billion compounds; “truly democratizes ultra-large-scale screening” (medicalxpress.com)</td> </tr> <tr> <td>AtomNet (Atomwise)</td> <td>n/a</td> <td>~8×10^9</td> <td>Deep learning model (parametric scoring)</td> <td>Reported screening ≈8 billion molecules/day (modelmedicines.com)</td> </tr> <tr> <td>GALILEO (Model Medicines)</td> <td>2025</td> <td>3.25×10^11</td> <td>Lightweight ML model (CPU inference on Google Cloud)</td> <td>Achieved 325 billion compound throughput/day; discovered MDL-4102 (modelmedicines.com)</td> </tr> </tbody> </table>

Sources: Model Medicines press releases (modelmedicines.com); Harvard Medical School (VirtualFlow) (medicalxpress.com); Model Medicines (Atomwise reference) (modelmedicines.com).

Artificial Intelligence and ML in Drug Design

Alongside library scale, **AI and machine learning** are revolutionizing medicinal chemistry. Beyond screening, AI can contribute to target identification, de novo molecule generation, ADMET prediction, and optimization. Model Medicines typifies this paradigm: their **GALILEO™ platform** integrates multiple AI modalities. It employs generative models to propose novel compounds, predictive models to forecast bioactivity and ADMET, and agent-based planners to coordinate complex objectives (www.biospace.com) (modelmedicines.com).

A key insight from recent analyses is that **quality over quantity** can aid generalization. Many AI drug teams have shown that training on excessively large, redundant datasets (“the more training data is better” fallacy) can disadvantage novel chemotypes (modelmedicines.com). Instead, Model Medicines emphasizes training on

diverse, out-of-distribution subsets to preserve extrapolation capabilities (modelmedicines.com). In practice, GALILEO is trained on carefully partitioned chemical spaces (guided by t-SNE clustering) so that the model learns to predict in unseen regions (modelmedicines.com). This tailoring makes the model more “creative” – able to rank molecules outside of common scaffolds at high screen throughput (modelmedicines.com).

For inference, the company deploys **lightweight neural networks** on CPUs to achieve high parallelism (modelmedicines.com). By training on GPUs but running on large CPU clusters (Google Kubernetes Engine on AMD EPYC machines), GALILEO attains unprecedented scale without prohibitive GPU needs (modelmedicines.com). This “train on GPUs, infer on CPUs” strategy ensures cost-effective throughput (modelmedicines.com). It also facilitates rapid iteration and containerized reproducibility, as all steps (featurization, inference) can be scaled independently on the cloud (modelmedicines.com).

With this AI-driven workflow, Model Medicines’ vertical integration spans the entire discovery funnel – from *target selection* to *virtual screening* to *lead optimization*. Separately, the company’s materials cite AmesNet, a deep-learning model for Ames mutagenicity prediction, as a platform capability; they do not state that it was applied to MDL-4102 (modelmedicines.com). By doing so, they aim to optimize multiple criteria (potency, selectivity, ADMET) in parallel – an inherently multi-parameter problem in drug design (www.biospace.com).

The 325-Billion Molecule Screen and MDL-4102 Discovery

Campaign Overview

In partnership with Google Cloud, Model Medicines launched an **unprecedented virtual screening campaign** targeting BRD4. In a report bylined October 30, 2025 (the page’s metadata shows a conflicting date), the company says the *GALILEO*™ engine achieved a throughput of **325 billion** compounds **per day** (www.biospace.com) (modelmedicines.com). According to the company, GALILEO uses lightweight models trained on GPUs and run on CPUs, with Google Kubernetes Engine scheduling and autoscaling thousands of containerized jobs. The company’s figure caption says that, running on 500 AMD EPYC CPUs, GALILEO achieved a record 325 billion compound throughput per day; the page does not describe the compound library or give a run time (modelmedicines.com). (Figure 1 illustrates the architecture.)

<img src="https://i.imgur.com/3EImYay.png" alt="Model Medicines GALILEO on Google Cloud (schematic)"> Figure 1: Model Medicines’ ULVS architecture. Thousands of containerized jobs run on Google Cloud (AMD EPYC CPUs) in parallel, reaching a record throughput of 325 billion molecules per day (modelmedicines.com).

The company’s page does not describe the composition of the screened library; it says only that “Cloud Storage holds input libraries, features, and all prediction outputs.” The model had been pre-trained on diverse binding data (including out-of-distribution examples) to enhance generalization (modelmedicines.com). Importantly, Model Medicines emphasized novelty: the model prioritized leveraging *rare chemotypes* rather than memorizing close analogs (modelmedicines.com).

The result of the screen was a ranked list of candidates with high predicted BRD4 affinity. Rather than docking each molecule, the ML model provided instantaneous scores. The company states that the campaign delivered MDL-4102; it does not describe any filtering steps between the screen and the candidate.

MDL-4102: Properties and Novelty

The company describes MDL-4102 as a “first-in-class, selective BRD4 inhibitor with no observable BRD2/BRD3 activity”; the release gives no assay type, potency values or methods, so this is an unverified company statement (modelmedicines.com). Such selectivity is rare: almost all prior BET inhibitors inhibited BRD2/3 similarly (pmc.ncbi.nlm.nih.gov). The company’s release uses the term “molecular choke points” for “the small set of functionally indispensable, evolutionarily conserved mechanisms that multiple diseases depend on”; it gives no structural or mechanistic explanation for the claimed selectivity (www.biospace.com).

Chemical similarity analysis shows that MDL-4102 occupies novel regions of chemical space. The company reports, based on ECFP4 Tanimoto, that MDL-4102 is chemically novel relative to clinical BET inhibitors, but it does not disclose similarity values or thresholds (modelmedicines.com). The company’s page states only that “MDL-4102 is chemically novel relative to clinical BET inhibitors (based on ECFP4 Tanimoto)”; any implication for toxicity or cross-resistance is speculation not supported by the company’s materials (modelmedicines.com). The company frames this novelty as evidence that large-scale AI screening can find “entirely new classes of drugs once thought unreachable” (www.biospace.com).

Pharmacological profiling is still ongoing (MDL-4102 is in preclinical development). The Synapse database (patsnap) lists MDL-4102 as a “small molecule inhibitor of BRD4” in preclinical status (synapse.patsnap.com). No formal in vivo or toxicity data have been published yet. Model Medicines has not published ADMET or safety data for MDL-4102. Separately, the company describes an Ames-test agent, AmesNet, which it says “outperformed all literature-reported Ames agents”; its materials do not state that MDL-4102 was assessed with it (www.biospace.com).

In summary, MDL-4102 represents a **proof-of-concept** for the GALILEO ULVS approach: a computationally-derived hit with high novelty, designed for a challenging target. In May 2026 the company stated that MDL-4102 is currently in IND-enabling studies, with IND submission targeted for 2027.

Comparison to Known BRD4 Inhibitors

Table 1 (above) lists other BET compounds that have been tested clinically. The company reports no observable BRD2/BRD3 activity for MDL-4102, but no peer-reviewed data are available to show whether it would behave differently from earlier BET inhibitors.

We note that MDL-4102 was discovered purely *in silico*, unlike traditional HTS leads. This means that rigorously validating its activity and drug-like behavior will require substantial experimental work. The company presents its identification of MDL-4102 as showing that ultra-broad virtual libraries can yield tractable leads, but without published assay data this remains unverified. It parallels Model Medicines’ earlier MDL-001 program, a broad-spectrum antiviral that the company says was found via GALILEO and that a preprint reports inhibits a range of RNA viruses in vitro (www.biospace.com). MDL-001 serves as a case study: its rapid

discovery and preclinical success (including activity against SARS-CoV-2 and HCV) suggest the GALILEO platform can produce actionable molecules (www.biospace.com). The company presents MDL-4102 as an extension of the platform into oncology, but no independent or peer-reviewed data validate it.

Data Analysis and Evidence

The evidence for Model Medicines' claims is drawn from company publications, presentations, and independent literature. Key quantitative points include:

- **Scale of screening**: The 325-billion figure is reported by the company and repeated in its BioSpace press releases; these sources are not independent. BioSpace notes Model Medicines "executed a 325-billion-compound ULVS in a day in 2025 in partnership with Google" (www.biospace.com). The company's own report details "GALILEO™ achieved 325 billion molecule throughput, marking the first in silico campaign at hundred-billion scale" (modelmedicines.com). By comparison, Atomwise's known benchmark was 8 billion per day (modelmedicines.com), and the company says laboratory HTS has the potential to evaluate one million compounds per day (www.biospace.com). On the company's own figures, 3.25×10^11 compounds per day is roughly 40 times the reported Atomwise rate (8×10^9) and about 325,000 times that HTS rate (10^6).
- **Hit novelty**: Model Medicines emphasizes the novelty of MDL-4102. They report ECFP4 fingerprint analysis showing MDL-4102 is **chemically dissimilar** from clinical BET inhibitors (modelmedicines.com). This novelty statement is the company's own; no independent analysis has been published, and it says nothing about patentability or freedom to operate. In contrast, many first-generation BET inhibitors shared common cores (e.g. triazolo [4,3-a]diazepine ring in JQ1-like compounds (pmc.ncbi.nlm.nih.gov)).
- **Selectivity**: The claim of "no observable BRD2 nor BRD3 activity" for MDL-4102 comes from the company's data (modelmedicines.com). No assay type, potency or fold-selectivity data are disclosed, so the claim cannot be independently assessed. Reviewers note that *even compounds explicitly designed to be BD1- or BD2-selective often still affect multiple BETs* (pmc.ncbi.nlm.nih.gov)
- **Comparative performance**: In the absence of direct IC50 or cellular data on MDL-4102, we rely on indirect comparisons. Table 1 lists compounds tested clinically but reports no response data. BRD4 is implicated in certain cancers (e.g. NUT carcinoma, myelofibrosis, leukemia), but no efficacy data for MDL-4102 are reported here (pmc.ncbi.nlm.nih.gov). Future combination strategies might also be needed, as seen with other BETi.
- **Technical throughput**: The company's figure caption states that, running on 500 AMD EPYC CPUs, GALILEO achieved a record 325 billion compound throughput per day (modelmedicines.com). The caption does not say whether "CPUs" means sockets, instances or cores, and that page gives no run time. The company's May 2026 release describes a 325-billion-molecule screen as run "in a single 24-hour period" (a company claim, not independently verified), so any per-CPU rate is only an indicative estimate given the unclear CPU definition. Taken literally, 3.25×10^11 compounds per day over 500 CPUs is about 6.5×10^8 compounds per CPU per day (roughly 2.7×10^7 per CPU-hour).

- **Generative impact**: Model Medicines also emphasizes generative design. Their January 2025 BioSpace release describes generating a vast candidate space ("over 50 trillion compounds") and identifying hits with "100% hit-rate" in certain tests (www.biospace.com). According to the company, these figures indicate that its pipeline is not just screening existing molecules but also **designing novel ones**. The sources reviewed do not say whether generative design contributed to MDL-4102.

Overall, the throughput, selectivity and novelty claims come from company press releases and have not been independently verified or peer reviewed. Objective performance metrics (IC50s, cell assays) for MDL-4102 have not been published.

Case Study: The MDL-001 Antiviral Program

As a **proof-of-concept** for their platform, Model Medicines previously announced MDL-001, a broad-spectrum antiviral discovered by GALILEO (www.biospace.com). The company describes MDL-001 as a broad-spectrum antiviral with "exceptional safety, ADME and potency properties", and a preprint with researchers at the Icahn School of Medicine at Mount Sinai and the University of California, San Diego reports that it inhibits a range of RNA viruses *in vitro* (www.biospace.com). Notably, MDL-001 was effective against multiple coronaviruses (including SARS-CoV-2 variants), hepatitis C, norovirus, and influenza (www.biospace.com). The company presents this as evidence that AI-led chemistry can deliver antiviral candidates. The company's release describes RdRp Thumb-1 as "a previously unrecognized, conserved allosteric and cryptic pocket of viral RNA-dependent RNA polymerase (RdRp)"; this is the company's own characterisation (www.biospace.com).

MDL-001's development parallels MDL-4102 in several ways. The two programs differ in scale: the company reports that the MDL-4102 campaign ran at a throughput of 325 billion compounds per day (modelmedicines.com). The MDL-001 program was achieved in collaboration with labs at Mount Sinai and UC San Diego, and pursued through partnerships (the resulting company Viromme is mentioned in the BioSpace release (www.biospace.com)). While MDL-001 addressed infectious disease, MDL-4102 shows the platform's reach into oncology. Together, they exemplify how modern AI pipelines can yield small-molecule leads across diverse disease areas (modelmedicines.com).

Discussion: Implications, Critiques, and Future Directions

Transforming Hit Discovery

The reported identification of MDL-4102 points to a potential shift in drug discovery. Ultra-large virtual screens can **amplify novelty**: by piercing deeper into chemical space, they find chemotypes unattainable by smaller screens or traditional medicinal chemistry intuition (modelmedicines.com) (medicalxpress.com). If MDL-4102 advances, it would be a notable step for AI-derived oncology candidates, but it has not entered clinical testing

and no peer-reviewed data are available. It could add to the case for targeting epigenetic regulators like BRD4 with selective agents. More broadly, the approach suggests that other “hard” targets – transcription factors, protein–protein interfaces, etc. – might be addressable by mining billions or trillions of candidates.

Beyond a single compound, the platform effect is profound. As Model Medicines CEO Daniel Haders puts it, trillion-scale ULVS “fundamentally change[s] what chemistry can be discovered, what diseases can be solved, and how many patients can be reached” (www.biospace.com). By integrating AI from target ID to candidate selection and ADMET forecasting, the traditional timeline could shrink. Indeed, others have reported dramatically reduced discovery timelines using AI (Insilico’s reported average of 12–18 months from discovery to IND-enabling studies across 22 benchmark programmes; Recursion’s 18 months from discovery to IND-enabling studies for one programme, REC-1245) (www.scientific-computing.com). Model Medicines reports 100% hit rates in some in vitro assays (a company-reported claim). Nevertheless, such integration of ultra-scale computation and ML-driven chemistry could eventually make early discovery a year(s) rather than a decade-long endeavor.

The immediate impact is on pipeline economics and competition. The Fierce Biotech Week release describes “choke points” as “the small set of functionally indispensable, evolutionarily conserved mechanisms that multiple diseases depend on” (www.biospace.com); it does not discuss the cost of discovery, and any implications for pipeline economics or business models are the author’s own speculation. The company says its 2025 325-billion-compound screen was run in partnership with Google, but its May 2026 release names no investors or commercial partners for these capabilities (www.biospace.com).

Limitations and Critical Perspectives

While promising, caution is warranted. Critics of AI drug discovery often note that **“not a single AI-discovered drug has completed clinical trials and reached patients”** (as of that article, which may now be out of date) (www.scientific-computing.com). A Scientific Computing World piece reminds us that accelerated preclinical development has not yet translated into new drugs reaching patients (www.scientific-computing.com). In that view, MDL-4102 is a hopeful candidate but must still prove safety and efficacy in vivo. It remains to be seen whether its reported selectivity will hold up in further testing.

There are also technical caveats. Large screening can amplify uncertainties: models trained on limited data may still make false predictions when extrapolating. For example, predicted mutagenicity scores are computational outputs and are not proof of safety. Even if MDL-4102 passes standard in vitro tests, metabolism, immunogenicity, or other liabilities could surface only in animals or patients. Moreover, ultra-large virtual screens often rely on *simplified* interaction models (e.g. 2D fingerprints or generic binding models), which may miss some key physics. Experts note that “docking remains the most validated approach” for ULVS (www.sciencedirect.com), suggesting that AI scorings should ideally be cross-checked by at least some structural modeling of the final hits. Model Medicines’ publications do not detail whether MDL-4102 was further evaluated by docking or other structural methods; presumably future peer-reviewed work will elucidate its binding mode.

Another concern is overhype. The enthusiasm around AI drug discovery has been tempered by recent failures of AI startups and unmet promises (www.scientific-computing.com) (www.genengnews.com). A commentary by Zhavoronkov (Insilico CEO) wrote in August 2024 that “during the past decade, I saw hundreds of overhyped

AI startups aiming to disrupt pharma emerge" (www.genengnews.com). Others argue that billions of investment in AI must still yield actual patient benefits to justify the excitement (www.scientific-computing.com). Model Medicines' achievement, while significant, is not a drug itself but a discovery step. Scientists will be watching whether MDL-4102 can survive real-world pharmacology.

On the other hand, skeptics have evidence on their side: Endpoints News reports that the first handful of AI-designed molecules have failed trials or been deprioritized after being brought into the clinic with considerable fanfare (endpoints.news). In its January 2025 release, the company says that, with IND-enabling studies underway, it is moving compounds towards the clinic, without tying the statement to MDL-001 specifically; the release gives no clinical phase, sample size or results for MDL-001. If MDL-4102 enters investigator-led trials (or partners with industry), it could become a bellwether for AI discovery.

Future Directions

Scaling further. Model Medicines has stated plans to reach trillion-scale screens by 2026 and even quadrillion (10^15) by 2030; in a March 2026 release it said it was constructing a one-trillion-compound screen, and no completed trillion-scale result has been found in the sources reviewed (modelmedicines.com). Achieving this will require even more parallelism and efficiency. The company's Figure 5 caption says GALILEO is "projected to reach trillion-scale screening by 2026 and quadrillion-scale by 2030"; no trillion-scale result is reported on that page (modelmedicines.com). As computational costs plummet and cloud adoption grows, such screens may become routine for large targets. We may see AI pipelines routinely sifting through 10^12–10^15 compounds, potentially uncovering unimaginable chemical motifs.

Integrated design. Beyond enumeration, future pipelines may entwine VS with **de novo** design in closed loop. In the coming years, one can envision *on-the-fly* generation: as screening data arrives, the model instantly proposes new molecules targeting unexplored chemotypes. Large language models (LLMs) trained on chemistry could even articulate multi-parameter objectives. In their presentations, Model Medicines has advocated "LLM use cases" and agentic systems to manage complex drug profiles (www.biospace.com). For example, an LLM agent might automatically iterate between screening, synthetic planning, and ADMET optimization, tracking alignment with the target product profile (www.biospace.com). If successful, this could blur the line between computational discovery and medicinal chemistry planning.

Beyond small molecules. While MDL-4102 is a small organic compound, the same principles could extend to larger modalities (e.g. peptides, oligonucleotides) by appropriate featurization. The fundamental idea is permutation of building blocks; thus, ultra-large combinatorial libraries of macrocycles or DNA-encoded libraries (DELs) could also be screened virtually.

Clinical and regulatory impact. The coming years will tell whether MDL-4102 and its successors can reach patients. Regulators will scrutinize whether AI-designed drugs have distinct risk profiles. Intellectual property issues will also arise: MDL-4102-like molecules may not be "invented" in the classic sense. Yet the promise of reaching "deeper, unexplored regions of chemical space" (www.biospace.com) to find cures for previously intractable diseases is compelling. If Model Medicines can deliver even one new drug out of ULVS, it could shift industry practices. Democratic access to their methods (or open-source analogs like VirtualFlow) means that academic labs and other companies will likely adopt similar strategies.

Conclusion

Model Medicines' MDL-4102 and its 325-billion-molecule virtual screen exemplify a new paradigm in drug discovery. By leveraging AI and cloud-scale computing, the company says it surpassed a threshold previously thought unreachable (modelmedicines.com). According to the company, MDL-4102 is the **first notable outcome** of this effort: a selective BRD4 inhibitor with claimed novel chemistry and no observable BRD2/BRD3 activity (modelmedicines.com). This breakthrough offers hope for targeting disease pathways that eluded conventional screening. The throughput, selectivity and novelty claims are the company's own statements, with no peer-reviewed data or disclosed assay results; the work aligns with the broader trend of ultra-large virtual screening and AI-assisted design (modelmedicines.com) (www.sciencedirect.com).

However, the real test will be in biology and the clinic. MDL-4102 must now navigate preclinical development to demonstrate that its AI-derived design yields tangible patient benefits. Concurrently, continued advancements in computational methods (e.g. more accurate binding models, multiobjective optimization agents) will refine and extend this approach.

In the near term, we anticipate Model Medicines expanding its "pipeline-in-a-pill" strategy to more targets. In the long run, the marriage of ULVS with generative AI may redefine how we scout chemical matter for medicines. This report has documented one of the first victories in this journey. If such AI-discovered compounds ultimately achieve regulatory approval, the pharmaceutical landscape will have irrevocably changed: breadth of chemical exploration will no longer be the limiting step.

References

- Model Medicines press releases and website (www.biospace.com) (modelmedicines.com) (www.biospace.com) (modelmedicines.com)
- Liu *et al.*, *J. Med. Chem.* 2017 (review of BRD4 as target) (pmc.ncbi.nlm.nih.gov).
- Qian *et al.*, *Cell Death Discov.* 2023 (BRD4 in cancer) (pmc.ncbi.nlm.nih.gov).
- Scientific Computing World 2024 (AI drug discovery critique) (www.scientific-computing.com).
- Dutchen S., *MedicalXpress* 2020 (VirtualFlow) (medicalxpress.com).
- Gorgulla *et al.*, *Nature* 2020 (ULVS platform) (medicalxpress.com).
- Drug Discovery Today 2026 (reviewing ULVS) (www.sciencedirect.com).
- Model Medicines / Synapse pipeline database (MDL-4102 info) (synapse.patsnap.com) (synapse.zhihuiya.com).
- Gorgulla *et al.*, *Annual Rev. Biomed. Data Sci.* 2023 (ULVS trends) (www.annualreviews.org).
- Chen *et al.*, *Signal Transduct. Target. Ther.* 2023 (BET proteins review) (www.nature.com).
- Model Medicines press release, January 28, 2025 (announcing the GALILEO preprint) (www.biospace.com).
- Scientific publications on known BETi (AbbVie, BMS, etc.) (pmc.ncbi.nlm.nih.gov).

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://www.biospace.com/press-releases/model-medicines-to-present-ultra-large-virtual-screening-and-agentic-ai-breakthroughs-at-ai-convergence-small-molecule-discovery-summit
2. https://modelmedicines.com/newsroom/record-scale-ai-screening-with-model-medicines-on-google-cloud-galileo-achieves-325-billion-molecule-throughput-for-oncology-drug-discovery
3. https://www.scientific-computing.com/article/why-ai-drug-revolution-hasnt-delivered
4. https://medicalxpress.com/news/2020-03-software-screens-billions-compounds-medicines.html
5. https://pmc.ncbi.nlm.nih.gov/articles/PMC5464988
6. https://www.sciencedirect.com/science/article/pii/S1359644626000218/pdf
7. https://pmc.ncbi.nlm.nih.gov/articles/PMC10746725
8. https://i.imgur.com/3EImYay.png
9. https://www.biospace.com/press-releases/model-medicines-to-present-end-to-end-ai-drug-discovery-platform-at-fierce-biotech-week-2026-in-boston
10. https://synapse.patsnap.com/drug/242fbbbc731f4836941d656b0283ce73
11. https://www.biospace.com/press-releases/model-medicines-unveils-generative-ai-multi-modal-therapeutic-pipeline-breakthrough-demonstrating-end-to-end-drug-discovery-capabilities
12. https://www.genengnews.com/topics/artificial-intelligence/is-generative-ai-in-drug-discovery-overhyped
13. https://endpoints.news/first-ai-designed-drugs-fall-short-in-the-clinic-following-years-of-hype
14. https://synapse.zhihuiya.com/drug/242fbbbc731f4836941d656b0283ce73
15. https://www.annualreviews.org/content/journals/10.1146/annurev-biodatasci-020222-025013
16. https://www.nature.com/articles/s41392-023-01647-6

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore AI enablement, pharma and biotech software development, data engineering and BI, and Veeva services. Contact IntuitionLabs to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: https://intuitionlabs.ai

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.