

Rapid, Single-Iteration Hit Identification for CLK1 using Direct-to-Biology

We discovered in just four weeks seven hits for CLK1, including one with subnanomolar activity.

Challenge

Small-molecule synthesis remains a central bottleneck in early drug discovery. Outsourced synthesis through CROs is costly, in-house chemistry often fails to scale, and commercially available or “on-demand” libraries frequently lack structural novelty or direct relevance to specific biological hypotheses. As a result, turning ideas into experimentally validated hits takes far too long.

Molecule.one’s solution

Using high-throughput experimentation (HTE) data as foundation, we built D2B-SpaceM1: the first generative, deep-learning chemical space trained on more than 300,000 experimentally observed reaction outcomes derived from Molecule.one’s MARIA™ platform. Unlike abstract virtual libraries, D2B-SpaceM1 is explicitly feasibility-anchored: every generated molecule is associated with a concrete, data-supported synthetic route, enabling immediate translation from *in silico* prioritization (AI predictive models, docking) to rapid direct-to-biology synthesis and testing experimental synthesis.

Notably, MARIA™ can operate at plate scale, enabling up to ~13,000 reactions per week with potential for further scaling. In addition, our synthesis models support uncommon yet effective conditions, such as Suzuki couplings in DMSO. This significantly broadens the range of makeable chemistry beyond standard constraints.

Target

CLK1 is a protein kinase that regulates RNA processing inside cells, making it an emerging therapeutic target in cancer and neurodegenerative diseases.



Virtual Hit Prioritization

39 compounds were prioritized from D2B-SpaceM1 in collaboration with Inventro using their proprietary machine-learning models. An additional 18 compounds were selected from a suggested list of BB-derived analogs accessible through available building-block combinations and were independently confirmed as potentially active by Inventro's AI Predictive Models. Finally, 37 further BB-derived analogs were added to expand the set.

D2B Synthesis

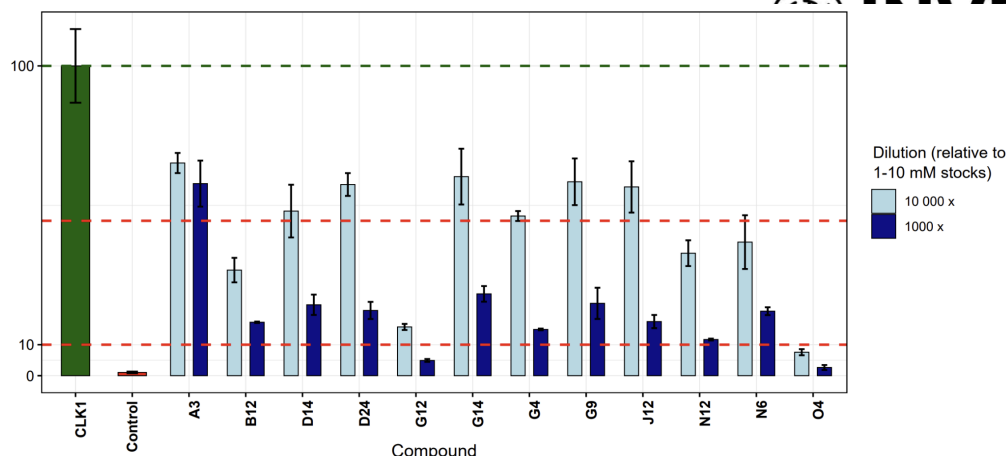
Selected molecules were advanced directly into microliter scale, plate-based synthesis, followed by minimal work-up and immediate biological testing. Across the CLK1 D2B grid ~93% of targeted products were successfully formed (87 out of 94), underscoring the robustness of feasibility-aware design at plate scale.

Primary and Secondary Screening (ADP-Glo)

Biological activity was assessed using the ADP-Glo™ luminescent kinase assay.

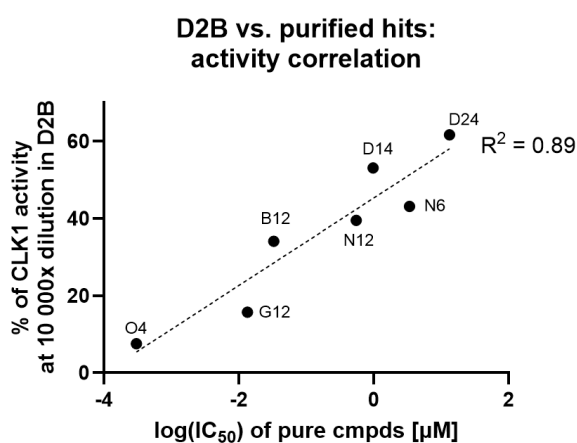
Primary screen: single-point measurement at 100× dilution identified 12 active compounds with <10% residual CLK1 activity.

Secondary screen: additional measurements at 1,000× and 10,000× dilutions confirmed 11 active compounds, corresponding to a hit rate of 18% for AI-prioritized compounds (12% across the full grid). ADP-Glo activity (% CLK1 residual activity) across 12 preselected D2B wells is depicted below.



Hit Confirmation

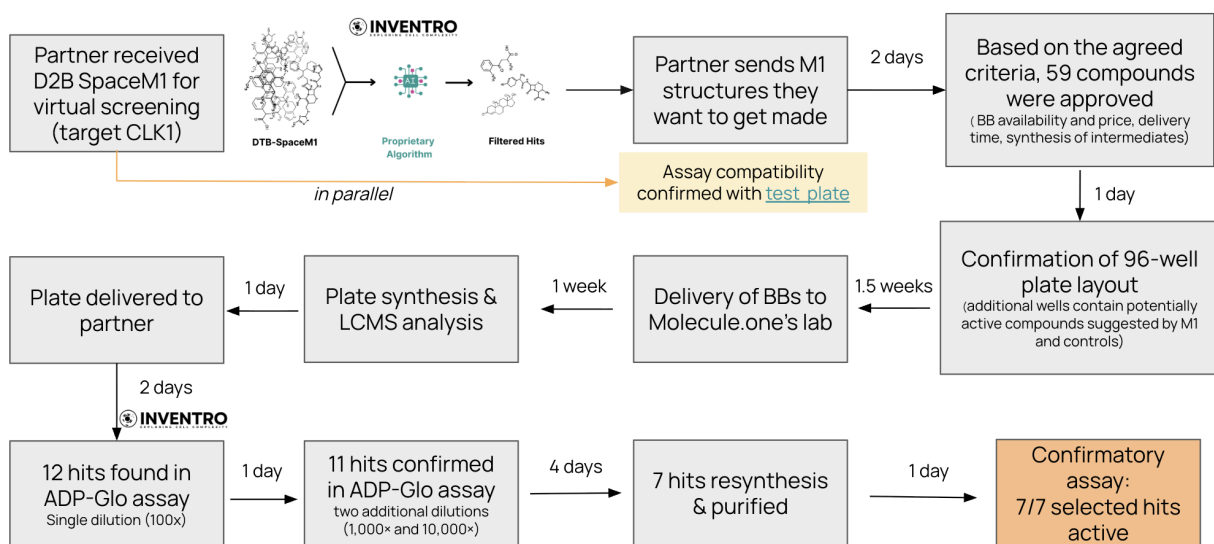
Based on activity and structural diversity, 7 compounds were selected for resynthesis, purification, and confirmatory testing. 7/7 purified compounds confirmed active in dose-response assays. Potency distribution spanned from μM to sub-nanomolar range.



A strong correlation ($R^2 = 0.89$) was observed between CLK1 inhibition measured directly from D2B crudes and IC_{50} values determined for purified compounds. This demonstrates that crude D2B data are sufficiently predictive to support confident hit selection without iterative purification cycles.

The most potent compound exhibiting sub-nanomolar activity is an analog suggested based on available building-block combinations (confirmed by the ML model). HitID was completed within 4 weeks.

Summary: Rapid Workflow



The Figure above summarizes the whole process. It took us 4 weeks to go from design to resynthesized and confirmed sub-nM hits.

	CLK1
Hit selection	ML predictive model by Inventro

Synthesis success rate	93 %
Hit rate	18 % (AI-prioritized), 12 % (full grid)
Potency of hits	Sub-nM top, followed by low μ M activity
Compounds made (in one week)	~100

The data show that our chemistry—diverse, feasibility-anchored SpaceM1 and rapid HTE execution—streamlines DMTA cycles and cuts resource use in early discovery allowing for hit identification within weeks not months, while requiring minimal FTE involvement.