SOLUTION BRIEF

# Pharmacelera Achieves Top Score of All 24 Teams in CACHE Challenge #5

**Scalable Ligand-Based Discovery of Novel MCHR1 Antagonists with exaScreen® in Ultra-Large Chemical Space**

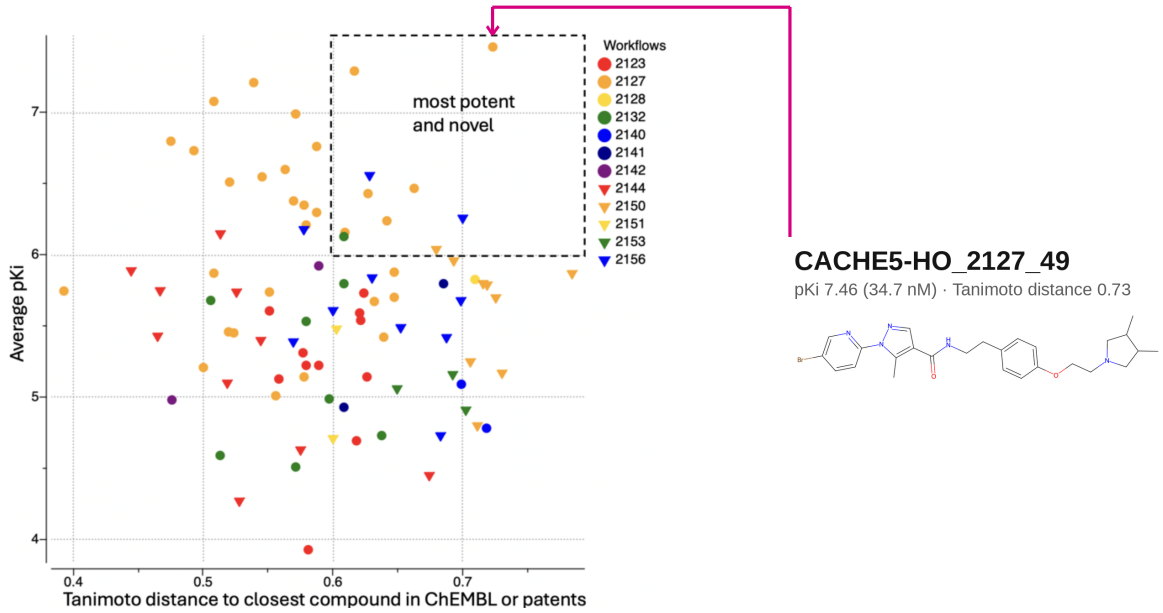



***Figure 1: Potency and novelty of Round 2 analogs**. pKi values are derived from competition experiments. Tanimoto distances were calculated with ICM (Molsoft). Source: CACHE Challenge #5 official results (cache-challenge.org/results-cache-challenge-5), chart reproduced unmodified; molecular structure and arrow annotation added by Pharmacelera to identify CACHE5-HO_2127_49. Pharmacelera (workflow 2127, orange) is the only team with six points inside the "most potent and novel" pKi ≥ 6, Tanimoto distance ≥ 0.6 box.*

## Introduction: Expanding Discovery Beyond Structural Constraints

Drug discovery teams are under increasing pressure to identify novel, potent compounds within rapidly expanding chemical spaces, often in the absence of target structural data. Traditional approaches struggle to balance scalability, chemical novelty, and predictive performance: two-dimensional similarity methods tend to converge on known scaffolds, while brute-force enumeration remains computationally impractical at a meaningful scale.

In CACHE Challenge #5, Pharmacelera set out to predict novel antagonists of the melanin-concentrating hormone receptor 1 (MCHR1), a GPCR target involved in metabolic disorders, including obesity, sleep disorders, depression, and anxiety. With no experimental structure available, the team applied a ligand-based, physics-grounded approach using its exaScreen® platform. This synthon-driven approach is designed to identify functionally relevant compounds across vast chemical space, delivering potency, novelty, and synthetic accessibility, without reliance on structure-based methods.

Pharmacelera earned the highest score of any of the 24 participating teams by employing a scalable ligand-based workflow centered on three-dimensional molecular field similarity to deliver novel, potent hits — including the most potent compound identified across the entire challenge — and was the only team to identify four compounds with pKi greater than 7. The following sections explain how.

## Challenge: Novelty, Scale, and Limited Structural Insight

CACHE Challenge #5, run by the Critical Assessment of Computational Hit-finding Experiments (CACHE) consortium, focused on identifying novel antagonists of the MCHR1 GPCR under conditions that reflect real-world discovery constraints. Participants were asked to predict up to one hundred chemically novel compounds that would test below a 30 µM IC50 threshold, working from a shared dataset of 2,605 ChEMBL compounds and 753 patented molecules — with no experimental MCHR1 structure available to guide the search.

Twenty-four teams from twelve countries took part, following a double-blind peer review of applications. The searchable chemical space spanned tens of billions of compounds, making conventional screening approaches inefficient or impractical. That scale is exactly what Pharmacelera's exaScreen® workflow was built to navigate: the team's strategy narrowed ~48 billion candidates down to the 38 compounds selected for Round 1 synthesis and testing, then carried the top hits into Round 2 lead optimization to reach four compounds with pKi greater than 7 — including the challenge's most potent compound, and the most novel among those with pKi > 6.

**From Scale to Selection**

**48B → ~1,000 → ~100 → 38 tested → 4 Top Hits**

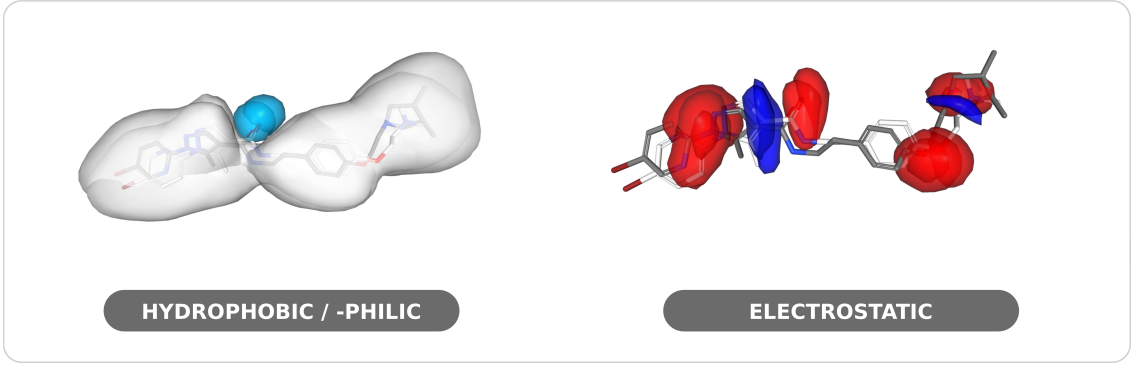
## Solution: Ligand-Based Discovery Guided by Molecular Field Descriptors

Pharmacelera addressed this challenge with exaScreen®, a ligand-based workflow centered on three-dimensional molecular field similarity. exaScreen®'s synthon-based approach enables rapid screening of ultra-large libraries (ULL) with the accuracy of Pharmacelera's physics-grounded hydrophobic and hydrophilic field descriptors — described in detail in the following sections. Rather than relying on structural models or 2D chemical similarity, exaScreen® represents each ligand by its quantum mechanics (QM)-derived 3D molecular field — the electrostatic, hydrophobic, and steric interaction pattern it presents to a binding site — and identifies compounds that share that pattern despite very different two-dimensional structures. That descriptor accuracy, not simply the choice of a ligand-based strategy, is what let the platform operate without any experimental MCHR1 structure.

These fields are generated by QaiM, Pharmacelera's AI-based QM descriptor technology — conformer-dependent models that accurately reproduce each molecule's hydrophobic and electrostatic profile at ultrafast calculation speed, without the computational cost of running first-principles QM directly. That combination of QM-level accuracy and practical speed is what makes exaScreen® workable across ultra-large libraries rather than a small, hand-curated set.

exaScreen® exploits the hydropathic profile of the ligands, which describes the complementarity of ligands with protein's pockets. Field complementarity goes beyond atomic or structural similarity, as it can identify isofunctional diverse scaffolds, pushing chemical novelty. Our workflow started with a cluster of the available active compounds, which was followed by the application of exaScreen®. Post-processing based on 2D dissimilarity to the full set of known MCHR1 ligands — from both patents and ChEMBL — combined with 3D field similarity to the reference, is the primary driver behind the chemical diversity seen across Pharmacelera's Round 1 and Round 2 hits.

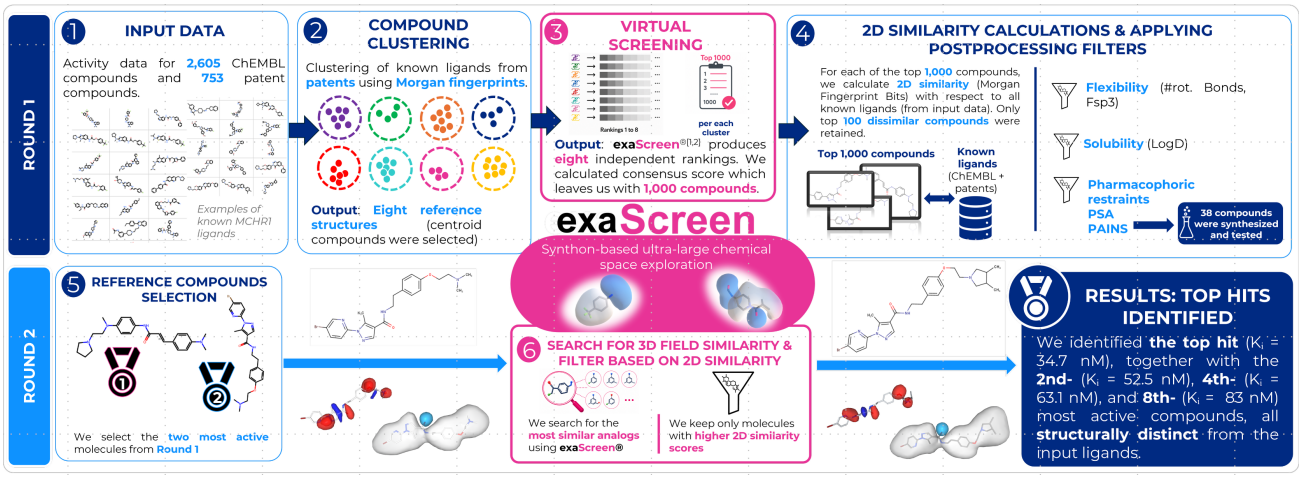
*Figure 2. QM-Derived 3D Molecular Field of CACHE5-HO_2127_49*




*White/Blue: hydrophobic field (Nonpolar/polar). Red/blue: electrostatic field (Negative/positive). Generated by exaScreen®, rendered by MolXPlore®, from the same QM-derived 3D field descriptors used throughout screening.*

The workflow began by clustering 753 known MCHR1 ligands provided by CACHE from the patent literature into eight representative reference structures, then exploring the Enamine REAL library (~48 billion compounds) using a synthon-based strategy native to exaScreen®, which preserves synthetic accessibility while avoiding brute-force enumeration. QM-derived 3D molecular field similarity, physicochemical filtering, and diversity selection enabled efficient prioritization of a final, high-quality candidate set. Starting with our two strongest Round 1 hits as reference compounds for Round 2, we then performed lead optimization by searching in exaScreen® for their closest analogs by 3D field similarity and then filtering based on 2D similarity.

*Figure 3. exaScreen® Application Workflow — CACHE Challenge #5 (Round 1 + Round 2)*




## Outcome: Top Results, Verified by CACHE's Independent Scoring

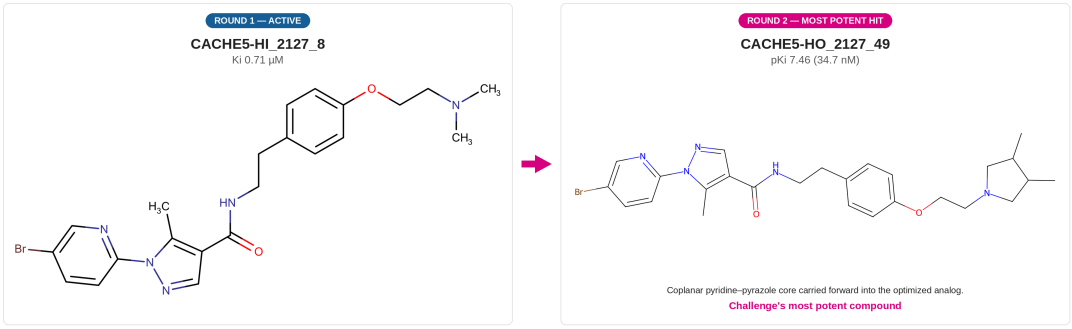
CACHE's Hit Evaluation Committee — independent experts from Bayer, Boehringer Ingelheim, Relay Therapeutics, AstraZeneca, and the University of North Carolina — blindly scored every Round 1 hit and its Round 2 analogs on Top Score, Aggregated Score, and a blind Merged Selection score. Pharmacelera's results led the field on multiple fronts.

- Highest Top Score of all 24 participating teams, per CACHE's official Table 1 — Pharmacelera's exaScreen®-identified hit outscored every other workflow's best compound.
- Highest score of any team on the blind Merged Selection task (89.9) — more than 25 points ahead of the next-best participant (62.0) — reflecting strong generalization on a task where team identities were hidden.
- The single most potent Round 1 hit across the entire challenge: 170 nM, sitting at the most potent end of CACHE's own reported range for all confirmed dose-response hits (170 nM–30 µM).
- Successfully predicted experimentally confirmed MCHR1 ligands with no MCHR1 structure — a scenario where most competitors turned to structure-based or hybrid methods instead, using AlphaFold structures or homology models.

### Round 1 — Primary Screening

- From the 1,455 compounds submitted across all 24 teams, Pharmacelera contributed just 38 (2.6% of the total) — yet delivered 4 of the 26 (15%) compounds that showed a full, confirmed dose-response across the entire challenge.
- Achieved a 10% dose-response hit rate (4 of 38 compounds) — roughly 1.5x the next-best team's 6.7% (6 of 89).
- Delivered 4 of the field's top-ranked most active compounds, placing 1st, 2nd, 4th, and 8th overall.

*Figure 4. Analog Expansion and Lead Optimization — Best Hit, Round 1 to Round 2*




## Round 2 — Analog Expansion & Lead Optimization

- Selected Pharmacelera's two most active Round 1 hits as anchors, then used exaScreen®'s 3D molecular field similarity-guided search and SAR analysis to propose up to 50 follow-up analogs, which produced the challenge's most potent hit, CACHE5-HO_2127_49 (pKi 7.46, 34.7 nM) — see Figure 4.
- The only team to deliver four compounds with pKi > 7 — at 34.7 nM, 52.5 nM, 63.1 nM, and 83 nM — including the single most potent compound identified across the entire challenge.
- Achieved the greatest chemical novelty of any submission among compounds with pKi > 6: an ECFP4 Tanimoto distance of 0.73 to the nearest known ligand for the top compound — well beyond the diversity achieved by competing structure-based and generative-AI approaches at that potency level.

Together, the results from Rounds 1 and 2 show that a purely ligand-based approach — without a target structure — can match or exceed structure-based and generative methods on potency, novelty, and hit rate.

## Why It Matters — exaScreen® at scale

Beyond this result, exaScreen® is built to scale, integrating seamlessly with existing medicinal chemistry, informatics, and data-management workflows. At its core is QaiM, Pharmacelera's AI-based QM descriptor engine, combined with a fragment-based, synthon search strategy that explores ultra-large chemical spaces without brute-force searching or excessive compound sourcing. Together, they deliver hits that are simultaneously potent, structurally diverse, and synthesizable — turning in silico predictions into an actionable, synthetically accessible starting point.

### Your Target Could Be Hit Next!

Pharmacelera's top-scoring result in CACHE Challenge #5 shows, under independent, blind validation, that exaScreen® delivers potent, novel, synthesizable hits at ultra-large scale.

**Request a Technical Discussion**

*Scan to connect*

contact@pharmacelera.com • www.pharmacelera.com