



When Colloids Look Like Leads: Preventing Aggregation-Driven False Discovery in Computational and Experimental Screening of Natural Product Compounds

Youssef Hariri¹, Hassan Nasser¹, Fatima Zahra²

¹Department of Colloidal Aggregation and Assay Artefacts, Faculty of Pharmacy, Mohammed V University, Rabat, Morocco.

²Department of False Discovery Prevention in Natural Product Screening, Faculty of Pharmacy, University of Casablanca, Casablanca, Morocco

ABSTRACT

Natural products occupy chemically rich space but can also present physicochemical behaviors that complicate screening interpretation. Among these, concentration-dependent colloidal self-assembly can transform an apparently potent molecular hit into a particle-mediated assay signal. This methodological framework examines aggregation as a false-discovery mechanism that spans experimental screening, virtual prioritization, and machine-learning data construction. The analysis distinguishes molecular structure from emergent colloidal state, apparent inhibition from target-specific pharmacology, colloid formation from precipitation or low free solubility, and rejection of an unsupported mechanism from rejection of the underlying scaffold. Evidence from biochemical and cellular screening, natural-product-derived compounds, biophysical studies, aggregation-specific machine learning, assay-liability prediction, and bioactivity-data curation is integrated into a proposed falsification-first strategy. The central proposal is that aggregation liability should enter discovery before mechanistic interpretation, computational flags should modify uncertainty and the next experiment, while condition-matched physical-state measurements and orthogonal counter-screens should test whether particle formation plausibly explains activity. A predicted aggregator is therefore not automatically rejected, and an experimentally observed aggregate is not automatically biologically irrelevant. Rescue through formulation represents a separate development hypothesis that must not retroactively validate an initial target claim. The framework is intended to reduce propagation of aggregation-contaminated activity labels into ranking systems and training sets while preserving chemically unusual candidates that survive falsification. The principal limitation is that the integrated workflow has not been prospectively validated across diverse natural-product libraries, assay formats, or model domains.

Key Words: Colloidal aggregation, Natural products, Assay interference, Machine learning, Virtual screening, Hit triage

eIJPPR 2025; 15(2):100-108

HOW TO CITE THIS ARTICLE: Hariri Y, Nasser H, Zahra F. When Colloids Look Like Leads: Preventing Aggregation-Driven False Discovery in Computational and Experimental Screening of Natural Product Compounds. Int J Pharm Phytopharmacol Res. 2025;15(2):100-8. <https://doi.org/10.51847/Hzib0val29>

INTRODUCTION

A primary screening signal is an observation, not yet a molecular mechanism. Small molecules can perturb assay readouts through aggregation, optical interference, chemical reactivity, reporter effects, and other nuisance

processes, each of which requires a different diagnostic response [1]. Colloidal aggregation is especially troublesome because the same nominal compound concentration can describe two physically different experimental situations: predominantly free molecular species or a suspension containing self-assembled

Corresponding author: Youssef Hariri

Address: Department of Colloidal Aggregation and Assay Artefacts, Faculty of Pharmacy, Mohammed V University, Rabat-Instituts, Rabat, Morocco.

E-mail: youssef.hariri@um5.ac.ma

Received: 15 December 2024; **Revised:** 27 March 2025; **Accepted:** 02 April 2025

This is an **open access** journal, and articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.



particles. If a particle interacts with protein, the resulting inhibition may look concentration dependent and reproducible while remaining disconnected from a defined ligand–target interaction. For natural products, where structural complexity can encourage strong mechanistic narratives, that distinction is essential: chemical complexity does not immunize a hit against mesoscale physicochemical behavior.

The concern is not confined to simplified enzyme assays. Colloidal aggregation has been shown to confound cell-based antiviral screening, has been observed during biological-target testing of minor cannabinoids, and has explained activity patterns in biochemical repurposing screens [2–4]. These studies do not establish a universal prevalence of colloidal false positives, nor do they imply that every aggregating compound lacks genuine pharmacology. They establish a narrower but consequential point: neither biochemical potency nor cellular activity is sufficient, by itself, to eliminate aggregation as an alternative explanation.

This article therefore treats aggregation control as an ordering problem in evidence generation. The proposed methodological principle is falsification first: before docking poses, target narratives, pathway hypotheses, or learned activity labels are allowed to carry mechanistic weight, plausible particle-mediated explanations should be challenged with condition-matched evidence. This ordering follows the wider logic of mechanism-aware assay-interference control [1], but extends it into computational prioritization and data curation. The aim is not to create a universal aggregator blacklist. It is to prevent an untested physical state from silently becoming a molecular mechanism and then propagating through screening, modeling, and lead selection.

Why natural product physicochemistry creates aggregation risk

Aggregation propensity is associated with molecular structure and physicochemical behavior strongly enough to support computational prediction, but those associations remain probabilistic rather than deterministic [5].

Hydrophobic and aromatic surface, ionization state, hydrogen-bonding patterns, amphiphilicity, conformational shape, and solvent exposure can collectively alter the balance between molecular solvation and self-association. Natural products frequently combine several such features in rigid, polycyclic, stereochemically rich, or polyfunctional architectures. The relevant methodological question is therefore not whether a scaffold “looks aggregating,” but whether its structure justifies raising an aggregation hypothesis under the concentration, buffer, cosolvent, temperature, and biomolecular environment actually used.

The consequence of self-assembly also depends on intent. Colloidal aggregation can undermine screening interpretation, yet related self-assembly behavior can sometimes be deliberately stabilized and exploited in formulation [6]. This creates an important distinction for lead triage: evidence that an initial target claim is particle-mediated may invalidate that mechanistic interpretation without automatically proving that the scaffold has no pharmaceutical value. Mechanism adjudication and scaffold disposition are therefore treated here as separate decisions.

Recent biophysical work further shows why simple one-descriptor filters are inadequate. Aromatic interactions and intermolecular hydrogen bonding can contribute to nanoaggregate formation, whereas structurally related compounds may remain soluble or proceed toward precipitation under different conditions [7]. Aggregated drug nano-entities can themselves also produce biological responses [8]. Thus, the relevant state space contains more than “active molecule” versus “inactive artifact.” It may include soluble species, colloidal particles, precipitated material, and biologically consequential particle states. The proposed framework treats these as competing physical-state hypotheses rather than fixed compound identities.

Table 1 organizes the methodological components that follow from this physicochemical argument and the interpretive boundary attached to each.

Table 1. Method components for interpreting aggregation risk before mechanistic attribution

Method component	What is interrogated	Interpretive use	Boundary
Structure-informed risk	Molecular features associated with aggregation	Decide whether aggregation merits early testing	Prediction does not establish particle formation
State-aware physicochemistry	Self-assembly versus soluble or precipitated behavior	Treat aggregation as condition dependent	No single feature or cutoff is universal
Biological-state check	Whether nano-entities alter biological response	Preserve particle-mediated biology as an alternative hypothesis	Biological response is not target engagement
Scaffold-disposition split	Whether self-assembly is nuisance or intentionally engineerable	Separate rejection of mechanism from rejection of chemistry	Formulation rescue requires independent validation

Note: The organization of these components is proposed here. The cited evidence supports the underlying physicochemical and biological distinctions, not a validated decision rule.

From soluble molecule to colloidal assembly

The transition from a soluble compound to a colloidal assembly is best understood as a change in experimental state, not as acquisition of a permanent molecular label. Amphiphilic drugs can display different biological responses as their self-assembled state changes [9]. Concentration and medium composition are therefore not merely assay settings; they can alter the physical object presented to the biological system. A compound may move among predominantly free species, dynamic assemblies, larger particles, and precipitated material as conditions change.

Once colloids form, their biological interface differs fundamentally from that assumed by conventional medicinal chemistry. Small-molecule colloids possess internal organization and can preferentially bind proteins at particle surfaces [10]. The resulting interaction is not equivalent to a ligand occupying a defined pocket with the stoichiometry represented by a docking pose. Molecular-dynamics analysis supports the plausibility of aggregate-protein interaction modes involving broader multivalent contacts rather than canonical one-ligand/one-site recognition [11]. Such simulations provide mechanistic plausibility rather than universal experimental proof, but they clarify why a monomer-centric computational representation can miss a relevant source of assay activity. Particle behavior is also constrained by mesoscale geometry. Experimental work indicates that surface crowding can limit protein access and inhibition by colloidal aggregates [12]. Aggregate-mediated activity should therefore not be reduced to a simplistic rule that more nominal compound necessarily yields proportionally more inhibition. Once the mechanistic unit becomes a particle, particle number, surface accessibility, aggregate architecture, protein adsorption, and free-compound concentration may all matter. Nominal concentration alone cannot identify which physical variable generated the observed screening signal.

How colloids produce apparently specific biological activity

Apparent specificity can emerge even when the intended molecular mechanism is wrong. Cellular screening is vulnerable to multiple nuisance mechanisms, including generalized cellular injury and other non-target perturbations, so colloidal aggregation should remain one hypothesis within a broader differential rather than the default explanation for every promiscuous hit [13]. This matters especially when a natural product produces a coherent phenotype: reproducibility, concentration dependence, or cellular activity can strengthen confidence that something biological occurred without establishing

which physical species produced it or which molecular target was responsible.

Mechanism-characterized nuisance compounds provide one way to interrogate that ambiguity. A curated collection of nuisance controls can be used to test whether an assay format is intrinsically susceptible to characterized interference behaviors [14]. Such controls cannot prove that a new natural product shares the same mechanism, but they convert assay integrity from an assumption into an experimentally examined property.

Orthogonal cellular readouts can further narrow competing explanations. High-content morphology reference compounds have been used to characterize patterns associated with cellular injury [15]. In an aggregation-aware workflow, these readouts are not target-identification tools; they are exclusion tools. A hit that loses apparent specificity when generalized injury becomes evident should not be promoted through a molecular-target narrative. Conversely, absence of a generic injury phenotype does not prove target engagement. The proposed framework therefore treats specificity as an evidence claim that must survive both particle-focused and nonparticle nuisance hypotheses.

Why standard computational screens cannot see the aggregate

Most virtual-screening and molecular machine-learning pipelines operate on representations of individual chemical structures. Aggregation-specific cheminformatics demonstrates that those structures contain statistical signals associated with colloidal liability [16], but the emergent particle itself is ordinarily absent from the representation. Concentration, cosolvent fraction, ionic strength, detergent, serum proteins, incubation time, and other state-defining variables are likewise rarely encoded in a standard fingerprint, molecular graph, or docking pose. A conventional computational score can therefore be internally consistent while answering the wrong physical question: how a monomer might interact, not whether the monomer remains the experimentally relevant species.

Aggregation-focused models partly address this blind spot. SCAM Detective, subsequent machine-learning approaches, and the BAD Molecule Filter show that aggregation liability can be estimated from molecular information with useful discrimination and expanding chemical-space coverage [17–19]. Their appropriate role is triage rather than observation. A model output represents a relationship to labeled aggregation behavior within its training domain; it does not measure whether the test compound has formed colloids in the current assay.

Explainable artificial intelligence adds a second interpretive layer. Model explanations can identify structural patterns associated with predicted aggregation

[20], which may help medicinal chemists understand why a compound was flagged. Yet attribution to a fragment or descriptor remains an explanation of model behavior, not experimental evidence for the intermolecular forces operating in solution. Benchmark construction also matters. Recent aggregation-classification work illustrates the use of explicitly balanced aggregator and nonaggregator datasets [21]. Such balance can be methodologically useful for model training or comparison, but it must not be mistaken for aggregation prevalence in a real screening library. The proposed computational role is therefore risk estimation with applicability-domain and uncertainty annotations, followed by empirical falsification.

Aggregators entering machine-learning training sets as “actives”

The aggregation problem becomes more durable when an experimental readout is converted into a reusable label. Contemporary computational systems can identify frequent-hitter and assay-interference liabilities through integrated false-positive screening, direct assessment of HTS artifacts, and expert-guided learning under sparse or imbalanced labels [22–24]. These approaches support an important data-engineering distinction: “active” and “free of known interference evidence” are not synonymous labels. Nor is “frequent hitter” synonymous with “colloidal aggregator.” Interference annotations should preserve the suspected mechanism, experimental basis, and uncertainty instead of collapsing different failure modes into one negative class.

Large bioactivity resources make this distinction consequential. Papyrus was constructed through substantial curation and standardization of activity records drawn from heterogeneous sources [25]. ChEMBL likewise spans multiple assay types, activity formats, and time periods [26]. These resources are indispensable for predictive modeling, but a historical activity record cannot be assumed to include an explicit examination of colloidal state. Missing aggregation metadata should therefore remain unknown, not silently become a nonaggregator label.

Data integration introduces another risk. Consistency assessment can materially affect molecular-property prediction when information from different sources is combined [27]. For aggregation-aware learning, an apparent contradiction could reflect experimental error, assay interference, altered physical state, differing assay conditions, or genuine context dependence. The proposed curation rule is therefore to preserve provenance and quarantine or down-weight unresolved interference-suspect observations rather than “clean” them into a single mechanistic truth.

Curcumin illustrates why such qualification matters in natural-product chemical space. Its extensive bioactivity literature has been critically reinterpreted in light of instability, reactivity, promiscuity, poor bioavailability, and assay liabilities [28]. Aggregation is not the sole explanation for curcumin’s behavior, and the case should not be generalized to all polyphenols or natural products. It nevertheless illustrates how an attractive bioactivity narrative can outpace evidential qualification. If such observations are reused uncritically as training labels, a model may learn to reproduce assay behavior while appearing to learn target pharmacology. The methodological response is not wholesale deletion of difficult chemistry, but explicit separation of activity evidence, interference evidence, assay context, and unresolved physical state.

Concentration, detergent sensitivity, and other diagnostic clues

Aggregation is most interpretable when physical-state evidence and assay perturbations converge. NMR relaxation can directly expose small-molecule nanoentities in solution [29], providing a readout conceptually different from asking whether an enzyme or cellular signal changes. Protocolized assessment of free-state solution behavior further emphasizes that self-aggregation depends on concentration, sample preparation, and medium conditions [30]. The most informative experiment therefore matches the physical-state measurement as closely as practical to the concentration, solvent fraction, buffer, incubation interval, and temperature used for the biological readout.

Detergent sensitivity can be informative because disruption of colloidal assemblies may attenuate particle-mediated inhibition, but it is not a universal verdict. Detergents can themselves alter proteins, membranes, or assay components. Serum is similarly nonabsolute: some drug aggregates can remain stable under high-serum conditions and produce pharmacokinetic consequences [31]. Loss of activity after a state-modifying perturbation strengthens an aggregation hypothesis only when the perturbation is assay compatible and appropriately controlled; persistence of activity does not automatically eliminate aggregation.

Cell-based aggregation evidence likewise favors paired interpretation over reliance on one diagnostic clue [2]. The proposed standard is convergence: concentration behavior, a direct particle or solution-state measurement, and at least one orthogonal assay perturbation should be interpreted together when aggregation is plausible. Negative evidence should be recorded explicitly. A compound that remains predominantly nonparticulate across the active range and retains activity under appropriate aggregation-disrupting

conditions has survived a stronger falsification test than one that merely lacks a structural aggregation alert.

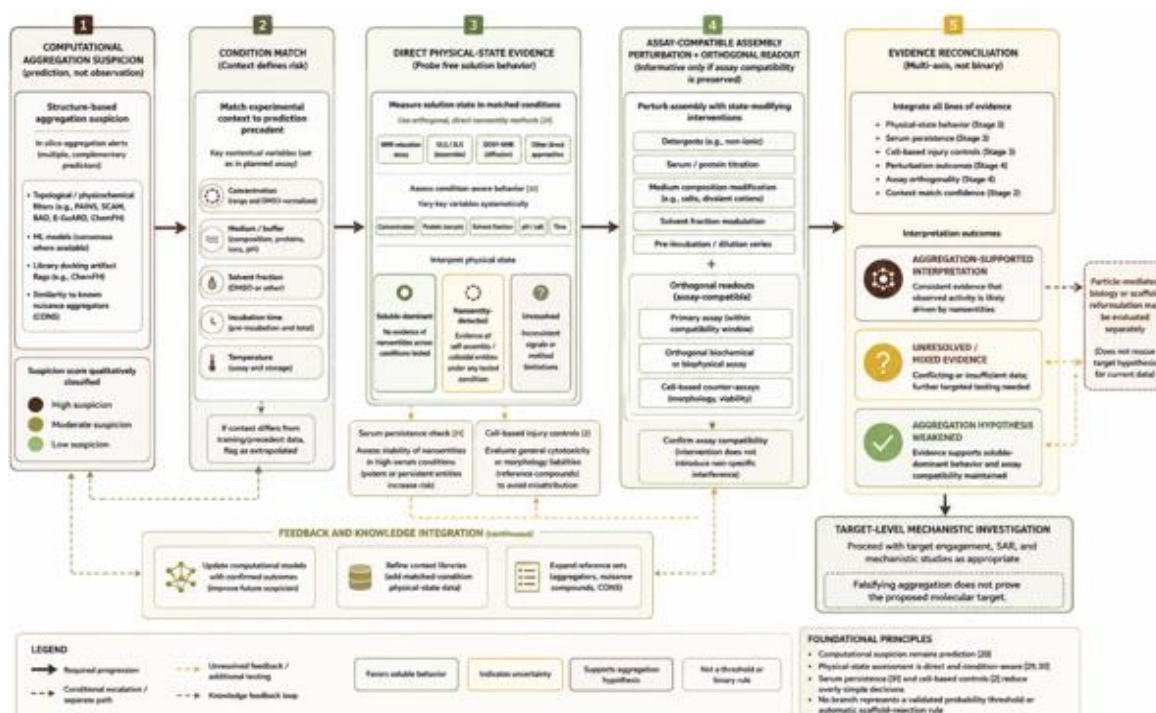
Table 2 converts these section-derived diagnostic principles into an ordered experimental sequence without imposing universal cutoffs.

Table 2. Proposed workflow steps for falsifying aggregation before assigning mechanism

Step	Question	Evidence sought	Interpretation
Match conditions	Is physical state measured near the biological assay condition?	Concentration- and medium-matched solution assessment	Avoids inference from irrelevant conditions
Detect particles	Are nanoentities observable?	Direct physical-state evidence such as NMR relaxation	Supports particle presence, not causal inhibition
Perturb assembly	Does activity change when aggregation is disrupted under assay-compatible conditions?	Controlled state-modifying comparison; cell-based precedent	Strengthens or weakens the aggregation hypothesis
Test persistence	Does protein-rich medium alter the state?	Serum or medium stability assessment	Identifies exceptions to simple disaggregation assumptions
Reconcile evidence	Do physical and biological observations converge?	Concordant or discordant findings across prior steps	Proposed trigger for reject, retain, or escalate decisions

Note: The workflow ordering is proposed. No single diagnostic is treated as sufficient, and no universal concentration, particle-size, serum, or detergent threshold is implied.

Figure 1 makes the manuscript-derived sequence from computational suspicion to experimental falsification visible as a process rather than a binary aggregator filter.



A pre-ranking aggregation-risk framework

The proposed framework places aggregation liability before experimental escalation rather than after a mechanistic narrative has already formed. Machine-learning-assisted hit prioritization demonstrates that computational information can reorganize which screening hits receive follow-up [32]. Aggregation risk can therefore be treated as one explicit uncertainty feature in ranking,

alongside potency, novelty, selectivity, and developability, without converting the risk score into a mechanistic verdict. A high-risk natural product should not be automatically discarded; it should be routed toward the experiment most capable of falsifying the aggregation hypothesis.

Three outputs are proposed for each candidate: an aggregation-risk estimate, an applicability-domain

statement, and an evidence-status field. The first summarizes predictive suspicion; the second records whether the compound resembles chemistry on which the predictor was trained; the third distinguishes predicted, experimentally supported, experimentally weakened, and unresolved aggregation. Chemical-space coverage remains a practical constraint for aggregation models [19], particularly for stereochemically dense or unusual natural-product scaffolds. Explainable model outputs may help identify which structural features drove a flag [20], but those explanations remain model-associated rather than causal.

The framework is intentionally asymmetric. A computational warning can justify additional testing, but it cannot establish aggregation. Conversely, experimentally weakening the aggregation hypothesis does not prove the proposed molecular target. It merely removes one important alternative explanation and permits stronger weight to be placed on subsequent target-specific evidence. This asymmetry is important for natural products because unusual scaffolds may fall outside the chemical domain of available aggregation models. An out-of-domain flag should therefore increase uncertainty rather than count as evidence of safety. Likewise, disagreement among predictors should not be averaged into false precision. In the proposed framework, disagreement becomes a reason to privilege direct physical-state testing. The ranking output is thus not a new potency score; it is an evidence-routing device that identifies which uncertainty should be challenged next.

Counter-screening before mechanistic interpretation

Counter-screening should be designed around plausible failure mechanisms rather than performed as a generic confirmatory ritual. Mechanism-aware assay-interference control supports testing the specific artifact capable of generating the observed readout before assigning target-level meaning [1]. For suspected colloids, direct physical-state measurement is especially valuable because it asks whether nanoentities are present rather than inferring them from loss of biological activity; NMR relaxation provides one such route [29].

Positive nuisance controls can reveal whether the assay is vulnerable to known interference behavior [14]. Orthogonal cell-injury profiling can separately challenge generalized toxicity as the source of an apparently coherent cellular response [15]. These controls should be interpreted jointly. Particle detection without altered activity may indicate coexistence of aggregation and genuine pharmacology; altered activity without directly detected particles may reflect an inadequately sensitive physical-state method or a different interference mechanism. The proposed rule is therefore evidential triangulation, not a

single pass/fail counter-screen. Discordant findings should remain unresolved rather than being forced into a binary classification. Repetition under matched conditions, alternative physical-state measurements, or redesigned assay controls may be more informative than simply increasing replicate number in the original format. This preserves the distinction between precision of a readout and validity of the interpretation attached to it.

Incorporating aggregation liability into computational prioritization

The scale of contemporary virtual screening makes this ordering consequential. ZINC-22 illustrates ligand-discovery libraries at multi-billion-compound scale [33], where exhaustive experimental artifact testing before ranking is impossible. Computational liability assessment is therefore most useful upstream, when it can change which molecules are purchased, synthesized, or advanced, and which controls accompany them.

Large-library docking itself can generate experimentally problematic artifacts, reinforcing that a favorable rank is not equivalent to lead validity [34]. Aggregation liability should consequently sit within a broader post-ranking quality layer rather than be confused with docking-score pathology. The proposed implementation preserves the original docking or activity score, adds aggregation-risk and applicability-domain annotations, and modifies the required follow-up evidence instead of applying an irreversible blacklist.

Prospective docking campaigns demonstrate the stronger evidential standard that computational hits can reach when synthesis, pharmacological testing, and mechanistic validation follow ranking [35]. That example does not validate the present anti-aggregation framework. It instead supports the principle that computational plausibility gains scientific weight through independent experimental escalation. For natural products, where isolation effort, scarcity, and structural complexity can make follow-up expensive, directing physical-state testing toward high-risk candidates before intensive mechanism work may improve evidential efficiency without assuming that the ranking scheme itself is prospectively optimized. The same logic applies to machine-learning models trained on historical activity data: an aggregation-liability annotation should alter confidence in the label used for training or benchmarking, not be silently converted into a new definitive class. Prospective testing must therefore evaluate both discovery efficiency and the false-negative cost of deprioritizing unusual chemistry.

Distinguishing colloid formation from poor solubility and precipitation

Colloid formation should not be used as a synonym for poor solubility. Experimental self-assembly studies distinguish soluble species, nanoaggregates, and precipitated material under defined conditions [7]. These states may share upstream physicochemical drivers, yet they differ in particle scale, dynamics, surface behavior, and consequences for assay interpretation. Visible precipitation may reduce free concentration; a colloidal suspension may remain visually unobtrusive while presenting protein-adsorbing surfaces.

Condition-aware solution analysis is therefore more informative than assigning a compound a permanent “insoluble” or “aggregator” label [30]. The manuscript proposes that physical-state classification be tied to concentration, medium, time, and measurement modality. Biological response may also change with self-assembled state [9], so activity should be re-examined when the physical state is experimentally altered. This distinction prevents two common interpretive shortcuts: treating visible insolubility as proof of colloidal inhibition, and treating the absence of visible precipitate as proof that the compound remained molecularly dispersed. Both errors collapse analytically different states into an observation that the experiment did not actually make.

Rescue versus rejection of aggregation-prone scaffolds

An aggregation-supported false mechanism does not necessarily imply chemically worthless material. Small-molecule nanoaggregation can sometimes be transformed deliberately into functional drug-delivery architectures [36]. Such rescue, however, is a new development hypothesis. It cannot retroactively validate the target mechanism inferred from the original screening assay.

State-dependent biological effects further argue against automatic structural blacklisting [9]. A scaffold may be unsuitable for a molecular-target program yet remain interesting for formulation or particle-mediated delivery. Serum stability can also be experimentally relevant when a rescue strategy is pursued [31]. The decision boundary proposed here is therefore explicit: first adjudicate whether the original activity claim survives aggregation testing; only then evaluate whether the scaffold merits an independently designed formulation program. Rescue requires its own evidence for stability, release, exposure, safety, and pharmacological purpose.

Minimum evidence for calling a natural product a genuine lead

A genuine lead should represent more than an assay signal that has acquired a plausible computational story. Mechanism-aware interference testing is required because a primary hit can be generated by multiple assay liabilities [1]. Where aggregation is plausible, condition-matched

physical-state evidence should determine whether nanoentities coincide with the active range; direct methods such as NMR relaxation provide one option [29]. Characterized nuisance controls can additionally probe whether the assay is susceptible to relevant interference mechanisms [14].

The proposed minimum-evidence standard does not require one instrument or universal threshold. It requires reconciliation among chemical identity, reproducible activity, physical state, assay-integrity controls, and orthogonal pharmacology. Only after those layers are mutually compatible should computational binding plausibility be promoted to a lead-level mechanistic hypothesis. Prospective structure-based discovery shows the value of escalating computational predictions through independent experimental validation [35], but does not establish a universal recipe for natural-product programs.

CONCLUSION

Colloidal aggregation is not merely an experimental nuisance that can be appended to late-stage hit validation. It is a physical-state alternative capable of distorting assay interpretation, computational ranking, and the labels from which later models learn. The framework developed here therefore moves aggregation control upstream: computational prediction generates suspicion, condition-matched experiments test the particle hypothesis, and mechanistic interpretation follows only after competing explanations have been challenged.

This ordering deliberately avoids two opposite errors. Treating every aggregation-prone structure as a false lead can discard unusual chemistry, whereas treating every potent, reproducible signal as molecular pharmacology can preserve a particle-generated mechanism. The proposed solution is neither blacklist nor permissiveness, but evidence escalation. Aggregation, precipitation, solubility, generalized nuisance behavior, genuine particle-mediated biology, and target-specific activity remain distinct hypotheses until the data justify collapsing them.

For natural-product discovery, eliminating the particle means eliminating it as an untested explanation—not necessarily eliminating the scaffold. Prospective validation across diverse libraries, assays, and computational domains is still required, including explicit measurement of false-negative costs, resource demands, and performance under natural-product-rich chemical-space shift.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

REFERENCES

- [1] Tan L, Hirte S, Palmacci V, Stork C, Kirchmair J. Tackling assay interference associated with small molecules. *Nat Rev Chem*. 2024;8(5):319-39. doi:10.1038/s41570-024-00593-3
- [2] Glenn IS, Hall LN, Khalid MM, Ott M, Shoichet BK. Colloidal aggregation confounds cell-based Covid-19 antiviral screens. *J Med Chem*. 2024;67(12):10263-74. doi:10.1021/acs.jmedchem.4c00597
- [3] Santillo MF, Paris L, Sprando RL, Glenn I. In vitro biological target screening and colloidal aggregation of minor cannabinoids. *J Nat Prod*. 2024;87(12):2881-6. doi:10.1021/acs.jnatprod.4c00608
- [4] O'Donnell HR, Tummino TA, Bardine C, Craik CS, Shoichet BK. Colloidal aggregators in biochemical SARS-CoV-2 repurposing screens. *J Med Chem*. 2021;64(23):17530-9. doi:10.1021/acs.jmedchem.1c01547
- [5] Reker D, Bernardes GJL, Rodrigues T. Computational advances in combating colloidal aggregation in drug discovery. *Nat Chem*. 2019;11(5):402-18. doi:10.1038/s41557-019-0234-9
- [6] Ganesh AN, Donders EN, Shoichet BK, Shoichet MS. Colloidal aggregation: From screening nuisance to formulation nuance. *Nano Today*. 2018;19:188-200. doi:10.1016/j.nantod.2018.02.011
- [7] Chen C, Wu Y, Wang ST, Berisha N, Manzari MT, Vogt K, et al. Fragment-based drug nanoaggregation reveals drivers of self-assembly. *Nat Commun*. 2023;14(1):8340. doi:10.1038/s41467-023-43560-0
- [8] Shahout F, Vanharen M, Saafane A, Gillard J, Girard D, LaPlante SR. Drug self-aggregation into nano-entities has the potential to induce immune responses. *Mol Pharm*. 2023;20(8):4031-40. doi:10.1021/acs.molpharmaceut.3c00196
- [9] Fioretto L, Ziacco M, Mercogliano M, Gallo C, Nuzzo G, d'Ippolito G, et al. The Janus effect of colloidal self-assembly on the biological response of amphiphilic drugs. *Pharmacol Res*. 2024;208:107400. doi:10.1016/j.phrs.2024.107400
- [10] Duan D, Torosyan H, Elnatan D, McLaughlin CK, Logie J, Shoichet MS, et al. Internal structure and preferential protein binding of colloidal aggregates. *ACS Chem Biol*. 2017;12(1):282-90. doi:10.1021/acschembio.6b00791
- [11] Ghattas MA, Al Rawashdeh S, Atatreh N, Bryce RA. How do small molecule aggregates inhibit enzyme activity? A molecular dynamics study. *J Chem Inf Model*. 2020;60(8):3901-9. doi:10.1021/acs.jcim.0c00540
- [12] Lak P, O'Donnell H, Du X, Jacobson MP, Shoichet BK. A crowding barrier to protein inhibition in colloidal aggregates. *J Med Chem*. 2021;64(7):4109-16. doi:10.1021/acs.jmedchem.0c02253
- [13] Dahlin JL, Auld DS, Rothenaigner I, Haney S, Sexton JZ, Nissink JWM, et al. Nuisance compounds in cellular assays. *Cell Chem Biol*. 2021;28(3):356-70. doi:10.1016/j.chembiol.2021.01.021
- [14] Hu H, Yi X, Xue L, Baell JB. A collection of useful nuisance compounds (CONS) for interrogation of bioassay integrity. *JACS Au*. 2024;4(12):4883-91. doi:10.1021/jacsau.4c00851
- [15] Dahlin JL, Hua BK, Zucconi BE, Nelson SD Jr, Singh S, Carpenter AE, et al. Reference compounds for characterizing cellular injury in high-content cellular morphology assays. *Nat Commun*. 2023;14(1):1364. doi:10.1038/s41467-023-36829-x
- [16] Yang ZY, Yang ZJ, Dong J, Wang LL, Zhang LX, Ding JJ, et al. Structural analysis and identification of colloidal aggregators in drug discovery. *J Chem Inf Model*. 2019;59(9):3714-26. doi:10.1021/acs.jcim.9b00541
- [17] Alves VM, Capuzzi SJ, Braga RC, Korn D, Hochuli JE, Bowler KH, et al. SCAM Detective: Accurate predictor of small, colloiddally aggregating molecules. *J Chem Inf Model*. 2020;60(8):4056-63. doi:10.1021/acs.jcim.0c00415
- [18] Lee K, Yang A, Lin YC, Reker D, Bernardes GJL, Rodrigues T. Combating small-molecule aggregation with machine learning. *Cell Rep Phys Sci*. 2021;2(9):100573. doi:10.1016/j.xcrp.2021.100573
- [19] Abou Hajal A, Bryce RA, Amor BB, Atatreh N, Ghattas MA. Boosting the accuracy and chemical space coverage of the detection of small colloidal aggregating molecules using the BAD Molecule Filter. *J Chem Inf Model*. 2024;64(13):4991-5005. doi:10.1021/acs.jcim.4c00363
- [20] Sturm H, Teufel J, Isfeld KA, Friederich P, Davis RL. Mitigating molecular aggregation in drug discovery with predictive insights from explainable AI. *Angew Chem Int Ed Engl*. 2025;64(29). doi:10.1002/anie.202503259
- [21] Wang L, Ji B, Zhai J, Wang J. Advancing promiscuous aggregating inhibitor analysis with intelligent machine learning classification. *Brief Bioinform*. 2025;26(3). doi:10.1093/bib/bbaf205
- [22] Shi S, Fu L, Yi J, Yang Z, Zhang X, Deng Y, et al. ChemFH: An integrated tool for screening frequent false positives in chemical biology and drug discovery. *Nucleic Acids Res*. 2024;52(W1). doi:10.1093/nar/gkae424

- [23] Alves VM, Yasgar A, Wellnitz J, Rai G, Rath M, Braga RC, et al. Lies and liabilities: Computational assessment of high-throughput screening hits to identify artifact compounds. *J Med Chem.* 2023;66(18):12828-39. doi:10.1021/acs.jmedchem.3c00482
- [24] Palmacci V, Nahal Y, Welsch M, Engkvist O, Kaski S, Kirchmair J. E-GuARD: Expert-guided augmentation for the robust detection of compounds interfering with biological assays. *J Cheminform.* 2025;17(1):64. doi:10.1186/s13321-025-01014-3
- [25] Béquignon OJM, Bongers BJ, Jespers W, IJzerman AP, van der Water B, van Westen GJP. Papyrus: A large-scale curated dataset aimed at bioactivity predictions. *J Cheminform.* 2023;15(1):3. doi:10.1186/s13321-022-00672-x
- [26] Zdrazil B, Felix E, Hunter F, Manners EJ, Blackshaw J, Corbett S, et al. The ChEMBL Database in 2023: A drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Res.* 2024;52(D1). doi:10.1093/nar/gkad1004
- [27] Parrondo-Pizarro R, Menestrina L, Garcia-Serna R, Fernández-Torras A, Mestres J. Enhancing molecular property prediction through data integration and consistency assessment. *J Cheminform.* 2025;17(1):163. doi:10.1186/s13321-025-01103-3
- [28] Nelson KM, Dahlin JL, Bisson J, Graham J, Pauli GF, Walters MA. The essential medicinal chemistry of curcumin. *J Med Chem.* 2017;60(5):1620-37. doi:10.1021/acs.jmedchem.6b00975
- [29] Ayotte Y, Marando VM, Vaillancourt L, Bouchard P, Heffron G, Coote PW, et al. Exposing small-molecule nanoentities by a nuclear magnetic resonance relaxation assay. *J Med Chem.* 2019;62(17):7885-96. doi:10.1021/acs.jmedchem.9b00653
- [30] LaPlante SR, Roux V, Shahout F, LaPlante G, Woo S, Denk MM, et al. Probing the free-state solution behavior of drugs and their tendencies to self-aggregate into nano-entities. *Nat Protoc.* 2021;16(11):5250-73. doi:10.1038/s41596-021-00612-3
- [31] Ganesh AN, Aman A, Logie J, Barthel BL, Cogan P, Al-Awar R, et al. Colloidal drug aggregate stability in high serum conditions and pharmacokinetic consequence. *ACS Chem Biol.* 2019;14(4):751-7. doi:10.1021/acscchembio.9b00032
- [32] Boldini D, Friedrich L, Kuhn D, Sieber SA. Machine learning assisted hit prioritization for high throughput screening in drug discovery. *ACS Cent Sci.* 2024;10(4):823-32. doi:10.1021/acscentsci.3c01517
- [33] Tingle BI, Tang KG, Castanon M, Gutierrez JJ, Khurelbaatar M, Dandarchuluun C, et al. ZINC-22—A free multi-billion-scale database of tangible compounds for ligand discovery. *J Chem Inf Model.* 2023;63(4):1166-76. doi:10.1021/acs.jcim.2c01253
- [34] Wu Y, Liu F, Glenn I, Fonseca-Valencia K, Paris L, Xiong Y, et al. Identifying artifacts from large library docking. *J Med Chem.* 2024;67(18):16796-806. doi:10.1021/acs.jmedchem.4c01632
- [35] Liu F, Wu CG, Tu CL, Glenn I, Meyerowitz J, Kaplan AL, et al. Large library docking identifies positive allosteric modulators of the calcium-sensing receptor. *Science.* 2024;385(6715). doi:10.1126/science.ado1868
- [36] Kim JH, Kulthe AD, Park SJ, Raj V, Lee WR, Nimse SB. Transforming small-molecule nanoaggregation into functional drug delivery platforms. *J Med Chem.* 2025;68(10):10384-98. doi:10.1021/acs.jmedchem.5c00562