

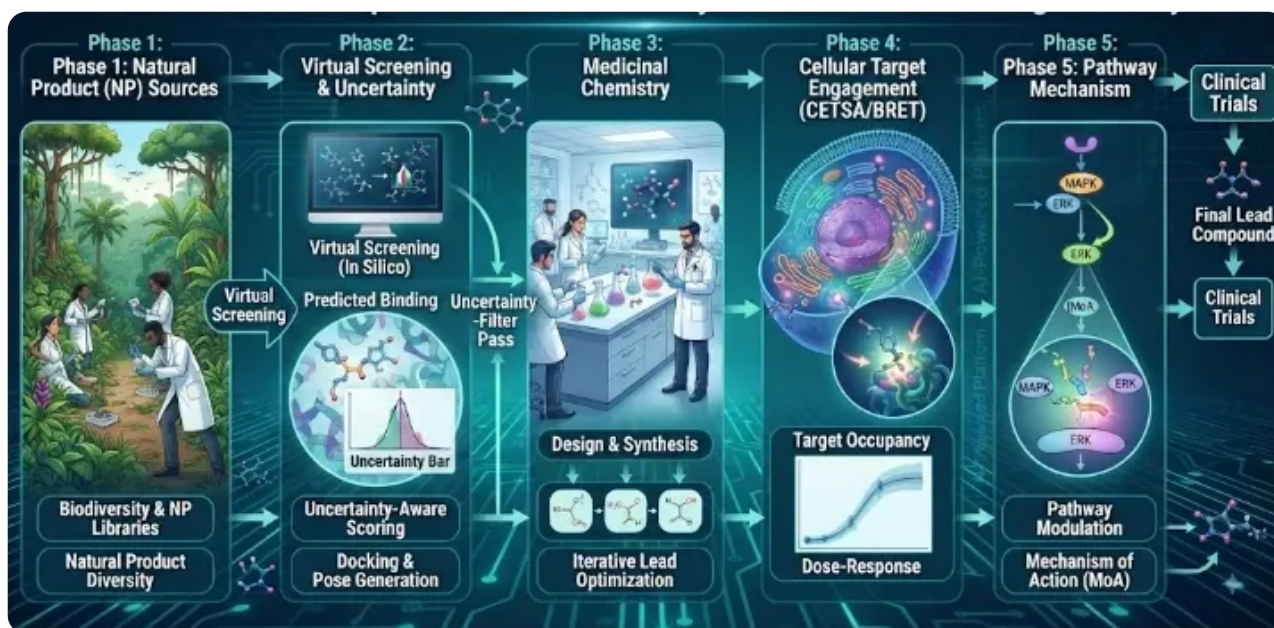
From Predicted Binding to Cellular Target Engagement: An Uncertainty-Aware Platform for Natural-Product-Inspired Medicinal Chemistry and Small-Molecule Drug Discovery

Md. Humayan Kabir

Mar 1, 2026

Abstract

This article examines how uncertainty-aware target validation can connect computational binding predictions with real cellular target engagement in natural-product-inspired medicinal chemistry. Complex natural-product scaffolds often contain unusual stereochemistry, conformational rigidity, and functional groups that challenge simplified docking and two-dimensional prediction methods. The article therefore separates predicted affinity from permeability, intracellular concentration, target occupancy, binding mode, exposure duration, and downstream pathway response. Thermal proteome profiling, chemoproteomics, cellular thermal shift assays, competition experiments, imaging, intact-protein mass spectrometry, genetic perturbation, and resistant mutants are presented as complementary ways to test whether a computationally prioritized interaction truly occurs in cells. Docking, molecular dynamics, graph learning, ensemble prediction, applicability-domain analysis, and conformal methods provide useful ranking power only when their uncertainty is visible. The article also integrates resistance, pharmacokinetics, off-target selectivity, manufacturing quality, and clinical practicality. Its central argument is that computational ranking becomes valuable when it is converted into a falsifiable biological program in which exposure, engagement, phenotype, and mechanism are experimentally connected.



The interaction prediction is just the first step in the mechanistic part of a natural-product-inspired drug discovery claim. However, the uncertainty-aware target validation in natural product inspired medicinal chemistry and in small molecule discovery is now no longer just about hit finding – when confronted with complex scaffolds and small amounts of training data, the system may provide confident binding predictions, but these are not actually in line with target engagement within the cell. The more important question is whether such a computationally prioritized interaction that is confirmed within relevant cells under pharmacologically realistic exposure is stable under realistic biological and developmental conditions. The turnover is from ranking to mechanistic program: discovery based on exposure, causality and reproducibility. As the publication deadline approaches, success in target-oriented natural-product discovery more and more relies on the ability to correlate molecular design with information obtained from experiments needed to understand the mechanism(s) for successful interactions with the target, as well as failures and constraints to be prioritized.

Natural-product molecules are often stereochemical, have unusual functional groups which are difficult for simplified docking and 2D prediction methods, and are also rigid in 3D which makes them difficult to dock. The history, and recent experience, indicate that a nice laboratory signal doesn't usually remain the same as the complexity of the biological system increases. A useful research methodology is to maintain the original activity signal whilst varying progressively the conditions under which the tests are conducted, to determine what features are still necessary.

The first step in mechanistic depth is to distinguish affinity from permeability,

intracellular concentration, target occupancy, covalent or reversible engagement, and response to the target in the pathway. This structure is not apparent in simple endpoint assays because it may happen through a variety of causal pathways and yield similar readouts. These methods allow for a more comprehensive understanding of the mechanisms of action beyond direct correlation with the downstream adaptation or other factors, such as non-specific stress or delivery failure: thermal proteome profiling, chemoproteomics, cellular thermal shift assays, competition studies, imaging and functional rescue. This type of measurement creates a beautiful response that can be tested and used to establish a viable link between chemistry, biological state and functional outcome.

Also, design variables cannot be considered independently, but as interacting variables. The effects of protonation, stereochemistry, conformation, solubility, permeability, subcellular localization, target abundance, and exposure time are intertwined and enhancement in one dimension may come with a price tag in another. A practical approach to optimization is to consider trade-offs rather than try to maximize the single goal. Where potency, selectivity, stability, tissue exposure and manufacturability all impact on translational value, multi-parameter profiles are particularly useful.

Computational leverage can be introduced in the form of docking, molecular dynamics, graph learning, ensemble prediction, applicability domain analysis and conformal methods, but prediction is only helpful when uncertainty is perceptible. Even with similar regions or only subtle differences in chemical or biological properties, trained models can fail when presented with scaffolds that differ or with unusual materials or underrepresented patient states. Expensive experiments do not need to be run until the models are calibrated, prospective testing, domain-of-applicability checks and disagreement among model families can help to reveal fragile predictions.

Orthogonal validation is a good method to use in experimental design. Intact-protein mass spectrometry, orthogonal biophysics, genetic perturbation, resistant mutants, probe competition, and dose-response engagement are all complementary ways of addressing various aspects of the causal question, and diminishing reliance on any single assay technology. Consensus between the chemistry, imaging, omics, genetics and functional biology builds one of the best mechanistic arguments than repeated measurement of the same endpoint. Conflict is also useful, as it can lead to discovery of obscure exposure thresholds, assay artifacts, compensatory pathways, or context-dependent pharmacology.

There are additional challenges with biological diversity. Target abundance can vary in the

state of protein, lineage of cell, metabolism, expression of transporters, target novelty, and immune tone without a change in nominal treatment, among other factors. Average responses may therefore obscure those response differences and exposure-response relationships among subgroups. Biological variation is a part of the design and not an explanation after the fact of inconsistent results.

Early attention should be paid to resistance and adaptation because the target may appear to be active but may actually be being missed due to target mutation, pathway redundancy, altered uptake or metabolic inactivation. Activity can be determined through the following methods: target mutants, resistant clones, sequencing, pathway rescue, and cross-resistance testing. A lower potency, but blocking common escape routes, may be more valuable than a fragile, high potency lead. Resistance-aware development also helps to focus mechanism since escaped mutations in recurrent escape changes often have the most causal weight in the proteins, pathways, or exposure processes.

Molecular promise to reachable biology – delivery and pharmacokinetics. The active form needs to be present at the right concentration and time at the intended site which depends on its free intracellular concentration, plasma binding, membrane permeability, presence of transporters, metabolism, and subcellular distribution. This relationship cannot be described using the nominal dose as it is modified by absorption, distribution, binding, transformation, clearance and barriers in the local tissue. Mechanistic exposure measurements should then be included with initial efficacy data from the earliest development stages.

Considering mechanism in safety assessment adds to the information. Distinguishing target-linked pharmacology from broad cellular injury and immune activation, reactive chemistry or formulation-driven toxicity is achieved using proteome-wide selectivity maps, healthy-cell engagement, reactive chemistry screens, organotypic assays, and exposure-margin analysis. Drug activity is determined against disease and healthy system comparison testing to be able to distinguish therapeutic activity from unwanted effects. This also allows a good efficacy signal to get further along a program without the parameter being inexplicably reduced by a similar mechanism.

Another scientific variable, not a downstream logistics issue is manufacturing. When the nominal composition is the same, the product characteristics, biological activity and exposure may vary because of changes in stereochemical purity, polymorphism, control of impurities, repro-ducibility of the route, solvate formation, and storage stability. Quality-by-design thinking correlates important material characteristics and process parameters

with measurable performance, rather than just at the end of the production process. Laboratory success is more believable with variations of raw material, larger equipment, storage, transportation and independent production batch runs.

Last but not least, realistic models of clinical use are needed for translation. The timing, dose frequency, route, complexity, diagnostic needs, and acceptable complexity are influenced by the availability of the biomarker, the prevalence of the target, the tissue penetration, the route of administration, duration of the treatment, and the presence of combination setting. While a technically beautiful platform is valuable, it becomes useless when it requires a great amount of administration, storage space, monitoring or costs are in conflict with the context where treatment is required. Thus, decisions about the development of the product provide a great advantage by considering both the mechanics and the deployability of the product before it is no longer easy to change course.

Data integration is a link between these layers. These affinity, confidence scores, target occupancy, proteomic selectivity, intracellular exposure, phenotype and resistance score outputs can be mapped to a common causal model instead of being stored as separate assay outputs. Iterative learning is reproducible thanks to traceable metadata, versioned models, standardized controls, and prospective decision rules. The end product is a development transcript that offers insight into which candidate did best, as well as which assumptions led to that determination and which uncertainties were not answered.

Strong platforms compare exposure-adjusted engagement, off-target burden, and phenotype rescue as well as prospective enrichment and consistency of exposure to the target. Strong platforms compare with the caliber error, prospective enrichment, target-occupancy consistency, phenotype rescue, off-target burden, and exposure-adjusted engagement, not docking rank. Success measures such as manufacturing robustness, biological breadth, exposure margin, durability, and reproducibility, as well as traditional potency, boost decision quality. These types of metrics will prevent optimization of a single assay and will favor candidates that are still "plausible" in separate assays. A clear score card too explains why a lesser dramatic outcome can be the better development option.

A good programme also sets a positive framework for failure. Negative results from (high-scoring) non-binders, cell-impermeable compounds, broad proteomic binders, unstable chemotypes and targets that are unrelated to phenotype define boundaries in the mechanism and limit re-investment in poor hypotheses. If the synthesis fails or the target is not engaged, exposure cannot be guaranteed, resistance emerges and/or the effects are

lost for a particular subpopulation, then the search space will be better represented going forward. Successful experiments teach as much as unsuccessful ones; and thus discovery is made more quickly.

There is a considerable amount of uncertainty in this process of target engagement that raises the computational ranking to a falsifiable biological program and not a replacement for mechanism. The main strength of the central advantage is that it is built on a series of explanations that are integrated across the issues of mechanism, exposure, biological diversity, safety, and production. That structure helps in the selection of the candidates because for each important development claim there is a test. The progress is no longer so much a series of amazing individual experiments, but a consistent series of results that are convincing, yet still carry over from the lab to real clinical situations.