

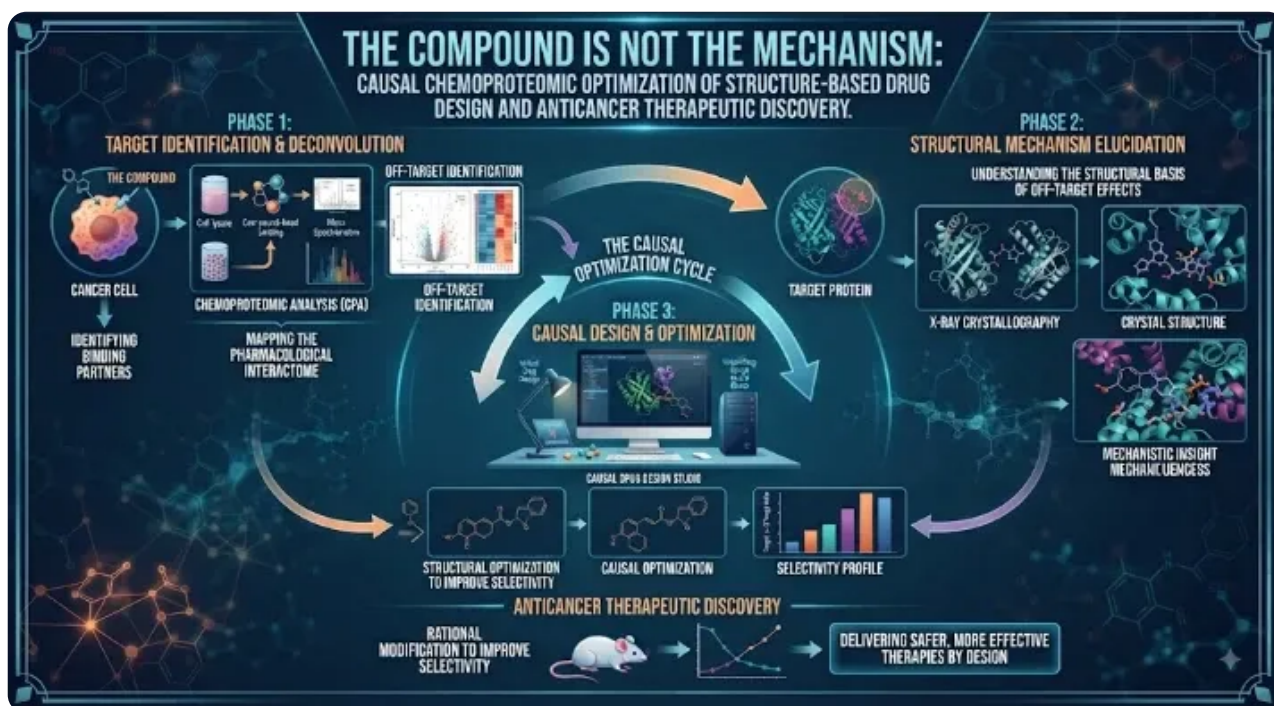
The Compound Is Not the Mechanism: Causal Chemoproteomic Optimization of Structure-based drug design and anticancer therapeutic discovery

Md. Humayan Kabir

May 11, 2025

Abstract

This article explains why a compelling structural binding pose does not prove that the same interaction drives anticancer activity inside living cells. It proposes causal chemoproteomics as a bridge between predicted binding, proteome-wide engagement, pathway perturbation, genetic dependency, and phenotype. CETSA and thermal proteome profiling can detect ligand-induced stability changes, activity-based protein profiling can map reactive engagement, and affinity or photo-crosslinking approaches can capture targets when probe design preserves parent-compound behavior. Kinase-focused platforms can also reveal broader cellular target landscapes than biochemical panels suggest. Probe design, free intracellular concentration, engagement kinetics, subcellular localization, covalency, target abundance, off-target profile, and cancer genotype are treated jointly. Resistance routes include alternative targets, polypharmacology, target mutation, bypass signaling, proteostasis adaptation, transporter changes, and lineage plasticity. Proteome-wide selectivity, healthy-cell engagement, reactive metabolites, mitochondrial effects, and exposure are incorporated into safety assessment. The article ultimately requires inactive probes, competition, target mutation, genetic rescue, orthogonal occupancy assays, and matched phenotypic response before a structural hypothesis is accepted as a causal mechanism.



A central issue in causal chemoproteomics in structure-based anticancer drug discovery is addressed in the Compound Is Not the Mechanism: a structure that is a good fit for the receptor does not necessarily mean that the same interaction is responsible for the activity within a live cancer cell. The good development strategy is based on a measurable mechanism – not just a single outstanding assay result. The change creates a new translation, examining chemistry, biology, exposure, resistance, safety, and production as one argument.

The scientific basis is based on a process that involves the use of causal chemoproteomics (predicted binding, proteome-wide engagement, pathway perturbation, genetic dependency, and phenotype). Everything is a piece of the puzzle, and none of the components on their own is enough to support the entire mechanistic statement. The connection of evidence mitigates the risk of mistaking potency, binding, formulation performance and/or model fit for understanding of a therapeutic relationship.

Structural hypotheses can now be tested in the cell using several means of modern chemoproteomics. The stability change (CETSA), the binding of the probe protein to its target can be reactive (ABPP), and affinity or photo-crosslinking can be used to capture the target if the target binding sequence of the probe is preserved (Target Design). Another example, though, of the different types of target landscapes of nominally selective inhibitors is provided by Kinase focused platforms like Kinobeads.

Concurrently, the performance of a probe can vary across design variables, such as the

design of the probe, the concentration of the free probe, how the probe engages the target, the location in which the probe is active within the cell, abundance of the target, covalency of the probe to the target, off target profile, and cellular genotype. Adopting optimization of a single parameter can impose an unseen penalty on other parameters, such as potency vs stability, potency vs selectivity, potency vs delivery or potency vs manufacturability. Considerable more realistic ranking of candidates can be achieved using multiparameter decision rules than a single numerical endpoint.

Having early measurements of failure routes is also a prerequisite for mechanistic confidence. There are important escape mechanisms such as: alternative targets, polypharmacology, target mutation, bypass signaling, proteostasis adaptation, transporter changes and lineage plasticity. Stress testing intentionally makes these routes into design information, which can be used to inform the next chemistry, formulation, dosing or model cycle.

A link between nominal dose and biological effect is exposure. The pertinent parameters are: uptake, unbound intracellular concentration, subcellular distribution, residence time of the target, metabolism and tissue pharmacokinetics. A candidate that is active in a simple assay but not active at the causal site can appear as a great candidate for assay, but be unsuitable for translation.

When interpreting the engagement or functional perturbation, it is recommended to do so at realistic concentrations. Side effects, off-target effects or assay artefacts may occur at high doses and give a false positive effect of the desired mechanism. Concentration-response analysis that correlates exposure and occupancy with the proposed causal process to discern the range in which the proposed process is likely to occur.

The biological context also influences interpretation of the properties of anticancer compounds identified by structural modeling and optimized in disease relevant cellular systems. Target dependence and effective exposure can be altered by differences in genotype, cell state, tissue architecture, immune tone, metabolic environment and previous treatment. Repeated testing in one convenient system is not as compelling as representative model panels.

Resistance and adaptation should be measured on the prospective and not the retrospective basis. Further selection, serial passage, genetic manipulation, repeated dosing or environmental stress can identify which mechanisms are able to withstand ongoing selection. A slightly weaker, but more blockbuster escape resistant candidate can have greater translational value than a weak screening leader.

A mechanism is a key element in safety assessment, especially to make it more informative. Some major issues are the proteome-wide selectivity, healthy cell engagement, formation of reactive metabolites, mitochondrial stress and engagement of essential off-target proteins. Mechanism-related pharmacology should be compared with unwanted biological responses to differentiate it from general cellular damage or formulation-related toxicity.

No longer can manufacturing be a downstream logistics exercise as analogue synthesis, probe reproducibility, linker effects, compound stability, purity, stereochemistry, and scalable medicinal chemistry routes all impact the compound which enters into biological systems. Changes in the size of the process can affect the product composition, physical state, impurity level, product stability, and/or the dose-reproducibility of the product even though the product nominal name is the same. The critical process attributes are directly connected to the mechanistic performance, adding robustness to development.

Proteomic dose-response modeling, target deconvolution, network propagation, structure-based analogue ranking, and causal integration of multi-omic evidence are tools applied to computational challenges and give a boost of leverage. Prediction is most useful when there are areas of uncertainty and applicability limits that are still visible, particularly in the vicinity of novel chemical structures, unusual biological states, and areas with poor data coverage. Models that allow for uncertainty will lead to better experiment selection than models that simply rank experiments with a high degree of confidence without revealing where their knowledge is low.

Prospective validation requires that the probes be inactive, there be competition experiments, target mutation, genetic knockout or knockdown, rescue, orthogonal occupancy assays, and matched response to the target. Repeated measurements in the same family of assays would introduce the same 'hidden bias' without providing greater evidence of causation because orthogonal tests are important. If the results are in agreement on independent designs it is a more robust chain from design choice to biological consequence and if they disagree, then there is an assumption to be modified.

Data transfer should maintain traceability of data from raw experiment to development decision. Chemical structures, process conditions, assay metadata, exposure measurements, omics features, resistance observations, and safety signals should all be versioned and have stable identifiers. This discipline helps to avoid subsequent interpretations getting "off the message" from what the result was.

The measurement of a decision must also account for translation and not laboratory

showmanship. Useful scorecards include not only reproducibility, but the amount of exposure normalized by activity, the mechanistic confidence, resistance frequency, safety margin, stability, and manufacturing robustness. This way, those who have a balanced approach to their candidates are rewarded, and compromises become apparent before a lot of costly development costs are piled up.

In this context especially negative results have high values. This boundary of a hypothesis is established by the failure to synthesize, failure to achieve a target engagement, unstable exposure, no rescue, emergence of resistance, lack of batch reproducibility, and loss of activity in realistic models. By capturing those failures, experimental design for future experiments is improved and the investment of time into appealing, but weak, explanations is avoided.

Reproducibility of a laboratory to laboratory and also batch to batch is another key filter. Even though statistical significance appears robust, a mechanism which relies on one operator, one preparation, one cell line, or one modeling choice is fragile, since that may be the only one sufficient to obtain significance. When conditions are matched, independent repeating differentiates repeatable biology from local experiments.

There is also a practical aspect of clinic or deployment that needs to be taken into account when doing translation. Technically sophisticated products may have a realistic value based on route, administration burden, diagnostic requirements, storage, monitoring, cost and treatment timing. When product design is designed to have a biological effect, it needs to be mechanistic excellence to make a difference.

Another added layer is temporal design which involves the response of biological systems, specifically short pulses, long exposure, washout, repeated dosing. Time-resolved measurements can distinguish between early activity that is due to engagement from later activity that is due to adaptation, and can also provide some indications of whether early activity continues after the initiating signal has been extinguished. The sampling schedule should thus reflect the expected kinetics of the causal chemoproteomics in the structure-based discovery of anticancer agents, not one convenient endpoint.

An integrated development record aids in improved communication between medicinal chemistry, formulation science, pharmacology, computation, manufacturing and clinical planning. The record does not have to tell one single story, but can differentiate between known connections, likely assumptions, unknowns, and the next experiment with the most information content. This transparency makes the iterations quicker and lessens the pressure to justify initial statements when conflicting data comes up.

Novelty is kept in check by a benchmarking against credible reference conditions. Whether a new approach will result in better prediction, better mechanism, longer duration or better manufacturability is determined by established compounds (inactive controls), standard formulations, baseline models, or accepted analytical methods. A head-to-head comparison also allows gains to be identified that are lost following exposure, assay sensitivity, batch sensitivity or experimental complexity normalization.

The most effective way of anticarcinogenic action is the sequence of chemistry-engagement-pathway-control-tumour phenotype. It is not simply a question of whether a candidate succeeds in one assay but whether he/she succeeds in all assays and whether the mechanism is coherent with regard to exposure, biological variation, resistance, safety, and production. That standard establishes a more disciplined process of discovery signal to a therapeutic concept that can withstand the complexity of the real world.