AI in Drug Discovery and Development

Prof. Rajnish Kumar
Dept. of Pharmaceutical Engineering and Technology
IIT-(BHU), Varanasi

Week-10
Lecture-49

Welcome to the course "AI in Drug Discovery and Development." In this session, we will talk about integrating multi-target drug discoveries. So, by the end of this lecture, you will be able to explain the rationale behind multi-target drug design and its advantages over traditional single-target strategies. Describe the evolution of multi-target-directed ligands and the shift toward polypharmacology. Compare and apply the primary medicinal chemistry strategies for MTDL development, and identify the computational approaches used in multi-target drug discovery. Interpret chemogenomics strategies for multi-target selection and validation, as well as evaluate the role of AI in designing, optimizing, and predicting multi-target drug candidates.

So, what is multi-target drug design? So, multi-target drug design refers to the process of designing and developing a single therapeutic agent that modulates two or more molecular targets involved in a disease network. With the goal of achieving enhanced efficacy, reducing side effects, and overcoming drug resistance compared to traditional single-target drugs. So, traditionally, what has been thought is that there is only one, you know, target in a disease, and one drug can engage the target and treat the disease. However, it is not true because many of the diseases you know are very complex in nature, such as when you talk about Alzheimer's disease, which is a neurodegenerative disease.

So, there are multiple pathologies working at the same time for the disease. In that case, targeting a single pathology will not be able to get rid of the disease. So, that is why we need the multi-targeting approach. So, the key characteristics of multi-target drug design are that it is polypharmacology-based and utilizes the principle that drugs often interact with more than one target. And then it is network-centric as well, focusing on disease-related networks and pathways rather than isolated targets.

and then balanced affinity where we aim for optimal, but not necessarily maximum binding to multiple targets. So you can see here, for example, there is one target, and the drug is hitting one target; however, in multi-target hitting, you have one drug that is acting on multiple targets. You can also see that there is a drug molecule interacting with protein A as well as protein B, where proteins A and B are interlaced in this complex network, affecting each other's actions and roles. So, that is how, if we wanted to, this could be a pathway for a disease. So, we have to block several of those points to get rid of the

symptoms and the disease as well.

So, if we talk about the history and evolution of multi-target-directed ligands. So, the early discovery of the initial MTDLs emerged serendipitously from high-throughput screening and not from computational design. So, people were screening compounds and then they found that the single molecule was inhibiting multiple targets. So, that was, you know, with the high-throughput screening, and then, however, the later use of X-ray crystal structures. Detailed structural equity relationships and computational chemistry enable the rational design of potent, selective multi-target drug ligands.

And it is obvious as well because many times those targets are homologous to other targets as well. They have sequence similarity, and thus their structures are also quite similar. So, it is highly likely that if a molecule is inhibiting one target, it is also inhibiting another. So, because of the similarity of that target to another target, we are going to hit the other target as well. So then, there was this kinase inhibitor milestone: since the approval of imatinib in 2001, 76 kinase inhibitors have been FDA-approved, and over 180 small molecule kinase inhibitors are in late-stage clinical trials.

So, that is from the single to the multiple target test shown, and it has also led to the rise of polypharmacology, where a single drug is acting on multiple targets. So, if you look at the systems network view of memantine's action, memantine is, you know, a drug to treat Alzheimer's disease. It is an FDA-approved anti-Alzheimer's drug. So, its primary target is, you know, NMDA; it acts as an NMDA antagonist. So, it is an uncompetitive low-affinity NMDA antagonist.

A little selectivity among the subtypes of NMDARs, which have fast on-off binding kinetics, exists; however, it also has secondary targets on which it acts, including the 5HT, nicotinic acetylcholine receptor, D2 dopamine receptors, serotonin, alpha 1 receptors, and voltage-activated sodium ion channels. And then it shows that by engaging with the primary and secondary targets, it demonstrates multiple activities, such as action against A-beta toxicity, inhibiting tau phosphorylation, reducing neuroinflammation and oxidative stress, and promoting neurogenesis effects as well. So, it acts to relieve the symptoms of Alzheimer's disease through its interaction with the primary and secondary targets. So, that then leads to therapeutic effects and well-tolerated side effects. Why do we need those, you know, multi-targeted drug lines? What could be the limitations of the one drug, one target paradigm? So, the first thing is the biological complexity of the disease, as I said, like many diseases, such as cancer, Alzheimer's, or diabetes.

So, they involve multiple pathways and interacting proteins, and targeting only a single protein often fails to modulate the entire disease network. For example, AD involves

ACHE, BACE-1, tau, NMDA, and oxidative stress; all of these contribute to the pathology; it's a multifactorial disease. And then there are, you know, compensatory mechanisms; biological systems can adapt via feedback loops or pathway cross-links, so inhibiting one target may lead to the upregulation of an alternative pathway, thereby reducing drug efficacy. It is like, because those cells are very smart by nature, right? So, if you block one pathway, the cell will activate another pathway to perform the same function. This leads to drug resistance or ineffectiveness of that drug molecule.

So, to beat that, you need to have molecules that can, you know, inhibit multiple targets at the same time. The limited efficacy is another important limitation where mono-target drugs may show initial benefits, but long-term outcomes are often suboptimal. A single-target approach might not cover all the disease drivers, especially in heterogeneous patient populations. And then drug resistance occurs when pathogens and cancer cells can mutate or bypass single target blockage, leading to resistance, and multi-targeting drugs reduce this risk by attacking multiple survival pathways simultaneously. Poor translation from bench to bedside, where a drug showing high in vitro potency on one target may fail in vivo due to system-level interactions that are not being addressed.

And then polypharmacology is natural; many approved drugs already interact with multiple targets, such as clozapine, aspirin, and imatinib. So, recognizing and leveraging this design of polypharmacology is more effective than trying to avoid it. So, coming to the rational design of multi-targeted ligands, commonly used strategies for obtaining MTDLs are merging, fusion, and linking two or more pharmacophores derived from known ligands of different targets. So you have, for example, these two drugs, drug A and drug B; these bind to two different targets. So one option is that I merge these, you know, molecules, which means I have some parts that are common between these two molecules.

So, I make another molecule that is, you know, a hybrid of these two molecules, and then this hybrid molecule has the possibility to bind to both targets. So, that is called merging. Another one is fusion, where I just attach these two molecules through a bond, and then that is called fusion. Thus, this part of the molecule will bind to one target, and this part of the molecule will bind to another target. And then another possibility is, you know, linking; it can be linked with a flexible, non-cleavable linker.

So, it will act as a chimera, where this part acts on one target, this part acts on another target, and this can even be a cleavable linker as well. Where it can get cleaved inside the cell, then they can get separated, and then they perform their action of binding to the receptors or targets for which they are designed. So, these are, you know, some of the strategies that are being used to design those multi-targeted ligands. So, in the merging approach, overlapping pharmacophoric elements from two ligands are integrated into a

single, compact molecule. and this strategy aims to preserve critical interactions with both targets using a minimalistic design improving molecular weight, lipophilicity and drug likeness.

So, however, it requires a detailed structure-activity relationship, pharmacophore mapping, and knowledge of shared binding-site features. So, the two strategies that are commonly used for merging are merging similar functional groups or merging similar skeletal groups. This is an example where similar functional groups have been merged; for example, this is one molecule that has, you know, an HDAC zinc-binding head group. So, then you have the lovastatin HMGR inhibitor, and this one was, you know, an HDAC inhibitor, and then using this group. This zinc-binding head group, which also shares similarities with the lovastatin functional group, is important.

So, by merging these two, you get a new molecule called a dual HDAC-HMGCR inhibitor. So, which has this zinc-binding group as well as the, you know, HMGR binding group, and then merging the similar skeletal groups. So, in this case, you have, for example, an example of a molecule that is an NAMPT inhibitor. So, in this case, you can see that it has this core which is, you know, responsible for the activity, and then you have another molecule, 21, which is CI994, an HDAC1 inhibitor with an IC50 value of 1.

2 micromolar. And then it has this functionality, which is the core that interacts with the receptor and has the activity. So, by merging the similar skeletal groups from these two, it led to the formation of this molecule 22. And then that had a good activity against NMPT as well as HDAC1, and after optimization, you get a compound that is optimal and binds to both of these targets. And then you have the fusion approach, where fusion involves combining two pharmacophores through a shared rigid scaffold that allows for interaction with both targets. So, it creates a structurally integrated system in which both pharmacophores are attached to a central molecular core.

So, you have, for example, rasagiline, which is an MAO-B inhibitor, and then you have rivastigmine, which is a cholinesterase inhibitor. And then you have M 30, which is a MAO A or B inhibitor, by just fusing all of them together. So, what you get is a MAO ACHE inhibitor and a pro-chelator molecule M 30 D. This design strategy results in multi-target molecules like M 30 D, which contain both a propargyl amine moiety. So, this propylamine is responsible for, you know, MAO A and B inhibitory activities.

And then it also has this dimethyl carbamoyloxy moiety, the red part, which comes from, you know, rivastigmine, which is an anti-AD drug. And then we have an example of a linking approach in which two distinct pharmacophores are joined by a flexible or rigid chemical linker. and this method preserves full pharmacophoric identity of each component

allowing independent binding to each target. So, the linker is crucial. Its length, flexibility, polarity, and spatial orientation affect the activity and pharmacokinetics.

So, you have galantamine, which is an AChE inhibitor, and you have memantine, which is an NMDA antagonist. So, by combining these two with a linker, you link them with a flexible alkyl linker. So, you get another compound, compound 3, which is, you know, a dual-acting drug that binds to both the ACHE and the NMDA. So, linking two FDA-approved drugs for AD leads to a new bifunctional drug candidate that acts against both ACH and the NMDA receptor. Okay, so that was, you know, the medicinal chemistry approaches and how we are using those methods to design multi-targeted drug ligands.

So let us take a look at the computational methods for multi-target drug design. So again, these can actually be classified into these four quadrants. So the top right one is the target-centric, where we can use de novo drug design, ligand-based drug design, and we can also use multiple pharmacophores and a common pharmacophore approach. And then we can see on the left-hand side that we can use the ligand-based methods, which are ligand-centric, where we use similarity search or machine learning methods. And then, we can use target-centric ligand-based methods, like de novo design, or we can use those pharmacophoric models, actually.

And then, if we are using the structure-based design, we can use target fishing, docking of pharmacophore modeling, or de novo drug design, which is, of course, structure-based. We can also use multiple docking, where we dock those molecules against multiple targets and then select the molecules that have good activity against those targets. So, this pharmacophore and docking-based multi-target drug design works like this. So, we generate a pharmacophore of protein A, then we generate the pharmacophore of protein B, and then we generate the pharmacophore for protein C. And then we are merging these three pharmacophores and trying to identify a common pharmacophore.

And once we have identified this common pharmacophore, what we can do is use this pharmacophore to identify ligands. And then we can either, you know, go for docking-based filtration or we can go for shape-based filtration. Then select the candidates and conduct the bioassay to ensure that the compounds we have identified are showing activity. Okay, so this is an example of a common pharmacophore-based design for a multi-target inhibitor of HNPSPLA2 and LTA4H. So where compound one fit a common pharmacophore generated by these two targets and the binding confirmation was confirmed using a docking method.

And then, compounds 1 and 2 also showed the desired multi-target activities. So, you can see that both compounds 1 and 2 have activity against these targets. So, this was how they

conducted the activity because both of these molecules showed similarity with the common pharmacophores. So, these were, you know, the pharmacophores for Target 1 and Target 2, and then this was the common pharmacophore. And then, compounds 1 and 2 showed similarity with the common pharmacophore, which led to the identification of these compounds as having both of those activities.

Okay, if we talk about the de novo base methods for multi-target drug design, it ideally starts with the initial seed structure, where we can select the initial fragments or the template ligands from the ligand base. However, if you are using structure-based methods, we can take the initial positions of the fragments into multiple targets, as those initial fragments interact with the target in the binding pocket. So, we get that information and then generate the new structures either using ligand-based enumeration for the new structure or structure-based fragment growing within multiple targets. So, we generate the structure with the help of docking interactions, and while we grow the molecule, we see whether it is optimally binding to the binding pocket; if not, we just drop it and generate new molecules. And then we perform multi-objective fitness evaluation and enumeration growth termination criteria, and we repeat this until we obtain the optimal molecules.

And then, if we get the optimal molecule, we go for candidate selection, synthesis, and bioassay because there is no alternative to biological validation; we have to validate them in biological assays once we design and synthesize those molecules. So this is a case study for ligand-based, you know, de novo multi-target design, where you can see that donepezil was used as a potent molecule. So de novo design using donepezil led to this molecule R3, and it further showed increased D2 activity. And then further design of this molecule, de novo design of the molecule, led to the development of molecule 4, which had decreased anti-target activity on the alpha 1 family. However, compound 5 is again a de novo design based on donepezil.

You know, 6. So, these were 7. Therefore, they had default selectivity, and the scaffolds were also new, actually. So, you can see how, based on the ligands in drug design, we can do de novo design and identify molecules that can have multi-target activity. Talking about the structure-based de novo multi-target drug design, we start with the known structure, the known ligands of target one, and the known ligands of target two. And then we extract the fragments and confirm with the buyer, and then we do the docking based on fragment positioning, and then we use LigBuilder 3, for example, for multi-target growing. And then, once we have designed a molecule, we synthesize that molecule, perform the bioassay, and if it has low potency, we design the molecule again, synthesize it, and then synthesize it again.

So, repeatedly this design make test and analyze DMTA cycle goes on until unless we

identify a high potency ligand and ultimately leading to the multi target directed ligands for you know further use. Okay, this is another case study in which fragment-based seeds are generated by extracting data from known ligands and confirming it experimentally. So, the initial seeding positions are determined by docking, and the LigBuilder 3 growth mode is used to design dual-targeted ligands. So, you can see here that you have this starting molecule, a fragment, for example, which has COX-2 inhibitory activity and LTA4H inhibitory activity. And further, it has grown into larger molecules, and finally, you can see that this molecule has, for example, both improved COX-2 and LTA4 inhibitory activity.

So, there are some notable dual-targeting molecules that have been developed. So, this figure shows dual inhibitors that were recently developed using in silico binding-mode analysis of target crystal structures and scaffold-merging strategies. The active warhead frameworks in these are shown in blue, and the red participates in key interactions with the target protein. Okay, let's come to the chemogenomic approaches to multitarget drug design. So, chemogenomics combines chemical biology and genomic data to systematically explore the interaction space between small molecules and biological targets.

So, it plays a critical role in rational multi-target drug design by enabling the prediction, selection, and validation of targets based on molecular and phenotypic data. So, there are three different approaches that are commonly employed: target-centric, ligand-centric, and network based. For example, we have the target space here; the row represents all the possible compounds. So, these are all the possible compounds in the chemical space, and the columns are showing us all the possible targets, okay. And then you know that the interaction of these molecules is predicted for these targets using the chemogenomic approach.

And then you see which of those molecules can engage with those targets, and based on that data, they are further experimentally validated. And then we have target-centric chemogenomics. So, it begins with well-characterized targets involved in a disease pathway, such as kinases or GPCRs. So, it uses a ligand-target interaction database, structured data, and bioinformatics tools to identify secondary targets and similar ligand mining profiles, and it facilitates target prioritization based on druggability and co-expression or co-localization data. So, an example is using kinase cross-reactivity profiles to design dual EGFR and VEGFR inhibitors.

Then here, the ligand-centric chemogenomics starts with identifying bioactive compounds and their known and predicted target profiles using chemical similarity, ML, and activity cliff analysis. So, it helps uncover off-target effects that can be repurposed for multi-target design or polypharmacology. It also supports scaffold hopping, pharmacophore merging,

or fusion based on shared binding motifs. An example is identifying that a known HDAC inhibitor scaffold also weakly interacts with the bromodomains, prompting the multi-target-directed ligand design. And then we can use the network-based chemogenomics as well, which integrates protein-protein interaction networks, disease gene networks, and compound-target networks to identify key intervention nodes.

So, the targets are selected nodes in isolation, but based on network centrality, synergy, and compensatory pathways. So, it enables the rational design of drugs that can modulate entire disease modules. So, an example is designing multi-target drugs for neurodegeneration by mapping disease-relevant protein networks and prioritizing nodes affected by inflammation, aggregation, and mitochondrial dysfunction. Okay, so talking about where AI stands in MTDL design, these are all the applications of AI's roles that we have already discussed in previous sessions. However, it can help us in multi-target design by assisting with target prediction, where it can predict the primary and off-target effects of molecules using structural activity data.

And some of the tools are like Swiss target prediction or deep affinity. It can help us with compound generation, like all those de novo generative models you know, such as Reinvent, MolGAN, or DrugEx. And then he/she can help us with the SAR analysis and activity prediction. So, where it can learn multi-target structure-activity relationships from large data sets, we can use DeepChem, ChemProp, or AutoQSAR. And then you can use it for, you know, multi-objective optimization where it balances potency, selectivity, and properties for multi-target hits.

Tools like Optuna can also be used. And then, you know, you can use it for de-risking and off-target profiling, where it can predict toxicity and unintended interactions early in design, like Protox2, AdmetLab, or PKCSM. So everything that comes with some advantages always has some, you know, challenges associated with it. So what are some of those challenges with multi-target drug design strategies? There is, you know, target selection complexity because selecting multiple targets is a challenge as well, unless we have clean, transparent, or understandable disease pathology. And then structure and binding site compatibility also need to be, you know, explored, and that is a big challenge as well. The pharmacokinetic and physicochemical balance of the properties is also, you know, important and challenging.

And add met and safety liabilities because now you have, like, if you are fusing two molecules, so that is making, you know, a larger molecule. With a higher molecular weight, higher log P, and then altering those physicochemical ADMET properties and toxicity data, as well as toxicity properties. So, how to deal with that is also a challenge, and there is a lack of validated polypharmacology models because, until now, we have been just focusing

on, you know, one disease, one pathology. The animal models we are developing or using in preclinical research are usually single pathology models. So, we need validated polypharmacology models where multiple pathologies are present, you know, and can be studied together.

And then synthetic feasibility is another issue because, as I said, we are merging two molecules, making them challenging to synthesize. And then regulatory and development hurdles are also another challenge where the regulatory bodies are, you know, a little bit, you know, skeptical about those multi-targeted drug ligands, and usually they need more data to, you know, validate those claims made by the submitters, actually. Okay, coming to the summary, this multi-target drug design is a promising approach that modulates multiple disease-relevant targets, improving efficacy and reducing resistance in complex diseases. And the medicinal chemistry supports these two strategies, such as pharmacophore merging and linking, to create multi-target directed ligands. Computational tools, including ligand-based modeling, structure-based design, and de novo generation, streamline scaffold discovery, and chemogenomic methods integrate chemical and biological data to identify synergistic targets.

AI enhances this field through predictive modeling, generative design, and system-level target identification. So, in the end, I have an open question for you. If a single AI model could design a compound that adapts its conformation and binding affinity dynamically inside the body, binding to different targets based on disease stage or tissue environment, could we be entering an era of self-optimizing living drugs for precision polypharmacology? And I have some of the suggested references that you can go through if you want to know more about, you know, this topic. And with that, thank you.