AI in Drug Discovery and Development

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

Week-04
Lecture-16

Hello, welcome to the course "AI in Drug Discovery and Development." Today, we will talk about drug targets. So, by the end of this lecture, you will be able to define drug targets and their significance. Describe the characteristics of an ideal drug target. Identify major classes of drug targets and explain enzyme inhibition mechanisms. So, let us have a look at what the drug targets are.

So whenever we talk about a drug, as we discussed earlier, a drug is supposed to bind to some biomolecule in our body, and that is called a drug target. So, a drug target is a specific biological macromolecule. Usually, it is a protein or a nucleic acid that a drug interacts with to produce a therapeutic effect. So, if you look at our cell, it has 23 pairs of chromosomes, which consist of around 3 billion base pairs.

So, our genetic material consists of 20 to 25 k protein-coding genes, which code for around 5000 potential druggable macromolecular targets. But the drug targets are, you know, the overlap between the disease-modifying genes and the druggable genome. So, if we look at the types of drug targets, we can classify them into different subclasses. For example, we have enzymes in our body. So, we have the enzymes in our body, and we can identify the molecules or drugs that act as their inhibitors, or they can also act as activators as well.

And then the second class of targets is the receptors. So, we have several kinds of receptors in our body. So, for which we can develop agonists, antagonists, their modulators, allosteric activators, or sensitizers. The next class of drug targets is the transcription factor, where we can develop its inhibitors or activators depending on what kind of function we want to have with those drug molecules. And then we have ion channels that are responsible for transporting ions in and out of the cells in our body.

So, we can modulate their action by developing inhibitors or openers of those ion channels to alter them to get the desired effect. And then we have transport proteins in our cells, which can also be inhibited with small molecules or biologicals, and we try to develop inhibitors for them as well. Another class of drug targets is the protein-protein interface, and we can also develop inhibitors of these protein-protein interactions and develop those molecules as drug molecules. And finally, we have nucleic acids as well, which can be modulated by developing molecules that can lead to their alkylation, complexation, and

intercalation; this is one of the major classes of drug targets used to treat cancer. And this is a small, you know, data; this is the percentage of different drug targets that have been used to develop drugs.

So you can see that a larger portion of the drug targets is composed of receptors, followed by the enzymes. And then we have the hormones and other factors that are being modulated to develop the drugs. And then we have a small percentage of the drug targets that consist of nucleic acid, around 2%. We have nuclear receptors and ion channels, with around 5 percent of them being utilized as drug targets for developing drugs or treating diseases. And then there are some drug targets that are unknown.

So, it means that we really do not know how those drugs act. Since they give the desired effect, those are being developed as drug molecules. So, if you further look at the enzymes, as they are biological catalysts, they speed up the chemical reactions in living organisms by lowering the activation energy required for the reaction to proceed. So, they play a very important role in metabolic pathways, attracting metabolism and disease mechanisms. And there are several classes of enzymes that are present in our body.

So these can be classified, for example, into oxidoreductases, which catalyze redox chemistry reactions, electron transfer, and often with the help of a cofactor. Some examples of oxidoreductases are HMG coenzyme A reductase and cyclooxygenase. And then we have transferases; they catalyze the transfer of functional groups. So, like tyrosine kinase or reverse transcriptase? Then we have the hydrolases, which catalyze the hydrolysis of a chemical bond. So, these are like HIV protease or tyrosine phosphatase.

And then we have lyase as well, which catalyzes bone cleavage, forming double bonds or rings without hydrolysis or oxidation. Some examples of lyases are adenylate cyclase and pyruvate decarboxylase. And then we have isomerases, which catalyze structural rearrangement to form isomers of the substrate. The examples are topoisomerase and retinol isomerase. And finally, we have the ligase enzyme which catalyzes the joining of large molecules with a chemical bond.

And then the examples are DNA ligase and RNA ligase. So, this is a clinically utilized drug target known as acetyl coenzyme A. Here, we see the structure of the acetylcholinesterase enzyme. So, which is targeted for the treatment of Alzheimer's disease, and here you can see that it is. bound with a drug called as donepezil, so which inhibits its action.

So, at acetylcholinesterase, it degrades acetylcholine at the neuronal junction. So, what this drug is doing is inhibiting its action, leading to the increased availability of

acetylcholine at the neuromuscular junction. So, here we see the enzyme called Acetylcholinesterase, short for ACHE. So, what you can see here is that there is a small molecule drug called Donepezil that is bound to the active site of this enzyme. And this enzyme plays a crucial role in Alzheimer's disease, and inhibition of this leads to improvement in the symptoms of Alzheimer's disease.

So, this is being used very frequently for the treatment of Alzheimer's disease. So let us take a look at how the enzyme works. So there are two hypotheses; basically, one is called the Lock and Key hypothesis. Where it is presumed that the enzyme has this binding pocket, the substrate comes to the binding pocket, binds, and the enzymatic reaction or catalysis occurs, converting the substrate into the product. Okay, and then the hypothesis is that the substrate perfectly fits into the binding pocket, like a key fitting into a lock.

On the other hand, the induced fit hypothesis states that the enzyme undergoes a conformational change upon substrate binding, adjusting its shape to fit the substrate more snugly. So, you can see here, for example, that the substrate binds to this enzyme binding pocket. Then the enzyme is adapting to the structure of the substrate, and this is what is called induced fit. And then the reaction happens, catalysis occurs, and then the engine functions, and the product is formed, which is released from the enzyme binding site. So then we can classify those enzyme inhibitions into two categories.

One is called reversible inhibition, where the inhibitor covalently binds to the enzyme. So, if you look at enzyme inhibition, it can be classified into two categories. One is called reversible inhibition, where the inhibitor binds non-covalently to the enzyme, and it can be removed, leading to the restoration of enzyme activity. On the other hand, we have irreversible inhibition as well, where an inhibitor can bind to the enzyme, permanently inactivating it, and then the enzyme's function cannot be restored. So, further reversible inhibition can be of multiple types, such as competitive inhibition, non-competitive inhibition, uncompetitive inhibition, and mixed inhibition.

Where the irreversible inhibition can be of two types: one is called mechanism-based inhibition, and the other is covalent inhibition. So, in the case of competitive inhibition, the inhibitor resembles the substrate and binds to the active site. However, it can be overcome by increasing the substrate concentration. So, one of the examples of the competitive inhibitor is Tamiflu, also known as Oseltamivir, which completely inhibits influenza neuraminidase, blocking viral replication. Non-competitive inhibitors are the inhibitors that bind to an allosteric site, and thus they reduce the enzyme activity without competing for the active site.

And it cannot be overcome by increasing the substrate concentration because it is not

binding to the pocket where the substrate is binding. So, it is difficult for you to overcome the inhibition by increasing the substrate concentration. So, one of the examples of non-competitive inhibitors is metals like lead, which inhibit enzymes by binding allosterically to the enzymes. Then the uncompetitive inhibitors bind only to the enzyme-substrate complex, preventing product formation, and increasing substrate concentration enhances inhibition. So, the example of an uncompetitive inhibitor is lithium, which inhibits inositol monophosphatase in bipolar disorder treatment.

And then we have mixed inhibition, where the inhibitor binds to both the free enzyme and the enzyme-substrate complex, but with different affinities. So, it lowers the enzyme efficiency like some of the kinase inhibitors in cancer therapy, which act in a mixed inhibition way. So, we have irreversible inhibitors; the first one is mechanism-based inhibitors, like suicide inhibition. So, these inhibitors mimic the substrate undergoing a partial reaction before forming a covalent bond with the enzyme. So, what they do is bind to the enzyme binding pocket, then form a covalent bond and permanently get attached to the binding pocket.

So, thus blocking it and then it cannot be reversed. So, one of the examples of this mechanism-based inhibitor is beta-lactamase like penicillin G, which irreversibly binds to PBPs, inhibiting bacterial cell wall synthesis. And then another one is covalent inhibition. These inhibitors directly form a covalent bond with the enzyme's active site residue. And an example is aspirin, which acetylates cyclooxygenase and inhibits prostaglandin synthesis.

So another type of drug target is G protein-coupled receptors, also known as GPCRs. So the GPCRs play a very significant role in signal transduction, and they are the largest class of human membrane-bound proteins. So if you look at the structure of GPCR, it usually contains seven transmembrane helices. You can see here; these are shown in this pink color: TM1, TM2 to TM7. So these are transmembrane helices, and here you can actually see the membrane.

So this is the cell membrane. So, you have the extracellular domains like EL1, EL2, and EL3, and then you have the intracellular domains. So, the extracellular domains bind with the ligands. However, the intracellular domains interact with the heterotrimeric G proteins G alpha, beta, and gamma. And those are responsible for further signal transduction, leading to the downstream signaling effect. So the GPCRs mediate cellular responses through different G protein-dependent signaling pathways.

So these are the two most common G protein-dependent signaling pathways, such as the cAMP signaling pathway, which is 3',5'-cyclic adenosine monophosphate that acts as a

secondary messenger. And then we have the phosphatidylinositol IP3/DAG signaling pathway. So, the IP3 inositol-1,4,5-triphosphate and DAG diacylglycerol act as secondary messengers from phosphatidylinositol-4,5-bisphosphate PIP_2. And then they further lead to signal transduction downstream. So, the drugs that act on G protein-coupled receptors modulate them through different mechanisms.

So, these are called, for example, full agonists. So, it is indicated by this blue line. So, what it shows is that it induces GPCR signalling equal to that of the endogenous ligand, leading to a maximal response. And then we have another category, which is partial agonist.

So, this is indicated by the red line. So, what they do is they activate the GPCR signaling, but produces a lower response compared to full agonist. And then we have neutral antagonists saying they do not induce GPCR activity on their own, but block the effect of an agonist. And then we have the inverse agonist shown in purple here. So, what they do is suppress the basal activity of GPCRs, thus reducing the signaling below normal levels. So, here you can see that this is at 0; this is the basal (constitutive) activity of GPCR.

And then these agonists and antagonists alter this basal activity either in the upward direction or in the downward direction. So, these are some of the drugs that target G protein-coupled receptors. So, for example, these are the natural ligands for the histamine receptor. So, histamine is a natural ligand, and then we have a histamine H1 receptor antagonist called loratadine.

So, it inhibits the action of the histamine. And then we have dopamine, which is a natural ligand for dopamine receptors. And then this olanzapine, the drug which is an inverse agonist, a D1D2 inverse agonist, can inhibit the action of dopamine. And then we have, similarly, serotonin, which is a natural ligand for the serotonergic receptor. So this methylsergide is the drug that inhibits the 5-HT2 receptor, and it acts as a 5-HT2 receptor antagonist. Further looking at the ion channels, these ion channels are transmembrane protein assemblies that regulate the selective flow of ions, such as sodium, potassium, calcium, or chloride, across biological membranes.

So they play a very important role in nerve impulse transmission, muscle contraction, and cardiovascular function, including the regulation of heart rate and rhythm. Here you can see the side view of an extracellular structure of Streptomyces lividans, a potassium channel. And then this is the top view obtained from the PDB ID: 1BL8, and here you can see the potassium ion that is bound to this channel. So this potassium ion is moving through this channel from inside the cell to the outside. So here it forms a kind of channel through which these potassium ions are being transported.

So, there are multiple gating mechanisms for these ion channels. So, there are, for example, ligand-gated channels that open in response to a ligand like a neurotransmitter binding. So, like the nicotinic acetylcholine receptor. So, you can see here that this is the cytoplasmic space, this is the extracellular space, and then we have the plasma membrane, with these ion channels embedded in it. So, even when this natural ligand is binding to this ion channel, only then does this ion channel open, allowing the ions to flow inside the cell.

So, this mechanism is called as these types of ion channels are called ligand-gated ion channels, which means the opening and closing of the ion channels is controlled by the ligands. And on the other hand, we have voltage-gated ion channels as well, so they open or close in response to the change in membrane potential. For example, the voltage-gated sodium channel, potassium channel, and calcium channel. So now the cell membrane is at its resting potential, during which time this ion channel is closed. So, when this activation happens, so the activation threshold potential has reached.

So, this leads to the opening of this ion channel, and once this ion channel opens, those ions can flow inside. And then again, if the further inactivation of this ion channel happens due to a change in the potential. So, this is how these voltage-gated ion channels act. So, their opening and closing are controlled by the membrane potential. So, then we have membrane transport proteins that are called transporters.

So, they facilitate the movement of molecules across the cell membrane. So, there are, for example, three types of them, like uniport, which transports a single type of molecule in one direction across the membrane, like the glucose transporter GLUT1, which facilitates glucose diffusion into the cell. Like uniport, which will only transport one single type of molecule inside or outside the cell. And then you have symport, which is also known as co-transport, that moves two or more molecules in the same direction across the membrane. For example, sodium glucose cotransporter SGLT1, which uses a sodium gradient to transport glucose into the intestinal cells.

So, symport will be like transporting two molecules, but in the same direction. However, the antiport acts as an exchanger that moves two molecules in opposite directions across the membrane. For example, the sodium-calcium exchanger exchanges three sodium ions for one calcium ion, regulating calcium levels in the cardiac cells. So, you can see here that these will be exchanged. So, one molecule will be transported in one direction in exchange for the other molecule, which is transported in the other direction.

So, the norepinephrine transporter, which is a sodium-dependent symporter that reuptakes norepinephrine into presynaptic neurons, regulates neurotransmission. So, we have the

glucose transporter GLUT1, which is a facilitated diffusion transporter responsible for glucose uptake, particularly in the brain and erythrocytes. And then we have ATP-binding cassette proteins. They utilize ATP hydrolysis to transport various molecules across membranes. Example: PGP, which is also involved in multidrug resistance.

So, we have transcription factors as drug targets. So they control gene expression by binding to DNA, influencing cellular processes like proliferation, differentiation, and apoptosis. So these TFs lack well-defined active sites for small molecules. So some of the key targets are nuclear receptors like estrogen receptors and androgen receptors. For example, tamoxifen for breast cancer, then p53, which is a tumor suppressor targeted for reactivation in cancer therapy. Then we have NF-kappa B, which regulates the inflammation and immune response targeted by the corticosteroids.

And then we have STAT proteins, which are targeted in cancer and immune disorders. For example, ruxolitinib targets JAK-STAT inhibition. And finally, we have nucleic acids as drug targets. So, nucleic acids like DNA and RNA play a crucial role in generating information storage and expression. So, targeting nucleic acids can regulate gene expression, inhibit replication, and disrupt cellular processes.

So, there are some drugs that target nucleic acids like DNA intercalating agents, also known as alkylating agents. Examples include cyclophosphamide, intercalating agents, and topoisomerase inhibitors. So then we have RNA-targeting drugs like antisense oligonucleotides and small interfering siRNA as well. And then we have G quadruplex stabilizers that target the G-rich DNA or RNA structures found in oncogene promoters and telomerase to inhibit cancer cell growth. And then we are using this CRISPR-based gene editing, which allows for precise gene modification for therapeutic applications.

So, in addition to these nucleic acids as targets, we have protein-protein interfaces as targets as well. So, these interfaces are crucial for cellular functions, making them attractive drug targets in diseases like cancer and neurodegeneration. So, some of the targeting strategies for protein-protein interfaces are small molecules that can disrupt specific interaction hotspots, like how Nutlin inhibits the MDM2 interaction with p53. And then we can also develop peptides and peptidomimetics that mimic the native interface region, like BH3 mimetics for the BCL2 family. We have monoclonal antibodies that block the extracellular PPIs, such as checkpoint inhibitors like pembrolizumab.

And then recently there has been a lot of, you know, interest in developing PROTAC inhibitors, which induce targeted protein degradation via protein-protein interactions. So this is one example of a potential drug target for treating many diseases. So this is the structure of FKBP51 bound to the HSP90 C-terminal peptide. So this complex is known as

the chaperone-cochaperone complex.

So here, HSP90 is a chaperone, and FKBP51 is a cochaperone of HSP90. So, together they stabilize several proteins that are involved in several diseases. So, one idea is that one can inhibit this protein-protein interaction to treat those diseases. So, it is involved in conditions like Alzheimer's disease, obesity, and depression. So, a lot of diseases are there where this this interaction has shown significance. So if we look at the summary, the drug targets are central to drug discovery and development as they determine the efficacy and specificity of therapeutic agents.

So the major drug target classes include enzymes, GPCRs, ion channels, and transporters, each playing a distinct role in physiological regulation. And then the protein-protein interface is a crucial drug target because it mediates key interactions in biological processes, offering opportunities to disrupt disease-related pathways with high specificity and efficacy. So, in the end, I have an activity for you. So, what you have to do is identify a few drug targets that were undruggable earlier, but now have become druggable. So, you can go through these nice articles if you want to learn more about drug targets. And with that, thank you.