NMR spectroscopy for Structural Biology
Prof. Ashutosh Kumar and Prof. Ramkrishna Hosur
Department of Chemistry
Indian Institute of Technology - Bombay

Lecture: 47
NMR in Drug Discovery

Good morning students. So this week we are discussing how we can use NMR spectroscopy in drug discovery. So based on the knowledge of protein-protein interaction, protein-drug interaction, we ventured into the use of NMR spectroscopy in drug discovery. In last week, we discussed protein-protein interaction. This week the first lecture was about how we can fish out some of the binders in the protein-protein interactions. So, we discussed two kind of experiment when we are detecting on ligand.

NMR experiments like STD, saturation transfer difference or water-LOGSY or linewidth analysis is of paramount importance where we can detect on ligand and find it out binders. So, it is kind of a fishing out experiment. There are so many small molecules, which binds to a receptor, which is a target and using these experiments we can find it out which molecule exactly binds and with what strength it binds to the receptor or the target. On the other hand, protein detected experiments can tell you the site on protein where ligand binds.

For doing that we need to isotopically label these proteins with 15N or 13C-15N both. So, if we do that and titrate ligand and whatever chemical perturbation happens or intensity decay happens using any of these probes we can find it out exact location where the drugs are binding. So, that is about drug discovery. We are discovering out of many compounds, whether it is coming from natural source or plant source or animal source or synthetic by our medicinal chemistry friend they have synthesized many molecules, and which one of them is binder, what is the potency, we can find it out using NMR spectroscopy. Now this week we are going to look at how we can start developing based on the knowledge, whatever we have.

So the first thing we can think of an SAR, Structure Activity Relationship. It is like a classical titration based experiment that we can think of. So here is my protein. And let us see that it has certain binding site. Here is kind of a binding site, which has a particular shape.

Here is a star shaped binding site and we need to find it out which molecules bind with what *K*d strength and using that knowledge can we start developing actually better drug. So SAR is a NMR method where protein chemical shift can be used to see where it binds. So there will be a change in the chemical shift and essentially that will be used to screen the low affinity binders. And we can use such information, the low affinity binders, where they are binding, to directly start attaching fragments and make a better affinity drug.

So that is the whole idea. So what we are going to do here taking a protein molecule, we are isotopically label this protein and say there is a binding site where we are starting with a binder. So, we see some shift that is happening here. So, suppose this peak shifted here, this peak shifted here and these peaks shifted here, other peaks say did not shifted. So, now we know that these are the binding epitopes on protein and we can find it out binder.

So even low affinity binders like this molecule binds with a *K*d of 17 mM. Now suppose there is another binding site, which is here. We can start again with another set of compound and find it out which is binding here, so different binding sites on same protein. Now next step, we can attach it to find a better binder, something like this. So here is our titration experiment, we took the N15 labelled protein, we titrated with a molecule, find it out all the binding position like here you can see.

The chemical shift shown here is directly telling this is not shifting. So, this is not shifting, here you are seeing that there is a shift here, these peaks are shifting here, this peak is shifting, this is almost not shifting. So utilizing these, we identified the binding sites. So here is my binding site, this molecule comes and binds like this. So exact binding site we can find it out for one binding site.

Suppose this protein has another binding site, which say here. Now similar experiment I need to do with another set of compound. Find it out that what are the residue involved here which are binding. So that is like here. So, if we are starting with another set of molecule you see a star shape molecule is coming and fitting here, the pentagon will not fit.

So, suppose this is the case now we have a better binder like here is say *K*d of micromolar. Suppose there is another ligand, again a low affinity ligand, which binds at different site, that is identified based on the chemical shift. Now these both ligands, which were discovered in the previous case here, which is say they shaped 1, 2, 3, 4, 5, 6, 7, heptagon shaped ligand, here is a star shaped ligand, this binds with a micromolar affinity, this binds with a millimolar affinity. Now we have found two ligand binding with two different affinity, can we use this to optimize a better binders like can we attach both of these and have an analog, which has like even more binding sites. So this is the idea that we will be using in drug discovery case. So using SAR, structure activity relationship by NMR, one can design a compound that binds to the catalytic domain of any protein like a stromelysin.

So here, if you look at this guy was the first binder and this was the second binder. We attached by some chemical bond. We use synthetic chemistry approach, a medicinal chemistry approach. Using NMR based methods, we find the first binding molecule that is binding at a millimolar range. Then we found another molecule that is binding at micromolar range. Now we attach these two and we find a molecule, which is by joining these now binding with a very high potency of nanomolar range.

Now the molecule that we were discussing, stromelysin is a metallo-MMP, matrix metalloprotease which has implication in cancer. So one can find it out the protein with a two different binding site, the two ligands separately by SAR by chemical shift perturbation and now we can join them by these bonds to find a more potent molecule that has a potency of essentially nanomolar. So here exposition of a simple titration experiment, where you use for screening weak ligands and structurally directing these ligands by chemical linkage to have a more potent compound and that is how one can approach in drug development stage.

So first thing what we did? We discovered the binders in the first approach by titration. Then using our chemical synthesis or medicinal chemistry, we attach these two weaker binder to come up with a strong binder of nanomolar affinity and that is how we are developing a more potent drug. So that is the simple concept. Let us discuss some of the advanced concept of the NMR based screening. So one of them is called SHAPES screening, how we can screen it.

So it is a basic strategy in the SHAPES screening to access the binding of a fairly small but diverse library of low molecular weight scaffold to have a drug target using NMR techniques that are based on observing the ligands. So what we are doing again let me dissect all those. In SHAPES screening, essentially we are taking a small set of library of various compounds and now we are detecting on the ligand side like using STD experiment, that we have seen and finding it out what are the binders and then probably we will start developing from there. So ligand based experiments, so I hope you remember STD.

What we are doing here is recording an NMR spectrum of this small molecule. We are having some peaks here. Then we added our receptor and we irradiated the receptor signal. The effect of that irradiation, RF irradiation is transferred to the binders. It will not transfer to the non-binders because they are not in contact of the protein.

And then some of the peaks, say these peaks show low in intensity. So if you take the saturation transfer difference, only this peak will be shown, these two peaks will be almost zero. So you know that this is the binding site on the ligand. So we are going to use this STD approach, and observe on ligand. And then we are screening the molecules that is called SHAPES screening.

So here using SHAPES screening, we find a chemical moiety that can bind to our ligand site. So here, our protein target is protein kinase P38. Now we find a moiety, which binds with a low affinity of *K*d of 2 mM, it is a quite weak. So then we find it out what is the essential moiety.

Can we start playing around using some chemistry and find it out some modification. Like here you attach something, like here you attach another group, and in this case you can change some of the group, like here nitrogen can be changed to say sulphur compound, then you attach few of these. So if we do such kind of chemistry by attaching few moiety and then again doing STD kind of experiment to find it out what is the next binding strength. So by doing this all chemistry, medicinal chemistry approach, again combining that with a SHAPES screening using ligand detected experiment, one can find it out a molecules that were developed can bind with a fairly strong strength which is going to be of 200 to 300 micromolar. So you see like almost 1000 fold or at least 100 fold binding capacity we have increased.

Now let us go ahead. So now what we did, the scaffold that are derived mainly from the shape. By shape what I mean, here is a binding site, you know the kind of a shape of a binding site, now you are finding a key that can fit here and that is we screened that shape and then shape or framework that are commonly found in the known drug. So, initially we had a weak binding hit say *K*d is in micromolar, but that is important to find the framework. And then in follow-up strategy, the SHAPES hits can be used for starting platform for virtual screening. We can start modifying using the virtual screen, like a computer based drug design and medicinal chemistry, what way I will fit it there that binding becomes stronger. So even looking at the structure, the shape of the binding pocket and the initial framework, you can start designing the molecule in computer which will tightly fit into that.

Then you go and synthesize this molecule and then again you come back to do NMR, find it out what is the real *K*d. So this iterative approach of finding a shape, where this molecule is binding. Finding initial platform or the shape of the molecule developing that into a better binder first using computational approach then using synthetic approach coming back again to STD NMR; finding it out what percentage or what fraction it has improved that is how you proceed in this kind of drug development. So by doing this, last example I showed that design of an inhibitor for mitogen activated protein kinase p38. So on the basis of this initial finding, a library of analog can be developed like just now we see lots of compounds can be synthesized. Here we can change even R2 group, we can add some ethyl at some position, methyl at some position or some other group and see how the binding is improving.

That is why we have a range of binding, it can vary depending upon how tightly it binds. So on the basis of these initial finding of binders using a STD approach, we can develop a library of analog with further derivatization and using this initial core structure we can obtain the high affinity ligand. So what substitution we can do? Substitute something. Core framework we are keeping same, we are attaching something like this. Also, you saw in the last slide, this is also kept intact, but there we are attaching another ring, here we are attaching another ring with fluorine. Now, fluorine is a very interesting molecule in the next classes I am going to tell you this can be used for tracing the drug, when you develop a drug you have to find it out where it is going.

So, fluorine is a very interesting nuclei in terms of NMR and it has many other role in the medicinal chemistry. So one can develop few of these fluorinated compound and you see all these derivatives that we have developed or synthesized, now my strength has increased, the potency has increased. So now it is 10 to 200 nanomolar. We started with a millimolar but developing on that framework, based on the shape of the binder, we have come up to a 100 to 200 of nanomolar and here the simple thing was used. Simple thing of one dimensional NMR was used for a screening, identifying, making it high affinity analog starting from a weak binding hit where we started developing weak binding was this 2 millimolar, these two rings, 5 member rings, pyridine rings and here 6 member ring with nitrogen. Now we are starting to develop and by developing all these we have found analogs that can bind with very high affinity of 10 to 200 nanomolar.

So that is a shape based screening. Now another approach, that one can think of an NMR SOLVE. So, SOLVE is called Structurally Oriented Library Valency Engineering. So, here we are now looking at the protein molecule and we are engineering a binder, so looking at the protein molecule. And even we are not totally labeling it, we will do some smart strategy to engineer a drug molecule, binding molecule. So how we are engineering? So first we need to have some structural information.

So structurally oriented library valency engineering. So here say this is my binding site. We need to have some proofs here. So NMR SOLVE is a drug design strategy that provides information for the construction of focus library. You are not going to do a blind screening like many approaches, you take it like a compound or 10,000 compound and just randomly screening based on the high throughput and you find some thousands then you come to 100 then you come to 10 then two molecules are working

at least in vitro and then you do not know what happens. So here, NMR can cut down that screening. So now you are designing based on this information. You are not blindly screening it. You are using some information that are available based on the protein binding site or the ligand binding site that we saw in the last slide.

Here this strategy provides information for the construction of focus library. So where the information is coming? From the binding site. So let us look at how binding site is providing. So basically this NMR SOLVE exploits the fact that there is a large family of protein that have an adjacent binding site and one of them is conserved.

So many of these are same family of protein. They can have a similar binding site or they have an additional binding site and that information needs to be conserved in the family proteins. Here you see one binding site looks common, but another binding site here is a star shape, here is some other shape, here is a triangular shape. They have at least one binding site, which looks quite conserved. So now, what we are going to do exploit which is conserved binding site.

Let us see what is there. Now for doing that we need to put some of our probes. What are our probes? So it can be amino acids, it can be any moiety, it can be protons which are there. Label them and we are looking only at these and this gives the idea about the binding site. So once we have the idea about the binding site, we can come up with a ligand that will probably fit here.

Now we already started with an initial here. Next we are going to develop that. So depending upon what is the second binding site, we can screen it and then start attaching. So let us see how it happens. So what we are going to do, here are my probes, they provided information and information is coming from the chemical shift change. So we can have either say N15 labeled protein or 13C labeled protein or both labeled protein and we just recorded the correlation spectrum like whatever we were discussing and then we find what are the residue that are involved here, so on a structure. So, we did titration we find it out the residue involved here and those residue are going to be conserved in other family protein as well.

Now, next we have to find it out what is a joining site between these two. So, we need to have a probe here as well. Now, once we have a probe we can design even the molecule that will fit here and next we will be designing for another binding site. So we find the orientation of the first binding site, common ligand that mimics. So we already by exploiting this binding site, by putting specific labels there, we know the idea of this binding site and these drugs at the binding site number 1 will fit to the all members of the family, so half job is done.

So one find a good binder here, we can just translate that binder to another family member of that family of proteins. So, the orientation of the most common ligand mimicking the binding site of a protein determined and that provides the basis for designing the linker

here. And for linker design, we need to have another probe which is coming from the amino acid that is located here. And then we will be essentially direct my ligand to this site. So now, we are developing a framework here based on the information, which is provided by this, what is the linker.

Now this will direct that my next binder will be coming here and attaching to a particular site. So now NMR data for protein 1 can be used to guide synthesis of a combinatorial bi-ligand library. So using this we can synthesize, so here is my first molecule, you also find a binder, sorry a linker and then we synthesize variety of molecules. We make a small library which has a bidentate and this bidentate like we have one time this shape, another time this shape, third time this shape. So two things helped us, one the information of the first binding site, second the linker, this exclusively came from the information that we obtained from the NMR.

Now utilizing these we synthesize a combinatorial library and then we did the activity assay to finding it out what can best bind to this place. So, such approaches has been used in fragment based drug design, this is called fragment based drug design. We have two binding sites. So here is a fragment that binds to the binding site; one can we find it out with HSQC. Binding site two can we find with HSQC experiment again and then we find these two molecules and we can map on the tertiary structure, then we synthesize a linker and using again N15 label you can find it out that binder. So this actually approach was used to identify a binder for BCL-xl and BCL-2 compound and this is the compound that is still in clinical trials ABT737, ABT199. So based on this fragment based drug design, like attaching these ligands, actually the drug was designed and now it is in the clinical trial.

So 3D structure based ternary complex, 3D structure was used for doing this. The another approach in the drug design can be seen is called interligand NOE, ILOE. So what we are doing here, say we have a two binding site, here I used in fragment based drug design. What I used, the information based on the protein side, we had a N15 labeled protein, we titrated it, find the two binders then attaching it using some information which is there and found a better molecule. In the second approach what we will be doing is using the NOE between two ligands.

So here is my first ligand that is binding to the first site. The F2 is a second ligand binding to a second site. If they are close enough, they can show the nuclear overhauser effect, because they are in close proximity. So, now here is my F1 ligand, here is my F2 ligand and if we record an NOE between them, we find the NOE which is coming here. So, in this case we are started with an unlabeled protein target, unlabeled ligand, we just look at this small molecule that is binders, the drug molecule, the NOE. Now this information provided that this particular portion of ligand 1 is in close proximity of a particular portion of Ligand 2. So this interligand NOE will help us in essentially joining this.

So structural analysis of the fragment based interligand NOE will help us that where we can do the chemistry to join this. And now this will be called bidentate ligand, you can again go for N15 labeled protein target and find it out how better it is binding. So if you look carefully, we are using the NMR spectroscopy with the chemistry, medicinal chemistry and some biochemistry to label the protein, some combinatorial chemistry to synthesize the ligand and after combining all these approach we can synthesize a better potent molecule the bidentate compound and one can validate again using NMR spectroscopy. So this is kind of approach was used to synthesize molecule and that gave us a strong binder molecule that are being used nowadays for the drug designs. To sum up now what we learnt basically in drug design, essentially we have to start with a framework either on protein detected experiment or on ligand detected experiment.

So we start with a framework and then start developing it. If needed, if your protein has two binding sites, you identify those two molecules, use some information coming from the protein site, attach them and make a better binder or a better potent drug. So we are starting with a millimolar binding affinity and one can reach up to nanomolar of binding affinity and that nowadays are used in designing and developing drugs. So I hope in the last two lectures, I gave you some idea that NMR has a profound impact on drug design and development. In the next class what I intend to do is to give you an overview how NMR can be used for monitoring the drug fate like once you consume or once anybody consume the drugs, how it is passing through different organs and can we use that information, the NMR based information to understand the drug metabolism.

So that is what I intend to do in the next class. Looking forward to have you in the next class. Thank you very much.