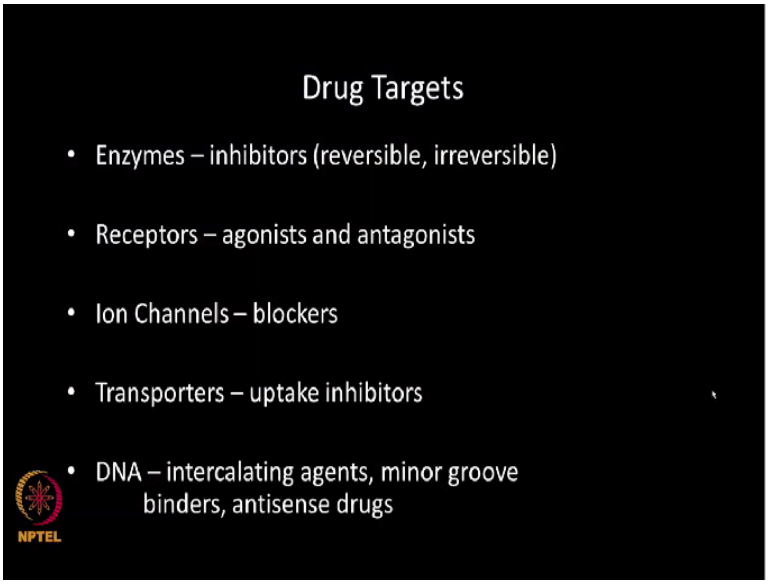
Computer Aided Drug Design
Prof. Mukesh Doble
Department of Biotechnology
Indian Institute of Technology – Madras

Lecture – 36
Target Based Drug Design

Hello, everyone, welcome to the course on computer in a drug design. We will continue on the topic of target based drug design that means looking at the docking of small molecules to the target. So, the target could be an enzyme.

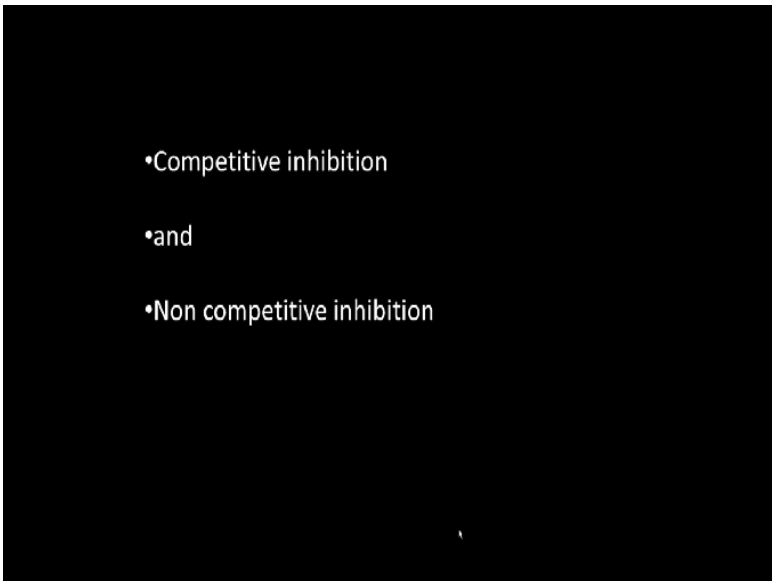
(Refer Slide Time: 00:29)



So, the drug could be inhibitors. It could be a reversible inhibitor, irreversible inhibitor. Target could be a receptor so the drug could be agonists or antagonist. Target could be an ion channel so you could be designing drugs an ion channel blocker. The target could be a transporter so we can design uptake inhibitors. Target could be a DNA so the drug could be an intercalating minor groove binders, antisense drugs.

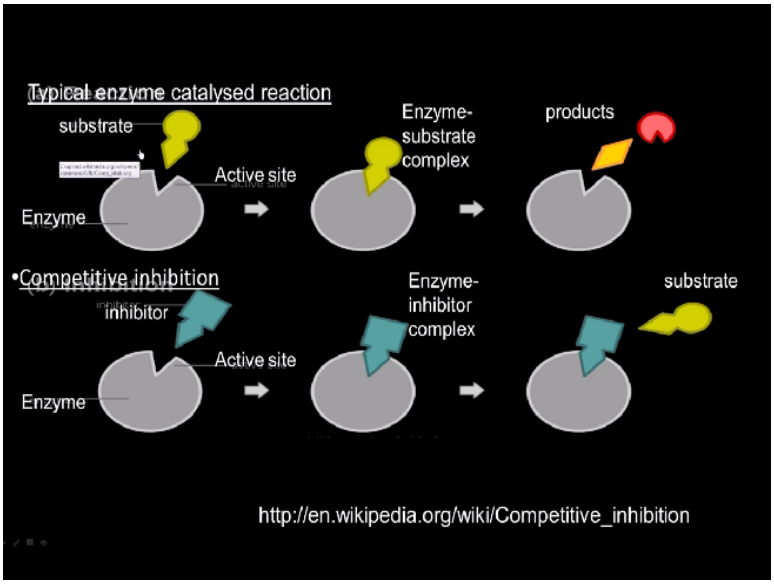
So, depending upon the target you call your drug as inhibitor or agonists or blockers or uptake inhibitors or intercalating agents or minor groove binders and so on actually.

(Refer Slide Time: 01:17)



If you look at the enzyme you could have a competitive inhibition, non-competitive inhibition, allosteric and so on. So, look at this.

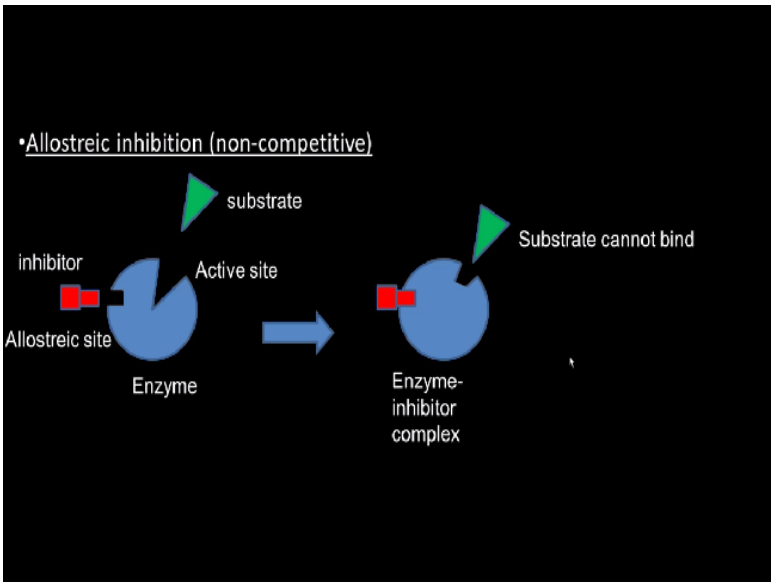
(Refer Slide Time: 01:26)



Your enzyme you must all have studied in your biochemistry or enzyme technology course. We have an enzyme we have the active side and then we have the substrate so the enzyme substrate complex is formed as you can see. Then it is producing the product. This is how things happen here and so when you have a competitive inhibitor, the inhibitor goes and binds to the active side here. So, it forms an enzyme inhibitor complex so the substrate they are not able to come and bind.

This is taken from this type of place or reference. So, this is competitive because your drug is competing for your active site like your substrate.

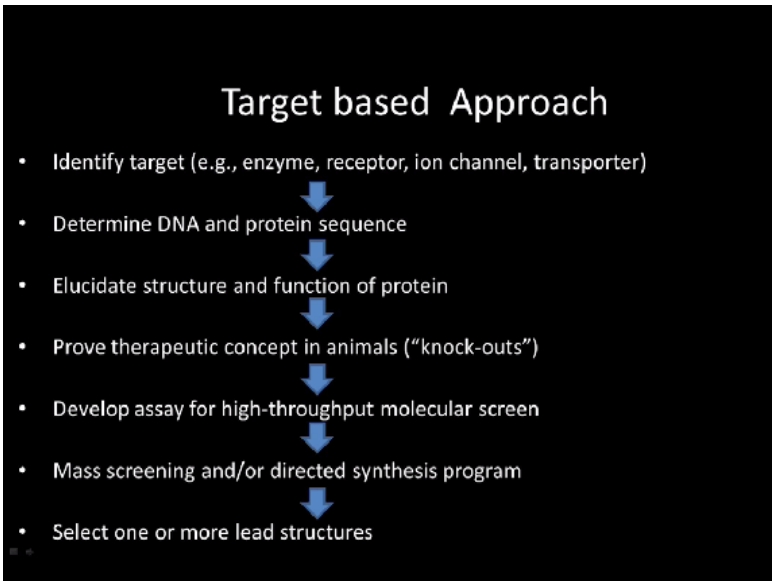
(Refer Slide Time: 02:26)



If you look at allosteric inhibition or non-competitive inhibition what is happening? The inhibitor may be binding to some other site, not the substrate site some other site. It is called allosteric site by doing that the active site gets disturbed or the structure of the enzyme or the protein gets disturbed so the substrate is not able to bind to that so that is another approach by which drugs can act. So, the drugs can go and bind.

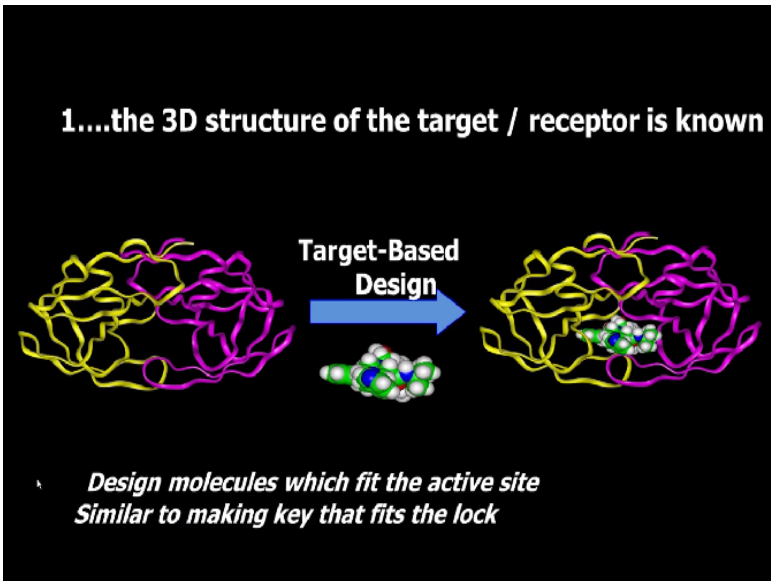
For example, sulphur drugs binds to the active site where the substrate like para aminobenzoic acid goes and binds. So, para aminobenzoic acid is not able to go and bind so sulphur drugs are competitive. Whereas if you take some other drugs, some other antibiotics they go and bind as an allosteric inhibitor and hence prevent the activity of enzyme and also prevents the binding of substrate to the active site.

(Refer Slide Time: 03:27)



So, in our target based approach what do we do? We identify the target. That means we know the disease mechanism, you know which enzyme I want to target or receptor or ion channel then we get the determine DNA and protein sequence, 2 dimensional sequence like we looked at then look at the elucidate structure and function of the protein. Prove therapeutic concepts in animals. Develop assay for high-through put screening. Mass screening, mass screening and then select one or more lead structures.

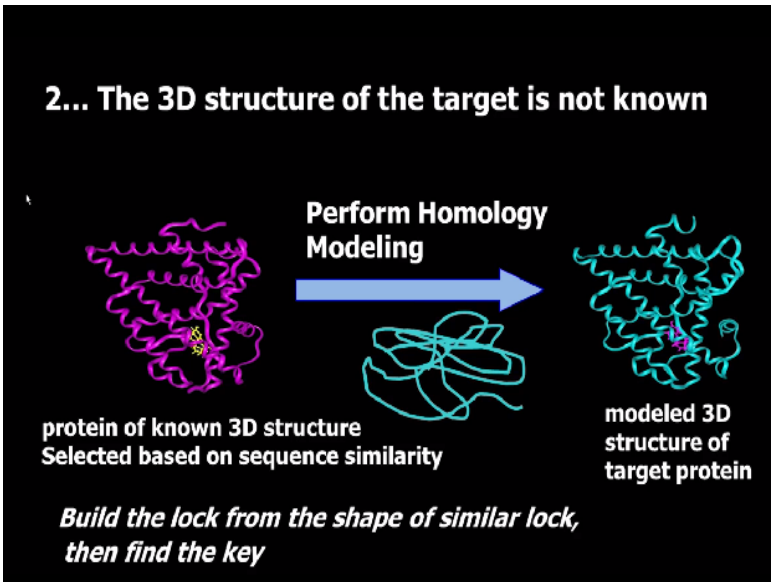
(Refer Slide Time: 04:03)



So, if I know the target protein. I know the active site and so I design or screen large number of molecules so that they go and bind very efficient. So, I it is like finding a key for the locks so we test 100 of keys and find that one particular key nicely fits into it and opens the lock. So life is

very simple. I have the 3 dimensional structure of my protein or target. I know the active site all. But life is not always so simple.

(Refer Slide Time: 04:35)

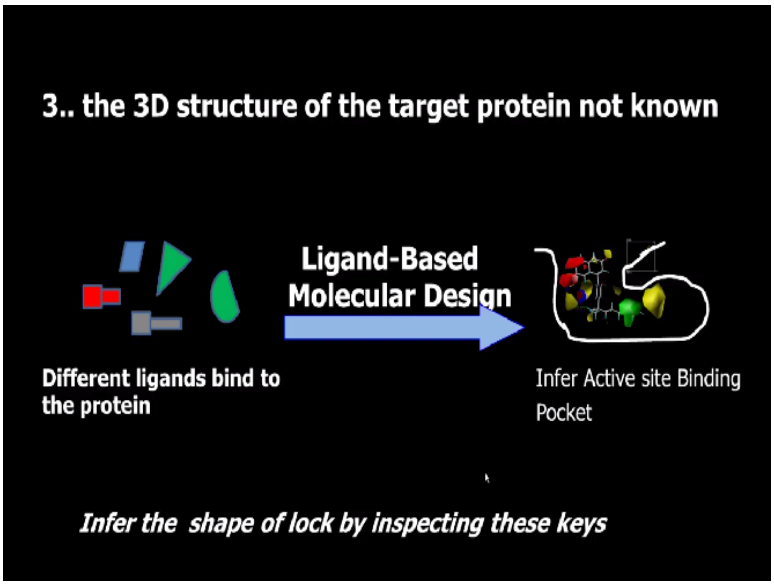


So, second approach is I know the protein sequence so that means I do not know the 3 dimensional structure. So, what do I do? I do something called homology modelling and then create 3 dimensional structure of the protein of my interest. The homology modeling is created with respect to another protein which has sequence similarity we talked about sequence similarity sometime back.

Okay, it is like building the lock from the shape of similar locks. So, there is one lock I have I want to build based on some other similar lock. Once I do that I have here target protein 3D structure is built. Then I make the key that means again I perform the ligand based docking and identify the key. So, I built the 3D structure of the protein target based on the protein sequence and based on similar protein structures and then I built the key.

The third approach could be I have lot of ligand which seem to work on the particular protein.

(Refer Slide Time: 05:55)



It is like I have lot of keys which seem to open the lock. So, I get the 3D structure of the active site of my binding based on structure of all these keys. So, I have lot of keys and I show based on the keys shape and size I will try to predict what will be the active site, shape and size. So, that could be another approach. So, these are all different approaches one follows in the target based design. So, this approach is generally common.

We may have the 2 dimensional sequence, amino acid sequence of the protein of our interest. So, we prepare the homology 3D structure of that and based on sequence similarity once I have the 3D structure I do the refinement and then I go into the docking of various ligands from database to the active site of the created protein.

(Refer Slide Time: 07:01)

Name	Developed by	Free for academia
AutoDock	Scripps Research Institute	Yes
DOCK	University of California, San Francisco	Yes
FlexX	BioSolveIT GmbH	No
FRED	OpenEye Scientific Software	Yes
Glide	Schrödinger, Inc.	No
GOLD	Cambridge Crystallographic Data Centre	No
ICM	Molsoft LLC.	No
Surflex-Dock	Tripos Inc.	No

So, there are many software, commercial software, not commercial software are there but they AutoDock. This is the one of the best. It is free for academia one of the best. Then we have the DOCK. This is also free for academia then we have BioSolve, FlexX, no FRED no. As Glide, no. GOLD again, no. ICM, no. So, there are many commercial software which we need to pay. They may have different rate commercial, rate for academia and industry.

But there are free softwares like AutoDock is one of the best software which can do the docking and it is widely used so we can use this software. I am glad that I will also you demo of how to dock, actually go along.

(Refer Slide Time: 07:54)

AUTODOCK

http://autodock.scripps.edu/

AutoDock is a suite of automated docking tools. It is designed to predict how small molecules, bind to a receptor of known 3D structure.

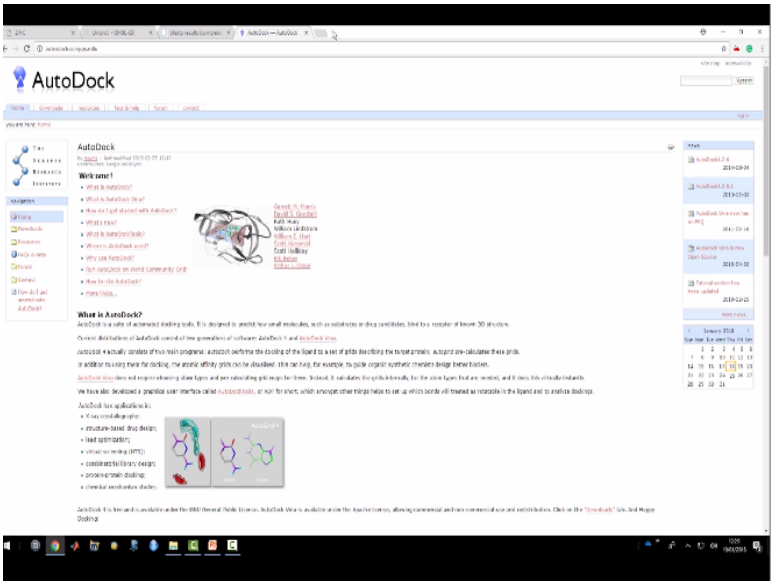
Current distributions of AutoDock consist of two generations of software: AutoDock 4 and AutoDock Vina

So, AutoDock this http://autodock.scripps.edu/. This is a suite of automated docking tools. I will show you later it is designed to predict how small molecules, bind to a receptor it can calculate binding energies. So, we have AutoDock 4, AutoDock Vina which can do the job.

(Refer Slide Time: 08:14)

Then we also have Swiss Dock, www.swissdock.ch this is another free software which can do the job of docking. Both are very good softwares which we can use. As I said, I will show you some examples of how to do docking using both AutoDock and Swiss Dock.

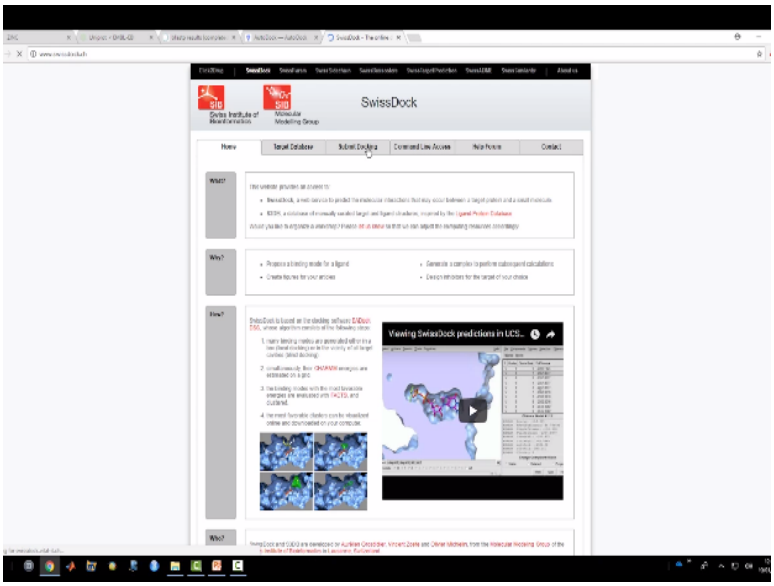
(Refer Slide Time: 08:44)



This is AutoDock. So we can download AutoDock version 4 and for academia as I said it is free so we can use that without any problem and it is one of the best. We also have something called

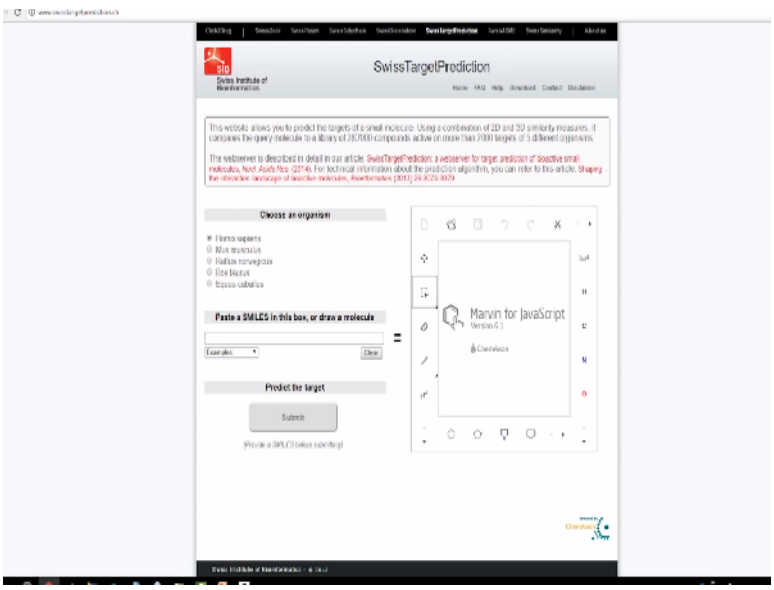
as Swiss Dock and that is again it is a free software. You can do docking.

(Refer Slide Time: 09:24)



As you can see we can do the docking. We can give both the ligand.

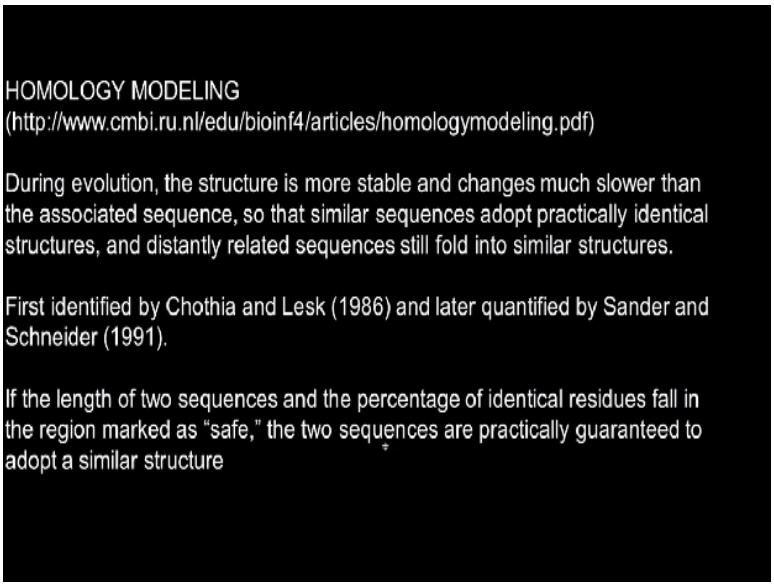
(Refer Slide Time: 09:33)



Here we can give the target protein. Here we can give the ligand and then we can give your job name. We can give the email id and then I will get the results. Where as in AutoDock once you download just stand-alone whereas Swiss Dock is through webserver. AutoDock is standalone we can download into our computer then set it up and then run it. We can get files docking log files, dlg files which give you the results of your docking.

And it gives you all the non-bonded interactions, give you the binding energy and so on.

(Refer Slide Time: 10:11)

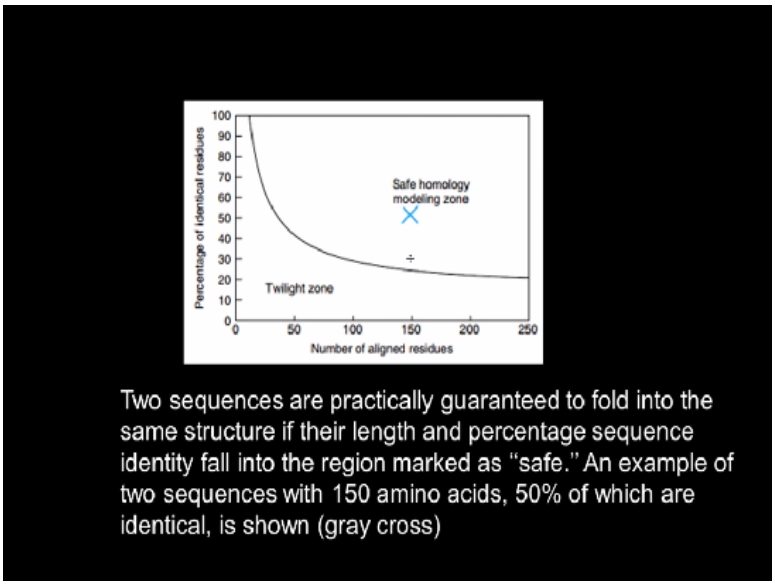


So, it is very powerful. Now, if you look at homology modelling, Homology like I said is the most important thing. We need to look at the 2 dimension sequences of, the protein of you interest with other proteins found in the protein data bank. What is the concept of homology modeling? So, this information is available in this PDF. I am just giving you summary of that. During evolution, the structure is most stable and changes much lower than the associated sequences.

So, the structure remains almost the same even if there are small changes in the sequences that is the finding from evolution. So, similar sequences adopt practically identical structures and distantly related sequences still fold into similar structure. So, that means if the similarity is even 30%, 40% we can say 3 dimensional structure will be almost the same. Even if sequences change the structure is most stable.

This is first identified by Chothia and Lesk and later quantified by Sander and Schneider. So, if the length of 2 sequences and the percentage of identical residues fall in the region marked as 'safe'.

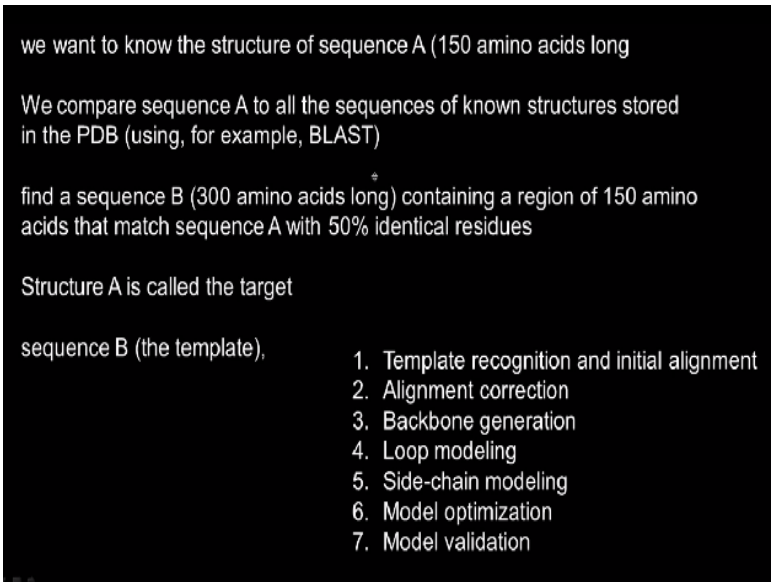
(Refer Slide Time: 11:36)



The 2 sequences are practically guaranteed to adopt a similar structure. So, if the number of aligned molecules is 150 then this region you can call safe homology model and percentage of identical residues, number of aligned residues. So if I have 100 adding it about 30% the chances are the 3 dimensional structure will be on this. So, 2 sequences are practically guaranteed to fold into the same structure if their length and percentage sequence identically fall into the region.

So, for example, if you have 150 amino acids and if it is 50% percentage sequence that is fantastic we can say both will have a similar structure.

(Refer Slide Time: 12:21)

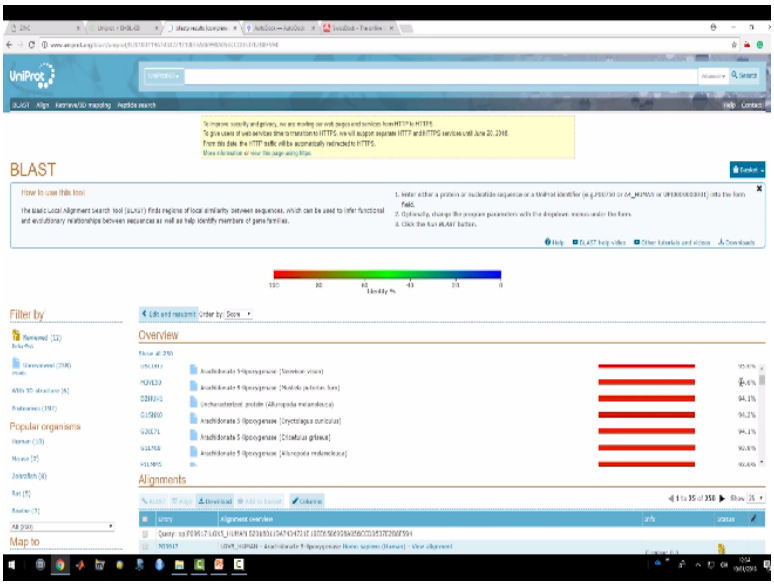


Even we can go down to 35%, still we can call it a, they will come to the same structure. And so

in homology remodeling what is it down? So we want to know the structure of sequence A amino acid is wrong. So, we compare is sequence A to all the sequences of known structure stored in the PDB using the blast command, you just type the BLAST command. Now, we find the sequence B which is 300 amino acids is long containing a region 150 match exactly.

So, the sequence identity is 50, sequence A is called the target, sequence B is called the template. So, we built structure of A based on the template B. But once we do the blast. In the previous class, I showed you how to run the BLAST. I will show you the BLAST. So, I will show you the blast here.

(Refer Slide Time: 13:26)



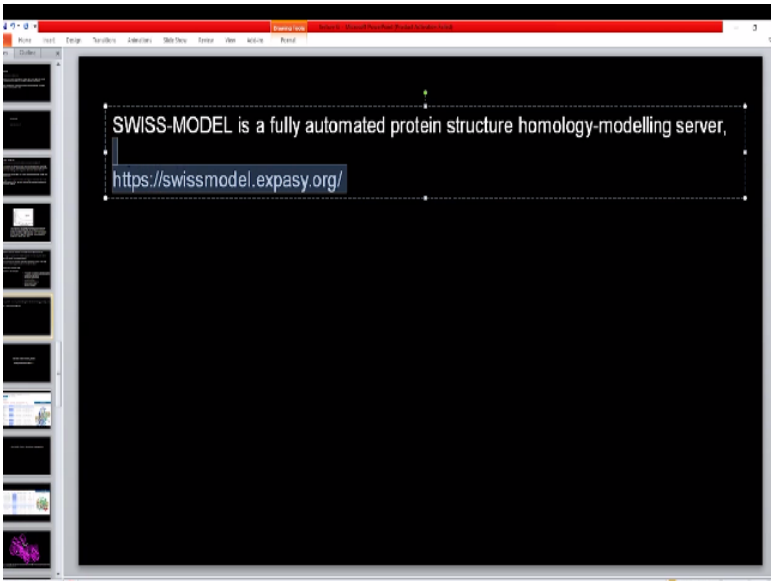
When you run the BLAST for example, I run for 5-lipoxygenase Blast. Tells you what are the various protein sequences which match starting 100%, 98% you can see 5-lipoxygenase from different animals gorilla, macca and so on. So, you can see quite high. So, we can take anyone of these proteins which we call it the template and then built the target protein. And as mentioned before, the concept of evolution that if the sequence percentage identity is greater than 35%, 40% chances are they will adapt similar structure so we can get the target.

So, even anyone of this protein also we can take because they are very high sequence identity as you can see. So, the blast is the first step which we run to identify what are the various proteins which have very high sequence identity. As you can see here it goes very high. So that is called

the template and the sequence.

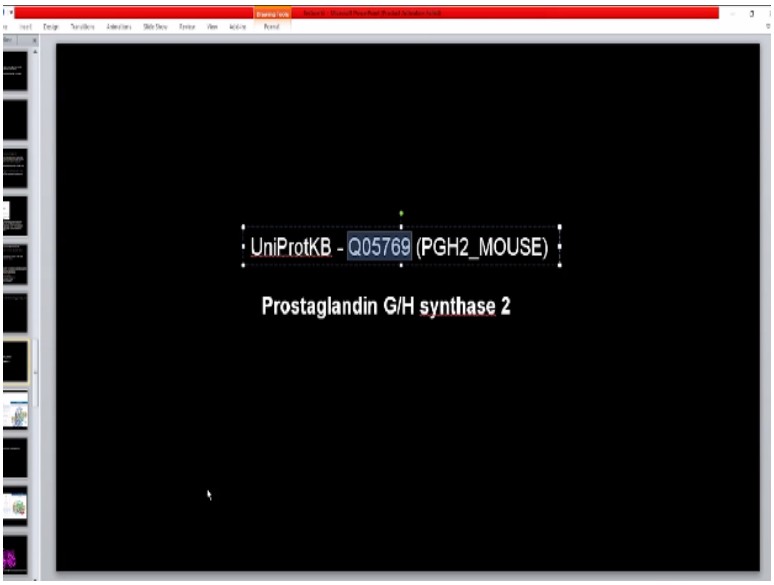
Using the sequence A, you are building, using the templates so that it is called target. So, we do the template recognition, initial alignment, alignment correction so we have used the backbone. Then we do the backbone generation, Loop modeling, Side-chain modeling, Model optimization, Model validation all these things are done when we do the homology modeling.

(Refer Slide Time: 15:19)



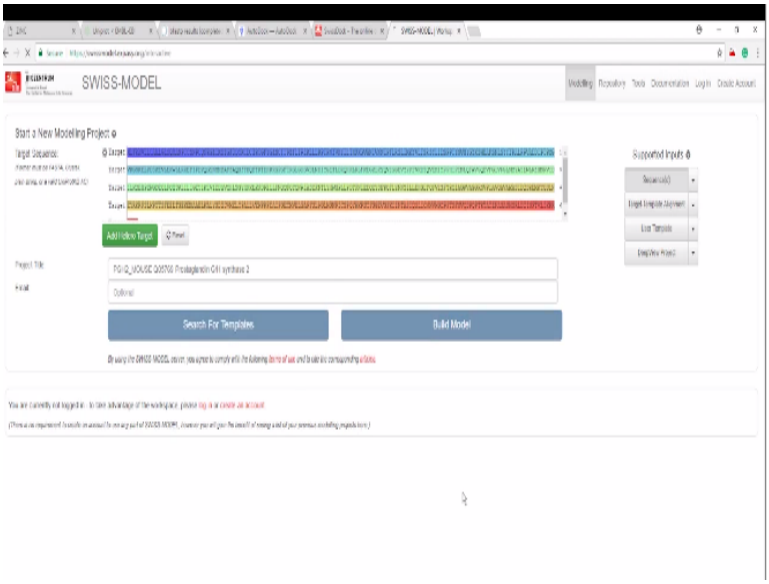
So, first you compare the sequences then we do the homology modeling. So, for example we can Swiss-Model software. It is a very good software for homology-modelling. Homology-modelling does a good job. So, we can look at Swiss Model. They can see start modeling so we can give the upload target sequence, we can upload target sequence then search for templates.

(Refer Slide Time: 16:05)



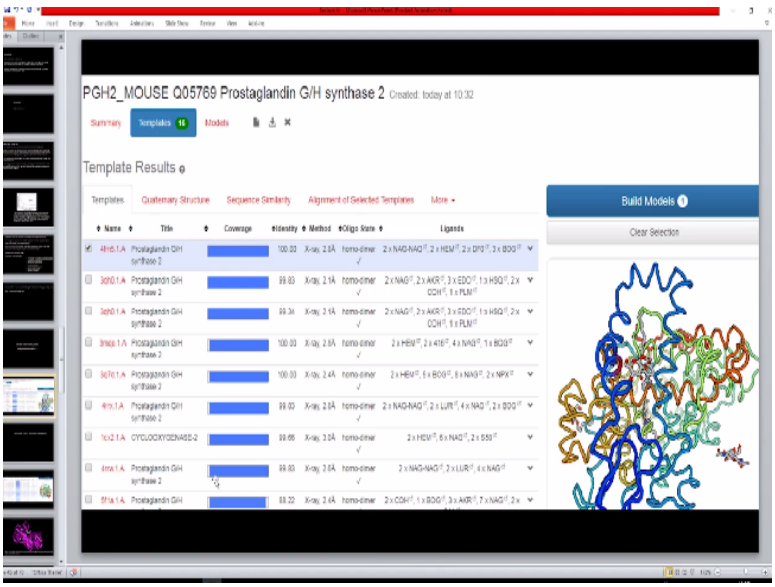
So, we can upload the target sequence. This is a prostaglandin PGH2 and GH2 synthase for the Mouse.

(Refer Slide Time: 16:22)



I can input that. This is the sequence of the target then you can search for template. So, it is doing the job, template results and it keeps doing that. Take little bit time.

(Refer Slide Time: 16:39)



So once it does it gives you the result here. So, it says different identity. We have a Prostaglandin synthase, another template, another template like that. It gives you the sequence identity also. It also tells you how the structure was obtained using x-ray, the resolution of 2.1 Angstroms, 2.83 like that. It takes little bit more time. It is taking time. So, it gives you lot of things. So, we can select one of them as our template and then model your system.

So if we click on any one of them, you get a 3D structure of your template. It is time as you can see. We can do for any protein. Look at this protein this has got UniProt number of again we can run this also no problem. We run this protein now. Start modelling. So we can give UniProt number. Now this protein is 5-Lipoxygenase human 5-Lipoxygenase which converts arachidonic acid into 5-htte. So, we can look at other proteins and with that also okay. This is the sequence and then we can search for templates.

This is the target and this is the template in the previous also. It is still running for a Prostaglandin synthase. We can search for template so it can do also, this also take time to come. Yeah it takes time so as you can see, keep doing it. Predicting alignment state conservation so it keeps doing that. Takes a little time because it is going to look at large number of proteins and then see how that sequence comparison is and then we can select one of them. So, we will do that we will continue.

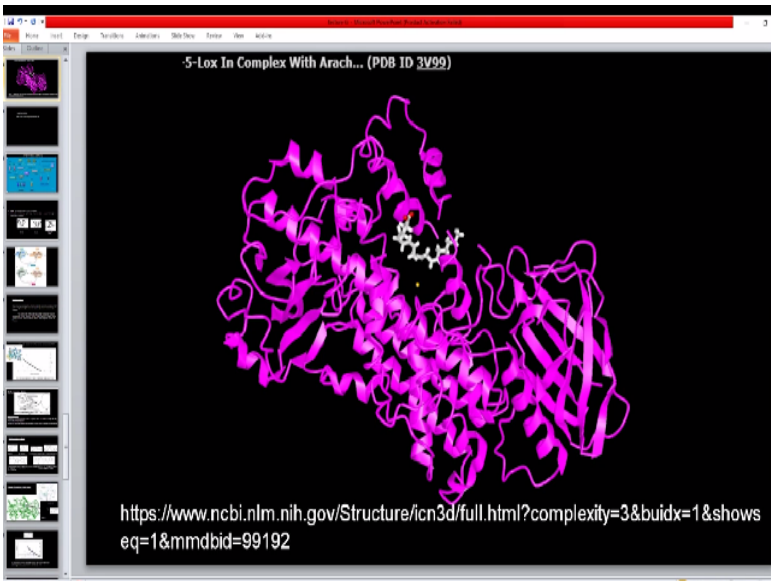
So, if you do for the 5-Lipoxygenase you may get a series of proteins now with the different percentage identity as you can see here. And the 3 dimensional structure how it is determined x-ray, (()) (19:48) with this is the method used. What are the ligands present? It contains ions because the Lipoxygenase contains ions also and then I think it contains arachidonate here.

The Prostaglandin synthase enzyme PGH2, it also gives you different ligands that are found here. That is a dimer and if you look at your mostly predominantly monomer. So, it gives you lot of information you can select the protein of our interest. Yes, we got some results here when we try to align the sequence of prostaglandin to synthase of course it is only one. We can look at this one also. For example, so, we got this protein as you can see here.

There is the protein at the template which you selected so we can give a homology model with that protein if you want. Then we can say build models with this particular protein. Build model this protein that will be who are in modelling job. That is your template, your target is your original sequence which you gave. So, it will build a homology model as you can see here. So, here it is also given for the 5-Lipoxygenase so it is given for 5-Lipoxygenase as you can see there will be an ion because Lipoxygenase contains ions.

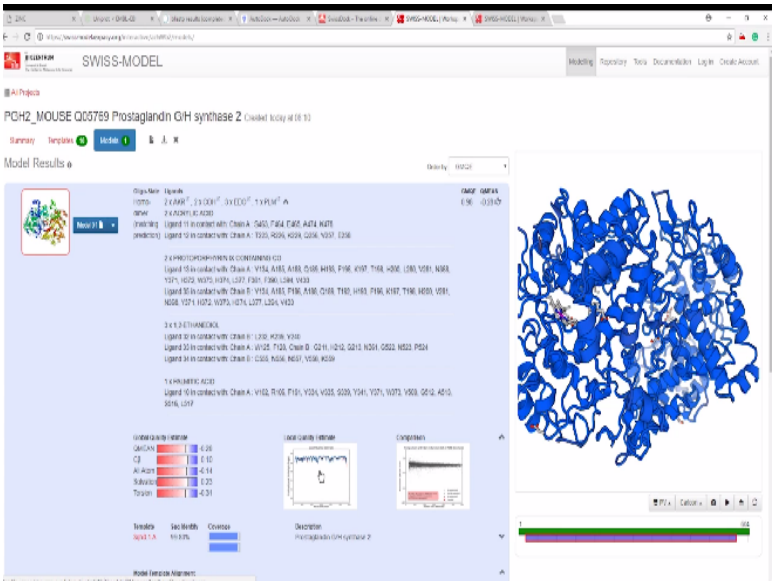
So, as you can see here, the ion is here and ion is here there. So, in this case we have said build model so if you start there building models we will see that little bit later. So, the first step is aligning your sequence with the known proteins in the data base. So, we can use this particular software is still good. This is model software once we take the best which we feel and then we can say build the model. So, we get homology model using that template.

(Refer Slide Time: 22:36)



Just got an arachidonic acid also.

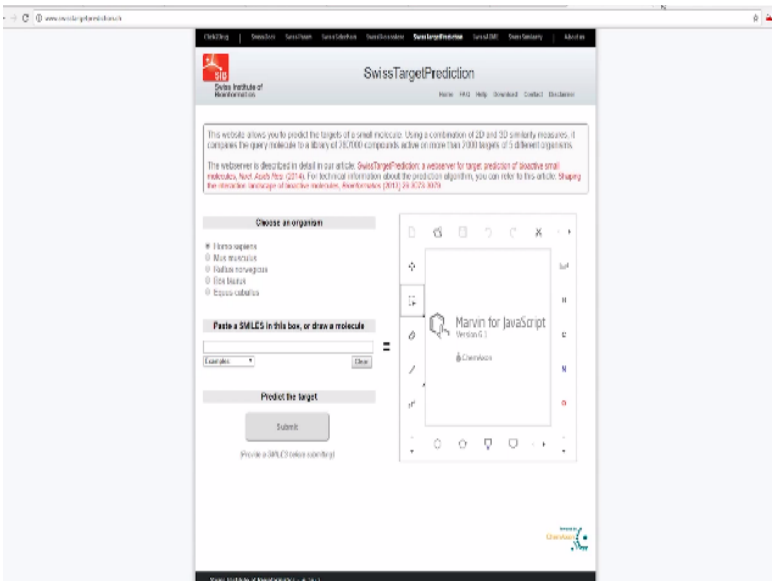
(Refer Slide Time: 22:55)



It is built a model. So it tells you the quality of the model. It tells you how the quality of the model. Both the local similarity with the target and template and then it also compares. Number of residues has been compared as Z score. So, it compares your target model with using the template and so we got a model based on the 2 dimensional sequence. So, Swiss model is very good software it makes use to identify and get 3D structure.

There is also another software that is called target prediction and that is called using Swiss target, given a ligand tells you what are the possible targets for the ligand.

(Refer Slide Time: 24:30)



So, in home sapiens for example. So, here we can place a smiles in this box and try a molecule. So, it tells you what are the different targets and possible targets making you the structure activity type of relationship. How do I get those smiles? for example, let me take Aspirin. So this is the smiles for the Aspirin. So, go to the target prediction database. So, we paste the smiles then this is the Aspirin as you can see acetylsalicylic acid. Submit.

So, it gives you possible target for Aspirin based on the structural similarities. So, this is also very useful data base because if you are looking at side effect profiles of molecules, we can see what could be the possible side effects because the molecule which you design may go and bind. So, it is given the targets and you can see prostaglandin because Aspirin comes in arachidonic acid path way of inflammation. COX-1 this is called prostaglandin synthase 1 and synthase 2, COX-1, COX-2 it binds to COX-1 and COX-2 enzymes and then other sodium dependant transporters.

It gives you many other possible, mostly transporters, close to lot of transporters here and enzyme predominantly 33% and then transporters 47%. So, it can go to lot of transporters as you can see. So, the server is also very useful because it gives you possible targets so probability almost > 50% here and you can see here of course COX-1 and COX-2 200%, tt is well known other transporters 50%.

So, it could be lot of transporters into which Aspirin may go and bind. So, this software is also very, very useful for us to make use of as we go along. So, we have seen many software homology models and then building your target 3D structure, making use of the template. So, we need to know how to make use of the docking software especially the AutoDock. So, we look at how to make use of AutoDock in the next class.

Because AutoDock is one of the software which can look at ligand protein docking and once we have the 3 dimensional structure so we can make use, get use of homology models software to align the sequence of your protein with the known protein 3D structures and then with those templates we can build the 3D structure of your target and then get a good structure and then we go to AutoDock and then they see how that docking takes place.

We can also estimate the binding energies and so on actually. So, we will continue more on this topic of docking in the next class. Thank you very much for your time.