Introduction to Dynamical Models in Biology

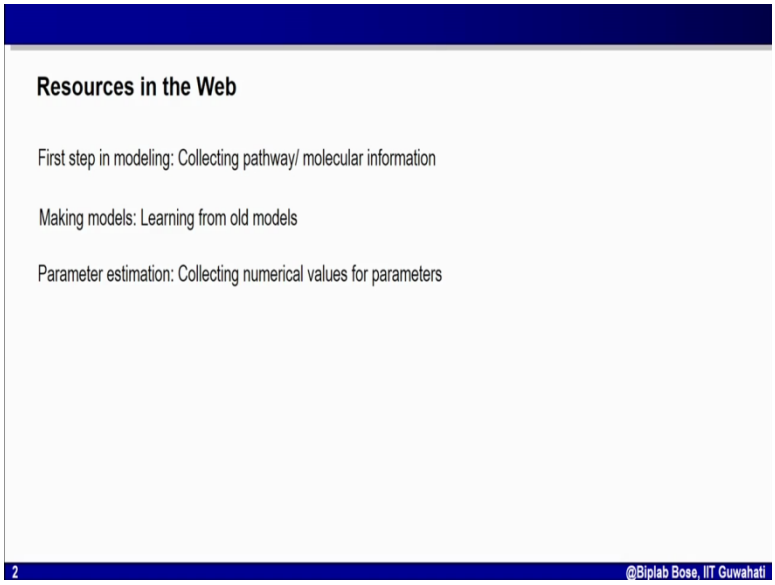
Dr. Biplab Bose
Associate Professor
Department of Biosciences & Bioengineering
Indian Institute of Technology, Guwahati

Lecture 24

Online resources for mathematical modeling in biology

Hello, welcome to the last module of our course. Through over this 4 weeks we have discussed how to use ODE's Ordinary Differential Equation to create mathematical module for biological system involving dynamics. Then we have discussed how to analyze those module. We have discussed how to use numerical symbolization to solve those problems and then we focused on particular type of module for example modeling, signaling path ways, involving positive feedback, negative feedback etc and also transcriptional network. In this last module I will discuss about certain useful resources in the web which you may find very useful while creating mathematical models for biological symbols particularly for cellular models involving cell signaling, transcriptional network.

(Refer Slide Time: 1:29)

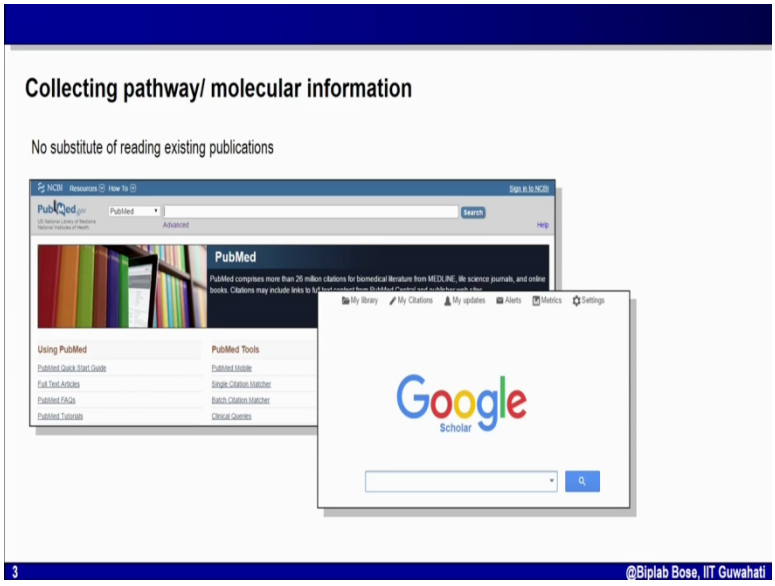


So let us start I will discuss on three specific topics for which you may have to look back into the internet to collect information. First one is to get information about the molecules and the rules involved in the processes. Suppose you want to create mathematical module for a particular signal transaction pathway for that you need to know which are the molecules involved in this

pathway. You have to know the cellular compartments where they are working at the same time you have to know who controls whom that means you have to know the rule of the game. We often begin to the web to collect this information. Secondly we may look into the web to look and find out old mathematical modules. Old mathematical modules are very useful thing to start our modeling we can recycle existing model to fit our problem and answer specific question.

So very often we look into existing model which are available in the net. And the third one is collecting information about parameters if you remember in the Ordinary Differential Equation I will have dependent variable I will have independent variable at the same time there will be some constant terms of the parameter values. For example in the enzymatic reaction I had to know the rate constant which are the parameter in the process so, I have to know the numerical value for this rate constant or parameters what do I get those parameter values. Often we go back to the web and use different web resources to collect those information.

(Refer Slide Time: 3:20)



I will start discussing with collecting information about molecules and pathways or the components of the model. Every model that is cellular level or at high level are made up of multiple components. At similar level these components are essentially molecules involved in the pathways and processes. Where do I get the information about all these components which are relevant for my modeling the best one you have to always remember is digging into existing

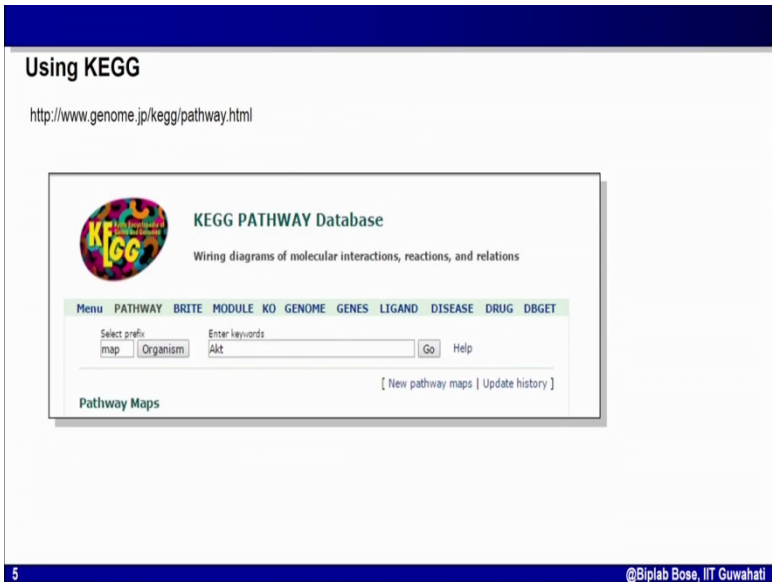
literature, there is no other alternative to it. We have to collect all the relevant publications and we have to rigorously read and distill information from this publication and one web resource which comes very handy for this purpose is obviously PubMed the screenshot of which I have shown here.

PubMed is a repository of abstracts of our publication over the ages and it also connects to the publication the paper if it is available freely and most of us who work in research in biology or will acquainted and use PubMed regularly. Some of us prefer Google Scholar to dig into the published literature and find out relevant paper. Whatever resource you use ultimately if you all want to create a successful model you have to read and distill from this published papers.

(Refer Slide Time: 4:49)

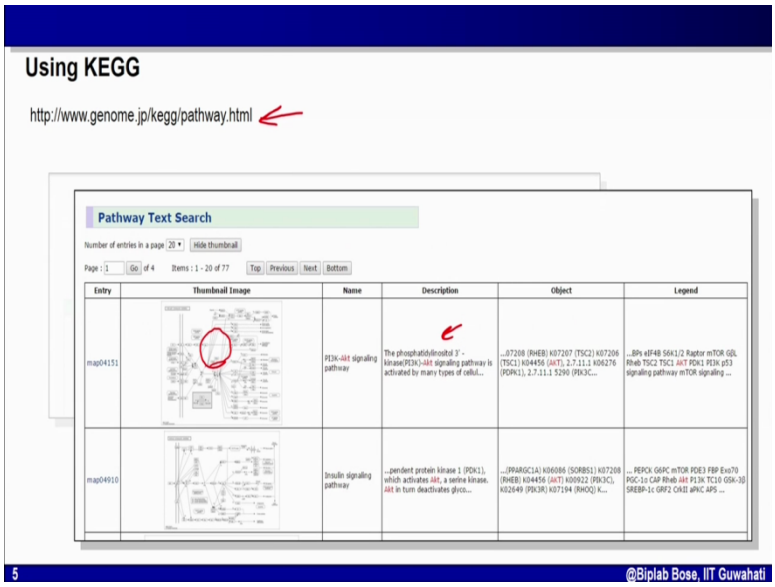
Now when we are trying to create mathematical models for cell signaling pathway or transcriptional network usually these pathways and networks involve multiple component and it becomes sometimes bit confusing to distill from literature and then create the model. And to help you in this purpose there are certain databases which actually curate manually information about different pathways. I have listed some of them here for example one is KEGG database I use it regularly, then there is Reactome which is also very rich curative database then you have Panther, WikiPathways and many more I have just listed 4 here and I will discuss KEGG database in details because I find it very useful in day to day work and modeling cell signal.

(Refer Slide Time: 5:40)



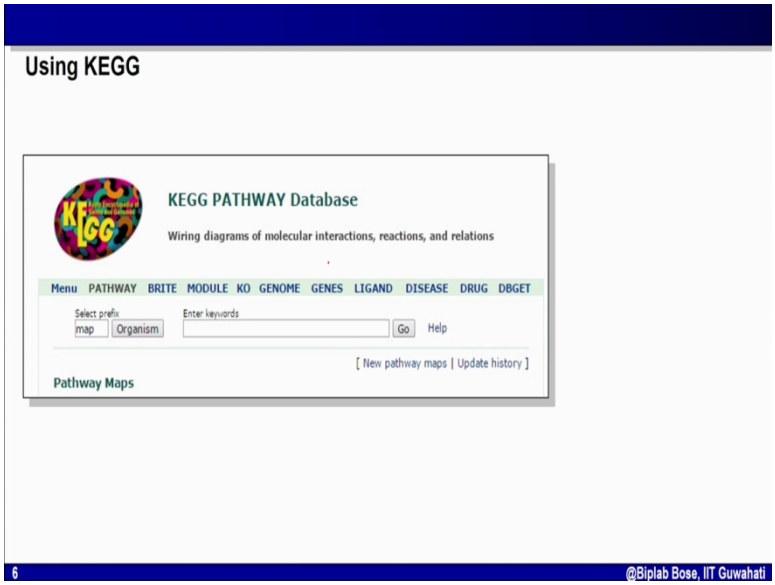
Let us look into KEGG database I have given the web link here if you go to this page you will find lots of information there for pathways. Remember KEGG is a manually curative database that means some people are reading to the published literature and distilling out information and drawing the pathway diagrams and curating that information in organized fashion. You as a user can search through this database and get relevant information. For example suppose I want to know all the molecular pathway in which the molecule Akt which the kinase is involved so I can write Akt in this search text box and then click go so that I get comprehensive result of all pathways where Akt is involved.

(Refer Slide Time: 6:35)



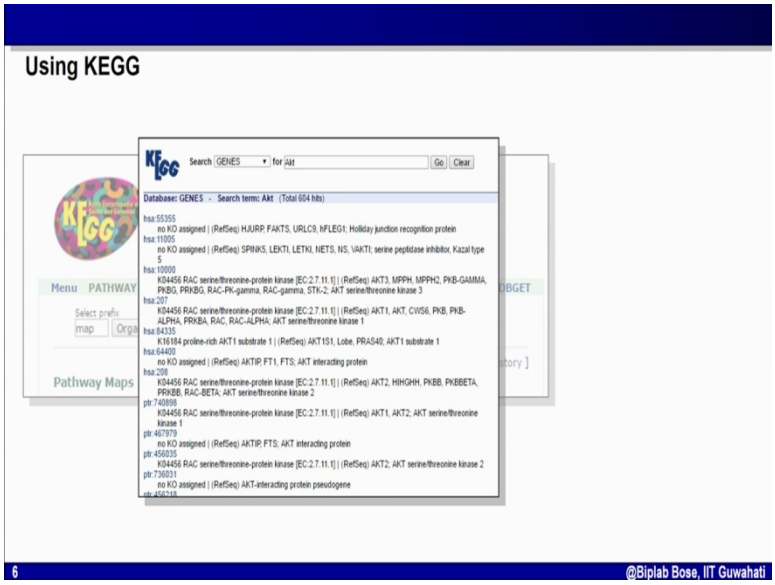
And the result will be like this in new page you will get a table where comprehensively all the pathway where Akt is involved as per the KEGG database will been shown. You can look into this for example this one is a particular pathway for example written here PI3K/AKT signaling pathway. If you click on it it will open up and you will get a expanded view of the pathway. I will show you separately.

(Refer Slide Time: 7:08)



The other way of searching is that from main page you notice there is something called genes so I can search for a particular gene so, if you click on that you will get a new window.

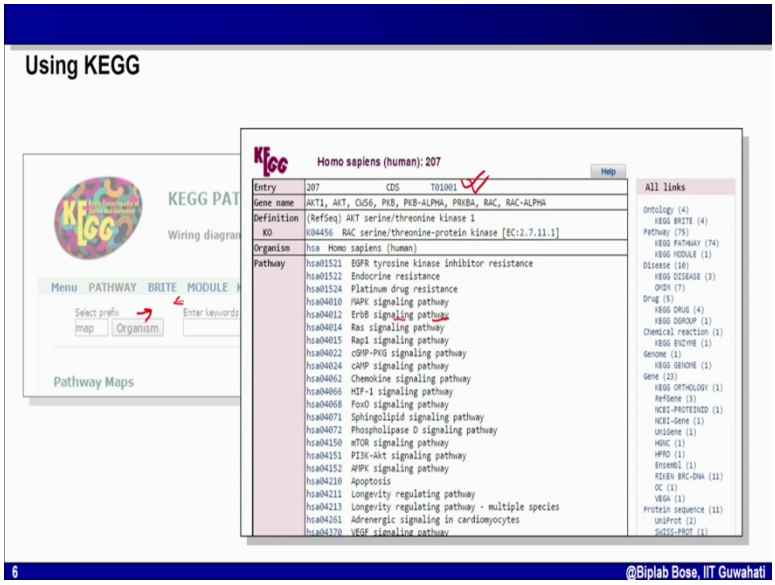
(Refer Slide Time: 7:15)



So, here you can type the name of the molecule for example I am type her Akt and you can also type the name of the organism for example I have written HSA that is human and then if you search by clicking this go you will get detail list of search result. Here in this window all the places in the database where Akt is present is listed so, it has listed all other molecules which are close to Akt or has similar name. If you find here this is what we are looking for. This the entry

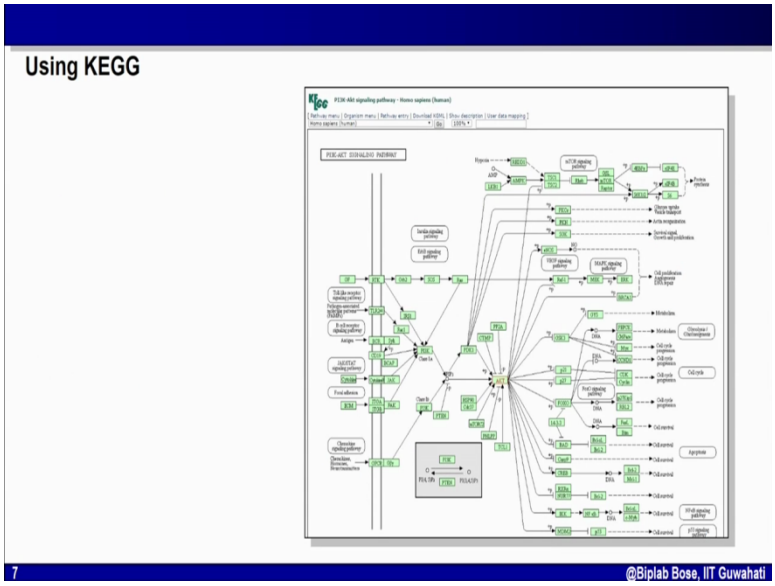
which I am looking for because we are looking for Akt or the synonym of that is PKB.

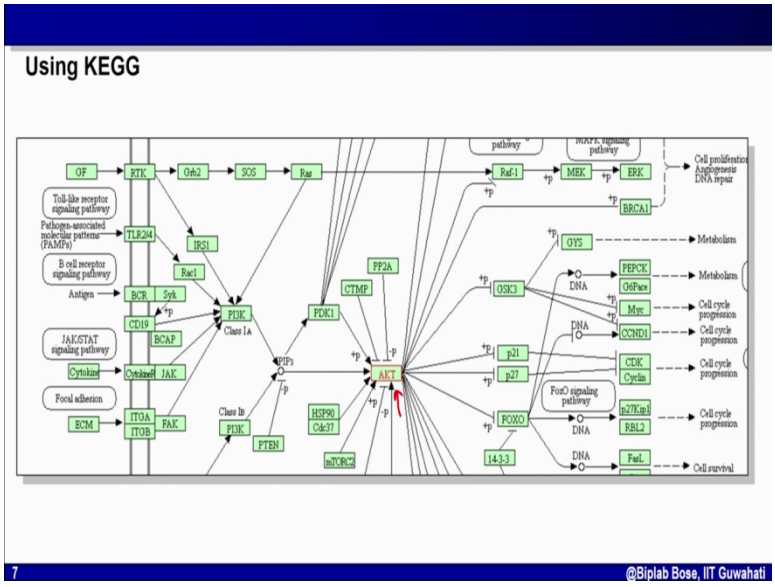
(Refer Slide Time: 8:09)



So, if I click on this entry it will give me a list where I will have detail information about this molecule as well as I will have information about the pathways in which these molecules are involved. So, here it has listed all the pathways in which the Akt molecule is involved, and suppose I am interested to know the PI3K/AKT signaling pathway so I will click on this particular entry so that I can a view of the pathway.

(Refer Slide Time: 8:43)

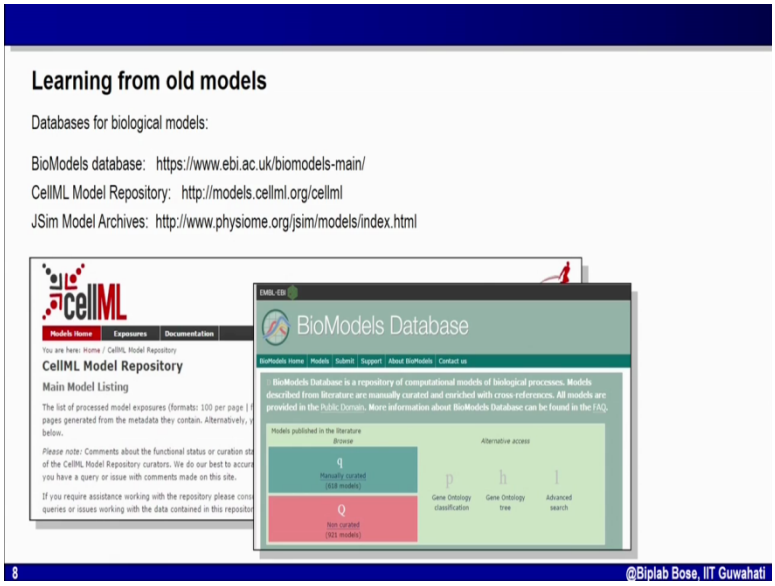




So let us see what will happen if I click on that. If I click on that a pathway map will appear and you can easily see I have Akt here in red. If I zoom you can see what we have here I have Akt that the molecule I am looking for the map is saying if I give a growth factor GF that activate RTK that is Receptor Tyrosine Kinase a particular group of receptors, that will activate IRS1 intermediate molecule that in turn will activate PI3K, that will activate and or rather produce PIP3 which activate by shown by this arrow by first phosphorylating Akt.

Once Akt is active it can inhibit GSK3 by phosphorylation this plus sign means phosphorylation this hammerhead means inhibition so active PAKT phosphorylate p21 and inhibit this one and this way you can easily see which molecule is controlling which molecule and how the control works. Either phosphorylation, dephosphorylation or the transcriptional process if the molecule is the transcription factor. So, in this way you can search through KEGG database and collect information about different molecules involved in the pathways and how they are connected with each other. So, whatever information you collect from literature gets supplemented by the information present in the KEGG database.

(Refer Slide Time: 10:23)



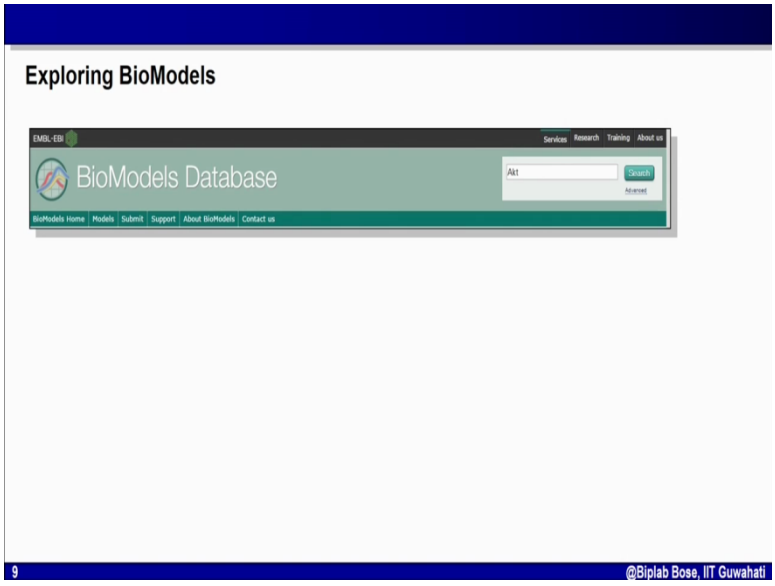
Now let us go to the second issue, if you remember the second issue for which we will prefer to go back to the web and search through it is for old models. Suppose I am creating a new model for a particular signal connection pathway say for insulin signaling in glucose metabolism. I have to write down some ODE representing for this model. It will be very helpful if I look into the literature where somebody else has previously created similar model and I try to use that type of equation so in our, my model. I sometime may prefer to collect information about parameter values used in those models.

I may sometime borrow the whole model and here and there and create a new model. That means I have to recycle the existing model. For this I have to go back to old models. One way of going

to old models in look into the literature but nowadays there are many databases which has come up where scientist store their model submit their model which they have published in the paper and those models are manually curated and share through the database so anybody can go back and search those model database and retrieve the existing model and use it for their own purposes. Many such database exist.

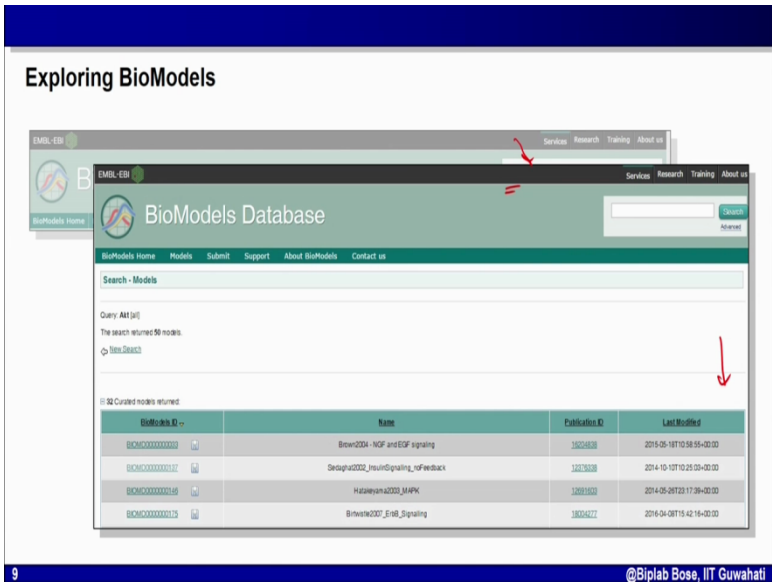
I have listed few databases here for example one is biomodel database I will discuss about it bit more. Then you have CellML database and then JSim web page if you remember it has archive of models. So, Instagram you go to the JSim web page you will find archive of models where you will find lots of model not just based on JSim tool but used by developed by other groups using other tools but available there in that archive and you can actually re-chip those models and run in JSim and play with them.

(Refer Slide Time: 12:27)



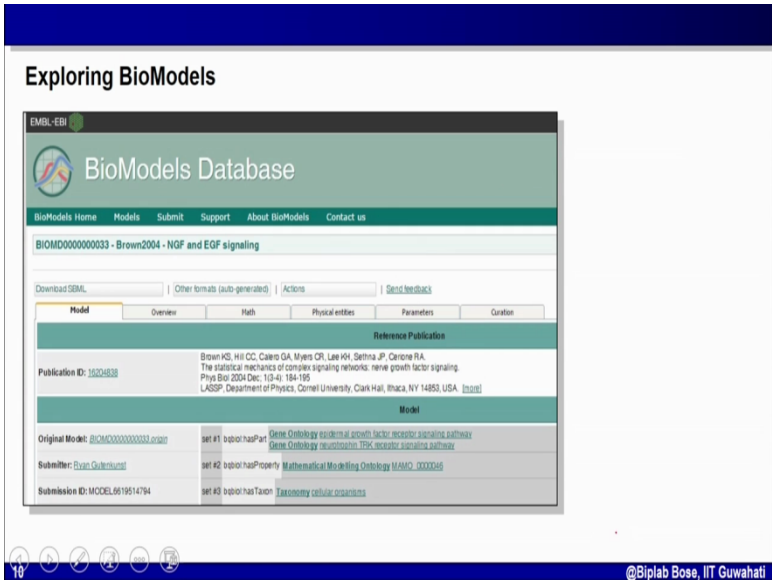
So, let us look into biomodel database. If you go to the biomodel database there is a large page the most used tool part for us is this search window where you can directly put a entry for example I have written Akt here because I want to search all the models submitted and curated in database which involve Akt.

(Refer Slide Time: 12:53)



So, if I do that I get search result which gives me a long list which I have truncated here and shown. In long list of all curated model where Akt is involved. So, for example suppose we want to see in detail about this model which involves NGF that is not growth factor and EGF signal.

(Refer Slide Time: 13:14)

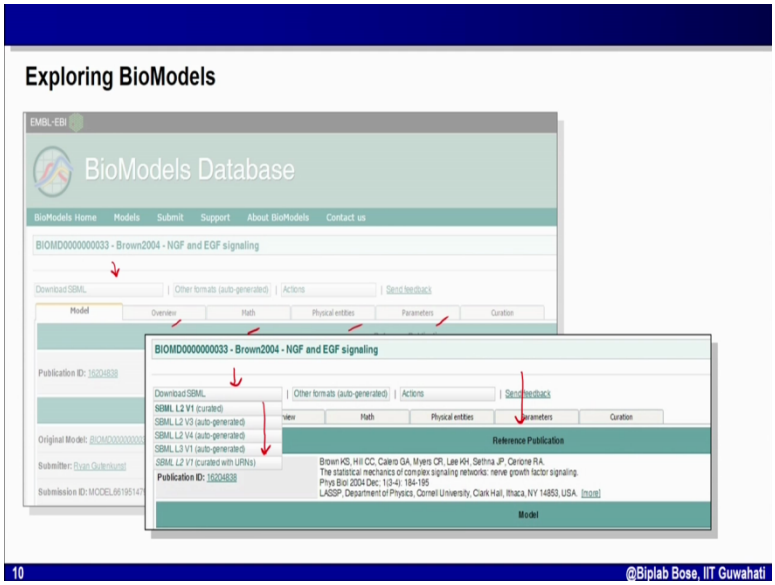


So if I click on this biomodel entry this is biomodel ID for this model I will get detail information about this model. So, what information I get I have detail information about the

people who created the model, publication where it has been published and all others. My focus would be primarily on this download option, overview, map, physical entities, parameters these tabs. These information present in the separate tabs are very useful. For example I may want to download this model and simulate it myself in my computer. Now remember there are not only one single modeling framework. If you are using Ordinary Differential Equation to create model you can use different tools to create that model.

We have extensively used JSim somebody maybe using C, somebody maybe using FORTRAN, somebody maybe creating MATLAB, using MATLAB to create their model. Now the issue is how will you share this models. One way of sharing this model is to convert them in a single language and then share it. And that single language is systems biology markup language so all model submitted in this biomodel database are actually converted into model made up of system biology markup language and you can download that model in SBML language and then you can export that to your software.

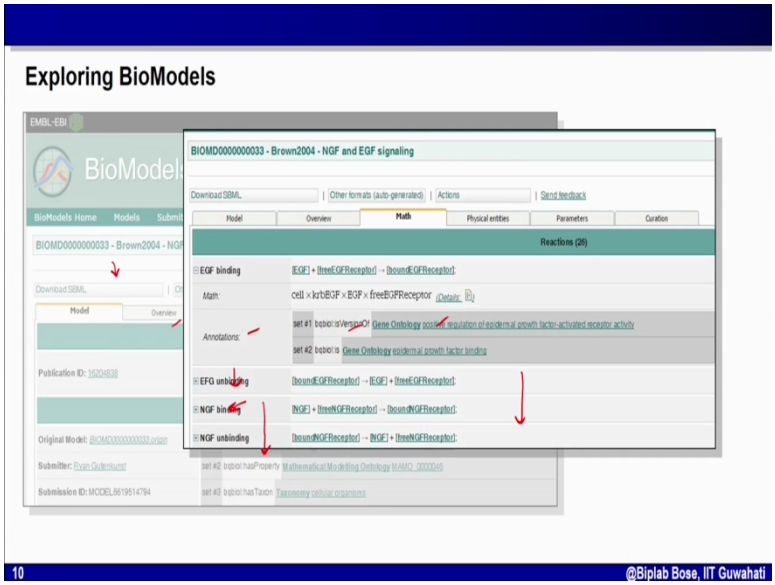
(Refer Slide Time: 15:03)



For example I can download the model from this database for this particular NGF and EGF signaling by clicking here, download SBML and you see I get down multiple option to download the same model in SBML format. Suppose for example I can download this one SBML L2/harshal1 curated that means it has been manually corrected and curated. So, if I download

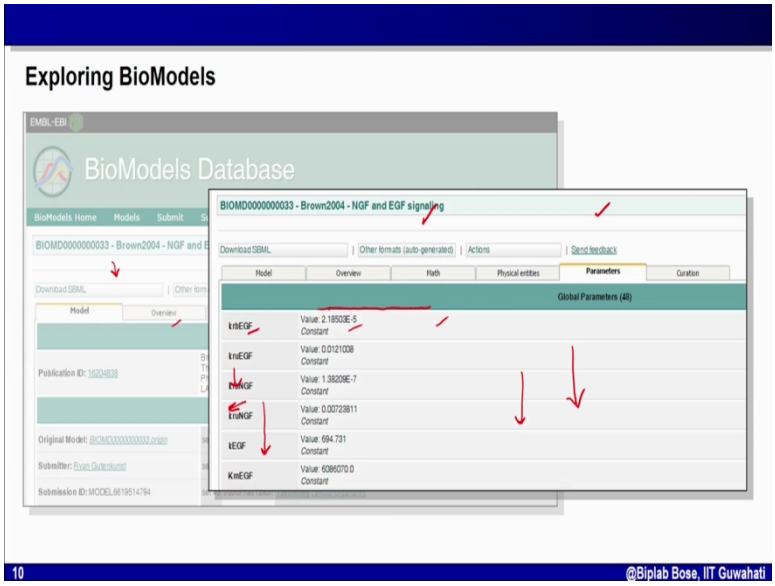
that file I will get model written in SBML and if you remember actually JSim can read SBML file and then it can be simulated using JSim. So, I can download this one and then open it in JSim and simulate myself and then I can manipulate if I require is very useful.

(Refer Slide Time: 15:41)



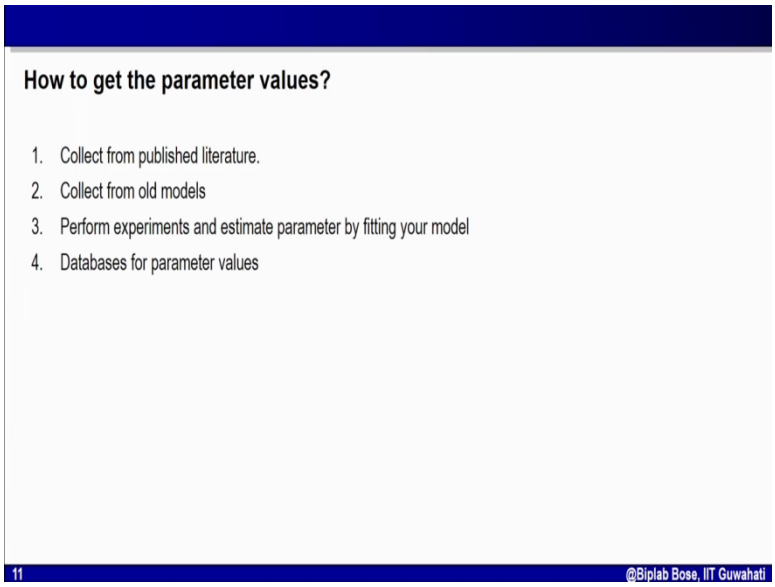
Let us look into other important tab for example the map tab it list all the unit processes involve and the mathematical reactions involved there. For example EGF binding is represented by these red constant into EGF into 3G receptor. So, all the mathematical formulation which are used to create the Ordinary Differential Equation representing weight of change of some molecules are listed in this map tab.

(Refer Slide Time: 16:17)



Then you have this parameter tab if you click on this parameter tab what you get? You get numerical values of all parameters using this model. For example this one, this one, this one are all kinetic parameter constant values used for this model. So, this tab is very useful. For example if you are looking for a parameter for a particular process in your model and you are not getting that information you can dig into biomodel database and if you find this suitable similar model you may find out the parameter value from this particular tab in a model. So, in this way I can get the model with information the mathematical equation that makes the model the parameter values and as a whole I can download the model in SBML format and use it as per my purpose.

(Refer Slide Time: 17:16)



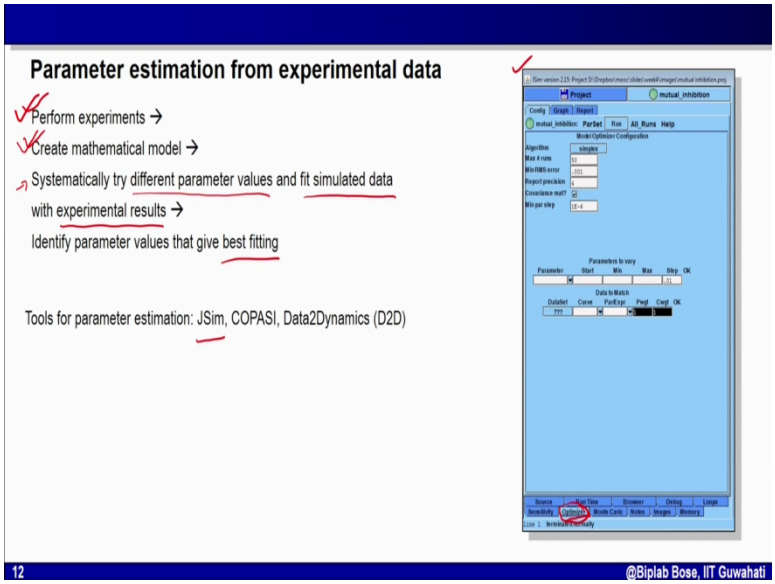
Let us come to the third issue for which we may want to go into the internet and use different web resources that is to get the parameter values. So, to run in model to simulate numerically I require the numerical value of the parameter the constants where do I get them? Obviously the first option is to look into published literature. So, suppose you are creating a mathematical model where you have enzyme which catalyze particular phosphorylation of protein and if that process is Michaelis-Menten kinetics you may look into old literature to find out if somebody has measured Michaelis-Menten constant for that reaction.

If you are lucky you may get that information in literature and you may use those numerical values in your model but most of the time these type of value numerical value for a particular process or a particular kinetic constants are not available. In that case many a time we go back to old models dealing with similar processes and we use some educated guess to use those parameter values which other people has used. For example if you are modeling a particular single transaction pathway where a kinase phosphorylate a particular protein y to phosphorylate it y. You have not got the kinetic rate constant for this phosphorylation reaction but you have seen another model where people have estimated experimentally the kinetic constant for a similar reaction but for a different kinase and subset.

But you have a reason to believe that your kinase and subset will also behave in a same fashion

so, you may reasonable take that numerical value used in that old model for your modeling and very often modular do that. The third option is obviously if you can afford to do is do your experiment yourself and form those experimental data you estimate the value of parameters that you have to use for modeling. I will discuss that in one slide. The other and the forth option where we may go back to the web to collect information is there are certain databases which stores this type of data collected from literature and helps you to find those numerical values easily.

(Refer Slide Time: 19:48)



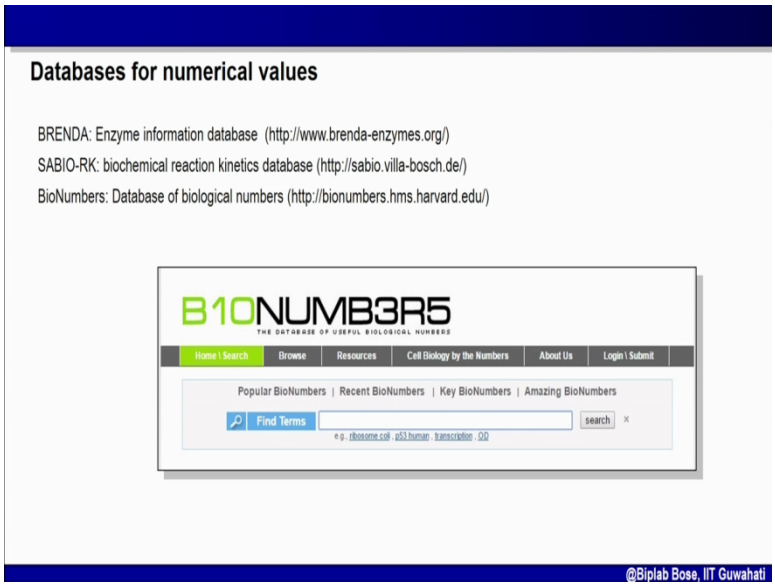
Let us first look into parameter estimation from experimental value. Sometimes people prefer to do their experiments themselves and generate experimental data in different experimental condition. From that data they want to estimate or generate their parameter value for their model. How do they do that? It involves multiple steps. Let us look into steps. First you perform organized experiments then you create a mathematical model using the Ordinary Differential Equation that will have dependent independent variables and parameters which you do not know. Then numerical values we do not know. Then you systematically try to use different parameter values that means for each parameter you systematically take different value and simulate your model and try to fit your simulated data with experimental result. So, what you are doing suppose you ODE has two parameters.

You have model using one single ODE and you have two parameters k1 and k2. You assume that k1 varies from 1 to 10 and k2 varies from 1 to 10. Take different combination of k1 and k2 in this value range and then simulate your model. The simulated model will give you some output some result. Now try to see how different those result are from the experimental result and try to find out the parameter value which gives you best fit. That means which gives you best simulation output which is closest to the experimental value. So, it is a iterative process and depends upon how systematically you search the parameter space to find out the optimum parameter value and it usually takes long period of time to do that and once you have identified the parameter value is be the best fit you use that for other purposes.

There are many readymade tools for this type of parameter estimation but you have to remember no parameter estimation algorithm can give you 100% assurance that you will always get the best parameter estimate you have to try and do trial and error and find out which one is working good for your particular type of problem. JSim itself has a option to do parameter estimation for example as I have shown in this figure here if you notice there is a tab in JSim call optimization if you click on that optimization tab you can actually run a parameter optimization algorithm to find out the best value of parameter which fit, generate simulated data which fits best with your experimental data. If you want to learn it more we have not discussed this part of parameter estimation in our course.

If you want to learn it more you may look into the JSim web page you will find lots of information about this optimization. Then many other readymade plug and play software like COPASI for modeling has this parameter optimization option. Then there is another tool called D2D (Data2Dynamics) which is written in MATLAB and you require MATLAB to run it but is a quite a powerful tool which gives a lots of algorithm different algorithm to do parameter estimation.

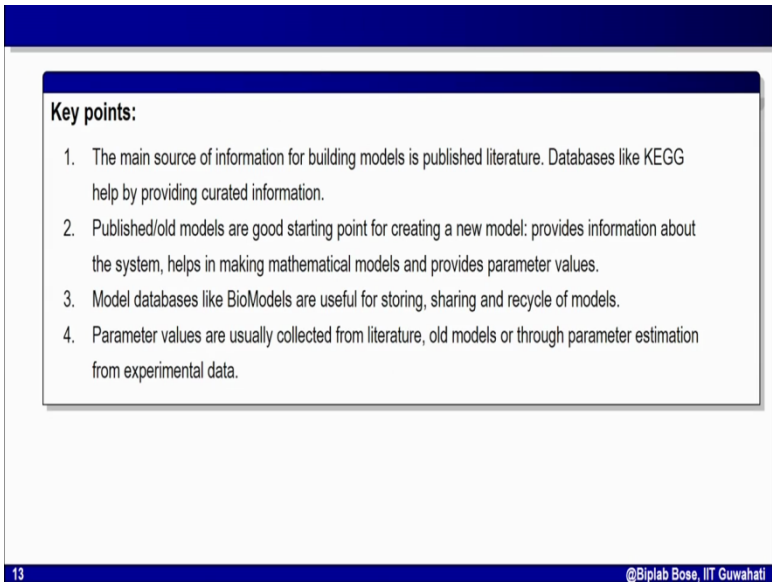
(Refer Slide Time: 23:28)



Now suppose you have done parameter estimation or have collected information from the old models but it will be always good if we have a database where numerical value of different constant kinetic constant or different rate values or other biological values are stored. There are couple of database which are for these purpose one of them is BioNumber which is a extensive database and we use it very frequently. Then you have BRENDA which curate information on enzyme kinetic data. Then you have SABIO-RK which stores reaction kinetic data in its database.

I will discuss a bit about BioNumbers the advantage with BioNumbers is that BioNumber stores lots of different types of data for example you may need to know the size on average size size of a Mammalian cell. You can go to BioNumber to search it. Suppose you need to know what is the number average number of insulin receptor present on a human cell, you can dig into BioNumbers to get that data. So, it is not just the parameter value for example rate constant that you used in your kinetic model but other numerical values which are very useful in our modeling are deposited and curated in BioNumbers. So, modular very frequently go and dig information in BioNumbers.

(Refer Slide Time: 24:59)



Let us jot down the key points the main source of information for building model is always published literature there is no other alternative to it. Database is like key help use by providing information in curated format collected from this published literature. Published of old models are good starting points for creating a new model. These model provide information about the system that mean which is the molecules involves in the pathway. The mechanism of the action of different molecules you get this information from old models. This old model help us to make mathematical equation that means it helps us to write down the mathematical model and it also provide us estimate of different parameter values.

Whereas certain model databases where existing model or published models are stored for example useful one is biomodel. These databases are for storing, sharing and for recycle of published models. Parameter values that are required numerical simulation of any ODE based model you may collect them from literature old model or by parameter estimation through your own experiments and sometimes you may get those values from numerical databases. Databases which stores biological values. That's all for this modules we are at the end of our course thank you for being with us.