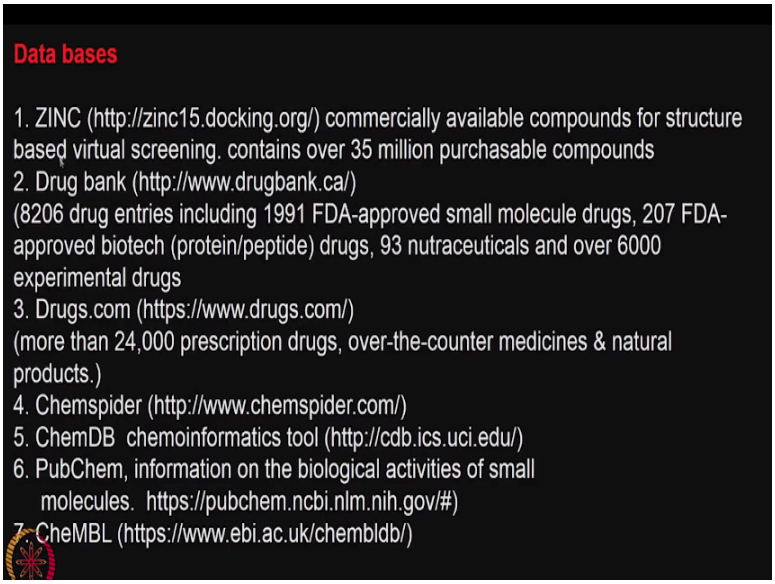
**Computer Aided Drug Design**
**Prof. Mukesh Doble**
**Department of Biotechnology**
**Indian Institute of Technology – Madras**

**Lecture – 04**
**Drug and Data bases**

Hello everyone, welcome to the course on computer aided drug design. In this class, we are going to look at some simple drugs and we are also going to look at some databases which gives a large number of structures. They are very useful for preliminary screening of molecules as I showed you in the first class.

**(Refer Slide Time: 00:34)**

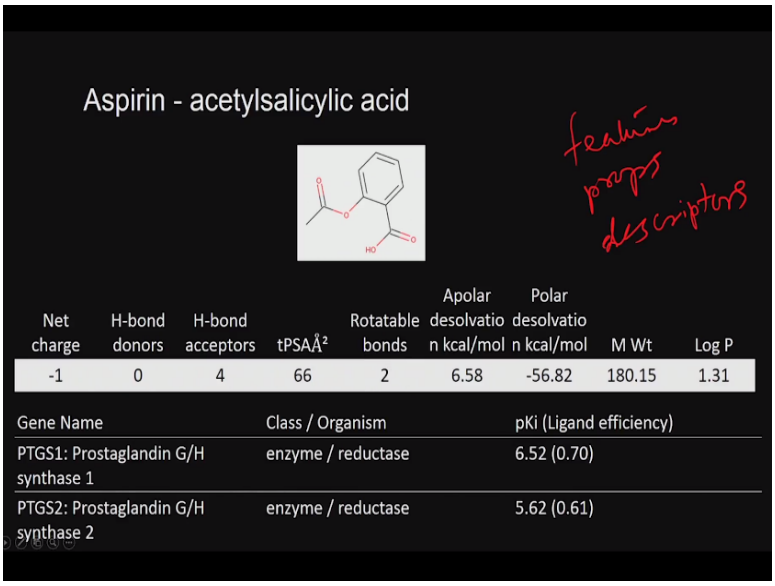



These are some of the databases, which contains structures of large number of molecules and the ZINC database contains 35 million compound structures with some properties. We can buy if you want then we have this drug bank and drugs.com. These databases contain structures of drugs that are currently available in the market over-the-counter drugs, FDA approved, and so on actually, biotech drugs and so on.

So if I want to know what is the drug available in the market for a particular disease the first step is I will go into these drug bank and drugs.com. In addition, we also have the Chemspider, ChemDB, PubChem, CheMBL all these contained structures of large number of molecules searchable and contains properties of all of them different types of properties, molecular weight,

solubility, and so on actually. So we will look at some of them in this class. Okay, the most well known widely used drug is aspirin okay it is called acetylsalicylic acid.

**(Refer Slide Time: 01:49)**




So we have this the acid so we have this salicylic acid. This is the acetyl group and that is present here okay. This is used for pain, fever. So it is very widely used and nowadays it is even being prescribed for reducing the viscosity of the drug I mean viscosity of the blood. So it is called a blood thinner. So aspirin is widely used some of the targets for aspirin is Prostaglandins. Prostaglandin synthase 1.

Okay then Prostaglandin synthase 2. So this is called a cyclooxygenase 1 and cyclooxygenase 2. These enzymes are reductase type of enzymes okay. This gives you the Ligand efficiency and these are some properties of this particular molecule. They are called properties and they are called descriptors, structural features. So we call them different names. features, properties. We also call descriptors.

We will use this quite often descriptor just like if you take a human being the properties could be height of the person, weight of the person, colour of the skin, colour of the hair, colour of the eyes. So many different properties we can have. Similarly, for each chemical structure, we can have large number of properties. We are going to talk about them later, but the molecular weight

is one easy property which you all know about it and it has got 2 rotatable bonds. So we will talk about these properties later.

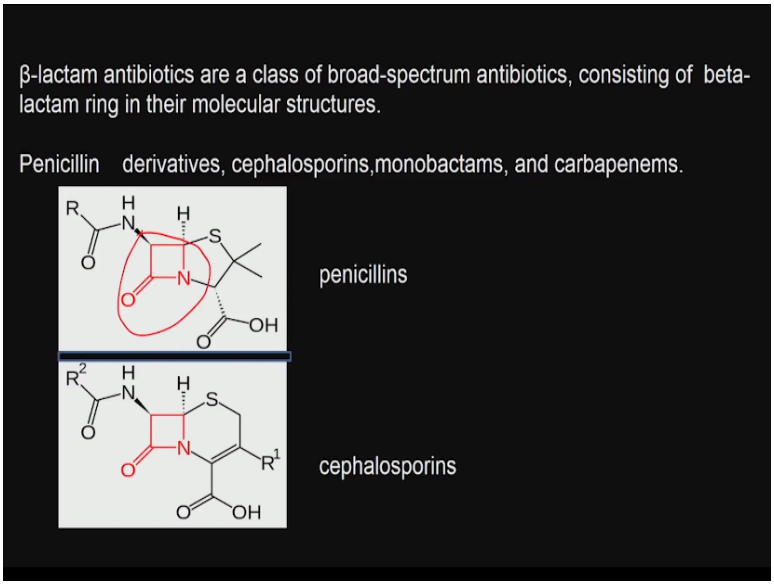
**(Refer Slide Time: 03:29)**


| Net charge | H-bond donors | H-bond acceptors | tPSA Å² | Rotatable bonds | Apolar desolvation kcal/mol | Polar desolvation kcal/mol | M Wt | Log P |
|---|---|---|---|---|---|---|---|---|
| 2 | 4 | 0 | 92 | 0 | 3.29 | -94.07 | 129.16 | -1.03 |

Now let us look at another one metformin. Metformin again everybody nowadays knows about metformin because diabetes has become a serious problem globally as well as quite a lot in India and it is widely used drug for controlling diabetes. So this is the metformin structure. It has got lot of nitrogens okay so it is highly polar so it is highly hydrophilic. Hydrophilic means it is highly water soluble as against hydrophobic which is lipid soluble.

So hydrophilic means it is more water soluble, hydrophobic means is more lipid soluble because of the presence of nitrogens it is highly water soluble similarly if you have lot of OH group, hydroxyl group it will become highly water soluble. So this is metformin used for diabetes again lot of properties of metformin are given and here we will spend more time later as we go long. Look at this.

**(Refer Slide Time: 04:33)**

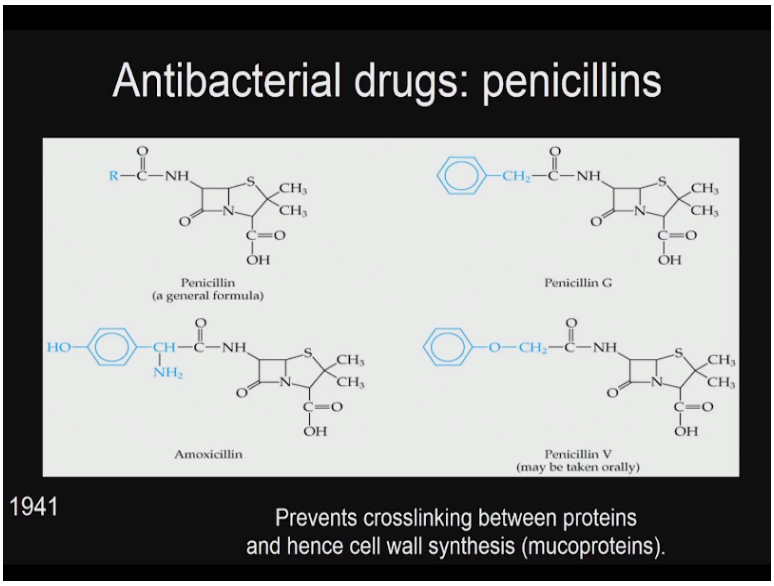



These are antibiotics. They are called beta-lactam antibiotics. These are broad-spectrum antibiotics okay and it is highly available in the market Penicillin one of the oldest antibiotics. It was given during World War 2 for soldiers who had infection. Many soldiers before the advent of Penicillin used to die because of infection and the Penicillin was a biggest boon okay. They are called beta-lactam antibiotics. There are many beta-lactam antibiotics.

So this called a beta-lactam ring okay. This is called a beta-lactam ring. The 4 member cyclic with a nitrogen and a ketonic group and many antibiotics which come under this class of beta-lactam antibiotics will have this particular ring. So we can have different substitutions. Here I have put R. We can have different substitutions here, here, but this particular ring is kept intact that is called the beta-lactam ring, 4-member cyclic with a nitrogen and a ketone.

So this is cephalosporins. Again you see this ring, the beta-lactam ring and then here different types of R groups are put in so we get large number of cephalosporins okay. So they are antibiotics, broad-spectrum antibiotics, and there are many, many types of beta-lactam antibiotics available currently in the market okay. So that is the beta lactam ring okay.

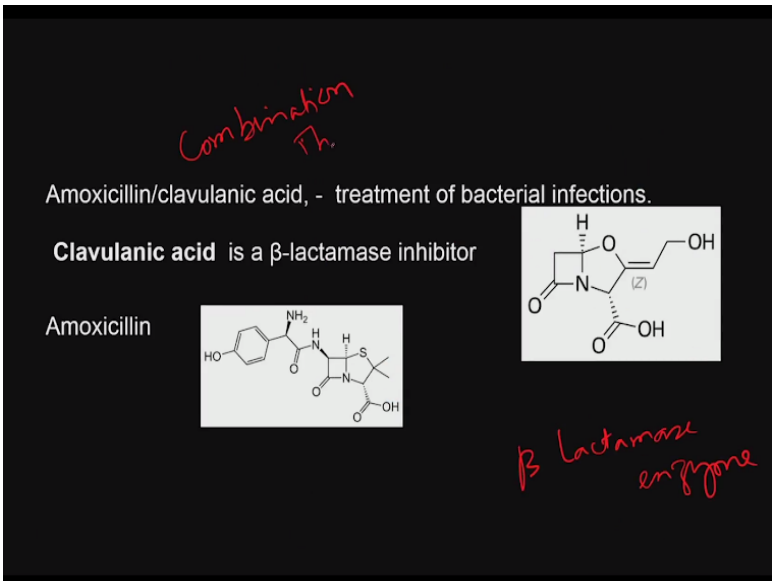
**(Refer Slide Time: 06:11)**




So look at this antibiotics Amoxicillin. So it has got this beta-lactam ring okay. Penicillin V, it is a oral this one Penicillin G. So this beta-lactam ring is intact. So there are lot of modifications to the side change to improve its solubility may be proved its stability to balance the hydrophilic and hydrophobic nature of the drug so many reasons at these extra substitutions to improve the drug likeness property. You are may be to improve okay. So this is the general Penicillin.

As you can see here the beta lactam ring is there, but there are so many different types of Penicillin, Penicillin G, Penicillin V, Amoxicillin, but all of them have this okay so how do they act? They prevent cross linking within proteins and hence cell walls synthesis so the bacteria cell does not grow because they prevent the cross linking of various proteins that are involved in the cell wall development.

**(Refer Slide Time: 07:17)**




So look at this combination of Amoxicillin and Clavulanic acid okay. Again they are used for treatment of bacterial infection okay. What happens is these bacteria became very smart and they started producing something called beta-lactamase enzyme okay that is called beta-lactamase enzyme. So that enzyme produced by the bacteria started breaking this beta-lactam thereby the antibiotic becomes ineffective or it loses its activity.

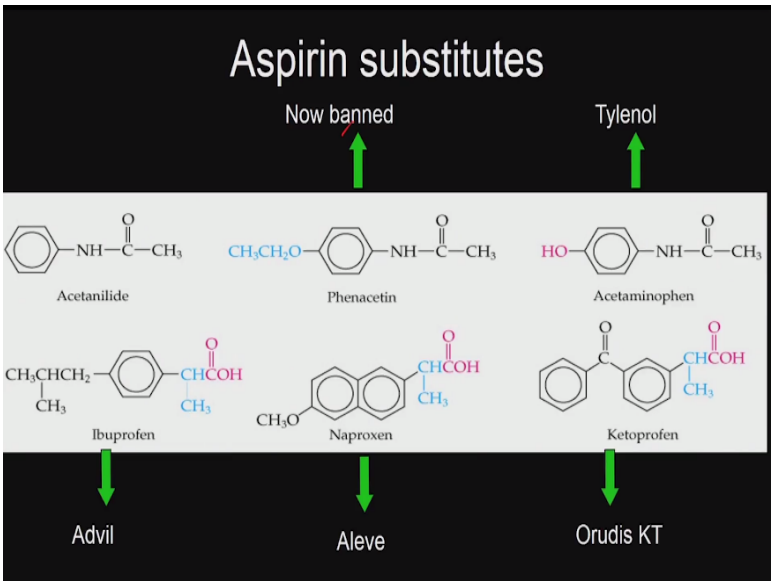
The bacteria started developing this particular enzyme and this enzyme will break this structure. So the antibiotic loses its activity. So what happened was they started introducing combination. One is called a Clavulanic acid/Amoxicillin. So this Clavulanic acid will go and bind to this enzyme beta lactamase and thereby that enzyme is not able to do any job whereas this antibiotic will start working and thereby it will kill the bacteria.

So this is called a combination therapy. Okay this is called a combination therapy where we give 2 drugs, 1 drug acts on 1 target and the added drug acts on the other target okay. So that is called a combination therapy. Because the beta-lactamase enzyme started degrading the beta lactam ring. They started giving 2 drugs in combination so this particular Clavulanic acid will go and bind to the beta-lactam enzyme that is beta-lactamase inhibitor whereas the Amoxicillin will start working on the cell wall thereby the bacteria will get killed okay.

So nowadays these bacteria become resistant. They are called resistant bacteria because they start doing different tricks. One approach is to produce these types of enzymes which will go and kill the, or destroy or degrade the drug that is 1 approach. Other approach is to produce something called efflux pumps that means whatever drug comes inside the microorganism or cell is thrown out that is efflux out okay so that is called efflux pumps.

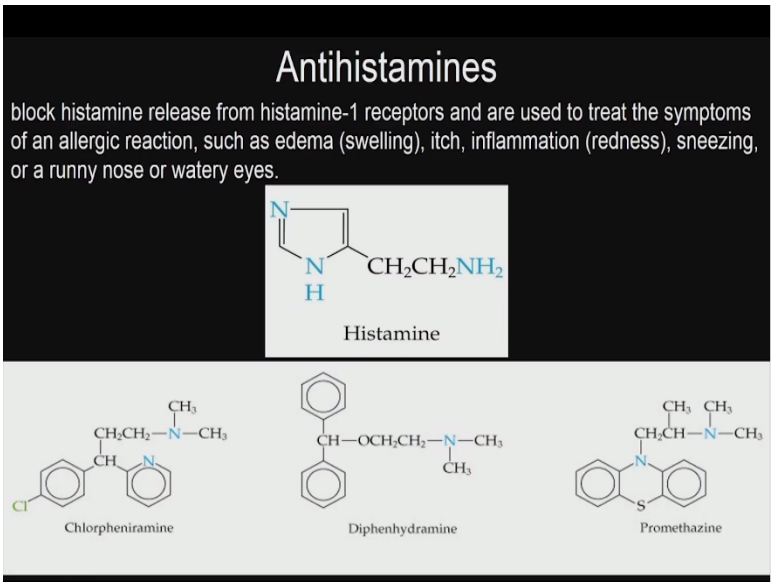
So the bacteria started developing these type of new tricks thereby they become resistant and nowadays you must have been hearing about resistance, bacteria, resistance, tuberculosis bacteria and so on. So these are the tricks they follow and so drugs are given in combination to overcome some of these problems. Aspirin okay we all know aspirin.

**(Refer Slide Time: 10:12)**




Okay we all know aspirin, but there are different types of aspirin just you can see Phenacetin, then Tylenol so all of them have different type of structural features as you can see here ibuprofen is also called Advil. These modifications are done in the structures as I said to improve the drug likeness property or to improve ADME or to make the molecule more stable or reduce toxicity and so on actually. We will look at some of them as we go along.

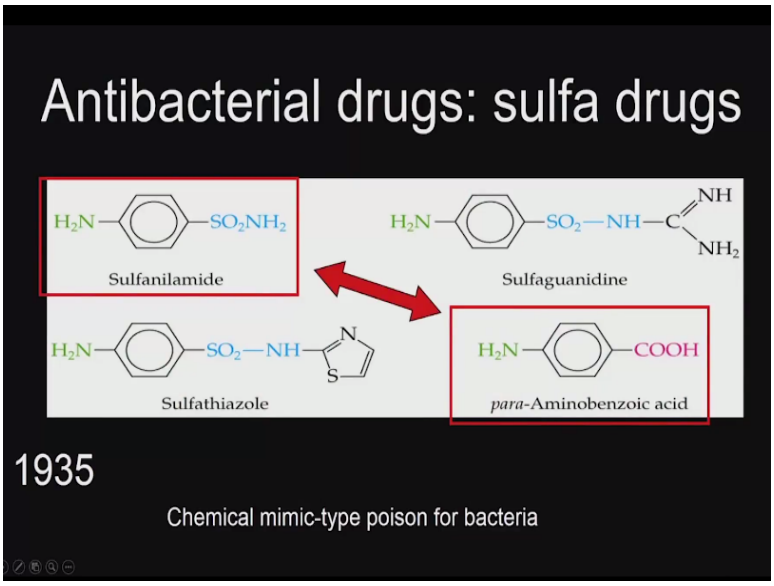
**(Refer Slide Time: 10:43)**




Antihistamines again all of us must have taken antihistamines. These are taken to overcome swelling or redness, inflammation, sneezing, runny nose, watery eyes, so how to they act? They block histamine release from histamine 1 receptor okay. So they block histamine when it gets release. We end up having swelling, itching, inflammation, runny nose, runny water eyes and so on actually. So they are called antihistamines okay they block the histamine release.

This is what is called the histamine the structure of the molecule and that is called histamine. So these are some of the antihistamine drugs which are in market Diphenydramine, Chlorpheniramine. So we will see that they have the nitrogen groups here okay. That is why this ending is amine okay. They have all the nitrogen groups here that is why it ends with amine. These are all antihistamines used for allergy type of situations.

**(Refer Slide Time: 11:48)**

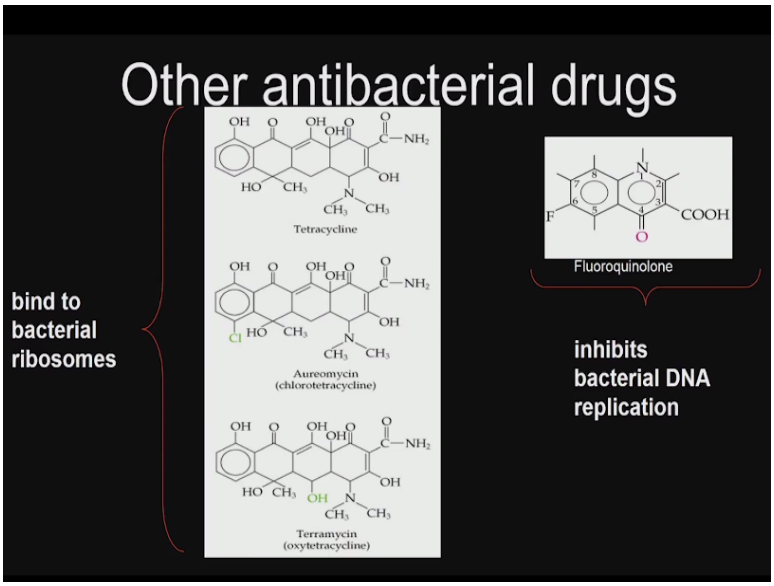



More of antibacterial these sulfa drugs were discovered very very early and they have this sulfur groups that is why they are called sulfa drugs so as you can see here Sulfanilamide so they have the NH2 so there is an amide group here okay that is why it is called Sulfanilamide okay. This is an aniline type of groups so different types of Sulfaguanidine, Sulfathiazole. This is called a thiazole structure.

A 5-member ring with nitrogen and sulfur is called a thiazole structure because thiamine sulfur zole is nitrogen and we have thiazole it means it is a 5-member nitrogen and sulfur is present. So you see different types of modifications to sulfa and these drugs nowadays are not used much because the bacteria have become more resistant to these drugs okay and the whole world have shifted towards different types of antibiotics like I showed you Penicillin. Again Penicillin also has become very meek bacteria become resistant to that.

So the medical fraternity has more into much more powerful antibiotics. So there is always a war between the microorganism and scientist as a scientist discovered new drugs and microorganism somehow get resistance and either through gene modifications or through design of the flux pumps or design of a new enzymes to degrade the drug. So again scientist starts discovering new molecules and so on actually. So these are sulfa drugs which are generally not used nowadays may be for some of the skin infections it is still being used in ointments.

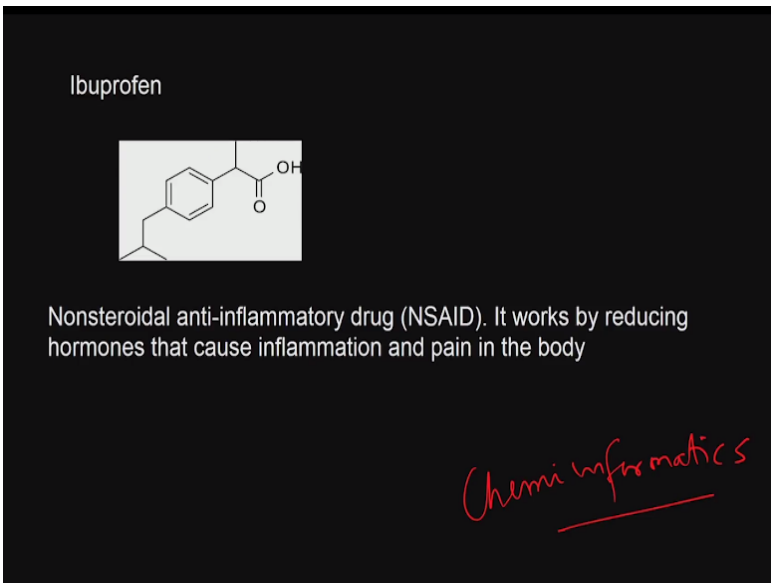
**(Refer Slide Time: 13:45)**




Then other antibacterial drugs we can see Tetracycline you must have heard all these names, Aueromycin okay again Terramycin, Fluoroquinolone. So these drugs they bind to the bacterial ribosomes whereas a Fluoroquinolones bind to the DNA replication. So the DNA does not get replicated. Okay so you can see different type of more of action antibiotics antibacterial drug.

If you look at Penicillin they prevent cell wall synthesis. If you look at these types of drugs Tetracyclines they bind to bacterial ribosome. If you look at Fluoroquinolones they prevent the DNA replication. So Pharma companies look at different possible targets and start designing molecules towards those targets.

**(Refer Slide Time: 14:38)**

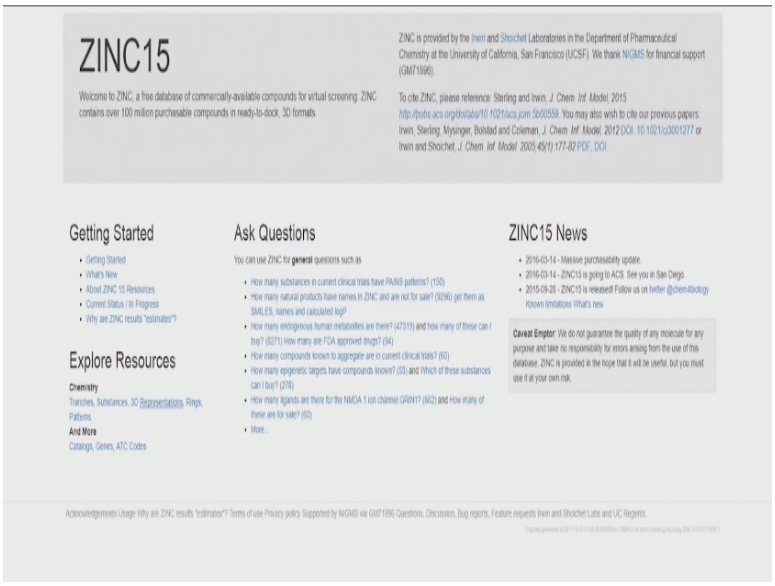


Ibuprofen okay this everybody would have taken because of inflammation and so on. The joints in the knee and so on so these are called nonsteroidal anti-inflammatory drug okay NSAID. Once upon a time steroids were given if there are swelling in knees or joints, but steroids as you know has lot of side effects so drugs were discovered, small molecule drugs were discovered and to overcome this inflammation as anti-inflammatory molecule so they are called nonsteroidal anti-inflammatory drug NSAID.

So how does it do? It reduces the hormones that cause inflammation and pain in the body. They go and bind to some enzymes which produce prostaglandins okay that is how they act. So large number of drugs, which we use very commonly I have shown which is very, very important for us to know the structural features.

We need to know the structural descriptors, which are normally described the molecular properties. So that is called chemiinformatics, which tries to find out the structural features and or structural descriptors of molecular structure. So let us look at some databases like I said before and see how we can make use of them.

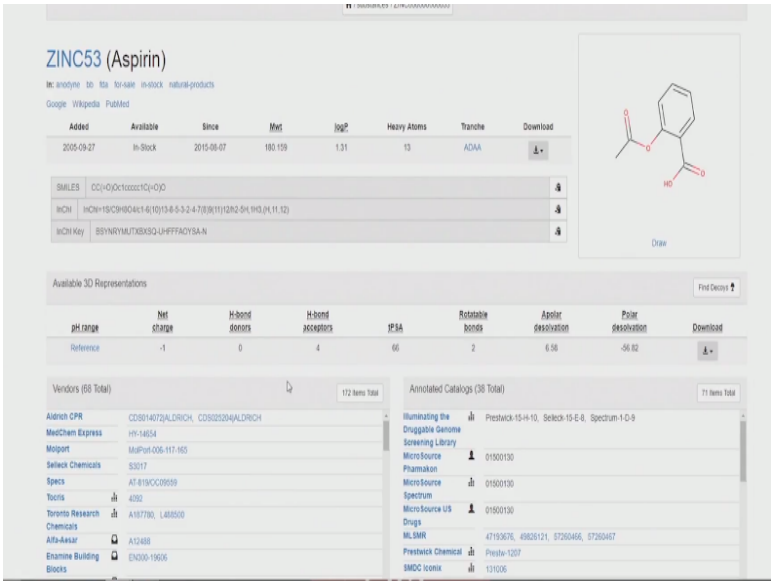
**(Refer Slide Time: 16:10)**




Okay let us first look at this database called Zinc15 database. This database contains structures of a large number of molecules as you can see 100 million. It keeps going up everyday purchasable that means if you think you have found a good possible lead compounds based on (()) (16:35)

you can even buy it okay that is the beauty of it. Now this zinc database how do I go about searching?

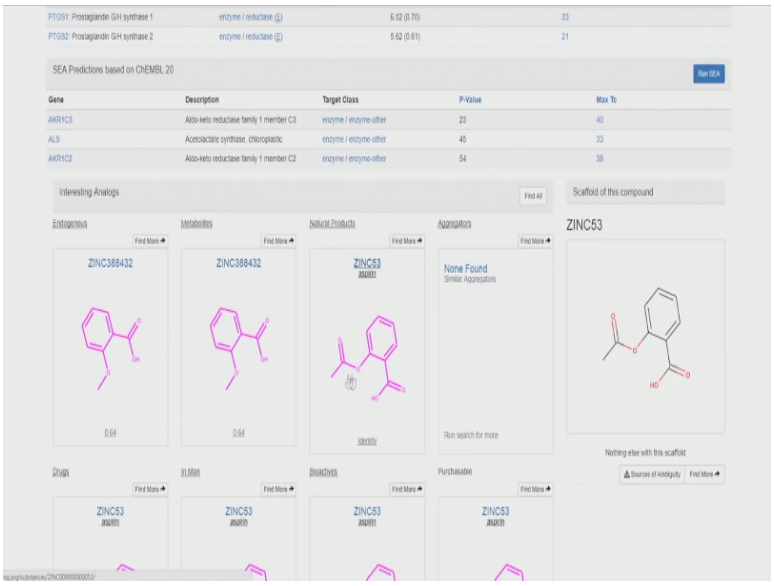
**(Refer Slide Time: 16:43)**




I can do click on this. So search for a substance so I can type say for example as ya, so we find structure. I click on it. So it gives lot of information. This structure as I said aspirin is nothing but acetyl. This is the acetyl group. This is the salicylic acid. So it has got a benzene and an acid here. It is available in stock and so on okay. The net charge on the molecule is - 1. Why does it get - 1 because the OH can easily dissociate so you can get a O - 1. It does not have any hydrogen bond donors okay because it gets the O and H gets okay dissociated as O -.

So hydrogen bond acceptors that means oxygen and nitrogen are called hydrogen bond acceptors. So 1, 2, 3, 4, so it has got 4 hydrogen bond acceptors. TPSA means total polar surface area. PSA means polar surface area okay. It has got 2 rotatable bonds. So 2 rotatable bonds. It has got then okay and so on. So lot of properties it is given and then in addition in tells which enzymes it goes and binds to if you click on these we can find out more details about those enzymes.

**(Refer Slide Time: 18:29)**



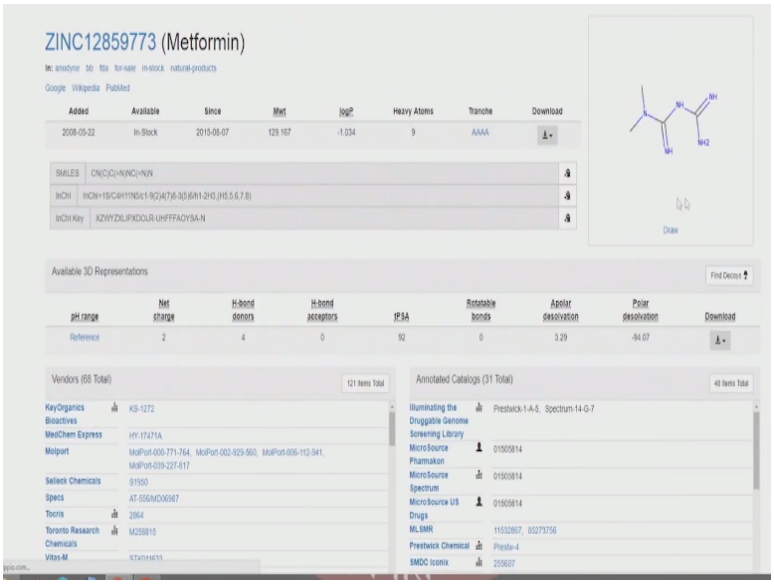

As you can see these are analogs of this okay. Different analogs of this as you can see here different analogs of aspirin and then what has to be done as any clinical trials been done. The details are given phase IV clinical trials antiplatelet therapy.

**(Refer Slide Time: 18:48)**


So as I said aspirin is nowadays being used for thinning of blood. It is also used for preventing platelet aggregation okay so strokes it is also being given now for strokes also and so phase IV clinical trials have been done. So it is in the market for treatment of blood thinning as well as for stroke. It is Warfarin or aspirin is being looked at it this term completed. like that you know what are the clinical trials that have been done with aspirin.

What is the current status 2005, 2006, 2007, 2011 those details are also given on this and so good amount of information is given about this particular molecule okay so if we are interested to buy aspirin we can always get it here. We can write to them and also get this information. Okay let us look at another structure metformin for example.

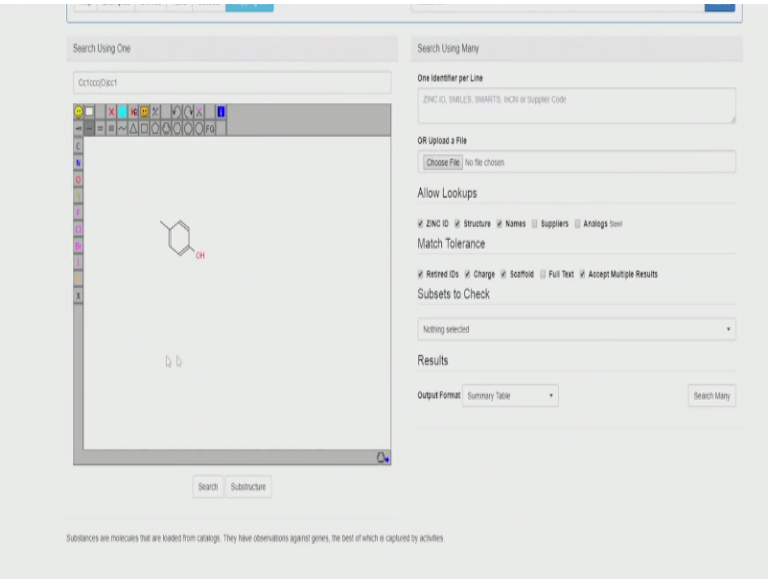
**(Refer Slide Time: 20:13)**




You have metformin here. So if you click on that. Okay so this is metformin. Again lot of properties of metformin and net charge is 2 because see you can see plus charges are there. So it is 2 hydrogen bond donors. It has got 1, 2, 3, 4 so 4 nitrogen groups which are hydrogen bond donors. There is no acceptors okay and lot of information about it is being given, metformin and Rosiglitazone. This is also another anti-diabetes drug okay for treatment.

So some clinical trials, details are being given here okay and then some of the properties of these molecule and then what type of studies are being done in clinical trials. So all these are given. So if I can click get more information on this. Yeah so I can get more information. So this is a very useful database. Okay in addition we can do lot of things here. We can draw structures here for example.

**(Refer Slide Time: 21:31)**




Imagine I am drawing a structure and here I can drug a chemical structure and okay so I am interested to look at molecules in the database which has got this substructure. So I can say substructure search. So I can draw different substructures here and search this database to find out whether there are any main structure with this type of substructure present. So because this is a very common thing it is searching and searching it will take long time.

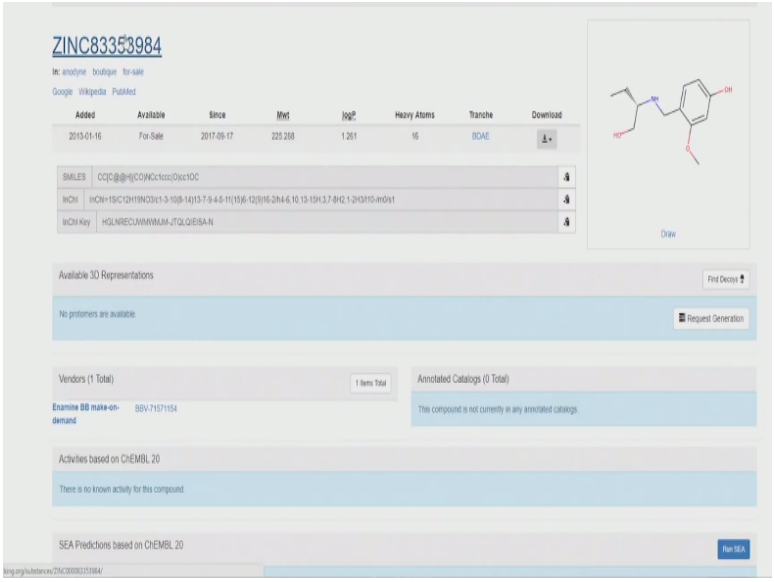
But if I have a big substructure and then ask you to search I can do that. So I can draw different structures using this particular template. You know I can draw benzene rings, 5-membered ring, heterocycles, carbon, nitrogen, oxygen, sulfur, different as you can see chlorine, bromine, like that I can draw. If I know the name of course I can give the name. Look at this.

**(Refer Slide Time: 22:54)**

So it has found so many molecules in the database which has this particular substructure as you can see this benzene ring with a O that is what I wanted right. So it finds you huge number of substructures, huge number of molecules with this particular substructure. I can take it further and do some studies or I can look at much more stringent filter and try to reduce them. As you can see here the large number of molecules have this particular substructure.

The benzene ring, there is an oxygen there is in the paraposition there is substitution groups as you can see more and more molecules are coming up. Imagine I just click on one of these small molecules are there. Imagine I just click on this particular molecule. So it gives you details of this molecule. Okay this is the molecule here. Okay.

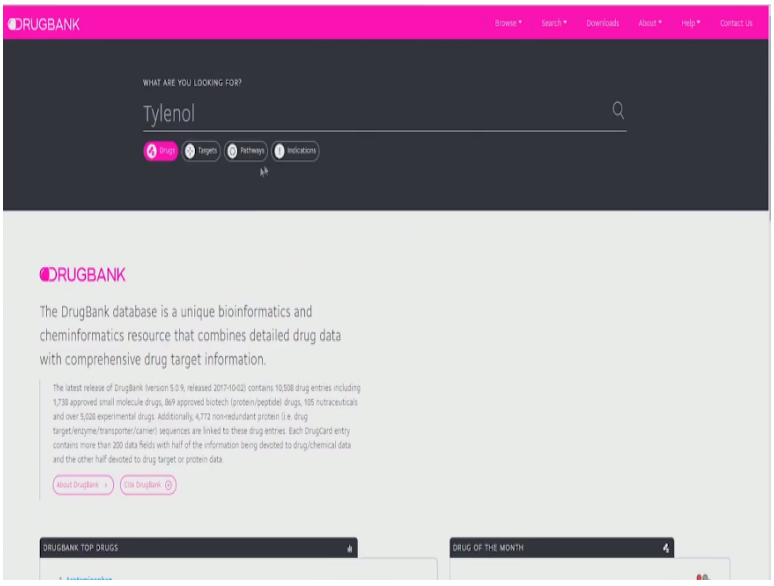
**(Refer Slide Time: 23:55)**




This is the zinc id of that molecule and then some properties you can find here. logP we will talk about all these molecular weight is given here. We can buy this molecule and then if you think it is really okay not much information about this molecule is there. So okay that is called a substructure search.

Okay so the beauty is we can do lot sort of search. Okay in addition there is another search or if you know the zinc ID that means this is called a zinc database so if you know the ID we can do that also. Okay so you can draw the structure that is the beauty of it. Now let us look at drug. Drugdb.

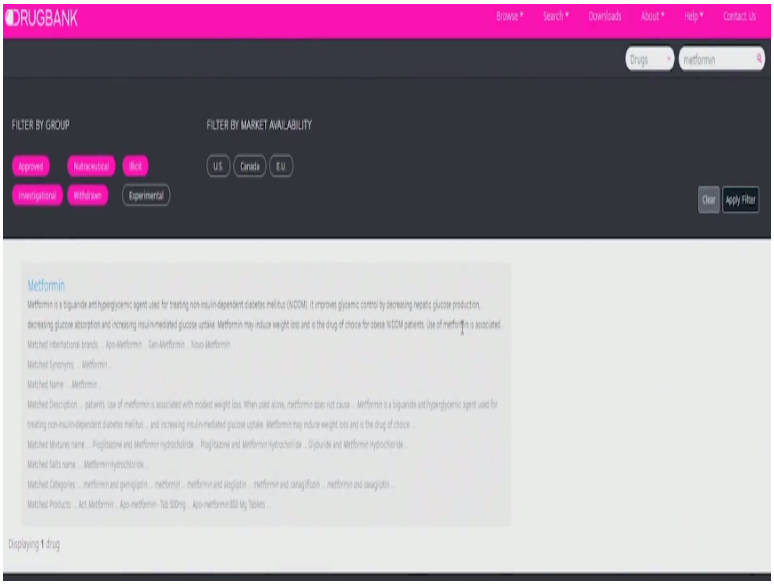
**(Refer Slide Time: 24:54)**



Okay there is this drug bank which contains large number of drugs that are currently in the market. So it contains 10,508 drug entries. That means these are approved by FDA and they are in the market including 1,738 small molecule drugs, so many biotech drugs okay so all these are there. Now we can start searching this also for example okay. This is drug of (()) (25:28) written top drugs Acetominophen, this is used for aches, pains. Acetylsalicyclic acid, metformin I talked about.

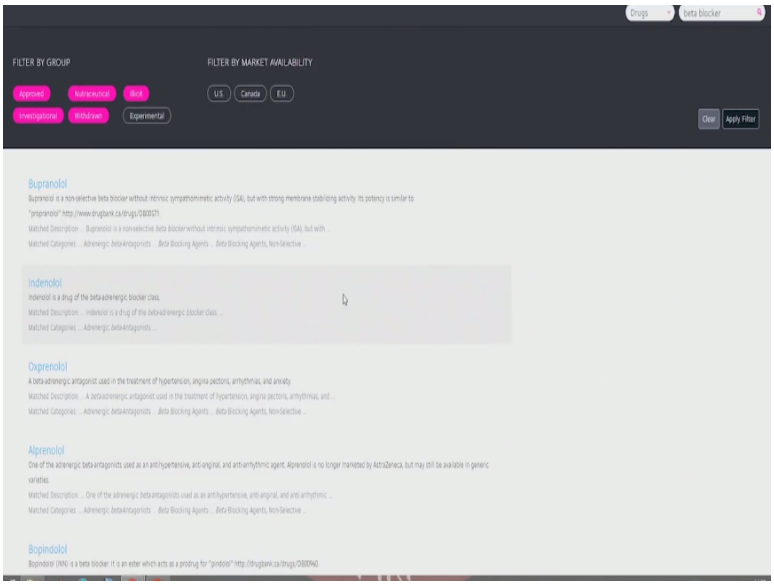
Aspirin I talked about morphine, morphine is a drug used for severe pain. It prevents it reduces the pain which you have okay. So this is a very useful database we can use for our study and you want to identify some new drugs. So we can see again metformin okay so we can search for metformin.

**(Refer Slide Time: 26:06)**



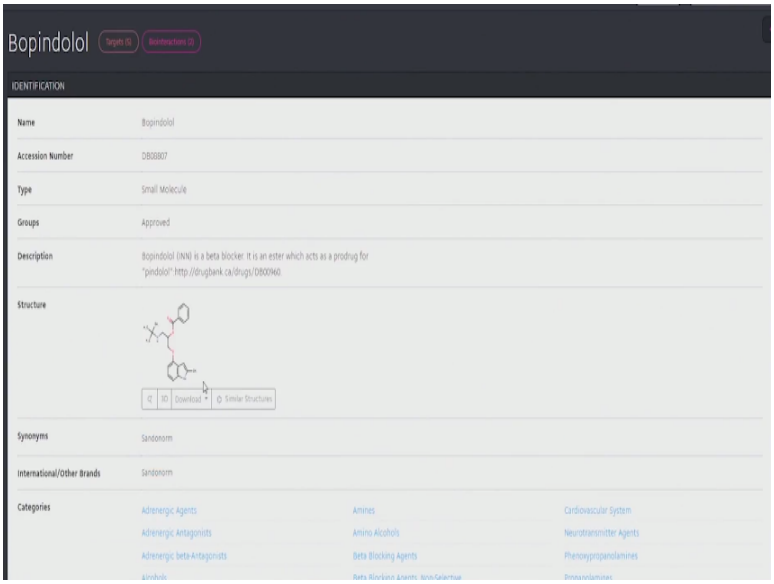

Metformin is a biguanide so okay it gives you all the details about this particular drug. So look at drugs currently for okay so I typed something called beta blocker.

**(Refer Slide Time: 26:36)**




These set of drugs are given for cardiovascular problems especially at high blood pressure there are certain beta adrenergic receptors which are blocked by these type of drug so large number of drugs. They all end up with this particular title Bupranolol, indenolol. So all these drugs.com under this category of beta blocker. This is nonselective beta blocker. So normally this is given for high blood pressure. For example, let me click on 1 of them.

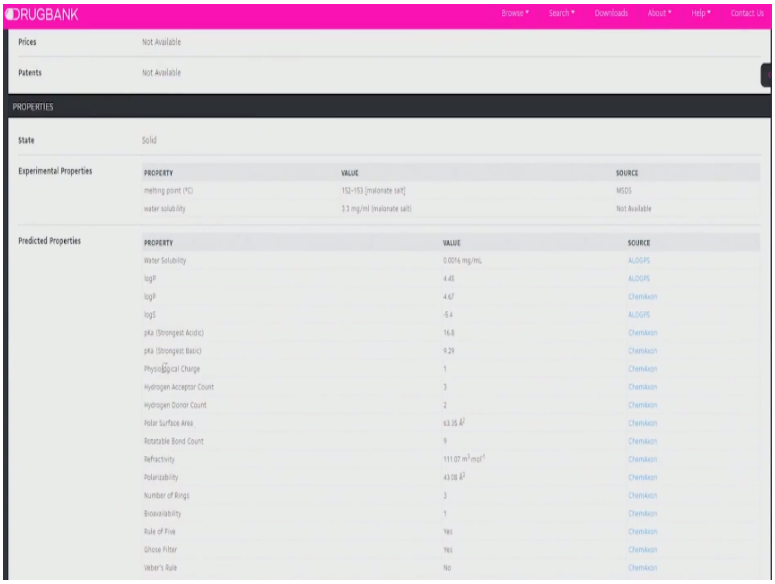
**(Refer Slide Time: 27:14)**




Okay so it gives you the structure of this molecule and then approve it is a beta blocker okay and then international other brand names because companies sell it as brand names okay what is the molecular weight, chemical formula, then pharmacological details okay what is the pharmacodynamics and how it acts how does it bind to and so on actually.

So lot of information of course details about drug interaction how it is interacting with other drugs because nowadays doctors may prescribe 2, 3 drugs so how to they interact with other drugs and so on. Then clinical trial details. Then some experimental properties melting point, water solubility, and then these are predicted property.

**(Refer Slide Time: 28:06)**




These are experimental, these are predicted properties solubility okay PKA that is dissociation acid and so on actually okay. Lot of information is being given number of rings bioavailability as a very important parameter. It tells you when I take a drug orally how much is actually available at the target site and so on actually. So it gives lot of information.

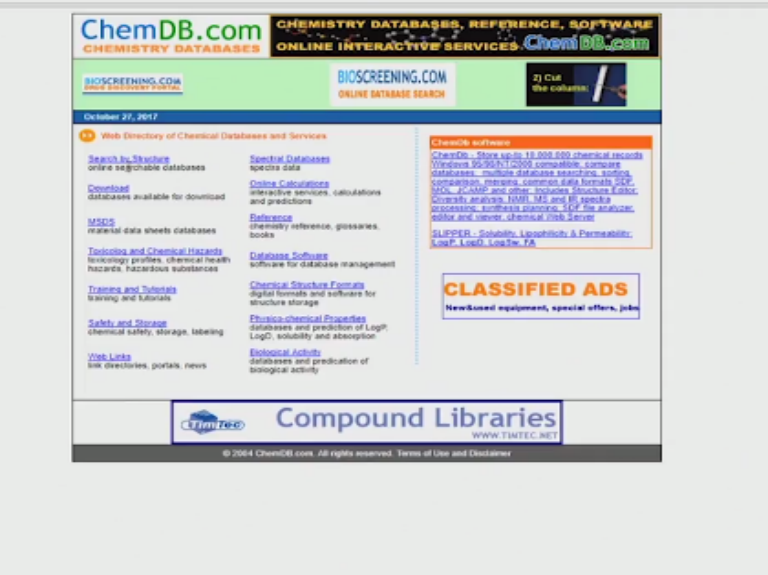
Then as you can see predicted this is that means computationally predicted ADMET properties: Absorption, distribution, metabolism, excretion, T, toxicity. So it gives you information about using software I think they predict lot of properties you can see different types of properties are being predicted and they are all given in this. So if I want to do some analysis, studies this type of information is very, very useful. So drug DB is another useful database. Then we also have as I said another database that is the drugs.com.

**(Refer Slide Time: 29:19)**


Okay So if you look at that also okay it gives you lot of information about different types of drugs that are currently in FDA approved. So we can do a drug index based on condition, class, generic drugs, over-the-counter drugs, international drugs, natural products, veterinary products, drug side effects and so if I am going to start a research on a particular disease first thing I will do is look at drugdb and drugs.com. So I will know all about currently existing molecules.

So I can know their properties. I can know mode of action. I will know how they are, what are the side effects? So that is a very good step to start from. Okay so you need to start from that particular step. Okay so other databases let us look at drugs.com and so you also have this ChemDB, that is another database.

**(Refer Slide Time: 30:29)**

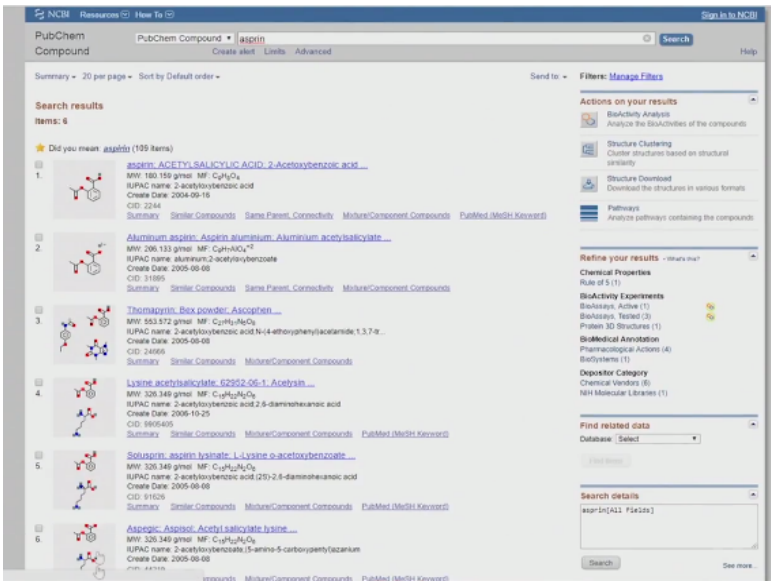



Okay so ChemDB is a chemical database online. Okay so you can see this is chemistry databases reference. We can search by structure. So I give this structure and then I again look at different type of structures and then we can start looking at okay so this is bio-screening. So it gives some we can start looking at structure search, then we also have the PubChem database which gives you information on various types of.

**(Refer Slide Time: 31:08)**

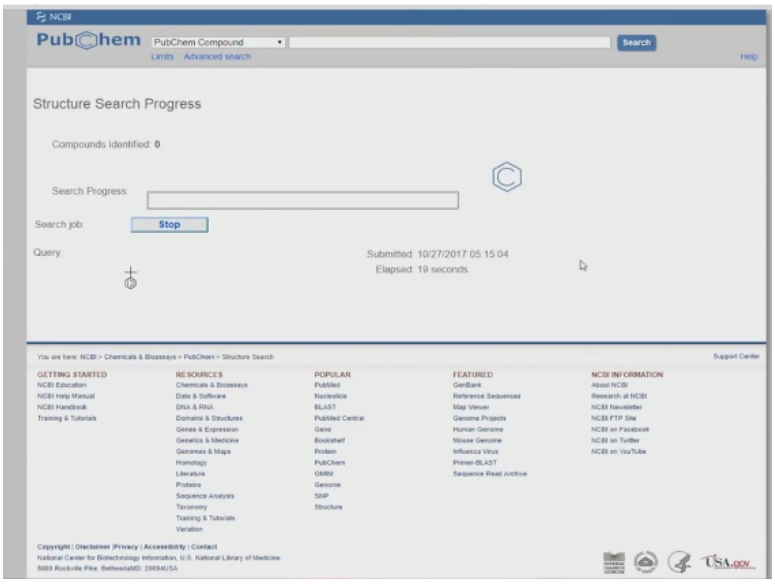

It is called a PubChem database. Again this has got large number of structure, searchable options, you can give properties searchable options, 3 dimensional structures searchable options and so on. So we can see compound aspirin. So different types of aspirin you can search. Look at this, it gives you lot of information.

**(Refer Slide Time: 31:40)**




It gives you the IUPAC name, that sort of things or we can. If I want to know some bioassays how do, I okay so it tells you some assay for aspirin various types of assays for aspirin either to detect aspirin or look at assays for its interaction with different targets. So that is sort of biochemistry is related information we can do. Also this also has a structure search options. So we can draw the structure. So this substructure, so we can draw structure here as you can see here. Lot of things can be done.

**(Refer Slide Time: 32:40)**




Yeah I can do lot of things with this okay so we can look at this and then we can say (()) (33:02) search so it is waiting so it is doing something. So it will look at substructure search for

molecules with that particular substructure. So this is called a PubChem database okay you people can start playing with all these types of different databases available in the web. Okay all these are very useful databases because it gives you structures and lot of properties which you can make use of for doing different types of calculations.

So because we will not have enough time to go through each one of the database in detail I would suggest that you people go through all these databases. We can look at it based on the name of the molecule or the drug. We can draw structures and also we can do a substructure search and we can do based on properties. We can do search based on disease conditions okay so lot of studies, data collection, and data analysis can be done with these type of databases and we will be using some of this databases as we go along. Thank you very much for your time.