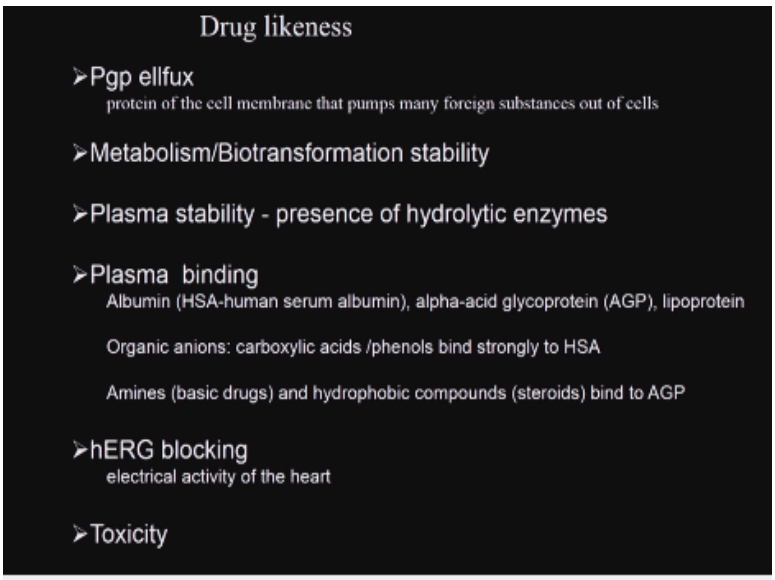
Computer Aided Drug Design
Prof. Mukesh Doble
Department of Biotechnology
Indian Institute of Technology - Madras

Lecture – 14
Drug Likeness

Hello everyone, welcome to the course on computer aided drug design, we will continue on the topic of drug likeness, so we saw many factors that contribute towards the drug likeness property.

(Refer Slide Time: 00:26)



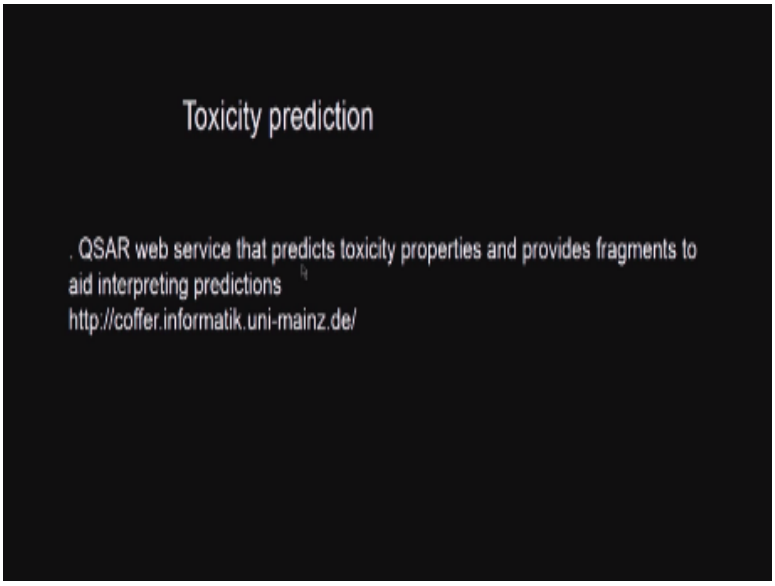
The first one was Pgp efflux, there these are proteins; P glucoprotein, proteins in the cell membrane that pumps many foreign substances out of the cells, so they act like an efflux pump, so the drug also gets efflexed, then you have to consider the metabolism and biotransformation that are taking place inside in the liver, enzymes like oxidoreductases, hydroxylation, so it affects the drug, drug may become more active or less active, the metabolites could be toxic.

So, we need to understand that, plasma stability; here we have a lot of hydrolytic enzymes in the plasma region again, they may be modifying the structure of the drug and making it inactive, plasma binding; there are a lot of proteins in the plasma like human serum albumin, alpha acid glycoprotein, lipoprotein, so organic anions like carboxylic acids, phenols binds strongly to human serum albumin.

Whereas, amines and hydrophobic compounds bind to AGP steroids, they bind to AGP, so the availability of the drug may go down, then there was another concept called hERG blocking; these are proteins which are involved in potassium signalling and again which is in turn combined with the electric activity at the heart and has the heartbeat, so any disturbance drugs blocking this particular protein will affect the heartbeat, heart rhythm okay.

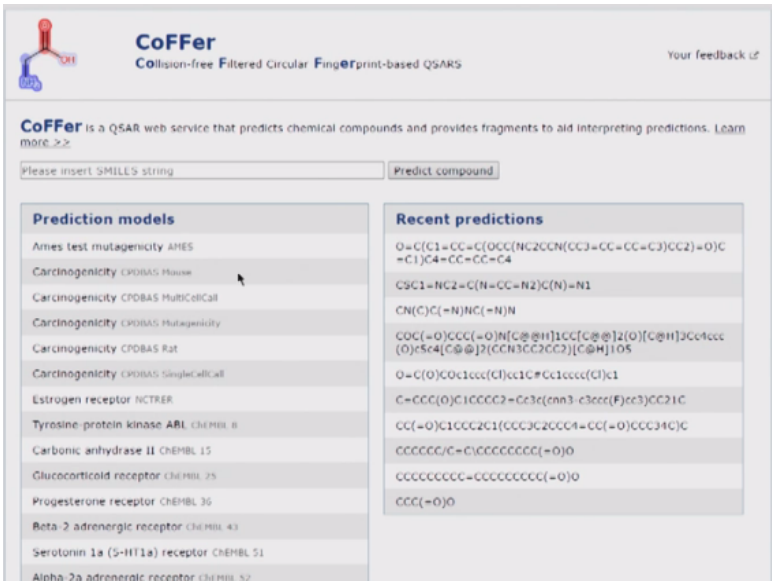
And hence you need to know whether the drugs have any effect on the cardiovascular activity, then toxicity; toxicity is related to whether the drug has short term, long term toxicity, genotoxicity, cytotoxicity and so on, so we need to understand those aspects as well. So, all these contribute to the drug likeness property.

(Refer Slide Time: 02:38)



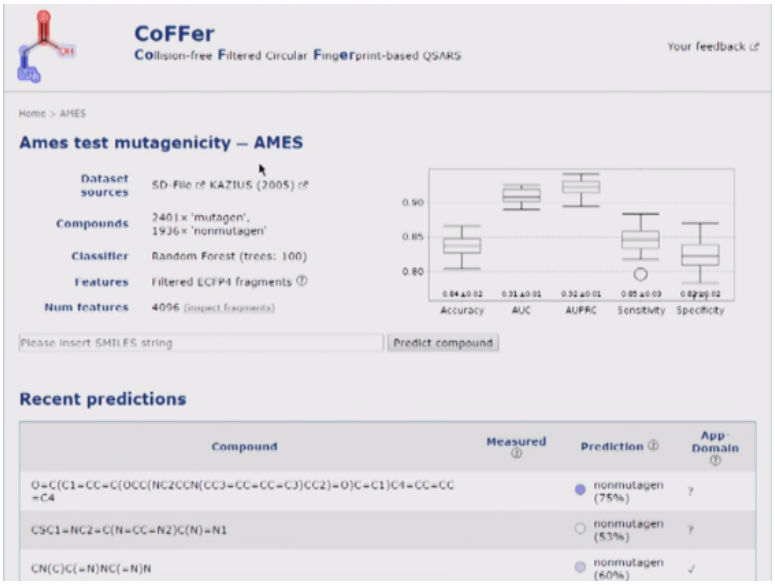
Like I mentioned there is a software which tries to predict toxicity properties, it provides fragments to aid interpreting prediction, this is the particular link for that I will just show you that.

(Refer Slide Time: 02:55)



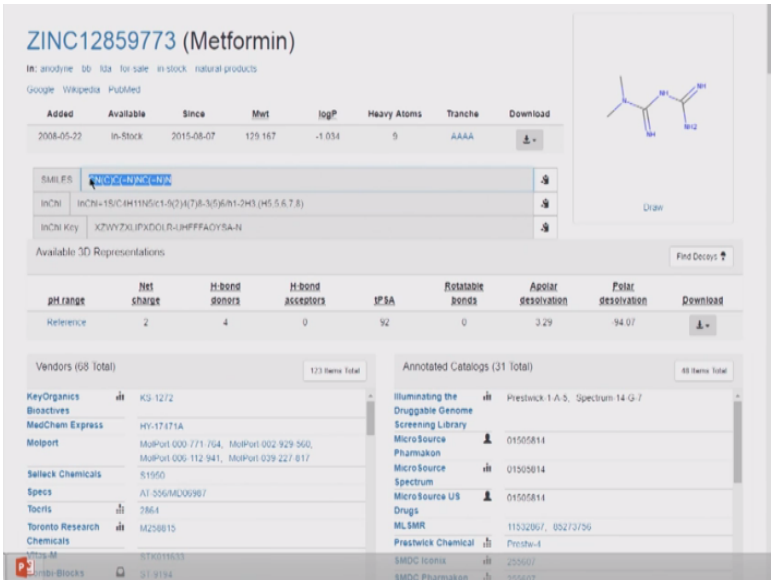
And so we have the; it is called coffer, so it can predict lot of toxicities, as you can see, Ames test mutagenicity, carcinogenicity, then estrogen, tyrosine protein kinase inhibition, cytochrome P 450, so lot of these.

(Refer Slide Time: 03:16)



For example, let me see Ames test mutagenicity okay, so you want to know whether your compound which you are trying to is mutagenic.

(Refer Slide Time: 03:25)

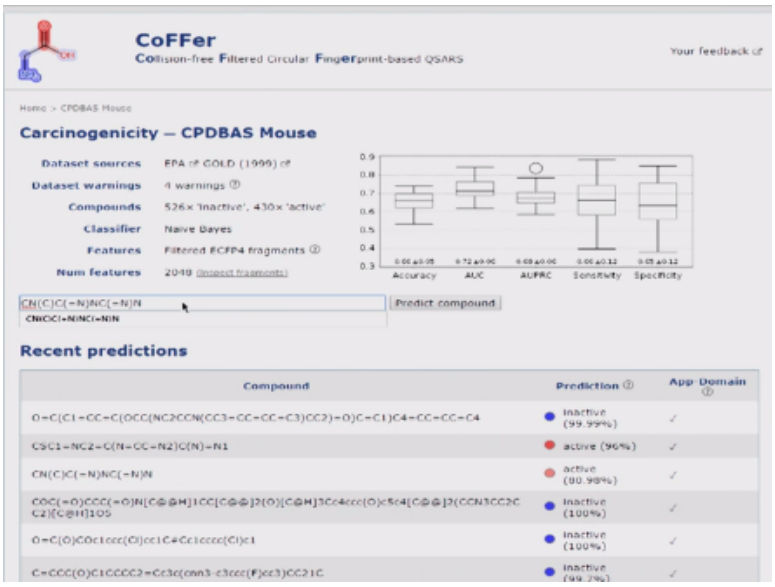


I went to zinc database, I picked up say metformin and then I copy the smiles notation and then I can put that here okay then I can say predict.

(Refer Slide Time: 03:43)

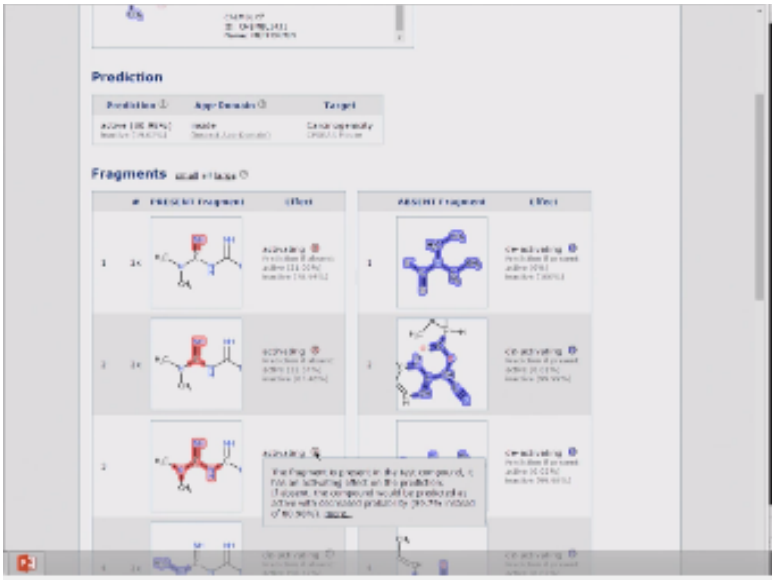
So, it gives you mutagen, non-mutagen okay, so it tells that some of these fragments are there in; present in the test compound okay they may be activating okay, they may be contributing towards mutagenicity as you can see some of these; some of these groups here. When you have the red colour here, it means that actually and it also gives a little bit information about the compound here information, so you can spend more time.

(Refer Slide Time: 04:26)



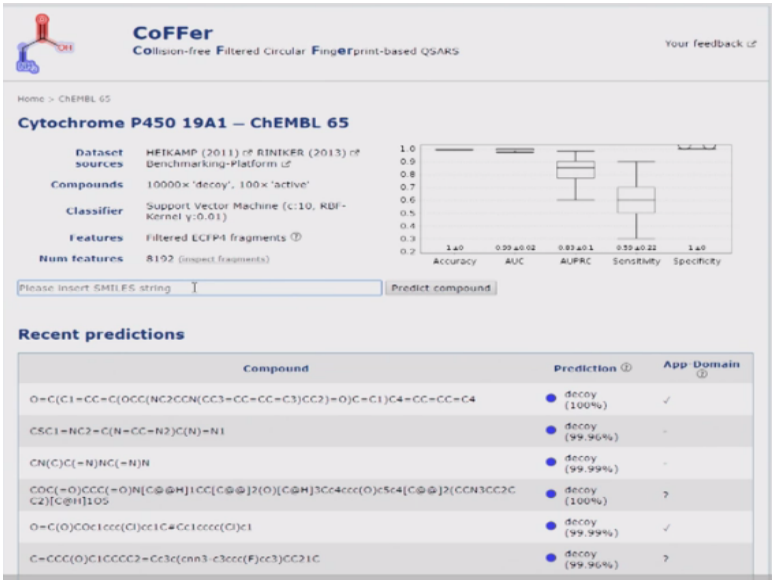
We can look at a lot of different carcinogenicity in mouse model okay.

(Refer Slide Time: 04:31)



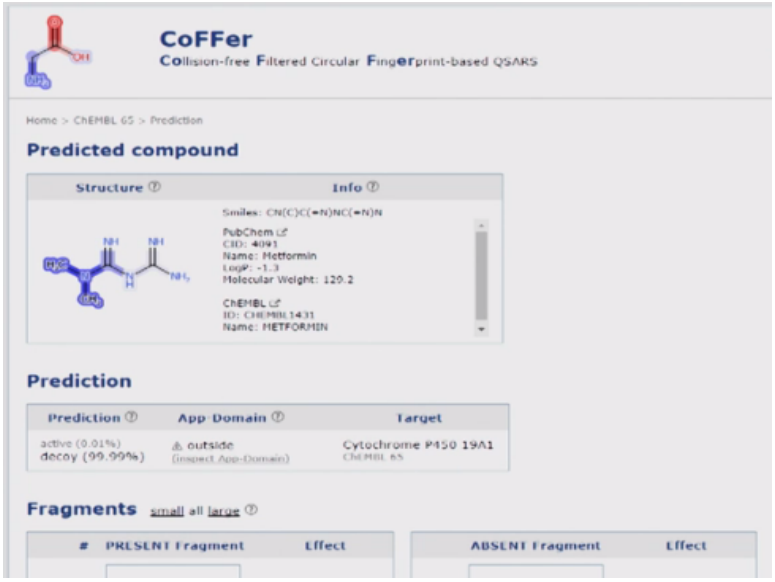
Again, we can put in the okay, it gives you a prediction, so you can see that a lot of these fragments are also part of the carcinogenic activity, it has an activating effect okay, okay, it has an activating effect, so lot of information is available, so they may be contributing to the carcinogenicity in some mouse model. So, we can look at a large number of predictions using this fingerprint based method okay.

(Refer Slide Time: 05:10)



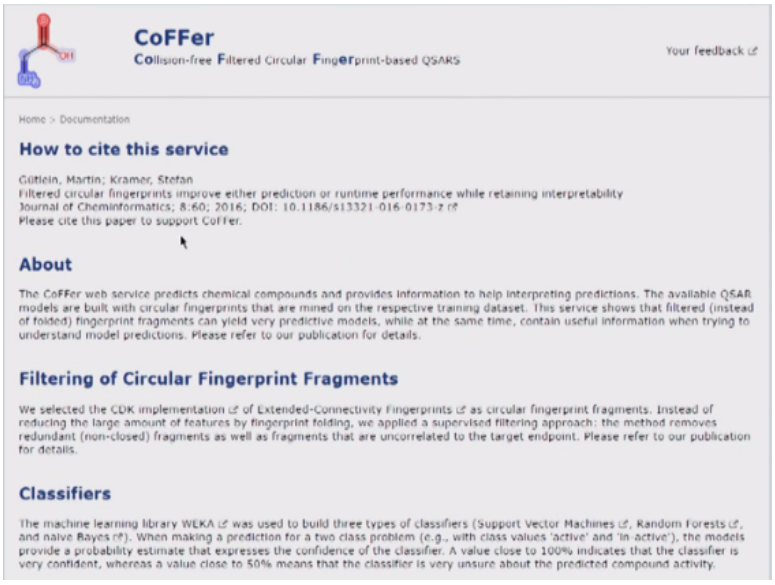
We can even look at say cytochrome P 450 as I had showed before cytochrome P 450 is involved in okay many of these activities.

(Refer Slide Time: 05:15)

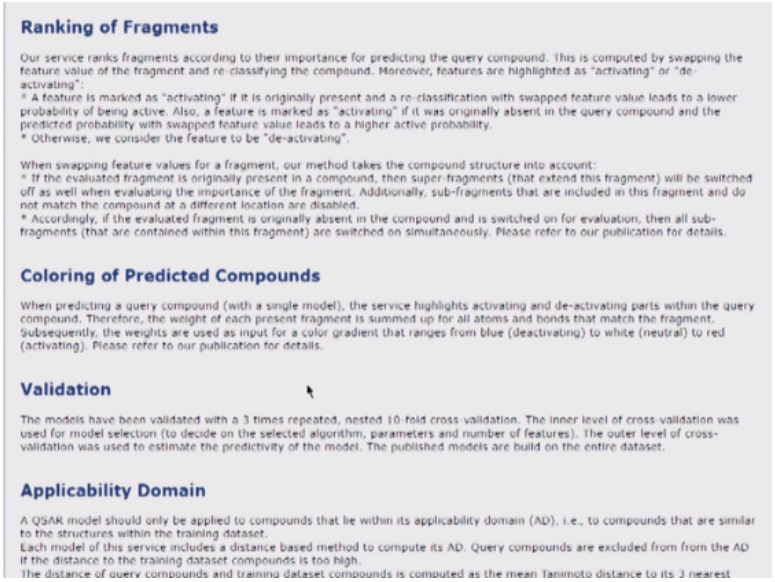


So, we can predict whether these compounds are part of the cytochrome P 450, activating so this software can give you a lot of information that predicts chemical compounds and provides fragments and to aid interpreting prediction okay.

(Refer Slide Time: 05:35)

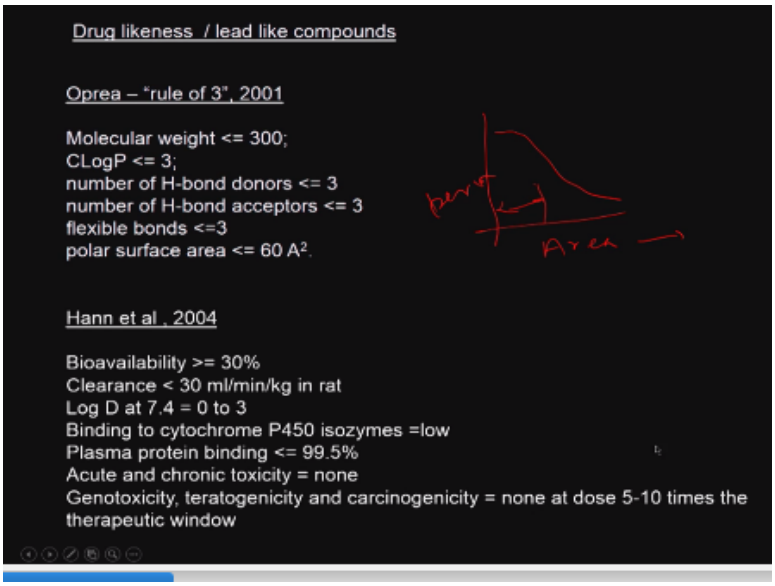


(Refer Slide Time: 05:36)



So, it not only looks at the whole compound it looks at the fragments that that could form from that particular compound and then tries to predict the properties of those fragments okay. So, this is a useful tool and there are many toxicity predicting tools available online free of charge, so you can test; check out your compound also okay.

(Refer Slide Time: 06:00)



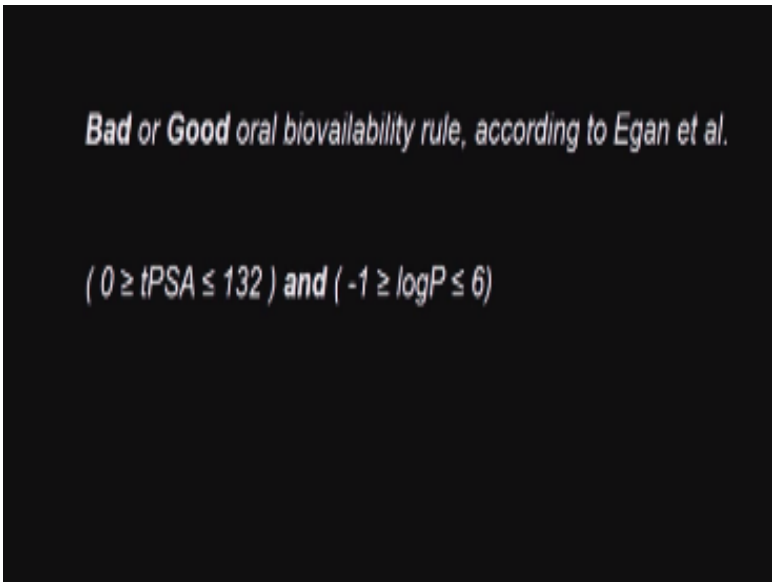
So, drug likeness, lead like compounds there are many rules just like we looked at Lipinski's rule, Muegge rule and things like that this is called Oprea rule of 3 at 2001 on molecular weight should be < 300 as you can see the Lipinski's rule, we said < 500, log P < 3, again log P also has become very stringent, number of hydrogen bond donors < 3, number of hydrogen bond acceptor 3, so they seem to have all 3 coming, flexible bonds 3.

They bring in another parameter called flexible bond, polar surface area < 60 angstrom, okay so according to them, the molecule should have this type of behaviour. Why this polar surface area; if you remember the graph, I showed you once, as the polar surface area increases the permeability through the GI goes down, so one would like to operate in this region, right that is why polar surface area is left low.

But this rule is more stringent then as you can see Lipinski's rule, there is another rule that is called Hann et al rule; bioavailability > 30, clearance < 30 ml per minute per kg in rat, log D at 7.4 should be this binding to cytochrome should be low, plasma protein binding should be < 99.5, acute and chronic toxicity; none, genotoxicity, carcinogenicity, teratogenicity; none at dose 5 to 10 times the therapeutic window, so this is very important.

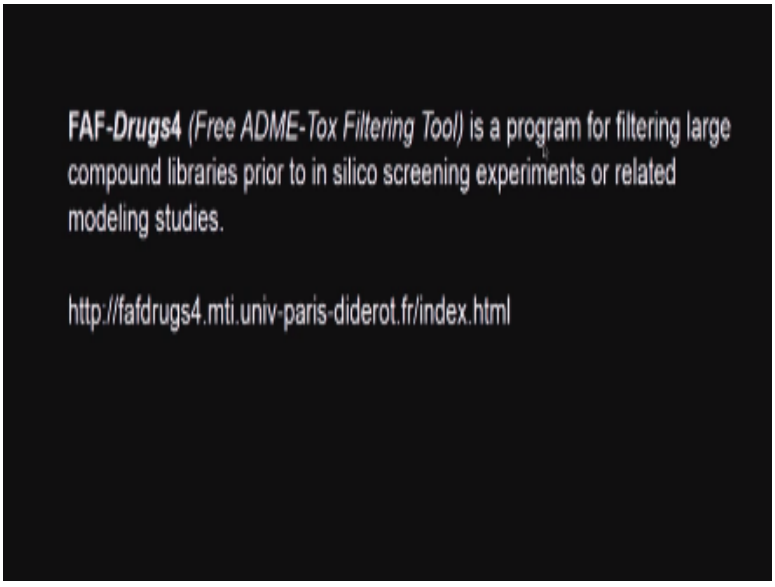
Generally, we always operate 1/10th of the toxic limits generally, okay so that is the Hann et al rule for a drug likeness property; here you have the Oprea rule of 3 for drug likeness.

(Refer Slide Time: 07:57)



There is another rule called Egan et al, they say bad or good oral bioavailability rule okay, so they say total polar surface area should be < 132 okay, so 132 may be falling here okay, little bit more whereas, Oprea rule is very stringent < 60, log P between -1 to < 6, so you see this is another rule which tells you what should be a good descriptor or parameters for a molecule to have good oral bioavailability and drug likeness property okay.

(Refer Slide Time: 08:39)

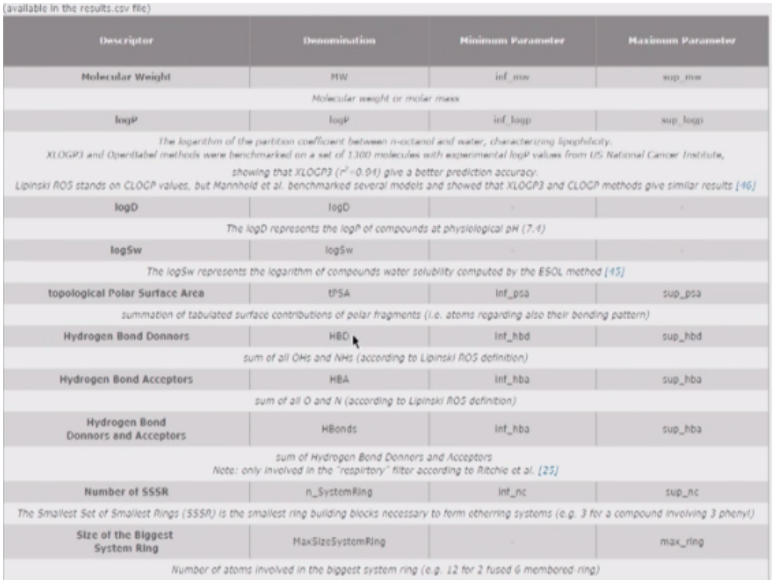


So, there is an another freeware, it is called ADME Tox filtering tool, FAF drugs 4, this is a program for filtering large compound libraries prior to in silico screening, so we can screen large number of compounds, this is the link for that. I will just show you that I will show you that link also for you to look at okay.

(Refer Slide Time: 09:06)

This is the yeah; FAF drugs 4, so it gives you a lot of calculations possibilities, we can look at ADME Tox.

(Refer Slide Time: 09:17)

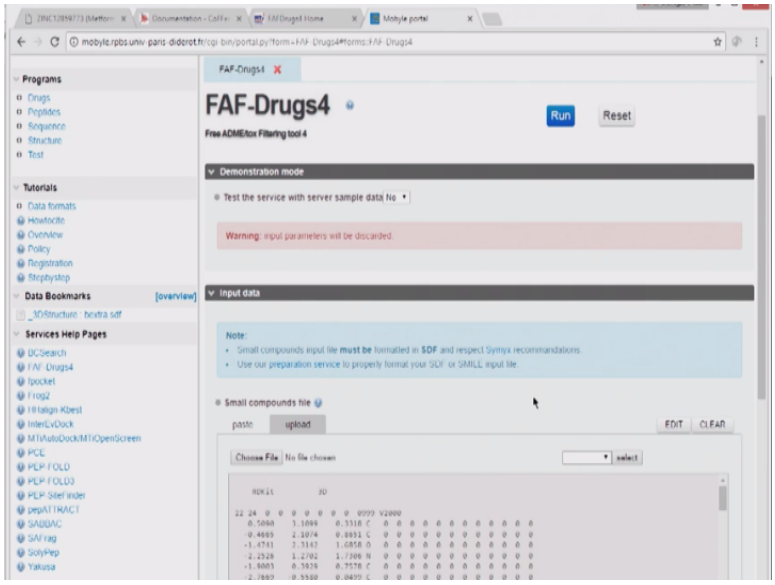


(available in the results.csv file)

Descriptor	Denomination	Minimum Parameter	Maximum Parameter
Molecular Weight	MW	inf_mw	sup_mw
Molecular weight or molar mass			
logP	logP	inf_logp	sup_logp
The logarithm of the partition coefficient between n-octanol and water, characterizing lipophilicity. XLOGP3 and OpenBabel methods were benchmarked on a set of 1,300 molecules with experimental logP values from US National Cancer Institute, showing that XLOGP3 (r^2=0.94) give a better prediction accuracy. Lipinski ROS stands on CLOGP values, but Mannhold et al. benchmarked several models and showed that XLOGP3 and CLOGP methods give similar results [46]			
logD	logD	-	-
The logD represents the logP of compounds at physiological pH (7.4)			
logSw	logSw	-	-
The logSw represents the logarithm of compounds water solubility computed by the ESOL method [45]			
topological Polar Surface Area	tPSA	inf_psa	sup_psa
summation of tabulated surface contributions of polar fragments (i.e. atoms regarding also their bonding pattern)			
Hydrogen Bond Donnors	HBD	inf_hbd	sup_hbd
sum of all OHs and NHs (according to Lipinski ROS definition)			
Hydrogen Bond Acceptors	HBA	inf_hba	sup_hba
sum of all O and N (according to Lipinski ROS definition)			
Hydrogen Bond Donnors and Acceptors	HBonds	inf_hba	sup_hba
sum of Hydrogen Bond Donnors and Acceptors Note: only involved in the "respitory" filter according to Ritchie et al. [25]			
Number of SSSR	n_SystemRing	inf_nc	sup_nc
The Smallest Set of Smallest Rings (SSSR) is the smallest ring building blocks necessary to form otherring systems (e.g. 3 for a compound involving 3 phenyl)			
Size of the Biggest System Ring	MaxSizeSystemRing	-	max_ring
Number of atoms involved in the biggest system ring (e.g. 12 for 2 fused 6 membered ring)			

We can use a PhysChem descriptors as you can see molecular weight; log P, log D, log solubility below surface area, hydrogen bond donors, hydrogen bond acceptors, total hydrogen bond, number of okay rings, small set of smaller rings, biggest system ring, number of rotatable bonds, carbon atoms, hetero atoms. Hetero means hydrogen not included number of heavy atoms, hetero to carbon atom ratio, Egan's rule, oral physical chemistry.

(Refer Slide Time: 10:03)



And so we can do a lot of descriptor calculations okay we can even do filtering based on the descriptors okay, so let us look at a system; FAF; so, when you say run, it gives you like this, so imagine I am uploading a structure, I can upload, choose a file, it needs a SDF file, so I am ruling SDF file, this is Bextra, this is a selective cyclooxygenase 2 inhibitor okay, so I have loaded this file and then I can do a lot of calculations here as you can go down.

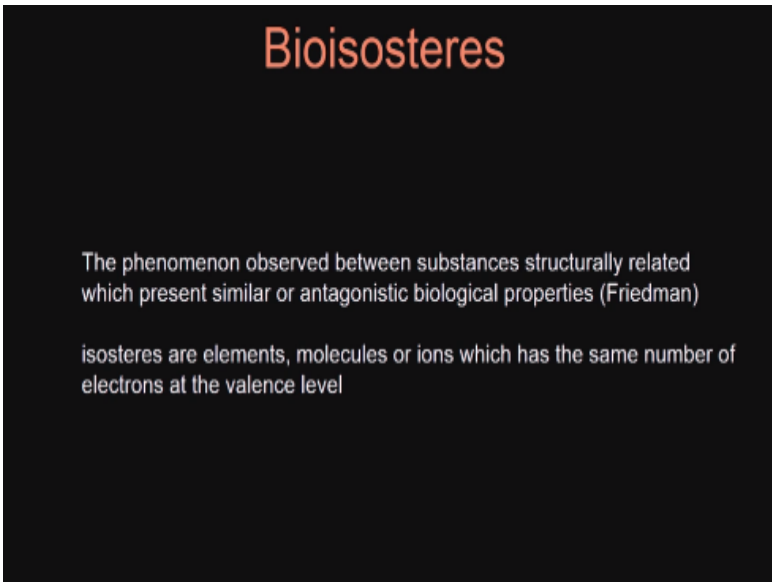
We can say log P computation then we can look at filtering possibilities here okay, no, no so, we can see just run, so it calculates, so it does lot of calculations and then it gives you a lot of properties, so if you have many molecules put in okay, so it does a lot of these calculations, molecular weight, log P, log D, solubility, rotatable bond, rigid bonds, flexibility, donors, acceptors, okay.

So, lot of information it gives about the compound as you can see, hydrogen bonds, hetero atoms, total charge, heavy atoms, carbon atom then it gives you oral bioavailability is good, Egan rule is also good okay, so if I have lot of 30 molecules or more, I can even prepare histograms showing a particular property, the histogram of a particular property, I can put a histogram of log P or I can put a histogram based on log S and so on actually okay.

So, this is a very good software for one to study the drug likeness property, try to collect information on various descriptors and it is a freeware okay, so we looked at quite a lot about the drug likeness, the oral bioavailability and what are the issues related when the drug travels from the oral cavity right up to the target site, how the pH changes or the solubility changes,

how it may get bio transformed because of presence of enzyme or it may be getting absorbed in plasma region.

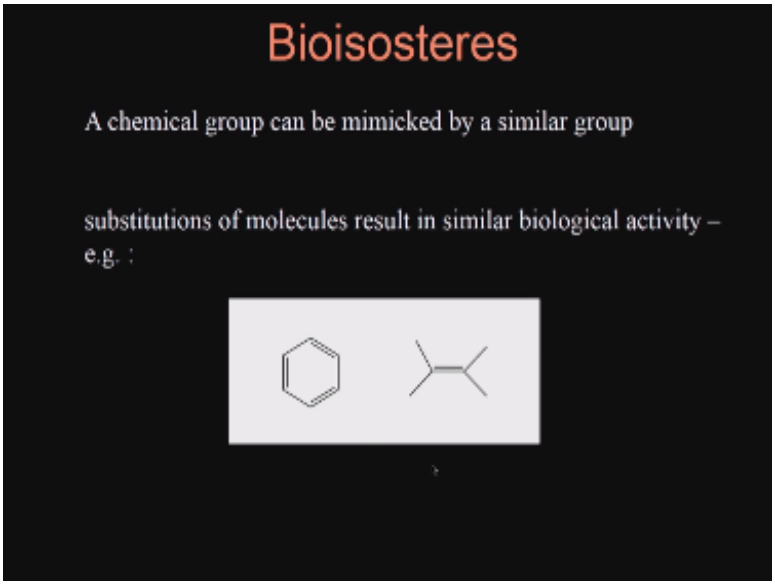
(Refer Slide Time: 13:31)



And what are the side effects it should cost especially to cardiac or create toxicity and so on okay, let us look at another important subject that is called Bioisosteres, okay this; what is this Bioisosteres? This is the phenomena observed between substances structurally related okay, so when they are structurally related, they will have similar or antagonistic biological properties, this is what is called Bioisosteres.

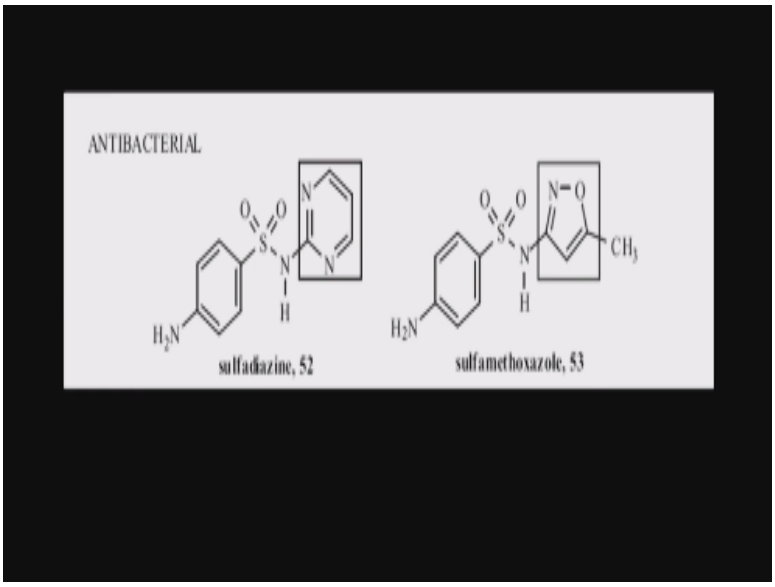
So, if they are structurally related chances are they will exhibit similar biological property, so these isosteres are elements, molecules or ions which have the same number of electrons at the valence level okay, so if they have same number of electrons at the valence level, the chances are they may exhibit similar property that is what is the concept of Bioisosteres and it is very useful when you are doing drug design okay.

(Refer Slide Time: 14:17)



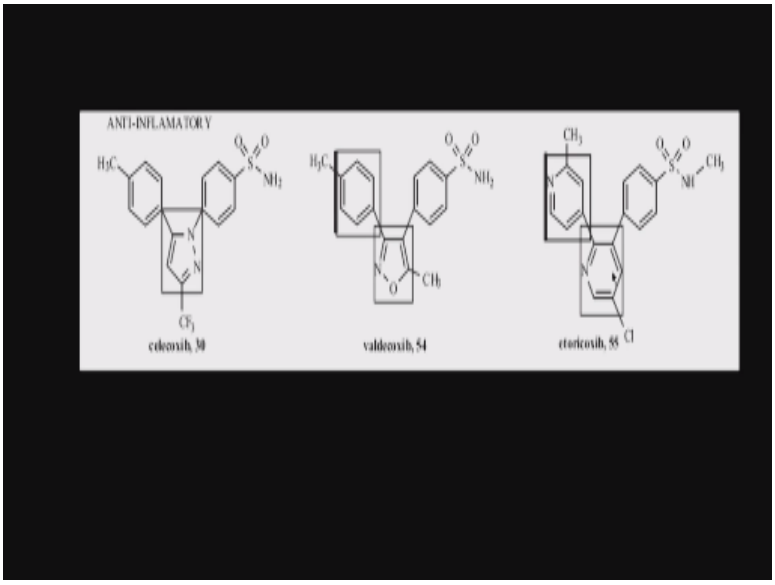
A chemical group can be mimicked by a similar group that is what is Bioisosteres, so this and this are Bioisosteres because you have here double bond here diode, so chances are they may have similar biological properties, so if you have a molecule with this as a group if I replace it with this group chances are it will have a similar biological activated okay.

(Refer Slide Time: 14:45)



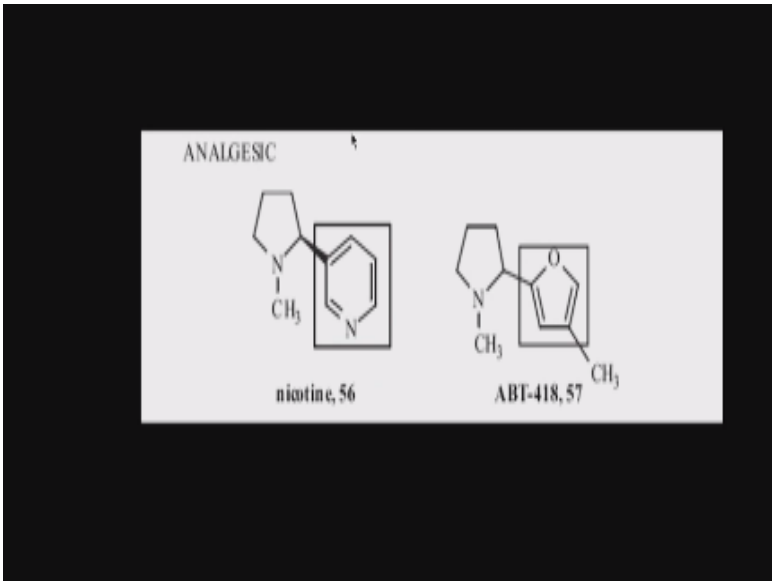
So, for example look at this, these are all anti-bacterial sulfa based drugs okay, so this portion is the same, so there are just a difference here okay; piperidine and then some hetero cycle; nitrogen oxygen hetero cycle okay, so these groups can be replaced against each other and you get a similar activity.

(Refer Slide Time: 15:08)



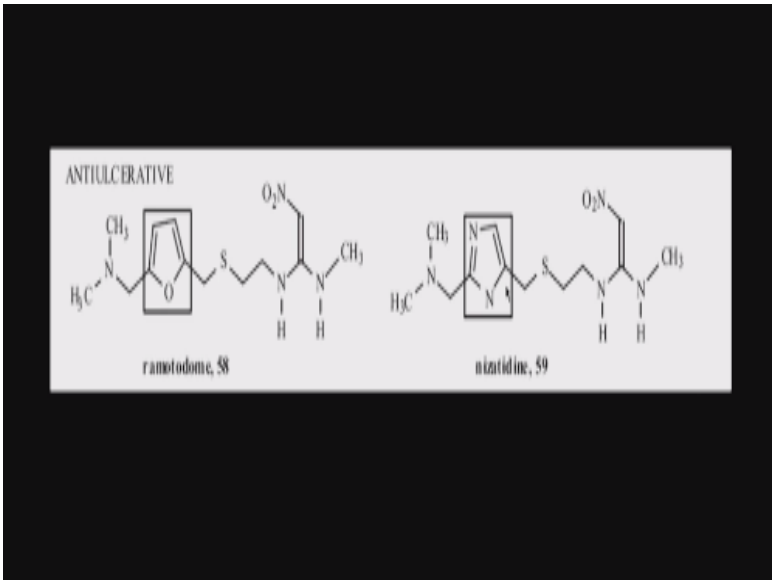
Look at this, I had shown this long time back, these are selectively cyclooxygenase to drugs, these are manufactured by Pfizer and Merck and so on, celecoxib, etoricoxib, valdecoxib look at this, they have; see the diaryl all of them and then there is a hetero cycle group here okay, 5 member or even 6-member hetero cycle, 6 members okay, so these compounds have similar activity not exact activity, similar activity.

(Refer Slide Time: 15:46)



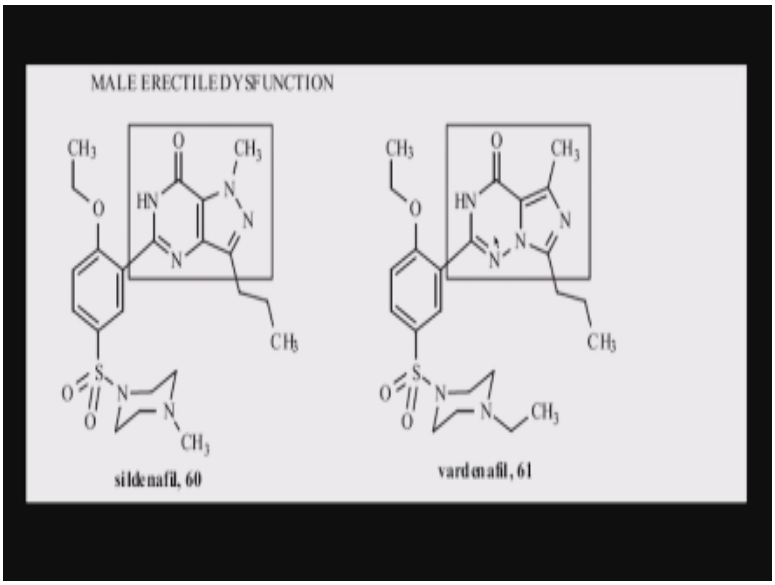
So, these can be called Bioisosteres; Analgesic; nicotine ABT- 418, 57, this is a clinical trial drug look at this okay, these are all; these 2 are Bioisosteres.

(Refer Slide Time: 15:57)



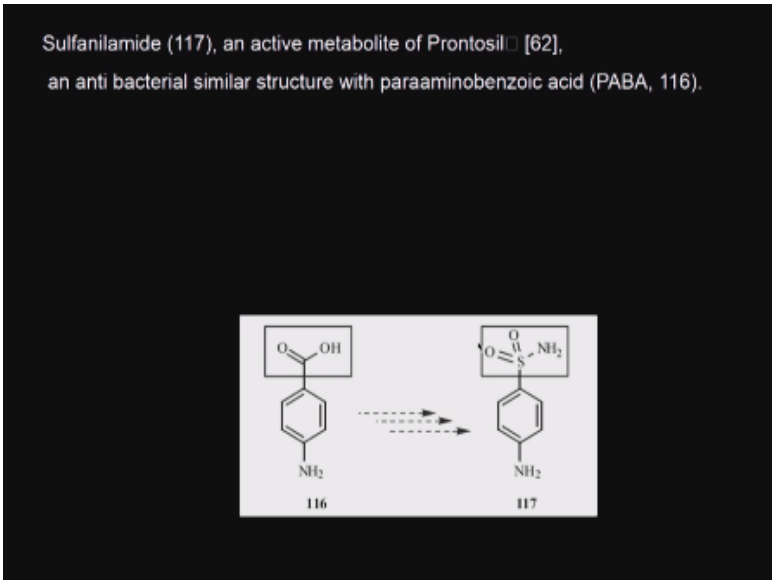
Antiulcer; this portion is the same, ramotodome and nizatidine; look at this, the hetero cycles are different, this is an oxygen 5 membered, this is the nitro 5 membered rings.

(Refer Slide Time: 16:16)



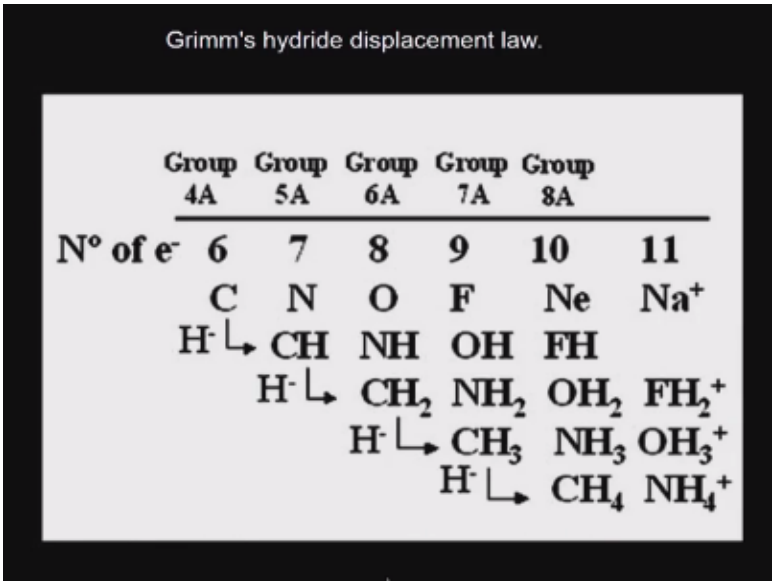
This is a male erectile dysfunction drug okay, sildenafil, vardenafil okay, look at this, this portion is the same, so this portion has been replaced and they exhibit similar activity.

(Refer Slide Time: 16:33)



So, sulfanilamide for example, an active metabolite of prontosil, an antibacterial similar in structure to para amino benzoic acid, so this is a substrate of bacteria, this is the sulfa drug which was designed to mimic the substrate, so it goes and binds to the same active site and kills the bacteria okay, these are sulpha drugs which were discovered long time back after the First World War, they are competitive inhibitors.

(Refer Slide Time: 17:10)

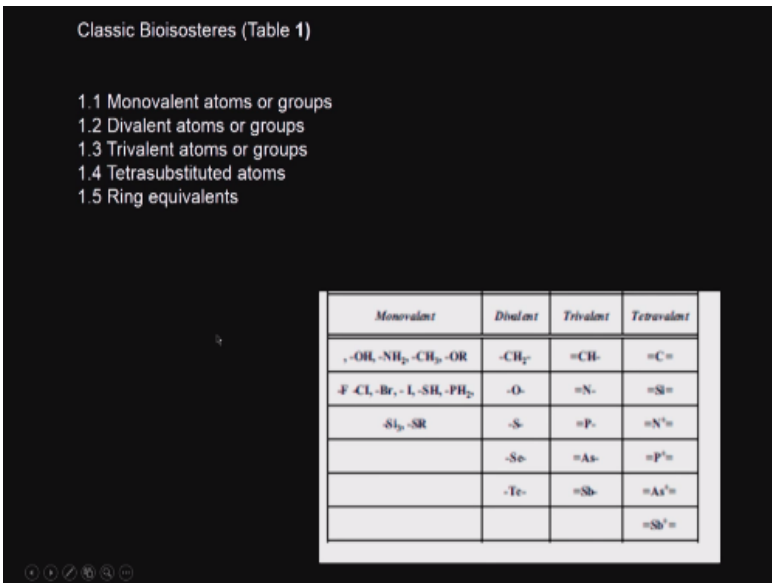


So, there is a Grimm's hydride displacement law, according to that law, these are Bioisosteres, these are Bioisosteres, these are Bioisosteres. So, what does it mean? If I have a molecule with the OH group, I can replace it with NH2 then I expect it to have similar activity or I can replace it with CH3 and expect it to have similar activity of course, the other properties may change; will change of course.

Because hydrophilic to hydrophobic, so solubility may change total polar surface area may change, hydrogen bond acceptor donation changes but we are talking about bioactivity; Bioisosteres talk only about bio activity. So, I could come up with new structures because maybe some properties are not good, so I want to make it more hydrophobic for example, so I may replace a NH2 with CH3 expected to have similar activity.

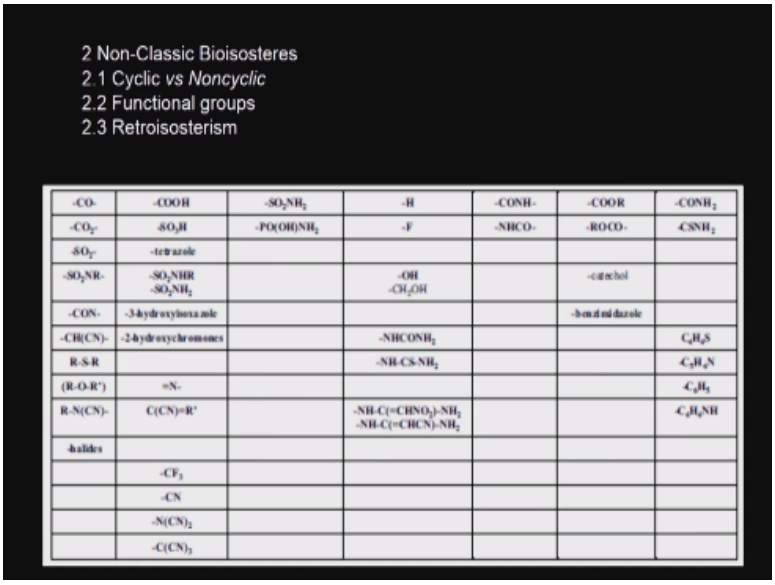
But as you know it is no more hydrogen bond acceptor, it is more hydrophobic as well, so some of the other drug likeness property may change, so it is same here. NH can be replaced by CH2 okay, so these are Bioisosteres based on these and hence we can design new molecules and expect to maintain the activity same but change the other properties like solubility, maybe log P maybe, hydrogen bond acceptor capability or hydrogen bond donating capability.

(Refer Slide Time: 19:00)



So, all these could be changed if I know, what are the Bioisosteres of certain functional groups, where I showed you a lot of examples of okay. Classic Bioisosteres; these are called classic bioisosteres, so like monovalent atoms or group, divalent atoms or group, trivalent atoms or group, tetrasubstituted atoms, ring equivalence all these are classic Bioisosteres as you can see here okay.

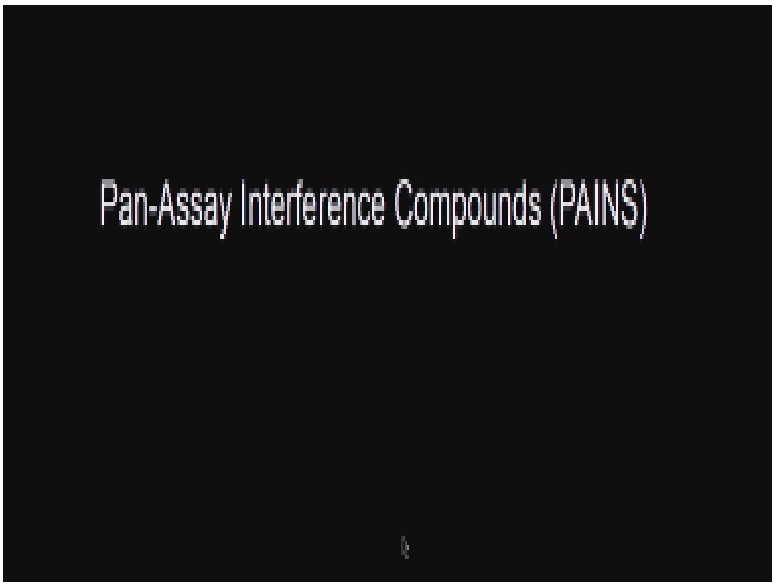
(Refer Slide Time: 19:18)



-CO-	-COOH	$-SO_2NH_2$	-H	-CONH-	-COOR	$-CONH_2$
$-CO_2-$	$-SO_3H$	$-PO(OH)NH_2$	-F	-NHCO-	-ROCO-	$-CSNH_2$
$-SO_2-$	-tetrazole					
$-SO_2NR-$	$-SO_2NHR$ $-SO_2NH_2$		-OH $-CH_2OH$		-catechol	
-CON-	-3-hydroxyisoxazole				-benzimidazole	
-CH(CN)-	-2-hydroxychromones		$-NHCONH_2$			C_4H_4S
R-S-R			$-NH-CS-NH_2$			$-C_5H_4N$
(R-O-R')	=N-					$-C_6H_5$
R-N(CN)-	C(CN)=R'		$-NH-C(=CHNO_2)-NH_2$ $-NH-C(=CHCN)-NH_2$			$-C_6H_4NH$
-halides						
	$-CF_3$					
	-CN					
	$-N(CN)_2$					
	$-C(CN)_3$					

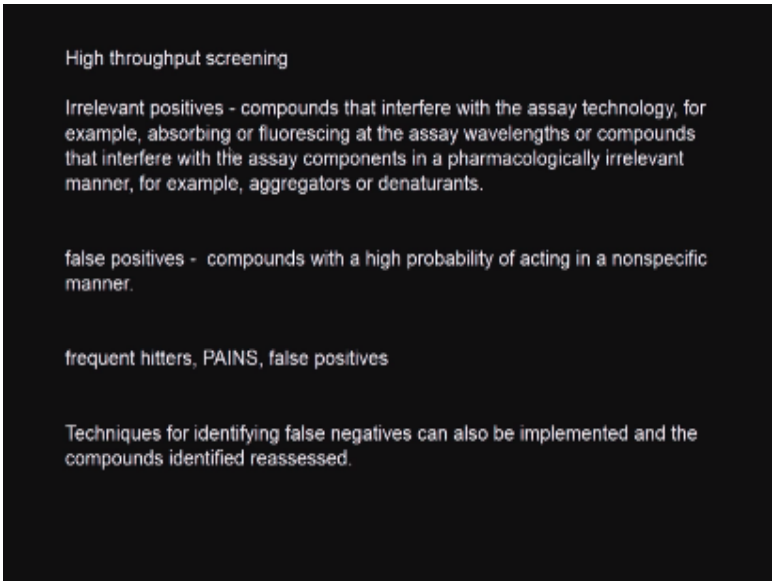
Look at this of course, you also have non classic Bioisosteres, cyclic verses noncyclic functional groups, retroisosterism, so we have groups which fall into this category okay, so these are called Bioisosteres, so the great advantage of it is; if I know which groups could be replaced by its own Bioisosteres, I expect the compound to have this similar activity but it will alter my other properties; the drug likeness property, physicochemical properties based on the type of functional groups I am replacing it with.

(Refer Slide Time: 20:25)



So, Bioisosterism is very nice trick to play on molecules to change these properties but maintain your activity similar. Now, another important concept which has become very important in the past 5 to 10 years I would say, this is called Pan assay interference compounds; PAINS. What does this Pan Assay interference compound means?

(Refer Slide Time: 20:33)

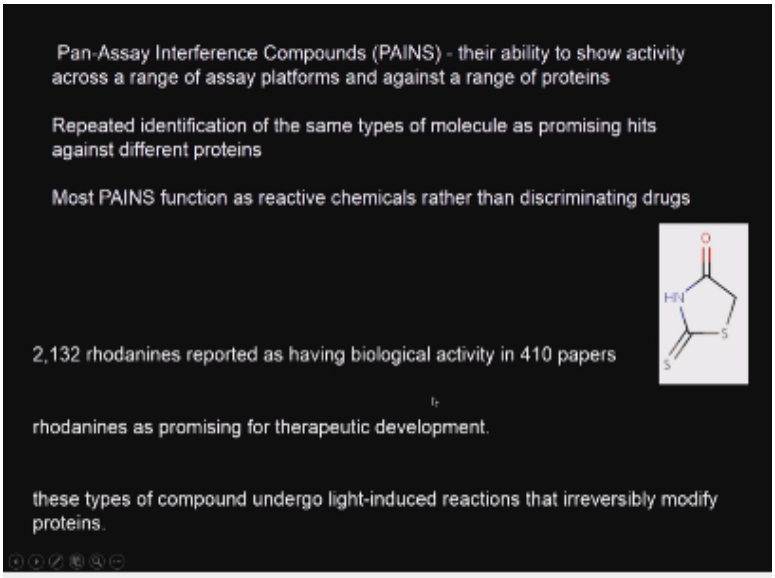


When we do high throughput screening, we get sometimes irrelevant positives, false positives frequent hitters, okay, compounds that interfere with the assay technology okay, so it is not actually doing a biochemical inhibition of an enzyme suppose, we are doing an enzyme assay with some compound, it is not actually working on the enzyme but it may be interfering with the assay.

For example, it may be absorbing or fluorescing at the assay wavelengths or compounds that interfere with assay components in a pharmacological irrelevant manner okay or it may be aggregated the protein, denaturing the protein, so it but it may appear as a hit when I do the screening. False positives; compound with a high probability of acting in a nonspecific manner okay; okay, it is not acting very specifically either as a competitive inhibitor or allosteric non-competitive.

Frequent hitters; okay it seems to be appearing in many irrelevant screens okay, they are called Pains, false positive. Techniques for identifying false negative can also be implemented and the compounds identified be reassessed okay especially, when we are doing this type of things okay.

(Refer Slide Time: 21:54)



Let us go more deeper into that, this Pan assay interference compounds; their ability to show activity across a range of assay platforms and against a range of protein, so you may be doing a study on proteins in the inflammation or you may be doing study on something else may be cardiovascular, the same compound may be appearing in many, many places not because they are inhibiting this protein.

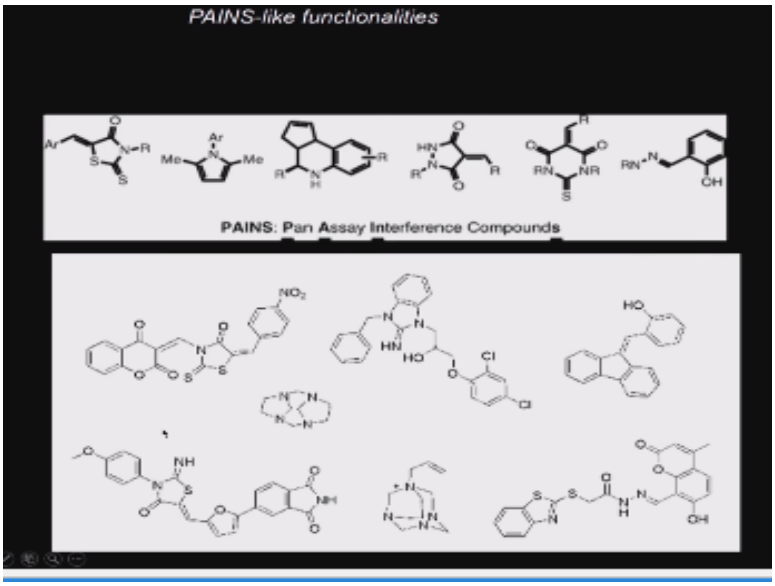
But maybe they are affecting the assay itself either fluorescing or absorbing in the same membrane, repeated identification of the same type of molecule as promising hits, so it may appear as a promising hit against different protein, so obviously in such situations are we careful, most pain function as reactive chemicals rather than discriminating drugs, so they may be reacting, they may be denaturing the protein or aggregating the protein rather than actually performing as a drug.

For example, this is a very, very dangerous candidate; rhodanine; there are 2132 rhodanines reported as having biological activity in huge number of publications rhodanines; rhodanines, so it may be thought as if they are very good for therapeutic development but actually they undergo light induced reactions that irreversibly modify the protein okay, so the protein gets denatured, they get modified.

Whereas, when we do the assay with in the presence of rhodanines, you may think they are acting on the protein, so they appear to be promising for therapeutic development and they seem to have a lot of biological activity. So, rhodanine there are almost 2000 of them, so they

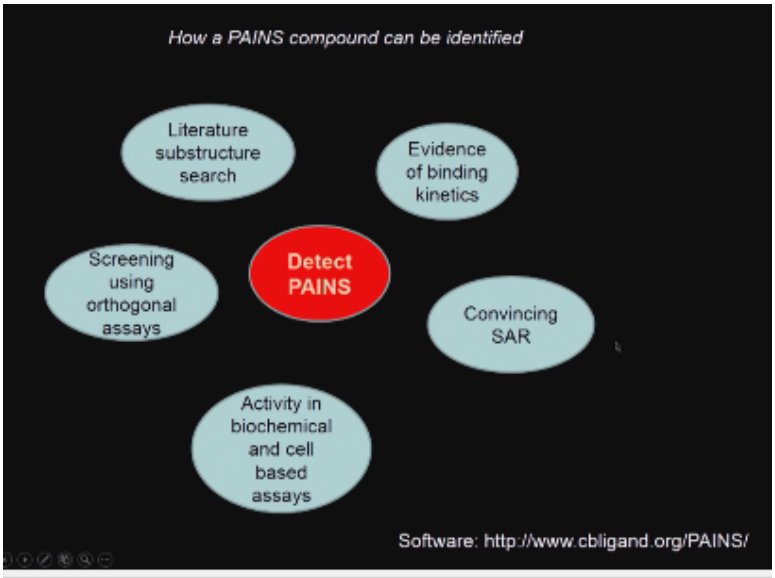
are not really compounds that may be taken up as a drug but these are compounds which affect your assay which affect your protein in an irreversible manner okay.

(Refer Slide Time: 24:12)



They are called the pan assay interference compounds, there are many compounds of that nature okay, many nitrogen containing with the double bonds and so on, look at this okay, lot of these 5-member nitrogen double bonds, nitrogen double bonds, huge number of compounds, there are softwares, freeware, where you upload your compound and it tells you whether your compound may be a interfering component; pan assay interfering component okay.

(Refer Slide Time: 24:43)



These; if your compound has these type of functionality is a better watch out, it may be a interfering component okay. So, this software; this is what I said, if you go to this software, you can upload your structures and be sure possibly whether it is a pan; pains compound or not. So,

how do we detect pains? We have to look at literature substructure search, we need to do some binding studies.

If you are doing a protein based biochemical assay, if you think one compound is showing good activity, it is good to do a binding kinetics okay that means you change the concentration of the substrate, concentration of the inhibitor and see whether it follows the Michaelis Menten, so that is the binding kinetics. So, from there you will exactly find out whether your compound comes under this category of pains.

Activity in biochemical and cell based assay; so you do the biochemical and then go to cell based assays also, screening using orthogonal assay; so okay you have found a hit in one particular assay, now try and I say which is totally different; orthogonally different and then see whether the compound shows, convincing SAR; so if you have a compound which shows activity then if I do structural modifications, electron withdrawing, electron donating or a hydrophobic hydrophilic.

Do I get a good structure activity relationship? If not then you better be careful that it could be a pains compound okay, so you need to be very sure that your compound does not fall under this category and it is a genuine inhibitor for a particular protein of which you are studying okay. So, this concept of pains has become very important in the past 10 years, so that you do not end up with some false positives in your screening protocols okay.

So, we will continue more in the next class, thank you very much for your time.