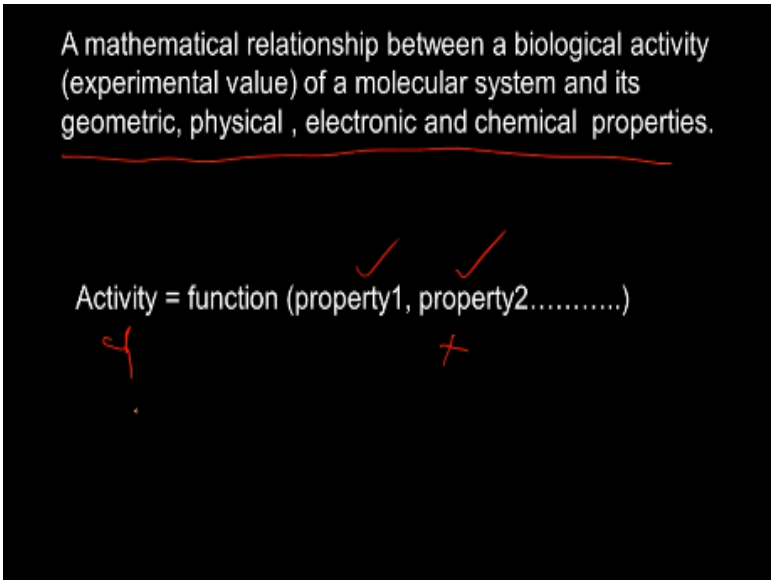
Computer Aided Drug Design
Prof. Mukesh Doble
Department of Biotechnology
Indian Institute of Technology - Madras

Lecture - 26
Quantitative Structure Activity Relationship(QSAR)

Hello, everyone, welcome to the course on computer aided drug design, we will continue on the topic of QSAR that is quantitative structure activity relationship, what is this?

(Refer Slide Time: 00:27)

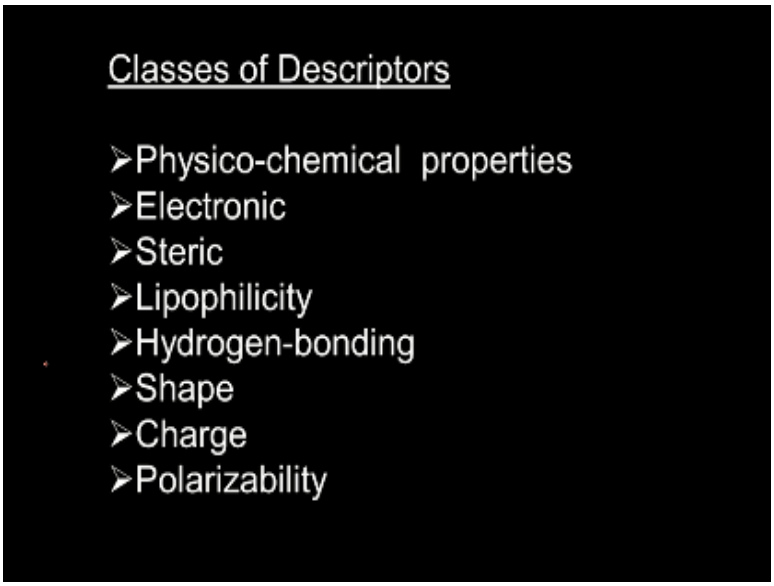


QSAR is just nothing but a mathematical relationship between the biological activity and that is the experimental data, or we can collect it from the literature of a molecular system and it is geometric, physical, electronic and chemical properties. So, the dependent variable will be the activity and the independent variable will be a parameter, descriptor or structural feature okay it could be a geometric activity.

Descriptor, physical descriptor, electronic descriptor, or chemical properties, so it could be any one of them could be any one of them. So, there are lot of descriptors available in the literature we will talk about some of them as you go along. So, basically you are developing a mathematical relation a linear regression or a nonlinear regression. So, the property is called the independent variables.

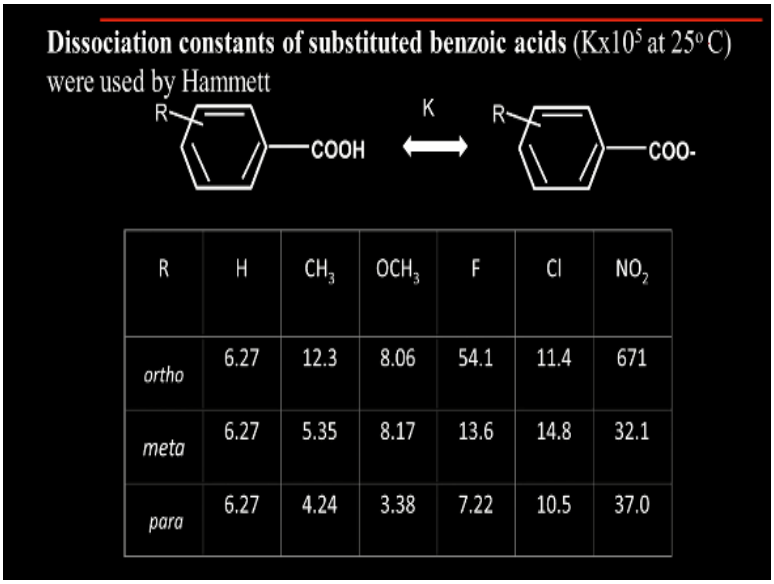
Okay or sometimes you call it x and activity to the dependent variable we call it y, so if you remember your statistics we get y=function of various x okay.

(Refer Slide Time: 01:39)



So, the descriptors could be physical chemical descriptors like your shape, size and so on electronic descriptors electronic features steric descriptors, volume, lipophilicity, solubility how lipophilicity is log p, hydrogen bonding capabilities okay number of hydrogen bonding formation, shape charge, polarizability. So many descriptors thousands and thousands of descriptors can be calculated for a molecule.

(Refer Slide Time: 02:10)



R	H	CH_3	OCH_3	F	Cl	NO_2
ortho	6.27	12.3	8.06	54.1	11.4	671
meta	6.27	5.35	8.17	13.6	14.8	32.1
para	6.27	4.24	3.38	7.22	10.5	37.0

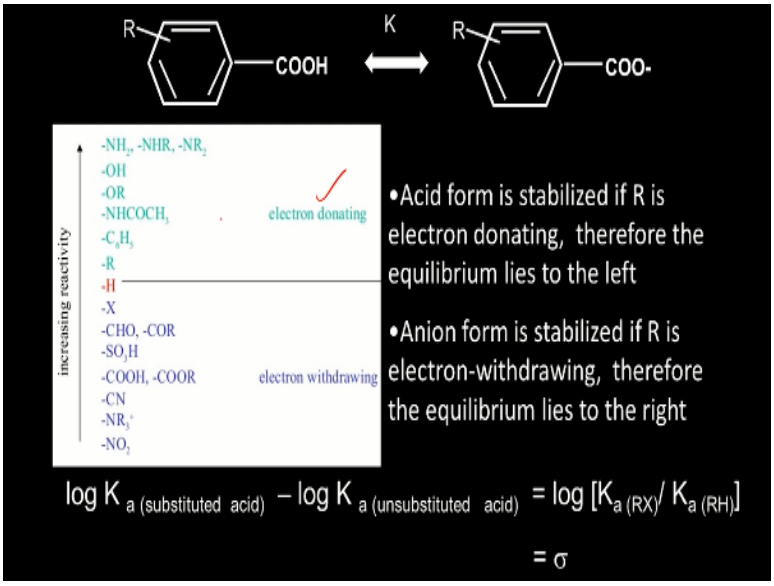
Okay we started with an electronic descriptors or it is also called used by Hammett okay

Dissociation constants of substituted benzoic acids okay these are called substituted benzoic acids and these benzoic acid can dissociate okay you have a dissociate constant and then dissociate into the acid forms COOH and COO and so on. Okay that is the substitute benzoic acid.

Now depending upon the R group and depending upon the position ortho meta para you can have different dissociation constant that is what this table gives. You can see fluorine that is there in ortho okay ortho it has a very high dissociation constant okay when you go down to CH3 type of group in the para okay here you can have a very low dissociation constant right you can see. Also depending upon the position, you can have and the type of group.

You can have different dissociation constants okay electron withdrawing or donating which stabilizes the anion that is called the anion, COO- is called the anion.

(Refer Slide Time: 03:24)

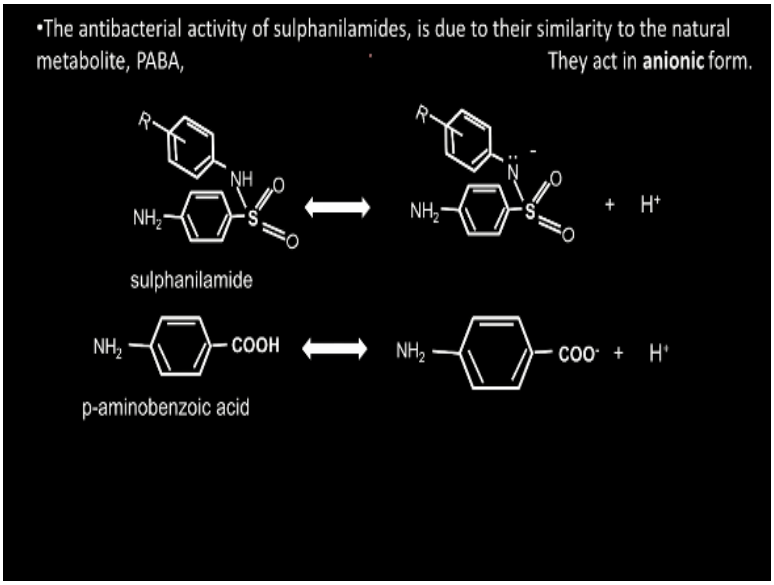


Okay, so electron donating groups the electron donating groups we have okay then OR, OH we have the electron donating groups whereas electron withdrawing groups we have here x x could be okay low CHO -COR, CN NR these are electron withdrawing groups acid form is stabilized if R is electron donating. So, all these groups so this is the acid this is the acid form this is the Anion form.

So, the acid form is stabilized when we have electron donating group okay the Anion form is stabilized if R is electron withdrawing group. So, if the electron is withdrawn like x and so on you get the Anion. Whereas acid form is formed when you are getting electron donated like OH, OR, NHCOCH3 and so on. SO, if the anion form is stabilized then you are going to have a large k if the acid form is stabilized you are going to have a small k okay.

So, log ka- log ka and substituted, log k is substituted that is called sigma.

(Refer Slide Time: 04:49)

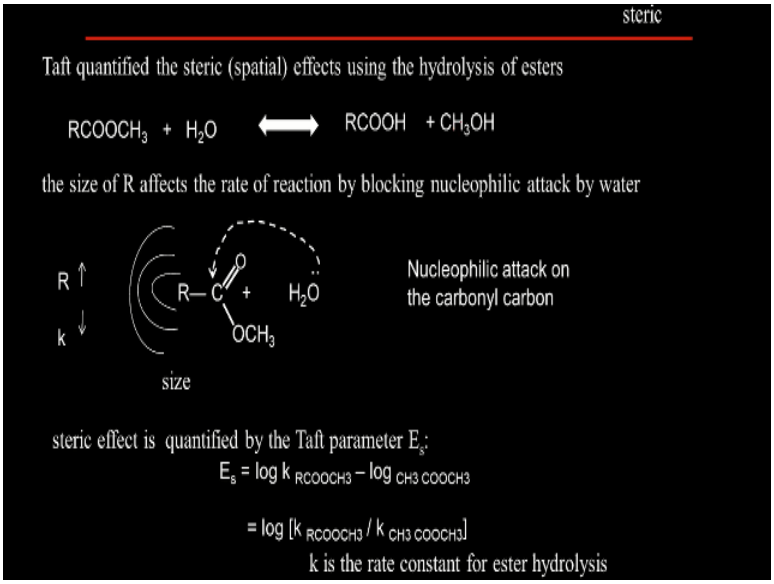


So, when does all this come into picture, if you look at anti-bacterial activity of sulfanilamide these are called sulfanilamide, they are called Sulphur globe and they form a- and H+ here okay depending upon the R either you will have this or you will have the Anion form, the Anion form or this form, but the anion form is active anti-bacterial agent not the neutral form okay because if you look at para amino benzoic acid this is which is the substrate for bacteria.

Okay and this Sulpha drugs also looks similar to that so there is a competitive inhibition that is happening and hence the sulpha drug act okay. So, the anion form is the active part and not the neutral form so that means we need to have this dissociation taking place. So, depending upon the R like we saw on the position of the R either the Anion or the neutral form is stabilized. So, if you have very high sigma we have very high sigma okay.

You have log 1/c also going up that means the activity R stabilize the anion by electron withdrawal, thus increasing biological activity okay whereas if we have an electron donating and then the biological activity goes down okay. If the electron with drawing, we have the biological activity also going up. Okay this is with respect to the electronic feature.

(Refer Slide Time: 06:33)

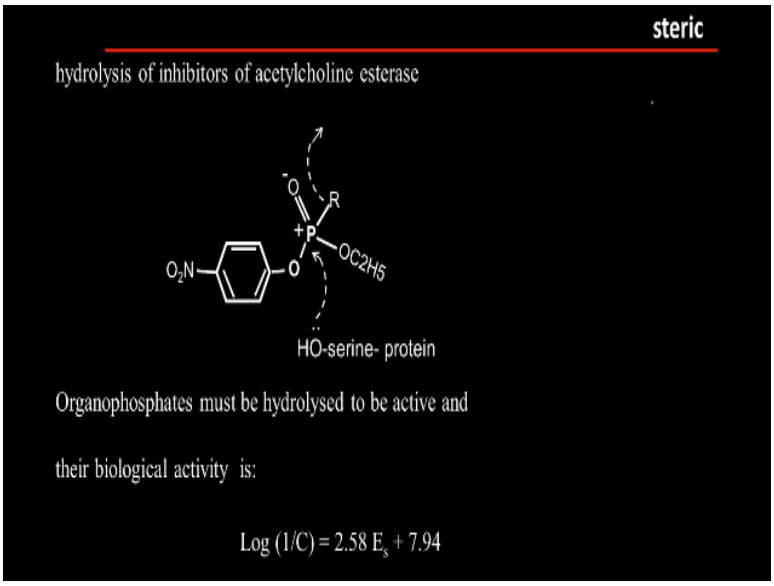


Now let us look at the other one that is called Steric feature if you look at the hydrolysis of the ester this is a ester RCO CH3 R could be anything and hydrolysis means water is acting, so you need a catalyst an enzyme or something like that, so it forms RCOOH+CH3OH that is the methanol okay methyl alcohol. So, the size of the R affects the rate of reaction because what happens is there is a Nucleophilic attack on the carbonyl carbon here R okay.

There is a Nucleophilic attack here okay which leads to the formation of RCOOH as well as C3 OH here. SO, the size of the R tells you how easy the Nucleophilic attack is or how difficult it is if the R is very large it is very difficult for the Nucleophilic attack, so the activity goes down. if the R is small activity goes down, this is called steric effect, or this is also called Taft parameter This is calculated by log k of RCOOCH3-log CH3 COOCH3.

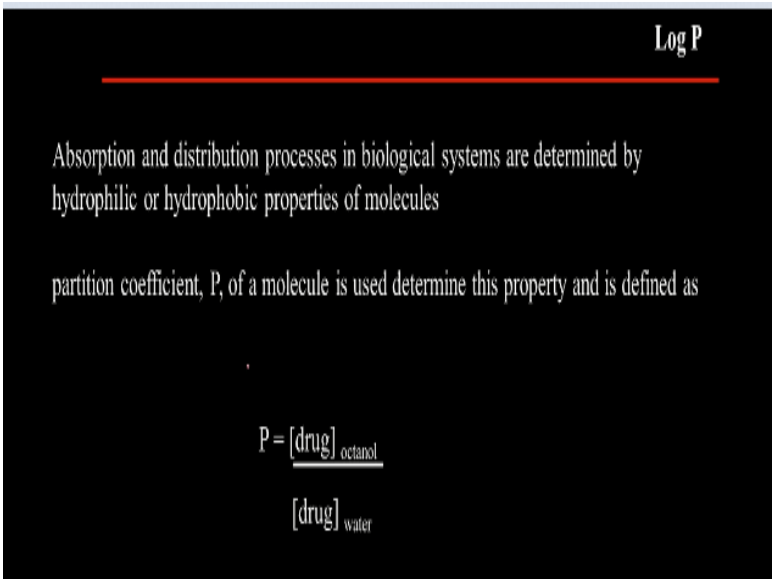
Okay k is the rate constant for ester hydrolysis. Okay So larger it is okay, so we have larger the R lower will be the rate constant okay.

(Refer Slide Time: 07:59)



Why does this important if you look at acetylcholine esterase there are inhibitors for acetylcholine esterase? So, hydrolysis of this acetylcholine takes place here okay it attacks here okay so organophosphates must be hydrolyzed to be active and their biological activity depends upon the size of this group the size leaving group here. SO, the activity is directly proportional to the size of leaving group here and it is called the Taft parameter.

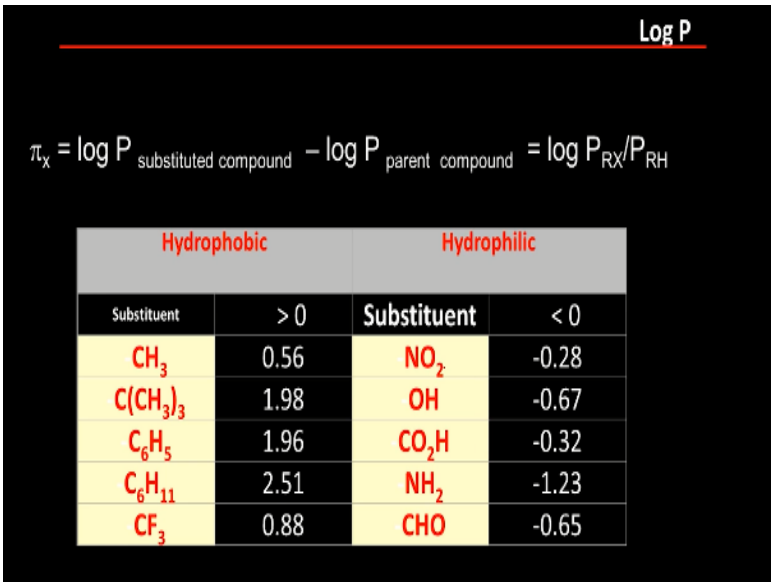
(Refer Slide Time: 08:40)



Another descriptor which is very, very important is the Log p here okay, that is the Log P. Absorption and distribution process in biological systems are determined by hydrophilic or hydrophobic properties of molecules. We looked quite a lot about this Log p, Log p is nothing but n octanol/n water the partition of this particular molecule okay ore in the octanol you call it

as hydrophobic less in octanol more in water we call it hydrophilic. So, many processes depend.

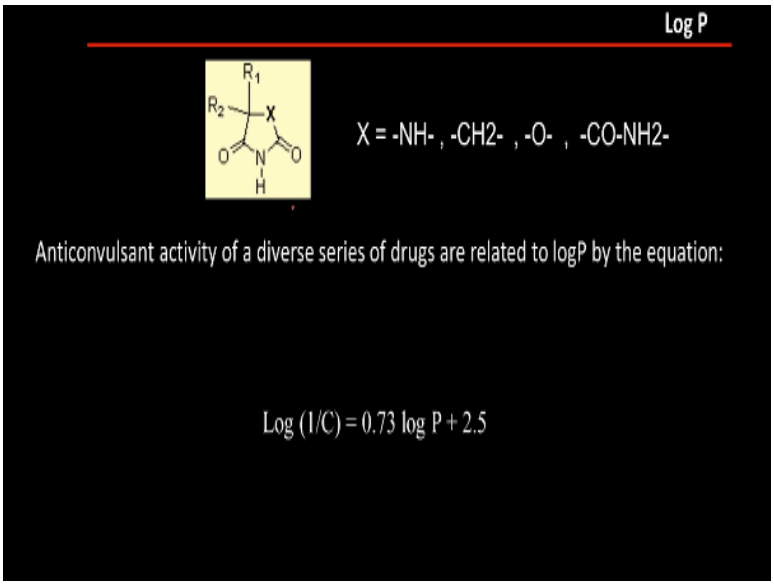
(Refer Slide Time: 09:12)



And as you can see depending upon the substitution, each substituting contributes towards log p, for example CH3 is 0.56 that is more hydrophobic if you have a say a treasury we have 1.98 okay Arial group 1.96 C6 H11 2.51, okay whereas if you look at nitro, OH , NH2 they are all negative because they all are hydrophilic. SO, if you have a molecule if you substitute with a particular functional group we can find out the contribution.

We can subtract the original log p of the parent compound, the log p of the parent compound. So, log p plays a very important role

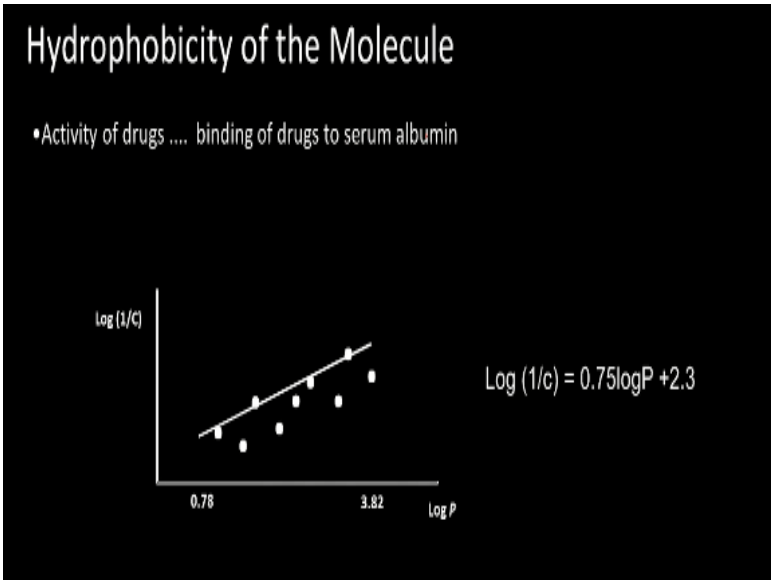
(Refer Slide Time: 10:00)



For example, look at Anticonvulsant activity of a diverse series of drugs okay these are the series of drugs we are talking about, they have R1. R2 groups here and then we have X here, X could be NH CH2 O CONH2, okay this is a five-member ring, so the activity of these drugs these are Anticonvulsant drugs okay they all depend upon the log p. I can see more hydrophobic more is the log p activity will be more, so log p plays a very important role.

We will look at log p more in detail because it plays a very important role because the absorption through depends on log p as we saw long time back.

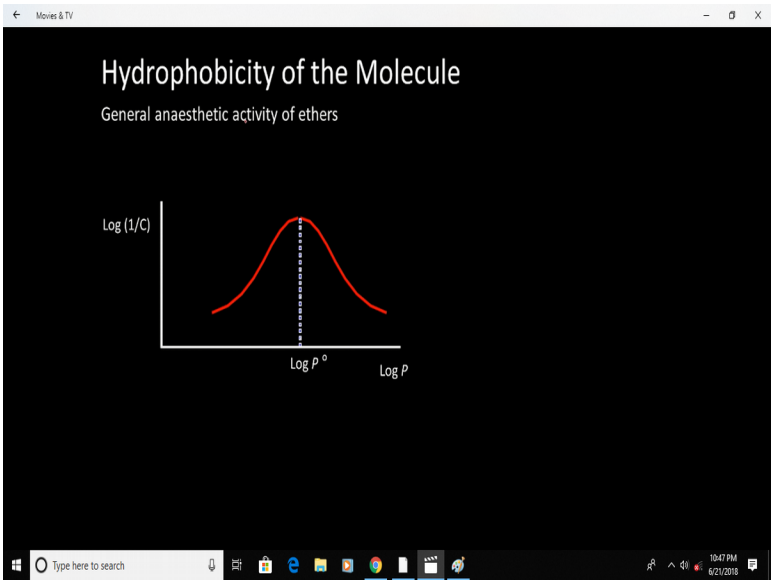
(Refer Slide Time: 10:45)



So, if you look at the drugs which binds to serum albumin okay so as log p increases because as

it becomes more hydrophobic and activity also increases by this is the QSAR this is repulsive relation so as the log p increases this linear regression you can see here the activity also increases this way. It need not be all the time linear okay greater for hydrophobic drugs binding increases as log p increases.

(Refer Slide Time: 11:23)

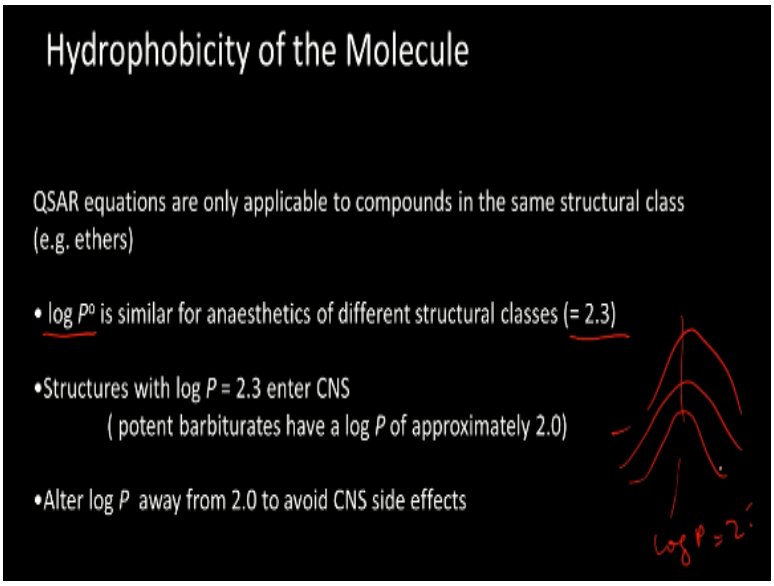


It need not be all the time linear it can be like this also, there can be an optimum log p, log p 0 when the activity is max and initially it increases which is the max and then comes down. So, generally for anesthetic activity of ethers it is followed like this. So, it is a nonlinear relation, so the regression relation could be nonlinear as you can see again it depends on log p, but it is log p square and log p.

So, we have a nonlinear regression relation okay, so we have a constant term here and then we have a squared term for log p here. And then we have a linear term for log p. So, the relation will go like this increases initially okay because of this it increases fast as it goes to a large p value this term starts taking place. So, the log p starts coming down okay, this is an optimum log p when you have the maximum activity.

This is applicable for the ethers only anesthetic activity. SO, you can have systems where this type of behavior may be observed okay so this is called the optimum value anesthetic activity okay.

(Refer Slide Time: 12:47)

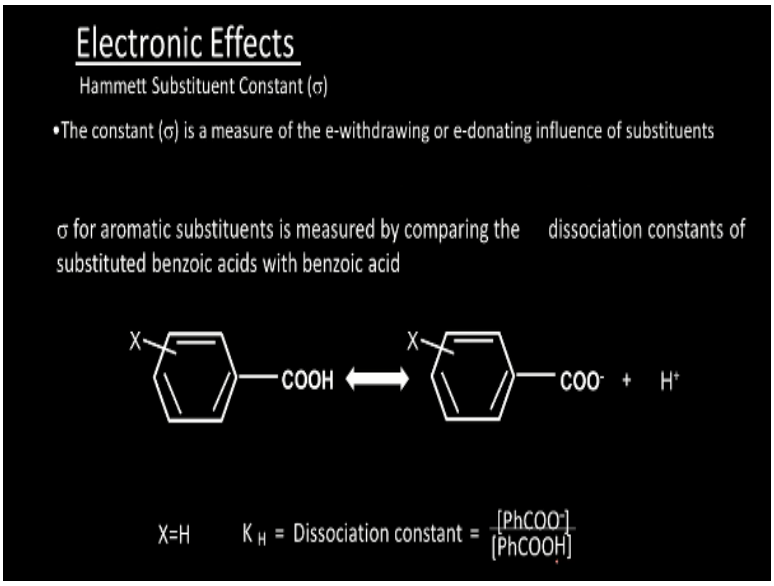


So, hydrophobicity applies quite a lot actually one thing you need to remember is QSAR equations are only applicable to compounds in the same structural class okay do not forget that. I perform a QSAR for one class of compounds like Benzoic acid and then I tried extending it to some other trial systems it might not work at okay. So, for a different structural classes also interestingly the optimum log p that is given here.

The optimum log p is given here seems to be 2.3 so around 2.3 you may get a very high activity for many of the structural classes also interesting. Structures with log p around 2.3 enter CNS that is your central nervous system as you remember blood brain barrier entering we need approximately 2.3to enter. If it is lower, it will be hydrophilic, so it might not enter. SO, for many classes for anesthetic drugs.

Even though you may have different types of graphs okay so approximately the log p comes over to be approximately 2.3 okay this is a very important point to keep in mind okay it is very useful to keep in mind. Barbiturate they have a log p approximately 2 so this number is quite interesting. So, if you want to avoid CNS side effects we need to be away from 2, so that it either comes down or down here so that it does not pass the blood brain barrier.

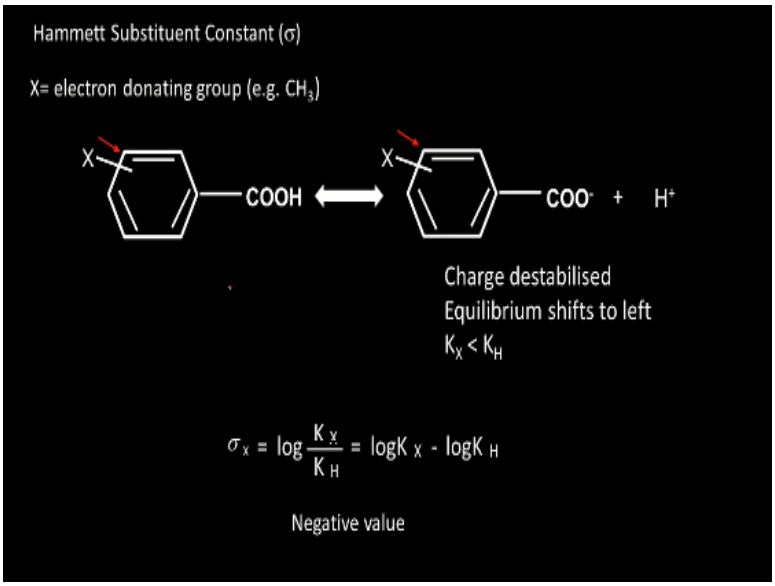
(Refer Slide Time: 14:31)



So, we looked at electronic effects the constant is called as the sigma it is called the Hammett substituent constant okay that is the measure of electron withdrawing or electron donating influence of substituents sigma for aromatic substituents is measured by comparing the dissociation constant of substituted benzoic acids with benzoic acid. So, sigma is a ratio of you can find out the substituted versus unsubstituted.

Okay and the dissociation constant is given by PHCOO-/PHCOOH COO- means Anion form COOH is the acid form. we have read that long time back.

(Refer Slide Time: 15:17)

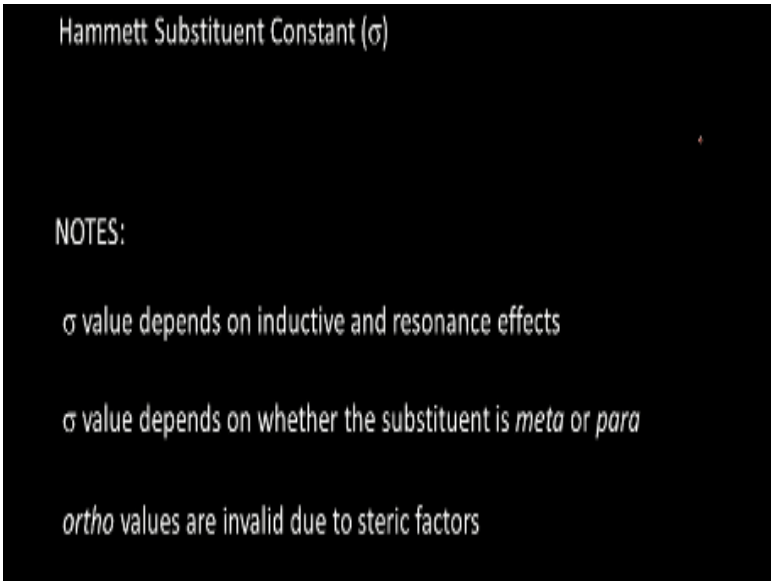


Agree, so the sigma x we have the electron withdrawing group for example NO2 okay the charge

is stabilized by X equilibrium shifts to right okay equilibrium shift to the right okay that is Kx>Kh means substituted Kh substituted. SO, electron and withdrawing group, equilibrium shifted here Kx> Kh, so sigma is a positive value. Electron donating group CH3 for example, okay it is donating then what happens charge is destabilized equilibrium shifts to the left here.

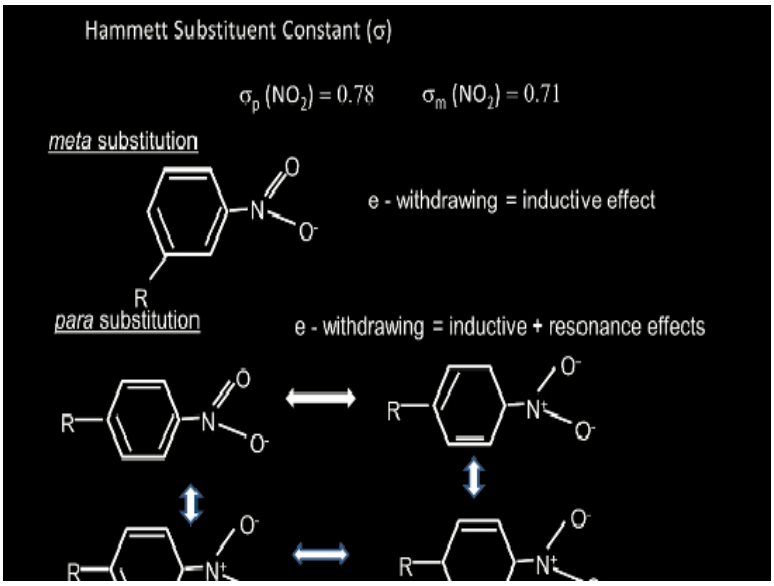
So, Kx is<Kh. So, sigma X will be negative value, it is called Hammett substituent Constant.

(Refer Slide Time: 16:21)



Okay sigma value depends on inductive and resonance effects okay sigma depends on whether the substituent is Meta or Para that is also very important. We saw long time back a table where Meta and Para have different values of dissociation constant. Ortho values are invalid due to steric effects, Ortho means the substitution is so close, so because of steric it might not work.

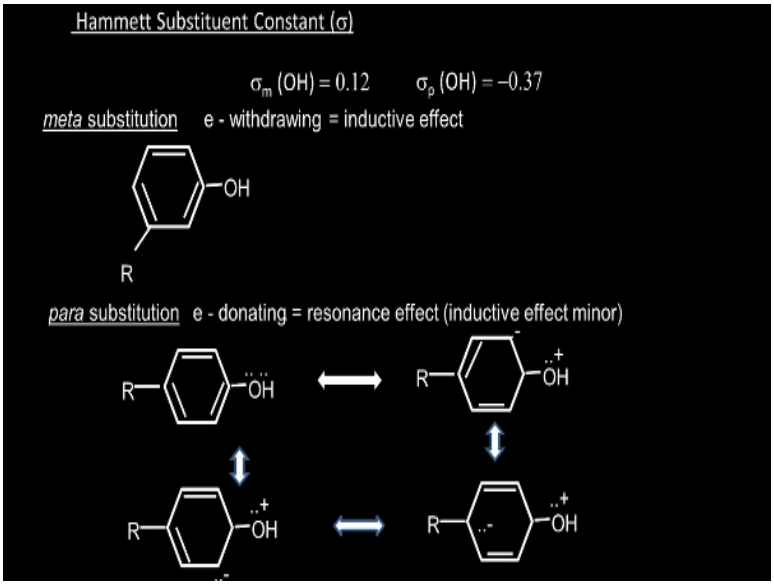
(Refer Slide Time: 16:45)



Okay so Meta substituent, look at this this is NO2 okay we have the R here, we have para NO2, sigma is 0.78 Meta NO2. So, sigma is 0.71 okay e-withdrawing = inductive effect. Okay when you have para substitution e withdrawing inductive+resonance effects because when we have the para substituting we have there is a resonance that is happening in the nitrogen as you can see and double bonds in the benzene keep moving like this, this is called resonance effect okay.

So, you have this like this. So, when the electron is withdrawing you can have inductive and resonance effect that is para substitution okay there is a meta substitution we will not have the resonance effect happening only the inductive effect will be happening as you can see here.

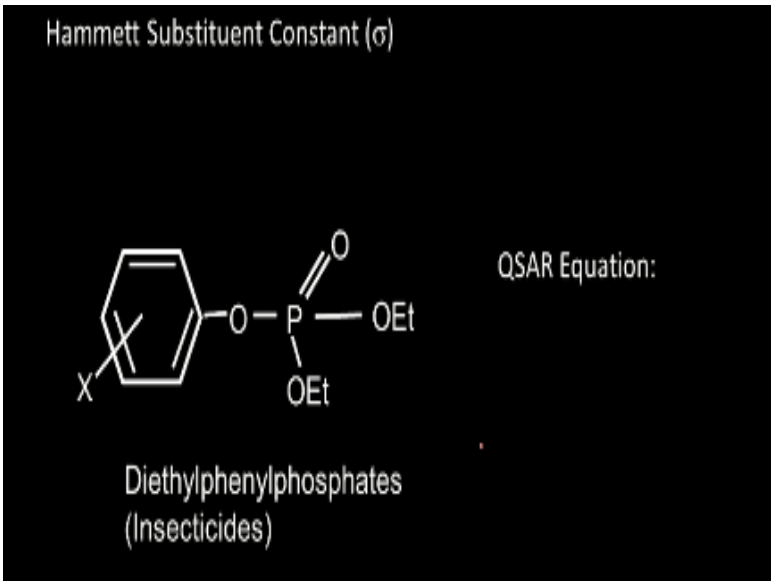
(Refer Slide Time: 17:54)



When we have OH type of group. We saw NO2 Meta we have sigma M meta 0.12 Para -0.37 okay there is only inductive effect. Whereas electron withdrawing whereas electron donating pre-dominant resonance effect inductive effect is minor. So, we have resonance effect happening as you can see how the double bond get shifted, shifted, shifted okay, so you have the lone pairs happening here.

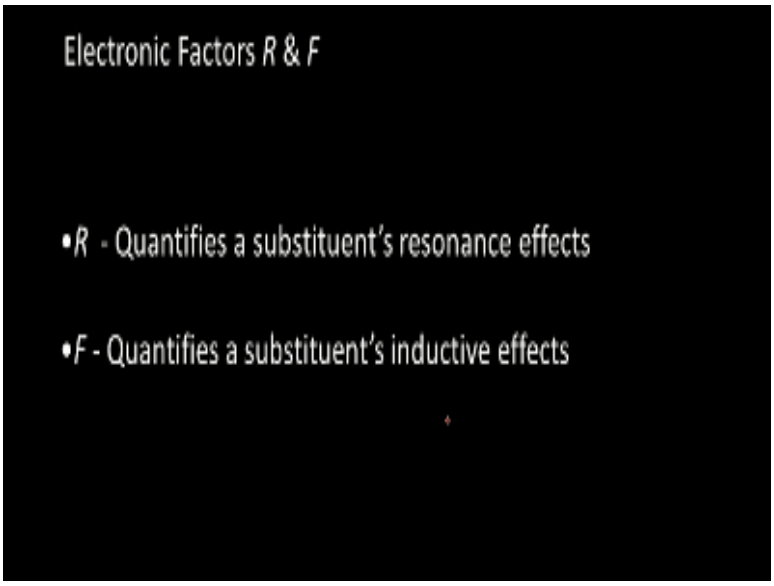
Okay, so it is -0.37 okay what does that mean a neutral form is more stable than the Anion form here the neutral form is the Meta substitution form is less stable in the Anion form okay. And whereas in Nitro Meta and Para almost same.

(Refer Slide Time: 18:50)



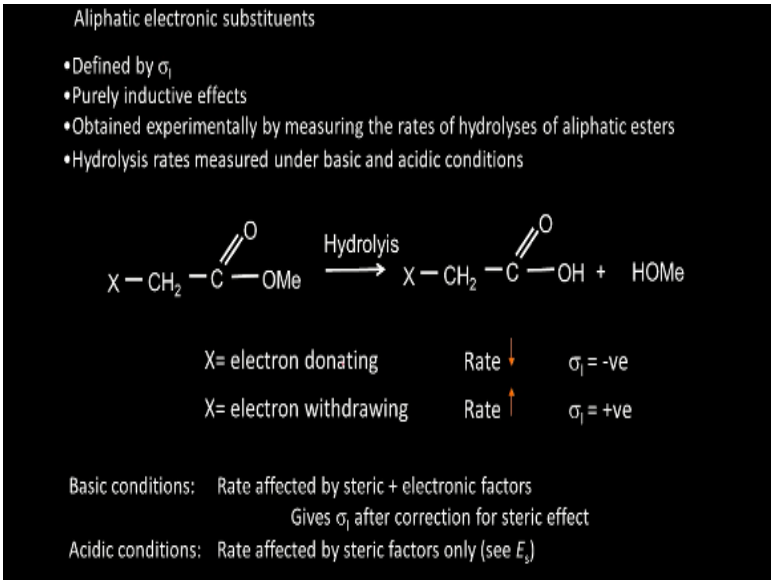
Diethyl phenyl phosphates you have to either group here or there used as insecticides we have equations like this 2.282 sigma-0.348 okay. So, electron withdrawing substituents increases the activity.

(Refer Slide Time: 19:11)



So, R quantifies a substituent resonance effects F quantifies a substituent inductive effects okay one is called the resonance and the other is called the inductive like I showed you here.

(Refer Slide Time: 19:28)



Aliphatic electronic substituents that is we have the Aliphatic group here, so you can have the hydrolysis like this Aliphatic, so we have Me OH how does they act they are defined by again sigma 1 you have only purely inductive effects and they are obtained experimentally by measuring the rates of hydrolyses of aliphatic esters okay hydrolysis rates measured under basic and acidic conditions.

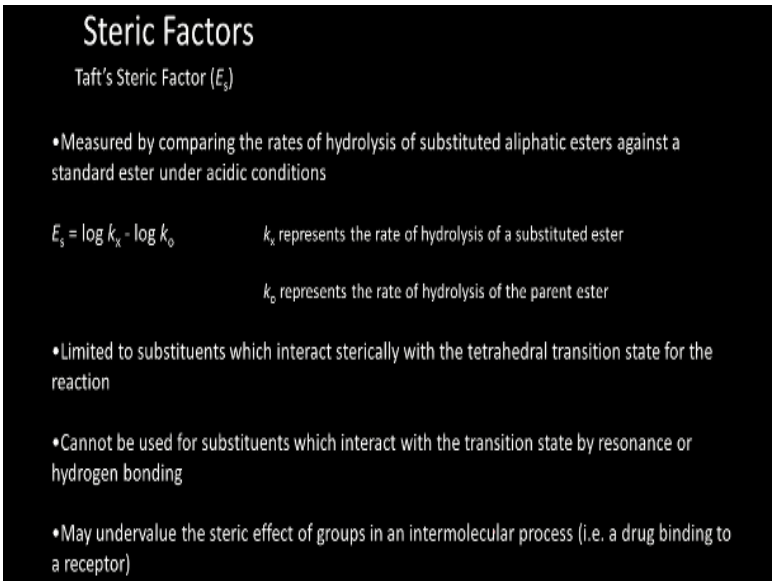
So, you do not have the resonance happening here you will always have induction happening

here because unlike the aromatic where we have the resonance bonds Aliphatic will not have. So, X electron donating rate goes down because sigma 1 is negative. Next electron withdrawing rate goes up because sigma 1 is positive okay did you understand. So, in Aliphatic systems there is only inductive effect.

There is no resonance effect how you measure we can measure by looking at the rates of hydrolysis of Aliphatic esters. Hydrolysis is measured under basic and necessity condition. So, if you have electron donating then we have the rate going down if you have electron withdrawing rate goes up, So, basic condition is rate affected by steric like I mentioned steric+electronic factors so give sigma1 after correction for steric effect.

And acidic condition rate is affected by steric factors only that is yes, we talked about years before Taft size factor do remember that. Under basic condition we will have Steric and Electronic effect. Under acidic condition we will have only steric effect that is why here we mentioned sigma L here because you can also have steric effect also coming into picture.

(Refer Slide Time: 21:31)

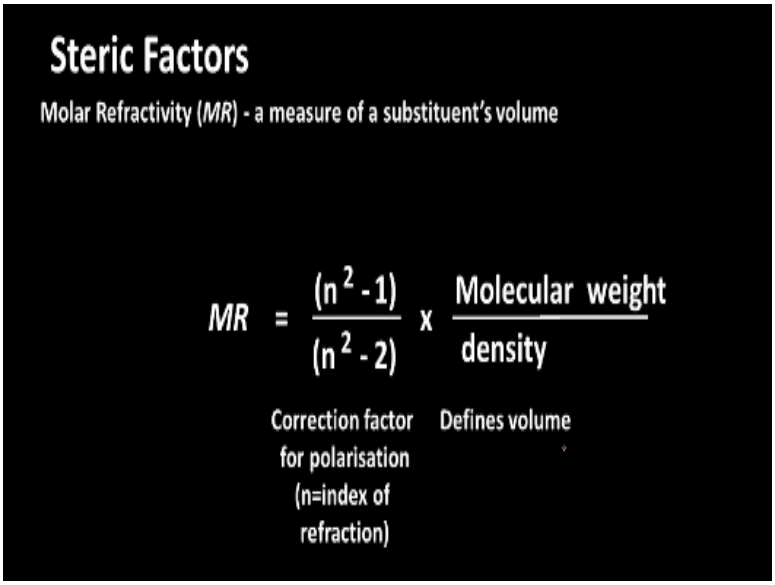


Steric factors Taft Steric factors we talked about is compared the rates of hydrolysis of substituted Aliphatic esters against a standard ester under acidic conditions okay if you bring basic then you are going to have a electronic effect coming in term okay. So, we can have a Kx log of Kx -log of K0 Kx represents the rate of hydrolysis of a substituted ester rate of hydrolyses

of the parent ester.

Okay limited to substituents which interacts sterically with the tetrahedral transition state for the reaction remember that. Cannot be used for substituents which interacts with the transition state by resonance or hydrogen bonding there it will start interfering the resonance effect and hydrogen bonding all this will affect your Taft parameter. May undervalue the steric effect of groups in an inter molecular process it is drug binding to a receptor sort of situation.

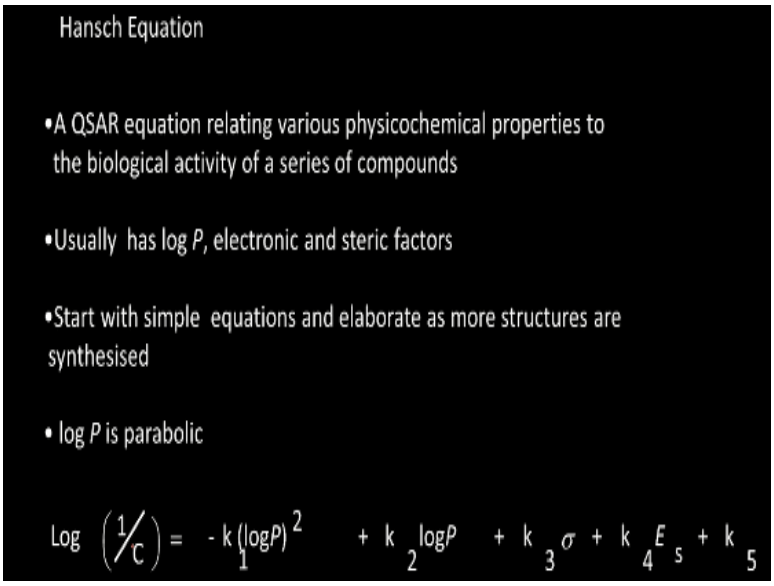
(Refer Slide Time: 22:27)



You also have a term called Molar Refractivity it is a measure of substituents volume this is a descriptor which we can calculate that is called Molar refractivity it is given by n square-1/n square-2 molecular weight/density okay this is a correction factor this is called index of refraction correction factor for polarization okay it is called index of refraction. So, it is a function of molecular weight function of density that is called Molar Refractivity.

When you do molecular weight with density of course you end up with volume right. So, this is a descriptor which we use many times.

(Refer Slide Time: 23:04)

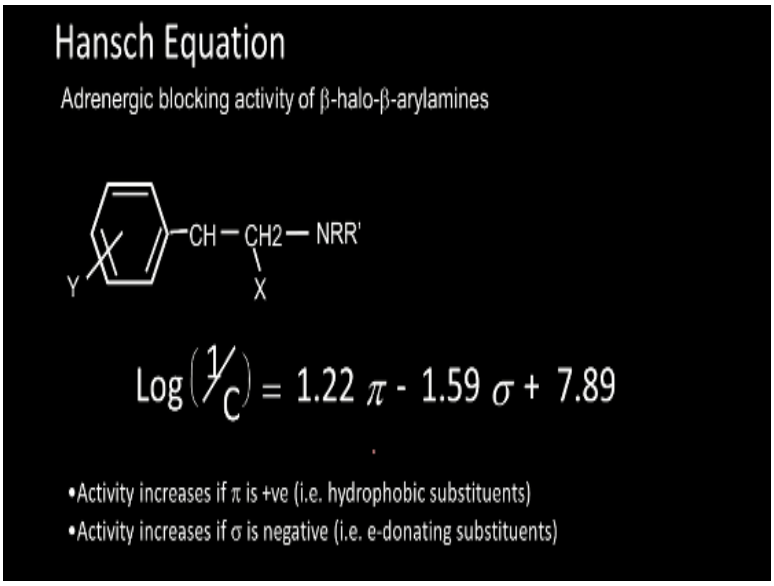


Then comes Hansch equation, what is Hansch equation A QSAR equation relating various physicochemical properties to the biological activity of a series of compounds. So, you will have log p coming into picture electronic factors coming into picture steric factor. Okay So we can have a QSAR activity on the left-hand side which may have terms related to electronic, steric and log p.

So, you start with the simple equation and you can elaborate as more structures are synthesized. Log p you can put it as a parabolic equation like this so log p square log p electronic term steric factor and this is a constant. So, typical Hansch equation okay remember QSAR we can develop only when we have some experimental data for the activity that is the left-hand side, or we get some data from literature.

Where someone else have estimated the activity for the series of compound okay.

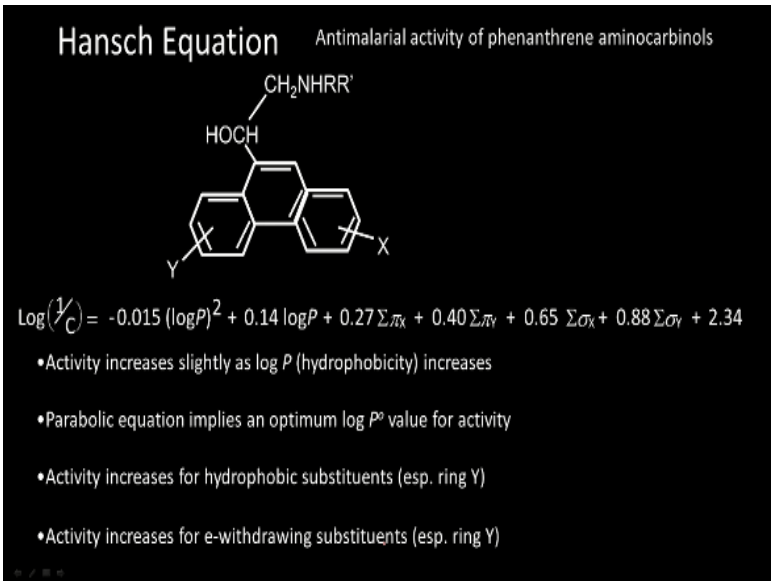
(Refer Slide Time: 24:14)



So, look at this Adrenergic blocking activity of beta halo okay this is beta halo beta aryl amines these are amines aryl amines as you can see here we have the nitrogen here and that is why we call Aryl amines then we have the halo compounds here So, typically this how the Hansch equation looks like 1.22 -1.59 1.22 pi -1.59 sigma pi represents the hydrophobicity sigma represents the electronic.

Okay, so it is a positive with respect to the hydrophobic substituents and it is negative with respect to electronic factor that is electronic donating substituents. This is called as a Hansch equation. So, the log p related term comes here the electronic comes here but there is no mention much about steric information here.

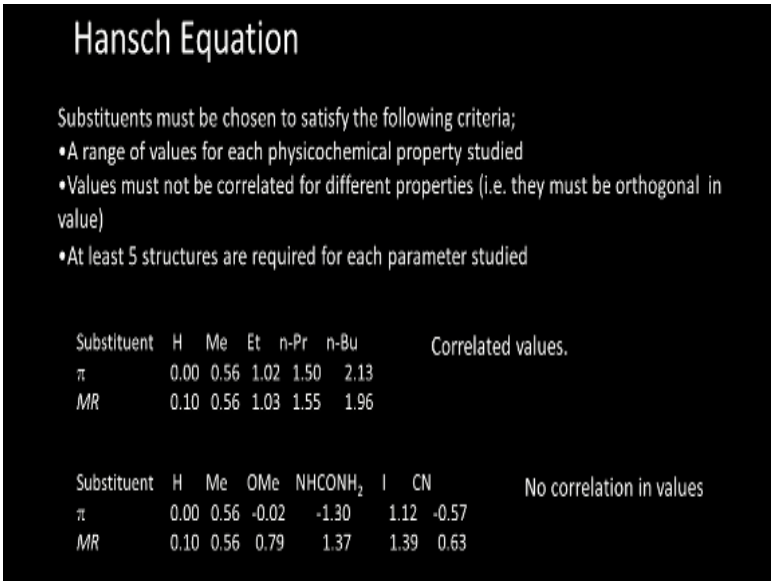
(Refer Slide Time: 25:21)



These are anti-malarial activity of phenanthrene amino carbinols okay, so we have 3 fused engine rings okay we have the Amino group here that is why Amino carbinols because we have the OH here look at them. We have log p square log p okay then comes the hydrophobic part of it and then we have the electronic part of it. Okay we have 2 terms one for sigma x and sigma y similarly for pix and pi y.

Activity increases slightly as log p. Parabolic equation implies an optimum log p value for activity. Activity increases for hydrophobic substituents especially ring Y here and then activity increases for electronic withdrawing substituents especially ring Y I think it should be y activity increase for electronic yeah and okay.

(Refer Slide Time: 26:37)

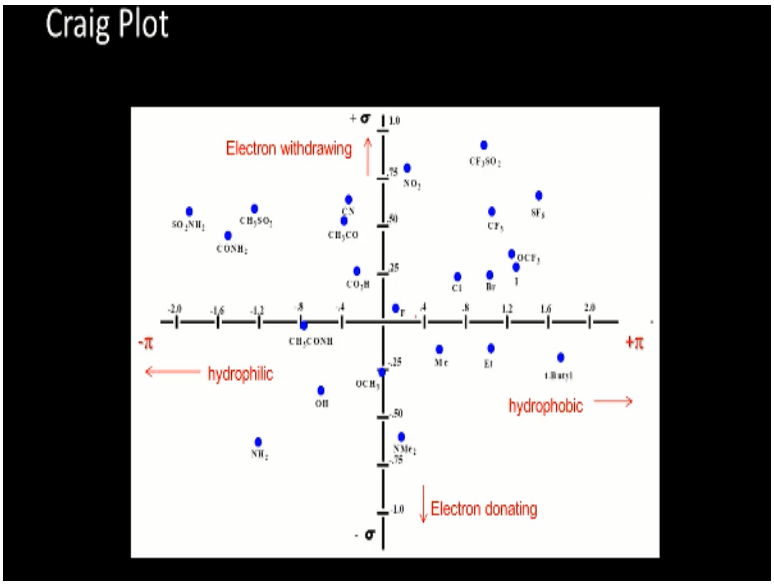


Substituents must be chosen to satisfy the following criteria, a range of values for each physicochemical property studied. So, when I synthesize new molecules I will try to substitute so that I get a different pi that means different hydrophobic hydrophilic values I get different sigmas and so on. Values must not be correlated for different properties means they must be orthogonal in value.

At least 5 structures are required for each parameter studied. So, this is is a very important rule of thumb okay for each parameter studied. So, for example I am looking at different I am substituting in H or ME Ethyl or N prophyll or N-Butal the five change like this. So, for H 0 it becomes more hydrophobic hydrophobic molar refractivity changes like this and I defined what is Molar refractivity before okay correlated values.

So, we have substituent H Me OMe pi MR as you can see some of them there is no correlation in values here there is no co relation in values. These are all co related Mr and the pi.

(Refer Slide Time: 28:12)



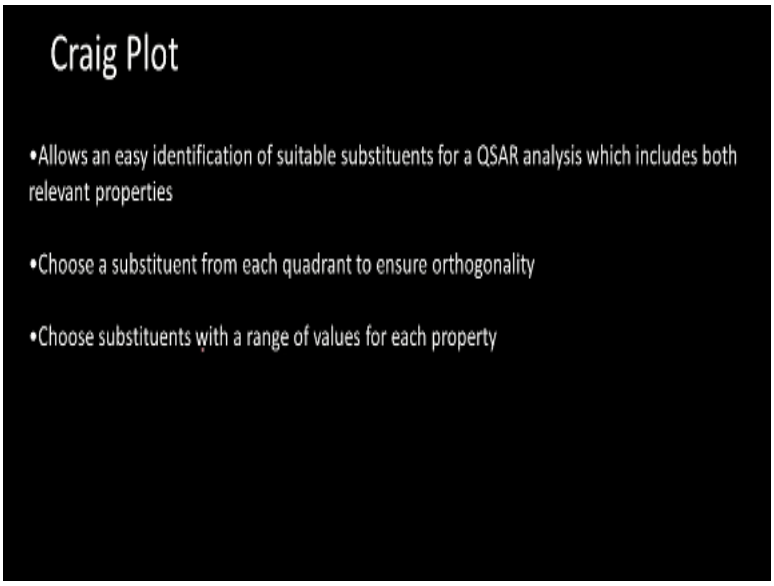
There is something called Craig plot which is also used in QSAR this is a Craig lot what is this we have 2 things happening one is the pi we have on the x axis okay pi that is the hydrophobic hydrophilic substitution sigma is the electronic substitution positive sigma positive pi this is negative pi negative sigma. Okay so different substitution groups as you can see NO2 falls here it has got a very high electronic and very low pi.

If you look at fluoro very, very low sigma and pi if you look at SF pi Sulphur with pi flow range it has got high pi as well as reasonably very high sigma. Okay this is electron withdrawing so here we have the electron donating, here it is more hydrophobic , here it is more hydrophilic , so if you have say COOH it is a hydrophilic and it is also electron withdrawing if you look here this is hydrophilic OH and it is electron donating.

So, this is a very interesting plot. So, we can decide which substituent to put into our parent molecule to our study. Do I want to look at electron withdrawing at effect by maintaining the hydrophobic hydrophilic balance or do I want to have a modifying the hydrophobic hydrophilic activity, nature by maintaining the hydrophobic hydrophilic activity y nature by maintaining the electron sigma?

So, by deciding I can synthesize different types of molecules and study their activity so that I will have a very fruitful QSAR relationship that is very, very important.

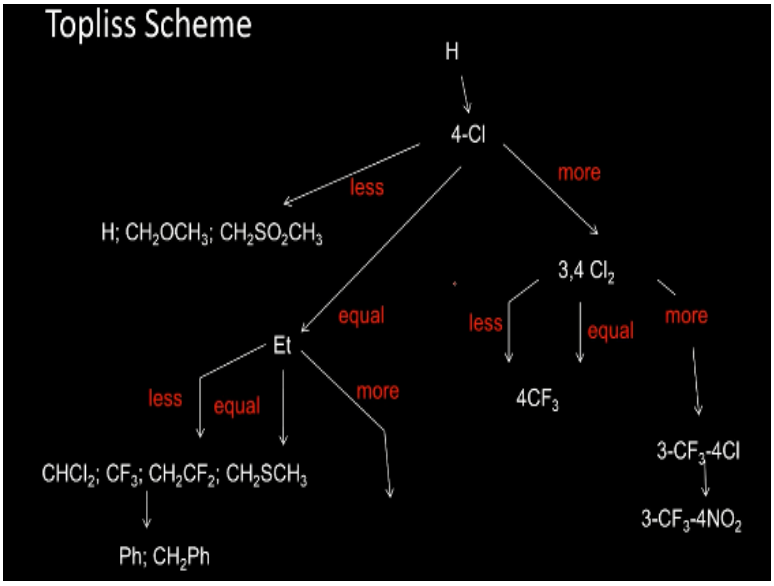
(Refer Slide Time: 30:33)



Allow an easy identification of suitable substituents for a QSAR analysis which includes both the sigma and the pi chose a substituent from each quadrant to ensure orthogonality that is very important. Choose substituent with a range of values for each property. Okay so I can choose substituents say from a groups here I can choose from here I can choose from bottom I can choose from this that way I keep orthogonality.

I can choose one low one high one low one high, so I can get the range also into the picture range of values that is the advantage of that particular figure.

(Refer Slide Time: 31:05)



Then we have something called the Topliss scheme okay which can help you to synthesize molecule in a very, very effective way so that you address both the hydrophobic as well as the electron features that is electron withdrawing and electron donating features okay, so we will continue this Topliss scheme in the next class okay. Thank you very much for your time.