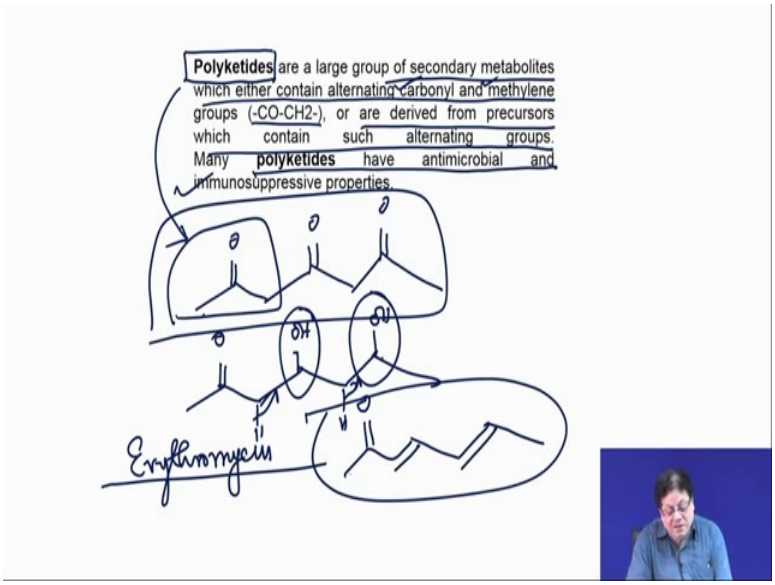
**Organic Chemistry In Biology And Drug Development**
**Prof. Amit Basak**
**Department of Chemistry**
**Indian Institute of Technology, Kharagpur**

**Lecture - 54**
**Polyketide Biosynthesis**

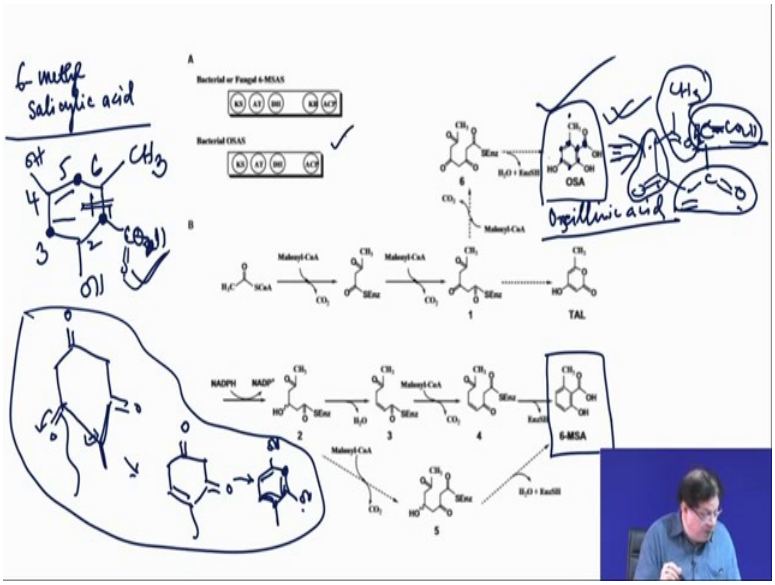
(Refer Slide Time: 00:31)




Welcome back, in the last session we have studied the biosynthesis of penicillin and we have seen how different structural variations can be done in the penicillin skeleton utilizing different tripeptides where the valine is changed to other amino acids. And, that can be converted into the different types of penicillin by the enzyme which is called Isopenicillin N Synthase (IPNS).

Now, let us talk about the other strategy where we manipulate the genes which are producing the enzymes, necessary for making the antibiotics. You manipulate at the genetic level and then get different types of antibiotics. Remember I told you that there are two ways you can do these changes, one is by forcing the enzyme to accept alternate substrates, the other is by changing the genetic map which will produce newer enzymes and then you can get newer types of metabolites.

(Refer Slide Time: 01:49)




Now, so we will talk about the second strategy and for that we will discuss the antibiotics which are also classified as polyketides. Now remember polyketides basically represent a large group of secondary metabolites. What are secondary metabolites? Like penicillin, cephalosporin, vancomycin all are secondary metabolites, not primary metabolites. What are primary metabolites? Primary metabolites are that without that which you cannot survive. Secondary metabolites are basically elicited for some other reasons like defense mechanism for the bacteria.

So, they are secondary metabolites, like plants produce terpenes, plants produce other different natural products, they are called secondary metabolites, monoterpenes sesquiterpene. Now, there must be some reason why plant is making those secondary metabolites, but without that terpene, the plant also survives. So, the primary metabolites are the ones which are required for the growth, for sustaining the system, the living system, secondary metabolites are made just to augment their defense or like it could be to attract the other partner; such secondary metabolites are called pheromones.

Now, out of these secondary metabolites, one class is polyketides. What are polyketides? Polyketides are a large group of secondary metabolites which either contain alternating carbonyl and methylene groups. That means, if I write this CO then $CH_2$ then CO then $CH_2$ then CO, like this.

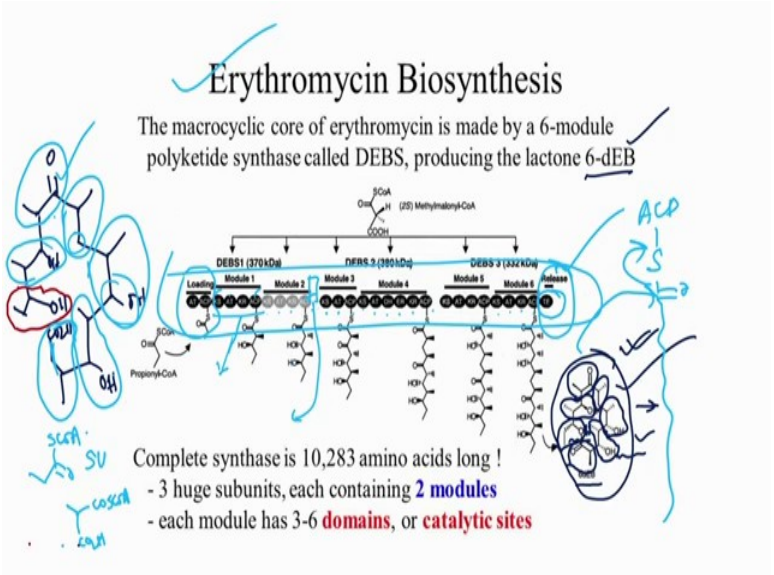
So, this is a polyketide. So, that will be a polyketide or they are derived from precursors which contain such alternating groups. Like some compound may be there, secondary metabolite which will be like this. Now, this is derived basically from this triketone, what happens in the process of generating the triketone, some of the ketone groups are reduced. So, this is also polyketide.

So, basically polyketides are either repeating units of $CH_2CO$ or it could be units which are derived from repeating units of $CH_2CO$. Many of the polyketides have good antimicrobial properties, some are anti-cancer drugs and some are immunosuppressive agents. In fact, the major proportion of the natural products which are used as anti-microbial agents or anti-cancer agents are basically derived from polyketide mechanism, or they belong to the class of polyketides.

We will take one polyketide for our purpose, like when we talked about the utilizing proteins to make new antibiotics; we took penicillin as our reference. Here we will take erythromycin; you know erythromycin used to be a very good drug for treatment of upper respiratory infections. Today erythromycin has lost most of its potency.

So, this has been the first generation; then after erythromycin it came the roxithromycin and then roxithromycin was replaced by azithromycin. Today, azithromycin are prescribed. But all are based on the skeleton of the erythromycin. Now, so we will take erythromycin as our example. Erythromycin is a macro cyclic compound.

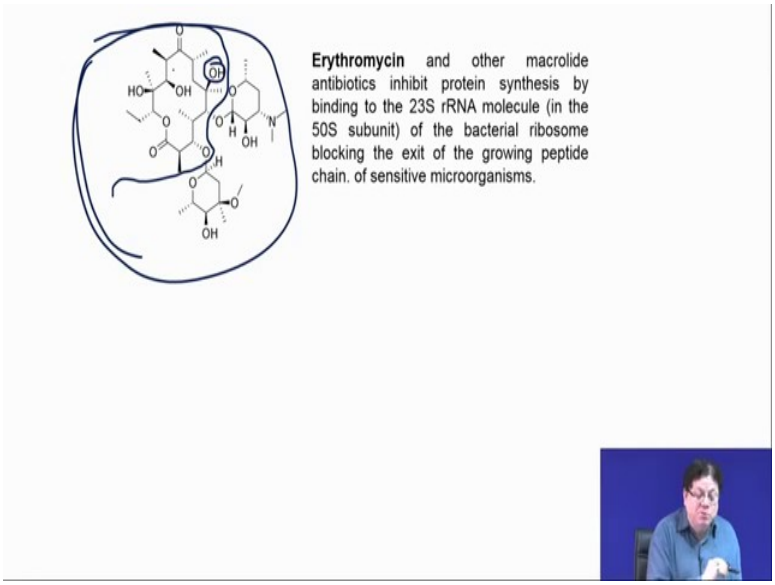
(Refer Slide Time: 06:22)



I will show you the structure of erythromycin, in some slide it is there. This is one of the precursors of erythromycin, called 6-deoxyerythronolide B (6-dEB). Like isopenicillin N is the metabolite that is formed first and then that α-amino adipoyl side chain is stripped off. So, you get 6-APN.

So, the primary metabolite from penicillin biosynthesis is isopenicillin N. And the primary metabolite for erythromycin biosynthesis is this one, then through post translational modification, it is modified a little bit, putting sugar moieties; that means, form glycosides; actually erythromycin is a glycoside; I think I have the structure somewhere.

(Refer Slide Time: 07:24)




See this is erythromycin, and you see the structure that I showed you is basically almost this one except lacking the OH group. So, this is what is called 6-deoxyerythronolide B, let us again go quickly to that structure. You see only the OH is missing here. So, that is called the deoxyerythronolide B. Now, this belongs to the class of polyketide. Now, polyketide chemistry or the biosynthesis of polyketides basically involves the repeating of the same type of reactions over and over again.

Initially it may look very quiet difficult, but the course progresses, you will see this is not difficult at all and you can always do retro biosynthesis. See if I give a natural product you can always break it down and that is much easier for polyketide to break it down to

the starting materials and you can then draw out how many enzymes are needed to make that polyketide.

So, I will progress little bit slowly, we now know that polyketides are either this type of carbonyl system or it could be hydroxy systems which are derived from the tricarbonyl or even it could be a double bond here like this, because the OH can undergo dehydration. So, basically if you see the structure actually these are polyketides because it has the skeleton which is derived from this tricarbonyl system.

Now, let us take a very simple compound to start with. We start with suppose this 6-methyl salicylic acid that is a natural product. 6-methyl salicylic acid is is one of the simplest polyketides. Now what is the structure? It is written here, but let me write again; there is a methyl, there is methyl and then there is this there is a methyl there is this acid and there is a OH here this is 6 methyl.

So, that is 6-methyl salicylic acid. Now, apparently you might not be able to figure out that how this belongs to a polyketide or what is the mechanism by which this molecule is formed. Now again I write that tricarbonyl compound in a different fashion. Suppose at some point of time this is made, now tricarbonyl compounds may cyclize; see they can undergo aldol condensation. Suppose this carbon attacks here, forms a OH and then the OH dehydrates.

So, if that happens what you will get? A carbonyl here, here and here is a dehydration because there is some reaction that has happened here. But this will now immediately tautomerize and since the aromatic system is more stable so this will tautomerize and form this aromatic ring. So, that means, the interesting point to note that when the tricarbonyl or tetracarbonyl, whatever be the case, when the polyketide chain is being made it can undergo cyclisation and form the aromatic rings. And in many of the polyketides, you will see generation of aromatic ring or generation of macrocyclic compounds like in case erythromycin.

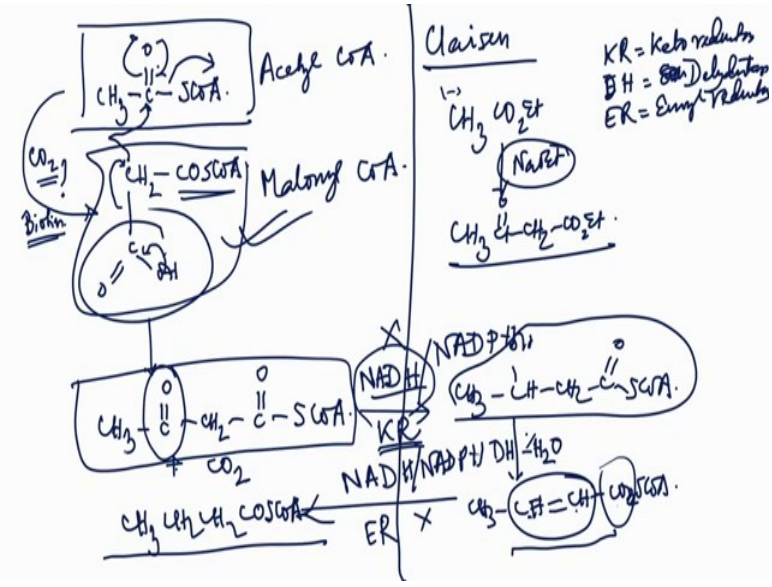
Now, 6-methyl salicylic acid is one polyketide, I will tell you how you arrive at that, but apart from 6-methyl salicylic acid, you have another compound which called orsellinic acid, Here there is another OH at the 4 position. So, earlier this was 6-methyl salicylic acid, now orsellinic acid.

So, in that compound, if you try to figure out the position of the carbon-oxygen bond; what you see there is a carbon-oxygen bond here because this is $CO_2H$, one carbon-oxygen bond and then there is a carbon then there is another carbon which has got a carbon-oxygen bond. So that means, you have basically a $COCH_2CO$ unit. CO unit was there which was somehow transformed into the OH, and then you proceed further, you have again another carbon-oxygen bond.

So, between all this carbon-oxygen bonds, there is a carbon there which is present. Now, the question is how do you know that there is a carbon here oxygen here? You have to know that or you have to learn little bit extra. Means what I have said that the polyketides are basically repeating units of carbon, if this carbon oxygen could be in the form of carbonyl or in the form of OH or could be in the form of double bond also, because earlier I showed that it could be in the form of double bond also.

So, double bonds are also suspicious functionalities which could be a carbonyl and then by cyclization chemistry, it became a double bond like I showed here. Now, ultimately what happened that there is a OH here then a carbon then another oxygen carbon oxygen, then another carbon. Now, the question is how this ring formation has taken place? So when there is a ring, you have to be careful that one of the carbon may be containing oxygen. Now, to make it a little bit more clear, we have to proceed a little bit further.

(Refer Slide Time: 15:06)



See in a polyketide, what happens? You have already been told the formation of acetyl coenzyme A which is a TPP catalyzed reaction, pyruvate to acetyl coenzyme A. Now you see the importance of acetyl coenzyme A. So, acetyl coenzyme A, during the formation of polyketide, this one of the starting points. Since it contains more number of carbon atoms so it needs another component so that there is a reaction between the two.

What is the other component? Usually the other component is $CH_2COSCoA$ and C double bond O OH; what is that? That is this is called acetyl CoA, and this is called malonyl CoA; because in malonic acid, only half of the carboxylic acid is converted into CoA. So, this is called malonyl CoA. Interestingly this reaction, generation of malonyl CoA is actually from acetyl CoA *via* a carboxylation reaction and this is, this requires biotin as a cofactor.

I leave the mechanism of this reaction for the students who are attending this course. So, this comes from the carbon dioxide; bicarbonate basically. Now these two are the two units which react together and form $CH_3COCH_2COSCoA$ plus carbon dioxide. So, what is the reaction here? That means, this goes here first, that forms an anion here and that anion reacts with this carbonyl, that goes here back and this goes out.

So, what it makes $CH_3COCH_2COSCoA$. Now, remember there is a reaction in organic chemistry called Claisen condensation, where you take ethyl acetate and add sodium ethoxide. What did you get? Ethyl acetoacetate; I think those who have studied organic chemistry, they remember it. So, what is the mechanism? One molecule forms anion and attacks the other molecule. Now nature's way of making the anion here is not using sodium ethoxide; nature does not have sodium ethoxide; in fact, this is unstable compound.

Any moisture will kill the sodium ethoxide and this reactions after all is taking place in water medium. So, nature's way of making the anion is *via* decarboxylation. So, malonyl CoA undergoes decarboxylation and forms this aceto acetyl coenzyme A, $CH_3COCH_2COSCoA$. Now usually what happens that there is a reduction step, $CH_3C(OH)HCH_2COSCoA$.

So, its reduction is carried out by NADP or NADH; that we have already learnt that the reductions are taking place by NADH or NADPH, so you get this alcohol. Then you get a dehydration step eliminating $H_2O$. So, that will form what? $CH_3CH$ double bond CH

COSCoA and then you get another reduction mediated by NADH or NADPH. And what you will get? You will get the saturation of the double bond.

This is exactly what is followed in case of fatty acid biosynthesis. In fatty acid biosynthesis, you start from acetyl coenzyme a that undergoes Claisen condensation with malonyl coenzyme A and then that gives acetoacetyl coenzyme A and then after these reductions, first NADH then dehydration then another reduction; so, you get $CH_3CH_2CH_2COSCoA$.

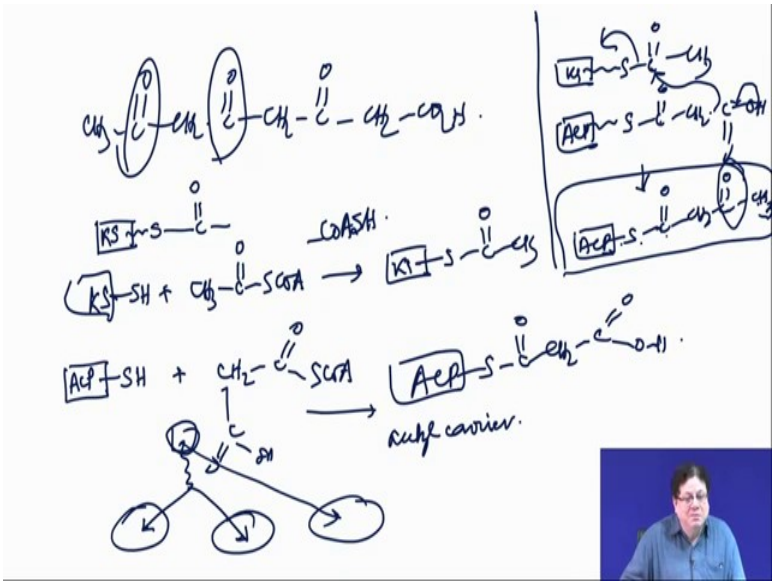
So, this is a butyryl coenzyme A so, you got the butyric acid. In fatty acid biosynthesis what happens? This cycle continues; is repeated over and over again. So, then this one reacts with another malonyl CoA and you will get hexanoyl CoA finally. And then hexanoyl CoA goes to C8 then C8 goes to C10, C12, C14, C16; palmitic acid is C16.

So, then it falls off from the enzyme, the thioester goes off and becomes the acid. So, that is palmitic acid. Now, so, there is a great similarity between fatty acid biosynthesis as well as the polyketide biosynthesis. In the sense that here the first reaction is the same reaction that happens in polyketide also. Actually this is the Claisen condensation after that this is a reduction step.

So, what is reduced? A keto carbonyl. So, the enzyme which does that will be called KR. What is KR? Ketoreductase; then this enzyme which causes dehydration will be called dehydratase DH and then the enzyme which does this reduction of the double bond is called ER. What is ER? Enoylreductase. So, you have KR that is ketoreductase; you have DH, that is dehydratase; and then you have ER that is Enoylreductase. Why enoyl reductase? Because that functionalities is an enoyl functionality; ene and then the acid enoyl, it is not acid; they are like acetyl.

So, this is enoyl reductase. So, these enzymes are there. Now in many of the polyketides, these steps will be used over and over again. In polyketides, some of these enzymes may not be present. See if you want to make a tricarbonyl compound then these type of enzymes should be absent in the system, because you do not want any reduction of the carbonyl.

(Refer Slide Time: 23:14)




If you want a compound which is say $CH_3COCH_2COCH_2CO$ and then $CH_2CO_2H$, if you want to make a compound like this your enzyme system that you are using should lack all these ketoreductase, dehydratase and enoylreductase because they are only carbonyls. And carbonyl group is found only in the first Claisen condensation.

So, they have to be processed without touching those other 3 enzymes. Now let us talk about the first reaction that is the Claisen condensation. How it happens? Basically there is an enzyme called ketosynthase which ends up as a thiol. So, this takes up the the acetyl; keto synthase has a thiol. So, ketosynthase SH plus acetyl coenzyme A, there is thiol exchange now. So, the ketosynthase sulfur holds the acetyl group and coenzyme A leaves.

And that is another enzyme protein which is called ACP (Acyl Carrier Protein). What it has? It has got an SH; That takes up the malonyl CoA; this is malonyl CoA. So, then ACP-S-CO-$CH_2$ then $CO_2H$; so, basically before the Claisen condensation takes place, first this has to happen a ketosynthase enzyme holds the acetyl group and an ACP acyl carrier protein holds the malonyl group.

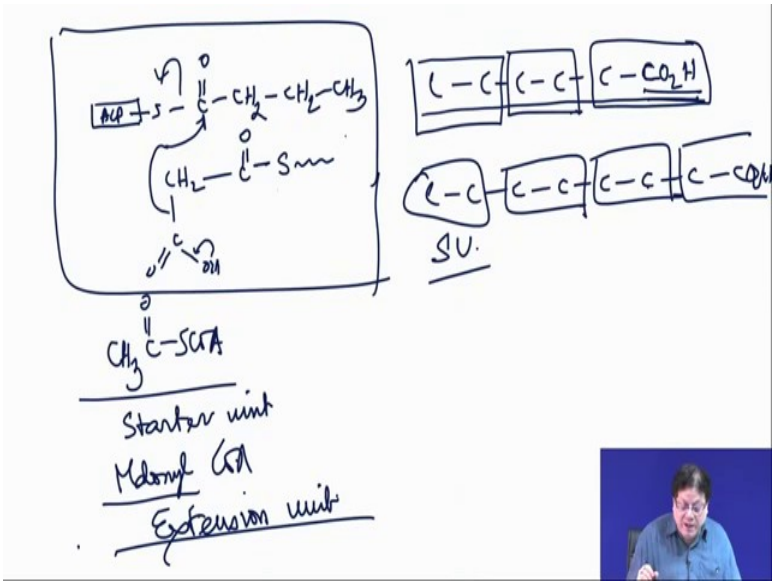
That means, what we are talking about is a ketosynthase here which takes the S-CO-$CH_3$ and an ACP side by side and that has got S-CO-$CH_2$-$CO_2H$. Now the reaction takes place that undergoes decarboxylation. So, that attacks here, the ketosynthase is free. So, what will happen? The ACP is now holding the whole thing, the acetoacetyl coenzyme A; CO,

not CoA; ACP; acetoacetyl ACP; because CoA has already left; $COCH_3$ this is the compound. So, $CH_3CO$ then $CH_2$ then COS-ACP. So, S-ACP is holding this whole four carbon chain; it shows it to keto reductase if it is present, then this carbonyl will be reduced; if there is the dehydratase enzyme then it will show to the dehydratase.

But all the time it is attached to the ACP; that means, you have several enzyme systems ketoreductase then dehydratase and then enoylreductase; suppose all are present then what will happen? The ACP is holding this chain. So, there is a conformational change all the time; if there is the ketoreductase, it puts the whole thing into the active site of the ketoreductase, so that reduces the carbonyl; then this is taken out.

And then this is shown to the dehydratase and it is put into the active site of the dehydratase; then the dehydration takes place and then if there is the enoylreductase, it will put to the enoyl reductase. If all the 3 are present, then this CO will be saturated, it will be $CH_3CH_2CH_2COS$-ACP.

(Refer Slide Time: 27:55)




If all that ketoreductase, dehydratase and enoylreductase are present, then we will end up with this butyryl coenzyme A, this was the $CH_2$ which was earlier CO.

Now, this has to be repeated. So, basically the reaction that takes place; you remember ACP always holds the malonyl group. Earlier in that reaction ACP was holding the malonyl group and the ketosynthase was holding the acetyl group. In the next reaction,

you again have another malonyl system. Now sorry CO CO O do not C O sorry this is; let's see.

So, that will be CO and then S. Now, we will talk about this whether it is ACP or whether this is keto synthase; we will talk about that little later; but a reaction again has to take place in order to extend that chain. So, now, you have extended the chain to 6 carbon, $CH_3CH_2CH_2COCH_2COS$ and then whether it is ACP or ketosynthase, will come to that point little later.

But this is the chemistry that takes place; that means, now I can figure out that acetyl CoA has been used how many times to make this 6 carbon chain? Acetyl CoA was used only one time, the first reaction and malonyl CoA was used the first reaction then the second reaction then the third reaction and the fourth reaction like.

So, acetyl CoA is called the starter unit and malonyl CoA is called the extension unit; because your 2 carbons are added from the malonyl system. So, if that is clear then we can slowly move into the earlier problem again where we got little bit stuck. See this was the orsellinic acid. So, that compound, remember in the how many carbons are there? Let us count; 8 are there. Now, I could see that there is oxygen carbon oxygen here then a carbon then there is there is this methyl.

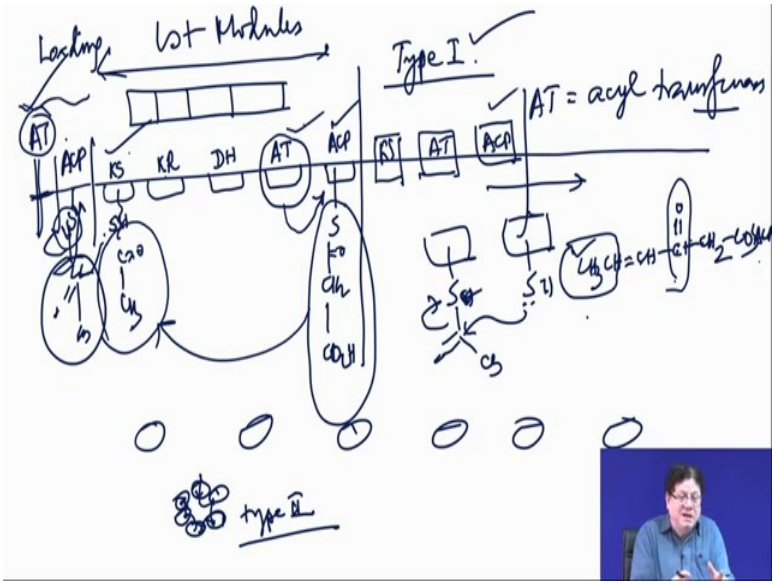
Let us again go back to that slide where I explained everything, see at the end of the reaction you end up with a terminal $CO_2H$; like if it is palmitic acid, that means you have to have 7 cycles of this reaction. So, that 7 multiplied with malonyl unit means 2 carbon units, that is fourteen plus acetyl CoA that is 2; that means, C16 is generated and then if it falls off from the enzyme system, you will get at end the $CO_2H$.

Now, if I say that how many malonyl units are present here? So, what you do you start from this from this $CO_2H$. So, this is one malonyl unit, then this is another malonyl unit and this is your starter unit. All cannot be malonyl unit; you can start from this side, first two carbons is your starter unit then the second two pair of carbons is coming from the extension unit. So, that is malonyl, this is also malonyl. So, if I have 8 carbons now, it is suppose $CO_2H$. So, this is one malonyl unit that is one malonyl that is one malonyl and this is the starter unit so; that means, malonyl CoA has reacted 3 times here.

So, again we go back quickly to that compound, I want to dissect it into a polyketide; I can start from the endpoint, see the $CO_2H$ is here, then there is a carbon then $CO_2H$ then there is a carbon then I go this way because there is oxygen here. I know that the two oxygens (here this one and that one) are in 1, 3 relationship. So, I do not go that way; what I do is $CO_2H$ then the carbon then there is this carbon which must be having a carbonyl or a derivative of the carbonyl then a $CH_2$, then another carbonyl, then $CH_2$ then another carbonyl and then methyl.

So, if you dissect it, then you should obtain a system like this, $CO_2H$ $CH_2$ then CO then a $CH_2$ then CO then a $CH_2$ then CO. So, now, what is the starting point? The starter unit is acetyl coenzyme A, then there are how many malonyl units here? What is the extension unit? There are 3 cycles of reaction that has taken place, that Claisen condensation followed by other enzymatic reactions.

(Refer Slide Time: 35:37)




See usually this is the picture of the protein, this polyketide can be synthesized or that is it has been found that there are basically 3 types of polyketides. But we will concentrate on only two types; one is called type I. What is type I? Type I is that you have a big protein, and then in the big protein you have domains of different activities in a single protein. Like suppose this is a ketosynthase, it has got a ketosynthase activity; suppose this is ketoreductase. Suppose this is a dehydratase; I will talk about this what is AT, I will come back to AT and then ACP.

So, basically this is called type I polyketide; in type I, you have a big protein where different domains of activities are there. And the type II is where you have different proteins like ketosynthase, ketoreductase; and at the time of biosynthesis, they all come together and then have a cooperative effect between them. So, first one reaction takes place here, that transfers it to here, then that transfers it to here, that transfers in a cyclic manner.

So, this is type II. So, type II is where different proteins assemble together at the time of biosynthesis. Type I is a single protein containing different functionalities at different domains. So, now, let me see which way we should go. The first reaction I told is basically between a ketosynthase holding an acetyl group and an ACP which is holding the malonyl group.

Now, the question is who is putting this acetyl group to the keto synthase? That is also done by another enzyme, see my hand is here, but now my left hand is attached to an acetyl group, but some enzyme has to come and then bring it alongside and then put it on my left hand. And on the right hand suppose this is ACP, so again the same problem here, this right hand is just sitting here, so another enzyme will bring.

So, there must be some carrier protein; protein which brings these starter and the extension unit and attack it to ketosynthase and ACP; I hope that is clear; so that means, all these polyketide synthase in the type I; we are talking about type I cases and we are not talking about these separate enzymes; one protein having all the activities. So, it starts with like this, it starts with what is called an acyl-transfer. This is nothing, but you are transferring the acetyl group from acetyl coenzyme A to the ketosynthase; you are transferring the malonyl group, that is also an acyl group and then putting it into the ACP.

So, this is what is called at AT; AT's job is to do this transfer, acetyl group transferred to the ketosynthase; malonyl group transfer to the ACP. So, AT is acyltransferase. So, initially there is an acyltransferase here. All polyketides synthesis starts with this acyltransferase because that takes up the acetyl group; acetyl coenzyme A and there is an ACP also here. So, first this AT puts the acetyl group on to the ACP, remember ACP has an SH.

So, first acetyl transfer takes place to the ACP; who does that? The AT. Now this is what is called the loading module because this is like a primer kind of thing. See loading module is that where you start the whole thing. So, the acyltransferase first takes the acetyl coenzyme A, gives it to the ACP, by a *trans*-thiolation reaction. The acetyl group is now in the ACP; next the ketosynthase has SH.

The SH is holding the $COCH_3$, now the acetyl group which was at the ACP that will be transferred to the ketosynthase. This is a little bit complicated, but once you know this that becomes easy I am sure.

So, there is a loading module, the AT transfers the acetyl group to the ACP; ACP also belongs to the loading module. Now the ACP transfers the acetyl group to the ketosynthase. So, basically it is a two-step process; directly it does not go to the ketosynthase, it was first given to the ACP; ACP shows it to the ketosynthase; remember ACP has a SH and ketosynthase has SH.

ACP first holds the acetyl group and then shows it to the ketosynthase; now ketosynthase attacks this acetyl and frees the ACP and thereby it is holding the acetyl group. So, first acyl-transfer is done by AT and then ACP to ketosynthase; so, it ends there, the loading, your starter unit is now being held by ketosynthase. Now there is another acyltransferase here, it puts the malonyl group into the ACP. Remember this ACP is different from that, this belongs to the loading module and this is what is called the first module, there maybe several modules and this is the first module.

So, ketosynthase has the acetyl group, ACP has the malonyl group. Who transfers the acetyl group to the ketosynthase? It is the acyltransferase in the loading module from the ACP which does this. So, now, there will be reaction between these two, that Claisen condensation. So, what will happen? ACP will hold the $S\text{-}COCH_2COCH_3$ after the reaction. Now because there is a ketoreductase, it will show the whole chain to the ketoreductase.

Ketoreductase now reduces the carbonyl; there is a dehydratase. So, the dehydration will take place, but there is no enoylreductase, if that is not there; that means, after this reaction you get the double bond compound. Now, what will happen? All the time this is being held in the ACP. In the second module, when it is the next cycle of reactions takes

place; there is a second module, it is not repeated so that this is not again thrown back to the ketosynthase; it is a modular process.

So, first module has this type of framework, this type of domains. So, once that is done then it is transferred to the ketosynthase of the next module. And suppose the next module has only ketosynthase acyltransferase and ACP, then what will happen? If I ask what the product at this end is, then what will be the product? I can write it $CH_3$ then CH double bond CH then CO then $CH_2$ then CO S ACP. How do I know this?

Because this is our starter unit; every module should have ketosynthase and acyltransferase and ACP that is mandatory. In the first module I see that there is a ketoreductase and dehydratase; that means, after the first reaction a double bond is formed. In the second module, I do not see any ketoreductase or dehydratase or anything. So, that ends up as the carbonyl and then $CH_2$-S-CO-S-ACP.

So, basically then we will end up by showing this. This is the deoxyerythronolide B. Now what is my starting point? The end point is a carboxy and the starting point could be acetyl, but do not take it sacrosanct that always the starting unit is acetyl, it could be propionyl, it could be butyryl, all these are possible. Similarly the extension unit can be malonyl or it could be methyl malonyl, there are different types of extension units that are possible.

Now, this 6-deoxyerythronolide; I see there is a lactone moiety; first check where is that acid group, because it has to end at the acid. So, I see that there is a lactone. If you hydrolyze this, what will happen? If you hydrolyze this that will become OH and this becomes $CO_2H$. Now you work backwards. So, this is your one extension unit, that is your another extension unit, this is your third extension unit. Actually we are going in the reverse direction, this is the fourth extension unit, this is the fifth extension unit, this is the sixth extension unit.

So, there are 6 extension units and now this is your starter unit. Let me write the structure again. So, here there is this bond what I actually deleted, because there is a carbonyl here. So, what happens? Maybe I put the OH here now and $CO_2H$ here, and there is ethyl; forget about the stereochemistry, there is ethyl here there is methyl, after every CH there is the methyl.

Let us see where the oxygens are, one oxygen is here OH, this is a carbonyl, this is $CH_2$, then there is a OH here. Again I repeat that we are not thinking about the stereochemistry here and then there is a OH here. Now, I can write what are the enzyme components in different modules. First of all how many modules it should have? That is number 1, apart from the loading module; loading module is separate; there should be 6 modules, because 6 extension units are here.

What are those 6 extension units, one is this one; this is your starter unit, this is the extension unit. Alongside are your extension units. So, 6 extension units; that means, you have to do 6 Claisen condensation reactions. So, there are 6 modules in this. So, it started all from here; this is the starter. So, what is the starter unit? It is not acetyl CoA, it is propionyl CoA because you have a ethyl group here. So, this is your starter unit. And what is your extension unit? It is not malonyl CoA, but it is a methyl malonyl CoA, there is a methyl group and then $CO_2H$.

Because everywhere you have a methyl extra, if this methyl was not there then malonyl CoA would be the compound, but it is methyl malonyl CoA. So, now, let us see. The first Claisen condensation has put a OH group here, but no dehydration, no saturation. So, in the first module, I can say what will be there; ketosynthase, acyltransferase and ACP are mandatory. Then we have to see what is present, ketoreductase, dehydratase or enoylreductase. Because, it lands up in OH so there is a ketoreductase. It does not matter what is the order of this, you can write anything and order is not important here, we are only trying to understand.

So, ketosynthase, acyltransferase has to be there, ACP has to be there and then ketoreductase so it ends up with OH only. Second one, again it ends up with OH. So, the second module again same ketosynthase, acyltransferase, ketoreductase, ACP. Third module it ends up at a carbonyl, so that means, there is no ketoreductase also.

So, third module ketosynthase, acyltransferase, ACP; fourth module I see there is $CH_2$, no carbonyl; that means, here domain activities are present. That means, it will have ketosynthase, acyltransferase, dehydratase, enoylreductase, ketoreductase and ACP. Again I repeat you can have any order, the most important is you should identify that domain activities in each module. Then the module, so this is the 4th module, then 5th module you end up as the OH; that means, it should have ketoreductase, ketosynthase,

acyltransferase, ketoreductase, ACP and final this is the other one; the last module. The last module has what? Last module again ends with OH.

So, it will have ketosynthase, acyltransferase, ketoreductase, ACP. Now, finally you have to release it. How do you release it? That is called the thioesterase, because after all your ACP is holding through S-CO and then the whole chain. So, now, you have to have an esterase enzyme, but that is the thioesterase enzyme which hydrolyzes this one. So, to release, always it uses thioesterase and for loading, always uses AT and ACP; see now it is more or less understandable I believe. If you can identify all these, where are the oxygens, and then try to find out the starter and the extension unit, you can tell what the modules are. So, next session we will start from here. See now we have these modules and there are scientists, like there is one Indian scientist, his name is Chetan Khosla, and he has many YouTube videos about polyketide synthase specially erythromycin biosynthesis. What he has done? Suppose I take this thioesterase domain and put it after module 2, then what will happen?

Module one reactions will take place, module two reactions will take place then the molecule will be released and then I can get a new type of compounds because sometimes this thioesterase can actually form macro-lactonization. Means like in this case erythromycin, this last OH actually is attacking this carbonyl and releasing the sulfur, but this is an enzyme catalyzed reaction. So, it is a macrocyclization; one OH and one carboxy, they reacted with each other forming the macrocycle.

So, that is the biosynthesis of erythromycin. So, now, you see you have the genetic map. So, you can cut up different domains, suppose I want an erythromycin where it is not carbonyl it is only $CH_2$. So, what I have to do? I have to add domains; I have to add a ketoreductase, a dehydratase and enoylreductase then I will get a $CH_2$ here. So, now, this is the trend of today that how to manipulate the gene especially after this is known that these are modular enzymes.

And then you can make new types of secondary metabolites and some of them may be very useful. In fact, different types of truncated erythromycin have been prepared by the same person, Chetan Khosla; he is a scientist at Stanford University. So, we will work out some problems in our next session.

Thank you.