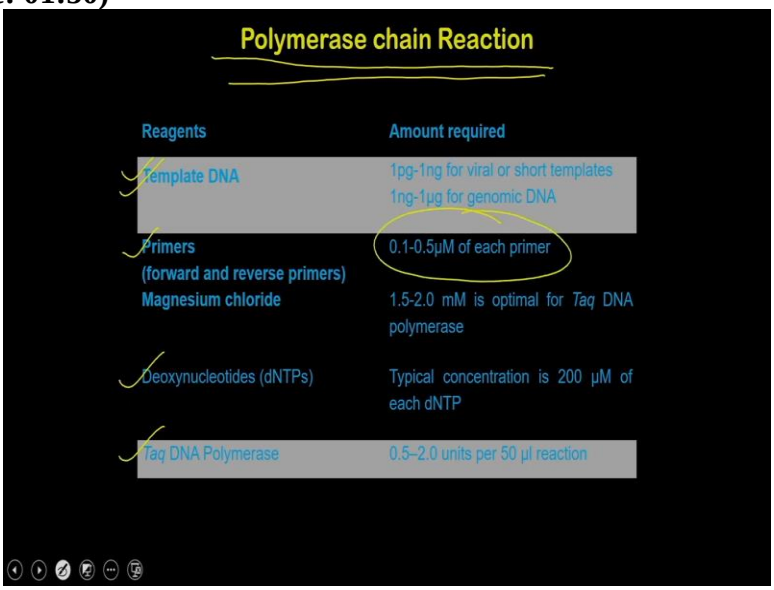
Experimental Biotechnology
Prof. Vishal Trivedi
Department of Bioscience and Bioengineering
Indian Institute of Technology, Guwahati

Lecture – 40
Polymerase Chain Reaction, Part - 3

Hello everybody this is Dr. Vishal Trivedi from department of bioscience and bioengineering IIT Guwahati. And in the previous lecture we have discussing about PCR. So for what we have discussed? We have discussed about different aspects related to PCR how the PCR is being developed and how you can be able to perform the polymerase chain reactions and in the previous lecture. We have also discussed about some of the practical aspect where we have shown you how to set up the reactions, how to you know make the reaction mixtures.

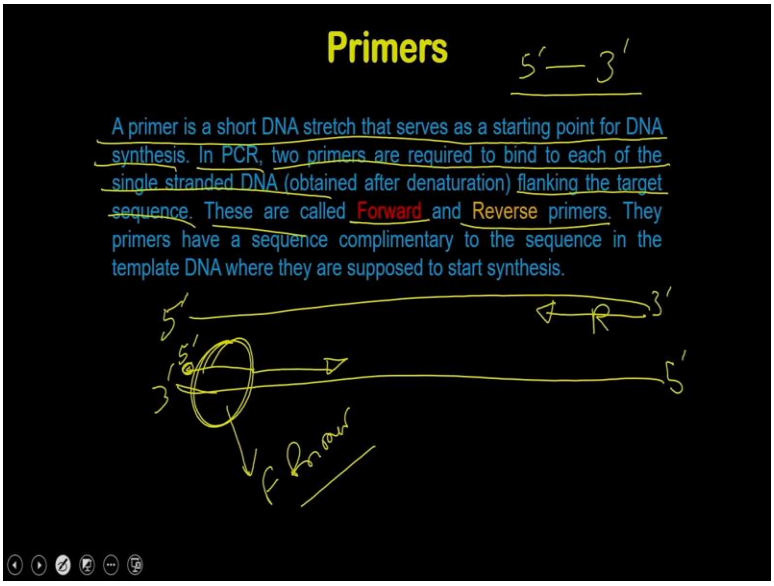
And how to perform the PCR? How you can analyze the PCR in the agro gel electrophoresis. So now in today is lecture we are going to continue with our discussion about the PCR. So let us discuss about the PCR.

(Refer Slide Time: 01:50)



So as you remember that in our previous lecture we will discussing about the setting up the PCR reactions and in a PCR reactions you have the few of these components like you have the template DNA then you have the primers you have the dNTPs and then you have the taq DNA polymerase. So out of these components the primers which are the very, very important component and for every target gene you have to synthesize a set of primers.

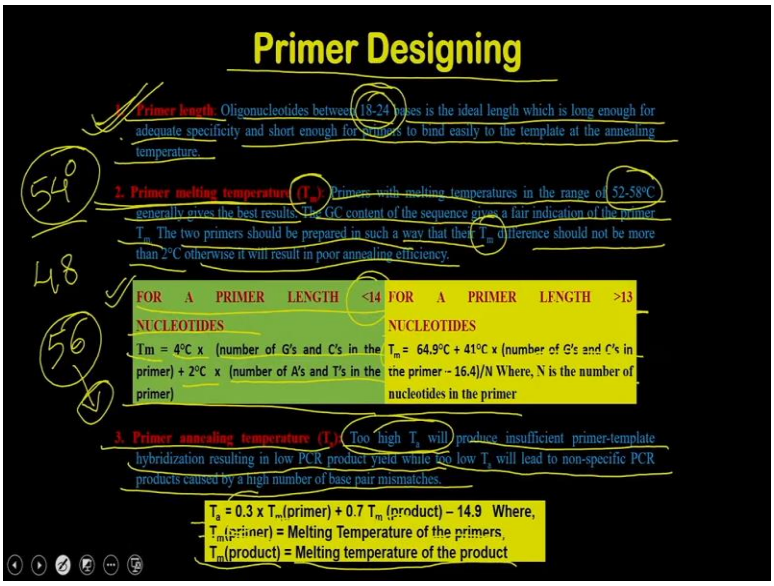
(Refer Slide Time: 02:25)



So the primer is a short DNA stretch that serves as a starting point for DNA synthesis in PCR you require the two primers to bind on the each single stranded DNA flanking the target sequence these are called as the forward as well as reverse primers. So if you have a gene and you know that the gene is always been notified as 5 prime to 3 prime and 3 prime to 5 prime, so the primer which actually is going to bind to this side because if you remember the DNA synthesis is always occurring in the direction of 5 prime to 3 prime.

So what will happen is the 5 prime of end of the primer is actually going to bind to the 3 prime end of the target gene. And that is how this will continue in this direction. So this primer is called as the forward primer whereas the primer which will bind on to the reverse strand is called as the reverse primer. So this is called as reverse primer, this is called as the forward primers. So now the question comes how you can be able to design the forward as well as the reverse primers.

(Refer Slide Time: 03:37)



So before you design the primers you have to consider many parameters. And considering those parameters only you can be able to design the primers because when you design the primer the purpose is that it should actually going to bind to the target DNA. And that complex is going to be recognized by the taq DNA polymerase and then it is actually going to initiate the elongation steps and that is why it is actually going to give you the complete amplifications.

So when you synthesize the primers or when you start designing the primers first thing what you have to remember is about a PCR primer length. So the oligonucleotides between 18 to 24 bases is the ideal length which is long enough for the adequate specificity. And the short enough for primer to bind easily to the template at the annealing temperature to the primer length of very short primers like if you have a 12 to 16 base pair primers.

It is actually going to lack the specificity because there is a probability that if you have a very small stretch of DNA it is actually could find multiple regions within the target DNA or if you are suppose using the genomic DNA properly there could be a possibility that the this particular primer could bind to multiple locations and that is how it is actually going to give you the not a desirable product but it is actually going to give you the non specific products or it may not give you the amplification at all.

Whereas if you give very long stretch of DNA for a stretch of primers like if you take a primer of 40 to 45 nucleotide long that will have a problem of secondary structures first of all because it is a very, very long so it is actually going to form the secondary structures and that is how it is actually not going to bind to the target DNA. Second that also has a problem that it actually may because a small portion of that primer may bind to the multiple places and it is actually going to give you the no amplification or the undesirable applications.

And that is why 18 to 24 base pair is a ideal length where you can have the primer length. The third point is if you are using a very long primers it is eventually going to end up in being very costly because the primer synthesis is always been done by the companies and that is how it actually may end up in you know you might be end up in spending a lot of money but at the end it is not going to serve the purpose.

Because the what you can get from 18 to 24 exactly the same or even worse probably would be possible if you are using the 40 or 45 base pair nucleotide long primers. Then you have to also consider the T m of the primer. So primer melting temperatures for primers with the melting temperature in the range of 52 - 58 degrees Celsius generally gives the best results the GC content of the sequence gives a fair indication of the primer T m.

The two primers should be paired in such a way that their T m difference should not be more than 2 degrees. Otherwise it will result in the poor annealing efficiency how you can be able to calculate the T m of a primer. So if you have a primer which is of less than 14 nucleotide then what you can do is you can calculate the T m simply by calculating the number of G and C and the number of A and T.

So T m is equivalent to 4 degree multiplied by number of G and C in the primer plus 2 degrees multiplied by the number of A and T in the primers. Now primer melting temperature is a very very important criteria because primer melting temperature is only going to decide at what temperature you are actually going to allow the annealing of these temperature and that is how the 52 - 58 degrees Celsius is been found to be ideal.

If you take very high temperature then it will actually having a problem in terms of annealing. And if you take very low number then it is actually going to anneal at multiple places in nonspecific sites. So that so it is actually going to give you the nonspecific multiplications. The second point is that the primer melting temperature or the T m of a primer between the two primer like the forward primer and reverse primers.

Should not have a difference of more than 2 degree because what will happen is if I have set it at primer anneal temperature of 54 degree Celsius and one of the primer has a melting temperature of 48, the other one is having a 56. So what will happen is that a 56 is going to anneal or because we are setting up at 54 degrees Celsius. So the 56 wall once will go and bind to the target DNA it will anneal but 48 degrees Celsius annealing T m is because this temperature is still bigger than the this temperature.

So because of that it will not going to anneal so because of that there will be no amplification of your target DNA. So if you have a primer of less than 14 base pair you can use this sequence or this particular formula. But if you have a primer length which is more than 13 nucleotide then what you can do is you can use calculate the T m by 64.9 + 41 degree Celsius number of GCs in the primer - 16.4 / N where N is the number of nucleotide in the primers,

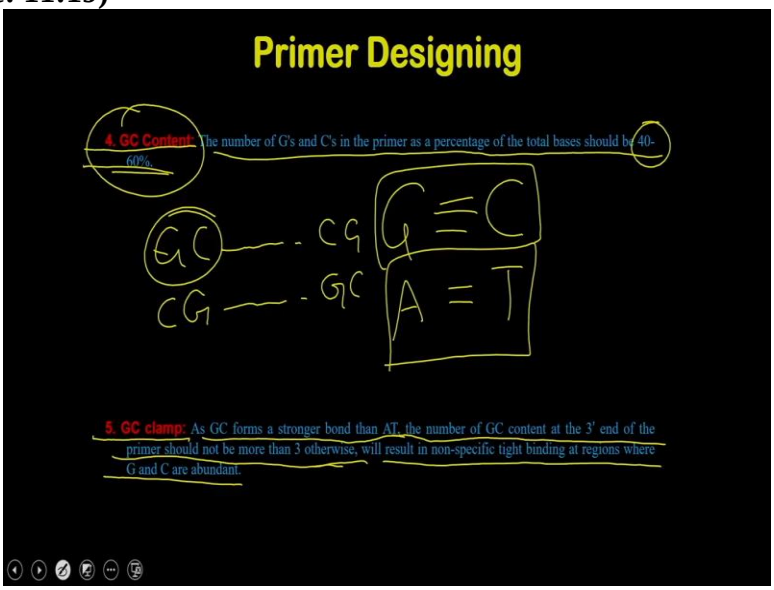
Apart from that you also have to consider so either of these formula you can use depending on the length of the primers to calculate the T m of the your primer and that value is actually going to be close between the forward as well as the reverse primer. So that is why you have to design the primer in such a way that it should have a very close T m value so that you can be able to very comfortably use the annealing temperature.

So that it to both the primers will actually go in and into the target DNA. Apart from that you also require the primer annealing temperatures so primer annealing temperature too high primer annealing temperatures provide insufficient primer template hybridization resulting in a low PCR yield whereas too low time primer leading temperature will give you non specific products caused by a high number of base pair temperatures.

And the T m is indirectly or directly been attached to the T m T a values or the primer annealing temperature for example how you can be able to calculate a primer annealing temperature 0.3 into T m the primer + 0.7 T m of the product -14.9 where the T m primer is the melting temperature of the primers that is the both primers. And the T m product is the melting temperature of the product.

So this primer annealing temperature is going to be deduced once you calculate the primer melting temperature and you can be able to calculate the melting temperature of the template as well because if you give the template sequence into the software it is actually going to calculate the template melting temperatures and the primer melting temperature as well. And that formula you can use put it into this the primer annealing temperature and that is how it is actually going to give you the annealing temperature. And that is the annealing temperature you can put it into the software into the into the PCR machine when you are setting up the PCR reactions.

(Refer Slide Time: 11:19)



Apart from that you also have to consider the GC content because primer melting temperature as well as the annealing temperature is depending on to the number of G and C present in the primers. And that is why the person GC content is also very important. The number of G and C in the primer as a percentage of the total basis should be between the 40 to 60 base pair 60% which means you should have a very high amount of G and C content because if that actually provides the binding of the primer to the template.

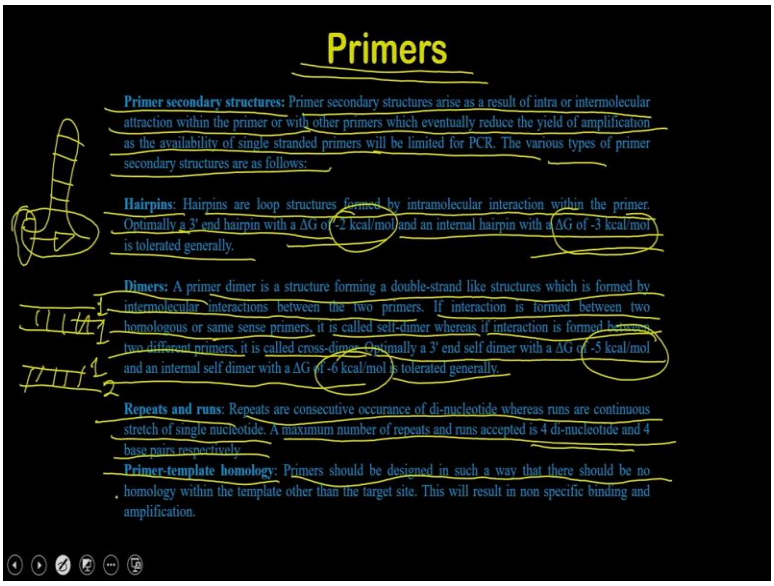
Because that actually is aligning very strong binding of the primer to the template and that how it actually provides the sufficient time for the taq DNA polymerase to sit onto this primer template complex and that is how it actually smoothly can initiate the synthesis whereas if you the GC content is very low the primer template annealing is going to be very very weak because if you remember the G is always making 3 base pair with C whereas the A is always making 2 base pair with T.

So if you have more of A and more of the you know AT resequence like for example what will happen in some of the pathogenic organism like malaria you are actually the through the primer template hybridization is going to be very weak. So when the enzyme will sit and try to you know synthesize the DNA it could possible that during that process the primer may actually dislodge or the primer is binding to the template but the interact the hybridization is very weak.

Either of that thing it is actually eventually going to reduce the efficiency of the enzyme mediated polymerizations and ultimately it may end up in giving you either the nonspecific product or the lesser efficiency of the enzyme. Then you also require the GC clamps so as the GC forms a stronger bond than the AT the number of GC content add that 3 prime of the primer should not be more than 3 otherwise it will result in a nonspecific tight binding at region where G and C are abundant.

So GC clamp is very actually going to give you a strong binding than the AT and because the GC clamp is to GC clamp is required. What is mean by GC clamp is that if you have a sequence like this something like this so if so in the target DNA the sequence is going to be like this. So this is actually going to provide the binding of the primer into the template in a very, very strong way. And that is how it actually ensured that that the enzyme will sit and synthesize but if you have more of these then it is actually going to give you the problem because then the primer template interaction is going to be very very strong.

(Refer Slide Time: 14:18)



Apart from that when you are designing a primer you also have to consider the formation of the primary secondary structures. So primary secondary structures arise as a result of intra or intermolecular attraction between the primer or with the primer with the eventually reduce the yield of amplification as the availability of single stranded primer will be limited for PCR. The various types of the primer secondary structures are as follows.

The hairpins so hairpins are the loop structure formed by the intermolecular interaction within the primer optimally a 3 prime end hairpin with a delta G of 0.2 kilo calorie per mole and an internal hairpin with a free energy of -3 kilo calorie per mole is tolerated generally. So hairpins are like a double stranded DNA. So the when the primers have the internal symmetry or when they are actually having the symmetry they will come and close together.

And that how it is actually going to give you the hairpin like structures like one part of the DNA is actually going to fold. So this is a single stranded primer but it actually going to fold and it will give you a hairpin like structure. So if this will happen first of all it is actually going to reduce the single stranded length of the primer and so that it will be able to having the interaction with the target DNA or the template DNA.

The second is because it having a hairpin like structure. So if the enzyme is sitting here and trying to do the amplifications it is actually not going to pass this particular hairpin and as a

result the hairpin is actually going to dislodge the DNA that the DNA polymerase or it is not going to allow the DNA polymerase to go for in to the elongation phase, as a thumb rule if you have a hairpin loop which has a delta G of -2 or -3.

That is well tolerated because you are going to have the hairpin which is going to be formed when you are putting the restriction enzyme restriction sites into the primer and that restriction site is always palindrome. So that actually is going to give you the hairpin but that hairpin is good enough for going to be broken up by an when you are putting into a temperature where they are going to anneal but that temperature is good enough to break these kind of interactions.

Then you are going to have the dimers a primer dimer is a structure formed by a double stranded structure which is formed by the intermolecular interaction between the two primers if the interaction is formed between the 2 homologous or same sense primer it is called self dimer whereas if interaction is formed between the 2 different primers it is called as the cross dimer optimally a 3 prime with a delta G of 5 kilo calorie per mole or an internal self dimer 6 kilo calorie is tolerated generally.

So what is mean by dimer is that you have this is the primer number one and this is your primer number one. So if it is both the sequence are actually having the some kind of homology or symmetry then they will actually going to form a double stranded DNA which means they are no longer be available as a single stranded primers. And that is how the purpose of heading the primers into the reaction is going to be over right.

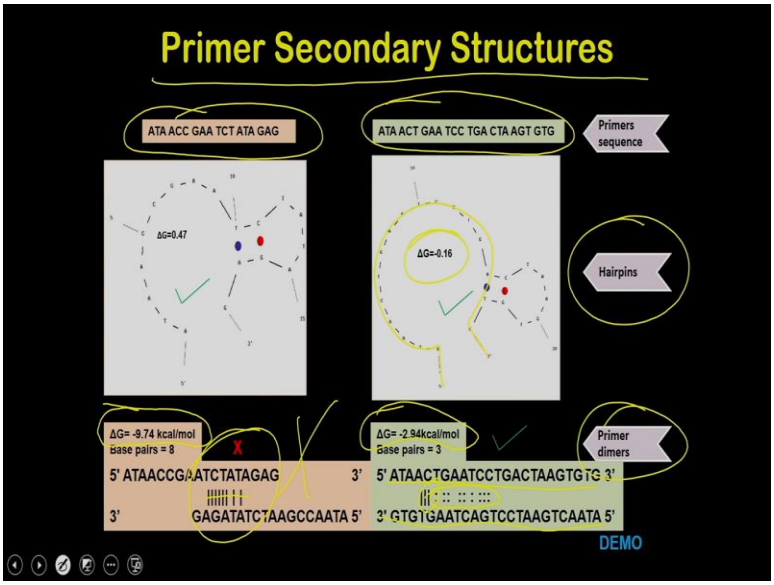
So this could be because of that or it could be like one is making a double stranded DNA with the second strand which means one either making a pair with one or one making a pair with 2 either of these things either it could be a homodimer or it could be a heterodimer whatever the case is actually going to sub sequester the single stranded primers but you do not have to worry about the dimers.

If the Delta G which is the free energy of this primer dimer is in the range of -5 to -6 because that is good enough that energy is going to be broken down once you keep this primers at a annealing

temperature because that is good enough because when you calculate the annealing temperature you are going to consider all these parameters. And that is how you will keep the annealing temperature in such a way that the primer will actual actually going to be still be able to anneal with the template but it will not going to be anneal with the next molecule of the primer to form the dimmers.

Then you are going to have the repeats and the runs for repeats are constitutive occurrence of the di-nucleotide whereas runs are the continuous stretch a single stranded nucleotide a maximum number of repeats and run accepted is a 4 di-nucleotides or the 4 base pair respectively then you have the primer template homology for the primer should be designed in such a way that there should be no homology within the template other than the target site this will results in a nonspecific binding and amplifications.

(Refer Slide Time: 19:18)



So these are the some of the primer secondary structures what you can see these are the so if you have these particular type of sequence like this is a 2 primers and you can have the hairpin loops. So you can have the like hairpin is been formed and this but you can see that this hairpin has a delta G of 0.16 and that is good enough when you are actually going to hit up and when you are bringing this parameter to annealing temperatures.

Then this is an example of the primer dimer where one sequence of the primer is making an interaction with the second sequence whereas what you see is this is actually a non acceptable primer dimers like in this case what you see is that only this region where only the 3 nucleotides are making an interaction with the nucleotides present into the other sequence and then rest all these sequences are very weak interaction.

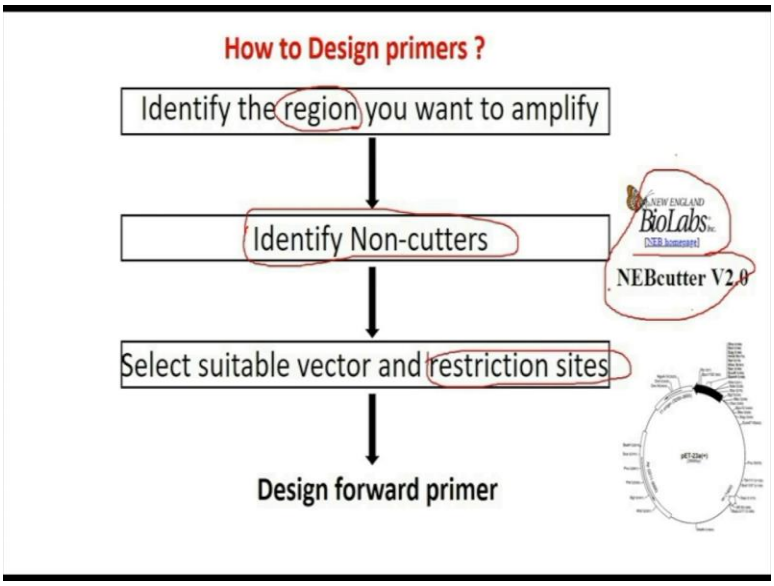
So this is actually going to be broken down because you have a delta G which is -2.9 whereas what you see here is actually a very strong binding of involving almost 6 to 7 nucleotides and what you see is the delta G is very high which is like -9.74. So that is why this primer when you are synthesizing this primer is not going to be useful because it is going to be ended up in giving you the primer dimmers.

And that is how it is actually going to you know it and this particular free energy is very high. So it is actually going to give you the nonspecific interactions or this primer will not give you the amplification at all. So this is a just a brief overview about how to design the primers let so to explain you how to utilize the software's because when you want to design the primers you can be able to use the software's for analysis of the multiple aspects like whether it is forming a hairpin whether it is forming a dimer what it will be a free energy and all that.

So you people are using different types of software's for all both for designing the primers as well as for analyzing the primers. So let me take you to my lab and where the students are going to explain you the different aspects of the primer designing and how you can be able to analyze those sequences to decide whether your primer is bad versus pure primary school because that is very important because the primary is the you know the crust of the whole story if you are designing a bad primer you are not going to get the amplification.

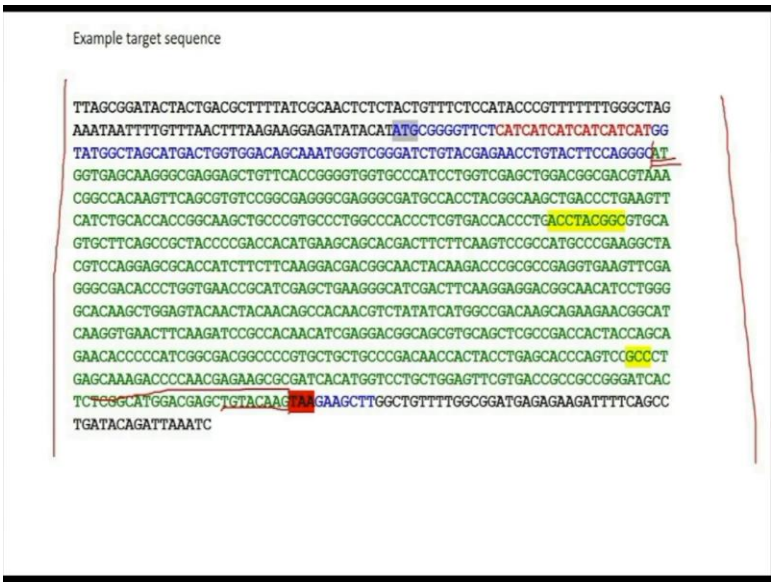
Because that is actually going to be you know give you give the starting point for your enzyme to go and sit and start the synthesis. So that is why the primer designing is a very, very crucial and very, very important to understand. Hello everyone, in this video I will show you how to design the primers and analyze them.

(Refer Slide Time: 22:19)



So for designing primers first you have to identify the region of interest your region of interest which you want to amplify from any vector or any sequence. So in second step you help to identify non cutters there are various software's available but we can use New England Biolabs NEB cutter version 2.O. After identifying non cutters you have to select a suitable vector in which you want to integrate this amplified region and suitable restriction sites you will get suitable restriction sites from non cutters after that you can go for designing forward primer.

(Refer Slide Time: 23:14)



So for understanding purpose I gave this sequence so I am using this sequence I will use this sequence to design the primers and analyze the primers. So this is the whole sequence but I do not want to amplify hold region I want to amplify the letters the sequence which is highlighted in

green. So I want to amplify starting from here to here. So now the question arises what are the non cutters. So you want to amplify this region and integrate into another vector for that you have to identify which are non cutting restriction enzymes. So what I will do? I will copy this sequence into NEB cutter and identify what are the non cutters.

(Video Starts: 24:20)

So I just copy the sequence paste here and I will last submit. So it will analyze the sequence and queue non cutter these are the enzymes cutting inside the sequence. But we are interested in which are non cutter so that means you can see here non cutters. So just click here. It will give number of enzymes which will not cut inside the sequence. So once getting this list we have to identify in which vector you want to integrate your amplified region.

So for that purpose so I have selected for easy of understanding I have selected pET 23a vector. So you can see this is that vector map. So this is the 5 prime side, this is the 3 prime side N terminal and this is the C terminal side. N terminal means forward primer, C terminal means reverse primer so I can use BamH1 in forward primer and Xho1 in reverse primer this is the details. So I have identified 2 restriction enzymes that are BamH1 and Xho1.

So I can use these enzymes in forward primer and reverse primer. So after identifying restriction enzymes and the vector will go for designing forward primer. So I will take this sequence I want to amplify from here to here. So I will copy the sequence here so for designing forward primer it is very easy you have to take the sequence whatever you are getting up to 15 to 20 basis you can take as this.

So if you want to insert a restriction enzyme suppose I want to insert a restriction enzyme this is the sequence as it is given from this whole sequence so I want to insert restriction enzyme that is BamH1. So this is the sequence for BamH1 here it cuts, so I can use this sequence here. So this is the our restriction enzyme here it will cut. So we cannot simply queue like this so there should be some more basis extra basis we have to add in the 5 prime side.

So I will use so this sequence I will use. So now this is 5 prime to 3 prime sides. So this is our forward primer is ready. So after designing this forward primer we have to analyze this

sequence. So this primer so what I will do is I just copy this sequence and I will use OligoAnalyzer software which is specially designed for this purpose only I will paste the sequence just ask analyze.

So here also you can see there are so many options are there like you can analyze hairpin loops, self dimer, hetero dimer. So these are the general details what is the length and GC content melting temperature, molecular weight. So these are normal details I will go for hairpin loop is there any hairpin loops. So we can see there are a number of hairpin loops we can see different different structures predicted by the software.

So if you want to explore this thing you can explore only 2 bases, 2 bases it is performing and the delta G value is -0.43 kilo calorie per mole. So this is fine up to -10 kilo calories per mole is fine. Those hairpin loops broken during the amplification process but above -10 kilo calorie per mole cannot be broken. So in that case what we will do either we redesigned the primers are we will add 5% 1% B10 or 5% is DMS form these are these chemicals disrupt these loops so that the amplification will be fine.

So next I will analyze for self dimer Is there any self dimers and what is the maximum delta G. So this is forming continuously 5 bases It is because of the restriction sites. So those are restriction site on those homodimers forming due to restriction site can be broken there is no issue but other than that this is also because of restriction site but other than that we have to look carefully. So is there any continuously 4 or 5 bases farming this homodimer then it is very difficult these interactions can be broken easily.

So here are some of the consecutive base pairs these are very weak interaction so they can be broken. So, other than that there is no significant self dimmer. So this sequence can be used and for heterodimer predicting heterodimer you need a complementary sequence with reverse prime reverse primer you need. So that we will discuss later so we got our forward primer here. So it is very easy to generate forward primer.

But in case of reverse primer it is somewhat difficult because not in terms of predicting things is somewhat tricky. So what I am saying is here we have sequence. So in case of forward primer we just take an as it is sequence 15 to 20 basis as these from sequences but here we have to take complimentary sequence not 3 prime to 5 prime or 5 prime to 3 prime sequence we have to take complimentary to this one.

Say this is the sequence we got from here. So what is the complimentary to this one? So just I will add here. So this is the complimentary to this particular sequence. So as you can see this is we have to keep from this direction 5 prime to 3 prime. So I will take like this so what we have to do is we want to insert a restriction site here. So we can insert a restriction site here directly. So in reverse primer we wanted to insert Xho1 site.

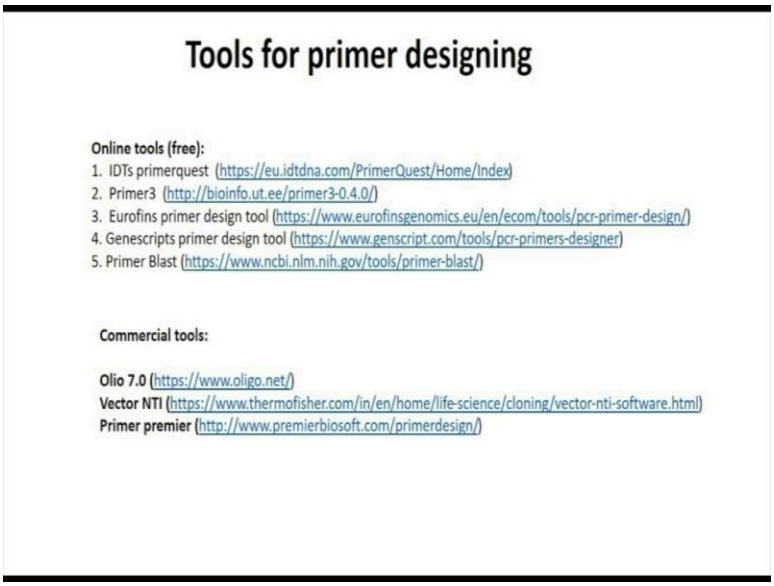
So this is the restriction site as usual we can use the how to insert TAA here. So this is the restriction site we added we can add flanking regions in between a flanking basis before this restriction site. So now we got our reverse primer so we have to go through same procedure like what I have shown in the case of forward 5 primers. So just I will copy paste here and analyze the reverse primer.

So is there any hairpin loops only one hairpin loop that is within the range of delta G. So there is no issue and self dimer. So we can see here continuously 4 bases are forming in this case we have to either change the sequence or remove the some of the bases we can ignore those restrictions those dimers farming through restriction site. So next heterodimer we have to analyze for heterodimer we need forward primer just copy paste here.

And calculate it will give Is there any heterodimers this is because of restriction site, this is also because of restriction site this can be broken those which are at the end of the sequence they can be broken but which is if those bases are middle It is very hard to disrupt those interactions and our amplification will be not good. So there is no amplification literally other kinds of interactions will be broken easily these are quick interactions. So this is how we can prepare design the primers and analyze the primers. We have done all these processes for designing forward and reverse primers.

(Video Ends: 38:55)

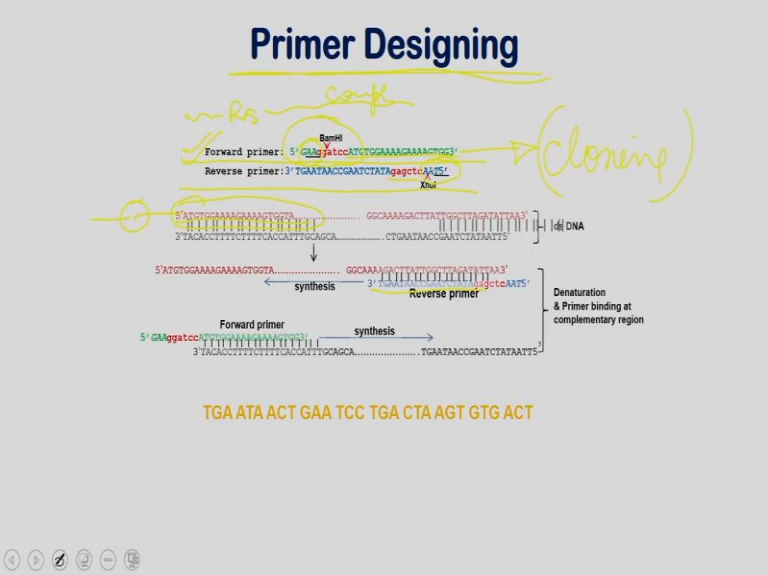
(Refer Slide Time: 38:56)



But instead of doing manually who can do it online we just have to submit the sequence and it will return the forward and reverse premise. These are some of the tools available online for freely. But there are commercial tools also available like Olio 7, Vector NTI, primer premier. So if you are interested in this software's or you can just go through these sites and submit your sequence you will get your primers.

So in this particular demo the students have discussed different aspects related to the primer designing and I hope you might have understood many aspects where the precautions what you have to take and what are the things you should avoid while you are designing the primers.

(Refer Slide Time: 39:54)



So this is like you know the classical example where I have we have synthesized 2 primers in the reverse primer as well as the forward primers. So when you want to synthesize the forward primer suit what you have to do is and this is this is a primer what we have synthesized for the cloning of the particular gene. So the cloning primers are different from the sequencing primers. So in the cloning primer what you have to do is the composition is very important that you take few nucleotides which are actually going to be only for a site for the enzyme to sit.

And then you are going to put the restriction site and then you are going to pit the complementary sequence like the sequence what has been identical to the initial site like this one right. So the forward primer synthesis of our primer is very simple because you can just take the if I prime end sequence you add the restriction side on this side and then you add some more sequences and that is ready for the forward primers.

Whereas in the reverse primers what you have to do is you have to just take this sequence you make it reverse and then you add the reverse the restriction sites and then you can be able to add some more sequences. So that the it will provide the docking site for the enzyme to sit so that because ultimately the you are going to use these primers for making a gene and then you are going to digest this with the restriction lines.

So these are those additional sequences what you are putting so that the restriction sites are going to sit in on this enzyme onto this particular target as amplified DNA and could be able to chew up there are multiple things what the student might have discussed how to analyze and how to verify and all that. So this is all about the primer designing and how you can be able to synthesize or design the primers utilizing the different software's.

And I hope you might have understood the whole process and it will actually going to help you in designing the primers in your lab. And with this I would like to conclude my lecture here in the second subsequent lecture we are going to discuss some more applications of the PCR. Thank you.