Metabolic Engineering
Prof. Pinaki Sar
Department of Biotechnology
Indian Institute of Technology-Kharagpur

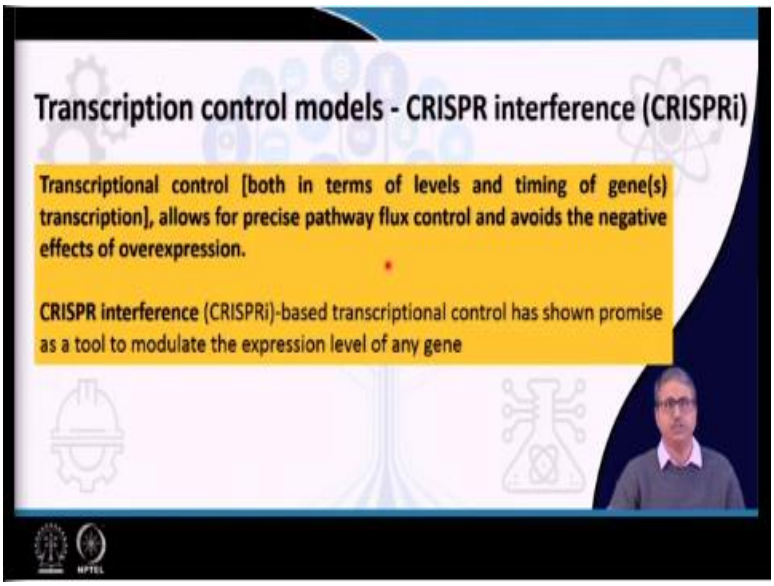
Lecture-32
CRISPR-Cas System and its Application in Metabolic Engineering-Part IV

So, we are going to discuss now the other methodologies which are available for genome editing through CRISPR-Cas9 based technologies. And this is going to be the part IV of my lecture on CRISPR-Cas system and it is application in metabolic engineering.

(Refer Slide Time: 00:48)

So, during this part the following concepts are going to be covered. So, we are continuing our discussion on genome manipulation models. And in this part we are going to discuss about the transcription control models.

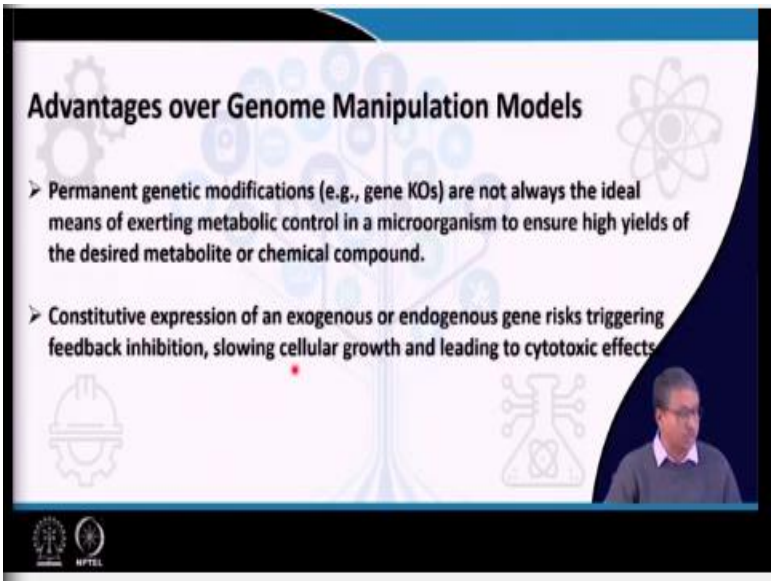
(Refer Slide Time: 01:03)



So, transcription control models are considered as very important methods or models for achieving the genome editing through CRISPR-Cas system. And they are conventionally referred as CRISPR interference or CRISPRi. So, transcriptional control both in terms of the levels and timing of genes transcription that is the expression of genes allows for precise pathway flux control and avoids the negative effect of over expression**.**

CRISPR interference or CRISPRi based transcriptional control has shown promise to as a tool to modulate the expression level of any genes.

(Refer Slide Time: 01:48)



Now the advantage over the genome manipulation models other genome manipulation models that we have discussed earlier. That permanent genetic modifications that is the gene knockouts are not always the ideal means for exerting the metabolic control in the microorganism to ensure high yields of desired metabolites or chemical compounds. And also the constitutive expression of an exogenous or endogenous gene like in HDR.

We have noticed that it is required or sometimes other plasmids are to be there in order to achieve maintenance of these things. So, is basically is risky because it often triggers the feedback inhibition, slowing cellular growth and leading to kind of cytotoxic effects.

(Refer Slide Time: 02:34)

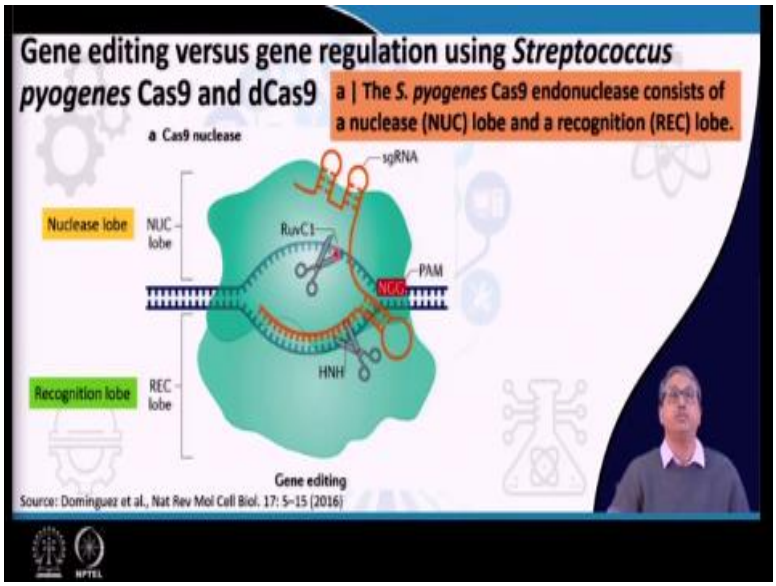
So, in this process the CRISPR-Cas base modification of the transcription deactivation of Cas9, this that is the nuclease is developed, and it is used. So, we are going to talk about that. So, the mutation of key residues, so one of the very fundamental basis of this concept of controlling the transcription of desired genes through CRISPR is basically the deactivation of Cas9 and using a deactivated Cas9.

And what is deactivated Cas9? So, the deactivated Cas9 is developed through mutation, and the mutation of some key residues in the 2 nuclease domains of the Cas9 protein, results in catalytically deactivated Cas9 or dCas9. So, 2 very specific mutations are made in 2 residues

which are involved in or responsible for nuclease activity. So, 2 nuclease activity for cutting both the strands are actually defunct or lost.

And the Cas9 is unable to catalyze the double stranded break formation. But it still retains the ability to bind to the specific region of protospacer, if it is guided by the guide RNA appropriately. So, the dCas9, as I mentioned retains the sequence specific binding ability of the nucleic acid target but ablates the new endonuclease activity, so it cannot create the double stranded break. However it can identify the specific region of it is target site and then bind over there.

(Refer Slide Time: 04:30)

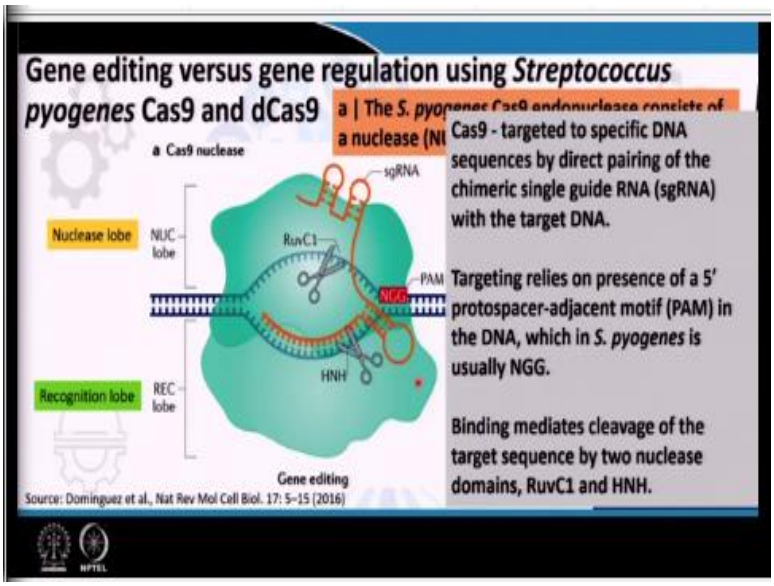


Now here is the structure of the CRISPR-Cas9 particularly and how it works generally. So, here we are going to see that the gene editing versus gene regulation using *Streptococcus pyogenes* Cas9 and dCas9. So, this is basically the Cas9 nuclease structure and on the background we have the DNA and the protospacer and also the spacer and the single guide RNA and also the PAM region is also indicated over here.

So, what is to be emphasized in this particular cartoon is that, it has actually it is showing 2 important nuclease lobes. So, you can see that it has the endonuclease consists of a nuclease lobe. So, this is the nuclease slope the upper dark relatively darker coloured it has called the nuclease lobe, and the recognition lobe or the REC lobe. So, these are the 2 lobes of the Cas9

nuclease**.** And also there are 2 very important sites of the Cas9 which are responsible for cutting the DNA at the appropriate location.

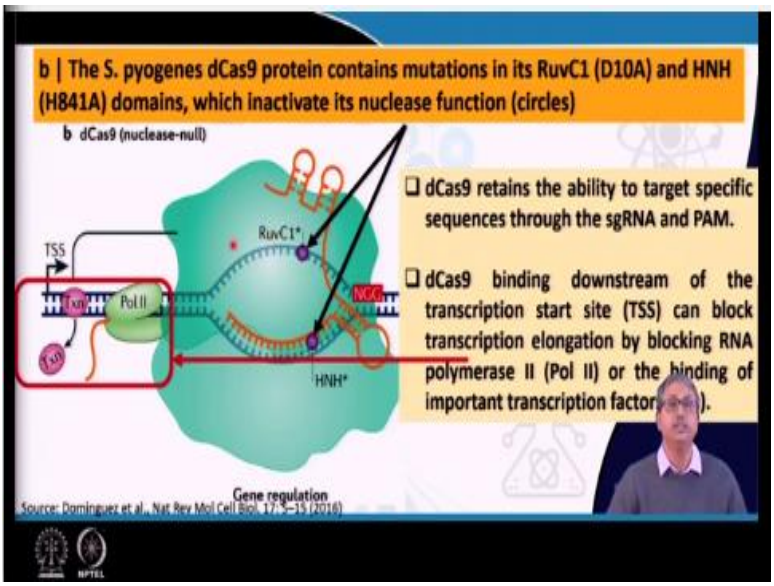
(Refer Slide Time: 05:43)



Now Cas9 targeted to specific DNA sequencing sequences by direct pairing of the chimeric single guide RNA that is the guide RNA with the target DNA. And targeting relies on the presence of the 5 prime PAM residue in the DNA which is in pyogenes is the NGG residue. And binding mediates the cleavage of the target sequence by the nuclease domains, 2 domains are there, that is RuvC1 and HNH region.

So, these 2 domains are involved these are called nuclease domain. So, any kind of mutation in these 2 domains like RuvC1 and HNH nuclease domains would actually lead to loss of catalytic activity, that is the nuclease activity of this Cas.

(Refer Slide Time: 06:36)



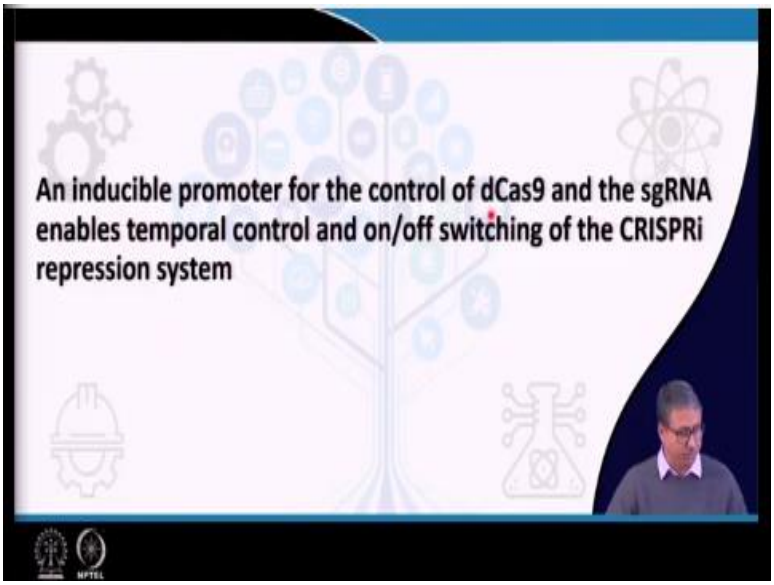
So, this particular dCas9 protein with mutation in the RuvC1 region and HNH region or domains will lead to inactivation of the nuclease activity. And thereby it is now incapable of carrying out any kind of cleavage or double stranded break. So, if both these domains are mutated then it will not be able to catalyze the DNA cleavage reaction. Now as we mentioned earlier, so now this dCas9 means deactivated Cas9 is unable to cleave the DNA in either of these strands, so double stranded break is not possible, now neither single stranded break.

But it binds to the target region because of the complementarity of the spacer and the protospacer and the recognition by the PAM. Now what are the consequences if the nuclease activity is lost but the Cas9 is able to bind to the target site? Now these are the 2 nuclease functions which are now lost, so the DNA cannot be cleaved. However it binds to the target site. Now dCas9 binding downstream of the transcription start site that is the TSS.

So, if we are capable of targeting this deactivated Cas9, so that it goes and it binds to the downstream of the transcription start site. That means it can block the transcription elongation process by blocking the RNA polymerase II or Pol II. So, as it is clearly depicted over here that if we can design a suitable spacer based on our choice that where we want to take this Cas9 deactivated Cas9. So, the deactivated Cas9 will reach to the target gene or target site and if this spacer is so designed, that it is just downstream of the transcription start site.

And then the deactivated Cas9 will bind and will sit there and it will block the transcription elongation by the Pol II or the RNA polymerase or the binding of the important transcription factor that is the Txn. So, neither the Txn will be able to bind nor the Pol II RNA polymerase would bind and function if deactivated Cas9 is sitting there in a transcription start site. So, it has a serious consequence on the cell that is basically the transcription of the targeted gene is inhibited.

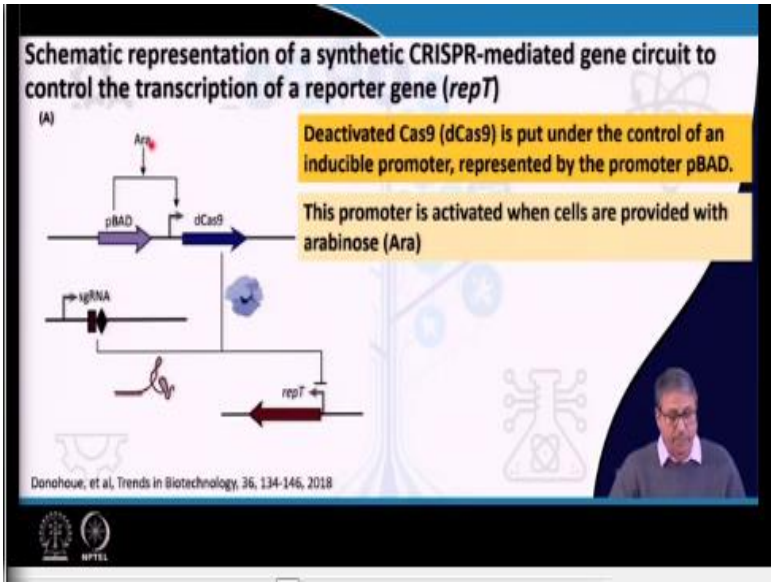
(Refer Slide Time: 09:32)



Now an inducible promoter for the control of the dCas9 and the single guide RNA enables temporal control and on, off switching of the CRISPRi represent system. So, we are going to see that. That if we can target the dCas9 to a specific gene of our interest or a specific location of our gene of interest and the single guide RNA is available. Then it will lead to interference of that expression of that particular gene.

And if the expression of the dCas9 or the expression of the single guide RNA is under the control of some kind of inducer. Then providing the inducer molecule or withdrawing the inducer or the lack of inducer molecule would act as a kind of or help the system to act as on, off switch.

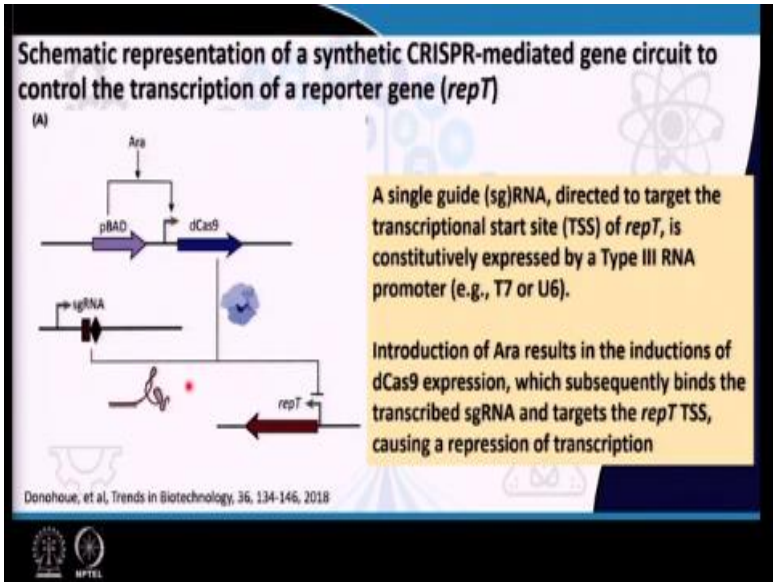
(Refer Slide Time: 10:32)



So, here is the schematic representation of the synthetic CRISPR mediated gene circuit which is developed to control the transcription of a reporter gene. So, this is repT gene, so here is the repT gene. So, now we are going to see that how the expression of this repT gene can be controlled very specifically by providing external stimulus. Now firstly the deactivated Cas9, the dCas9 is put under the control of an inducible promoter which is in this case is a pBAD.

So, as we know that pBAD is actually under activated when cell is provided with arabinose. So, the dCas9 nuclease protein gene is in a plasmid and it is under the control of pBAD and pBAD is only activated when the cell is supplied with arabinose. So, that means if the cell is provided with arabinose the pBAD will allow the transcription and then the dCas9 will be produced. So, the production of dCas9 within the cell engineered cell or the availability of this dCas9 within the engineered cell is under the control of availability of arabinose or non availability of arabinose.

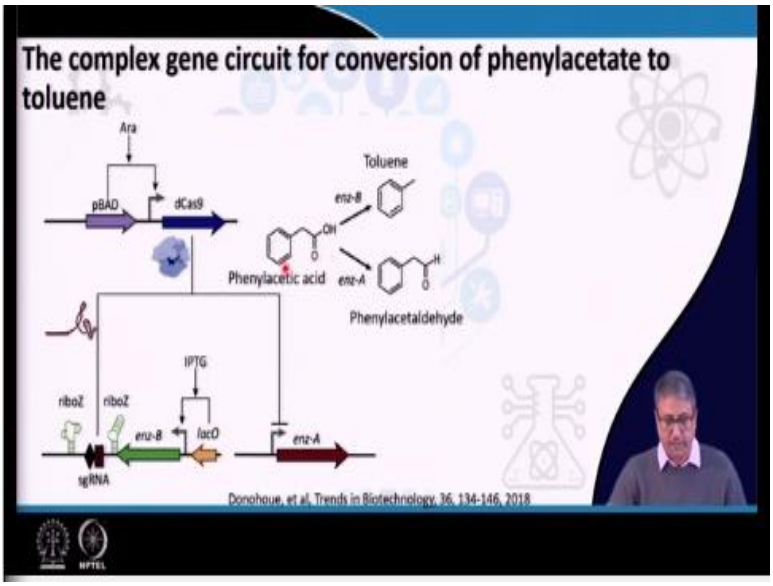
(Refer Slide Time: 11:49)



Now a single guide RNA to target the transcriptional start site TSS of repT. Now we know that the what is the target region transcription binding transcriptional start site of repT. So, a single guide RNA can be produced which can target the TSS of the repT. And it can be constitutively expressed by type III RNA promoter that is T7 or U6. Now the moment we introduce arabinose into this culture system which will lead to the introduction or induction of the dCas9 expression.

So, if arabinose is available then dCas9 will be expressed and that will subsequently bind to the transcribed sgRNA. Because sgRNA single guide RNA is already there because it is constitutively express. So, the dCas9 will now be taken to the TSS site of the repT, because the single guide RNA is specific to this particular site of the TSS of repT. So, the entire dCas9 and sgRNA will bind to the TSS of repT site, repT and thus the expression of the repT will be inhibited or stopped.

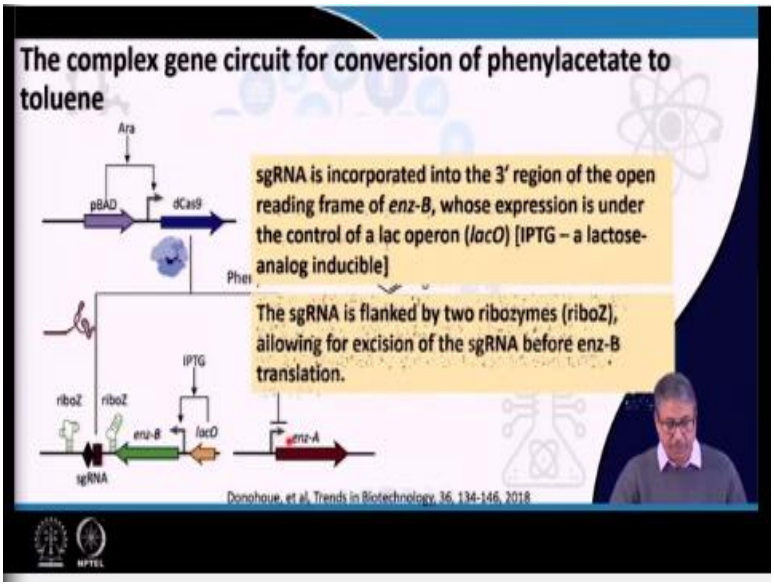
(Refer Slide Time: 13:12)



Now the complex gene circuit for the conversion of phenyl acetate to toluene is achieved by extending this particular technology further. Now in the initial phase we observed that how the inducible promoter can be used to control the activity of the dCas9 while the single guide RNA remains under the control of constitutive expression. So, any time present of arabinose or this inducer will enable the cell to express dCas9 or not expressing the dCas9.

Now the conversion of the phenyl acetic acid to either to toluene or phenyl acetaldehyde is under the control of 2 genes the enzyme B and enzyme A. So, if you have a high concentration of enzyme A, then phenyl acetic acid is going to convert into phenyl acetaldehyde. But if we have higher the enzyme A is not operating properly then enzyme B will be able to take the flux towards the toluene production. So, in this case the system is engineered to produce toluene from phenyl acetic acid, we will see how it is achieved.

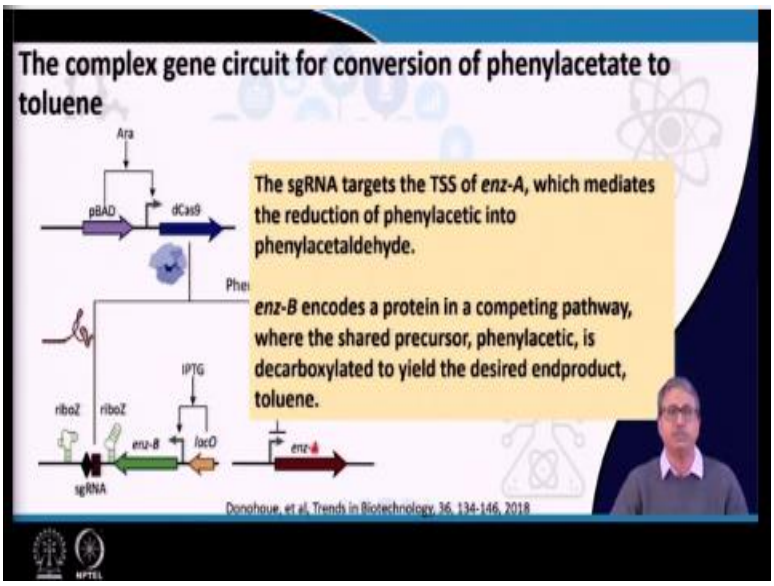
(Refer Slide Time: 14:29)



Now here the single guide RNA is incorporated into the 3 prime region of open reading frame of enzyme B which is responsible for converting this compound to toluene, whose expression is under the control of lac operon. And we know that IPTG is the inducer for this lac operon. So, now this single guide RNA, and this single guide RNA is actually these spacer of this guide RNA is complementary to the particular region of the enzyme.

And the single guide RNA is also flanked by two ribozymes that is riboZ is abbreviated here allowing for excision of the single guide RNA before the enzyme B translation**.** Now it is the enzyme B is over here, we have the IPTG inducible lac operon over here, lacO system is there and so PAM finally we have the single guide RNA. And this guide RNA is actually responsible for binding to the some region of the enzyme A.

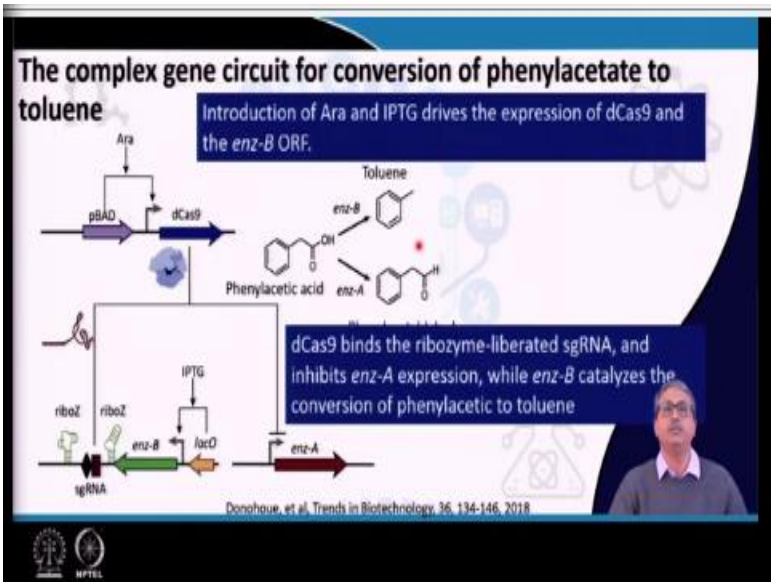
(Refer Slide Time: 15:54)



Now the single guide RNA targets the TSS of the enzyme A. Now the complexity of the entire process is such that this guide RNA is specific to the transcription binding or transcription start site for enzyme A. And enzyme A is the kind of enzyme which is not desirable in this case because it is supposed to be taking the flux not towards the toluene but towards the phenyl aldehyde.

Now the single guide RNA targets the TSS of the enzyme A which mediates the reduction of the phenyl acetic acid into the phenyl aldehyde. Now enzyme B encodes the protein in the competing pathway that is conversion of the phenyl acetic acid into the toluene which is the desirable product in this case.

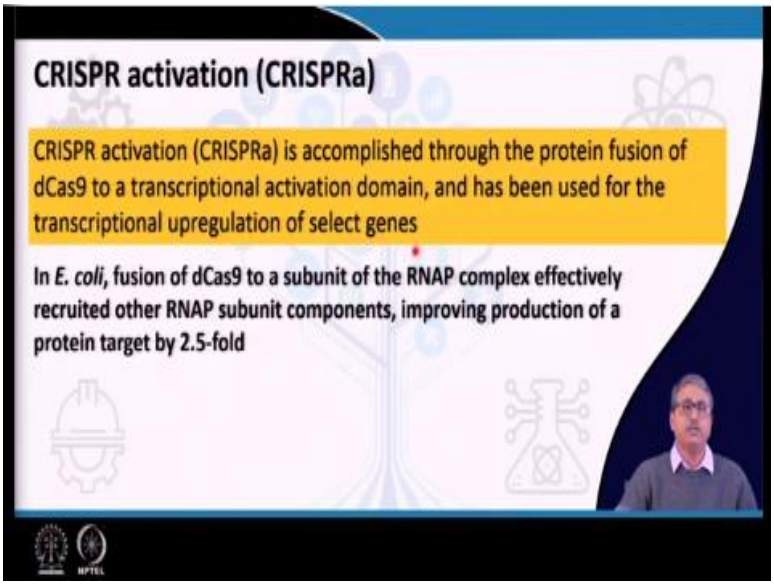
(Refer Slide Time: 16:45)



Now the introduction of both the inducers, that is the Ara and the IPTG drives the expression of dCas9 and the enzyme B ORF. So, this Ara will induce the production of the dCas9 and the IPTG will lead to the production of the enzyme B. And it will also produce the sgRNA which is actually complementary to the TSS of enzyme A**.** And then dCas binds to the ribozyme liberated sgRNA and inhibits the enzyme A expression.

So, it goes there and finally bind to this enzyme A TSS, and thereby reducing the or inhibiting the expression of enz-A while enz-B or enzyme B is fully expressed and it is capable of converting the phenyl acetic acid into toluene.

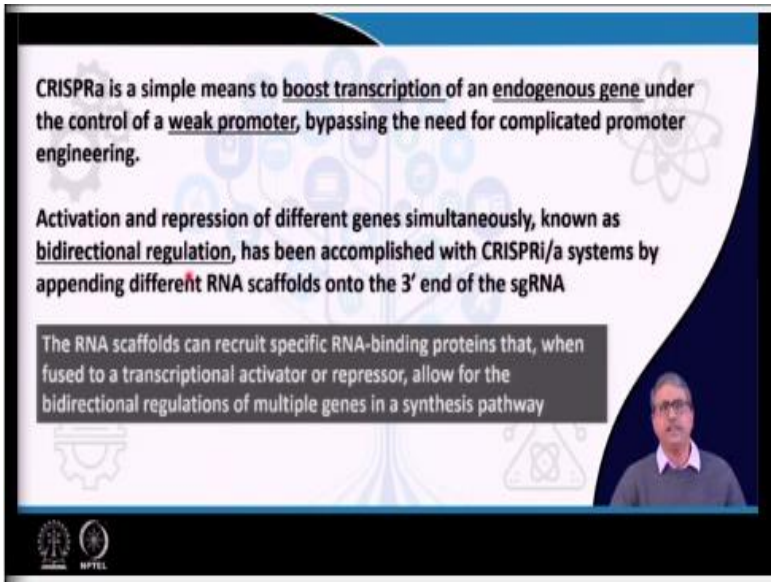
(Refer Slide Time: 17:43)



Now the next advancement is made through some more modification within this interference. But it is not interference it is rather activation of the CRISPR system. Now in this CRISPR activation or we call it CRISPRa is accomplished through the protein fusion of dCas9 to a transcriptional activation domain and has been used for the transcriptional upregulation of selected genes.

Now in E.coli the fusion of dCas9 to a subunit of the RNA polymerase complex, effectively recruited other RNA polymerase subunit components improving the production of the protein target by more than 2.5 fold.

(Refer Slide Time: 18:29)

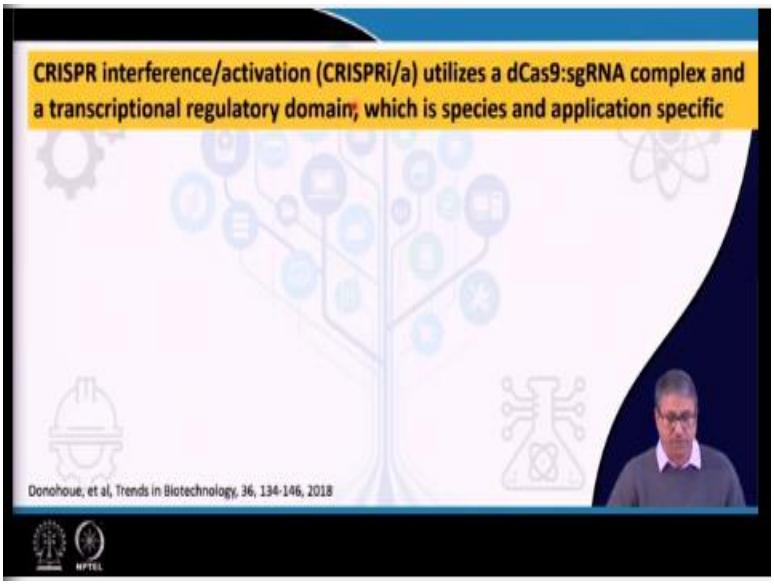


Now the CRISPR inference CRISPR activation system is considered to be a simple means to boost the transcription. So, it allows the enhanced kind of transcription of the endogenous gene under the control of a weak promoter, bypassing the need for complicated promoter engineering. So, if we are capable of doing this it would actually help us to avoid a number of complicated genetic engineering procedures.

And the RNA scaffolds can recruit specific RNA binding proteins that when fused with the transcriptional activator or repressor allow for the bidirectional regulation of the multiple genes in a synthetic pathway. Particularly the activation and repression of different genes simultaneously known as the bidirectional regulation, and it has been accomplished with the

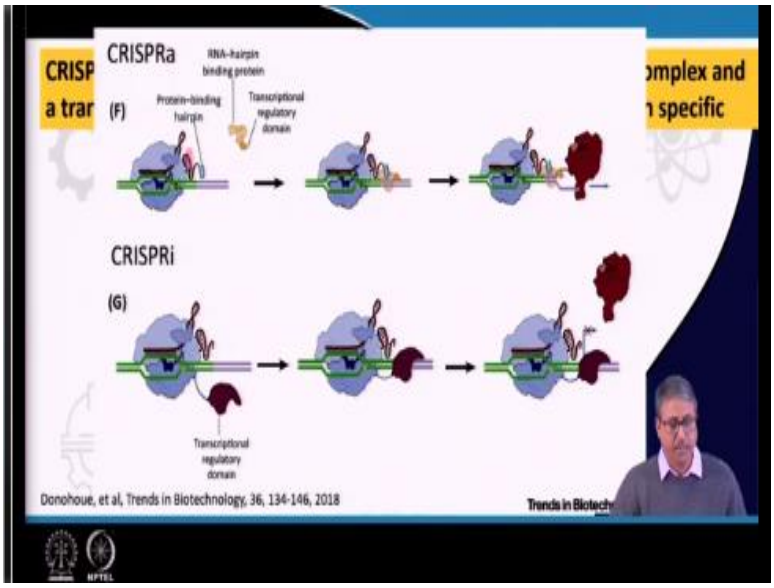
CRISPRi or CRISPRa system by appending the different RNA scaffolds onto the 3 prime end of the single guide RNA.

(Refer Slide Time: 19:31)



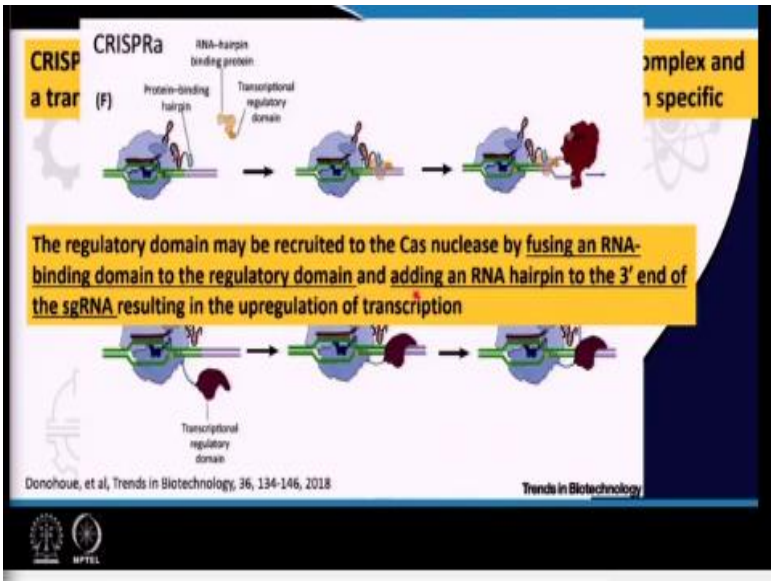
Now CRISPR interference or activation that is CRISPRi or CRISPRa utilizes the dCas9 single guide RNA complex and transcriptional regulatory domain which is a species and application specific.

(Refer Slide Time: 19:46)



So, here is a kind of a general layout of the entire process, so in case of CRISPRa.

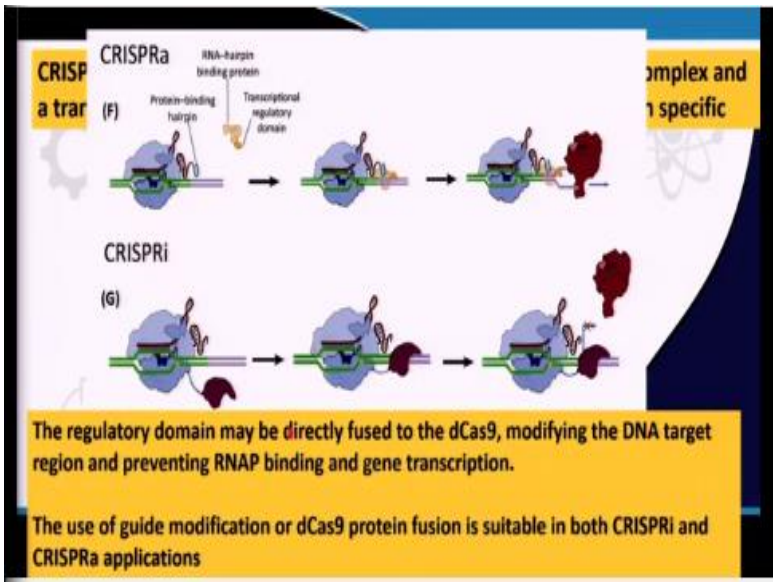
(Refer Slide Time: 19:52)



As you can see over here the regulatory domain may be recruited to the Cas nuclease by fusing an RNA binding domain to the regulatory domain. And adding an RNA hairpin that is the hairpin over here, RNA hairpin to the 3 prime end of the single guide RNA resulting into the upregulation of the transcription. So, you have the DNA hairpin binding protein which is connected to the transcriptional regulatory domain.

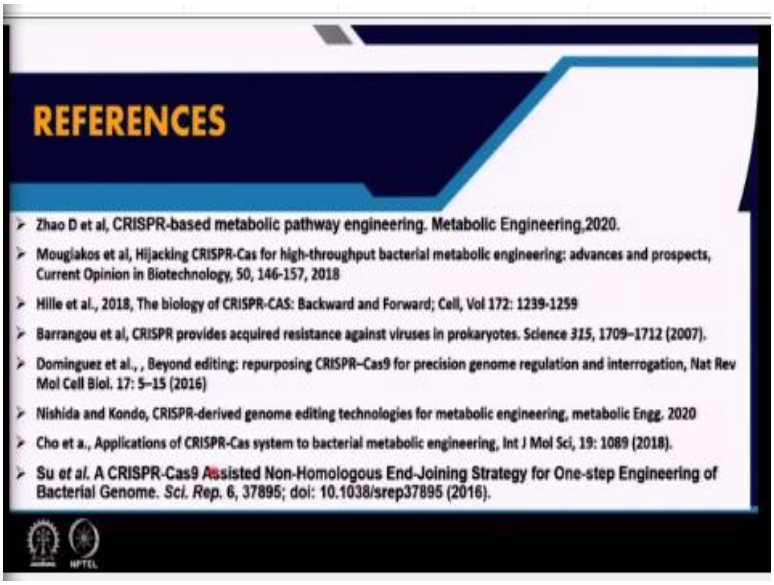
And you have the protein binding hairpin connected to the single guide RNA and following their binding this transcriptional regulatory domain is now attached to the entire deactivated Cas9 complex. And that facilitates the binding of other polymerase molecules which will enable the Cas9 or the expression of the gene or activate the expression of the gene.

(Refer Slide Time: 20:40)



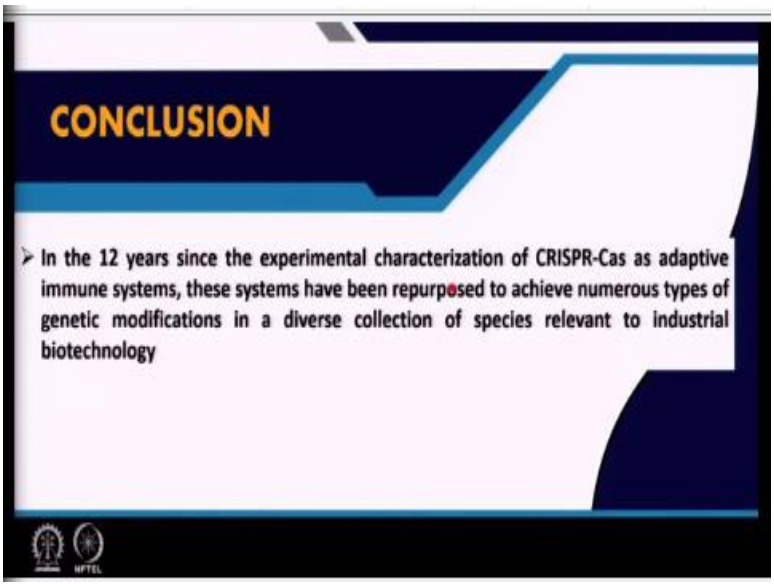
Similarly the regulatory domains will be directly fused to the dCas9, modifying the DNA target region and preventing the RNA polymerase binding and gene transcription. Now the use of guide RNA guide modification or dCas9 protein fusion is suitable for both CRISPRi and CRISPRa application.

(Refer Slide Time: 21:02)



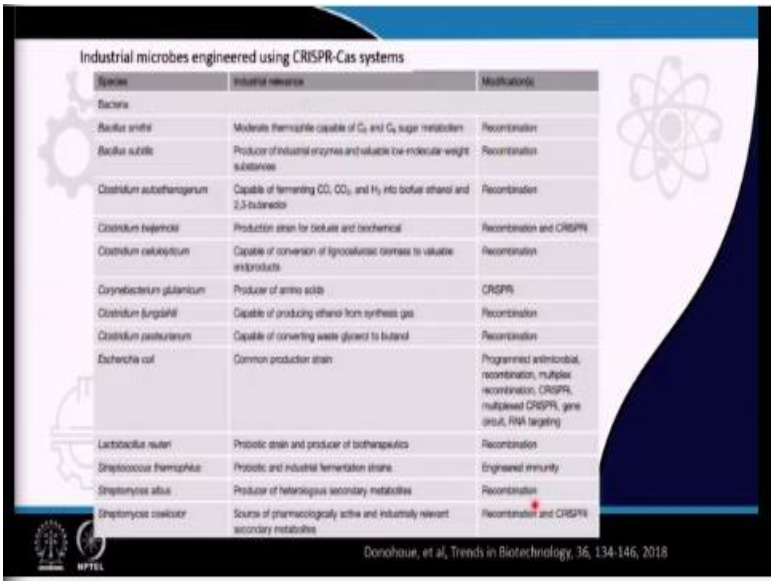
And for this particular portion of this part 3 and as well as part 4 of my lecture the following the papers would be very useful. Particularly some of the papers like this CRISPR provides the acquired resistance against viruses. The biology of CRISPR-Cas backward and forward and the hijacking the CRISPR-Cas for high throughput bacterial metabolic engineering etcetera, would be are very useful.

(Refer Slide Time: 21:32)



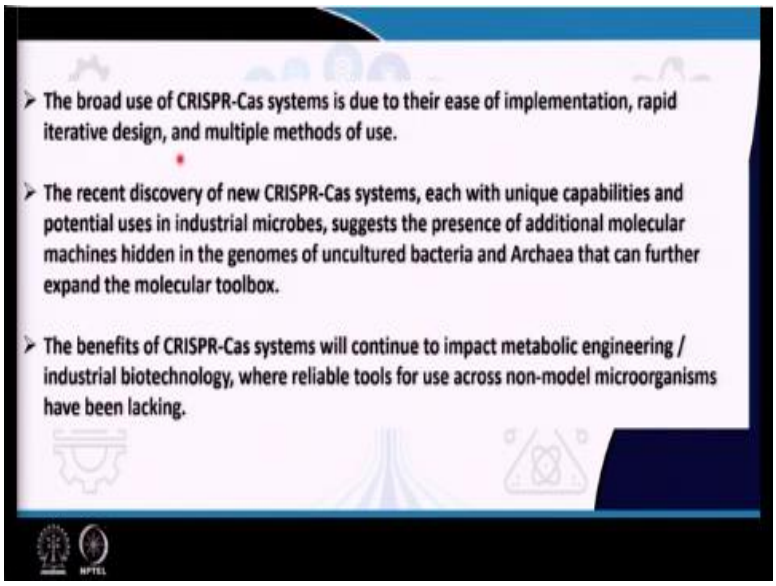
So, in conclusion in the last 12 years or so, since the experimental characterization of CRISPR-Cas as adaptive immune system is achieved. These systems different CRISPR-Cas systems have been repurposed to achieve numerous types of genetic modification in a diverse collection of species relevant to industrial biotechnology in general or metabolic engineering in particular.

(Refer Slide Time: 21:58)



And here is a kind of a brief list of the organisms and the industrial relevance and the kind of modifications which have been incorporated.

(Refer Slide Time: 22:08)



The broad use of CRISPR-Cas system is basically due to their ease of implementation, rapid iterative design and multiple methods of use. And recent discovery of the new CRISPR-Cas system each with unique capabilities and potential uses in industrial microbes, suggests the presence of additional molecular machines hidden in the genomes of uncultured bacteria and archaea that can further expand the molecular toolbox.

The benefit of CRISPR-Cas system will continue to impact metabolic engineering or industrial biotechnology, where reliable tools for use across non-model microorganisms have been lacking so far, thank you.