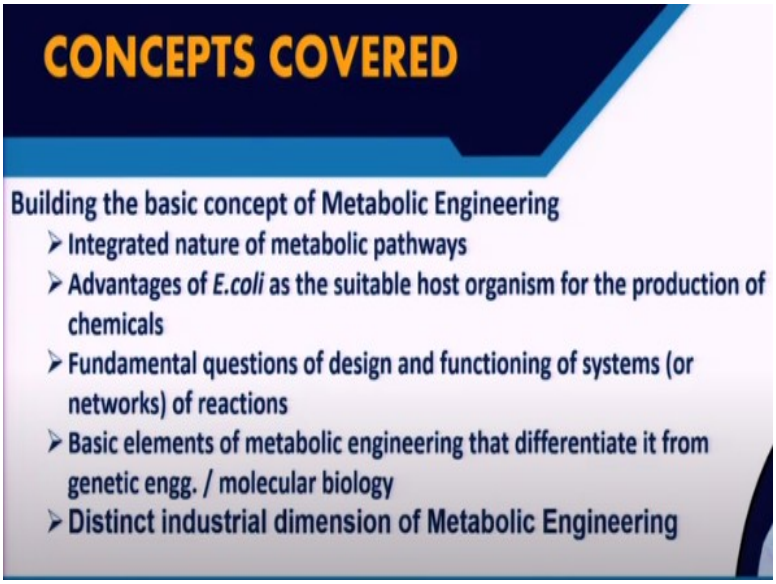
Metabolic Engineering
Prof. Pinaki Sar
Department of Biotechnology
Indian Institute of Technology-Kharagpur

Lecture - 03
Essence of Metabolic Engineering - Part B

Welcome to the NPTEL course on metabolic engineering and today we are going to continue our discussion on the essence of metabolic engineering. And during this introductory lecture, which is continued from my previous lecture, today we are going to discuss few other very important points of metabolic engineering.

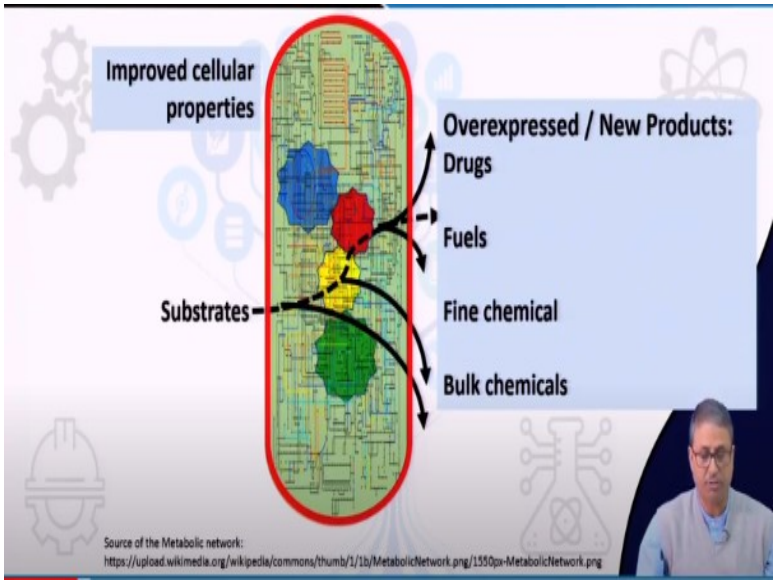
(Refer Slide Time: 01:11)



One of the major issue with the metabolic engineering concept is the incorporation of the fact that it is of highly integrated in nature in terms of the metabolic pathways present within the host organisms. In today's lecture, I am also going to highlight the advantages of *E. coli* as the suitable host organism for the production of different chemicals through metabolic engineering.

Fundamental questions of design and functioning of systems or networks of reactions will be introduced. Basic elements of metabolic engineering that differentiated from genetic engineering and molecular biology will be highlighted. And finally, the distinct industrial dimension of metabolic engineering will be briefly discussed.

(Refer Slide Time: 02:17)

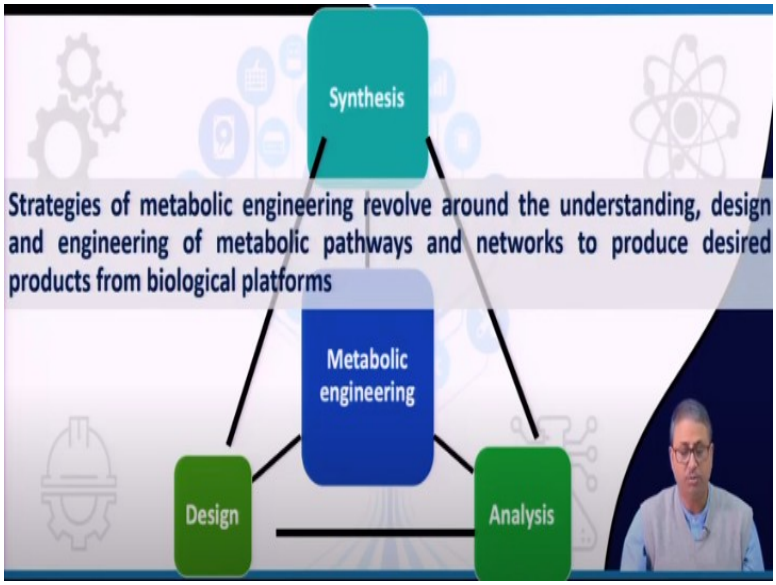


So as we understood from the previous lecture, that metabolic engineering aims to improve the cellular properties of the products produced by the host organism by specific modifications through metabolic pathways and these modifications are done using a set of well-defined recombinant DNA technology.

Now while we try to produce a number of new products or products which are natural to the host organisms that include drugs, fuels, fine chemicals, bulk chemicals, etc. as well as we try to improve different type of cellular properties of the host microorganisms, we need to understand that the pathway that allows the conversion of the substrate to different products are not always very simple and straightforward.

And they are often highly interconnected and they are the part of interconnected and complex metabolic network.

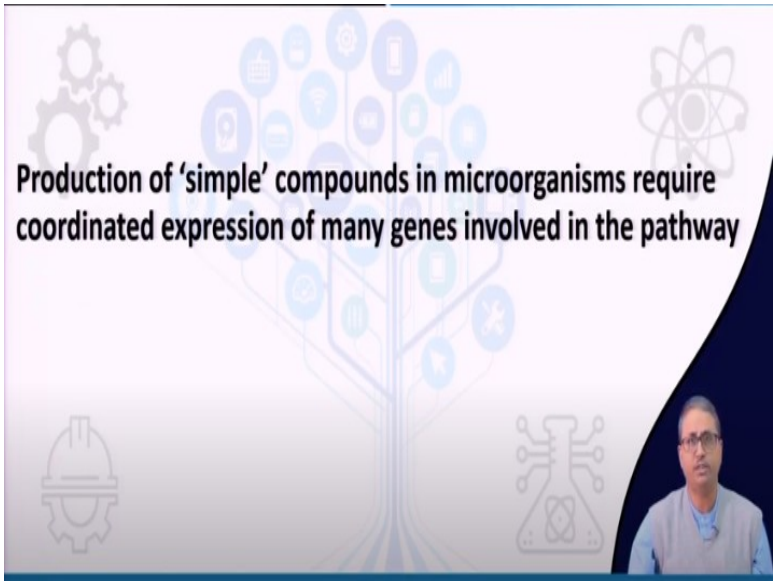
(Refer Slide Time: 03:44)



We also discussed in our previous lecture that metabolic engineering is basically composed of three important components. The components are synthesis, analysis and design. The strategies of metabolic engineering revolve around the understanding of these designing and engineering of metabolic pathways. So based on our goal, we need to select that whether we start with the synthesis step or we start with the design step.

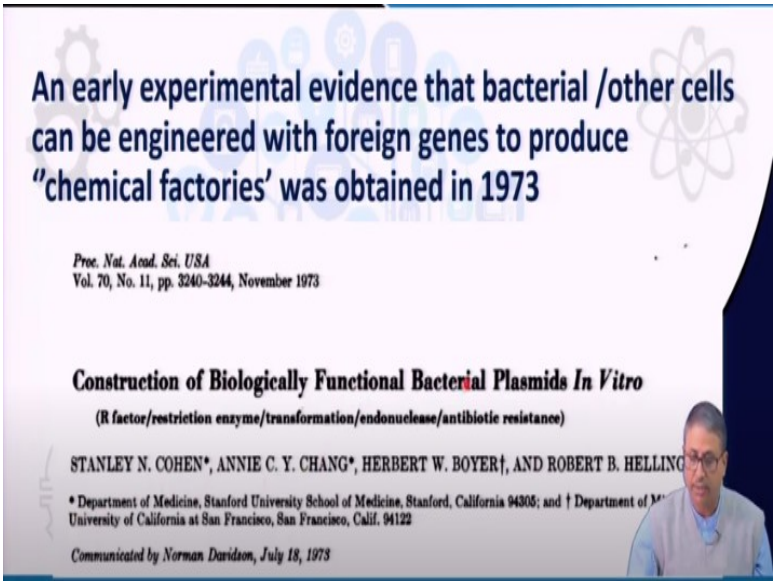
But one of the basic and fundamental requirement for metabolic engineering with any kind of host organism is the analysis of the metabolic pathways and metabolic reactions, which are leading towards the specific product formation.

(Refer Slide Time: 04:46)



Now in today's class, we are going to discuss on a particular aspect of metabolic engineering that is that metabolic pathways present within the microbial cells or other organisms are often highly coordinated and they act as a kind of system. Now the production of simple compounds like ethanol and many other similar type of compounds in microorganisms require a coordinated expression of many genes involved in the pathway or pathways.

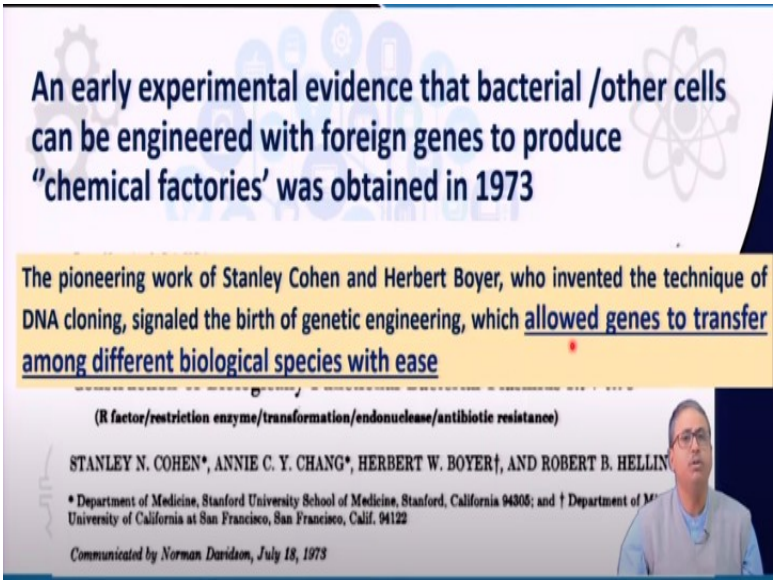
(Refer Slide Time: 05:28)



So one of the early experimental evidence, which possibly paved the path for genetic engineering in the subsequent time showed that bacterial or other cells can be engineered with foreign genes. Genes obtained from other organisms, other microbial cells or even cells of eukaryotic nature like even the human cells. And these foreign genes can be brought into *E. coli* or other bacterial system and they can be utilized to produce a number of chemical compounds.

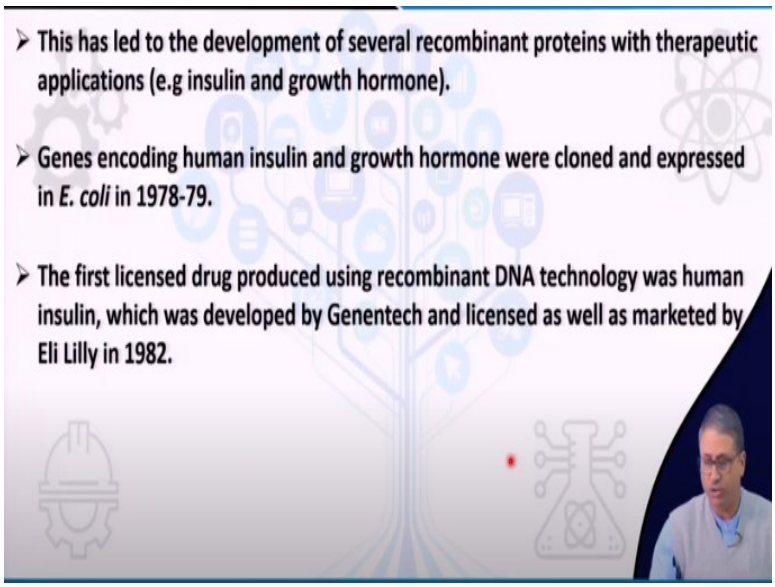
And thereby, the host microbial cell like the *E. coli* cell for example might work like a chemical factory. So this concept was initially shown by the research team of Cohen and Boyer mainly. And it was way back in 1973 when the construction of biologically functional plasmids were shown, that the two different type of genes can be combined in vitro and then they can be successfully transferred to a living cell allowing their expression within the host organism.

(Refer Slide Time: 06:57)



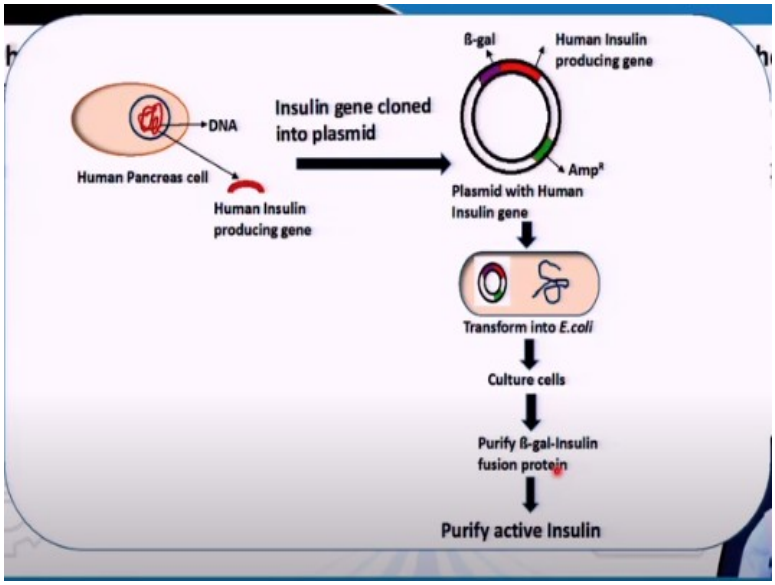
Now the pioneering work of Cohen and Boyer, they invented the technique of DNA cloning, which we refer to, signal the birth of genetic engineering which allows genes to transfer among different biological species with ease. So following this successful experimentation, there are numerous experimental success, where foreign genes have been transferred to E. coli or other host organisms for the production of numerous chemical compound.

(Refer Slide Time: 07:36)



So with the experimental success of Cohen and Boyer, the development of several recombinant proteins having therapeutic applications were initially emphasized, and that includes insulin and different growth hormones. Genes encoding human insulin in particular, were cloned and expressed in *E. coli* in 1978, 1979.

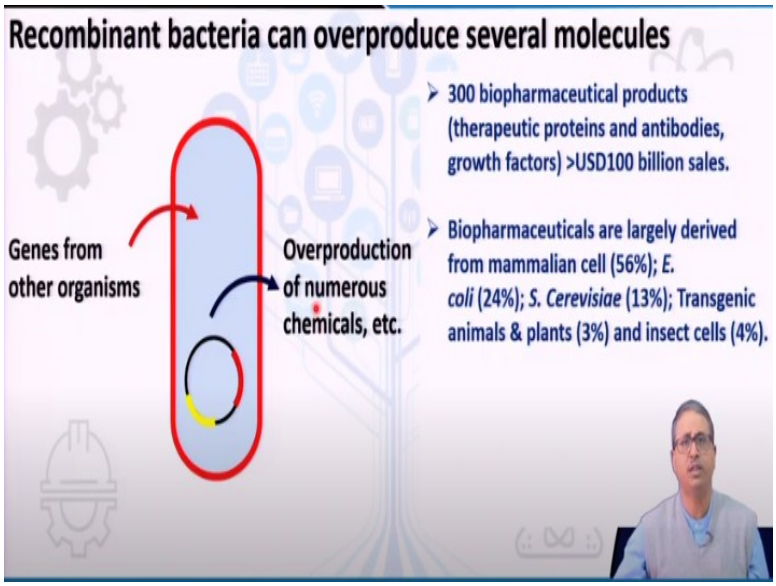
(Refer Slide Time: 08:10)



And this diagrammatic representation shows that how the human insulin gene can be cloned from human cell and then transferred to an *E. coli* host organism and those transformed *E. coli* cells having the recombinant plasmid having the insulin, human insulin gene, will eventually produce the particular insulin protein and that protein can be purified and then following other modifications and processes can be used as therapeutic agent.

Now the fast licensed drug which was produced out of these recombinant DNA technology was the human insulin using the strategy that generally laid out in my previous diagram, which was developed by Genentech and then licensed and marketed by Eli Lilly in way back 1982.

(Refer Slide Time: 09:27)

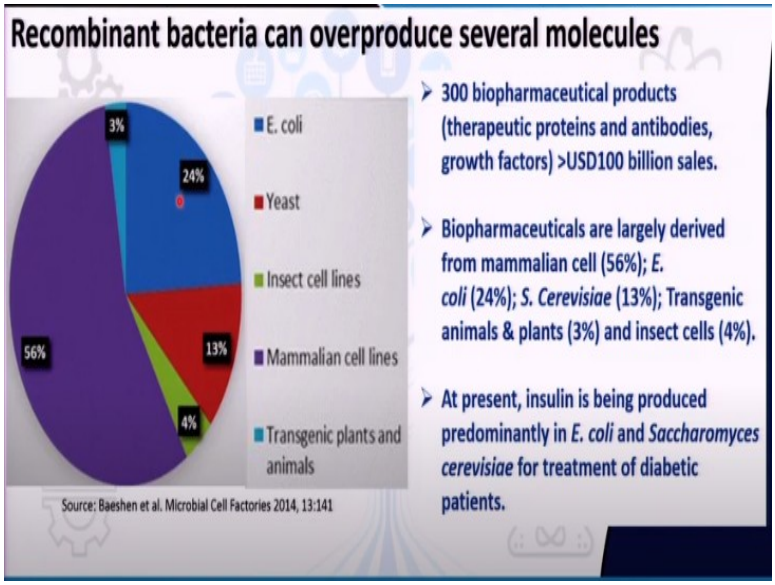


So now this recombinant bacteria have been found to be an excellent tool to produce several or rather over produce several molecules of industrial or biotechnological significance. So genes from different organisms, not necessarily of bacterial nature but from fungal and Archaea and other organisms, including the eukaryotic organisms like humans also can be cloned, brought into the *E. coli* cell system or a suitable host system and then over production of the chemicals can be possible.

Now in the past 10 years or so we have seen that around 300 biopharmaceutical products including the therapeutic proteins and antibiotics, antibodies, different growth factors have been successfully produced using the concept of metabolic engineering where the recombinant DNA technology played a very important role.

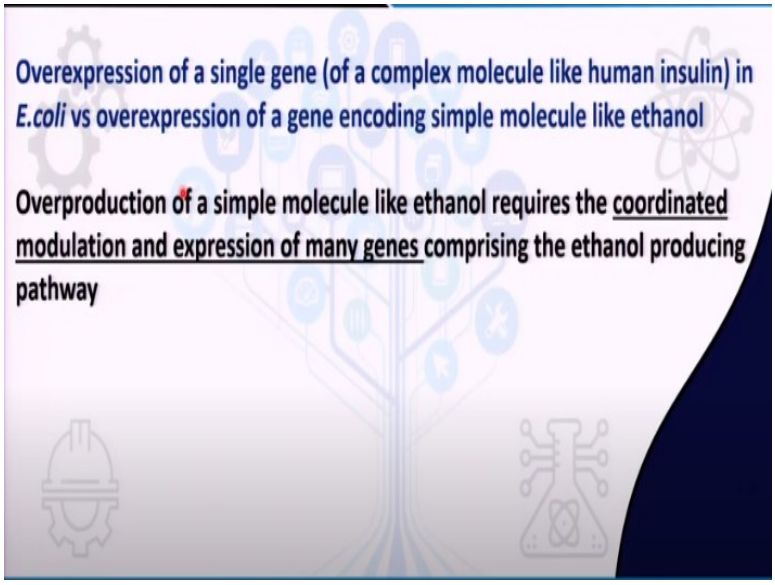
Biopharmaceuticals are largely derived from mammalian cells around 56%, *E. coli* around 24%, and *Saccharomyces cerevisiae* around 13% and transgenic animal and plants they constitute around 3%, insect cell around 4%.

(Refer Slide Time: 10:54)



So at present insulin is being produced predominantly in *E. coli*. So this is the distribution pie chart of the source of different molecules, which are therapeutic molecules mainly including the therapeutic proteins and antibodies, growth factors etc. And you can see that *E. coli* represent a significant portion of these production of these molecules including in particular the insulin.

(Refer Slide Time: 11:21)



Now overexpression of a single gene like insulin of complex molecule, complex nature in *E. coli* versus the overexpression of a gene encoding simple molecule like ethanol. So why we are discussing about this comparison that the expression or overexpression of an eukaryotic gene into *E. coli* and an existing prokaryotic gene, which is already present in *E. coli* system into *E. coli* itself, it is like just like over production or over expression.

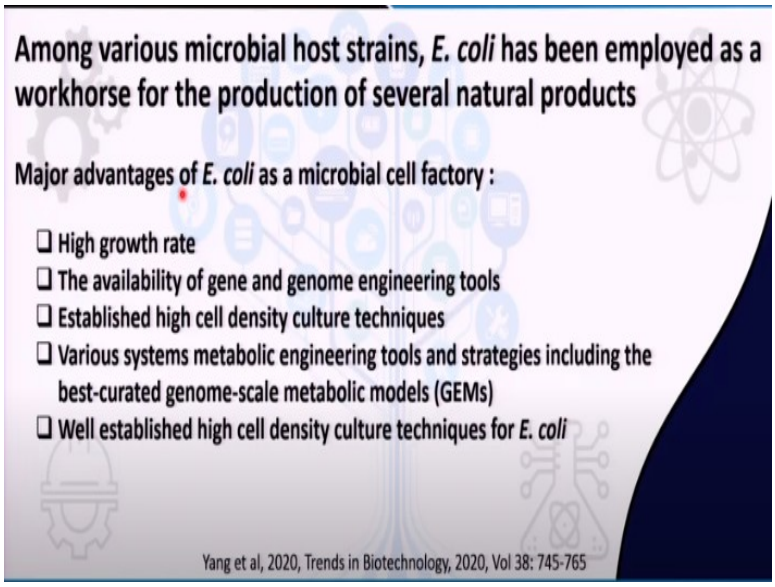
We are comparing because, when the scientist tried to engineer organism for the purpose of complex molecules like insulin, which are encoded by one or two genes, and when they try to adopt the similar strategy for different other molecules, which are apparently very simple molecules like ethanol.

But they are encoded in many cases by multiple genes or multiple genes are actually found to be involved in their production we experienced a very interesting or rather contradictory results. So overproduction of a simple molecule like ethanol requires the coordinated modulation and expression of many genes comprising the ethanol producing pathway.

So in case of the natural product formation, so *E. coli* cells for example, are naturally endowed with the ability to produce ethanol, although the concentration level is generally very low. So when the scientist who tried to overproduce ethanol in *E. coli* cell using the different strategies of recombinant DNA technology, they observed that

unlike the cloning and expression of insulin genes, the ethanol overproduction requires the coordinated modulation and expression of a large number of genes.

(Refer Slide Time: 13:49)

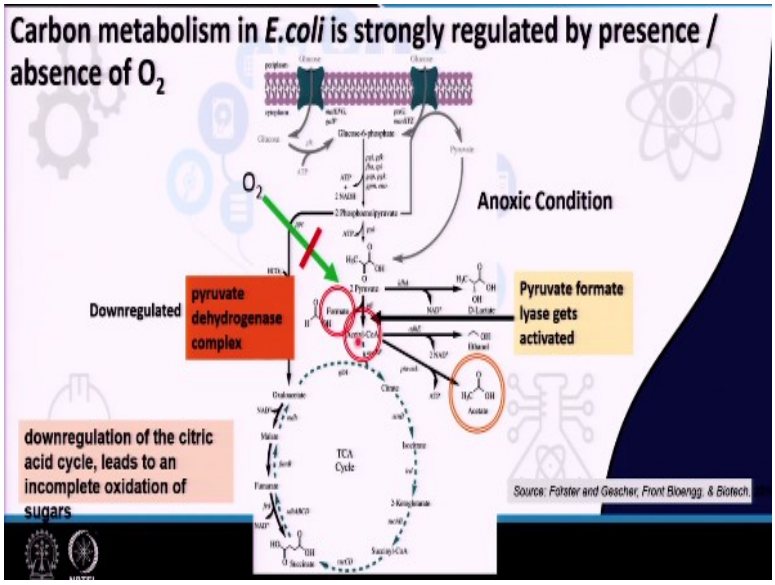


That was very interesting. So among various microbial host, why we or the scientists who are working on metabolic engineering are selecting *E. coli* because *E. coli* has been employed as a workhorse for the production of several natural products. Now the major advantages of *E. coli* being the microbial cell factory are the high growth rate, the availability of genes and genome sequencing tools, because these tools are quite well developed for *E. coli* system and the genes are available.

Established high cell density culture techniques for large scale propagation of the cells. Various systems, metabolic engineering tools and strategies including the best curated genome scale metabolic models or GEMs. And well established high cell density culture techniques are also there for E. coli.

So *E. coli* has been identified as one of the main organisms for carrying out the experiments and using them as the production strain thereby, they represent one of the major organisms as microbial cell factories.

(Refer Slide Time: 15:15)



Now before we enter into the details of the interactions process, let us try to look at the carbon metabolism in *E. coli* very briefly or very generally. This is important because in contrast to an eukaryotic gene like insulin when we tried to do that, when we were trying to take the similar strategy for ethanol production, we experienced lot of troubles. The scientist experienced lot of difficulties.

And these difficulties were realized only when the carbon metabolism within *E. coli* system was investigated thoroughly. So as depicted in this picture, we can see that the glucose following its uptake by different transport system is converted to glucose 6-phosphate or to pyruvate and eventually it is converted to phosphoenolpyruvate. This is a major step of the glycolytic carbon metabolism or glucose metabolism.

Now this phosphoenolpyruvate is further oxidized to pyruvate and this pyruvate is now representing a very important intermediate or metabolite to power the different types of metabolic reactions further. Now if we look carefully, under oxic condition, when oxygen is supplied sufficiently or the cells are growing under oxic condition, this pyruvate formation or rather pyruvate to acetyl-CoA formation is highly stimulated because of the functioning of the pyruvate dehydrogenase complex.

So pyruvate dehydrogenase complex is the enzymatic complex which is responsible for converting the pyruvic acid to acetyl-CoA. So during oxic condition, this enzyme complex is highly activated, and they carry out the conversion of pyruvic acid to acetyl-CoA. Now this acetyl-CoA is subsequently processed by the TCA cycle.

And as the acetyl-CoA molecules are oxidized further through the TCA cycle, a large number of ATP molecules and other reducing equivalents are potentially generated like NADH, H+ or FADH2 and the Gibbs free energy are produced during this TCA cycle. So this is a very known phenomenon for microbial metabolism of glucose through this Embden-Meyerhof-Parnas pathway or glycolytic pathway.

Now under anoxic condition, when there is no oxygen in the system, so the first thing happens is the down regulation of the pyruvate dehydrogenase complex so which was earlier very activated is now down regulated. So the moment it is down regulated, the pyruvate dehydrogenase the moment it is down regulated, different type of phenomenon start happening inside the cell.

The down regulation of this pyruvate dehydrogenase complex is followed by the down regulation of the entire citric acid cycle or TCA cycle, because the supply of acetyl-CoA is decreased, because the original enzyme which was responsible for the conversion of pyruvate to acetyl-CoA is down regulated so the acetyl-CoA concentrations are low.

And as the acetyl-CoA concentrations are low so naturally the citric acid cycle is also down regulated and as a result the formation of the reducing equivalent or the formation of the Gibbs free energy etc., are all very low. Now at the same time when the anoxic condition is there in the system, some other interesting things also happen. This includes that the acetyl-CoA is not entering through TCA cycle but is now able to move towards acetate.

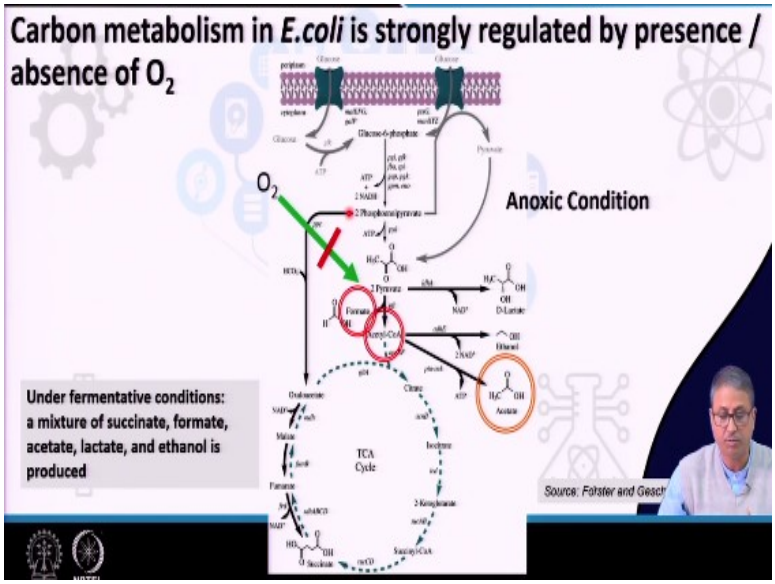
So acetate is formed during the anoxic condition or anoxic growth of the E. coli cell. So from oxic to anoxic you see the shift in product formation. When there was sufficient oxygen it was mainly the complete oxidation of the pyruvic acid or glucose molecule, but as the system enters into an anoxic condition, it is the acetate which starts accumulating in the system.

Now this anoxic condition also leads to activation of the pyruvate formate lyase, which is a very important enzyme during anoxic condition for E. coli that allows the

conversion of pyruvate to formate and acetyl-CoA. Now on the one hand E. coli, which is growing under anoxic condition has the down regulated pyruvate dehydrogenase complex resulting the formation of acetate and TCA cycle is already down regulated.

At the same time the pyruvate formate lyase enzyme is activated under this anoxic condition resulting into conversion of pyruvic acid to formate and to some extent of acetyl-CoA.

(Refer Slide Time: 20:49)



And this acetyl-CoA is further adding into the acetate pool of the system. So now if we try to see that how the carbon metabolism is actually adjusted between the two conditions like oxic condition and anoxic conditions, we will be able to find the very sensitive situation of the pathway control.

In one case, the oxic condition, it is completely oxidized with large amount of Gibbs free energy and reducing equivalent while under anoxic condition it is primarily the accumulation of the sugars because of the oxidation is inhibited. The complete oxidation is rather inhibited through TCA cycle, but at the same time conversion of acetate, conversion of acetyl-CoA to acetate.

And this is also due to the fact that the enzyme responsible for a conversion of pyruvate to formate and acetyl-CoA is activated and at the same time the pyruvate dehydrogenase complex which was responsible for feeding the pathway towards TCA

cycle is down regulated. Now there could be a third condition. The third condition is basically the fermentative condition.

So under the fermentative condition, there could be a mixture of different compounds, which will be produced by these metabolic pathway that includes succinic acid because in that condition, the phosphoenolpyruvate can lead to production of more oxaloacetate through a different pathway. And then this oxaloacetic acid can be converted to succinic acid.

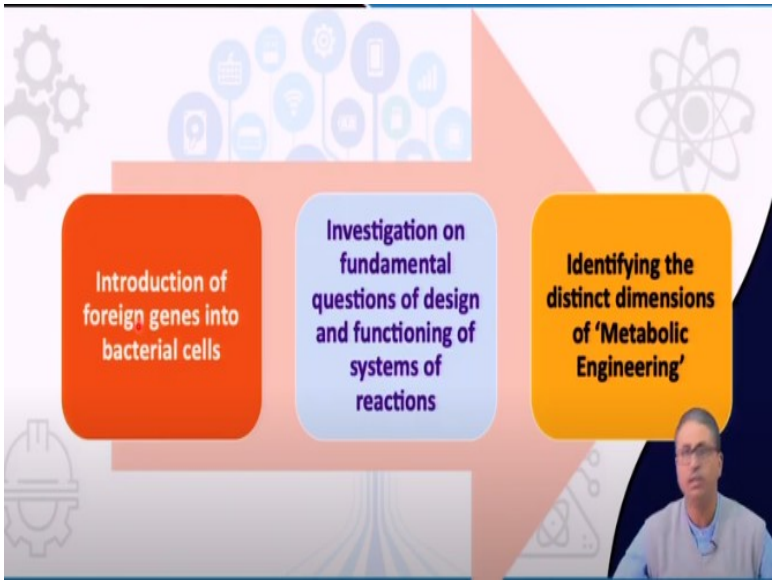
The pyruvic acid which is produced is converted to lactic acid by lactate dehydrogenase or the acetyl-CoA, which is produced due to the activity of pyruvate formate lyase can be converted to acetaldehyde and then ethanol. Or the acetyl-CoA can be converted to acetate. So we eventually have a mixture of succinic acid, formic acid, acetate, lactate and ethanol under the fermentative condition.

So now the overall behavior of the carbon metabolism with reference to conversion of glucose to different products under the three conditions reveal a very interesting, but at the same time complex nature of the metabolic reactions.

Now this information will be in the background of our mind, because from this type of interactive metabolic processes or metabolic pathways, which are also evident in different other part of the carbon metabolism or other metabolism including the amino acid metabolisms in *E. coli*, we see that the individual reaction or individual set of reactions are highly coordinated.

And they are also regulated by a number of growth factor or environmental factors like the presence of oxygen or absence of oxygen.

(Refer Slide Time: 24:04)

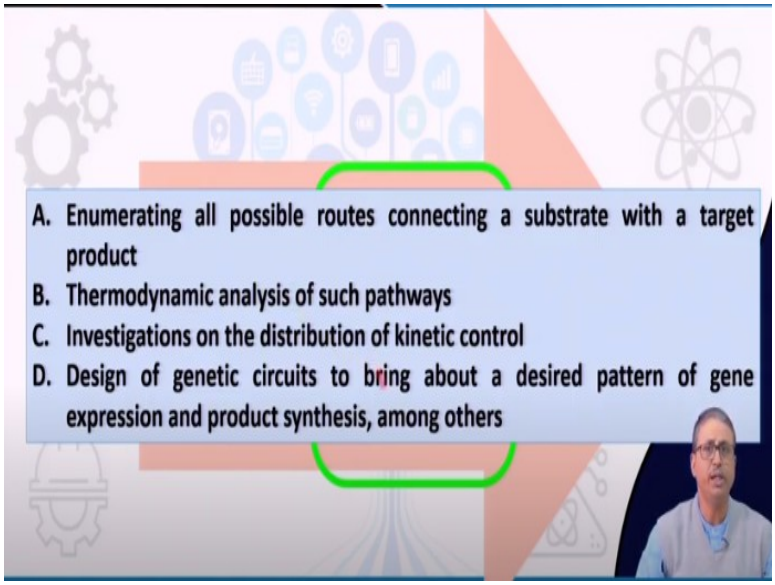


Now so the first part of the 1980s, 1985 we see a dramatic rise in application of recombinant DNA technology in getting superior strains or improving the product concentration within the natural host organisms like *E. coli* as you can see with the introduction of different foreign genes into the bacterial cells through the established genetic engineering protocols.

Now very soon, we encountered failures or the scientist encountered failures. Why the scientist encountered failures or the results were not up to the mark? Because we started realizing that the metabolic pathways within the systems like the *E. coli* systems or other microbial systems are highly complex and coordinated and they represent a part of a very complex network of metabolic reactions or what we refer as metabolic network.

So instead of considering individual reactions or genes responsible for individual reactions, emphasis was soon given to understanding the fundamental issues or fundamental questions about the how do we then design and improve and then work on the functioning of the network of metabolic reactions. So with respect to investigating the fundamental questions related to these networks of metabolic pathways or metabolic networks, a set of very well defined questions are proposed.

(Refer Slide Time: 25:57)



These questions include enumerating the all possible routes connecting a substrate with a target product. So there could be numerous routes, it is not that you have a substrate like A and then the first intermediate is B and then B is going to be converted to your desired product with two or three step reactions.

There could be more than one or maybe more than several numbers of reaction pathways, which could lead or which might lead to the product formation from the same set of substrates. So enumerating all the possible routes, connecting the substrate to the product of the target molecule was one of the target area or the main area of research with respect to understanding of the first step in understanding the metabolic network.

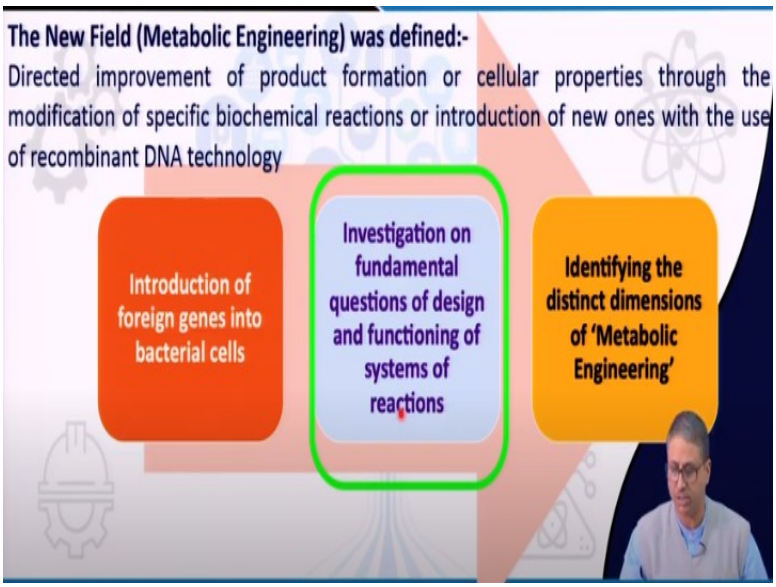
The second is the thermodynamic analysis of such pathways. We may have for example, from glucose to ethanol, five pathways. We may just for example, that we may have five pathways in a particular bacterial system like an E. coli system or three pathways. But out of these three pathways, what are the thermodynamic criteria for all these pathways?

Next comes the investigation on the distribution of kinetic controls regarding each of these pathways, more than one pathways are involved for particular product to a particular substrate to product formation. Then how the kinetic controls are distributed among these different set of pathways, which are actually responsible for producing the same product, but through different pathways.

And then finally, design of genetic circuits based on this information, which can bring about a desired pattern of gene expression and product synthesis among others. So a systematic knowledge about all these steps starting from enumerating the possible, all the possible routes through which this particular substrate can be converted to a particular product, evaluating the thermodynamics and also the kinetic control.

And based on these thermodynamic analysis and kinetic control data, genetic circuits or genetic maps or genetic backgrounds can be developed or should be developed, which will be used to bring about the desired pattern of gene expression and then the product synthesis within the organism.

(Refer Slide Time: 28:48)

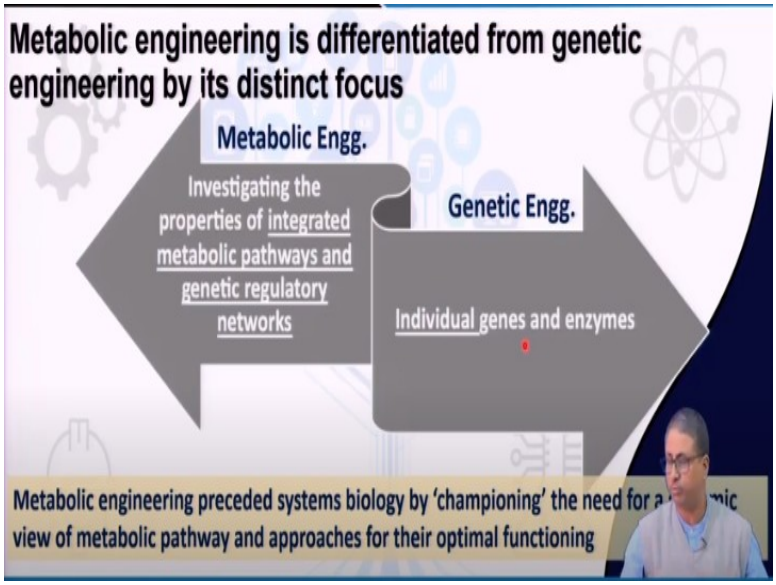


Now as soon as this particular criteria of incorporating these questions of metabolic networks, the fundamental questions related to metabolic networks were incorporated into this concept of production or overproduction of molecules or chemical compounds through microbial hosts, we entered into a new dimension of metabolic engineering.

Before that, we had this kind of impression that or a definition in our mind, because 1991 this definition was coined that directed improvement of product formation of cellular properties and the requirements for this thermodynamic feasibility or the kinetic control and identifying all possible routes were going on maybe just before that, before coining the term metabolic engineering.

So when we define metabolic engineering, actually, the metabolic engineering as a scientific field or a research area or subject emerged with multiple distinct dimensions.

(Refer Slide Time: 29:57)



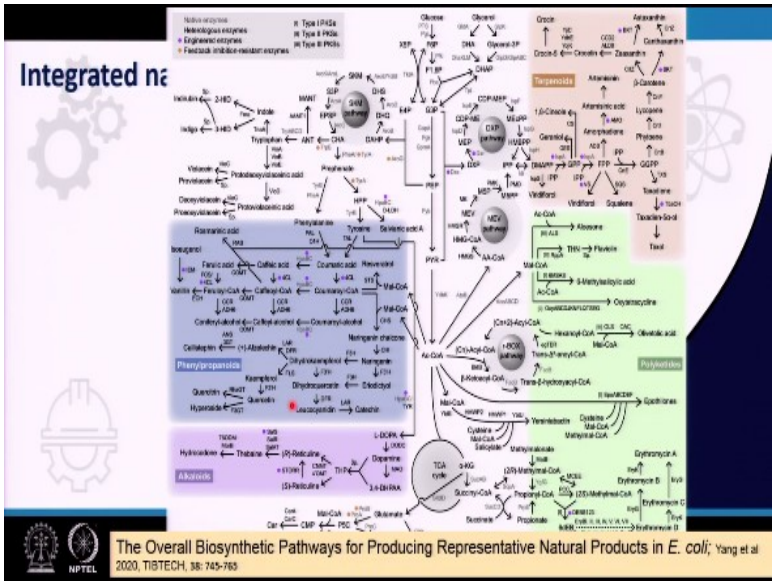
So one of the major important aspect of these multiple dimensions is whether it is just a manifestation of genetic engineering or an application of genetic engineering or something more than that. So metabolic engineering is distinctly different from genetic engineering because of its focus. In metabolic engineering, the principle or main focus is on investigating the properties of integrated metabolic pathways, the metabolic networks and genetic regulatory networks.

So there could be actually multiple networks involved in controlling the formation of a product or a particular set of products or a number of products or physiological conditions, but we need to investigate those integrated metabolic pathways and the genetic regulatory networks which are controlling or responsible for such product formation or such phenotypic properties.

Whereas the genetic engineering is basically based on individual genes and enzymes. It rarely addresses the integrated nature of the metabolic network or the regulatory network as well. So metabolic engineering, in that sense, actually preceded the systems biology by championing the need for a systemic view of metabolic pathway and approaches for their optimal functioning.

So instead of looking into a single gene or a couple of genes responsible for producing a particular product out of a substrate, we started looking into all the reactions involved or likely to be involved or controlling the production of a particular product including the regulatory controls also. Because there could be many reactions, which are not directly involved in the product formation, but they are indirectly involved and might be having a very strong role, because they are the part of the different regulatory modules.

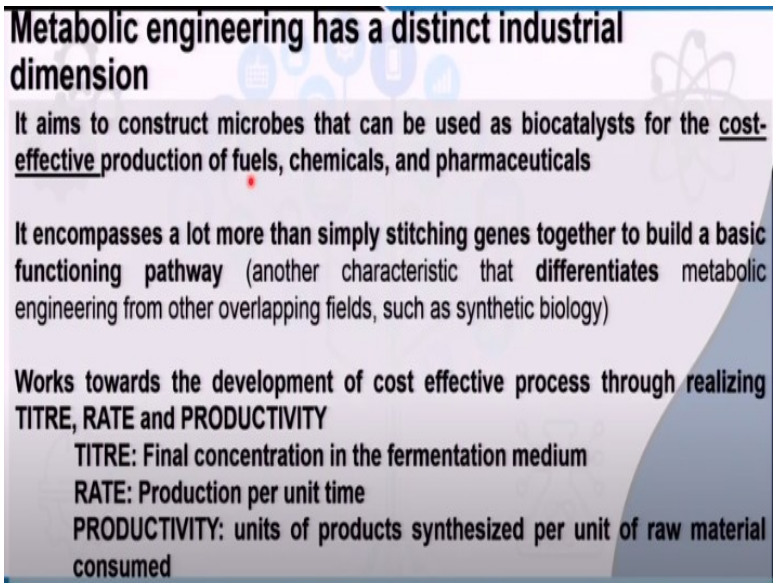
(Refer Slide Time: 32:08)



Now the integrated nature of metabolic pathways are revealed. So this integrated metabolic nature of metabolic pathways are visualized and represented by many investigators and many authors. And this particular image or diagram is portraying actually the overall biosynthetic pathway for producing representative natural products. As you can see here, here is the central metabolite like acetyl-CoA or the pyruvic acid.

And from this central metabolite like acetyl-CoA a numerous metabolic pathways or a large number of metabolic reactions are starting and they are leading to different type of molecular production including the terpenoids production, polyketides production, phenylpropanoid production, alkaloid production, and of course the ethanol and other butanol and other molecule production.

(Refer Slide Time: 33:05)



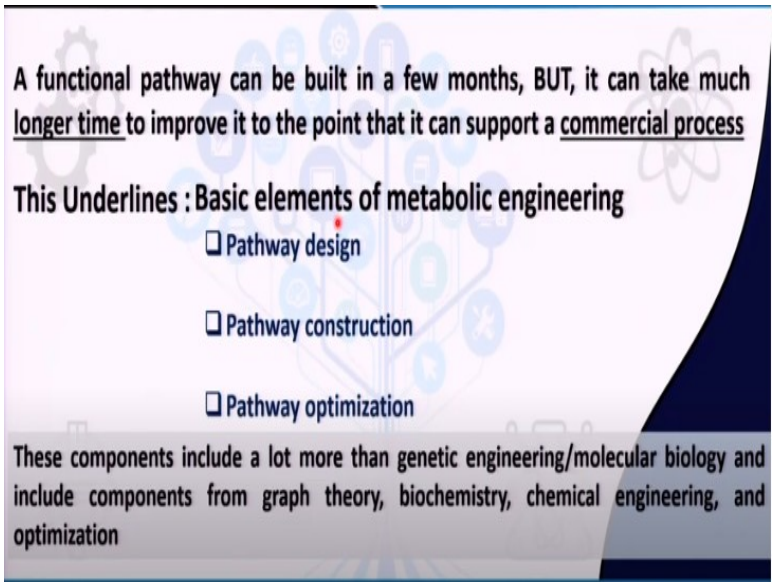
So metabolic engineering is not only different from recombinant DNA technology based approaches or the genetic engineering, but it has also a distinct industrial dimension. It aims to construct microbes that can be used as biocatalyst for the cost effective production of fuel, chemical or pharmaceuticals or any other molecules of importance.

So one of the major point over here that we want to develop the microbial biocatalyst for the cost effective production of the molecules, it is not about stitching a few genes together and build a basic functioning pattern.

Stitching a few genes through genetic engineering tools could be a very attractive research topic, but we need to understand that for metabolic engineering, we need to address the cost effectiveness of the entire process, particularly in terms of three important parameters which are called the titre, rate and productivity. The titre is the final concentration in the fermentation medium of the growth medium, where the organism is growing.

Rate is the production of the compound the target product molecule per unit time. And productivity is the units of product synthesized per unit of raw materials consumed. So when we incorporate the concepts of these titre, rate and productivity into our planning and designing, and then implementing those plans for improvement of the strains, we make progress towards a kind of cost effective implementation of those improved strains or organisms.

(Refer Slide Time: 34:58)



Now a functional pathway can be built in a few months like the stitching of genes from different organisms or suitable host organism can be done in a few months' time. But it can take much longer time to improve it to the point that it can support a commercial process.

So we have seen that in most of the cases the commercial production has taken sufficient time like 8 to 10 years' time, because lot of parameters, lot of factors were standardized to attain the titre, rate and productivity.

Now this underlines, this particular point that it is not merely joining few genes through a plasmid vector or through using some other recombinant DNA technology method, it is actually addressing towards the, cost effective use of those recombinant organism or those biocatalyst that we are trying to build for the production of a particular product or for the improvement of some physiological properties.

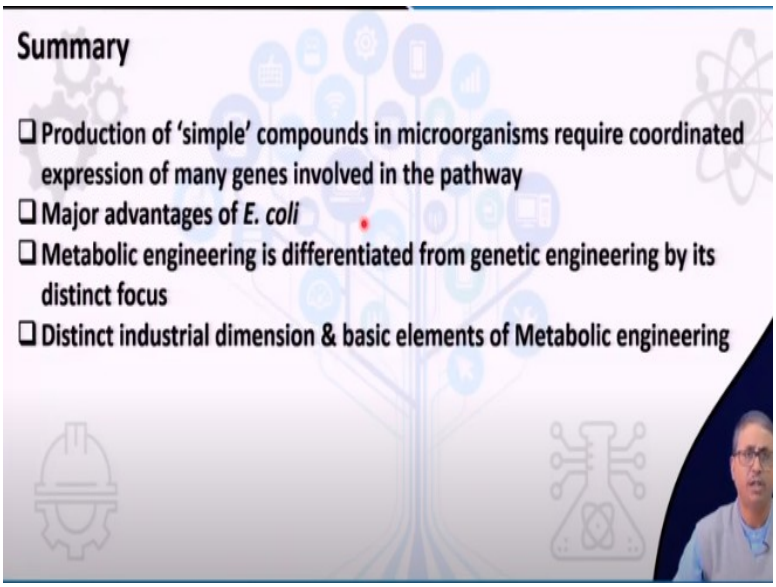
These includes three basic elements of metabolic engineering. These are pathway design, pathway construction and pathway optimization. So this the change in concept or the improvement in concept is very important. So from genetic engineering to metabolic engineering, and then within that, we also include the or we also cover the scope for the cost effective production of the molecule that includes the pathway design, pathway construction and pathway optimization.

These components include a lot more than genetic engineering and molecular biology analysis and include components from different and diverse fields including the graph theory, biochemistry, chemical engineering and optimization.

(Refer Slide Time: 36:55)

So with this, the today's lecture is ended and for this lecture, these are the following references which were used.

(Refer Slide Time: 37:07)



And to summarize today's lecture, here are the points like the production of simple compounds in microorganisms require coordinated expression of many genes involved in the pathway. Major advantages of *E. coli* are emphasized. Metabolic engineering is differentiated from genetic engineering by its distinct focus. And lastly,

the distinct industrial dimension and basic elements of metabolic engineering are also briefly discussed. Thank you.