MICROBIAL BIOTECHNOLOGY

Prof. Utpal Bora
Department of Biosciences and Bioengineering
Indian Institute of Technology Guwahati

Lecture-34
Lec34:Production of biopharmaceuticals(Enzymes,antibodies&therapeutic proteins)using rDNA technology

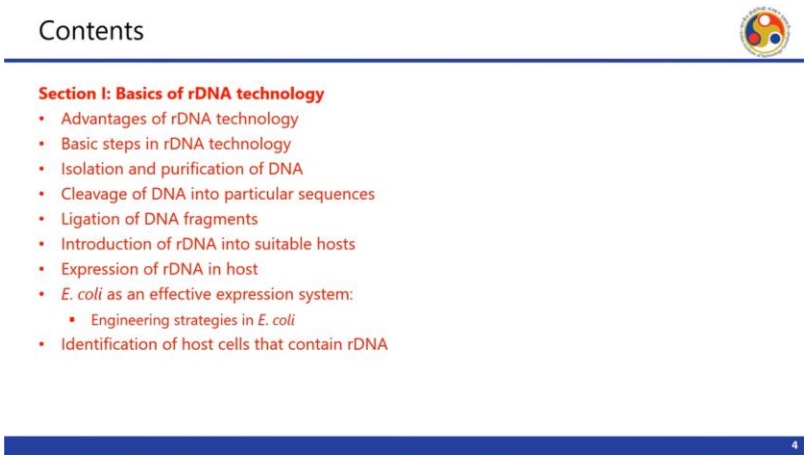
Hello friends, welcome back to my course on microbial biotechnology. Today, we are going to start module number 10, which is about microbes in medical biotechnology. In this lecture, we are going to discuss the production of biopharmaceuticals using recombinant DNA technology. We will mostly focus on the production of enzymes, antibodies, and other therapeutic proteins. This lecture is divided into two broad sections.

The first section will deal with the basics of recombinant DNA technology for those who are new to this field. Section number two will cover the application of this recombinant technology in pharmaceutical production. So, let's start with the basics of recombinant DNA technology. In this figure, you can see bacteria from Staphylococcus aureus, where phages are used as vectors in recombinant DNA technology.

What are these vectors? We will discuss them soon. This particular section has the following content. First, we will begin with a discussion about the advantages of using recombinant DNA technology. Then, we will cover the basic steps of this technology.

And then, isolation and purification of DNA, which is the most important first step in recombinant DNA technology. Then, cleavage of DNA into particular sequences—how do

we cut down these large DNA molecules into smaller fragments, and also vector molecules, because engineering is all about cutting and joining. Here, we will show you how we cut DNA of interest and paste it into carriers called vectors, which are also basically DNA molecules. After cutting these fragments, we paste them into these vectors. This is known as ligation.

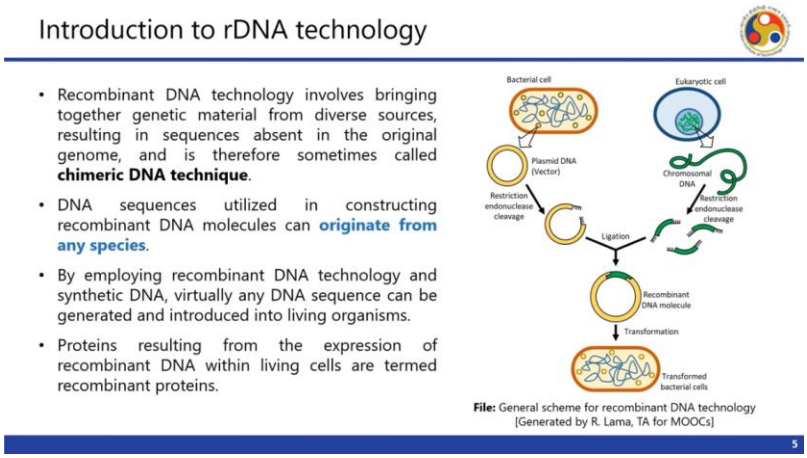


So, we'll discuss the ligation of these DNA fragments. Then, once this vector is constructed—which is basically a carrier containing a DNA or gene of interest—we need to introduce these recombinant DNA products into serial hosts. Hosts are basically living cells where these vectors will multiply, thereby also multiplying the recombinant DNA. For the production of protein therapeutics, the gene that we clone into the vector needs to be expressed—first transcribed and then translated. So, we'll also discuss the expression of recombinant DNA in the host.

Then, we will discuss E. coli as an effective expression system and the various engineering strategies involving E. coli, as well as the identification of host cells containing the recombinant DNA. So, let's start with a very basic introduction to recombinant DNA technology, or rDNA technology. This involves bringing together genetic material from diverse sources, resulting in sequences absent in the original genome, and is therefore sometimes called the chimeric DNA technique. DNA sequences used in constructing recombinant DNA molecules can originate from any species. By employing recombinant DNA technology and synthetic DNA, virtually any DNA sequence can be generated and introduced into living organisms.

Proteins resulting from the expression of recombinant DNA within living cells are termed recombinant proteins. So, let's have a look at this general scheme for recombinant DNA technology. So here we have a bacterial cell which has extra-chromosomal DNA called a

plasmid. Using these plasmids, we construct the vectors, or some of these are already ready-to-use vectors. Now, on the other hand, let us take an example of a eukaryotic cell which also has DNA, but this is the nuclear DNA, and we call it chromosomal DNA.



File: General scheme for recombinant DNA technology [Generated by R. Lama, TA for MOOCs]

Now, both this vector DNA and chromosomal DNA are subjected to restriction digestion with the help of restriction endonuclease. So here, you can see that due to the digestion by restriction enzymes, this vector has opened up. We made a cut in the vector. Similarly, this DNA will be digested into small pieces. Okay, and then by the restriction endonuclease activity.

Now, these small fragments of DNA can be inserted into this vector by a process called ligation. So here, these are the products as a result of this entire procedure, and these particular vector molecules carrying the gene of interest are known as recombinant DNA molecules. Now, these particular recombinant DNA molecules are sent inside a host cell, which is, for example, a bacterial cell, and by the process of transformation, we get the transformed bacterial cell. Now, these vectors having the loaded DNA inside may simply keep on multiplying because the plasmid factors are autonomously replicating units, and while the vector multiplies, the DNA which is cloned into it will also multiply.

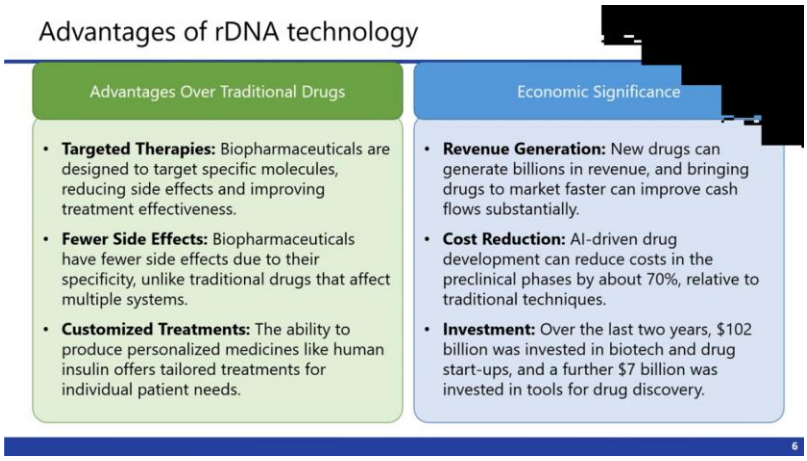
So, therefore, we will be able to increase the copy number of these genes of interest by this process. If we add certain DNA sequences, which help in the expression of the gene—for example, if it is a protein-coding gene—we can also further produce the protein resulting from this process. Depending on the interest or the needs, we may use simply cloning vectors, which help us increase copy numbers, or we may also use expression vectors, which help us first clone the DNA into the vector and then subsequently express the gene product. So, this is, in brief, the general scheme of recombinant DNA technology.

Now, what are the advantages of using recombinant DNA technology over traditional drugs? Let us focus on therapeutics only. For example, it is helpful in targeted therapies. Biopharmaceuticals are designed to target specific molecules, reducing side effects and improving treatment effectiveness.

There are fewer side effects. Biopharmaceuticals have fewer side effects due to their specificity. Unlike traditional drugs, which affect multiple systems. And then, we can also go for customized treatments. The ability to produce personalized medicines, like human insulin, offers tailored treatments for individual patients' needs.

And then there is also economic significance. For example, in the case of revenue generation, new drugs can generate billions in revenue and bringing drugs to market faster can improve cash flows substantially. Then we can also go for cost reduction. For example, AI-driven drug development can reduce costs in the preclinical context phases by about 70% relative to traditional techniques.

And then over the last two years, around $102 billion was invested in biotech and drug startups, and a further $7 billion was invested in tools for drug discovery. We have precision in gene manipulation when we are using this particular technology. Recombinant DNA technology allows for precise manipulation of genetic material, enabling the creation of specific gene sequences and the study of gene function. Then there is the advantage of scalability and cost effectiveness. The technology is scalable



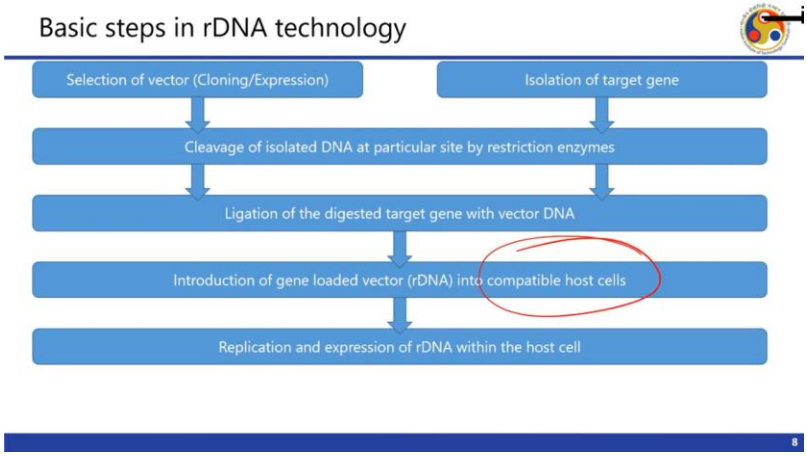
and cost-effective, making it possible to produce large quantities of proteins for research and therapeutic purposes. It offers versatility in protein engineering. RDNA technology enables the production of a wide range of proteins, including human insulin, epinephrine alpha, and human growth hormone, which are significantly improved treatments for

various medical conditions. It helps us in achieving targeted therapies. It has led to the development of targeted therapies, reducing side effects and improving treatment effectiveness, which I've also discussed earlier.

Then we are able to offer customized treatments. Now let us discuss the basic steps in RDNA technology. We have to select a vector. It may be a cloning vector or an expression vector. On the other hand, we also have to think about isolating the target gene.

Then we proceed with cleaving the isolated DNA at a specific site using a restriction enzyme. We make a cut in the vector and also need to digest the DNA chromosome where our gene of interest may be located. This has already been briefly described in the general scheme. Then we proceed with ligating the digested target gene with the vector DNA.

Next, we introduce the gene-loaded vector into a compatible host cell. This is very important. The host cell must be compatible to receive the gene-loaded vector. There are certain treatments by which we can make the host cells compatible or competent. This step is also known as the competent cell preparation step.

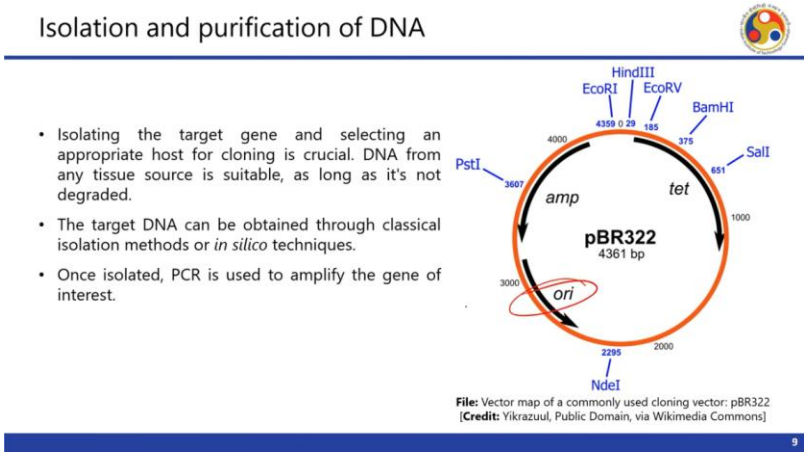


Then we go for replication and expression of our DNA within the host cell. We also have to then go for screening of the transformed cells because not all of them will be transformed, which means not all host cells, even though they are competent, may be able to receive the gene-loaded vector. So, we need to separate the host cells that have the gene-loaded vector from those that do not. In subsequent steps, we will use only the cell population that has the gene-loaded vector.

So, we start with the isolation and purification of DNA. Isolating the target gene and selecting an appropriate host for cloning is crucial. DNA from any tissue source is suitable as long as it is not degraded. Then we have the target DNA, which can be obtained through

classical isolation methods by established protocols, or we can also design the DNA sequence in silico and subsequently synthesize it in the laboratory. We obtain the DNA either through traditional isolation techniques or by in silico and in vitro synthesis.

Then we perform polymerase chain reaction to amplify the gene of interest. This figure shows the map of a commonly used cloning vector, pBR322. Some of the important features here are that each vector has an origin. This is the point or sequence from which the replication of the vector starts. Then, it has many marker genes.



For example, this is an ampicillin-resistant gene, and this is the tetracycline-resistant gene. These antibiotic resistance genes help us in screening the vector-loaded host cells. So, we use this in various steps: first, for the successful entry of this vector into the host cell, and then for the selection of a gene-loaded vector inside a host. Then, you can see here certain sites for restriction enzymes. You have ECO R5 here.

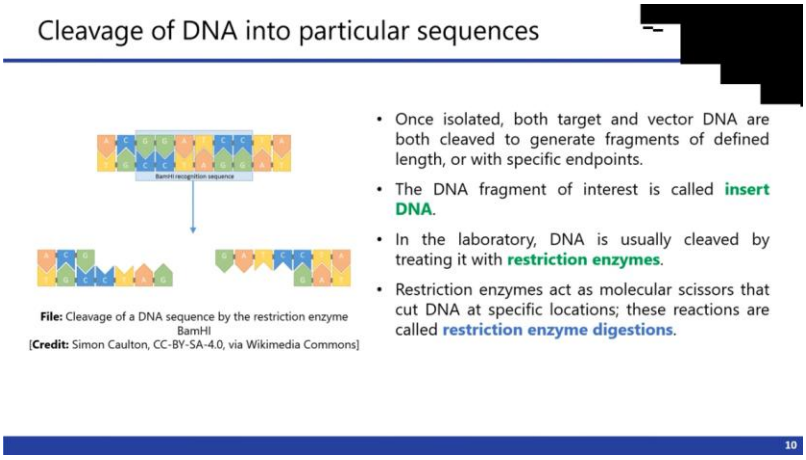
Then, you have ECO R1, BAMH1, SALH1, then HIN3, and PST1. So, these are the sites along with the numbers. So, these particular sequences will cut the vector at these particular sites to open it up. So, we may use a BAMH1 single enzyme to open it up. So, if we use BAM H1, the vector will open at this site at 375.

But we can also use a combination of restriction enzymes. For example, if we use ECO R1, and SALH1, we will have a vector ranging from 4359 to 651, and then we will have another fragment here, this smaller fragment. So, this fragment will be lost because it does not have the origin, but this fragment will be useful for us. So, we create a big gap over here, and in this gap, we can clone a very large gene sequence.

So, this is basically what a vector is all about—the various restriction sites, the antibiotic-resistant genes which help us in the screening process, as well as the origin of replication.

So, let us look into the cleavage of DNA into particular sequences. So here, the example is for the BamHI recognition sequence, which is basically a palindromic sequence. So, when BamHI cuts any given DNA, it will create a site like this, as you can see over here. So, once isolated, both target and vector DNA are cleaved to generate fragments of defined length or with specific endpoints, as I was showing you—from 4300 to somewhere around 400 or 300.

The DNA fragment of interest is called the insert DNA, and then the DNA is cleaved by treating it with the restriction enzyme, which I have already shown you. These restriction enzymes act as molecular scissors that cut DNA at specific locations. These reactions are called restriction enzyme digestions. Then, we go for the ligation of the DNA fragments. Target DNA is treated with a restriction enzyme to generate fragments, which are then mixed with vector DNA pretreated with the same enzyme and DNA ligase. So here, due to the complementarity of these two ends over here, they will pair with one another, and then the two molecules will be joined by DNA ligase.



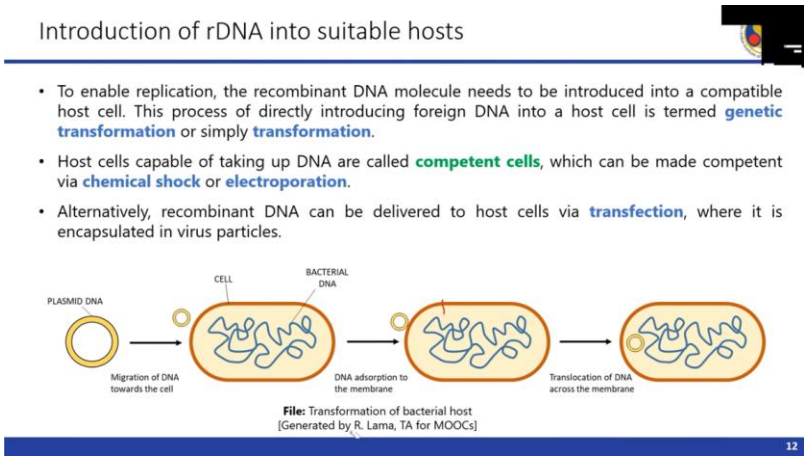
File: Cleavage of a DNA sequence by the restriction enzyme BamHI
[**Credit:** Simon Caulton, CC-BY-SA-4.0, via Wikimedia Commons]

DNA ligase covalently joins the DNA fragments and vector DNA through 3' to 5' phosphodiester bonds. The ligation mixture contains various interconnected DNA molecules, including the vector bonded to the target DNA. Once the vector is loaded with the target DNA or insert DNA, we proceed with the introduction of the rDNA into a suitable host. To enable replication, the recombinant DNA molecules need to be introduced into a compatible host cell. This process of directly introducing foreign DNA into a host cell is termed genetic transformation or simply transformation.

So, this is the plasmid DNA, and there is basically loaded DNA over here. Then this plasmid will migrate toward the cell, and the DNA will adsorb to the membrane here, as you can see, and then it will enter by the process of translocation across the membrane. So,

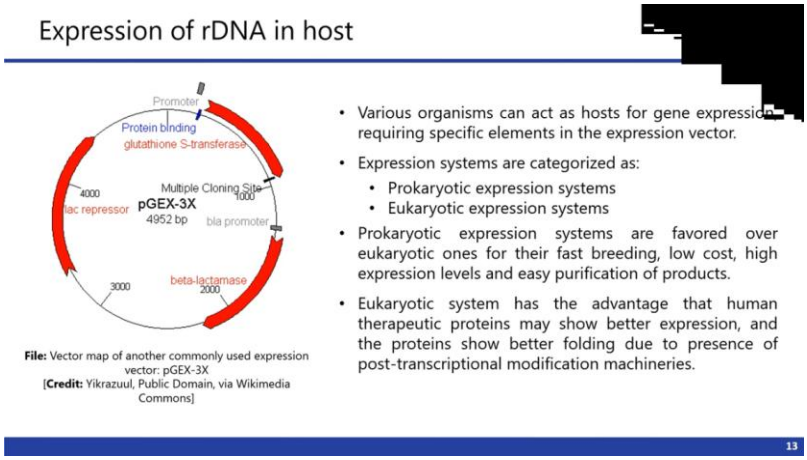
many times this is difficult, so we have to create the conditions for plasmid DNA to enter this host cell. So, host cells capable of taking up DNA are called competent cells.

So, one of the facilitations we do is treat the cells with certain chemicals which make them competent via chemical shock or even electroporation. So, when we give a very high voltage for a short pulse or a short time, or even a low voltage for a slightly extended time, because of the fluidity of the cell membrane, there will be an opening or creation of pores in the membrane, and these plasmids can enter through these small pores into these competent host cells. Alternatively, recombinant DNA can be delivered to host cells via transfection, where it is encapsulated in virus particles. So, let us now study a little bit about the expression of our DNA in the host. So, various organisms can act as hosts for gene expression, requiring specific elements in the expression vector.

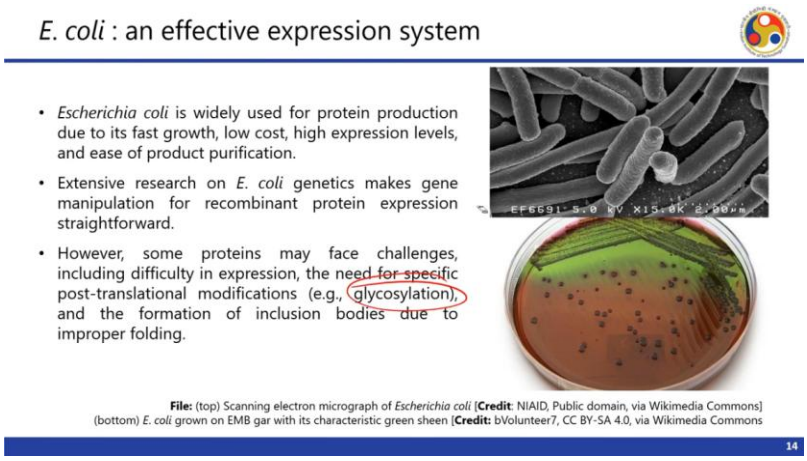


So, this is a vector map of a commonly used expression vector, Psex3x. So, here we basically have a promoter, which is very essential for an expression vector. Then we have a multiple cloning site over here, and this is a marker gene for beta-lactamase, and this is a lac repressor. So, these are some of the basic features of an expression vector, as shown here. We can classify the expression systems as prokaryotic expression systems or eukaryotic expression systems.

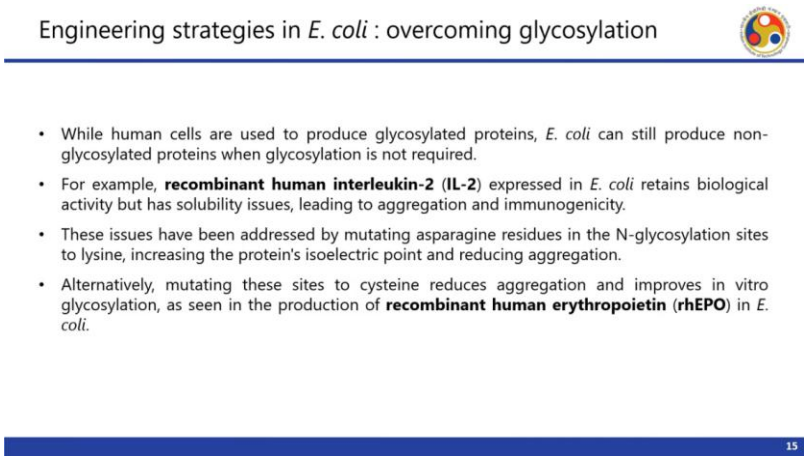
Prokaryotic expression systems are favored over eukaryotic ones for fast breeding, low cost, high expression levels and easy purification of products. Eukaryotic systems have the advantage that human therapeutic proteins may show better expression and the proteins show better folding due to the presence of post-transcriptional modification machinery, which is absent in prokaryotic expression systems. So, let us look into the application of E. coli as an effective expression system. Here you can see the image of E. coli taken under a scanning electron microscope.



As you can see, E. coli is widely used for protein production due to its fast growth, low cost, high expression levels, and ease of product purification. Extensive research on E. coli genetics makes gene manipulation for recombinant protein expression simple and straightforward. However, some proteins may face challenges, including difficulty in expression, the need for specific post-transcriptional modifications, for example, glycosylation and the formation of inclusion bodies due to improper folding. So, we have certain engineering strategies in E. coli for overcoming glycosylation.

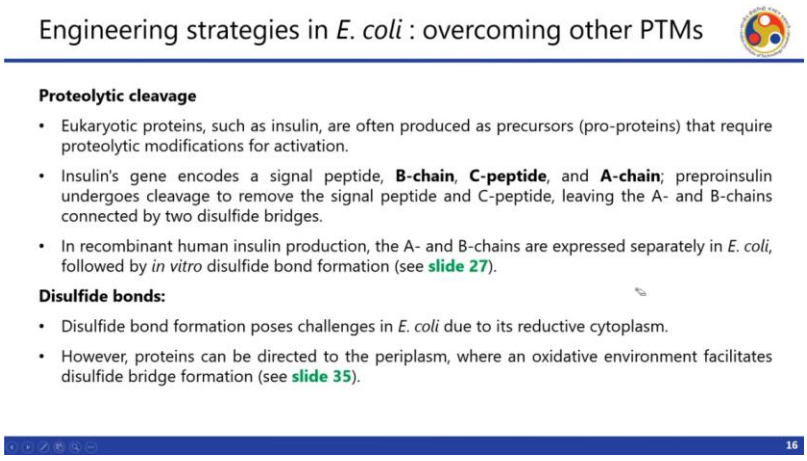


While human cells are used to produce glycosylated proteins, E. coli can still produce non-glycosylated proteins when glycosylation is not required. For example, we have recombinant human IL-2 expressed in E. coli which retains biological activity but has solubility issues leading to aggregation and also immunogenicity. This issue has been addressed by mutating asparagine residues in the N-glycosylation sites to lysine, increasing the protein's isoelectric point and thereby reducing aggregation. Alternatively, mutating these sites to cysteine reduces aggregation and improves in vitro glycosylation, as in the production of recombinant human erythropoietin in Escherichia coli.

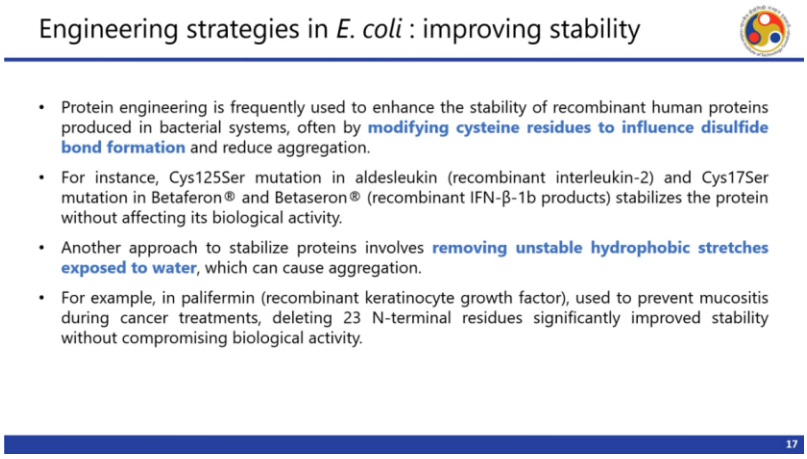


Let us now discuss other engineering strategies for overcoming other kinds of post-translational modifications. Let us begin with proteolytic cleavage. Eukaryotic proteins such as insulin are often produced as precursors, which we call pro-proteins, that require proteolytic modification for activation. The insulin gene encodes a signal peptide, beta-chain, a B-chain, C-peptide, and A-chain. Pre-pro-insulin undergoes cleavage to remove the signal peptide and C-peptide, leaving the A and B-chains connected by two disulfide bridges.

In recombinant human insulin production, the A and B chains are expressed separately in E. coli, followed by in vitro disulfide formation, which we will discuss in a later slide. Another important PTM is disulfide bond formation. This poses a challenge due to the reductive cytoplasm. However, proteins can be directed to the periplasm, where an oxidative environment facilitates disulfide bridge formation. We will discuss this further in slide number 35.



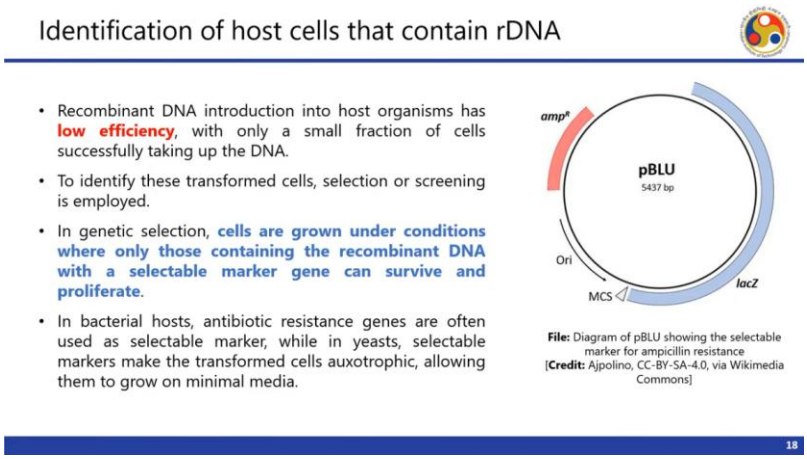
Protein engineering for improving stability is another engineering strategy. Protein engineering is frequently used to enhance the stability of recombinant human proteins produced in bacterial systems, often by modifying cysteine residues to influence disulfide bond formation and thereby reduce aggregation. For example, the Cys125-to-serine mutation in recombinant interleukin-2 and the Cys17-to-serine mutation in beta-interferon, which are recombinant IFN beta-1B products, stabilize the protein without affecting its biological activity. Another approach to stabilize proteins involves removing unstable hydrophobic stretches exposed to water, which can cause aggregation. For example, in palifermin, which is a recombinant keratinocyte growth factor used to prevent mucositis during cancer treatments, deleting 23 N-terminal residues significantly improves stability without compromising biological activity.



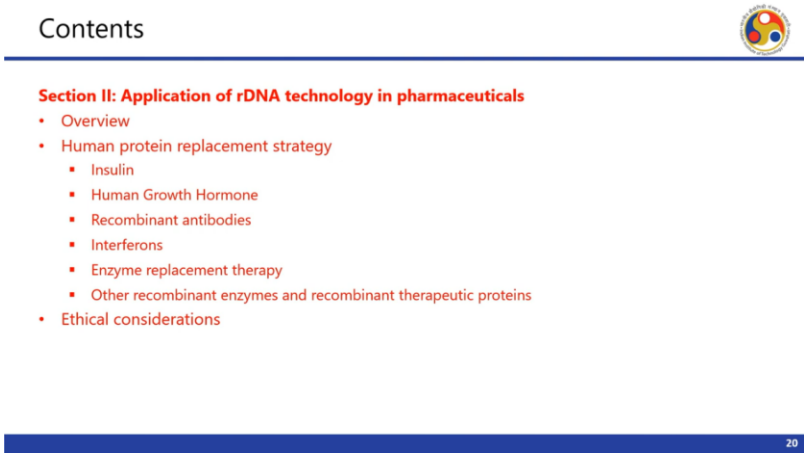
The next important step is the identification of host cells that contain our DNA. After the vector is loaded with the gene of interest and the recombinant DNA is introduced into the host cell by the method of transformation or transfection, it is essential to identify the host cells that contain this recombinant DNA molecule. So, this recombinant DNA introduction into the host organism is a low-efficiency process, with only a small fraction of cells successfully taking up the DNA. So, we have to separate out these transformed cells, which we will multiply and increase in population. So, for doing this, we need to go for a selection or screening process.

Now, here in this genetic selection, cells are grown under conditions where only those containing the recombinant DNA with a selectable marker gene can survive and proliferate. For example, here we are using ampicillin resistance, so these media will have ampicillin. And the cells which have this vector will be able to survive in this media. So, bacterial host antibiotic genes are often used as the selectable markers, while in yeast, selectable markers

make the transformed cells auxotrophic, allowing them to grow on minimal media. Let us now discuss the application of rDNA technology in pharmaceuticals.



In this section, we have an overall overview of the application of rDNA technology in pharmaceuticals. Then we discuss the human protein replacement strategy, like insulin, human growth hormone, recombinant antibodies, interferons, enzyme replacement therapy, and other recombinant enzymes and therapeutic proteins. And we'll also have a small discussion on the ethical considerations because this is all about human therapy. So, in this picture, we can see Novolin is a kind of insulin. And then we have Twinrix, a vaccine comprising an inactivated hepatitis A vaccine and a recombinant hepatitis B vaccine.

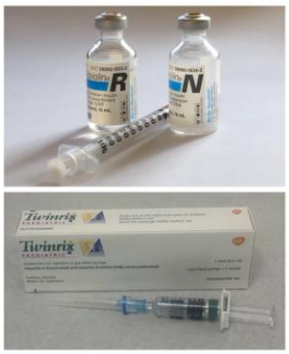


So, recombinant DNA technology has revolutionized medicine by enabling the production of over 400 recombinant proteins, with around 50 approved for human use already. This technology is also used to identify new genes, understand their roles in diseases, and in gene therapy to treat genetic disorders, cancer, and other diseases. In India, nine products have been approved for marketing, such as insulin; alpha interferon, which is a cancer drug;

hepatitis B vaccine; GM-CSF; G-CSF; blood clotting factor 7; erythropoietin, a drug used in kidney failure; streptokinase, a drug administered in heart attacks; and human growth hormone. Some of the key milestones in recombinant drug development started in 1982 when recombinant human insulin was made, which was the first recombinant biologic to be approved. And insulin, as we know, is a therapy for diabetes.

Overview

- Recombinant DNA technology has revolutionized medicine by enabling the production of over 400 recombinant proteins, with around 50 approved for human use.
- This technology is also used to identify new genes, understand their role in diseases, and in gene therapy to treat genetic disorders, cancer, and other diseases.
- In India, nine products have been approved for marketing – **insulin** (diabetes drug), **alpha interferon** (cancer drug), **hepatitis B vaccine**, **GMCSF**, **G-CSF**, **blood clotting factor 7**, **erythropoietin** (drug used in kidney failure), **streptokinase** (drug administered in heart attacks) and **human growth hormone**.



File: (top) Novolin®, an insulin medication [**Credit:** Melissa Johnson, CC-BY-2.0, via Flickr] (bottom) Twinrix, a vaccine comprising of inactivated hepatitis A vaccine and recombinant hepatitis B vaccine[**Credit:** Whispyhistory, CC-BY-SA-4.0, via Wikimedia Commons]

21

Then, four years later, muromonab CD3, which was the first approved monoclonal antibody (or murine antibody), was used for the prevention of organ rejection. Then, in 1989, epoetin was used for the treatment of anemia. This was the first erythropoietin to be approved. Then, Filgrastim in 1991 was the first colony-stimulating factor to be approved, used for chemotherapy-induced neutropenia. Then, we have Abciximab, which was the first chimeric monoclonal antibody used against glycoprotein 2B/3A for antiplatelet therapy.

Key milestones in recombinant drug development (*cont...*)

Date	Product	Target	Indication
1982	Recombinant human insulin	Diabetes	First recombinant biologic approved
1986	Muromonab-CD3	Prevention of organ rejection	First approved monoclonal antibody; murine antibody
1989	Epoetin	Treatment of anemia	First erythropoietin approved
1991	Filgrastim	Chemotherapy-induced neutropenia	First colony stimulating factor approved
1993	Abciximab	Glycoprotein IIb/IIIa antiplatelet therapy	First approved, chimeric monoclonal antibody
1997	Daclizumab	Anti-rejection drug	First approved, humanized monoclonal antibody
1998	Infliximab	Biologic disease modifying anti-rheumatic drug	First approved, tumor necrosis factor inhibitor (chimeric antibody)

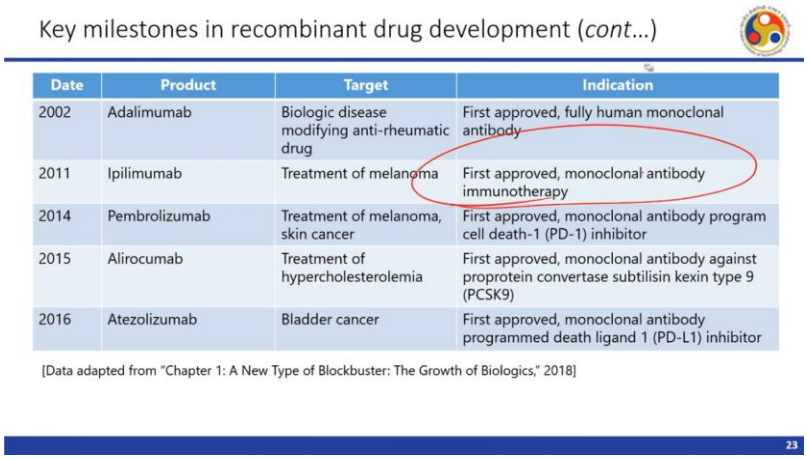
22

Then, we have Daclizumab in 1997, which was the first approved humanized monoclonal antibody. So, this is a murine antibody, and this is a humanized antibody. So, a lot of the

sequences— Similar to human genetic sequences—have been replaced here. This is basically an anti-rejection drug.

Then we have infliximab. This is the first approved tumor necrosis factor inhibitor, and it is a chimeric antibody. This is a therapy for biologic disease-modifying anti-rheumatic drugs in 1998. Then we have, in 2001, adalimumab, the first approved fully human monoclonal antibody. So, this is a humanized monoclonal antibody, but this one is a fully human monoclonal antibody.

So, we have ipilimumab, which is the first approved monoclonal antibody immunotherapy. Then we have pembrolizumab, which is a treatment for melanoma skin cancer, and it is the first approved monoclonal antibody programmed cell death-1 (PD-1) inhibitor. Then we have Alirocumab, which is a treatment for hypercholesterolemia. This is the first approved monoclonal antibody against proprotein convertase subtilisin/kexin type 9.



Key milestones in recombinant drug development (*cont...*)

Date	Product	Target	Indication
2002	Adalimumab	Biologic disease modifying anti-rheumatic drug	First approved, fully human monoclonal antibody
2011	Ipilimumab	Treatment of melanoma	First approved, monoclonal antibody immunotherapy
2014	Pembrolizumab	Treatment of melanoma, skin cancer	First approved, monoclonal antibody program cell death-1 (PD-1) inhibitor
2015	Alirocumab	Treatment of hypercholesterolemia	First approved, monoclonal antibody against proprotein convertase subtilisin kexin type 9 (PCSK9)
2016	Atezolizumab	Bladder cancer	First approved, monoclonal antibody programmed death ligand 1 (PD-L1) inhibitor

[Data adapted from "Chapter 1: A New Type of Blockbuster: The Growth of Biologics," 2018]

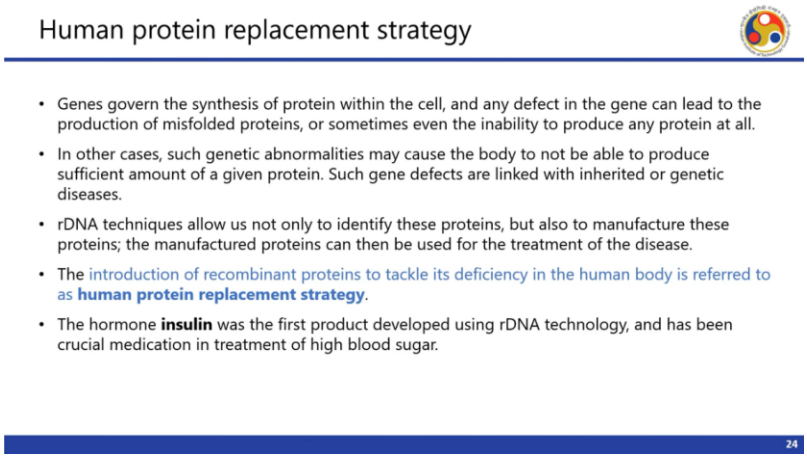
23

Then, in 2016, we have atezolizumab, which is a therapy against cancer. And this is the first approved monoclonal antibody programmed death-ligand inhibitor. So, this table has been adapted from 'A New Type of Blockbuster: The Growth of Biologics,' published in 2018. So, we can see from 1982 to 2016 and even now, later on, many antibodies, which are all recombinant, produced by recombinant DNA technology, have been coming to the market regularly.

This actually showcases the importance of this technology, our DNA technology in pharmaceuticals. Now let us discuss the human protein replacement strategy. Genes govern the synthesis of proteins within the cell, and any defect in the genes can lead to the production of misfolded proteins or sometimes even the inability to produce any protein at all. In other cases, such genetic abnormalities may cause the body to be unable to produce

a sufficient amount of a given protein. Such gene defects are linked with inherited or genetic diseases.

rDNA techniques allow us not only to identify these proteins but also to manufacture them. The manufactured proteins can then be used for the treatment of the disease, and you can already see many of those in the earlier slide. The introduction of recombinant proteins to tackle this deficiency in the human body is referred to as the human protein replacement strategy. The hormone insulin was first developed using rDNA technology and has been a crucial medication in the treatment of high blood sugar. In medical terms, insulin refers to any pharmaceutical preparation containing the hormone insulin, which is used to treat high blood glucose conditions such as type 1 diabetes, type 2 diabetes, gestational diabetes, complications of diabetes, and hyperkalemia, which is high blood potassium levels.



So, this is Humulin R, and this is Ectropede. These are two synthetic insulin drugs that are available in the market. Traditionally, before these synthetic drugs, insulin was derived from pigs and cows and was used for the treatment of high blood sugar levels. But this has now been replaced with human insulin production using rDNA techniques. In 1983, Humulin, which is a brand name for recombinant human insulin produced in E. coli, developed by Genentech and licensed to Eli Lilly, became the first rDNA drug to be manufactured.

Insulin

- In medical terms, insulin refers to any pharmaceutical preparation containing the hormone **insulin**, used to treat high blood glucose conditions, such as type 1 diabetes, type 2 diabetes, gestational diabetes, complications of diabetes, and hyperkalemia (high blood potassium levels).
- Traditionally, insulin derived from pigs and cows were used for treatment of high blood sugar levels, but this has now been replaced with human insulin produced using rDNA techniques.
- In 1983, **Humulin®**, brand name for recombinant human insulin produced in *E. coli*, developed by Genentech and licensed to Eli Lilly, became the first rDNA drug to be manufactured.

File: Humulin® and Actrapid®, two synthetic insulin drug available in the market
[**Credit:** Wesalius, CC-BY-4.0, via Wikimedia Commons]

25

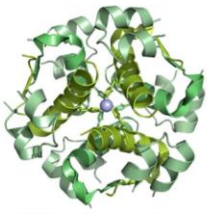
We have various types of insulin. Variations of the human insulin molecule, called insulin analogs, are also now available as alternative medications. Those bind to insulin receptors, bringing about the same reactions as insulin. Many types of insulin are widely available as medications. Number one is fast-acting or rapid-acting, which includes insulin analogs such as Aspart, Lispro, and Glulisine.

They take effect within 5 to 15 minutes and remain active for 3 to 4 hours. Then we have short-acting, which is basically regular insulin, which starts working in about 30 minutes and lasts for 5 to 8 hours. Then we have intermediate-acting, which includes NPH insulin, which becomes active within 1 to 3 hours and has a duration of 16 to 24 hours. Then we have long-acting insulin, which includes analogs like Glargine U100 and Detemir. These begin working in one to two hours and provide steady activity for around 24 hours, though this can vary between individuals.

Types of insulin

- Variations of human insulin molecule, called insulin analogues, are also now available as alternate medications: **these bind onto insulin receptors** bringing about the same reactions as insulin.
- Four main types of insulin are widely available as medications:

1. **Fast-acting** (**Rapid-acting**), includes insulin analogs such as ***aspart***, ***lispro***, and ***glulisine***. They take effect within 5 to 15 minutes and remain active for 3 to 4 hours.
2. **Short-acting**, refers to regular insulin, which starts working in about 30 minutes and lasts for 5 to 8 hours.
3. **Intermediate-acting**, includes ***NPH insulin***, which becomes active within 1 to 3 hours and has a duration of 16 to 24 hours.
4. **Long-acting**, includes insulin analogs like ***glargine U100*** and ***detemir***, which begin working in 1 to 2 hours and provide steady activity for around 24 hours, though this can vary between individuals.



PyMol diagram of an insulin lispro dodecamer
[**Credit:** Fvasconcellos, Public domain, via Wikimedia Commons]

26

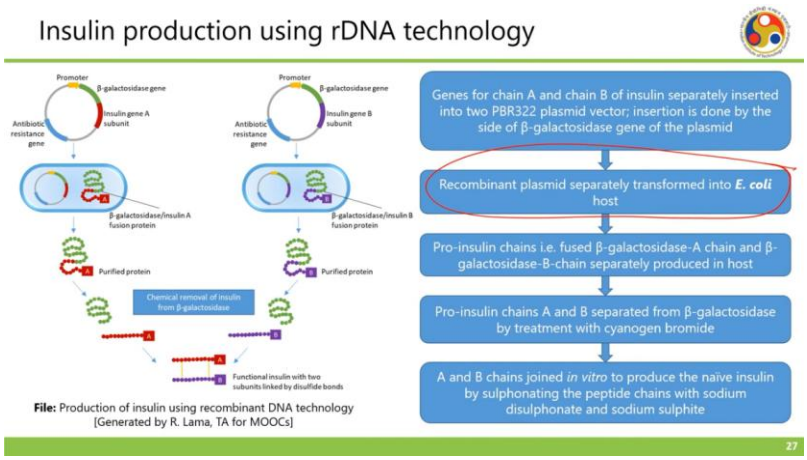
This is a PIMOL or computer-generated diagram of an insulin Lispro duodecamer, which is basically a model and not a real picture of a protein. So, how is insulin produced using

our DNA technology? Diabetes is a very common disease nowadays across the globe, and in India, the cases of diabetes are increasing every day. In fact, we have a large diabetic population, and India is often referred to as the diabetes capital of the world. The current approach for insulin-dependent diabetes is the application of synthetically prepared insulin, basically through the process of our DNA technology.

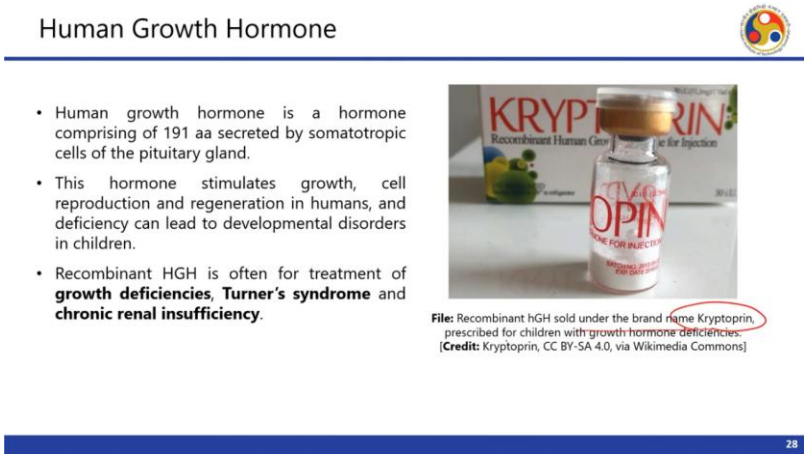
So, let us try to understand how this recombinant DNA technology is useful in producing insulin. So basically, we take an expression vector, which is a promoter, and then the insulin gene is cloned into this vector, and we clone the A chain and B chain into two separate vectors. So in this, we are cloning the insulin gene A subunit, and in this, we are cloning the insulin gene B subunit. And both have antibiotic-resistant genes, which help us in the selection process. Once the vector is internalized into a competent cell, the insulin A is produced as a fusion protein along with beta-galactosidase, which is also under the control of the same promoter.

So, in one population, we will be getting the fusion protein that contains the insulin A gene, and in the other, we will get the insulin B fusion protein. So, these are purified separately, and beta-galactosidase is removed by a chemical process. The method releases the A-chain and the B-chain. These A-chain and B-chain are then cross-linked with one another by disulfide bonds to form a functional insulin molecule. So, this is a very popular method now and also an economical method to produce insulin in bulk quantities.

So, basically, you can see the recombinant plasmid separately transformed into E. coli host cells containing the A and B chain genes. And the pro-insulin chains, the fused beta-galactosidase A chain and beta-galactosidase B chain, are separately produced in the host. Then these are separated from beta-galactosidase by treatment with cyanogen bromide, and the B-chain is joined in vitro to produce the native insulin by sulfonating the peptide chain with sodium disulfonate and sodium sulfite. Let us now discuss the human growth hormone. Human growth hormone is a hormone comprising 191 amino acids, secreted by somatotropic cells of the pituitary gland.



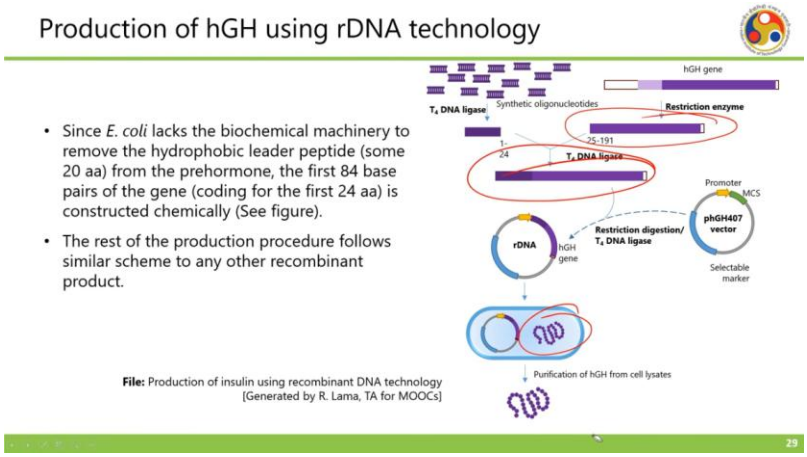
These hormones stimulate growth, cell reproduction, and regeneration in humans, and deficiency can lead to developmental disorders in children. Recombinant human growth hormone is often a treatment for growth deficiencies, Turner syndrome, and chronic renal insufficiency. So, here you can see recombinant human growth hormone sold under the brand name Cryptoprine. This is prescribed for children with growth hormone deficiencies. So, here is the process for the production of human growth hormone using our DNA technology.



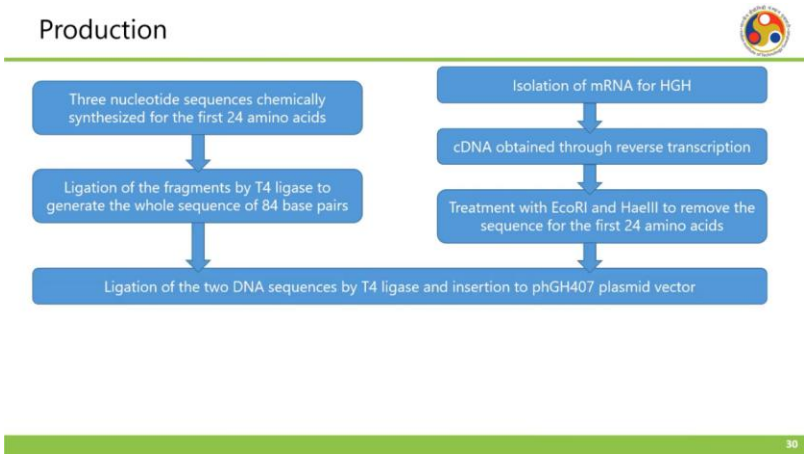
E. coli lacks the biochemical machinery to remove hydrophobic leader peptides, which are some 20 amino acids from the pre-hormone. So, you have this human growth hormone gene, and we undergo restriction digestion. And these are the synthetic oligonucleotides. So, we use T4 DNA ligase to ligate these 24-nucleotide residues with this portion of the human growth hormone. And then these are put into a vector, as in the case of insulin.

So, this vector has a promoter and a multiple cloning site, and then we use restriction sites in this MCS to open up this vector so that these constructs can be loaded. So, these are then

transformed into the cells, and then the expression host produces the HGH, and then we purify the human growth hormone from the lysates. The rest of the production procedure follows a similar scheme to any other recombinant product, as in the case of insulin. So, three nucleotide sequences are chemically synthesized for the first 24 amino acids, and ligation of the fragments by T4 ligase generates the whole sequence of 84 base pairs. Then we go for the isolation of mRNA for human growth hormone.



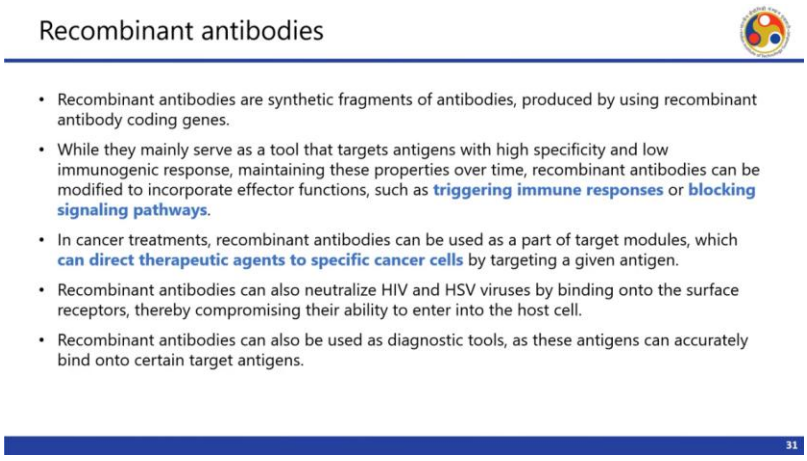
cDNA is obtained through reverse transcription. It is treated with EcoR1 and H3 to remove the sequence for the first 24 amino acids. Ligation of the two DNA sequences by T4 ligase and insertion into the plasmid vector, then the recombinant plasmid is transformed into an E. coli host where recombinant human growth hormone is produced. Now, let us discuss the production of recombinant antibodies. Recombinant antibodies are synthetic antibody fragments produced using recombinant antibody-coding genes.



While they mainly serve as tools that target antigens with high specificity, they also elicit a low immunogenic response. Maintaining these properties over time, recombinant

antibodies can be modified to incorporate effector functions, such as triggering immune responses or blocking signaling pathways. In cancer treatments, recombinant antibodies can be used as part of target modules, directing therapeutic agents to specific cancer cells by targeting a given antigen. Recombinant antibodies can also neutralize HIV and SSV viruses by binding to surface receptors, thereby compromising their ability to enter host cells.

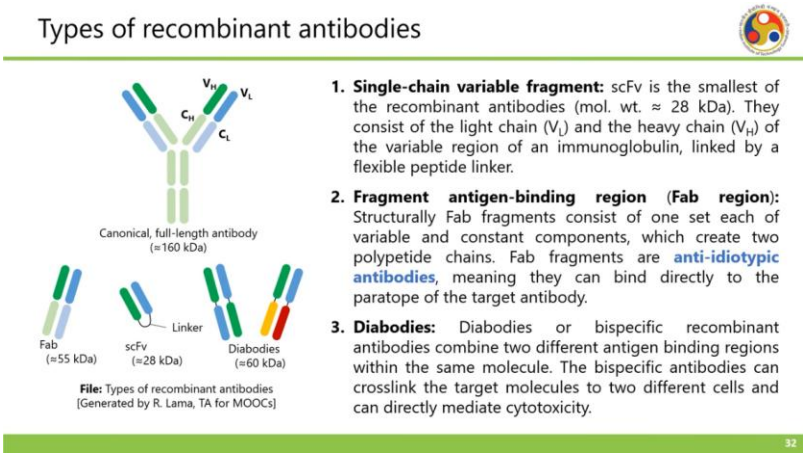
Recombinant antibodies can also be used as diagnostic tools, as they can accurately bind to certain target antigens. So, what are the types of recombinant antibodies? This is a canonical figure of a full-length antibody, roughly around 160 kilodaltons. It consists of the variable heavy chain, variable light chain, constant heavy chain, and constant light chain in those portions. Single-chain variable fragments (scFv) are the smallest recombinant antibodies, with a molecular weight of around 28 kilodaltons.



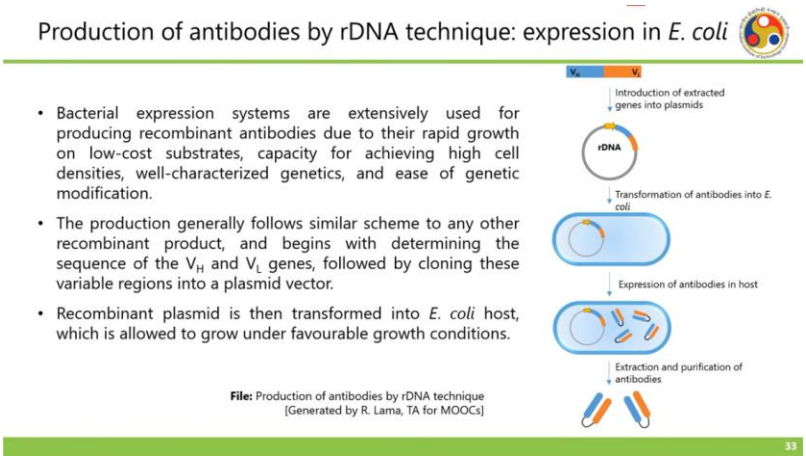
So, here if you consider this particular portion and then you link them covalently through a linker, you have this single chain variable fragment. It consists of the light chain VL and the heavy chain VH here as I have shown you, which is the variable region of an immunoglobulin linked by a flexible peptide linker. Then you have the fragment antigen binding region, FAB region. Structurally FAB fragments consist of one set each of variable and constant fragments components which create two polypeptide chains. FAB fragments are anti-idiotypic antibodies, meaning they can bind directly to the parotrope of the target antibodies.

So when we are taking this entire thing, this is the FAB, which is roughly around 55 kilodentons. Then we have diabetes or bispecific recombinant antibodies combine two different antigen binding regions within the same molecule. The bispecific antibodies can cross-link the target molecules to two different cells and can directly mediate cytotoxicity.

So, how these antibodies are produced using RDNA technique in E. coli? So, let us have a small discussion on that.



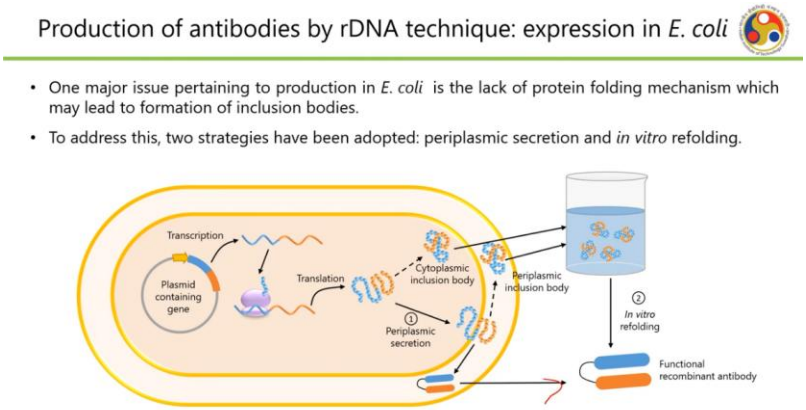
Bacterial expression systems are extensively used for producing recombinant antibodies due to their rapid growth on low cost substrates, capacity for achieving high cell densities, well characterized genetics, and ease of genetic modification. The production generally follows similar scheme to any other recombinant product and begins with determining the sequence of the variable heavy and variable light genes followed by cloning these variable regions into a plasmid vector. So, you have this variable heavy and variable light and these are introduced into plasmid and these are transformed into E. coli, then these are expressed and then you have these followed by extraction and purification of the antibodies. So, after transformation we grow them under favorable growth conditions.



One of the major issues pertaining to production in E. coli is the lack of protein folding mechanisms, which may lead to the formation of inclusion bodies. To address these, two strategies have been adopted: periplasmic secretion and in vitro refolding. So, here, this is

the plasmid containing the gene. So, after transcription, it is translated, and the protein is produced. So, here, this particular protein is allowed to go for periplasmic secretion,

and/or here, it gives the functional recombinant antibody. Then, in the other case, it goes into cytoplasmic inclusion bodies, which we can also separate out. So, once it is in the plasm, we go for separation, isolation, and then we allow for in vitro refolding. A second approach to obtain this functionality involves mimicking eukaryotic antibody secretion in E. coli, where proteins with a signal sequence are directed to the plasm. The periplasmic space, having an oxidizing environment opposed to the reductive cytoplasm, oxidizes the sulfhydryl groups, enabling the formation of the disulfide bonds over here, as you can see.



However, high-level expression may lead to insoluble aggregates due to folding intermediates. Protein aggregations can be resolved in vivo by co-expression of periplasmic chaperones like SKP and OMPH. Aggregations may also be resolved in vitro using the methods as follows. The first approach to the production of recombinant antibodies in E. coli is to produce antibody proteins as cytoplasmic inclusion bodies, followed by in vitro refolding. In this case, the protein is expressed without a signal sequence under a strong promoter.

Expression of in *E. coli* by secretion

- A second approach to obtain functional antibody fragments involves mimicking eukaryotic antibody secretion in *E. coli*, where proteins with a **signal sequence are directed to the periplasm**.
- Periplasmic space, having an oxidizing environment opposed to the reductive cytoplasm, oxidizes the sulfhydryl groups, enabling the formation of disulfide bonds.
- However, high-level expression may lead to insoluble aggregates due to folding intermediates.
- Protein aggregation can be resolved *in vivo* by coexpression of periplasmic chaperones like **Skp/OmpH**.
- Aggregations may also be resolved *in vitro* using methods discussed in the next slide.

35

The inclusion bodies contain the recombinant protein in a non-native and non-active conformation, which is then refolded using any of the following methods. We may have dilution refolding, redox refolding, or disulfide-restricted refolding. In the case of dilution refolding, we have around 12.6% efficiency; redox refolding, 40%; and disulfide-restricted refolding, 50%. So, from the point of view of reinitiation type, dilution refolding involves reinitiation prior to disulfide bond formation. The various steps involved here are: denatured proteins are first solubilized in a high denaturant concentration.

Then, the denaturant concentration is gradually decreased through dialysis. Then, we have redox refolding, where reinitiation and disulfide bond formation happen simultaneously. This utilizes a glutathione redox coupling system to catalyze disulfide interactions as the protein refolds into its native state. Then, we have disulfide-restricted refolding, where disulfide bond formation prior to reinitiation takes place. Here, denatured proteins are first solubilized with agents that catalyze disulfide bond formation, after which they are refolded by the slow removal of the denaturant. Now, let us look into some of the FDA-approved recombinant antibodies produced in E. coli.

In vitro refolding of cytoplasmic inclusion bodies

- The first approach to production of recombinant antibodies in *E. coli* is s to produce antibody proteins as cytoplasmic inclusion bodies followed by *in vitro* refolding.
- In this case the protein is expressed without a signal sequence under a strong promoter.
- The inclusion bodies contain the recombinant protein in a nonnative and nonactive conformation, which is then refolded using any of the following methods:

Method	Renaturation type	Steps in refolding	Efficiency
Dilution refolding	Renaturation prior to disulfide bond formation	Denatured proteins are first solubilized in a high denaturant concentration, then, the denaturant concentration is gradually decreased through dialysis	12.6%
Redox refolding	Renaturation and disulfide bond formation simultaneously	Utilizes a glutathione redox coupling system to catalyze disulfide interactions as the protein refolds into its native state	40%
Disulfide-restricted refolding	Disulfide bond formation prior to renaturation	Denatured proteins are first solubilized agents that catalyze disulfide bond formation, after which it is refolded by the slow removal of denaturing agent	50%

36

We have the generic name in the first column, followed by the molecular format, then expression condition, the disease against which this was developed, and the year of approval. So, we have Certolizumab Pegol. This is basically a pegylated Fab' fragment soluble in the periplasm. It is used against Crohn's disease. It was developed in 2008.

List of FDA approved recombinant antibodies produced in *E. coli*

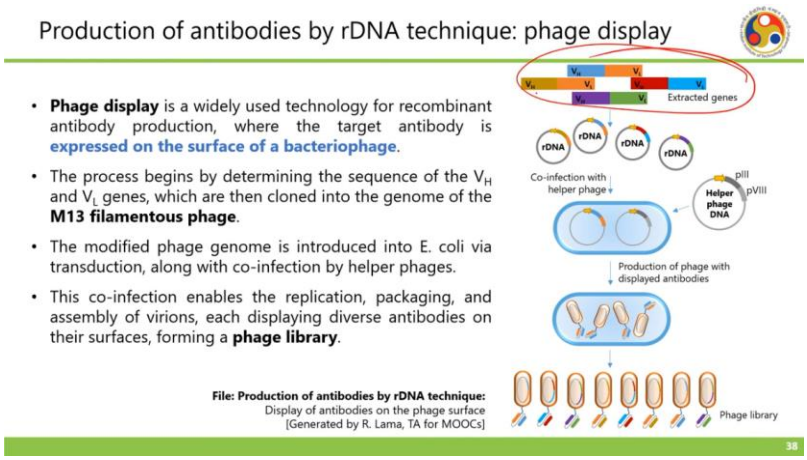
Generic (Brand) names	Molecular format	Expression condition	Diseases	Approval year
Certolizumab pegol (Cimzia®)	PEGylated Fab'	Soluble in periplasm	Crohn's disease, RA	2008
Ranibizumab (Lucentis®)	Fab	Soluble in periplasm	Wet AMD	2006
Romiplostim (Nplate®)	Peptide-Fc fusion	Inclusion body refolding in cytoplasm	ITP	2008
Moxetumab pasudotox (Lumoxiti®)	dsFv-PE38 fusion	Inclusion body refolding in cytoplasm	HCL	2018
Caplacizumab (Cablivi®)	Tandem NB	Extracellular soluble	aTTP	2019
Brolucizumab (Beovu®)	scFv	Inclusion body refolding in cytoplasm	Wet AMD	2019
Tebentafusp (Kimmtrak®)	Bispecific TCR-scFv fusion	Inclusion body refolding in cytoplasm	Metastatic uveal melanoma	2022

[Data adapted from Rashid., 2022] 37

Then we have Ranibizumab, which is basically a FAB soluble in periplasm. This is for wet AMD, developed earlier than the pegylated FAB in 2006. Then there are others, which you can simply go through in this table. And the latest ones are the Tabenta-Fast or Chemtrak. This is a bispecific TCR-scFv fusion.

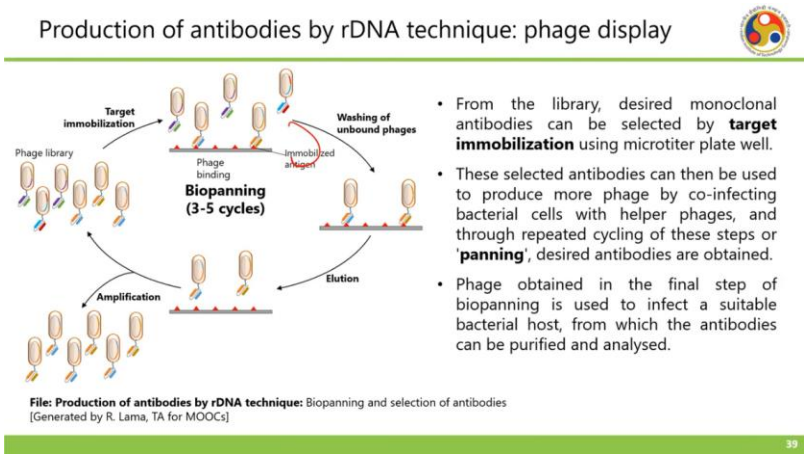
Its expression condition involves inclusion body refolding in the cytoplasm, used against metastatic uveal melanoma. Let us discuss the production of antibodies by the rDNA technique, the phage display method. So, phage display is a widely used technology for recombinant antibody production, where the target antibody is expressed on the surface of a bacterial phage. The process begins by determining the sequence of the VH and VL genes, which are then cloned into the genome of the M13 filamentous phage.

The modified phage genome is introduced into E. coli via transduction, along with co-infection by helper phages. This co-infection enables the replication, packaging, and assembly of variants, each displaying diverse antibodies on their surfaces, forming a phage library. So, here are the extracted VH and VL genes. So, we clone them, and then there is a co-infection with helper phage DNA over here, and then there is the production of phage with displayed antibodies, and then we get a phage library where you see various antibodies being displayed. From this library, the desired monoclonal antibody can be selected by target immobilization using a microtiter plate well.



So, here you have the FAS library. Then we proceed with the target immobilization. Then washing of the bound FAS occurs here. And then the unbound FAS will be removed. The unbound FAS is washed off, while those captured by this plate will be retained and then eluted out. These can then be further amplified, or we can repeat this panning process for around three to five cycles of biopanning.

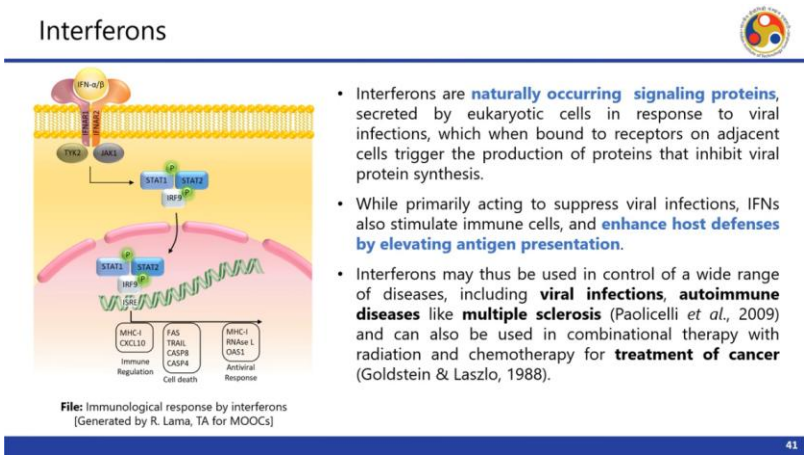
The selected antibodies can then be used to produce more FAS by co-infecting bacterial cells with the help of FASs. Through repeated cycling of these steps or panning, desired antibodies are obtained. The FAS obtained in the final step of biopanning is used to infect a suitable bacterial host, from which the antibodies can be purified and analyzed. So, here is a list of recombinant antibodies produced using the FAS display technology. In the first column, you can see the name, for example, Adalimumab. And then in the second column, you have the company or the developer.



This is FY. And then you get the target, which is TNF alpha, and the indication against which it is useful, such as rheumatoid arthritis, Crohn's disease, psoriatic arthritis, plaque

psoriasis, and so on. The status or the trade name is proof that it is Humira. So, you can go through this table and see many such antibodies have been developed, whether for systemic lupus, squamous lung cancer, asthma, or Alzheimer's disease. And also as a prophylaxis and treatment of anthrax. Some of them are already approved, while some were in phase 3 trials when this table was compiled.

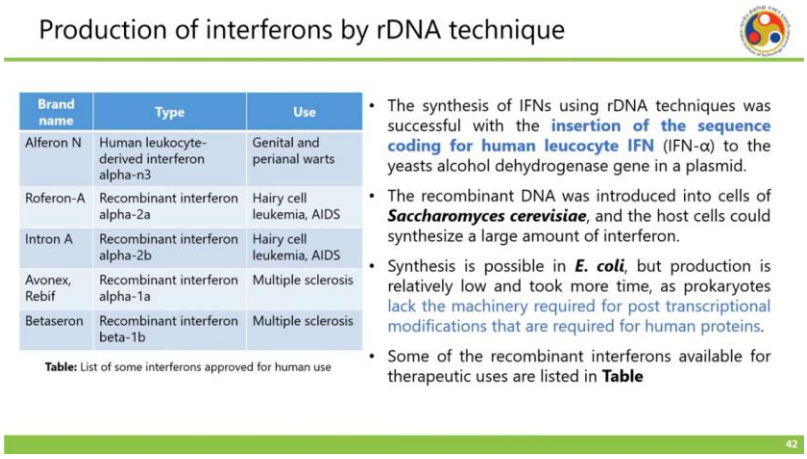
Let us now discuss interference, which are naturally occurring signaling proteins secreted by eukaryotic cells in response to viral infections. When bound to receptors on adjacent cells, they trigger the production of proteins that inhibit viral protein synthesis. While primarily acting to suppress viral infections, IFNs also stimulate immune cells and enhance host defense by elevating antigen presentation. Interferons are used in the control of a wide range of diseases, including viral infections and autoimmune diseases like multiple sclerosis. They can also be used in combination with radiation and chemotherapy for the treatment of cancer. So, here we discuss the production of interferons using the rDNA technique. The synthesis of IFNs using the rDNA technique was successful with the insertion of the sequence coding for human leukocyte IFN, or IFN-alpha, into the yeast alcohol dehydrogenase gene in a plasmid.



The recombinant DNA was introduced into cells of Saccharomyces cerevisiae. In the earlier case, we were mostly using E. coli, but here we are using yeast. The host cells could synthesize large amounts of interferon. Synthesis is possible in E. coli, but production is relatively low and takes more time, as prokaryotes lack the machinery required for post-transcriptional modifications needed for human proteins. Some of the recombinant interferons available for therapeutic uses are listed in this table here.

So, you have the brand name and the type in the second column, and the use. The brand name is Alferon N. Here, this is human leukocyte-derived interferon alpha, and it is used

for genital and perineal warts. Then you have Roferon A, which is recombinant interferon alpha-2A. It is used against hairy cell leukemia and also in acquired immunodeficiency syndrome. Then you have Intron A, which is recombinant interferon alpha-2B, used against hairy cell leukemia and AIDS.



Then Evonex Revif, which is recombinant interferon alpha 1A used against multiple sclerosis. And against multiple sclerosis, there is another drug, which is the beta serone. Basically, it's recombinant interferon beta 1B. So, next we go to discuss about enzyme replacement therapy ERT which is a medical treatment device in the 1960s by Christian D. Duvet and Roscoe Bradley. This involves the intravenous administration of an enzyme for which the patient is deficient or it is absent in the body of the patient.

ERT does not treat the genetic disorder causing the enzyme deficiency. It increases the concentration of the enzyme that is lacking in the patient. ERT treatments are now available for some lysosomal storage diseases such as Gocher disease, Fabry disease, MPS1, MPS2, Hunter syndrome, MPS4 and Pompe disease. So here in this picture you can see some drugs using enzyme replacement therapy. In the top you can see Febrazyme, a medication for Fabry's disease.

Then you have this Cerezyme over here, a medication for Gotcher disease. ERT has also been used to treat patients with severe combined immunodeficiency resulting from an adenosine deaminase deficiency. So, we have numerous recombinant enzymes produced in microbes against various diseases. And the expression system ranges from E. coli, P. pasteris, Saccharomyces cerevisiae, then Ogechia minuta, then Yarrowia lipolytica, and then we have also Saccharomyces pombii. And then we have various different enzyme targets.

Enzyme replacement therapy

- Enzyme replacement therapy (ERT) is a medical treatment devised in the 1960s by Christian de Duve and Roscoe Brady, which involves the **intravenous administration of an enzyme that is deficient or absent in the body**.
- While ERT does not treat the genetic disorder causing the enzyme deficiency, it increases the concentration of the enzyme that is lacking in the patient.
- ERT treatments are now available for some lysosomal storage diseases, such as: Gaucher disease, Fabry disease, MPS I, MPS II (Hunter syndrome), MPS VI and Pompe disease.
- ERT has also been used to treat patients with severe combined immunodeficiency (SCID) resulting from an adenosine deaminase deficiency (ADA-SCID).

File: Drugs used in enzyme replacement therapy:
(top) Fabrazyme, a medication for Fabry's disease [**Credit:** Genzim Company, Public Domain, via Wikimedia Commons]
(bottom) Cerezyme, a medication for Gaucher disease[**Credit:** Genzyme Company, Public Domain, via Wikimedia Commons]

43

For example, we have adenosine deaminase against ADA deficiency. Then in Gaucher's disease, we have beta-glucocerebrosidase. In Fabry's disease, we have alpha-galactosidase. In Pompe's disease, we have alpha-glucosidase and so on. Apart from these, there are many other recombinant therapeutic enzymes like recombinant uricase or urate oxidase, which is an enzyme that catalyzes the conversion of uric acid to 5-hydroxyisourate and can therefore be used in treating conditions related to high amounts of uric acid in gout and in rhabdomyolysis with kidney failure.

List of recombinant enzymes produced in microbes

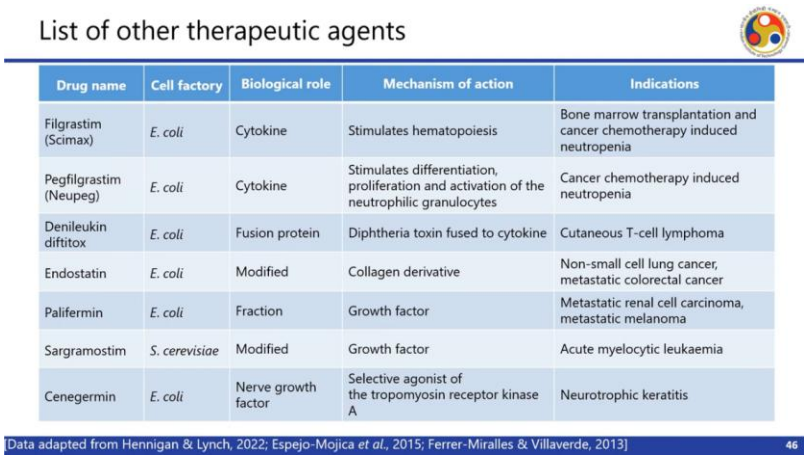
Disease	Enzyme target	Expression system
ADA-SCID	Adenosine deaminase	*Escherichia coli*
Gaucher's disease	β-glucocerebrosidase	*Pischia pastoris*
Fabry's disease	α-galactosidase	*E. coli, P. pastoris, Saccharomyces cerevisiae, Ogataea minuta*
Pompe disease	α-glucosidase	*Yarrowia lipolytica*
MPS II (Hunter syn.)	Iduronate-2-sulfatase	*E. coli, P. pastoris*
MPS IV A (Morquio A syn.)	N-acetylgalactosamine 6-sulfatase	*E. coli, P. pastoris*
MPS VII (Sly syn.)	β-glucuronidase	*E. coli*
Lysosomal acid lipase deficiency	Lysosomal acid lipase	*P. pastoris, Saccharomyces pombe*
GM2 gangliosidosis	N-acetyl-hexosaminidase A N-acetyl-hexosaminidase B	*O. minuta, P. pastoris*
Congenital Sucrase-Isomaltase Deficiency	Sucrase	*Saccharomyces cerevisiae*

[Data adapted from Silva *et al.*, 2022; Espejo-Mojica *et al.*, 2015; Hennigan & Lynch, 2022]

44

In humans, however, the enzyme is absent, and therefore recombinant uricase is widely used in the treatment of hyperuricemia. Then we have rasburicase. This is produced in Saccharomyces cerevisiae, and pegloticase produced in E. coli are two variations available for commercial use. And we have recombinant carboxypeptidase G2. Glucarpidase or voraxaze is a medication used for the treatment of elevated levels of methotrexate during treatment of cancer patients who have impaired kidney function.

Glucarpidase, a recombinant form of the bacterial enzyme carboxypeptidase G2, converts methotrexate into less toxic glutamate and 2,4-diamino-N-methylpteroic acid. So in this table, we can see many therapeutic agents. The first column we have the drug name, and then the cell factory or the organism in which it is produced. Mostly they are all majorly produced in E. coli, and then we also have Saccharomyces cerevisiae. And then this column tells you about the mechanism of action.



List of other therapeutic agents

Drug name	Cell factory	Biological role	Mechanism of action	Indications
Filgrastim (Scimax)	*E. coli*	Cytokine	Stimulates hematopoiesis	Bone marrow transplantation and cancer chemotherapy induced neutropenia
Pegfilgrastim (Neupeg)	*E. coli*	Cytokine	Stimulates differentiation, proliferation and activation of the neutrophilic granulocytes	Cancer chemotherapy induced neutropenia
Denileukin diftitox	*E. coli*	Fusion protein	Diphtheria toxin fused to cytokine	Cutaneous T-cell lymphoma
Endostatin	*E. coli*	Modified	Collagen derivative	Non-small cell lung cancer, metastatic colorectal cancer
Palifermin	*E. coli*	Fraction	Growth factor	Metastatic renal cell carcinoma, metastatic melanoma
Sargramostim	*S. cerevisiae*	Modified	Growth factor	Acute myelocytic leukaemia
Cenegermin	*E. coli*	Nerve growth factor	Selective agonist of the tropomyosin receptor kinase A	Neurotrophic keratitis

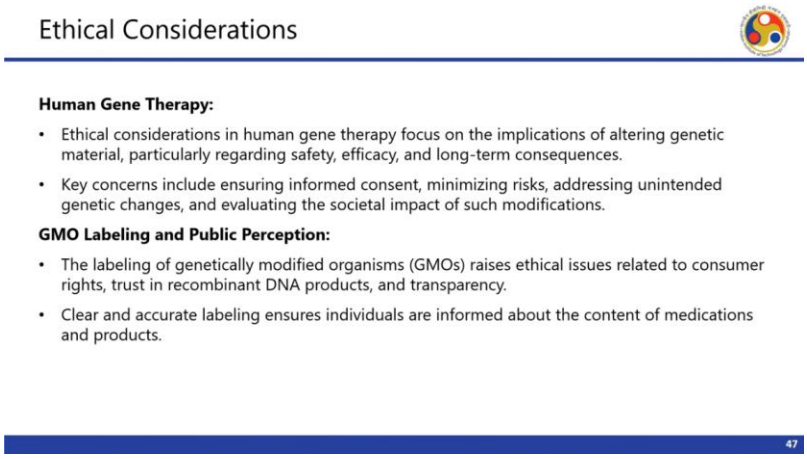
[Data adapted from Hennigan & Lynch, 2022; Espejo-Mojica *et al.*, 2015; Ferrer-Miralles & Villaverde, 2013] 46

For example, this filgrastim stimulates hematopoiesis. And it is used in the bone marrow transplantation and cancer chemotherapy induced neutropenia. Then we have PEG filgrastim. This is also known as NeuPEG. This stimulates differentiation, proliferation, and activation of the neutrophilic granulocytes used in cancer chemotherapy induced neutropenia.

Then we have danilucan-diphtytox, which is a fusion protein. This is diphtheria toxin used, fused to cytokine. Indications is cutaneous T-cell lymphoma. So, kindly go across this table, which have several other drugs like endostatin, danilucan-diphtytox, we have already discussed, then pelifamine and so on. Now let us discuss about one very important issue, which is ethics.

Ethical considerations, particularly human gene therapy, focuses on the implications of altering genetic material, particularly regarding safety, efficacy and long-term consequences. Key concerns include ensuring informed consent, minimizing risk, addressing unintended genetic changes and evaluating the societal impact of such modifications. And GMO labelling and the public perception is also an important ethical issue. The labelling of genetically modified organisms raises ethical issues related to consumer rights, trust in recombinant DNA products and transparency.

Some people may not prefer GMO products, and they should be given the opportunity to identify the product. Clear and accurate labeling ensures individuals are informed about the content of the medications and products, and they can therefore make an informed decision. Other important issues are regulatory oversight and responsible innovation. Regulatory frameworks are essential for ensuring safety, scientific integrity, and ethical standards in gene therapy and GMO technologies. Oversight must balance innovation with patient safety, societal impacts, and accountability.



Responsible innovation requires transparency, stakeholder engagement, and consideration of broader societal implications. So, with this, we come to the end of this lecture. Thank you for your kind attention. Thank you.