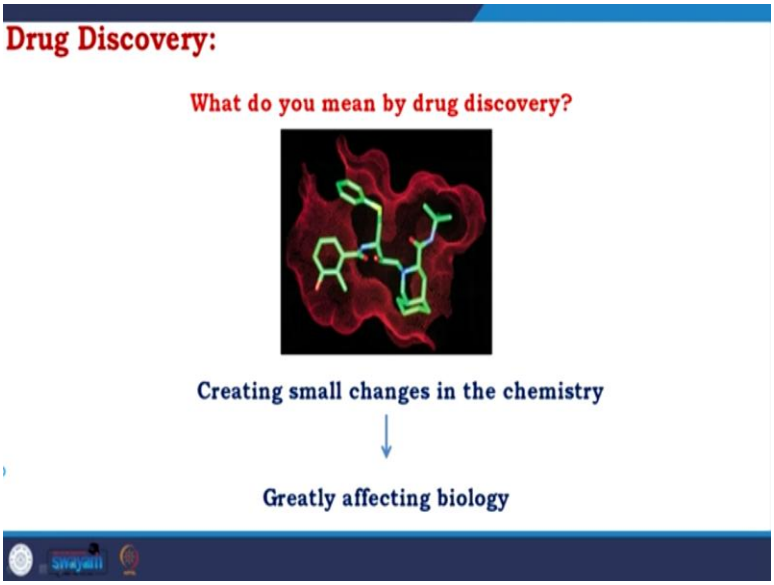
Structural Biology
Prof. Saugata Hazra
Department of Biotechnology
Indian Institute of Technology – Roorkee

Lecture – 56
Introduction to Structure-Based Drug Discovery (SBDD)

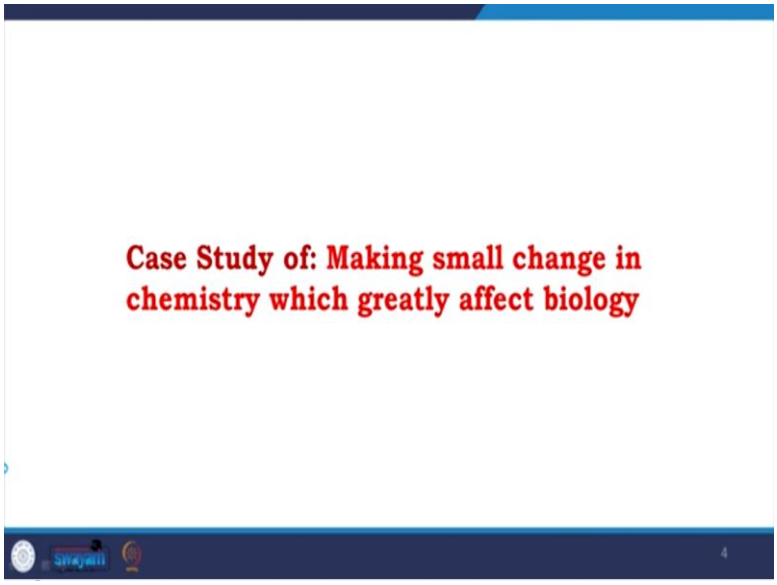
Hi everyone, welcome again to the course on structural biology. We are at the end module, where the topic is structure-based drug discovery. In the drug discovery, we will see the methodologies which are helping us to identify a potent small molecule that could be used as a drug.

(Refer Slide Time: 02:21)



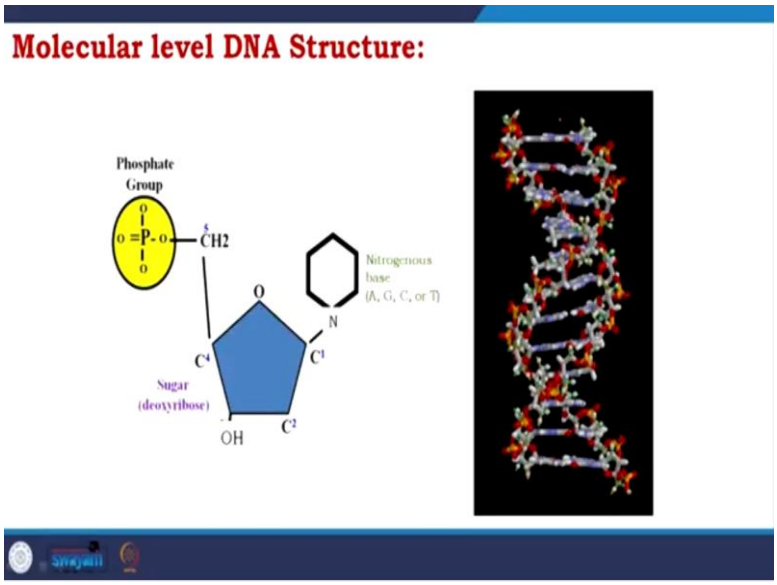
What do you mean by drug discovery? Creating small changes in chemistry and by making that small change, there is a huge effect on biology.

(Refer Slide Time: 03:13)



To explain what I mean by this, I take the help of a case study of drug designing, which is my research area, and try to explain the whole concept.

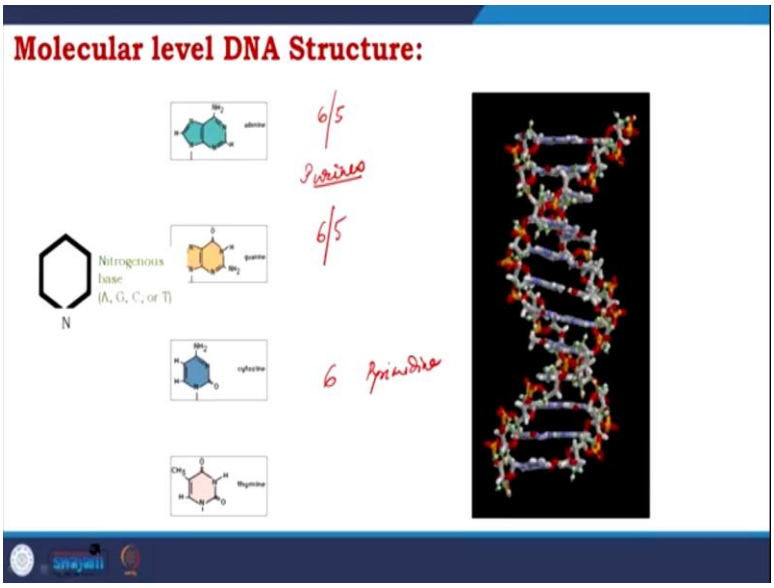
(Refer Slide Time: 03:23)



In my introductory classes, I have already talked about the molecular level of the DNA structure. If you look, you will see the red molecules, red and yellow molecules are phosphate, the bases are hydrogen-bonded, and they are stacked. If you look at the molecular component, you could see three major components in DNA or even in RNA. There is sugar in DNA. It is called deoxyribose sugar, whereas, in RNA, ribose sugar. There is a phosphate group when there is a synthesis of DNA. It comes with three phosphates, releases two phosphates, and gets the energy to make the polymerization reaction go into a thermodynamically favorable direction. Then the

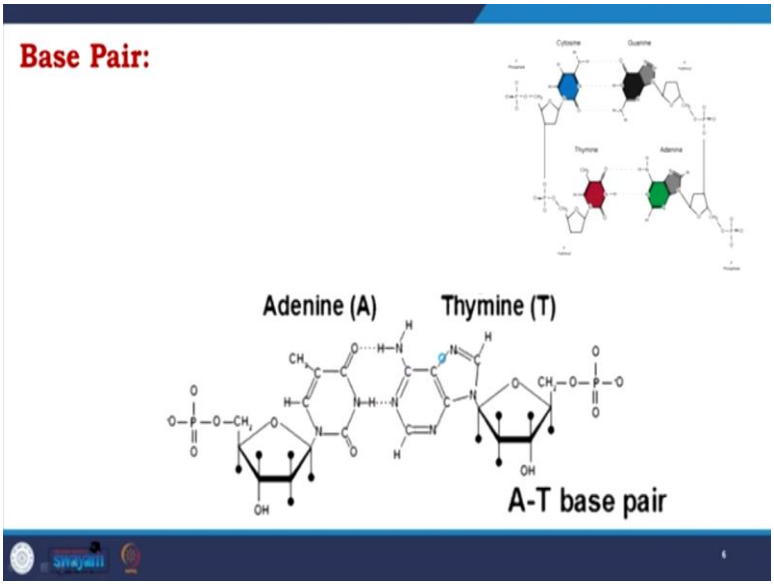
third one is the nitrogenous bases, and in the case of DNA, it is A, G, C, T, whereas, in RNA, the thymine is replaced with uracil. So, if you look at the entire composition, the only difference happened in the nitrogenous bases.

(Refer Slide Time: 05:10)

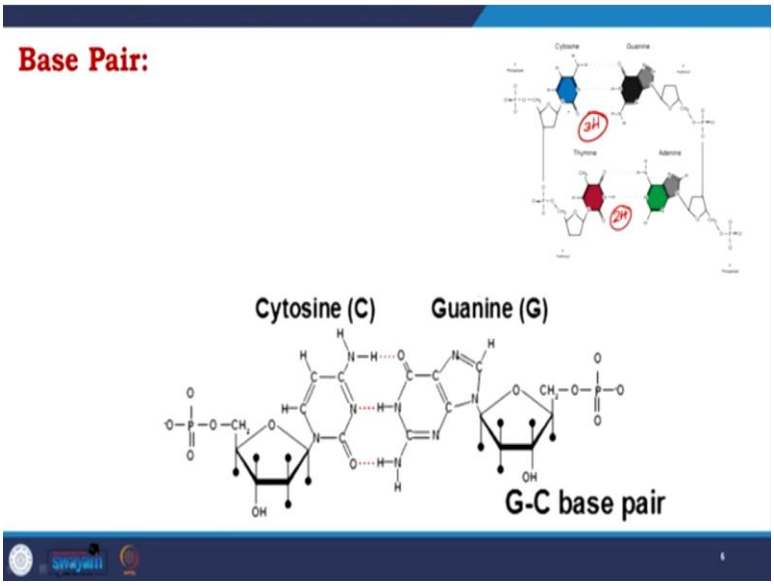


The nitrogenous bases are of 4 types adenine, guanine, cytosine, and thymine, but among those four, they are also divided into two classes. Adenine and guanine are called purines, cytosine and thymine are called pyrimidine.

(Refer Slide Time: 05:53)

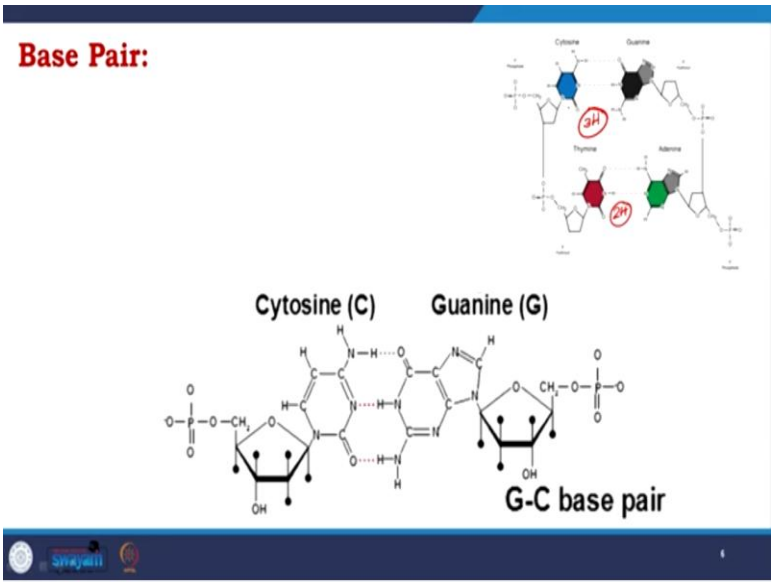


(Refer Slide Time: 06:03)



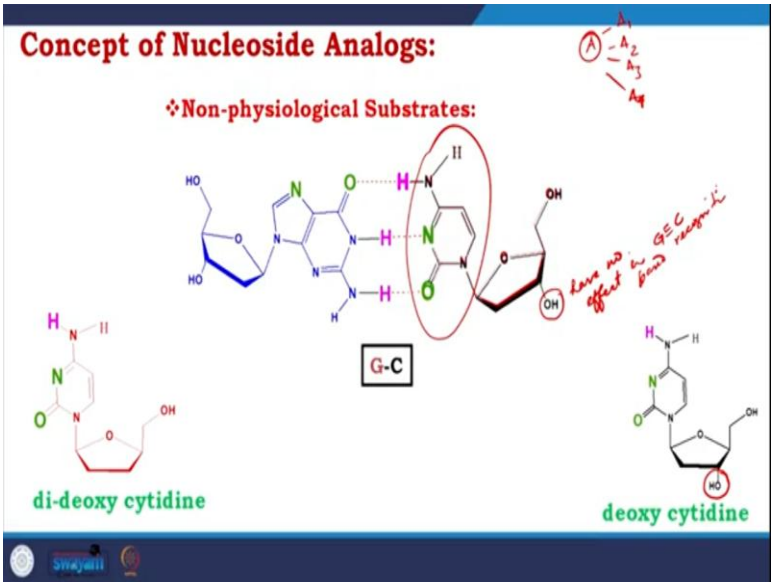
Where two hydrogen bonds are formed between adenine and thymine and cytosine with guanine, there are three hydrogen bonds.

(Refer Slide Time: 06:21)



Hydrogen bonds are intrinsic characters used in living organisms to identify A over T and G over C. So, they are very critical. And with that background in mind, now, I am going to introduce something which is called nucleoside analogs.

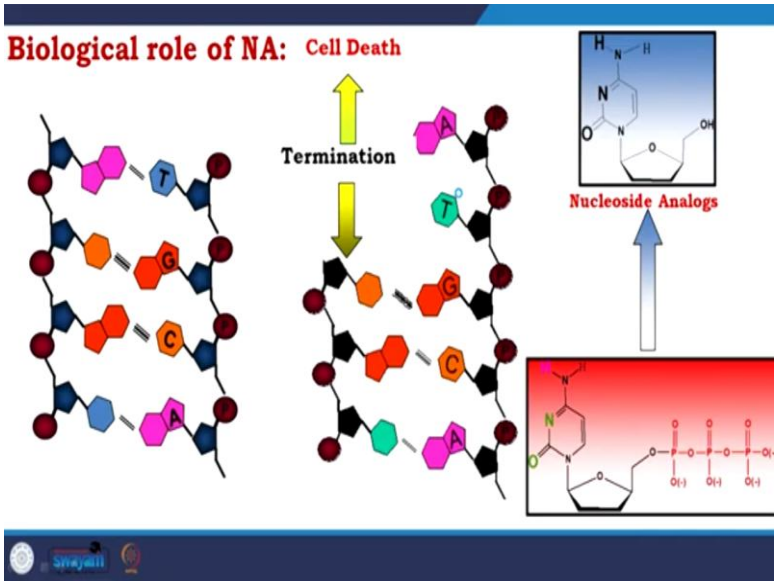
(Refer Slide Time: 07:08)



So, you know about nucleosides. Now, I am talking about analogs. What are analogs? If you have a compound A, and then you make variations A 1, A 2, A 3, A 4, they are called analogs. So, by definition, nucleoside analogs are variants of these nucleosides. Let us start with a simple one. You see deoxy cytidine; just now, we see that deoxy guanosine forms 3 hydrogen bonds with deoxy cytidine.

Now, if you see, only this portion is involved in a hydrogen bond. So, now, if I cleverly change the sugar part's other part, what will happen? Before going into details, I could show you that we designed a compound and did not design get rid of this hydroxyl group. This hydroxyl group does nothing to recognize hydrogen bonds, so it did not affect G-C-based recognition. It could have the same hydrogen-bonding pattern as deoxy cytidine, it could easily insert into the DNA by the recognition process through G-C hydrogen bonding. But, once it will insert there, what will happen? Let us see.

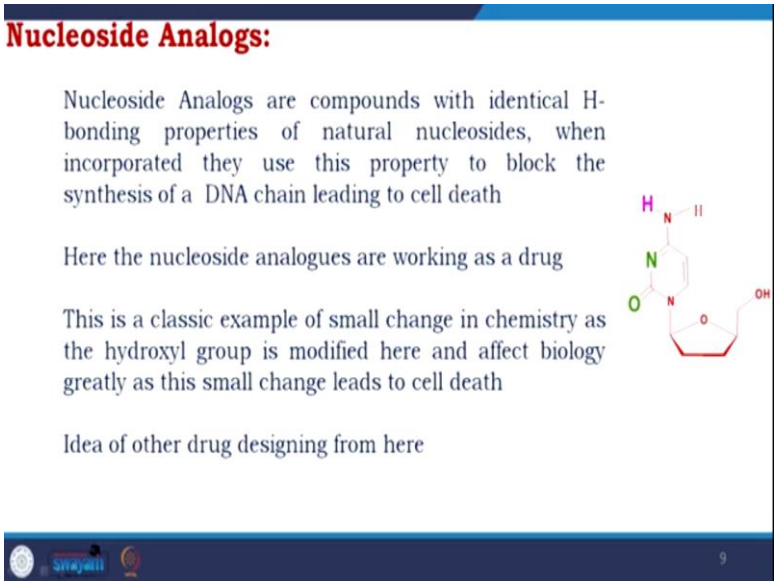
(Refer Slide Time: 09:45)



So, this is a normal DNA synthesis, where A is there, so T is coming. C is there, so G is coming. G is there, so C is coming. And T is there, so A is coming in that way. A newly formed chain is continuing to form a new DNA molecule. Whereas, instead of this deoxy, what will happen if we have di deoxy? Somehow, if a di deoxy could incorporate, it would never allow the DNA to extend further, which will lead to the termination of DNA synthesis.

So, this DNA would be stopping its elongation in the process of polymerization. So, when you do not have DNA, your cell cannot survive. So, it could lead to cell death. That is what I call drug designing. So, this is nucleoside analog, which is their critical role in biology.

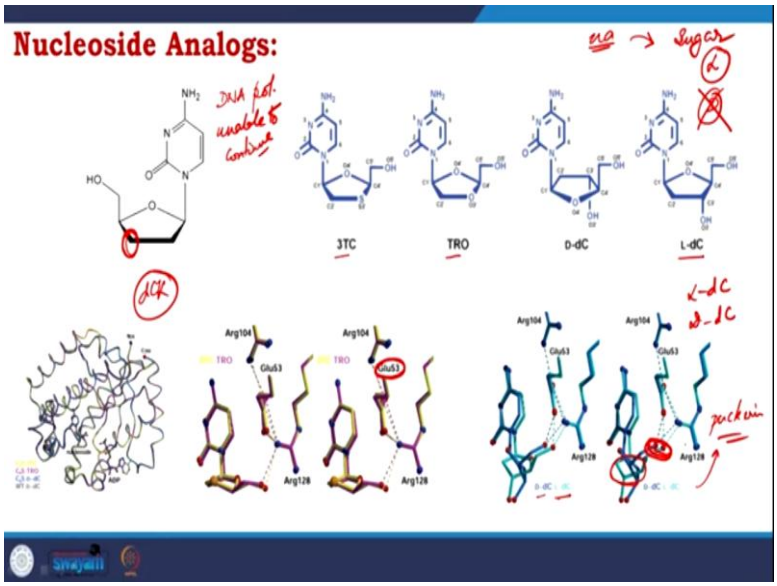
(Refer Slide Time: 11:58)



Nucleoside analogs are compounds with identical hydrogen bonding properties. You have already seen that in the picture where I put the G-C hydrogen-bonded pattern and then brought the di deoxy one. So, nucleoside analogs are compounds with identical hydrogen bonding properties to natural nucleosides. When incorporated, they use this property to block the synthesis of a DNA chain, leading to cell death. This nucleoside analog could cause cell death. You could use these as a drug, how you could use this as a drug. If a person has cancer, cancer cells divide with an uncontrolled condition. Somehow, if you could incorporate these nucleoside analogs, they would kill the normal cell, but because this process is speeded up, they will kill more and more cancer cells. And that is the thing you probably observed when we said that a person is affected or diagnosed with cancer, getting cancer chemotherapy.

In most cases, they are treated with nucleoside analogs. We take advantage of a cancer cell, which grows much faster. So, these nucleoside analogs enter and stop DNA synthesis and then kill the cells. So that tells us that these nucleoside analogs could be working as a drug as a therapeutics. If you see people getting cancer chemotherapeutics, you see that they vomit, become tired, and lose their hair. This is because this nucleoside analog kills the cancer cell, but simultaneously, they are killing the normal cells. So, we want specific therapeutics, and for that, you have to work on this compound and improve the molecule that is the process of drug optimization. So, we have to go through the next cycles. So, here this molecule would be considered a lead molecule, and then the further level of drug discovery would be going on to make that drug better.

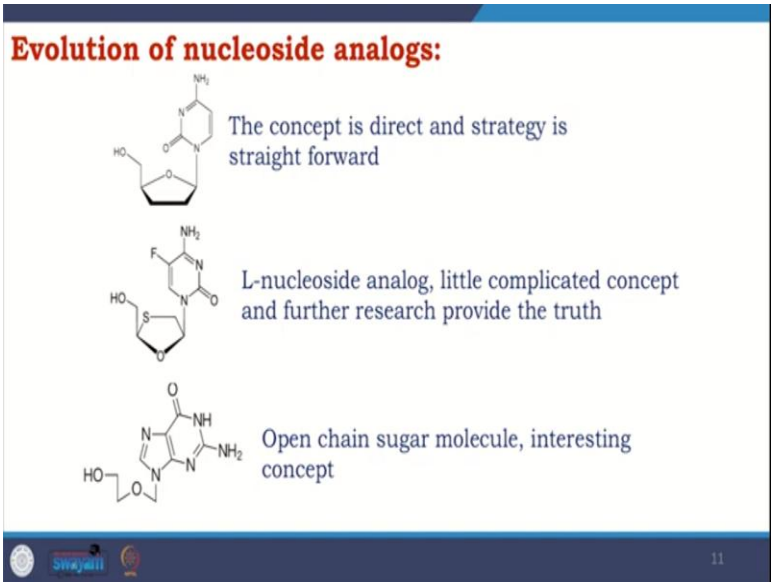
(Refer Slide Time: 17:26)



So, nucleoside analogs, if I am talking about a di deoxy cytidine, di deoxy uracil, di deoxy thymine, the case is clear and easy to understand for you. This does not have this hydroxyl group. So, once it would introduced DNA, DNA polymerization would not continue. I will introduce new molecules 3TC: lamivudine, throxocytamine, L-di deoxy cytidine, and D-di deoxy cytidine. So, there are three nucleoside analogs where the sugar loses the hydroxyl group. But here, the sugar is having alkydility instead of normal d. That was an enigma for a long time, though, in reality, as drug lamivudine, throxocytamine. L-deoxy cytidine all work much better than those straightforward analogs. They kill the natural cells at a speed much less than di deoxy compounds. But the mechanism was in question. And in my Ph.D., I got the opportunity to work with this. So, we crystallize all of those compounds with a target enzyme called deoxy cytidine kinase. And we have clearly shown the molecular details of how these molecules interact with the enzyme and how the catalytic residue Glu53 is coming close to the sugar.

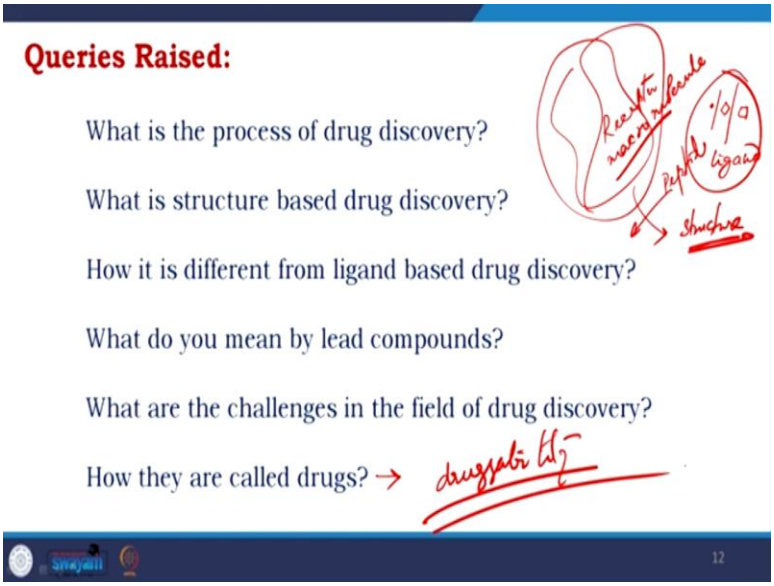
The most interesting part of this design is comparing L-deoxy cytidine and D-deoxy cytidine. In the case of D-deoxy cytidine, you will see that a huge puckering (sugar puckering) happened. And these sugar puckering helps the L nucleoside present its atoms close to the Glu53 as that D-deoxy cytidine did not.

(Refer Slide Time: 22:32)



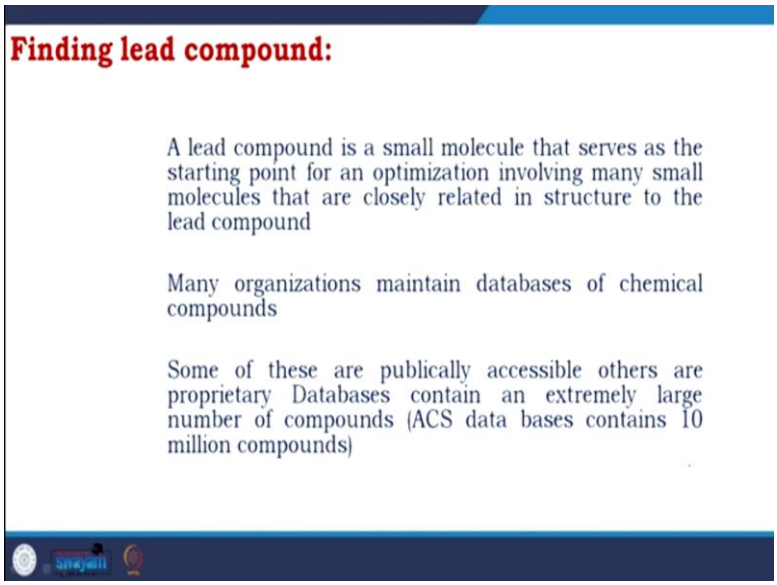
And as I was talking about how drug designing was evaluated, we talked about these. The concept is direct, and the strategy is straightforward: the absence of this hydroxyl group does not allow the DNA polymerase to elongate the ring. But when you come to the L nucleosides. So, it is a little bit complicated. But the research comes up with the truth. Even if it had a hydroxyl group, it would not present that to the enzyme L DNA polymerase to further elongate and then come to the third type of molecule, which is wonderful anti-viral, specifically working against herpes simplex virus. This is called acyclovir, ganciclovir, penciclovir, the base is taken out, so instead of being a close ring, this is an open ring.

(Refer Slide Time: 25:47)



So, if you follow me, you will now have questions, you have got few answered, what is the process of drug discovery, I have already shown you what is done discovery, but how to do it, what are the processes, it should be in your mind then, as I told this is a structural biology course. So, our drug discovery would also be in addition to other methods, and we have a major focus on structure-based drug discovery.

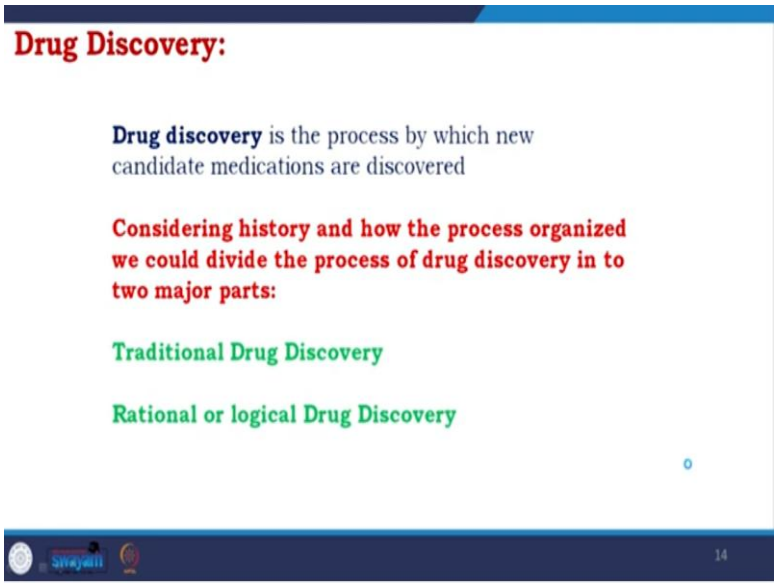
(Refer Slide Time: 30:17)



A lead compound is a small molecule that serves as the starting point of an optimization involving many small molecules that is close to related in structure to the lead compound. Many organizations maintain databases of chemical compounds. Some of these are publicly accessible, and others are proprietary databases that contain an extremely large number of compounds. The

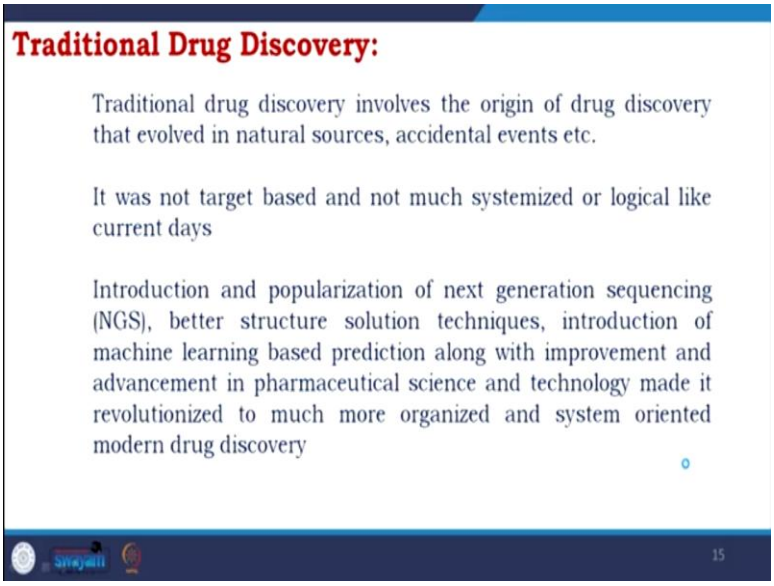
most popular one which is accessible is ACS database, which contains 10 million compounds. There is a database called zinc. When I go to structure-based ligand design, I will give you a table of the popular databases of small molecules.

(Refer Slide Time: 31:14)



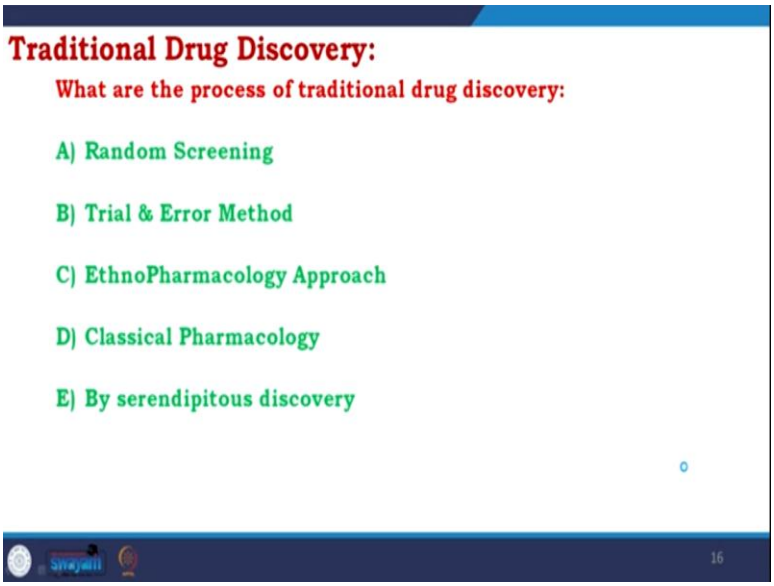
Drug discovery is the process by which new candidate medications are discovered. Considering the history and how the process is organized, we could divide the process of drug discovery into two major parts: traditional drug discovery and rational or logical drug discovery.

(Refer Slide Time: 32:42)



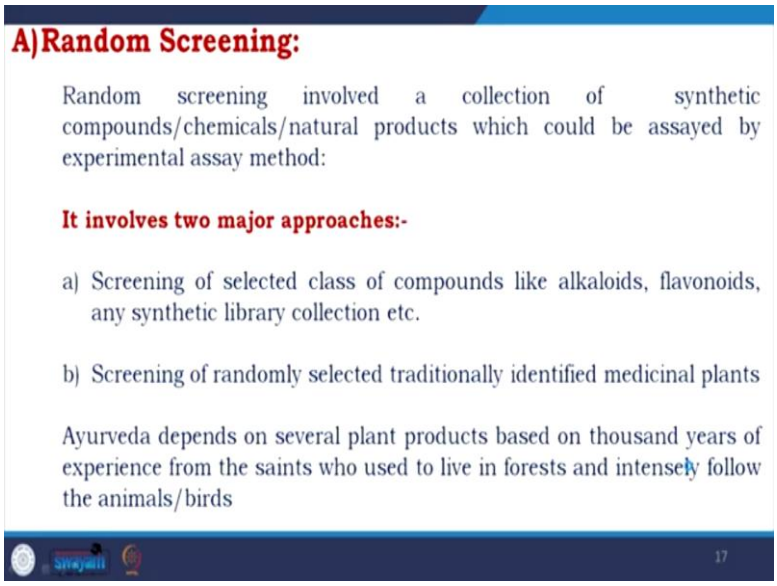
The traditional drug discovery involves the origin of drug discovery that evolved in natural sources, accidental events, etc. It was not target-based and not much systemized or logical, like current days. Introduction and popularization of next-generation sequencing, better structure solution techniques, machine learning, and improvement and advancement in pharmaceutical science and technology made it revolutionized to much more organized and system-oriented, modern drug discovery.

(Refer Slide Time: 33:32)



You could divide traditional drug discovery majorly into five topics, random screening, trial and error method, EthnoPharmacology approach, classical pharmacology, By serendipitous discovery.

(Refer Slide Time: 34:17)



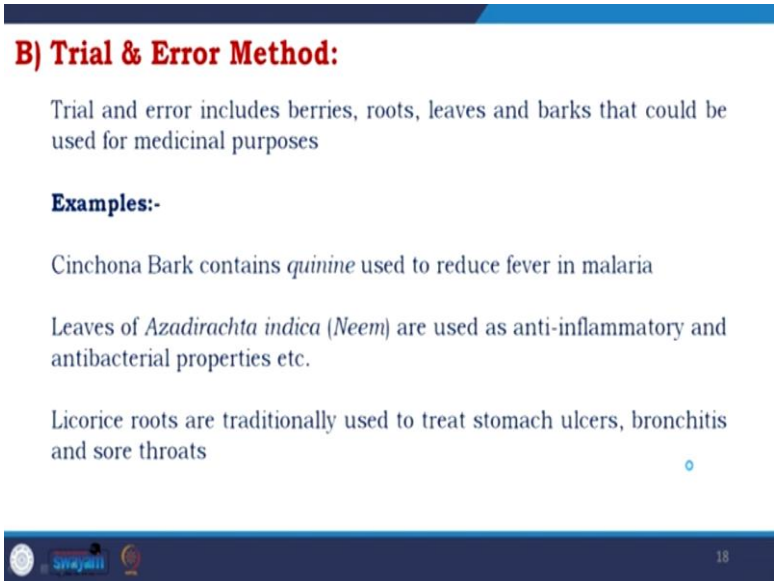
Random screening involves a collection of synthetic compounds, chemicals, natural products, which could be assayed by the experimental assay method. So, suppose you want to develop a drug against diarrhea. You have an assay system whereby by adding the drug, you could differentiate between a small molecule that is affecting in a good manner or which is not affecting.

It involves two major approaches.

1) The screening of selected compounds like alkaloids, flavonoids, any synthetic library collection.
2) Screening of randomly selected traditionally identified medicinal plants.

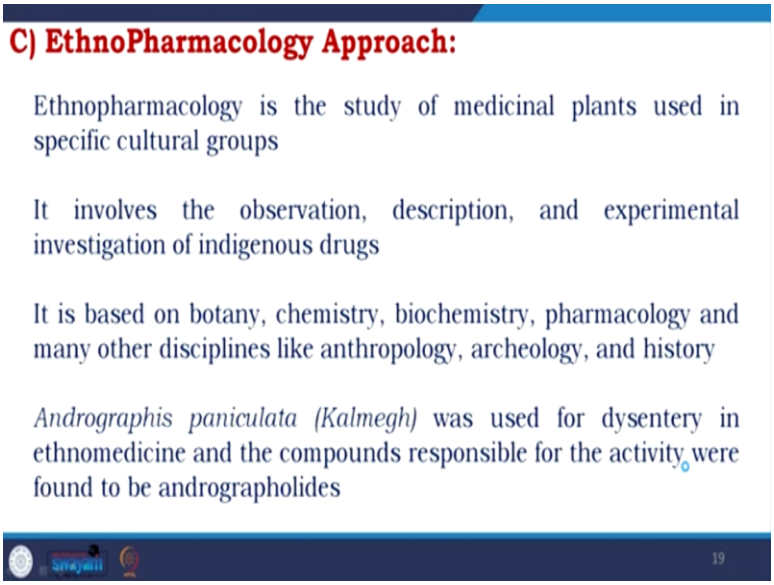
In traditional drug design, the medicinal plant has a huge role. As we know from ancient Indian civilization, they used to culture the process called Ayurveda; the Ayurveda depends on several plant products based on 1000 years of experience from the saints who used to live in the forest and intensely follow the animals or birds.

(Refer Slide Time: 36:27)



Trial and error include berries, roots, leaves, and barks used for medicinal purposes. A few examples are Cinchona bark content quinine, which reduces fever in malaria. Leaves of Azadirachta indica (neem) are used as anti-inflammatory and antibacterial properties. Licorice roots are traditionally used to treat stomach ulcers, bronchitis, and sore throats.

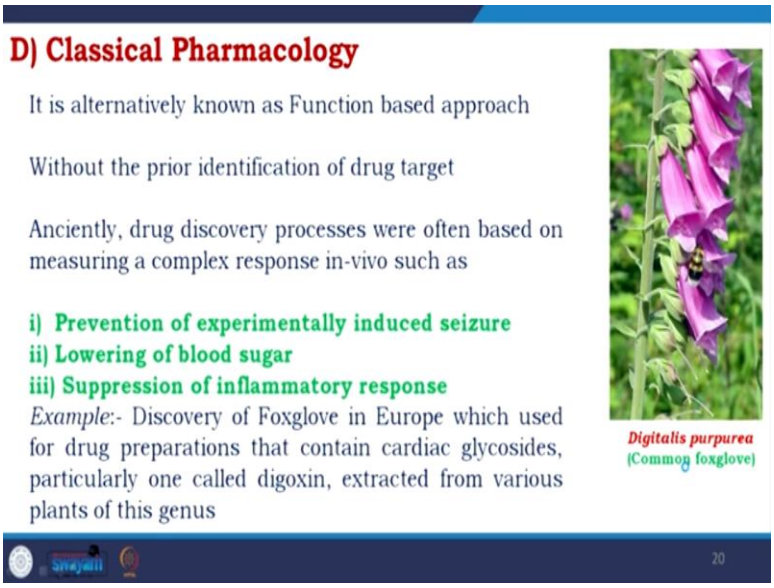
(Refer Slide Time: 37:17)



EthnoPharmacology is the study of medicinal plants used in specific cultural groups. It involves the observation, description, and experimental investigation of indigenous drugs. It is based on botany, chemistry, biochemistry, pharmacology, and many other disciplines like anthropology, archaeology, and history. Andrographis paniculata (Kalmegh) was used for dysentery in

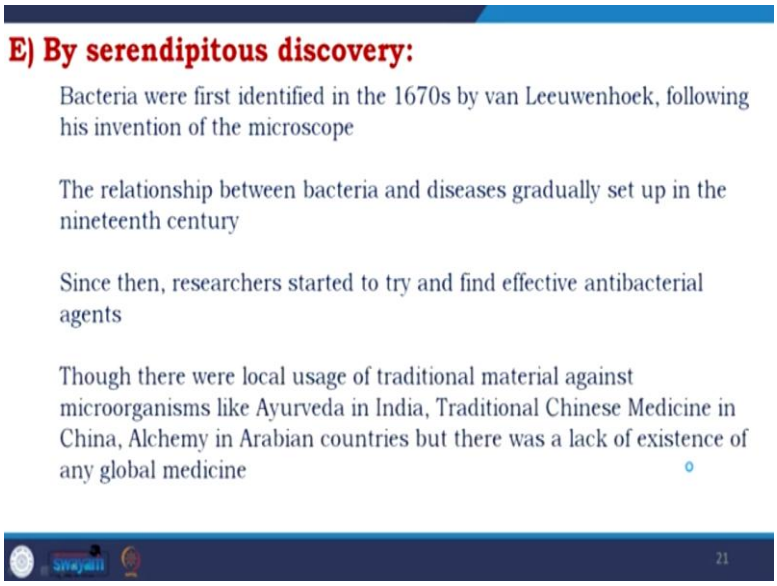
ethnomedicine, and the compounds responsible for the activity were found to be andrographolides.

(Refer Slide Time: 38:16)



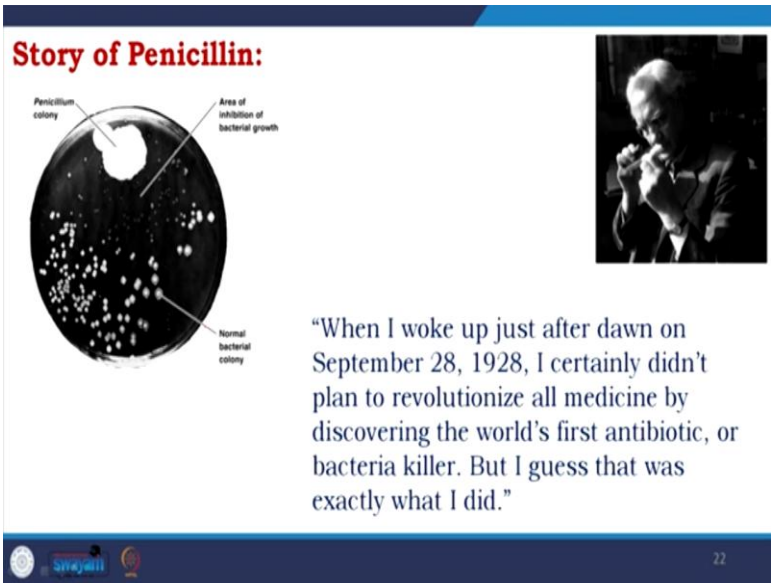
Classical pharmacology is alternatively known as the function-based approach. Without prior identification of drug target anciently, drug discovery processes are often based on measuring a complex response in vivo, such as preventing experimentally induced seizures, lowering blood sugar, and suppressing the inflammatory response. One very good example is the discovery of Foxglove in Europe, which was used for drug preparation. That content cardiac glycoside, particularly digoxin, is extracted from various plants of this genus. This is digitalis purpurea, which is common Foxglove.

(Refer Slide Time: 39:15)



Bacteria were first identified in the 1670s by Van Leeuwenhoek. I have already told you following his invention of the microscope, the relationship between bacteria and diseases gradually set up in the 19th century. Since then, researchers today to try and find effective antibacterial agents. Though there was local usage of traditional material against microorganisms like Ayurveda in India, traditional Chinese medicine in China, Alchemy in Arab countries, there was a lack of global medicine.

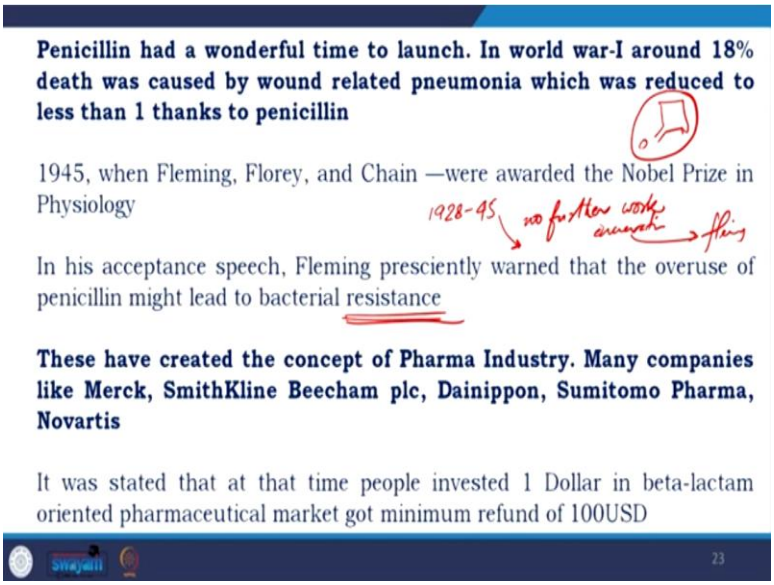
(Refer Slide Time: 40:33)



And the concern was somehow solved by a discovery in 1928 by Alexander Fleming, he was a doctor who was working on staphylococcus, and once he went to a vacation, you could see the plate, this is a bacterial Petri plate, when he returned from vacation, with his surprise, he realized

that there is a mold growing in the Petri plate. And very interestingly, in the region where the mold is grown, the bacteria is not present there. They are already killed, the colonies are gone or minimized. After some following up experiment, what he found, I want to use his experience 'when I woke up just after dawn on September 28, 1928. I certainly did not plan to revolutionize all medicine by discovering the world's first antibiotic or bacteria killer, but I guess that was exactly what I did'. So, in one moment, the worry of lack of communication around the world was solved, because somehow the world caught a beautiful answer, the magic bullet which could kill any and every bacteria.

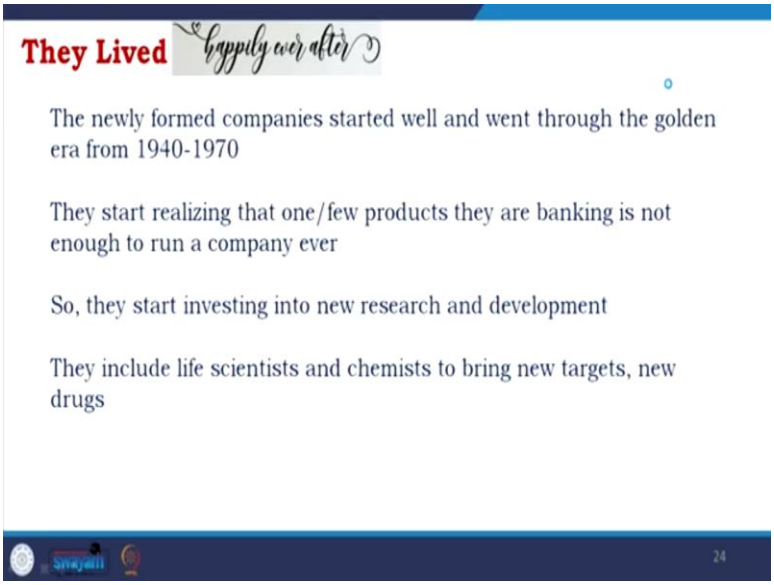
(Refer Slide Time: 43:28)



Penicillin had a wonderful time launching it with a beautiful sense of the debate. Because when penicillin was invented, World War 1 was already finished. And World War is coming. In World War 1, if you see, around 18% of the soldiers were not killed in the war. They had a minor wound, but that only led to bacterial infection. They were following up on the killing of the soldier. 18% of the soldiers were killed there. Second World War, which is more fierce in terms of intensity, is only less than 1% of people infected or killed thanks to penicillin. And you would be more surprised to know that penicillin was not even a drug at that time. It was the mold, and the mold was put under the squeezer. The juice coming out from the mold is allowed to put on the wound, which was good enough to save them.

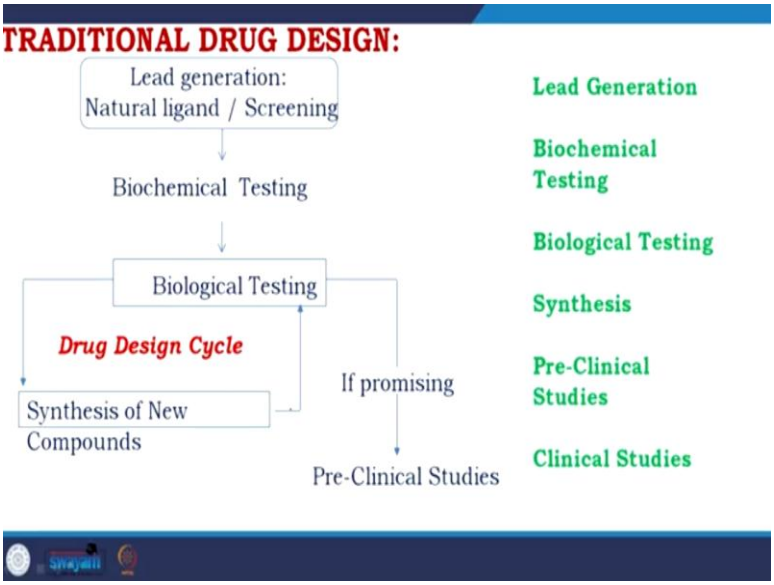
Suddenly, all the countries involved in the war got the utility of penicillin, and it was accepted globally. In 1945, Fleming, Florey, and Chain were awarded Nobel Prize. Florey and Chain were at the University of Oxford, and they had done the actual molecular level research and come up with the structure. Fleming presciently warned the overuse of penicillin might lead to bacterial resistance. The first time the world listened to the term bacterial resistance, today you look at the entire world a burning problem is bacterial resistance.

(Refer Slide Time: 51:08)



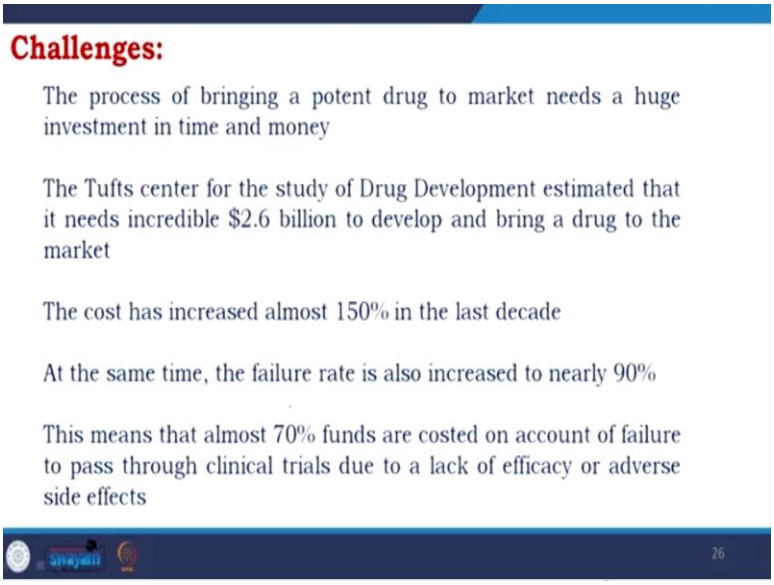
The newly formed company started well and went through the golden era from 1940 to 1970. They start realizing that one or a few products they are banking on is not enough to run a company forever. So, they start investing in new research and development. They include life scientists and chemists to bring new targets and new drugs to patent and continue the industry.

(Refer Slide Time: 51:40)



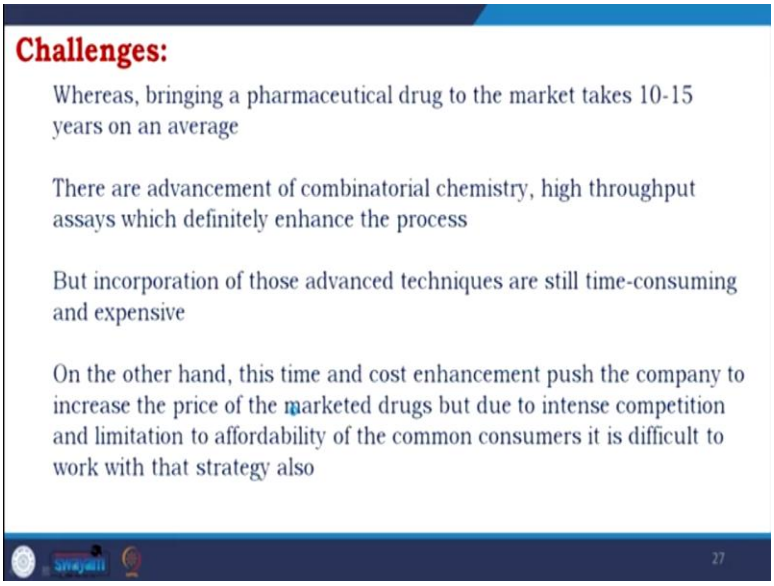
Before going into challenges, let us look at traditional drug design. You have this lead generation natural ligand, and you do a screening. You get the biochemical testing, and if you get positive results in the biochemical testing, you go to biological testing. You get success, and you synthesize new compounds on one side, which is a drug design cycle. On the other side, if it gives promising results, you go to preclinical and clinical studies. So, what involved traditional drug designing, lead generation, biochemical testing, biological testing, synthesis of small molecules, peptides, preclinical studies, and if that would work, clinical study.

(Refer Slide Time: 52:40)



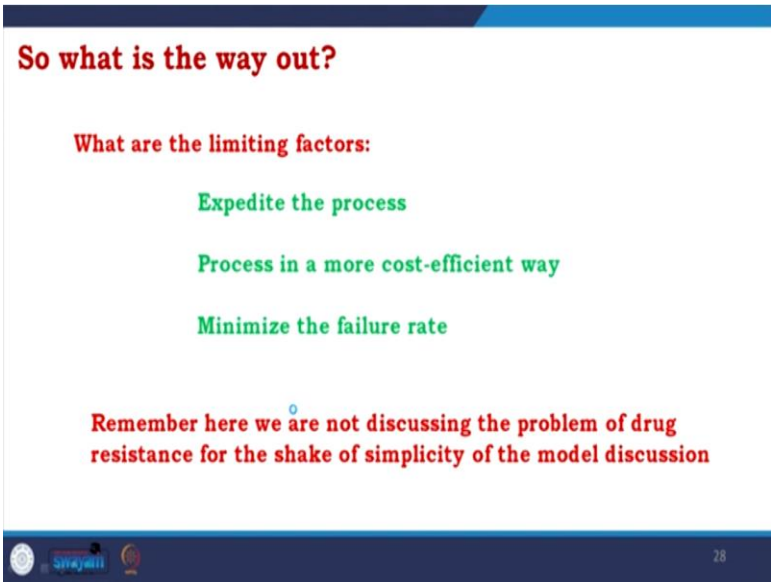
The process of bringing a potent drug to market needs huge investment in time and money. The tufts center for the study of drug development estimated that it needs an incredible 2.6 billion dollars to develop and bring a drug to the market. The cost has increased almost 150 percent in the last decade, making it more problematic. At the same time, the failure rate is also increased to nearly 90%. This means that almost 70% of funds are costed for failure to pass through clinical trials due to a lack of efficacy or adverse side effects on the preclinical or clinical trials.

(Refer Slide Time: 53:36)



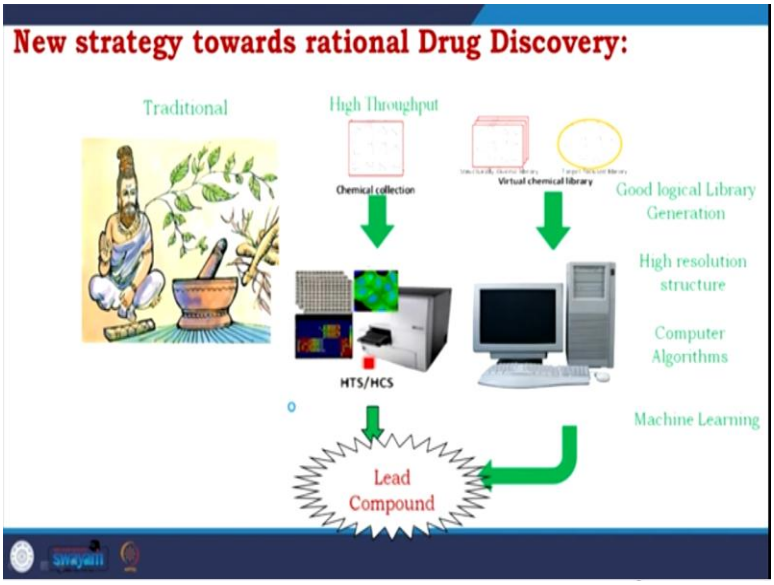
Whereas bringing a pharmaceutical drug to the market takes 10 to 15 years on average, advancements in combinatorial chemistry high throughput assays enhance the processing. Incorporation of those advanced techniques is still time-consuming and very expensive. On the other hand, this time and cost enhancement boost the company to increase the price of the marketed drugs. But due to intense competition and limitations to the affordability of common consumers, they will not be getting proper sell targets.

(Refer Slide Time: 54:48)



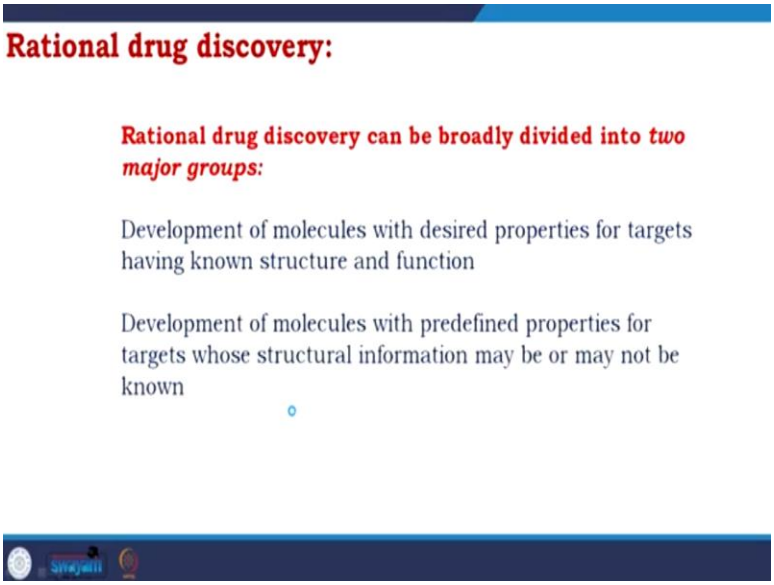
You have to expedite the process of drug development, make the process a more cost-efficient way, and minimize the failure rate.

(Refer Slide Time: 55:52)



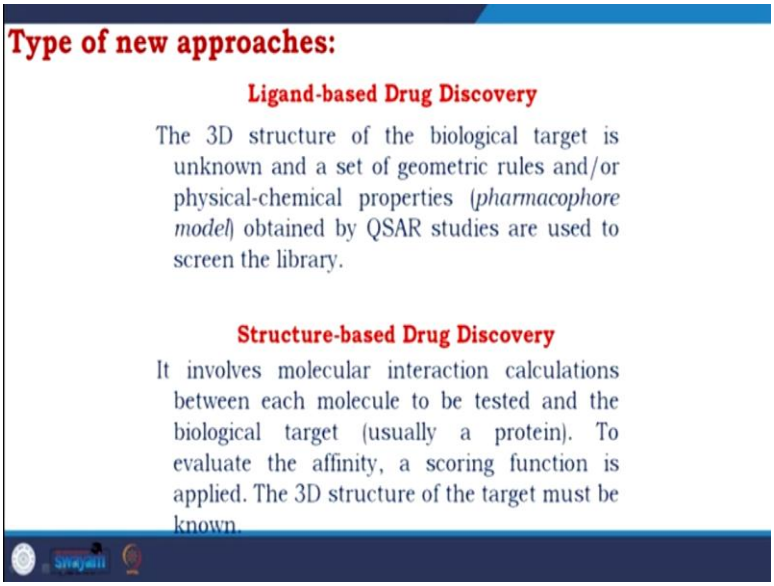
The new strategy is to take the traditional method, which is converting to high throughput, replacing it somehow by computational method, computer to achieve the dream goal, lead development replacing the experimental way as much as possible using the computer. So, we need good logical library generation high-resolution structures, computer algorithms, and incorporate the modern techniques of machine learning.

(Refer Slide Time: 56:30)



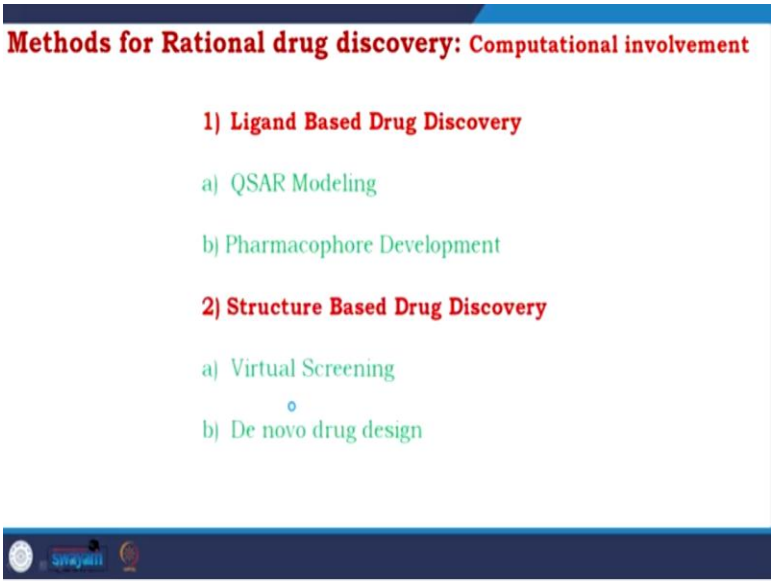
Rational drug discovery can be broadly divided into two major groups: the development of molecules with desired properties for targets with known structure and function. Development of molecules with predefined properties for targets whose structural information may be or may not be known. The unknown target information can be found by global gene expression data if not known.

(Refer Slide Time: 57:02)



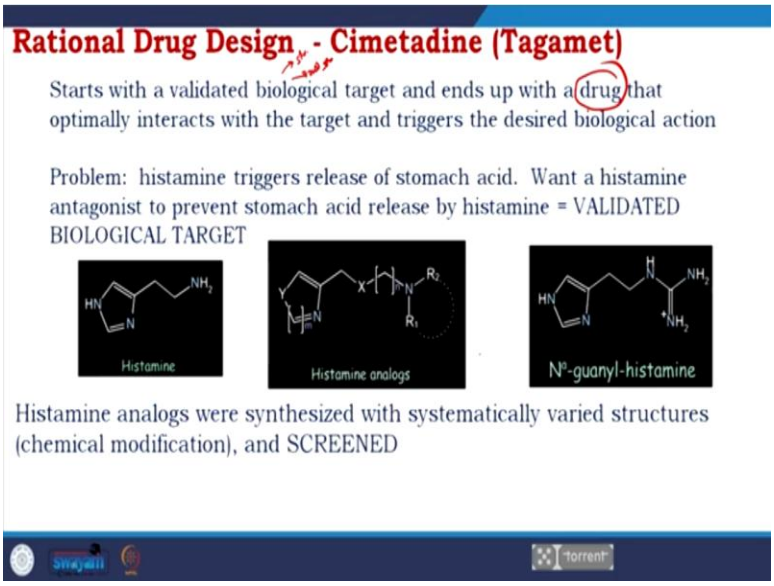
So, there are two new approaches one is ligand-based drug discovery, and another is structure-based drug discovery. The ligand-based drug discovery, the 3D structure of the macromolecule's biological target is unknown, and a set of geometric rules and or physical, chemical properties (pharmacophore model) is obtained by QSAR (quantitative structure activity relationship) studies are used to screen the library. In structure-based, it involves molecular interaction calculation between each molecule to be tested, and the biological target, usually a protein to evaluate the affinity, a scoring function is applied. The 3D structure of the target must be known.

(Refer Slide Time: 57:56)



So, my thoughts of rational drug discovery as I told it would be more computational involvement, ligand-based drug discovery, where it involves QSAR modeling, Pharmacophore development, whereas in structure-based drug designing a virtual screening and de novo drug designing.

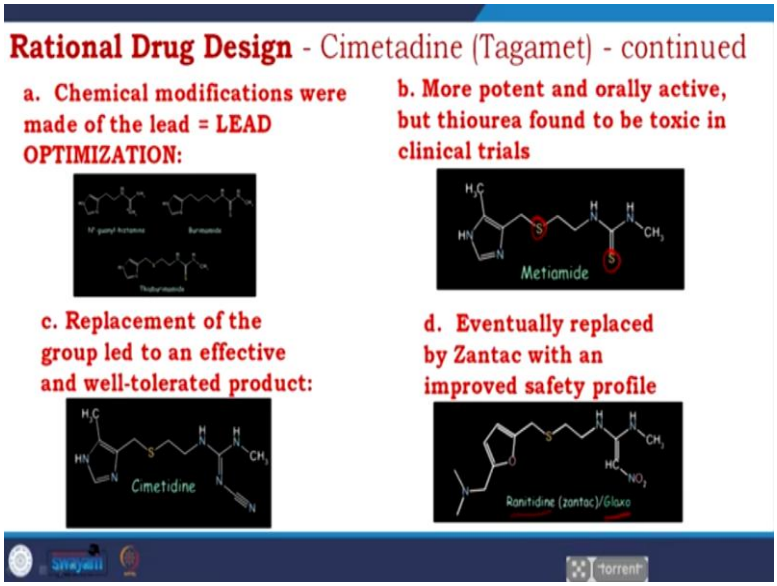
(Refer Slide Time: 58:23)



What is rational drug designing? Let us take an example using Cimetidine which is also called Tagamet. So, it starts with a validated biological target and ends with a drug that optimally interacts with the target and triggers the desired biological activity.

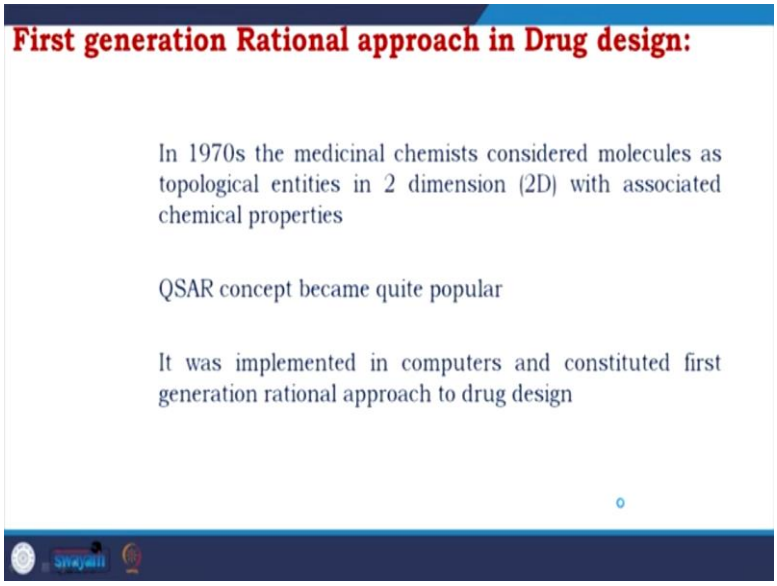
We are talking about a problem where histamine triggers the release of stomach acid. Want histamine antagonists to prevent stomach acid release by histamine (validated biological target). So, it takes histamine, and then we look at the molecule and want to make the histamine analogs. Now we get N-alpha-guanyl histamine which is already a good analog having desired biological effect. So, histamine analogs were synthesized with the systematically varied structure of the chemical modification and then screened. In N-guanyl histamine showed some antagonists' properties become the lead compound. So, this became the lead compound.

(Refer Slide Time: 01:00:19)



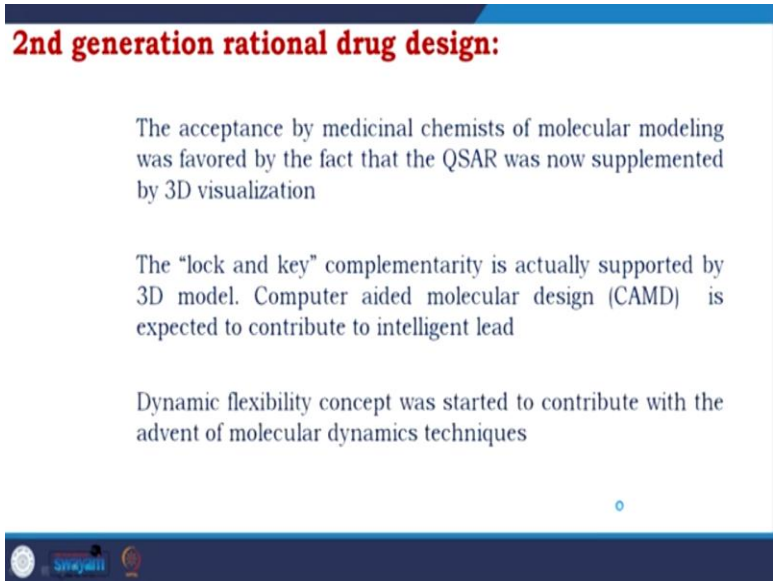
We make the chemical modification of the lead compound called lead optimization. From N-guanyl-histamine, we make the modifications which are Burimamide and Theoburimamide. Then we make it more potent and orally active. Still, we include thiourea because of the sulfur it found toxic, and people are commonly allergic to sulfur. So, that is the problem. Replacement of the group led to an effective and well-tolerated product. So, one sulfur is replaced by a chemical group that gives the drug Cimetidine and eventually replaced by Zantac with an improved safety profile.

(Refer Slide Time: 01:02:16)



In the first generation rational approach, in 1970, the medicinal chemists considered molecules as topological entities in two dimensions with associated chemical properties. QSAR concept became quite popular, and it was implemented in computers and constituted a first-generation rational approach to drug designing.

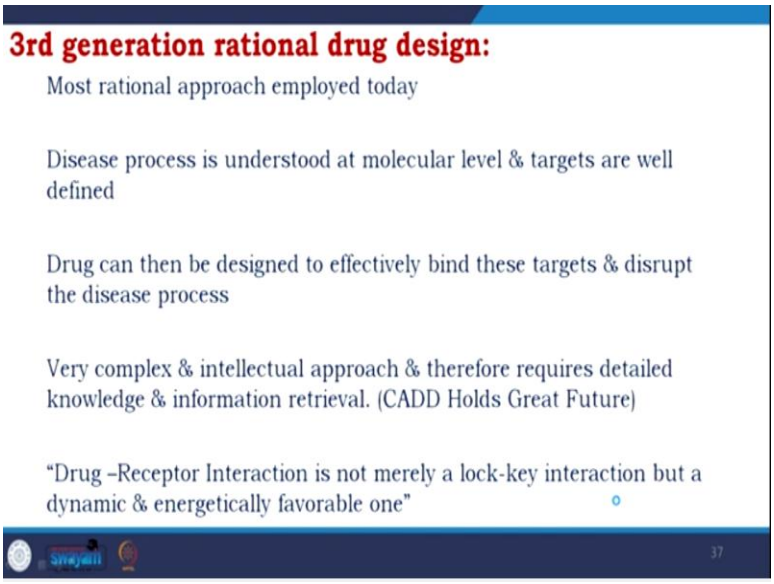
(Refer Slide Time: 01:02:43)



The second generation, the acceptance by medicinal chemists of molecular modeling, was favored because QSAR was now supplemented by 3D visualization. From 2D to 3D, you get a lot of visual effects in terms of sterical effects or interactions not realized in 2D. The lock and key complementarities supported by 3D model, Computer-Aided Molecular Design (CAMD), is

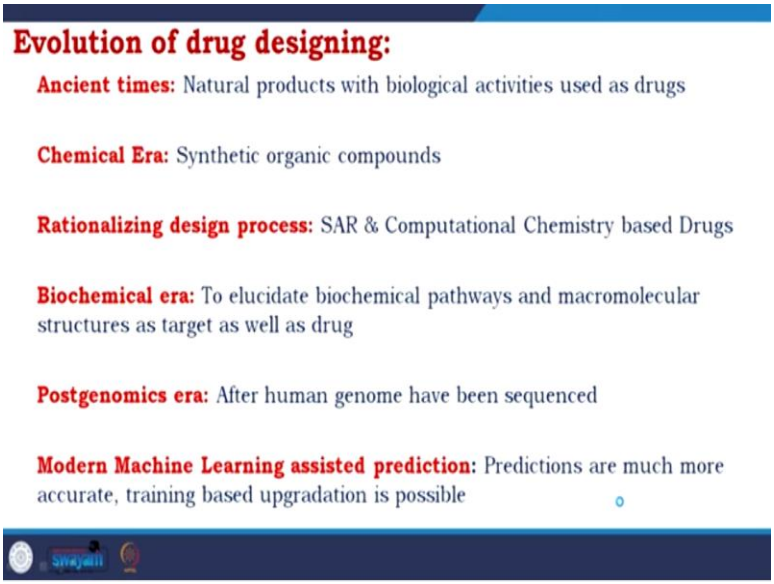
expected to contribute to intelligent lead. The dynamic flexibility concept contributed to the advent of molecular dynamics simulation techniques.

(Refer Slide Time: 01:03:24)



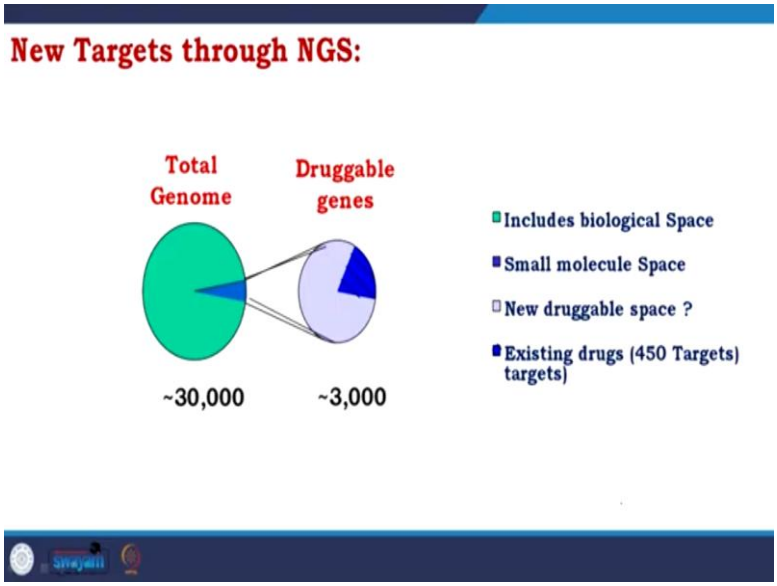
The third generation rational drug designing is the most rational approach employed today. Disease processes are understood at the molecular level. So, we are not only looking at the target drug but also understanding the opposite part of the main biological entity that is the key to success. So, the disease process is understood at the molecular level, and targets are well defined. Drugs can then be designed to bind these targets and disrupt this disease process effectively. A very complex and intellectual approach requires detailed knowledge and information retrieval. CADD (Computer-Aided drug designing), when you say structure aided drug designing or structure-based drug discovery, you cannot put the computer. You utilize it to develop more potent ones using a computer. Drug-receptor interaction is not merely a lock and key interaction but a dynamic and energetically favorable one.

(Refer Slide Time: 01:05:09)



There is an evolution of drug designing, in ancient time natural production with biological activities as a drug. Chemical era: synthetic organic compounds. Rationalizing design process: structur, activity relation (SAR), even computational chemistry with drugs. Biochemical era: to elucidate biochemical pathways and macromolecular structures as the target and drug. Post genomics era: after human genome had been sequenced. Modern machine learning assisted prediction: predictions are much more accurate, training based up-gradation is possible.

(Refer Slide Time: 01:05:50)



NGS is now providing us with new targets, we had the total genome, but now we look into that druggable genes. So, the NGS is not only giving us the genome sequence, but it also gives us the transcript from the protein of the condition, how a genome is affected with the change of conditions, how disease people protein or disease people gene is differently behaving compared to a normal healthy person, all those are giving us a new concept in the drug designing era.

(Refer Slide Time: 01:06:46)