

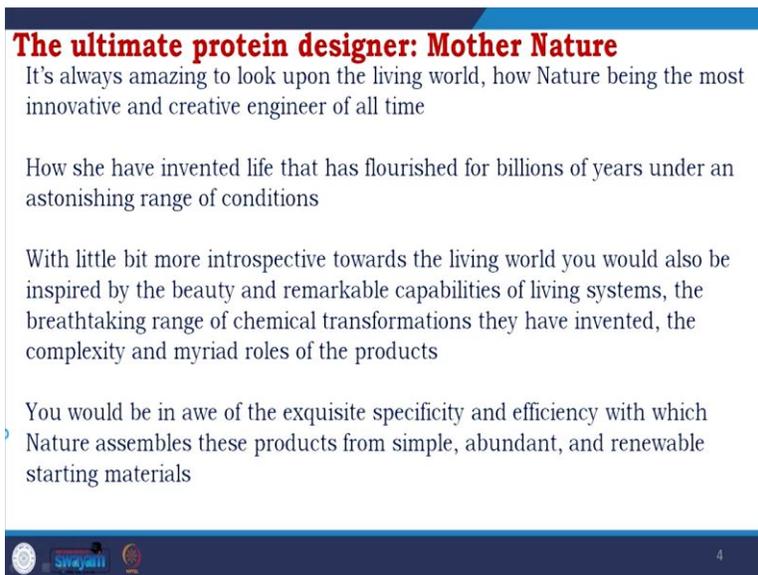
**Structural Biology**  
**Prof. Saugata Hazara**  
**Department of Biotechnology**  
**Indian Institute of Technology - Roorkee**

**Lecture – 54**

**Designing Protein by Mimicking Nature: Process of Directed Evolution**

Hi everyone, welcome again to the course of structural biology, we are going to the module of protein engineering. We have talked about the method of rational design, very unique type of designing which is called de novo designing which could even give you a protein which is not present at all. Today we will continue from there and we will talk about something which is mimicking nature. So, the process of such type of protein engineering is called as directed evolution. As I told earlier directed evolution is a process where we do mimic the nature. To go about the details of the method, first let us agree to an amazing fact which I talked about earlier but you want to talk about these repeatedly.

**(Refer Slide Time: 02:01)**



**The ultimate protein designer: Mother Nature**

It's always amazing to look upon the living world, how Nature being the most innovative and creative engineer of all time

How she have invented life that has flourished for billions of years under an astonishing range of conditions

With little bit more introspective towards the living world you would also be inspired by the beauty and remarkable capabilities of living systems, the breathtaking range of chemical transformations they have invented, the complexity and myriad roles of the products

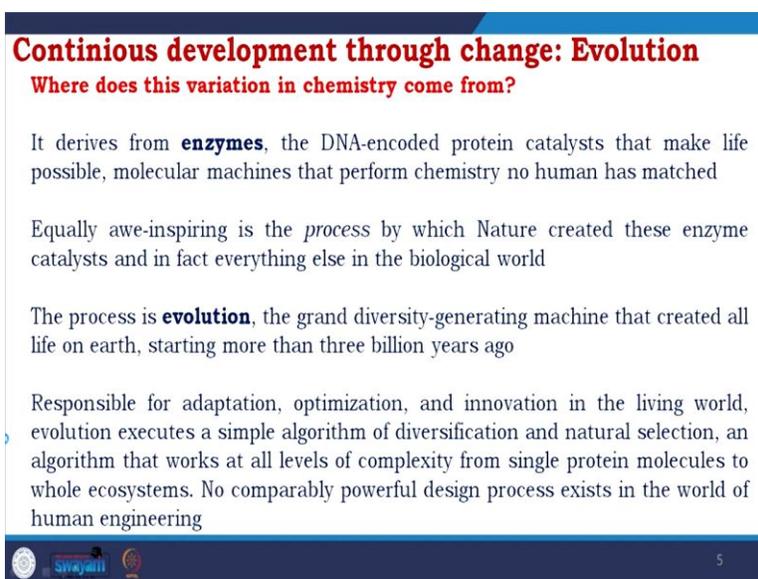
You would be in awe of the exquisite specificity and efficiency with which Nature assembles these products from simple, abundant, and renewable starting materials

4

The ultimate protein designer the ultimate biological macromolecule designer the ultimate designer of life is Mother Nature. So, as a person if you start thinking about the universe about the living system going on, it is always amazing to look up on the living world how nature being the most innovative and creative engineer of all time. How she have invented life that has flourished for billion of years under an astonishing range of conditions, variations diversity. With a little bit more introspective if you go into deep and all towards the living world you would also

be inspired by the beauty and remarkable capabilities of living system. The breathtaking range of chemical transformations they have invented everywhere every day every minute every second. So, many reactions, so many transformations are continuously going through giving the life in a living mode the complexity and the myriad roles of the products. You would be in awe of the exquisite specificity and efficiency with which nature assembles these products from simple, abundant and renewable starting materials. To maintain these for billions of years what nature is continuously doing is development through change. There is a term it is called evolution Mother Nature continuously adapting the changing situation by the tool which is called evolution.

(Refer Slide Time: 04:06)



**Continuous development through change: Evolution**  
**Where does this variation in chemistry come from?**

It derives from **enzymes**, the DNA-encoded protein catalysts that make life possible, molecular machines that perform chemistry no human has matched

Equally awe-inspiring is the *process* by which Nature created these enzyme catalysts and in fact everything else in the biological world

The process is **evolution**, the grand diversity-generating machine that created all life on earth, starting more than three billion years ago

Responsible for adaptation, optimization, and innovation in the living world, evolution executes a simple algorithm of diversification and natural selection, an algorithm that works at all levels of complexity from single protein molecules to whole ecosystems. No comparably powerful design process exists in the world of human engineering

5

Where does this variation in chemistry come from? It derives from enzyme, the DNA encoded protein catalysts that make life possible, molecular machines that perform chemistry, and no human creator could have matched. Equally awe-inspiring is the process by which nature created these enzyme catalysts and in fact everything else in the biological world, the assembling, the complexity, the hierarchies, the maintenance of hierarchies starting from one point, how everything is regenerated? How everything is continuously operating without being stopped? The process is evolution, the grand diversity generating machine that created all life on earth starting more than 3 billion years ago. Responsible for adaptation, optimization and innovation in the living world. Evolution executes a simple algorithm of diversification and natural selection. An algorithm that works at all levels of complexity from single protein molecule to the whole ecosystems. No comparably powerful design process exists in the world of human engineering.

(Refer Slide Time: 05:57)

**Introduction to the process of Directed Evolution:**

Natural evolution of enzymes has existed since the emergence of life on Earth

Genes have mutated and proteins have evolved to improve the fitness of an organism to tackle conditions in new environments

For thousands of years, humans have been breeding animals and plants through the selection of organisms with desired properties

For most of this time without even knowing they were doing it, humans evolved and optimized enzymes and binding proteins over many generations



6

Introduction to the process of directed evolution: The natural evolution of enzymes has existed since the emergence of life on earth. Genes have mutated and proteins have evolved to improve the fitness of an organism to tackle condition in new environments. For example in some cases we see there are 3 enzymes in bacteria, which comes under one complex machinery in human, and they become organized to tackle life. For thousands of years humans have been breeding animals and plants through the selection of organisms with desired properties. For most of this time without even knowing they were doing it, human evolved and optimized enzymes and binding proteins over many generations.

(Refer Slide Time: 09:07)

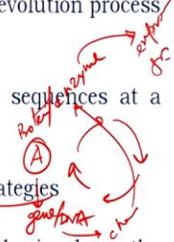
## Introduction to the process of Directed Evolution:

Directed evolution of enzymes and binding proteins is a manmade procedure built on molecular insights, which moves the evolution process into the laboratory and speeds it up

The procedure relies on intended variation of protein sequences at a defined level of randomness

This is coupled to engineered screening and selection strategies

Directed evolution is an iterative procedure which involves the identification of a starting state protein, diversification of its gene, an expression and screening strategy, re-diversification, re-screening, and so on until a satisfactory performance level in terms of enzymatic activity, binding affinity or specificity is reached



Directed evolution of enzymes and binding protein is a manmade procedure built on molecular insights which moves the evolution process into the laboratory and speeds it up. The procedure relies on intended variation of protein sequences at a defined level of randomness. This is coupled to engineered screening and selection strategies. So, If you even go to the ancient time, Gregor Johann Mendel taken the pea plants and break them the criteria, the selection was probably been on vision probably there was not instrumental aid to look at the quality of the piece. But it was looking at the color, the size and all those things. That is engineered screening and ultimately that leads to selection strategies. Directed evolution is an iterative procedure which involves the identification of a starting state protein, diversification of its gene, an expression and screening strategy re-diversification, re-screening and so on until a satisfactory performance level in terms of enzymatic activity, binding affinity or specificity is reached.

**(Refer Slide Time: 13:12)**

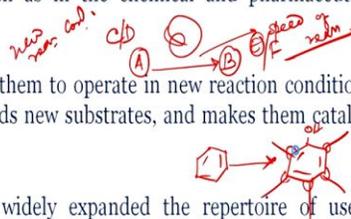
## Introduction to the process of Directed Evolution:

Directed evolution of enzymes and binding proteins has become a widely used strategy in academic research as well as in the chemical and pharmaceutical industries

Directed evolution of enzymes tailors them to operate in new reaction conditions, optimizes their catalytic activity towards new substrates, and makes them catalyze new chemical reactions

Directed evolution of enzymes has widely expanded the repertoire of useful biocatalysts

The evolved enzymes offer efficient and environmentally-friendly alternatives to metals and organic catalysts in chemical and biotechnical industries



Directed evolution of enzymes and binding proteins has become a widely used strategy in academic research as well as in the chemical and pharmaceutical industries. Directed evolution of enzymes tailors them to operate in new reaction condition, optimizes their catalytic activity towards new substrate and makes them catalyze new chemical reaction.

Directed evolution of enzymes as widely expanded the repertoire of useful biocatalyst. So, you have number of enzymes, but by doing directed evolution you get many different variations. So, you make diversification. The evolved enzyme offer efficient and environmentally-friendly alternatives to metals and organic catalysts in chemical and biotechnological industries. Just to give you more idea. You know whenever you do something where a CH bond breaks and new functional bond forms, that are challenging in chemistry, like if you convert benzene to phenol apparently this is very simple reaction, but if you want to do it chemically it needs so many steps, it is so expensive. Now you could use the enzyme to do that because already we know about enzyme, about the specificity of the enzyme. Also, it could not only make it phenol, it could also help you introducing functional group in a desired position without doing any blocking of those groups. So that is why the evolved enzymes offer efficient and environmentally friendly alternatives to metals and organic catalysts in chemical and bio technological industries.

**(Refer Slide Time: 16:30)**

## Directed evolution of enzymes and binding proteins – in theory:

In 1984, Manfred Eigen published a theoretical paper outlining a possible workflow for directed evolution of enzymes

Eigen noted that such optimization becomes an interesting challenge because the genotype and the phenotype are dependent on different molecules

He reasoned that finding rare improved variants in large libraries would be hard if not impossible

Instead he proposed the use of smaller libraries and several generations of mutagenesis and screening as a procedure that would more likely lead forward

So, directed evolution of enzymes and binding proteins in theory, what was the theory based on? In 1984 Manfred Eigen published a theoretical paper outlining a possible workflow for directed evolution of enzymes. Eigen noted that such optimization becomes an interesting challenge because the genotype and the phenotype are dependent on different molecule. What is genotype? Genotype is the characteristics at genetic level and phenotype is the expression level, and what you are looking at the functional level. He reasoned that finding rare improved variant in large libraries would be hard but not impossible. Instead he proposed the use of smaller libraries and several generations of mutagenesis and screening as a procedure that would more likely lead forward.

**(Refer Slide Time: 19:53)**

## Directed evolution of enzymes and binding proteins - in theory:

Another theoretical prediction is found in a patent that describes evolution of binding proteins through iterative diversification of libraries between selection rounds

Directed evolution of enzymes was also briefly introduced as a possibility in an abstract of a more classical protein-engineering study of enzyme optimization

Controlled optimization by the above procedure really becomes an interesting challenge when genotype and phenotype are different molecules.

Let us therefore expand the procedure as follows -

- 10 PRODUCE A MUTANT SPECTRUM OF SELF-REPRODUCING TEMPLATES
- 20 SEPARATE AND CLONE INDIVIDUAL MUTANTS
- 30 AMPLIFY CLONES
- 40 EXPRESS CLONES
- 50 TEST FOR OPTIMAL PHENOTYPES
- 60 IDENTIFY OPTIMAL GENOTYPES
- 70 RETURN TO 10 WITH A SAMPLE OF OPTIMAL GENOTYPES



10

Another theoretical prediction is found in a patent that describes evolution of binding proteins through iterative diversification of libraries between selection rounds. Directed evolution of enzyme was also briefly introduced as a possibility in an abstract of a more classical protein engineering study of enzyme optimization. This is the form of where control optimization by the above procedure really becomes an interesting challenge.

(Refer Slide Time: 22:24)

## Directed evolution of enzymes and binding proteins - in practice:

One decade after Eigen's theoretical work, the first experimental work appeared that described successful implementation of directed evolution of enzymes in a laboratory setting to improve enzyme function and versatility

**Frances H Arnold** reported the directed evolution of subtilisin E to obtain an enzyme variant which was active in a highly unnatural (denaturing) environment, i.e. at high concentrations of the polar organic solvent dimethylformamide (DMF)

After four sequential rounds of mutagenesis and screening in the presence of DMF, an enzyme variant with 256-fold higher activity than the wild-type enzyme in 60% (v/v) DMF was created

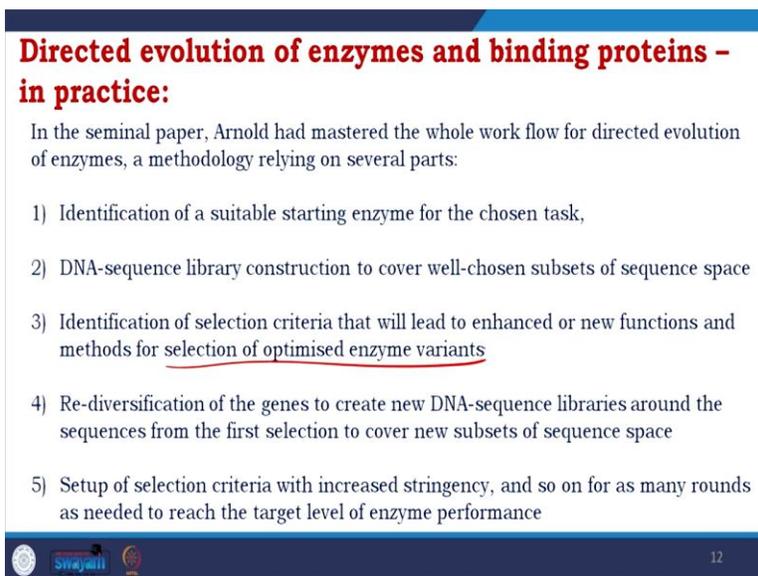


11

One decade after Eigen's theoretical work, the first experimental work appeared that described successful implementation of directed evolution of enzymes in a laboratory setting to improve enzyme function and versatility. Frances Arnold reported the directed evolution of subtilisin E to obtain an enzyme variant which was active in a highly unnatural environment which is a

denaturing environment at high concentration of the polar organic solvent dimethylformamide or DMF. After four sequential rounds of mutagenesis and screening in the presence of DMF, enzyme variant with 256 fold higher activity than the wild type enzyme in 60% DMF was created. So, this was the first enzyme engineering in reality using directed evolution.

**(Refer Slide Time: 26:15)**



**Directed evolution of enzymes and binding proteins - in practice:**

In the seminal paper, Arnold had mastered the whole work flow for directed evolution of enzymes, a methodology relying on several parts:

- 1) Identification of a suitable starting enzyme for the chosen task,
- 2) DNA-sequence library construction to cover well-chosen subsets of sequence space
- 3) Identification of selection criteria that will lead to enhanced or new functions and methods for selection of optimised enzyme variants
- 4) Re-diversification of the genes to create new DNA-sequence libraries around the sequences from the first selection to cover new subsets of sequence space
- 5) Setup of selection criteria with increased stringency, and so on for as many rounds as needed to reach the target level of enzyme performance

12

In the seminal paper Arnold has mastered the whole workflow for directed evolution of enzymes a methodology which was continued by others in the following years. So, the protocol was having different steps, the first step was identification of a suitable starting enzyme for the chosen task. DNA sequence library construction to cover well chosen subset of sequence space. Identification of selection criteria that will lead to enhanced or new functions and method for selection of optimized enzyme variants. One of the major challenges of directed evolution lies here. So, you get a starting material, you get the gene, and you make changes. Now you kept them in a library. You have to screen from those library to the candidates which are better. And for each and every enzyme checking betterment is different. Re-diversification of the genes to create new DNA sequence libraries around the sequences from the first selection to cover new sub set up sequence space. And the last step, set up of selection criteria with increased stringency and so on for as many rounds as needed to reach the target develop enzyme performance.

**(Refer Slide Time: 28:50)**

### Directed evolution of enzymes and binding proteins – in practice:

Each of these five steps has since been further developed and optimized over the years in the Arnold lab and several other labs

In addition to a first set of four combined single mutations, the first work used error-prone PCR to create and re-diversify DNA-sequence libraries through three rounds of random mutagenesis and screening to evolve subtilisin E

The selection criterion was hydrolysis of the milk protein casein. Active enzyme variants created visible halos on agar plates with casein

Enzymes secreted by bacterial colonies were thus transferred to agar plates containing both DMF and casein, to enable identification of the most active enzyme variants in the presence of the organic solvent

Plasmid DNA was isolated from clones secreting an enzyme variant that produced a halo larger than those surrounding the parent enzyme, and subjected to further rounds of mutagenesis



Each of these five steps has since been further developed and optimized over the years in the Arnold Lab and several other labs. In addition to a first set of four combined single mutation, the first work use error prone PCR to create and re diversify DNA sequence libraries through three rounds of random mutagenesis and screening to evolve substilisin E. The selection criterion was hydrolysis of the milk protein casein; active enzyme variants were created visible halos on agar plates with casein. So, you need something which you could see or which you could differentiate that is the first criteria of selection, also you need something which would make difference between one successful mutants. Now enzymes secreted by bacterial colonies were thus transfer to agar plates containing both DMF dimethylformamide and casein to enable identification of the most active enzyme variants in the presence of organic solvent. Then the plasmid DNA would be isolated from clones secreting an enzyme variant that produce a hello larger than those surrounding the parent enzyme and subjected to further rounds of mutagenesis.

**(Refer Slide Time: 30:49)**

### Directed evolution of enzymes and binding proteins - in practice:

The directed evolution of subtilisin E to improve its activity in a polar organic solvent was a benchmark achievement that opened the field of directed evolution of enzymes

This work became the starting point for continued technical development of the methodology for directed evolution

The field expanded towards improving and reshaping enzymes for numerous chemical reactions, old and new, leading to applications of importance for research in organic synthesis, as well as for the chemical and pharmaceutical industry and beyond

The directed evolution of subtilisin E to improve its activity in a polar organic solvent was a benchmark achievement that opened the field of directed evolution of enzymes. This work became the starting point for continued technical development of the methodology of directed evolution. The field expanded towards improving and reshaping enzymes for numerous chemical reactions, old and new, leading to application of importance for research in organic synthesis as well as for the chemical and pharmaceutical industry and beyond.

**(Refer Slide Time: 32:00)**

### Molecular insights guide library design:

It is not possible to randomize every position in an enzyme, the typical size of which is 200-300 amino acid residues or more. Indeed, only a small fraction of the amino acid positions can be varied if the aim is a library with full sequence coverage.

The reason is simple combinatorial mathematics and the quickly growing number of variants relative to the number of clones that can be handled in any laboratory setting, or even using the joint capacity of all laboratories in the world.

Still, a wealth of studies makes it clear that mutations in and nearby the active site as well as more distant substitutions on the enzyme surface may contribute to optimized catalytic activity.

Arnold and co-workers have shown by many examples that library design must be based on molecular insight and knowledge-based choices of which amino-acid positions to vary, combined with some element of added randomness, e.g. through error-prone PCR.

Molecular insights guide library design: It is not possible to randomize every position in an enzyme the typical size of which is 200 to 300 amino acid residues or more. Indeed only a small fraction of the amino acid position can be varied if the aim is a library with full sequence coverage. The reason is simple combinatorial mathematics and the quickly growing number of variants relative to the number of clones that can be handled in any laboratory setting or even using the joint capacity of all laboratories in the world. Still a world of studies makes it clear those mutations in and nearby the active site as well as more distance substitution on the enzyme surface may contribute to optimize catalytic activity. So, we start understanding if all the amino acids are equally important or the amino acids for enzyme which is taking part in the enzyme reaction are important. So, we get a set up limited number of amino acids to change. So that is called knowledge based selection. For a binding protein we are trying to understand, we are trying to gain the information regarding the number of amino acid residues plays a critical part in binding or other residues which helps the critical residues to be in optimal Position. Arnold and co workers have shown by many examples that library design must be based on molecular insight and knowledge based choices of which amino acid positions to vary combined with some element of added randomness for example through error prone PCR.

**(Refer Slide Time: 37:14)**

## Molecular insights guide library design:

A prominent early contributor to the development and implementation of methodology for directed evolution was the late William (Pim) Stemmer (2013)

Stemmer introduced a DNA recombination strategy termed "DNA shuffling" to the evolution of enzymes. This was an efficient way to propagate beneficial mutations while increasing the size of a DNA library through random fragmentation and re-assembly of genes

 p53 → tetramer

He showed that the use of DNA shuffling, i.e. recombination of DNA from similar genes from several organisms, introduces more variation than many other methods and can thus improve the chances to reach a substantial activity increase in the evolved variants

In a proof-of-principle study Stemmer and co-workers set out to increase the activity of the enzyme  $\beta$ -lactamase (an enzyme responsible for antibiotic resistance)

A prominent early contributor to the development and implementation of methodology for directed evolution was the late. William Pim Stemmer. Stemmer introduced a DNA recombination strategy which was termed as DNA shuffling to the evolution of enzymes. This was an efficient way to propagate beneficial mutations while increasing the size of a DNA library through random fragmentation and reassembly of genes. So, mostly the concept of DNA shuffling, if I want to explain in a very simple way, takes a protein called P53. P53 is that gene response responsible for cancer, it generally form tetramer. Now if you take the tetramer developing region and fuse it to an enzyme which form a monomer this will help the enzyme to form a tetramer and increase its efficiency. So, you have different fragments, you make pieces and then you paste randomly, what you do is DNA shuffling.

He showed that the use of DNA shuffling which is recombination of DNA from similar genes from several organism. So, for initial experiments of DNA shuffling, he brought different homologous enzymes from different organisms. And when you do DNA shuffling of, it would introduce more variation than many other methods and can thus improve the chances to reach a substantial activity increase in the evolved variance. In a proof of principle studies Stemmer and co workers set out to increase the activity of the enzyme beta lactamase.

**(Refer Slide Time: 41:29)**

### **Molecular insights guide library design:**

Three cycles of DNA shuffling and screening on plates with successively higher concentration of the antibiotic cefotaxime led to evolution of an enzyme with significantly increased activity

Gene shuffling had been reported as a means to increase variation and improve antibody affinity for a target, sometimes called affinity maturation, and provided early examples of directed evolution of binding proteins

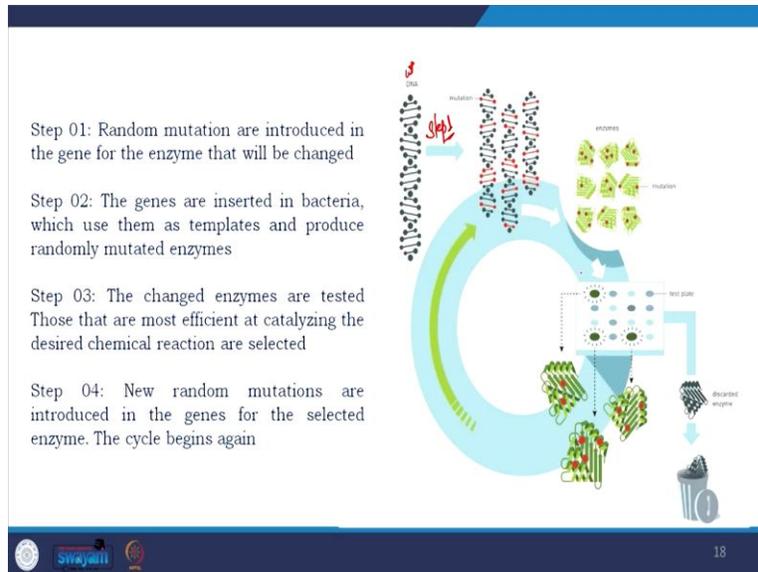
In these cases the shuffled gene segments corresponded to light chains or to the variable loops of heavy and light chains of immunoglobulins



Three cycles of DNA shuffling and screening on plates with successively higher concentration of the antibiotic cefotaxime (which is a cephalosporin) led to evolution of an enzyme with significantly increased activity. Gene shuffling had been reported as a means to increase variation and improve antibody affinity for a target sometimes called affinity maturation which has happened in the real organism real cell and provided early example of directed evolution of binding proteins.

In these cases the shuffle gene segment correspond to light chains or to the variable loop up heavy and light chain of immunoglobulins. So, where the variation happens those are more important for improvement.

**(Refer Slide Time: 42:19)**



In step one, you have the DNA which is wild type DNA, the first step is random mutation introduced. Then the genes are inserted in bacteria which used them as template and produce randomly mutated enzymes. In the step 3, the change enzyme are tested, those that are most efficient at catalyzing the desired chemical reaction are selected. Step 4 new random mutations are introduced in the gene for the selected enzyme and the cycle begins again.

**(Refer Slide Time: 43:31)**

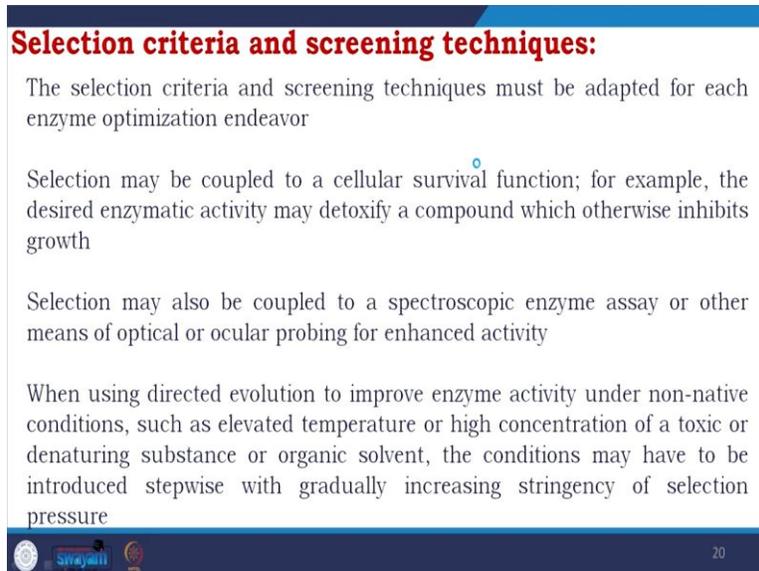
DNA shuffling, staggered extension process (StEP) and other methods for library generation were further developed in Arnold's and Stemmer's laboratories for use in the directed evolution of enzymes during the second half of the 1990s

Since the end of last century, the progressively decreasing costs for *de novo* synthesis of genes with degenerate codons, or fully designed DNA libraries, have opened a new path towards efficient and affordable production of sequence libraries with tailored diversity.

DNA shuffling staggered extension process which is called StEP and other method of library generation were further developed in Arnlod's and Stammer's laboratories for use in the directed evolution of enzymes during the second half of the 1990's. Since the end of last century the progressively decreasing costs for *de novo* synthesis of genes with degenerate codons or fully

designed DNA libraries have opened a new path towards efficient and affordable production of sequence libraries with tailored diversity.

**(Refer Slide Time: 44:39)**



**Selection criteria and screening techniques:**

The selection criteria and screening techniques must be adapted for each enzyme optimization endeavor

Selection may be coupled to a cellular survival function; for example, the desired enzymatic activity may detoxify a compound which otherwise inhibits growth

Selection may also be coupled to a spectroscopic enzyme assay or other means of optical or ocular probing for enhanced activity

When using directed evolution to improve enzyme activity under non-native conditions, such as elevated temperature or high concentration of a toxic or denaturing substance or organic solvent, the conditions may have to be introduced stepwise with gradually increasing stringency of selection pressure

20

Selection criteria and screening technique are the most important part in the success of directed evolution. The selection criteria and screening techniques must be adapted for each enzyme optimization endeavor. Selection may be coupled to a similar survival function for example the desired enzymatic activity may detoxify a compound which otherwise inhibits growth. Selection may also be coupled to a spectroscopic enzyme assay or other means of optical or ocular probing for enhanced activity. When using directed evolution to improve enzyme activity under non native conditions. Such as elevated temperature or high concentration of a toxic or denaturing substance or organic solvent the condition may have to be introduced step wise with gradually increasing stringency of selection pressure otherwise you could lose intermediates more importantly you probably lose everything because you have to understand that. At what level of toxicity, at what temperature, at what criteria we could make selection at least few of them would be there. If you set in such standard when no one is selected then your process would fell.

**(Refer Slide Time: 46:38)**

## Selection criteria and screening techniques:

This ensures that the enzyme template in each round has at least some rudimentary starting activity under the conditions used in that round

Stepwise increase of the stress factor, with intervening diversification between selection rounds, makes it possible to derive successively more effective and more tolerant enzymes or enzymes with new catalytic properties

Selection may be performed on agar plates, using a filter-lift assay or using flow cytometry

Dan Tawfik showed that directed evolution of enzymes can be set up without use of living cells, for example, using *in vitro* compartmentalization in water-in-oil emulsion droplets containing ribosomes and library mRNA

This ensures that the enzyme template in each round has at least some rudimentary starting activity under the condition used in that round. Otherwise you cannot do the fact that. Stepwise increase of the stress factor with intervening diversification between selections rounds makes it possible to derive successfully more effective and more tolerant enzymes are enzymes with new catalytic properties. Selection may be performed on agar plates using a filter lift assay where you put filter and you could leave and check or using flow cytometry in modern days. Dan Tawfik showed that directed evolution of enzyme can be set up without use of living cells. For example using *in vitro* compartmentalization in water in oil emulsion droplets containing ribosome and library of mRNA.

**(Refer Slide Time: 47:36)**

### New reaction conditions:

The early applications of directed evolution of enzymes aimed to optimize the stability and the performance under new reaction conditions such as high fractions of organic solvents, further rounds of directed evolution were added to reach a 471-fold activity increase over wild-type

Many methods exist for increasing the thermostability of enzymes and other proteins

When directed evolution of enzymes is used with an aim to increase their thermal stability, the evolutionary process may be set up as interleaved heat treatment and activity assays, or alternatively the activity assays may be performed at elevated temperature

In one study, the thermostability of *Bacillus subtilis* *p*-nitrobenzyl esterase was increased by over 14°C (increase in  $T_m$ ) after six generations of random mutagenesis, recombination via DNA shuffling, and screening with interleaved heat treatment and activity assays



The early application of directed evolution of enzymes aimed to optimize the stability and the performance under new reaction conditions such as high fraction of organic solvents further rounds of directed evolution were added to reach a 471 fold activity increase over wild type. Many methods exist for increasing the thermo stability of enzymes and other proteins. When directed evolution of enzymes used with an aim to increase the thermal stability the evolutionary process may be set up as interleaved heat treatment and assays or alternatively the activity assays may be perform at elevated temperature or higher temperature. In one study the thermo stability of *Bacillus subtilis* *p*- nitrobenzyl esterase was increased by over 14 degrees centigrade after six generation of random mutagenesis, recombination via DNA shuffling, and screening with interleaved heat treatment and activity assays.

**(Refer Slide Time: 49:22)**

### New reaction conditions:



This work showed that it is possible to improve the thermal stability of an enzyme without compromising its catalytic activity at lower temperatures, if both properties are constrained

If not, the evolution of one property may come at the cost of the other, regardless of whether the two properties are inversely correlated or not correlated at all

This work showed that it is possible to improve the thermal stability of an enzyme without compromising its catalytic activity at lower temperature if both properties are constant. The evolution of one property then may come at the cost of the other regardless of whether the two properties are inversely correlated or not correlated at all.

(Refer Slide Time: 50:22)

### New reaction conditions:

Nature usually provides organisms adapted to cold or warm environments with two different enzymes having optimal catalytic properties at low or at high temperatures, respectively

Arnold showed that directed evolution can produce a single enzyme with high catalytic activity at both high and low temperatures

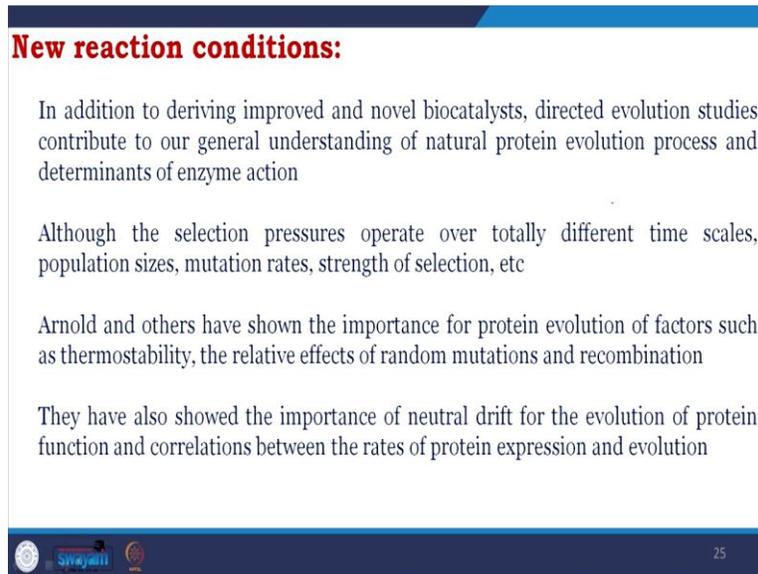
Another directed evolution strategy relied on structural information in the form of crystallographic B-factors, a measure of which regions are more or less ordered in a crystallized protein

By focusing the library of mutations to the 10 positions with highest B-factors, a large increase in enzyme stability was achieved

Nature usually provides organism adapted to cold or warm environment with 2 different enzymes having optimal catalytic properties at low or at high temperatures. Arnold showed that directed evolution can produce a single enzyme with high catalytic activity at both high temperature as well as low. Another directed evolution strategy relied on structural information in the form of crystallographic B factors a measure of which regions are more or less ordered in a

crystallized protein. If you remember I have talked about that if you remember my explanation of the coordinate files the pdb file you will see the thermal factor I talked about that is called B factor. By focusing the library of mutations to the 10 positions with highest B factor a large increase in enzyme stability was achieved.

**(Refer Slide Time: 52:07)**



**New reaction conditions:**

In addition to deriving improved and novel biocatalysts, directed evolution studies contribute to our general understanding of natural protein evolution process and determinants of enzyme action

Although the selection pressures operate over totally different time scales, population sizes, mutation rates, strength of selection, etc

Arnold and others have shown the importance for protein evolution of factors such as thermostability, the relative effects of random mutations and recombination

They have also showed the importance of neutral drift for the evolution of protein function and correlations between the rates of protein expression and evolution

25

In addition to deriving improved and noble bio catalysts directed evolution studies contribute to our general understanding of natural protein evolution process and determinants of enzyme action. Although the selection pressures operate over totally different time scales population size mutation rate strength of selections etcetera. Arnold and others have shown the importance of protein evolution of factors such as thermal stability the relative effect of random mutations and recombination. They have also showed the importance of neutral drift for the evolution of protein function and correlation between the rates of protein expression and evolution.

**(Refer Slide Time: 53:33)**

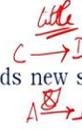
### Choice of starting state:

Arnold and co-workers have repeatedly shown that it is possible to evolve enzymes to improve their activity under new conditions in terms of solution composition, temperature, etc

It is also possible to change their catalytic activity towards new substrates and reactions

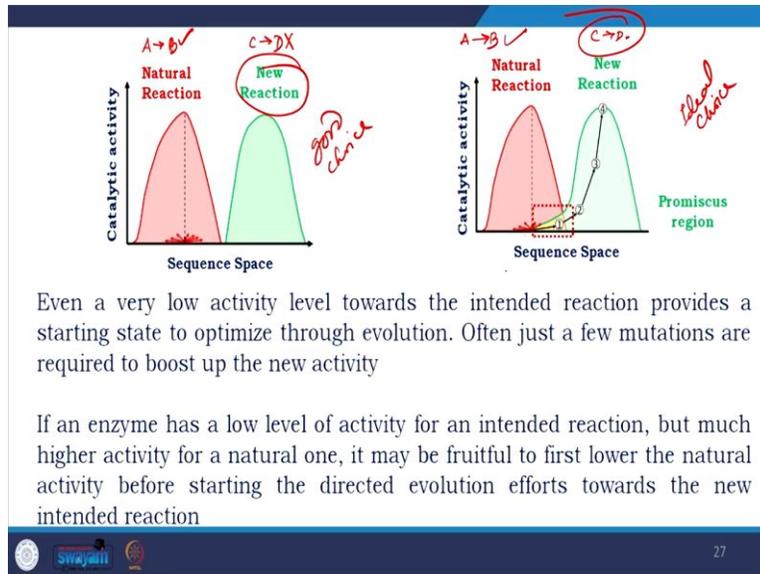
This is possible as long as the enzyme that is chosen as a starting point has at least some low level of activity for the intended reaction, i.e. some level of catalytic promiscuity

An inactive scaffold is not a suitable choice; directed evolution requires some low level of activity



Arnold and co workers have repeatedly shown that it is possible to evolve enzyme to improve their activity under new conditions in terms of solution composition temperature and all. It is also possible to change their catalytic activity towards new substrates or a completely new reaction. This is possible as long as the enzyme that is chosen as a starting point has at least some low level of activity for the intended reaction that is some level of catalytic promiscuity. An inactive scaffold is not a suitable choice directed evolution requires some low level of activity.

**(Refer Slide Time: 55:25)**



Even a very low activity level towards the intended reaction provides a starting state to optimize through evolution often just a few mutations are required to boost up the new activity. If an enzyme has a low level of activity for an intended reaction but much higher activity for the natural one it may be fruitful to first lower the natural activity before starting the directed evolution effort towards the new intended reaction.

**(Refer Slide Time: 58:07)**

### **New chemical reactions:**

As a recent example of this latter strategy, the activity of tryptophan synthetase from *Pyrococcus furiosus* was first reduced by 95% through the removal of the non-catalytic domains of the enzyme

The isolated catalytic domain was subject to three rounds of directed evolution to introduce new catalytic activities towards synthesis of tryptophan analogues

In a series of studies, Arnold and co-workers changed the activity of cytochrome P450 to catalyze a set of reactions for which no specific enzyme was previously available, for example, cyclopropanation

Cytochrome P450BM3 has a catalytic promiscuity and an ability to catalyze, with very low efficiency, the cyclopropanation of styrene by ethyl-diazoacetate (EDA)



A new chemical reaction, as a recent example of this latest strategy the activity of tryptophan synthetics from *pyrococcus furiosus* was first reduced by 95% through removal of the non catalytic domain of the enzyme. The enzyme has another domain, which is a non catalytic domain, you remove it that would reduce the original reaction catalysis and then engineer it. The isolated catalytic domain was subjected to three rounds of directed evolution to introduce new catalytic activities towards synthesis of tryptophan analogues. In a series of studies, Arnold and co workers changed the activity of cytochrome P450 to catalyze a set of reactions for which no specific enzyme was previously available. For example they had achieved a process called cyclopropanation which was not done ever by any biological enzyme because there was no biological importance of cyclopropanation. But industrially it is very important cytochrome P450M3, has a catalytic promiscuity and an ability to catalyze, with very low efficiency, the cyclopropanation of styrene by a ethyl-diazoacetata.

**(Refer Slide Time: 59:41)**

### **New chemical reactions:**

Much more specific and efficient enzymes were evolved and only a small fraction (0.2%) of the amino acids in the enzyme needed to be changed to optimize the new catalytic activity

This included a change of the iron-ligating residue from Cys to Ser or His, leading to a shift in the characteristic 450-nm Soret peak in the absorbance spectrum of the enzyme to 411 nm, the evolved enzymes was called as cytochrome P411

Other examples of reactions for which no natural enzymes have evolved are nitrene transfer reactions

In one case, Arnold and co-workers started from a cytochrome P411 variant performs azide reduction about 100 times more efficiently than nitrene transfer to sulphide



Much more specific and efficient enzymes are evolved and only a small fraction 0.2% of the amino acids in the enzyme needed to be changed to optimize the new catalytic activity. These included a change the iron-ligating residue from Cys to Serine or His, leading shift in the characteristic 450 nanometer, which is because of the Soret peak comes at the 450 that is sifted to 411 nanometer, and they call the model modified enzyme was called as cytochrome P411. Other example of reactions for which no natural enzymes have evolved our nitrene transfer reactions. In one case Arnold and co worker started from a cytochrome P411. The modified one variant and perform azide reduction about 100 times more efficiently than nitrene transferred to sulphide.

**(Refer Slide Time: 1:00:43)**

### **New chemical reactions:**

Using directed evolution they produced an enzyme variant that instead efficiently promotes the desired nitrene transfer process

There are several other examples of directed evolution of enzymes for carbene and nitrene transfer reactions

Reactions with aliphatic and aromatic CH bonds are another tractable goal

Using directed evolution of cytochrome P450 monooxygenase, an enzyme was created that catalyzes intermolecular amination of benzylic C-H bonds

The biocatalyst is enantioselective and lasts for up to 1,300 turnovers, thereby providing an efficient biocatalyst for synthesis of valuable benzylic amines



30

Using directed evolution they produced an enzyme variant that instead efficiently promotes the desired nitrene transfer process. There is several other example of directed evolution of enzymes carbon and nitrene transfer reactions. Reactions with aliphatic and aromatic CH bonds are another tractable goal. Using directed evolution of cytochrome P450 monooxygenase an enzyme was created that catalyzes intermolecular amination of benzylic C-H bonds. The biocatalyst is enhanced your selective and lasts for up to 1300 turnovers thereby providing an efficient biocatalyst percentage is valuable benzylic amines.

**(Refer Slide Time: 1:02:23)**

### **Other examples of evolution towards new reactions include,**

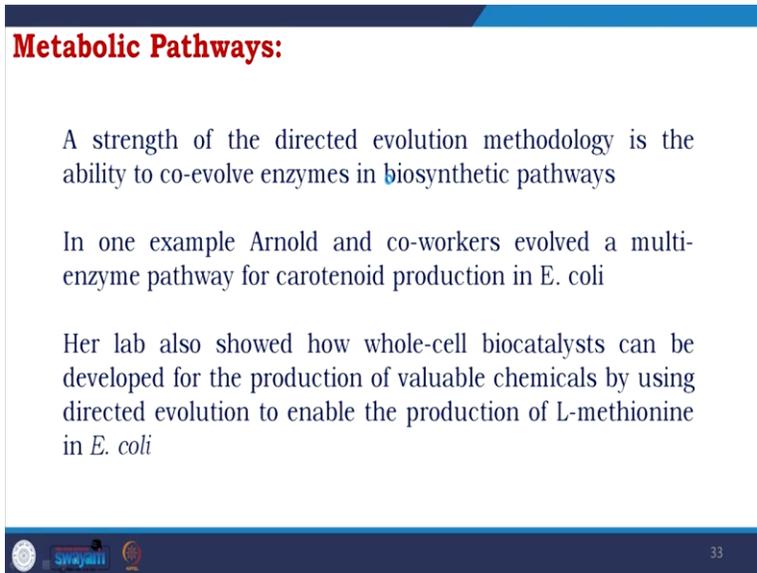
- A) Generation of enzymes that catalyze arsenate detoxification
- B) The production of highly strained carbocycles
- C) The switching of an enzyme from a galactosidase to a fucosidase



31

There are many examples but few to mention, generation of enzyme that catalyze arsenate detoxification which is industrially important. The production of highly strained carbocycles, and the switching of an enzyme from galactosidase to fucosidase.

**(Refer Slide Time: 1:02:57)**



**Metabolic Pathways:**

A strength of the directed evolution methodology is the ability to co-evolve enzymes in biosynthetic pathways

In one example Arnold and co-workers evolved a multi-enzyme pathway for carotenoid production in *E. coli*

Her lab also showed how whole-cell biocatalysts can be developed for the production of valuable chemicals by using directed evolution to enable the production of L-methionine in *E. coli*

33

The strength of the directed evolution methodology is the ability to co evolve enzyme in biosynthetic pathways. In one example, Anrold co workers evolved a multi enzyme pathway of carotenoid production in *E. coli*. Her lab also showed how wholesale bio catalysts can be developed for the production of valuable chemicals by using directed evolution to enable the production of L-methionine in *E.coli*.

**(Refer Slide Time: 1:03:41)**

## Biofuels:

One challenge for mankind is finding suitable replacements or supplements for fossil fuels, which can be produced in a sustainable and environmentally-friendly manner

Here, one seeks to produce alcohols from short-chain alkanes and a leading candidate biofuel is 2-methylpropan-1-ol (isobutanol)

Isobutanol can be produced using a biosynthetic pathway in recombinant *Escherichia coli*

Two enzymes in the pathway, however, require reduced nicotinamide adenine dinucleotide phosphate (NADPH) as a cofactor, while glycolysis, the normal metabolism during growth of *E. coli* produces reduced nicotinamide adenine dinucleotide (NADH)



One challenge for mankind is finding suitable replacement or supplements for fossil fuels which can be produced in a sustainable and environmentally friendly manner. This is something which we all are now focusing on, because we have enhanced population day by day enhancing demanding more food, demanding more resources, using sustainable resources are the challenges in development of biofuels as you need to replace the fossil fuel. The target is to produce alcohol from short-chain alkanes and a leading candidate biofuel is 2 methylpropane-1-ol, also known as isobutanol. Isobutanol can be produced using a biosynthetic pathway in recombinant *Escherichia coli*. Two enzymes in the pathway, however, require reduced nicotinamide adenine dinucleotide phosphate NADPH as a cofactor, while glycolysis the normal metabolism during growth *E. coli* produces reduced nicotinamide adenine dinucleotide.

**(Refer Slide Time: 1:05:04)**

### **Biofuels:**

To resolve this obstacle, Arnold and co-workers used directed evolution to alter the co-factor dependence of the enzymes so they can instead rely on NADH, making the enzymes and thereby the organism suitable for biofuel production

In India, we also have groups involved in metabolic, enzyme engineering based biofuel production

One such group is there in IIP involving scientists like Dr. Debasis Ghosh, Dr. Diptarka Dasgupta etc.

To resolve this obstacle, Arnold and co workers use directed evolution to alter the cofactor dependence of the enzymes. So that they instead rely on both rely on NADH making the enzymes and thereby the organism suitable for biofuel production. In India we also have a lot of groups working in metabolic enzyme engineering based biofuel production. One such group is there in IIP Dehradun, Indian Institute petroleum involving scientists like Dr. Dabasis Ghosh, Dr. Diptarka Dasupta and team.

**(Refer Slide Time: 1:05:42)**

## New Chemical Bonds:

Carbon-silicon bonds are common in human-made chemicals but absent in biology

Nature has not evolved enzymes that catalyze the formation of carbon-silicon bonds

However, directed evolution can be used as a strategy to ensure that such chemistry invented by humans can also be conducted by help of enzymes

Arnold and co-workers noted that heme proteins can catalyze non-natural carbene-insertion reactions

After screening a number of heme proteins from various organisms, they decided to use cytochrome c from *Rhodothermus marinus* as a starting point

This protein catalyzes the formation of carbon-silicon bonds with low efficiency, but with 97% enantiomeric excess



Carbon-silicon bond is also common in human made chemical but they are not present in biology. Nature has not evolved enzymes that catalyze the formation of carbon silicon bonds. However directed evolution can be used as a strategy to ensure that such chemistry invented by humans can also be conducted by the help of enzymes. Arnold and co workers noted that heme proteins can catalyze non natural carbene insertion. After screening a number of heme proteins from various organisms, they decided to use cytochrome c from *Rhodothermus marinus* as a starting point. This protein catalyzes the formation of carbon silicon bonds with low efficiency but with 97% enantiomeric excess.

**(Refer Slide Time: 1:07:08)**

### New Chemical Bonds:

A small library of variants was screened during heat treatment and in catalytic activity assays, and the best candidate was subjected to further mutagenesis and screening

The result of this work is an enzyme that catalyzes silicon-carbon bond formation 40 times better than the starting enzyme and with 99% enantiomeric excess

The evolved enzyme had 15 times higher turnover number than the best non-enzyme catalyst known for the same reaction

This example shows that it is possible to expand the scope of enzyme-catalyzed reactions in terms of which kinds of bonds are formed by the engineered enzyme

Other examples of bonds and reactions not catalyzed by any enzyme found in nature, but for which directed evolution was used to create efficient enzymes, are carbon-borane bonds and enantio-selective intramolecular C-H amination



A small library of variant was screened during heat treatment and in catalytic activity assays and the best candidate was subjected to further mutagenesis and screening. The result of this work is an enzyme that catalyzes silicon carbon bond formation 40 times better than the starting enzyme and 99% enantiomeric excess. The evolved enzyme had 15 times higher turnover number than the base non enzyme catalysts known for the same reaction. This example shows that it is possible to expand the scope of enzyme catalyzed reaction in terms of which kinds of bonds are formed by the engineered enzyme. Other example of bonds and reactions not catalyzed by any enzyme found in nature but for which directed evolution was used to create efficient enzymes are carbon borane bond and enantio-selective intramolecular C-H amination.

**(Refer Slide Time: 1:08:07)**

## Enantioselectivity of Enzymes:

Directed evolution is an efficient way to improve the enantioselectivity of enzymes, i.e., enhancing their performance in asymmetric catalysis

The evolved enzymes are used in the production of chiral substances with high enantiomer purity

An early example of directed evolution with the aim of improving the enantioselectivity of an enzyme was reported by Matcham and Bowen concerning transaminases in the catalysis of chiral amine production

This work started with an enzyme with low level of S-selectivity (65%) in the conversion of the ketone  $\beta$ -tetralone to amino tetraline, the corresponding amine

Directed evolution is an efficient way to improve the enantioselectivity of enzymes, enhancing their performance in asymmetric catalysis. The evolved enzymes are used in the production of chiral substances with high enantiomer purity. An early example of directed evolution of the aim of improving the enantioselectivity of an enzyme was reported by Matcham and Bowen concerning transaminases in the catalysis of chiral amine production. These work started with an enzyme with low level of a S-selectivity 65% in the conversion of the ketone beta tetralone to amino tetraline the corresponding amine.

**(Refer Slide Time: 1:08:53)**

## Enantioselectivity of Enzymes:

A library of mutants was generated and screened for enhanced activity on the S-isomer but not the R-isomer

The result was a biocatalyst that produced the S-amino tetraline with greatly enhanced selectivity (94%), which was further improved by additional rounds of mutagenesis and screening

Manfred Reetz and co-workers reported another early example that has led to improved enantioselectivity of lipases in ester hydrolysis

Through directed evolution via four cycles of random mutagenesis, the selectivity factor of a bacterial lipase from *Pseudomonas aeruginosa* was first increased from 1.1 to 11 and then to 35 after further diversification of the library



The library of mutant was generated on screened for enhanced activity on the S- isomer but not on the R-isomer. The result was the biocatalyst that produced the S-Amino tetraline with greatly enhanced selectivity 94%, which was further improved by additional round of mutagenesis and screening. Manfred Reetz and co-worker reported another early example that has led to improved enantioselectivity of lipases in ester hydrolysis. Through directed evolution via four cycle of random mutagenesis, the selectivity factor of a bacterial lipase from *Pseudomonas aeruginosa* was first increased from 1.1 to 11, and then to 35 after further diversification of the library.

**(Refer Slide Time: 1:09:43)**

## Directed evolution in organic synthesis and industry:

Directed evolution quickly made its way from the academic setting to industrial applications

Enzymes developed using directed evolution are used in industry in the production of biofuels, materials, bulk and fine chemicals, detergents, consumer products, laboratory reagents and pharmaceuticals, as well as intermediates for the pharmaceutical industry

Several of the enzymes developed in the Arnold lab are used in industry. Many companies have their own scientific teams applying directed evolution strategies to improve catalysts or protein-based therapeutics in terms of stability, activity, specificity or other properties

Specific examples of evolved enzymes and products are taste enhancers, drugs against diabetes and vascular plaques, as well as lipid-lowering pharmaceuticals. Some enzymes produced by directed evolution are made on very large scale

This includes lipases used in detergents. Industrial chemicals are made in enormous quantities with the help of biocatalysts produced by directed evolution



Directed evolution quickly made its way from the academic setting to industrial applications. Enzymes developed using directed evolution are used in industry in the production of biofuels, materials, bulk and fine chemicals, detergents, consumer products, laboratory reagents and pharmaceuticals as well as intermediates of the pharmaceutical industry. Several of the enzymes developed in Arnold lab are used in industry. Many companies have their own scientific teams applying directed evolution strategies to improve catalysts or protein based therapeutics in terms of stability, activities, specificity and many other properties. Specific example of evolved enzymes and product are taste enhancers drugs against diabetes and vascular plaques as well as lipid lowering pharmaceuticals some enzyme produced by directed evolution are made on very large scale. This includes lipases used in detergents industrial chemicals are made in enormous quantities with the help of bio catalyst produced by directed evolution.

**(Refer Slide Time: 1:10:56)**

### Summary about Directed Evolution:

The use of directed evolution as a strategy for engineering novel enzymes and catabolic pathways has yet to reach its full potential

There are certain challenges for future directed evolution that include developing workable strategies for evolution of new catalytic functions evolving biosynthetic or degradative pathways, evolving single and new enzyme pathways for large scale chemical production and evolving enzymes that are difficult to handle in terms of functional expression, stability and assay development

There is a need to clearly understand what one means by a “good” library for evolution and we certainly need high throughput ~~screening~~ tools and new assay formats to address the challenging ~~problems~~ to which these tools are now being applied



The use of directed evolution as a strategy for engineering novel enzymes and catabolic pathways has yet to reach its full potential. There are certain challenges for future directed evolution that includes developing workable strategies for evolution of new catalytic functions evolving biosynthetic or degradative pathways evolving single and new enzyme pathways for large scale chemical production and evolving enzymes that are difficult to handle in terms of functional expression stability and assay developments. There is the need to clearly understand what one means by a good library for evolution, and we certainly need high throughput screening tools, and new assay format to address the challenging problems to which these tools are now being applied.

**(Refer Slide Time: 1:12:36)**

### Summary about Directed Evolution:

*in vitro* evolution has been reported to be a powerful and efficient tool for tuning enzyme performance in a wide range of applications

Directed evolution needs to be exploited to the maximum in certain areas such as designing novel allosteric biocatalysts, enzymes capable of chiral resolution, i.e. to direct the enzyme to produce the desired enantiomer

This technology needs to be used in active site protection of the chiral building blocks

Directed evolution can also be used in designing enzymes that are able to recognize certain recalcitrant entities and are able to modify them to decrease or eliminate their toxicity and hence can have a major role in decontaminating the environment



In vitro evolution has been reported to be a powerful and efficient tool for tuning enzyme performance in a wide range of applications. Directed evolution needs to be exploited to the maximum in certain areas such as designing novel allosteric biocatalysts enzymes capable of chiral resolutions that is to direct the enzyme to produce desired enantiomer. This technology needs to be used in active site production of the chiral building blocks. Directed evolution can also be used in designing enzymes that are able to recognize certain recalcitrant entities and are able to modify them to decrease or eliminate the toxicity and hence can have a major role in decontaminating the environment.

**(Refer Slide Time: 1:13:26)**

### Summary about Directed Evolution:

One of the most critical areas is the ability of the enzymes to show activity under non-aqueous or anhydrous conditions

Laboratory evolution can help us in obtaining these types of **“super enzymes”** that can catalyze certain industrially important reactions

As a reliable catalyst improvement technology, directed evolution is expected to play a major role in removing bottlenecks of biocatalysis process technology



43

One of the most critical areas is the ability of the enzymes to show activity under non accurate or anhydrous conditions. Laboratory evolution can help us in obtaining this type of super enzymes that can catalyze certain industrially important reactions. As a reliable catalyst improvement technology directed evolution is expected to play a major role in removing bottlenecks of bio catalysis process technology.

(Refer Slide Time: 1:13:55)

### Summary about Directed Evolution:

Directed evolution of binding proteins is an efficient way to identify variants with high affinity and selectivity for a given target

Also it could be used to map the sequence requirements for high-affinity and high selectivity protein interactions

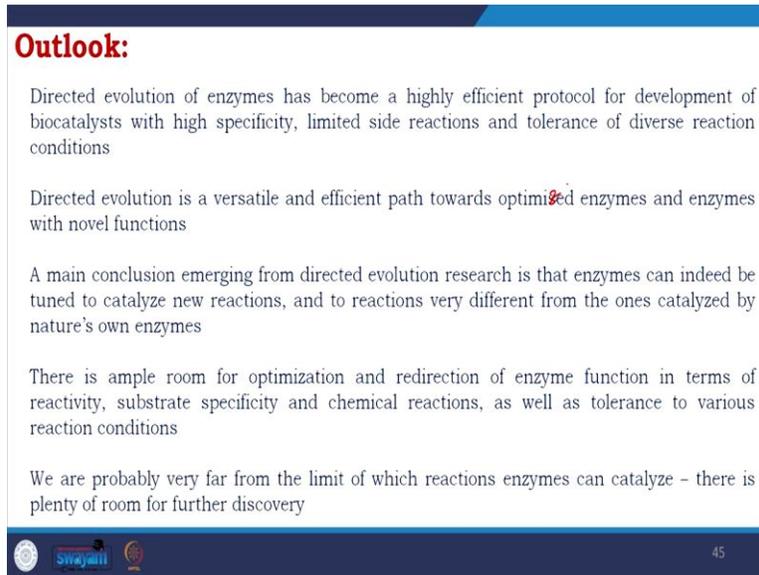
Directed evolution of human antibodies leads to useful therapeutics



44

Directed evolution of binding proteins is an efficient way to identify variants with high affinity and selectivity for a given target. Also it could be used to map the sequence requirements for high affinity and high selectivity protein interactions. Directed evolution of human antibodies leads to useful therapeutics.

(Refer Slide Time: 1:14:17)



**Outlook:**

Directed evolution of enzymes has become a highly efficient protocol for development of biocatalysts with high specificity, limited side reactions and tolerance of diverse reaction conditions

Directed evolution is a versatile and efficient path towards optimized enzymes and enzymes with novel functions

A main conclusion emerging from directed evolution research is that enzymes can indeed be tuned to catalyze new reactions, and to reactions very different from the ones catalyzed by nature's own enzymes

There is ample room for optimization and redirection of enzyme function in terms of reactivity, substrate specificity and chemical reactions, as well as tolerance to various reaction conditions

We are probably very far from the limit of which reactions enzymes can catalyze - there is plenty of room for further discovery

Directed evolution of enzymes has become a highly efficient protocol for development of bio catalyst with high specificity limited side reactions and tolerance for diverse reaction conditions. Directed evolution is a versatile and efficient path towards optimized enzymes and enzymes with novel functions. A main conclusion emerging from directed evolution research is that enzymes can indeed be tuned to catalyze new reaction and to reactions very different from the ones catalyzed by nature's own enzymes. There is ample room for optimization the direction of enzyme function in terms of reactivity substrate specificity and chemical reaction as well as tolerance to various reaction conditions. We are probably very far from the limit of which reaction enzyme can catalyze. There is plenty of room for further discovery and our future. It is a very bold statement but our future is depending on expanding that reparture. With this I will finish today's class on directed evolution. Thank you.