Site-Directed Mutagenesis and Protein Engineering

Dr. Sanatan Majhi

Assistant Professor
Department of Biotechnology
Utkal University, Vani Vihar, Bhubaneswar

Mutagenesis

- Mutation-random changes in the genetic material.
- Mutagen-an agent causing mutations.
- Directed mutagenesis-a process whereby specific changes are brought about in the code for a protein with the purpose of improving it for a specific application.

….cont

- It is possible to clone a gene, produce the protein and analyze the results in a variety of ways.
- It is however only possible to produce naturally occurring proteins.
- These proteins may not be well suited for an industrial task.

….cont

- The variety of enzyme properties available by searching for enzymes from organisms in extreme conditions.
- For example, α-amylase from *Bacillus stearothermophilus* or *Taq* polymerase from *Thermophilus aquaticus*.
- Still limited to naturally occurring proteins.

….cont

- Temperature or pH stability
- Reactivity in non-aqueous solvents
- Cofactor requirements
- Substrate binding or specificity
- Stability to proteases

Directed Mutagenesis Procedures

- It is not a simple matter to create a new protein.
- However, it is quite feasible to modify the existing properties of known proteins.
- It can be carried out at the protein level but chemical modifications of proteins are generally harsh, non-specific, and required repeatedly, for each batch of protein.
- It is preferable to manipulate the DNA sequence of a cloned gene to create an altered protein.
- Directed mutagenesis is a trial-and-error strategy.

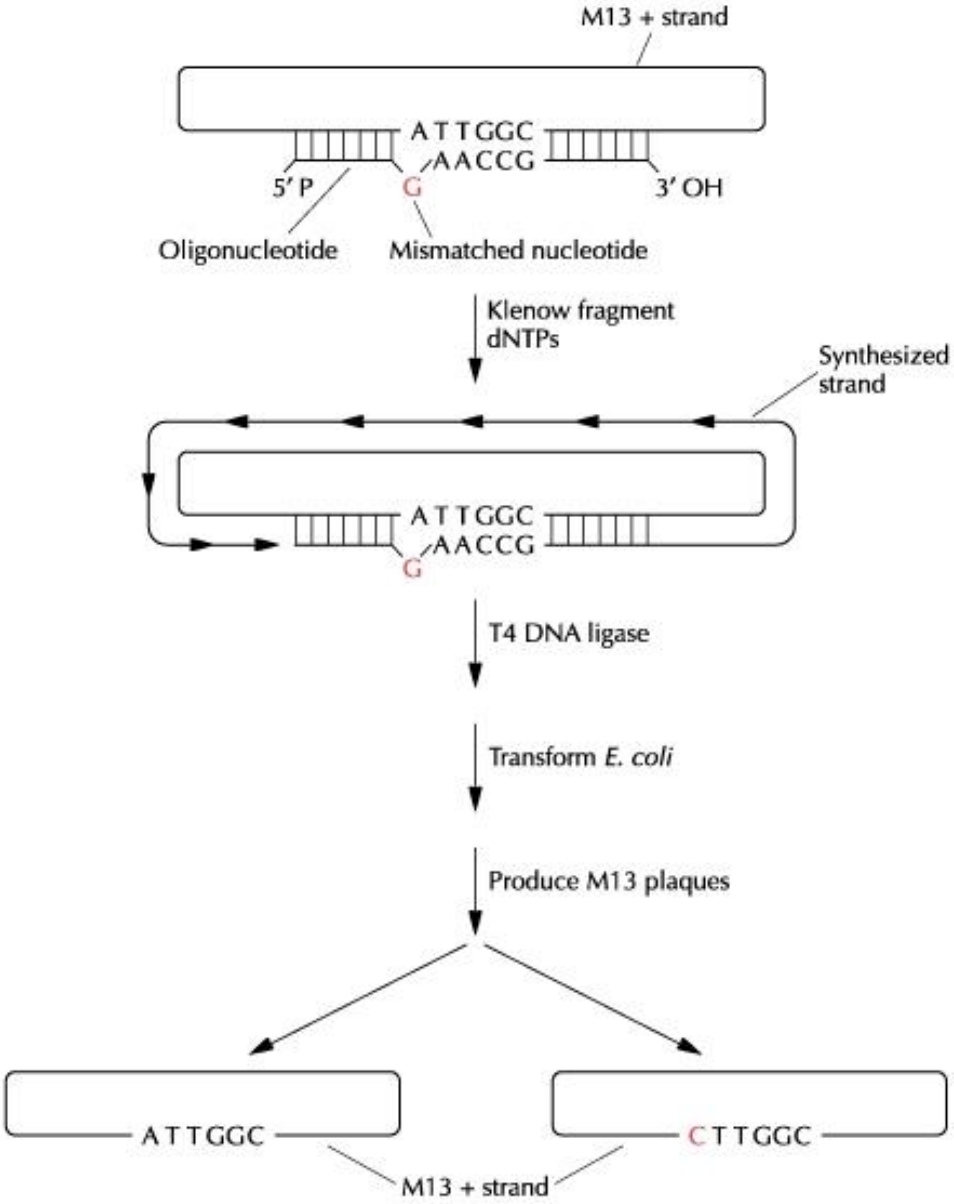
Directed Mutagenesis Procedures

Oligonucleotide-directed mutagenesis with M13 DNA

- Single-stranded bacteriophage M13 (+ strand), carrying a cloned gene, is annealed with a complementary synthetic oligonucleotide containing one mismatched base.
- DNA synthesis is catalyzed by Klenow fragment of *E. coli* DNA polymerase I.
- The newly synthesized DNA is circularized by T4 DNA ligase.
- The ligation mixture is used to transformed *E. coli*.

Oligonucleotide-directed mutagenesis with M13 DNA

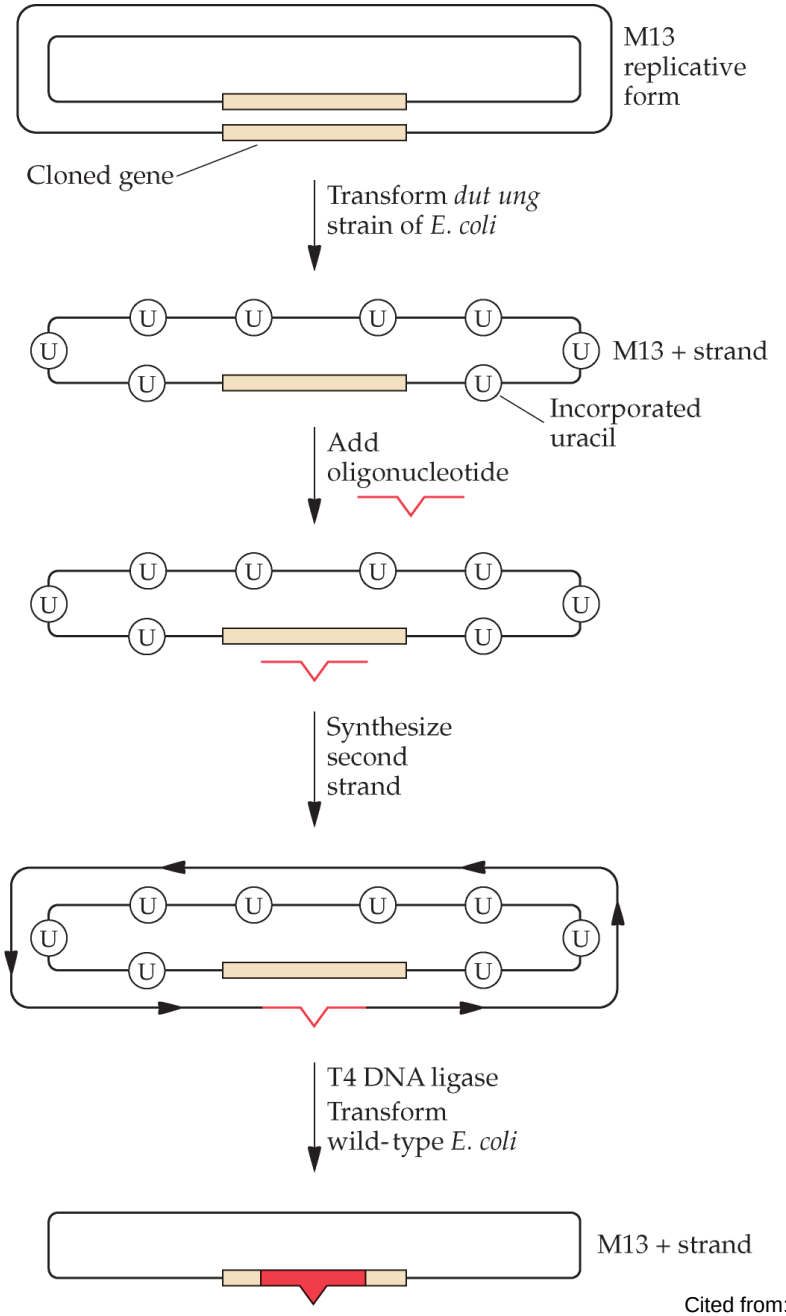
- Half of the produced phage will have the mutation (theoretically).
- The mutated clones are identified by DNA hybridization under highly stringent conditions.
- The original nucleotide containing mismatched sequence is used as a probe.
- Must know the sequence of the gene and the amino acid changes that are being introduced.
- Only 1% of the plaques actually contain phage carrying the mutated gene.



Cited from: Molecular Biology of the Cell by Bruce Albert

Oligonucleotide-directed mutagenesis using M13 DNA

- One approach is to introduce M13 viral vector carrying the gene that is to be mutagenized into an *E. coli* strain that has two defective enzymes.
- The *dut* mutation causes the intracellular level of dUTP to be elevated.
- The *ung* mutation prevents the removal of any incorporated uracil residues.
-
- Following the directed mutagenesis, the cloned is used to transform wild-type *E. coli*.
- In this way, the yield of M13 bacteriophage carrying a gene with a site-specific mutation is increased.



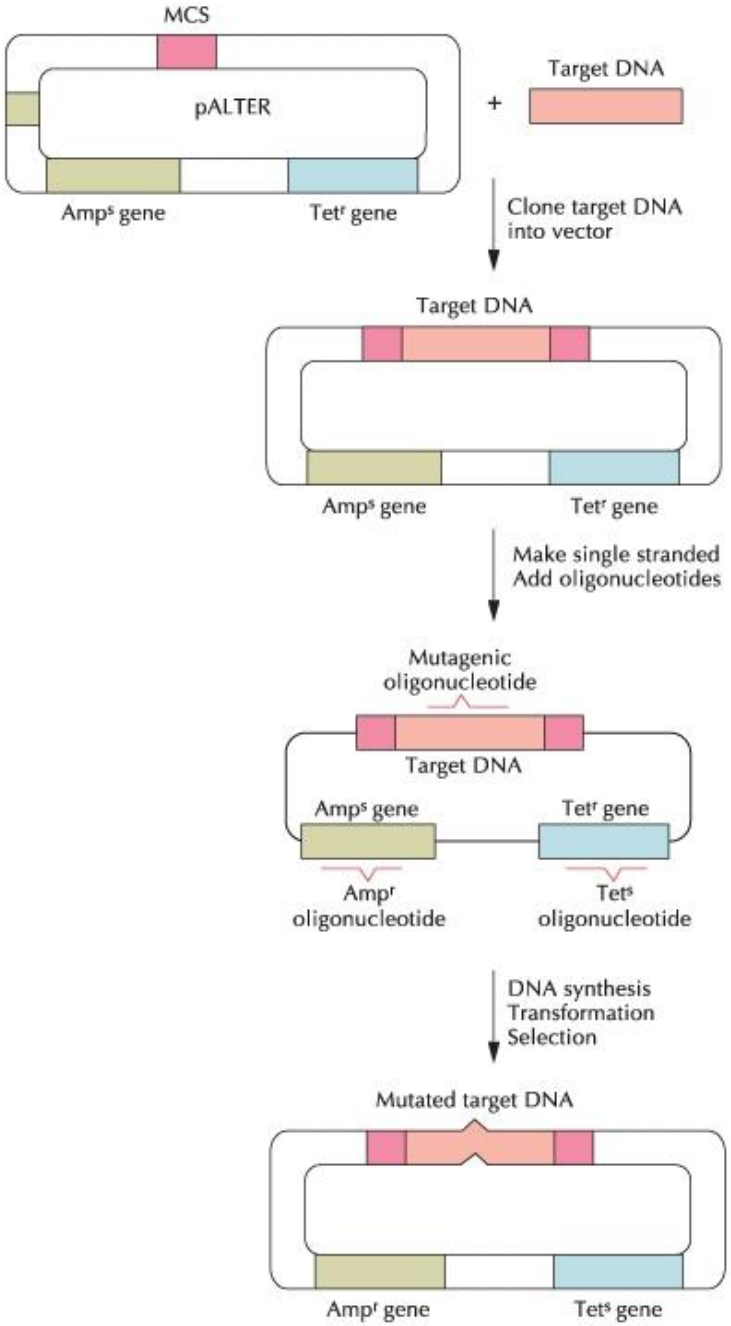
Cited from: Molecular Biology of the Cell by Bruce Albert

….cont

- One approach is to introduce M13 viral vector carrying the gene that is to be mutagenized into an *E. coli* strain that has two defective enzymes.
- The *dut* mutation causes the intracellular level of dUTP to be elevated.
- The *ung* mutation prevents the removal of any incorporated uracil residues.
- Following the directed mutagenesis, the cloned is used to transform wild-type *E. coli*.
- In this way, the yield of M13 bacteriophage carrying a gene with a site-specific mutation is increased.

Oligonucleotide-directed mutagenesis with Plasmid DNA

- The drawback to M13 mutagenesis protocols is time-consuming.
- Clone into plasmid with a non-functional amp^R gene and a functional tet^R gene.
- Mutagenize with three separate primers:
  - One primer mutates cloned gene
  - One primer inactivates amp^R
  - One primer reactivates tet^R
- Transformants are selected for ampicillin resistance and tetracyclin sensitivity.
- The cells with specified mutation in the target gene are identified by DNA hybridixation.



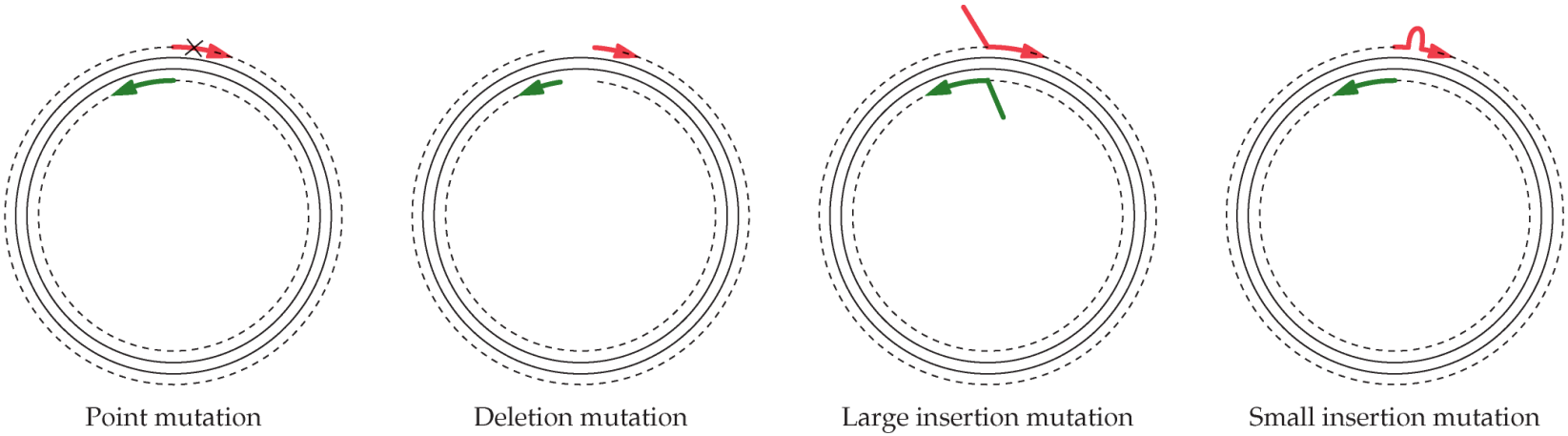
Cited from: Molecular Biology of the Cell by Bruce Albert

PCR-Amplified Oligonucleotide-Directed Mutagenesis

- PCR can be exploited both to introduce the desired mutation and to enrich for the mutated gene.
- This procedure uses forward and reverse PCR primers.
- Each pair of primers has one that is completely complimentary to a portion of our gene and one that has a single nucleotide change for point mutation.
- Following the PCR, a high percentage of the plasmids produced will have the desired mutation.
- No specific plasmids required, all we need to know is the sequence of the cloned gene.

- To create deletion mutations, primers must border the region of target DNA to be deleted on both sides.

- To create mutation with long insertions, a stretch of miss-matched nucleotides is added to 5’ end of one or both primers.

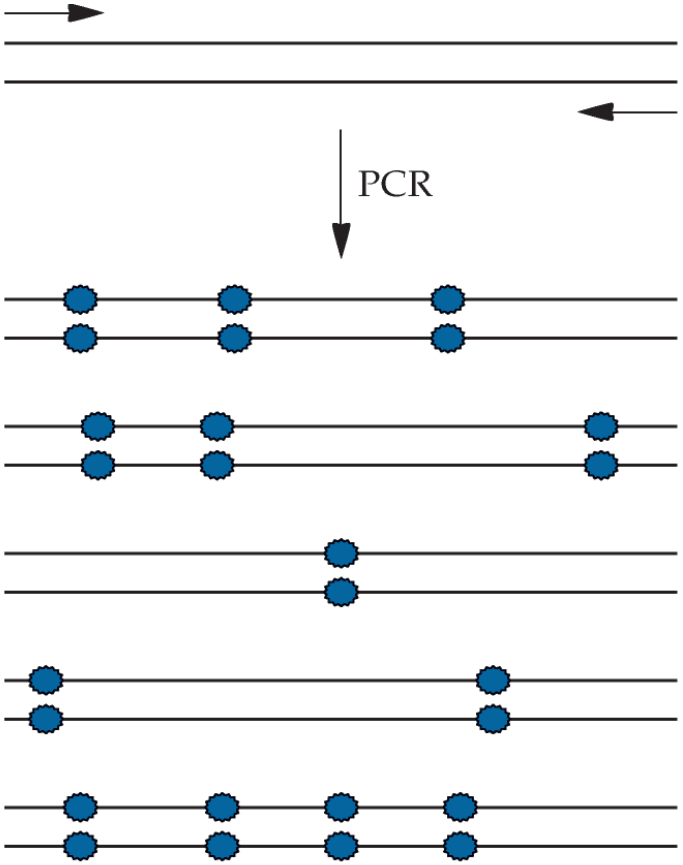


Cited from: Protein Engineering by Owe T Bornschuer, Mattihas Honhe

Error-Prone PCR

- It is a powerful method for random mutagenesis.
- *Taq* DNA polymerase lacks proof reading activity .
- Adding Mn^{2+}, increasing the concentration of Mg^{2+}, and adding unequal amount of 4 dNTPs will increase the error rate.
- Following the PCR, the randomly mutagenized DNA is cloned into expression vectors and screened for altered or improved protein activity.
- The desired cloned are then isolated and sequenced so the changes can be elaborated.

- Error – prone PCR of a target gene yields a variety of mutated forms of the gene.



Cited from: Protein Engineering by Owe T Bornschuer, Mattihas Honhe

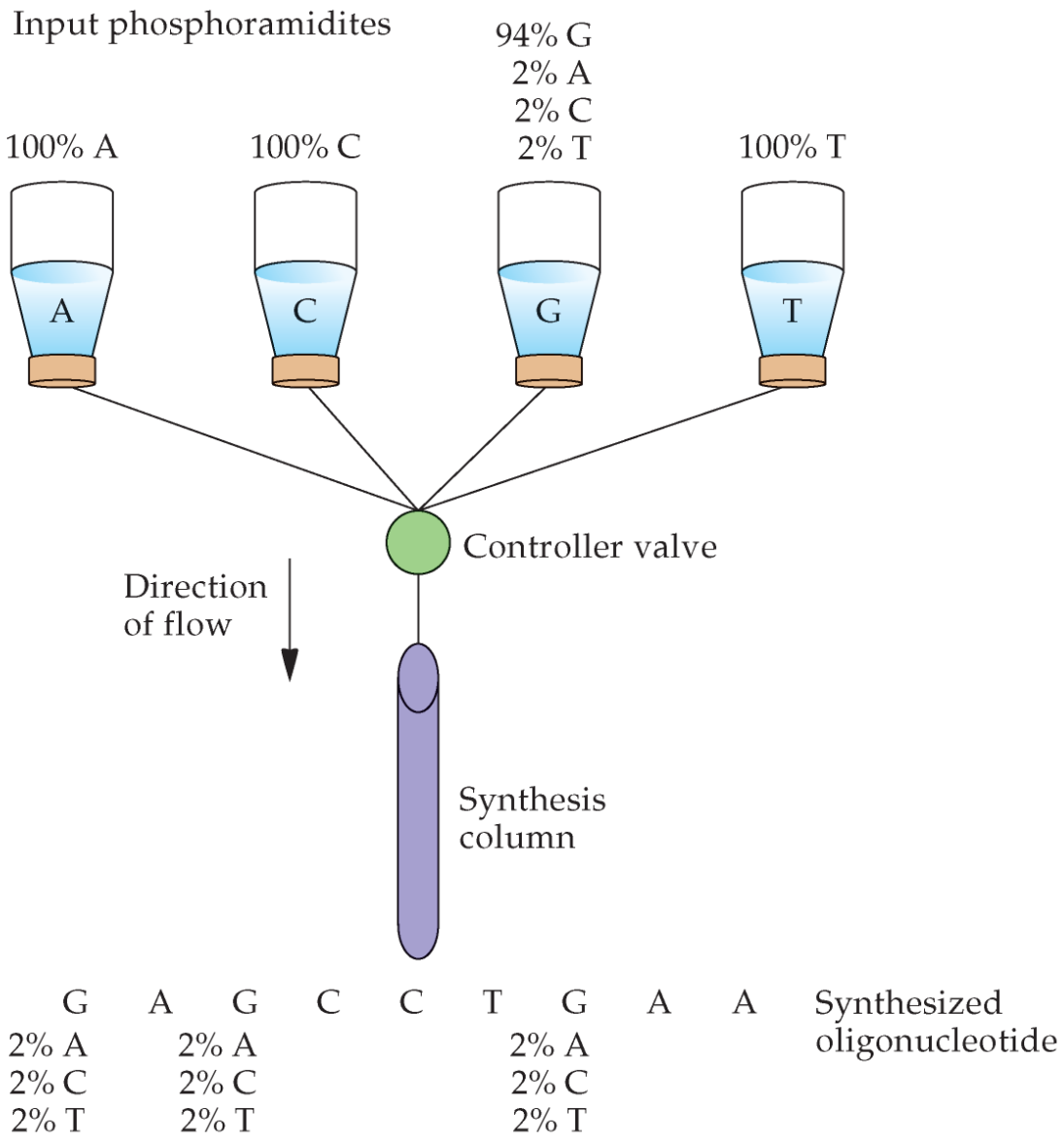
Random Mutagenesis with Degenerate Oligonucleotide Primers

- This approach is used if the protein in question is less well characterized.

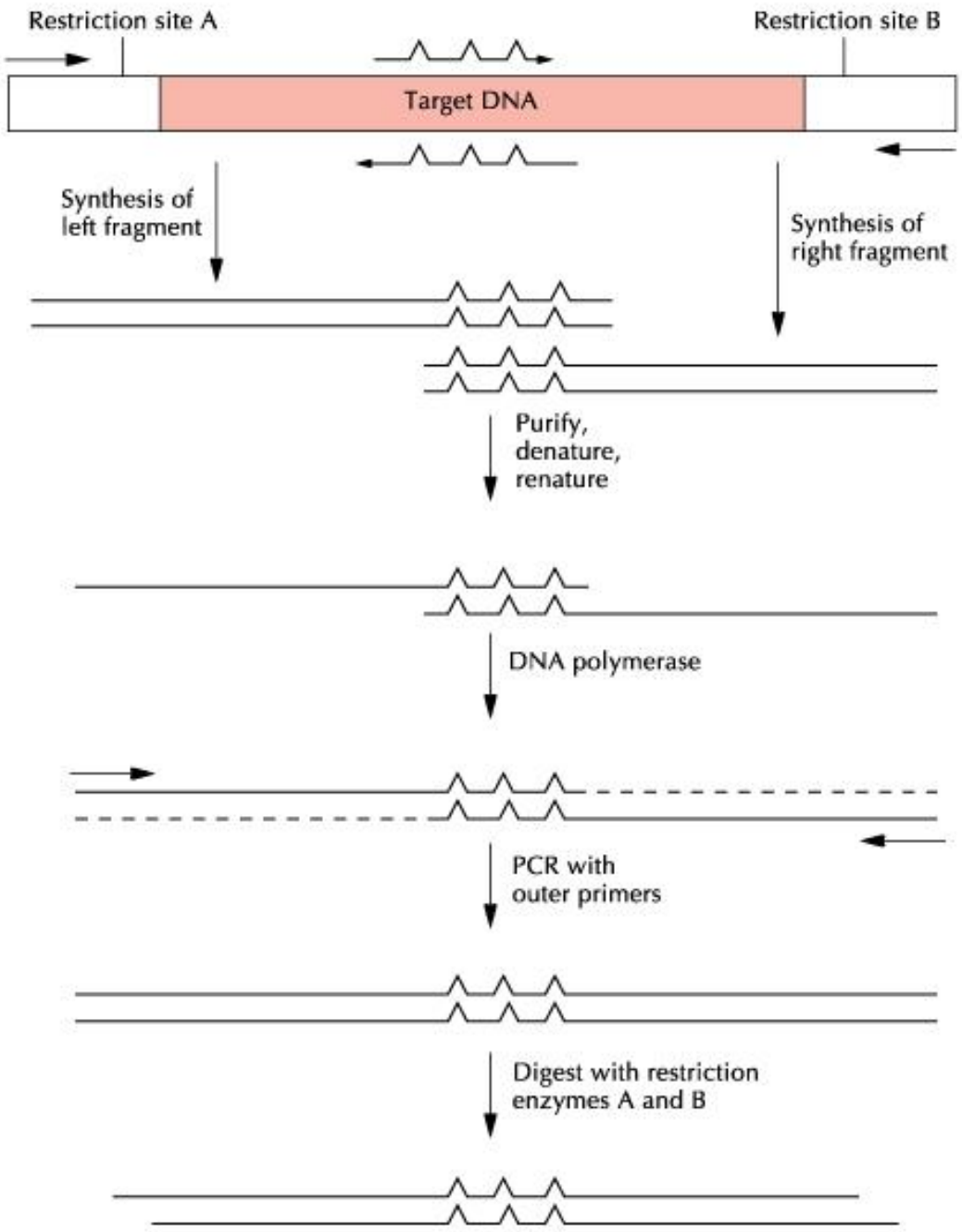
- This approach has two advantages:
  - The researcher does not have to know which amino acid is important or how to change it.
  - The range of mutants produced is broad with lots of interesting possibilities.

- If, however, none of the mutants is a desirable one, the entire process needs to be repeated.

- To start this procedure, degenerate primers are made. These are then used in PCR to generate random mutations in our gene of interest that has been cloned into a plasmid.
- The left and right portions of the target DNA are amplified separately by PCR.
- The amplified fragments are amplified, denatured to make them single stranded, and then re-annealed.
- Complementary regions of overlap are formed.
- The DNA polymerase is used to amplify the second strand and then the entire fragment is amplified.
- The product s digested and cloned into the vector.



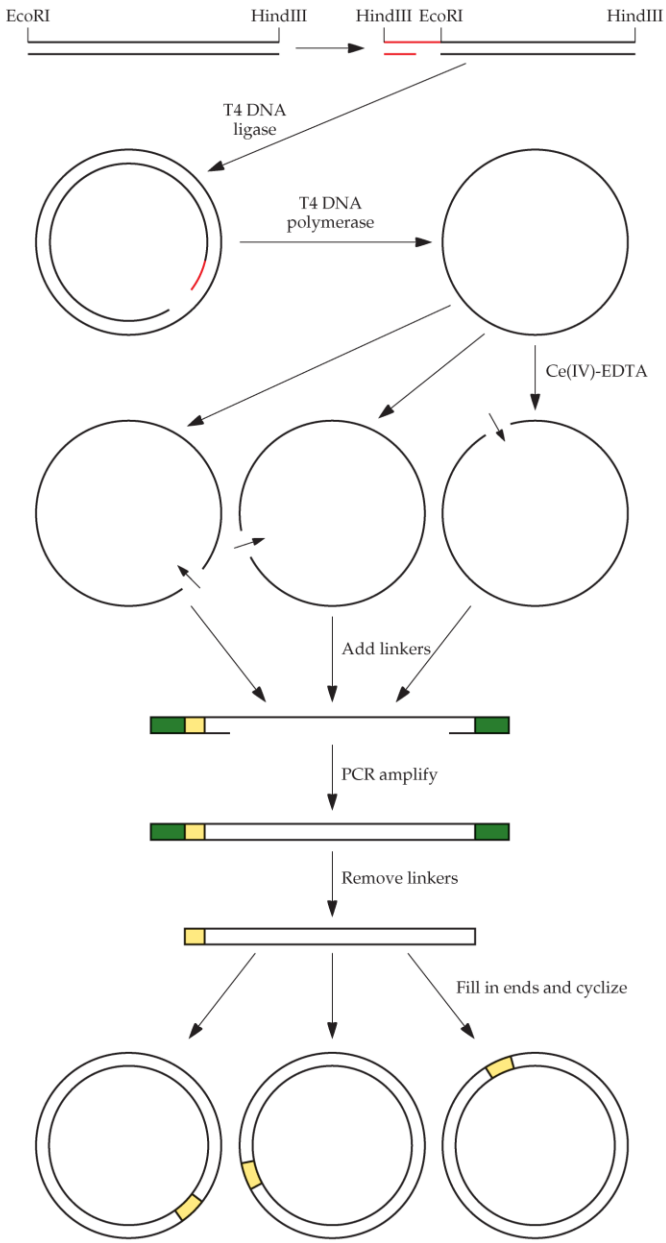
Cited from: Molecular biology, 2nd edition by David Freifelder



Cited from: Molecular Biology of the Cell by Bruce Albert

Random Insertion/Deletion Mutagenesis

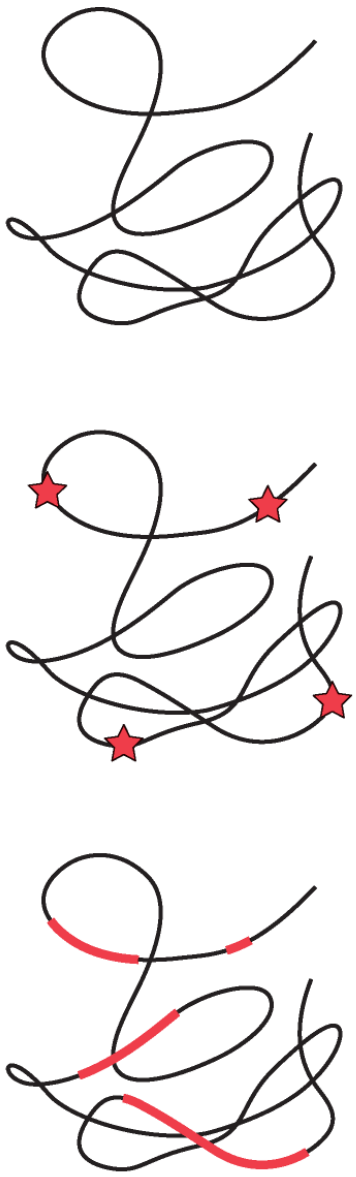
- This approach is an alternative to error-prone PCR.
- With this approach, it is possible to:
  - Delete a small number of nucleotides at random positions along the gene.
  - Insert either specific or random sequences into that position.



Cited from: Protein Engineering by Owe T Bornschuer, Mattihas Honhe

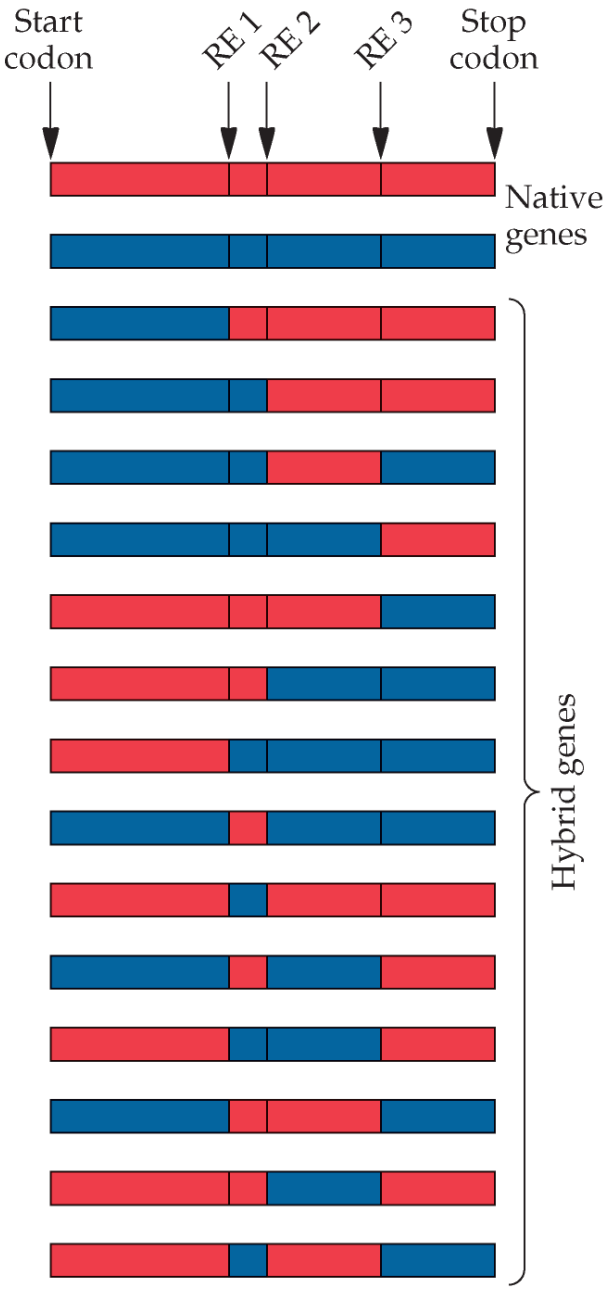
DNA Shuffling

- Some proteins are encoded by a group of related genes called a gene family.
- It is possible using common restriction site to recombine domains from different members of the family to look for proteins with unusual characteristics.
- Also, some of the hybrid proteins may combine important attributes of two or more of the original proteins, e.g., high activity and thermostability.



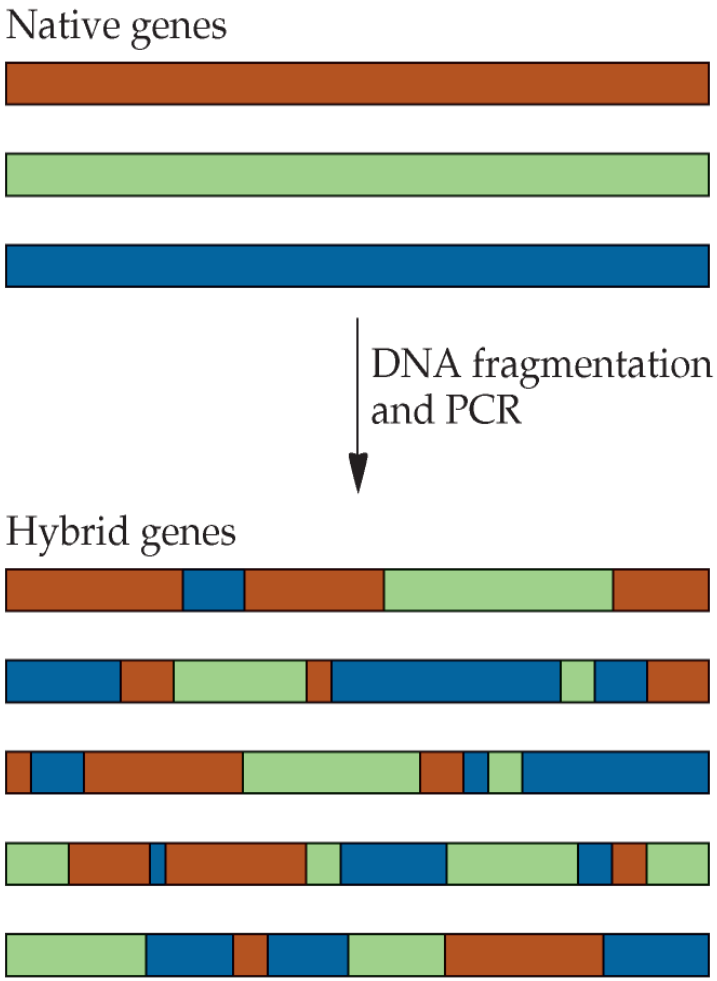
- Random mutagenesis or error-prone PCR causes single-amino-acid substitutions.
- DNA shuffling, in which genes are formed with large regions from different sources.

Cited from: Protein Engineering by Owe T Bornschuer, Mattihas Honhe



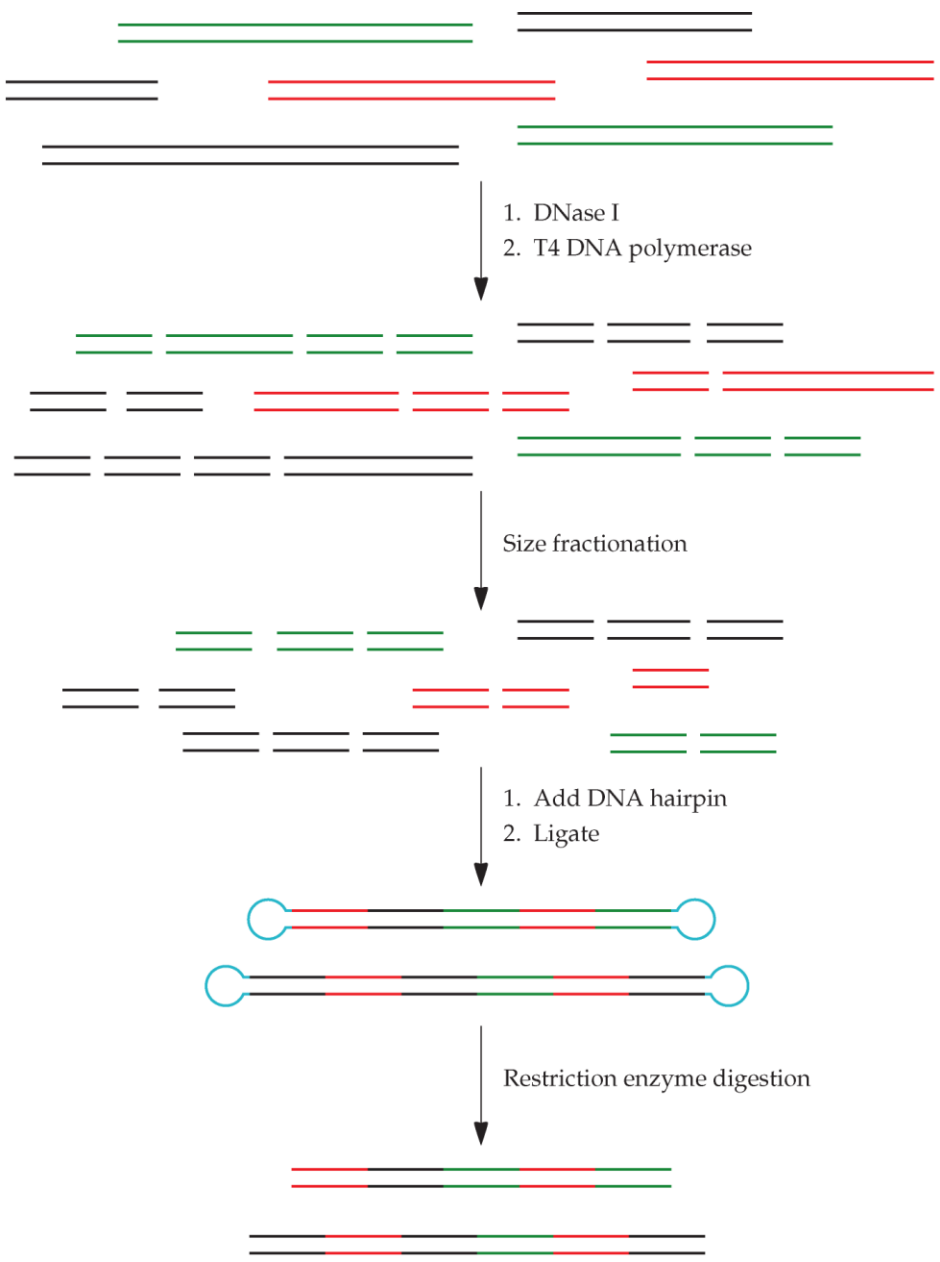
- Digestion of two or more of the DNAs that encode the native forms of similar proteins with one or more restriction enzymes that cut the DNA in the same place.
- Followed by ligation of the mixture of DNA fragments, a large number of hybrids are generated.

Cited from: Protein Engineering by Owe T Bornschuer, Mattihas Honhe



- Some of the hybrid DNAs that can be generated during PCR amplification of three members of a gene family.
- The mixed DNA is fragmented by DNase I, selecting smaller fragments, and PCR.
- Different gene fragments from a gene family will cross-prime each other.
- It is not useful with little or no homology.

Cited from: Protein Engineering by Owe T Bornschuer, Mattihas Honhe

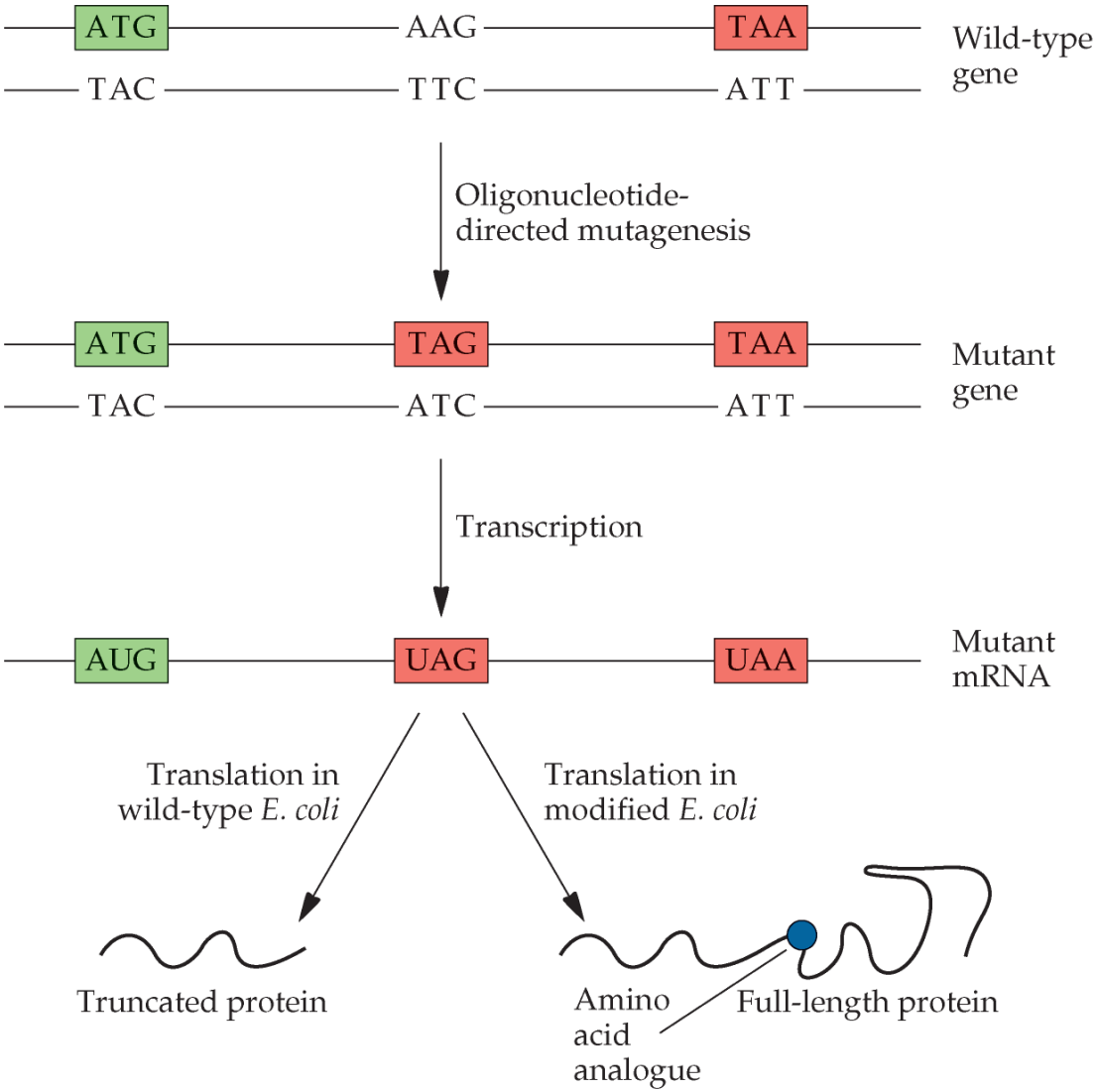


Cited from: Protein Engineering by Owe T Bornschuer, Mattihas Honhe

Mutant Proteins with Unusual Amino Acids

- Any protein can be altered by substituting one amino acid for another using directed mutagenesis.
- However, it is limited to the 20 amino acids.
- One way to increase the diversity of the proteins is to introduced synthetic amino acids with unique side chains at specific sites.
- It is possible using modified tRNAs and tyrosine-tRNA synthetase to introduce unusual amino acid into mutagenesis amber stop codon.
- An attempt to alter the final protein activity in ways that aren't normally possible.

Mutant Proteins with shifted Amino Acids



Cited from: Protein Engineering by Owe T Bornschuer, Mattihas Honhe

Protein Engineering

- Of the many thousand enzymes characterized, only about 20 or so are used greatly in industrial processes.
- The major barrier is finding enzymes with the desired activities and well suited for a highly specialized industrial application.
- These naturally occurring enzymes are easily denatured by high temperature and organic solvents used in many processes.

….cont

TABLE 8.1 Some industrial enzymes and their commercial uses

Enzyme	Industrial use(s)
α-Amylase	Beer making, alcohol production
Aminoacylase	Preparation of L-amino acids
Bromelain	Meat tenderizer, juice clarification
Catalase	Antioxidant in prepared foods
Cellulase	Alcohol and glucose production
Ficin	Meat tenderizer, juice clarification
Glucoamylase	Beer making, alcohol production
Glucose isomerase	Manufacture of high-fructose syrups
Glucose oxidase	Antioxidant in prepared foods
Invertase	Sucrose inversion
Lactase	Whey utilization, lactose hydrolysis
Lipase	Cheese making, preparation of flavorings
Papain	Meet tenderizer, juice clarification
Pectinase	Clarifying fruit juices, alcohol production
Protease	Detergent, alcohol production
Rennet	Cheese making

Disulfide bonds addition

- There is generally a direct correlation between the number of disulfide bonds and the thermo-stability of a protein.
- To increase stability of a variety of enzymes disulfide bonds have been added to their structure.
- The problem is whether extra disulfide bonds perturb the normal function.

Native protein

Engineered protein

Cited from: Protein Engineering by Owe T Bornschuer, Mattihas Honhe

….cont

- T4 lysozyme- 6 variants were created using oligonucleotide-directed mutagenesis of amino acid whose R-groups were in close proximity to incorporate more cysteine residues.
- The variants were tested and were generally found to be more thermostable.
- Some variants however lacked activity.
- It is a trial-and-error process.
- However, it is clear that increasing disulfide bonds to enhance protein stability is feasible.

….cont

TABLE 8.2 Properties of T4 lysozyme and six engineered variants

Enzyme	Amino acid at position:							No. of -S-S-	% Activity	T_m (°C)
	3	9	21	54	97	142	164			
wt	Ile	Ile	Thr	Cys	Cys	Thr	Leu	0	100	41.9
pwt	Ile	Ile	Thr	Thr	Ala	Thr	Leu	0	100	41.9
A	Cys	Ile	Thr	Thr	Cys	Thr	Leu	1	96	46.7
B	Ile	Cys	Thr	Thr	Ala	Thr	Cys	1	106	48.3
C	Ile	Ile	Cys	Thr	Aka	Cys	Leu	1	0	52.9
D	Cys	Cys	Thr	Thr	Cys	Thr	Cys	2	95	57.6
E	Ile	Cys	Cys	Thr	Ala	Cys	Cys	2	0	58.9
F	Cys	Cys	Cys	Thr	Cys	Cys	Cys	3	0	65.5

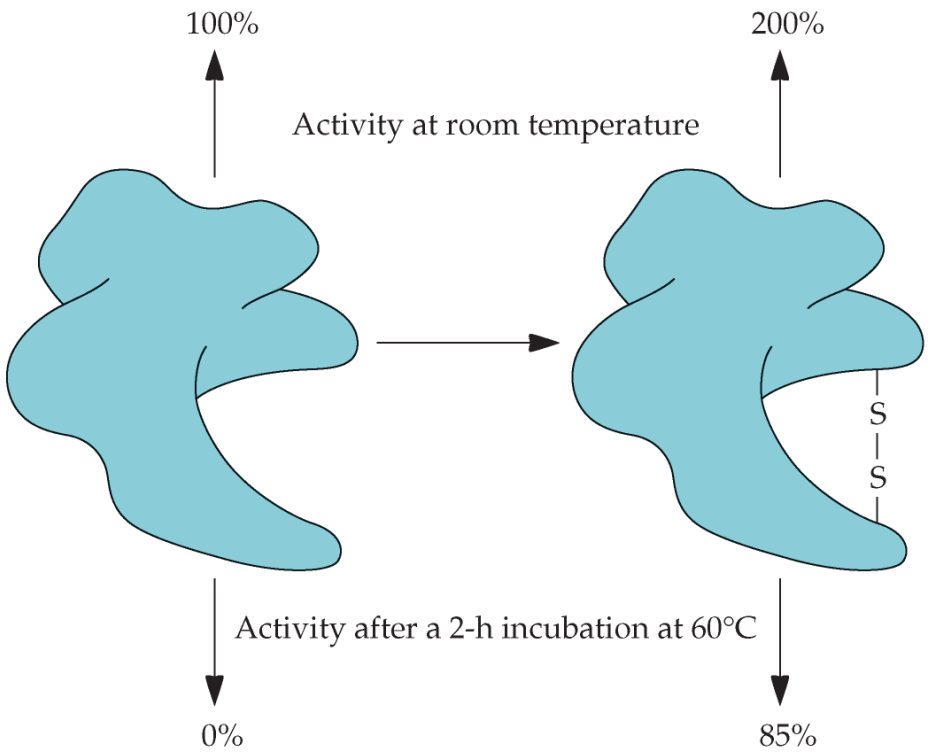
Adapted from Matsumura et al., *Nature* **342:**291–293, 1989.

wt, wild-type T4 lysozyme; pwt, pseudo-wild-type enzyme; A through F, six engineered cysteine variants; -S-S-, disulfide bonds; T_m, "melting" temperature (a measure of thermostability).

….cont

- Similar experiments were used to produce a more thermostable xylanase.
- An enzyme used in the degradation of hemi-cellulose, the compound generally darkens paper.
- The use of the variant xylanase that function efficiently at high industrial temperature allows the use of less bleach in paper production.
- Less pollution

- Addition of a disulfide bond stabilizes the protein.
- Its activity is doubled at room temperature, and protected against heat inactivation.

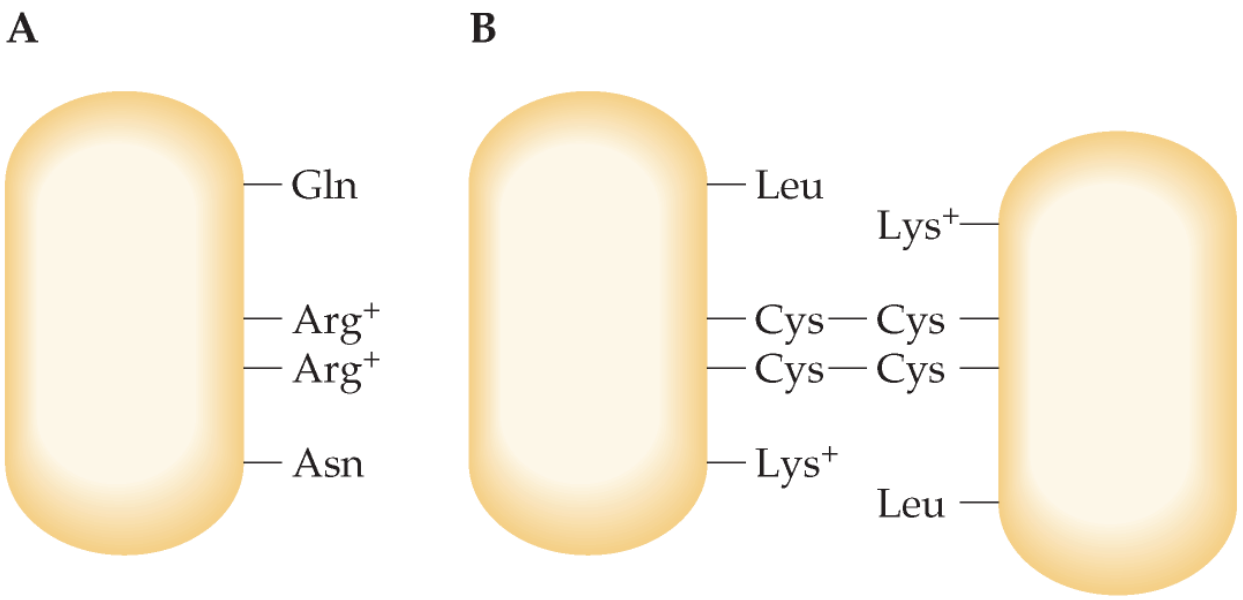


Cited from: Enzymes Engineering, Method and Protocol by James C. Samulelson

….cont

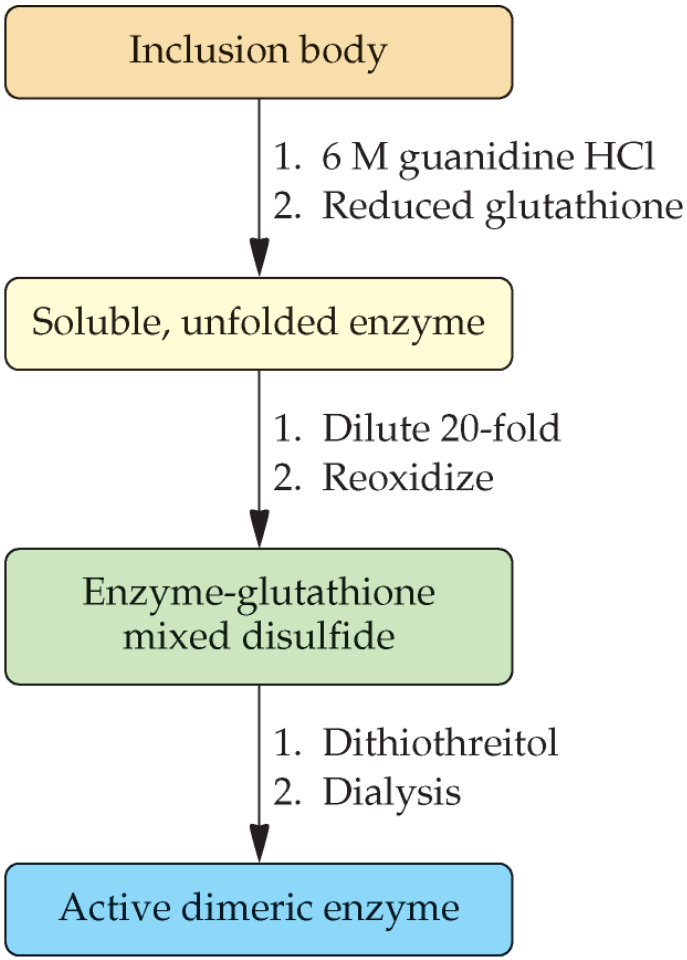
- RNase from bull semen has been used as an antitumorigenic agent.
- It is taken up into tumor cells and decreases protein synthesis by degrading rRNA, thereby blocking protein synthesis and causing cell death.
- The antitumor activity is dependent on the dimeric structure.
- Human antibody against the bull semen RNase are often produced.
- To eliminate this, human pancreatic RNase has been engineered to increase its functional life and activity using disulfide bonds.

….cont



Cited from: Protein Engineering by Owe T Bornschuer, Mattihas Honhe

….cont



Cited from: Enzymes Engineering, Method and Protocol by James C. Samulelson

Changing Asparagine to another Amino Acids

- Asparagine and glutamine residues can undergo deamination at high temperature becoming aspartic acid and glutamic acid, respectively.
- The change could lead to improper folding and loss of enzyme activity.
- A yeast enzyme served as the model and showed increased thermostability when Asn was change to threonine and isoleucine.
- Similarly, a modified long-lasting insulin was produced by changing Asp to Gly (approved for therapeutic use).

….cont

TABLE 8.3 Stability at 100°C of the yeast enzyme triosephosphate isomerase and its engineered derivatives

Enzyme	Amino acid at position:		Half-life (min)
	14	78	
Wild type	Asn	Asn	13
Variant A	Asn	Thr	17
Variant B	Asn	Ile	16
Variant C	Thr	Ile	25
Variant D	Asp	Asn	11

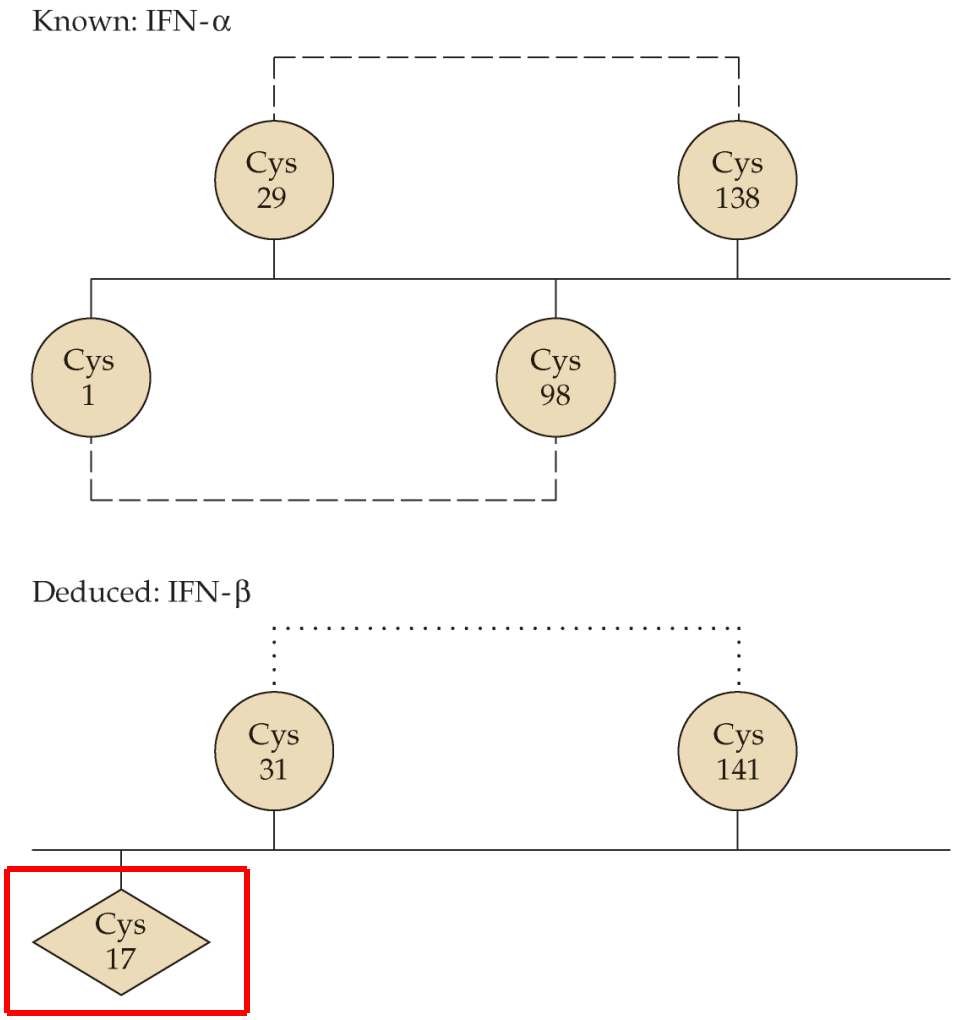
Adapted from Ahern et al., *Proc. Natl. Acad. Sci. USA* **84**:657–679, 1987.

Enzyme stability is expressed as the half-life, or rate of enzyme inactivation, at 100°C. A longer half-life indicates a more stable enzyme.

Reducing the Number of Free Sulfhydryl Residues

- In many early expression studies, the expressed protein showed much less activity than would have been predicted.
- Interferon showed only 10% of the antiviral activity of authentic glycosylated form.
- In *E. coli* production, the unglycosylated protein exist as dimers and higher oligomers that were inactive.
- Activity was greatly increased by changing different cysteines to serine.

….cont



Cited from: Molecular Biology of the Cell by Bruce Albert

Increasing Enzymatic Activity

- Experiments were performed using an Tyrosyl-tRNA synthetase for which the active site was well characterized.
- A threonine at position 51 of the protein was targeted for modification.
- If it was replaced with an alanine, the Km of the enzyme increased 2X without altering the rate.
- If it was replaced with proline, the enzyme bound ATP 100X more tightly than normal and increased its efficiency.

….cont

TABLE 8.4 Aminoacylation activity of native (Thr-51) and modified (Ala-51 and pro-51) tyrosyl-tRNA synthetases

Enzyme	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} M^{-1})
Thr-51	4.7	2.5	1,860
Ala-51	4.0	1.2	3,200
Pro-51	1.8	0.019	95,800

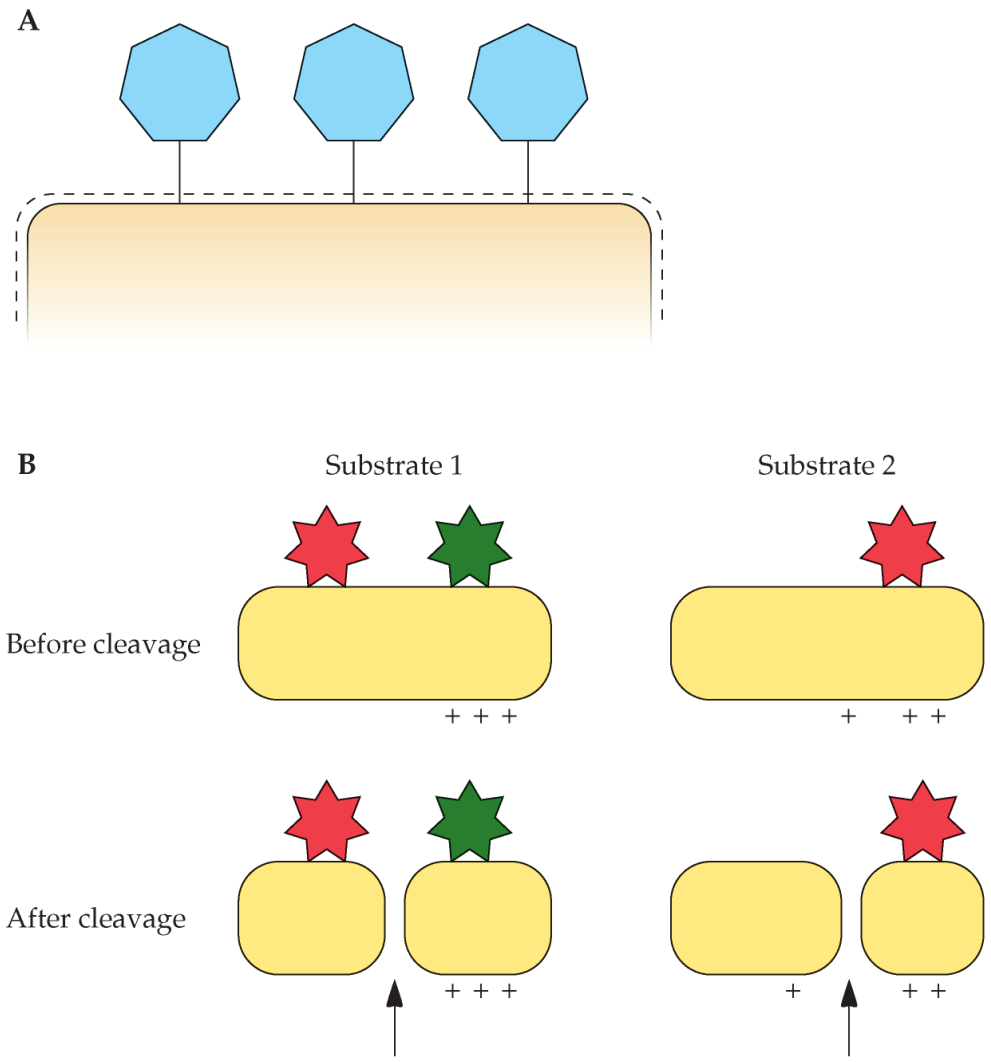
Adapted from Wilkenson et al., *Nature* **307:**187–188, 1984.

The units for K_m, the binding constant of the enzyme for ATP, are millimolar units (mM); the units for k_{cat}, the catalytic rate constant, are reciprocal seconds (s^{-1}); and the units for k_{cat}/K_m, the catalytic efficiency, are s^{-1} M^{-1}.

….cont

- It is necessary to simultaneously select for the new modified activity against the original enzyme activity.
- The endoprotease gene was first subjected to error-prone PCR, and modified genes were then cloned into E. coli displaying endoprotease on its surface.
- Two different substrates containing fluorescent dyes with Ala-Arg and Arg-Arg cleavage sites are added.
- After cell sorting, cells that showed both increased in green fluorescent and a decrease in red fluorescent were isolated and tested further .

Increasing Enzymatic Activity



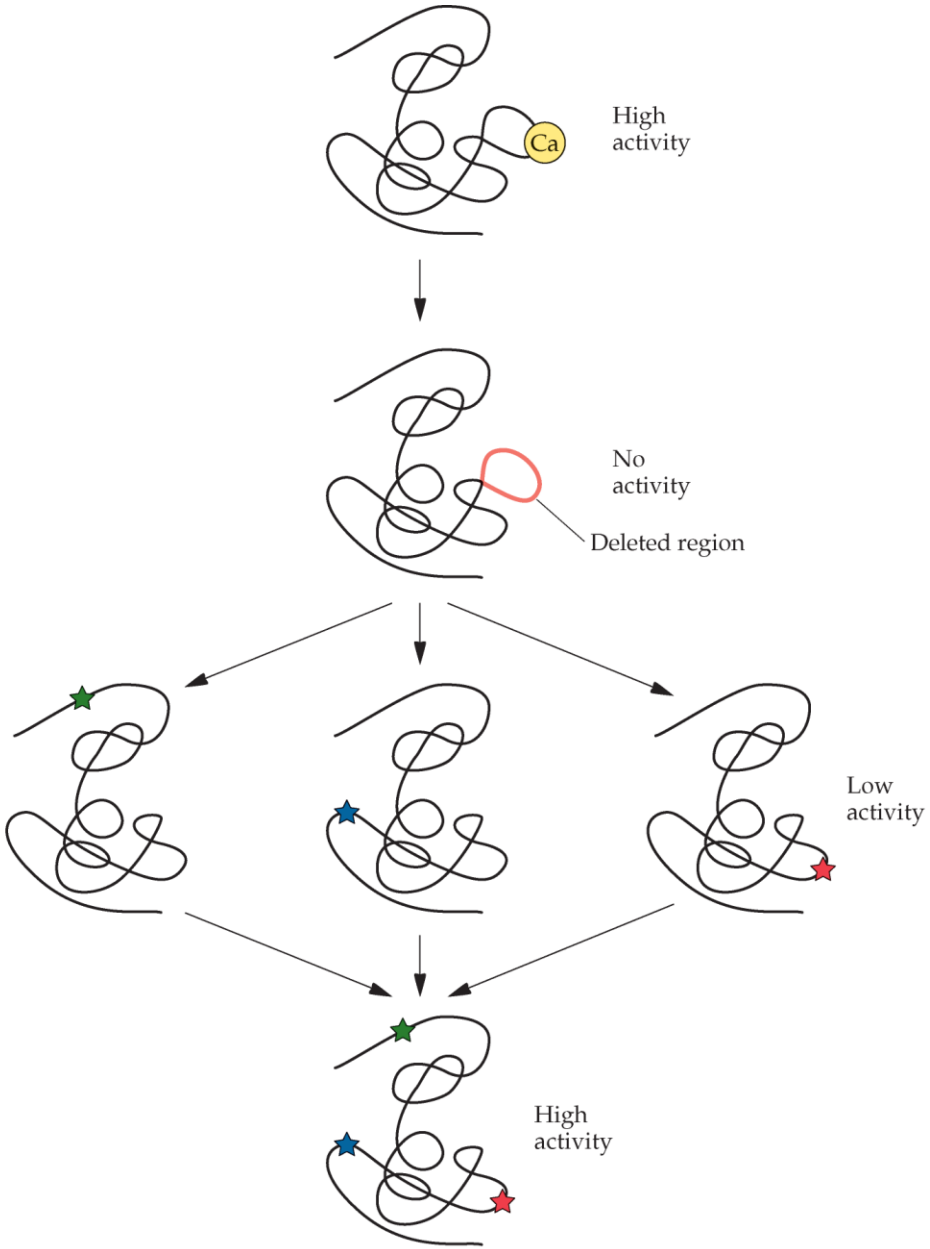
Cited from: Protein Engineering by Owe T Bornschuer, Mattihas Honhe

Modifying Metal Cofactor Requirements

- Subtilisin was studied. This enzyme from gram-positive bacteria and are widely used as biodegradable cleaning agents in laundry detergent.
- It requires Ca^{+2} as a cofactor to stabilize the enzyme.
- The enzyme is used in many industrial settings where there are a large number of metal-chelating agents that can bind to and effectively remove calcium.
- These enzymes are rapidly inactivated under these conditions.

….cont

- To avoid this, the portion of the enzyme required for Ca^{+2} binding was determined and eliminated from the protein (= inactive protein).
- The researchers began changing amino acids to restore enzyme activity.
- The result was a modified subtilisin that retains its activity but no longer requires Ca^{+2} ions for improved industrial use.



Modifying Metal Cofactor Requirements

Cited from: Protein Engineering by Owe T Bornschuer, Mattihas Honhe

Modifying Metal Cofactor Requirements

TABLE 8.5 Effects of random mutations of selected amino acid residues on the stability of subtilisin BPN lacking a calcium-binding domain

Region of protein	Amino acid residue	Stabilizing mutation	Fold increase in half-life
N terminus	2	Gln→Lys	2.0
	3	Ser→Cys	17.0
	4	None found	None
	5	Pro→Ser	1.2
Omega loop	41	Asp→Ala	1.2
	44	Lys→Asn	1.2
α-Helix	73	Ala→Leu	2.6
	74	None found	None
β-Pleat structure	206	Gln→Cys	17.0
	214	None found	None

Adapted from Strausberg et al., *Bio/Technology* **13:** 669–673, 1995.

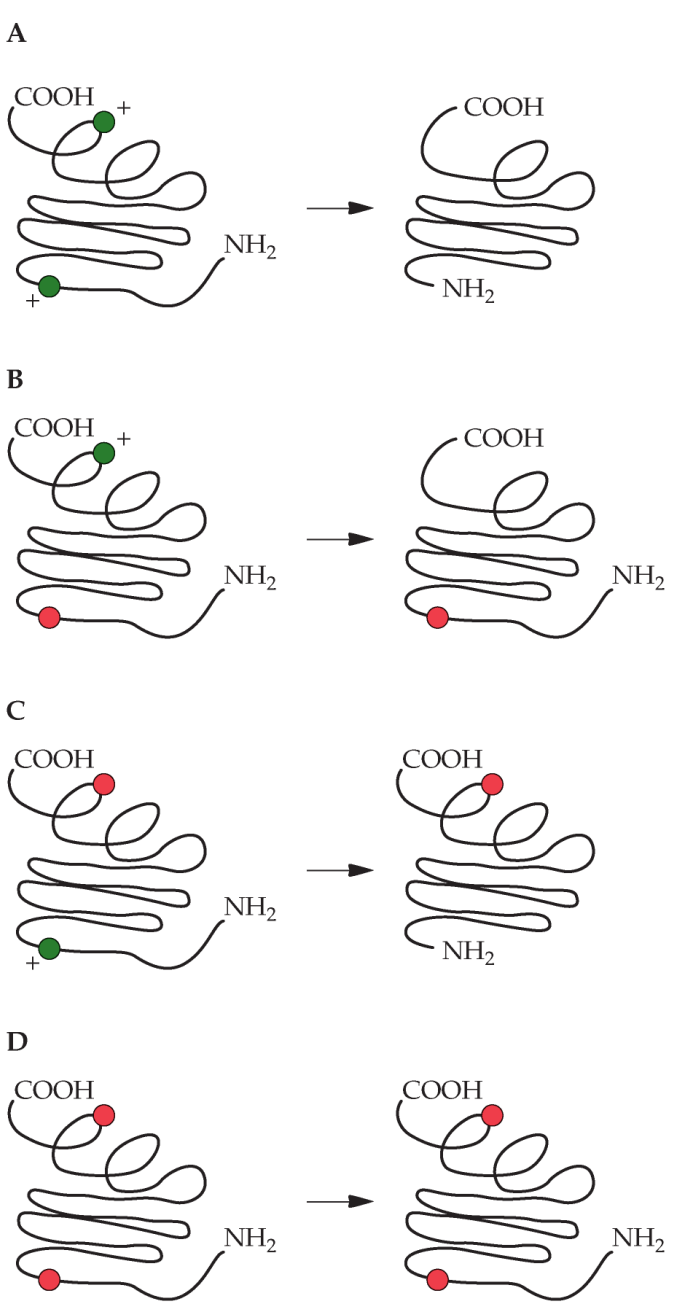
The mutations at positions 3 and 206 to Cys occur in the same clone and provide such a high level of stability because of the formation of the disulfide bridge between these residues.

Decreasing Protease Sensitivity

- Demonstrated with streptokinase; a protein produced by pathogenic strain of *Streptococcus* bacteria that is a blood clot dissolving agent.
- The protein complexes with plasminogen converting it to plasmin; the protease that degrades the fibrin in a blood clot.
- Unfortunately, plasmin also degrades streptokinase decrease its useful life in medical treatment.
- A long-lived streptokinase could be administered as a single injection before a person with heart attacked is transported to a hospital.

….cont

- Plasmin is a trypsin-like protease cleaving the peptide bond adjacent to a lysine or arginine residue.
- Plasmin cleaves streptokinase twice after 2 separate lysine residues, lysine 59 and lysine 386.
- To make streptokinase less susceptible to plasmin, the two lysine residues were changed to glutamine by oligonucleotide-directed mutagenesis.
- Glutamine does not have positive charge and the length of its side chain is similar to lysine thus not changing the three dimensional structure.
-
- The double mutant showed 21X more protease resistance.



Processed size (kDa)	Activity
37	Low
42	Low
40	Low
47	High

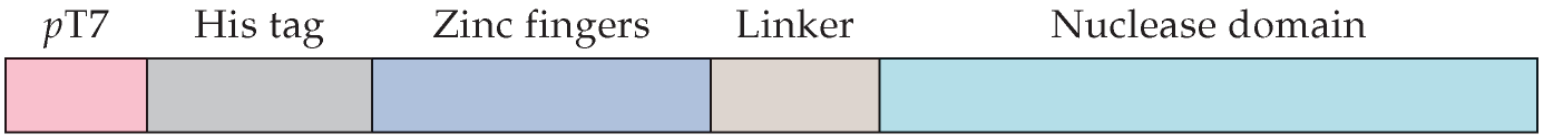
Cited from: Enzymes Engineering, Method and Protocol by James C. Samulelson

Modifying Protein Specificity

- *Fok*I restriction enzyme from *Flavobacterium okeanokoites* was used as a model. It is a relatively non-specific nuclease.

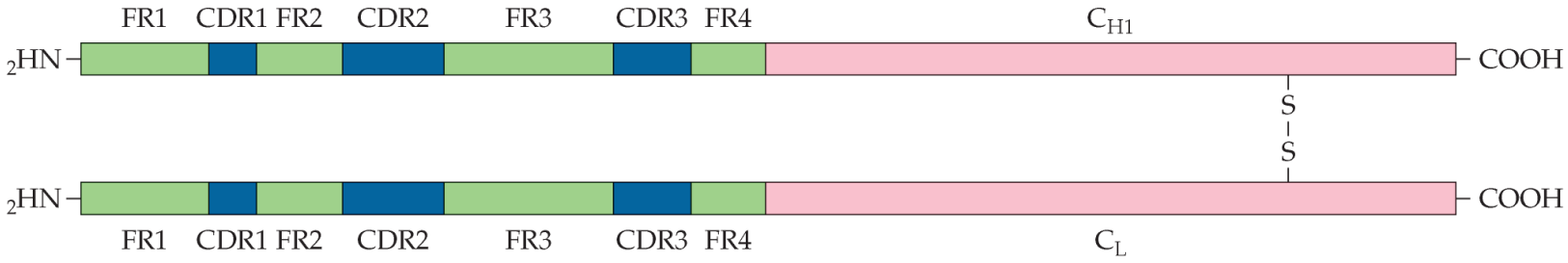
- To add specificity to the enzyme, the portion of the gene encoding the catalytic domain was fused to another segment of DNA encoding histidine, three zinc finger domains, a linker peptide to confer flexibility.

- The modified enzyme was shown to recognize the expected site and to a lesser degree another two sites.



Cited from: Molecular Biology of the Cell by Bruce Albert

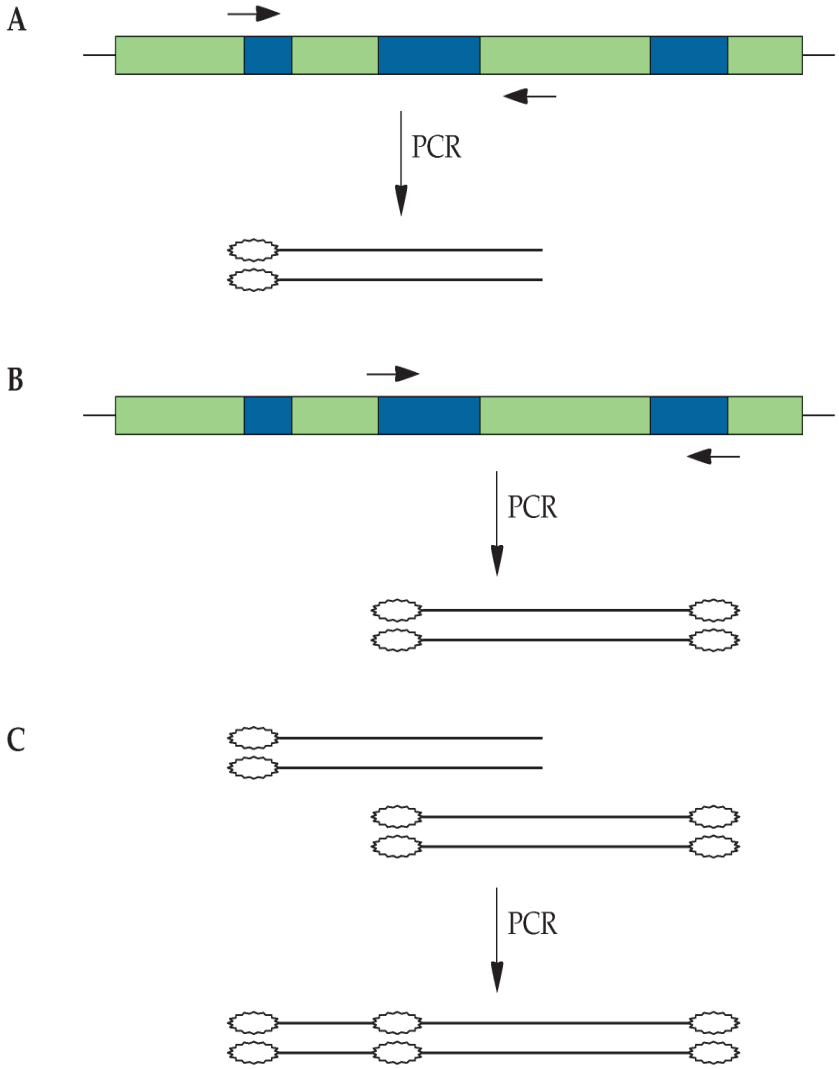
- Antibodies have also been used to demonstrate this.
- By modifying the hypervariable regions of the protein, it should be possible to generate antibodies that are directed against a wide range of antigenic determinants.
- Done by introducing mutations in the hypervariable complimentarity-determining regions (CDR). Together, the six of these CDRs determine the specificity of an antibody.



Cited from: Molecular Biology of the Cell by Bruce Albert

- Three CDR on the heavy chain were modified using random mutagenesis with mixed degenerate primers.
- The first PCR modified the CDR 1.
- The second PCR modified CDR 2 and 3.
- The third PCR combined the three altered CDR into a single fragment.
- In one instance, a Fab fragment of a monoclonal antibody that was specific for the compound 11-deoxycortisol was altered to be specific for cortisol.
- This approach can be used to create Fab fragments directed toward any antigenic determinant depending on the method used in screening the library.

Modifying Protein Specificity



Cited from: Molecular Biology of the Cell by Bruce Albert

Increasing Enzyme Stability and Specificity

- Tissue plasminogen activator is a serine protease useful in breaking down blood clots. Like streptokinase, it is eliminated from the body quickly by degradation.
- If high concentrations are given, non-specific internal bleeding can result.
- Thus a long-lived tPA with increased fibrin specificity that doesn't cause non-specific bleeding is desirable.

- Three separate directed mutations were introduced to achieve these three properties.

- Moreover, combining these three mutations in a single construct allows all three activity to be expressed simultaneously.

TABLE 8.6 Stabilities and activities of various modified versions of tPA

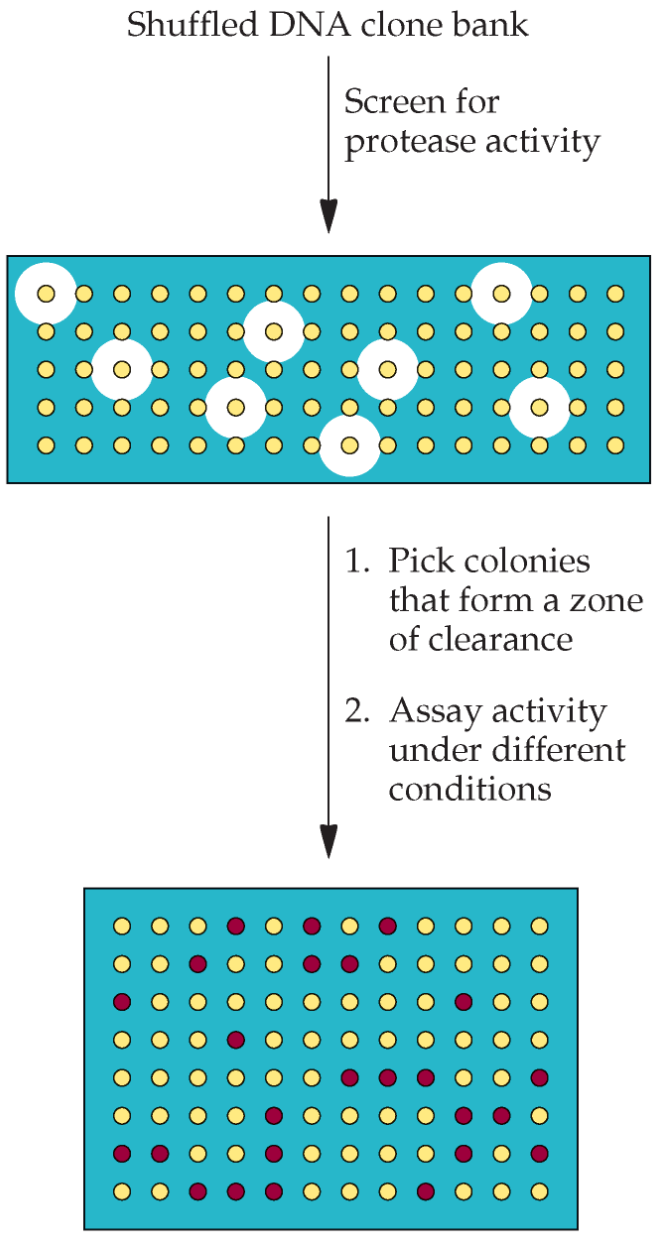
tPA variant	Modification(s)	Stability in plasma	Fibrin binding	Activity in plasma	Activity vs. clots
1	Thr(103)→Asn	10	0.34	0.68	0.56
2	LysHisArgArg(296–299)→AlaAlaAlaAla	0.85	0.93	0.13	1.01
3	Thr(103)→Asn,KysHisargArg(296–299)→AlaAlaAlaAla	5.3	0.33	0.13	0.65
4	Thr(103)→Asn, Ans(117)→Gln	3.4	1.0	1.13	1.17
5	LysHisArgArg(296–299)→AlaAlaAlaAla, Ans(17) Gln	1.2	1.33	0.16	1.38
6	Thr(103)→Asn, LysHisArgArg(296–299)→AlaAlaAlaAla, Asn(117)→Gln	8.3	0.87	0.06	0.85

Adapted from Keyt et al., *Proc. Natl. Acad. Sci. USA* **91:**3670–3674, 1994.

All of the values shown are normalized to the wild type. Plasma stability is the reciprocal of the time it takes for plasma clearance, larger numbers indicate a more stable derivative. Fibrin specificity is reflected by a high activity versus clots and a lows activity in plasma.

Altering Multiple Properties

- Often changing one property often disrupts other important characteristics.
- Using subtilisin, DNA shuffling was used to created a large number of chimeric enzymes with unknown properties.
- These were tested for activity at 23˚C, thermostability, solvent stability, and pH dependence.
- Of the 654 clones tested, 77 performed better than native subtilisin.
- This approach is called “molecular breeding.”



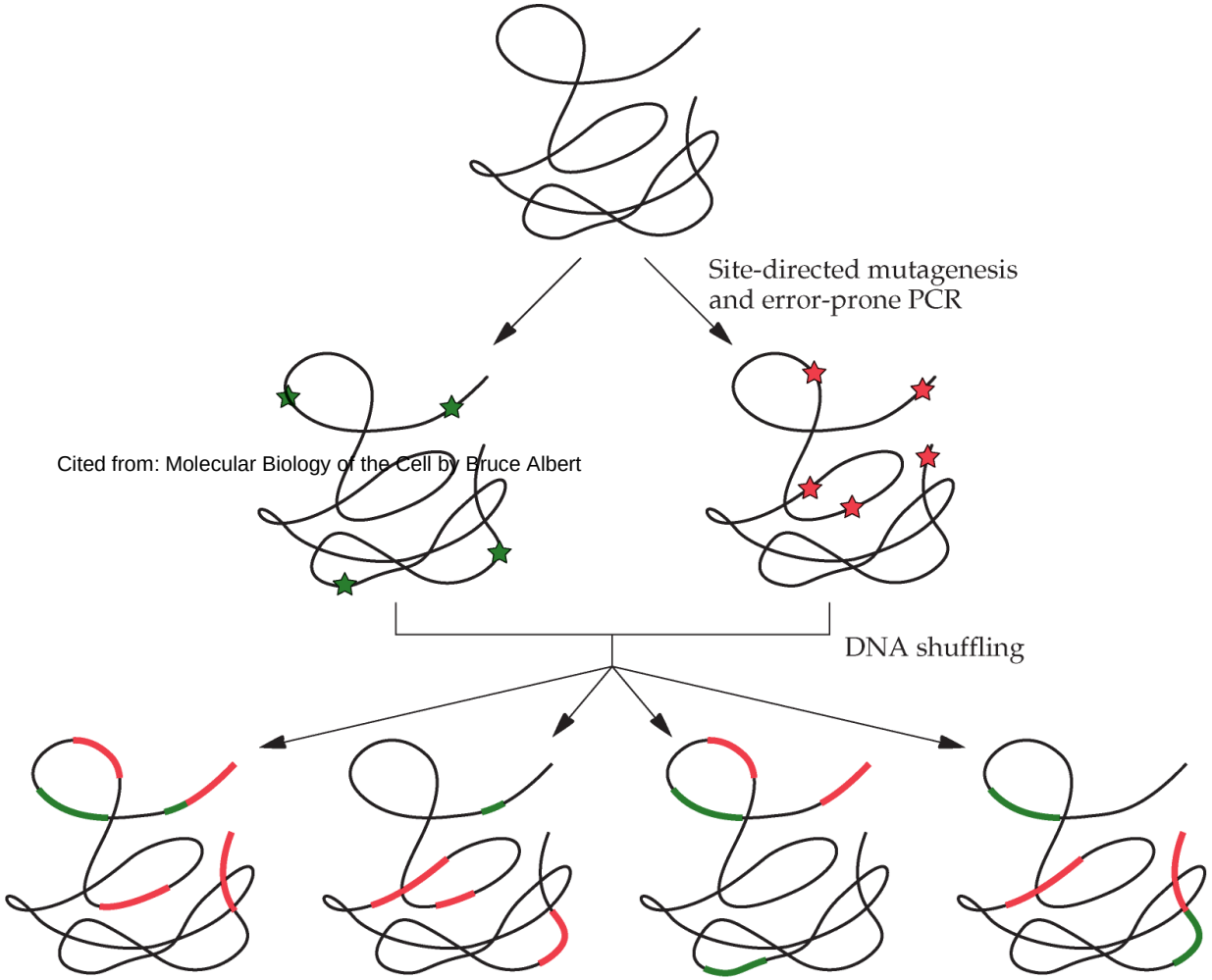
- The entire library is screen for protease activity.
- The clones that generate clear zone are selected.
- The chosen clones are grown under different conditions.
- Enzyme activity is quantified by the intensity of the color that is formed.

Cited from: Enzymes Engineering, Method and Protocol by James C. Samulelson

- Peroxidase from the inky cap mushroom was also altered for use in preventing dye transfer in laundry detergent.

- Using a combination of oligonucleotide-directed mutagenesis, error prone PCR, and DNA shuffling, researchers produced variants with better solvent stability, pH stability, and life span.

….cont



Cited from: Molecular Biology of the Cell by Bruce Albert

Cited from: Enzymes Engineering, Method and Protocol by James C. Samulelson

Thank you