

INTRODUCTION TO

# BIOTECHNOLOGY

# GY BIOTECHNOLOGY

# TECHNOLOGY BI

# NOLOGY BIOTECH

# TECHNOLOGY B

# BIOTECHNOLOGY

William J. Thieman and Michael A. Palladino

THIRD EDITION

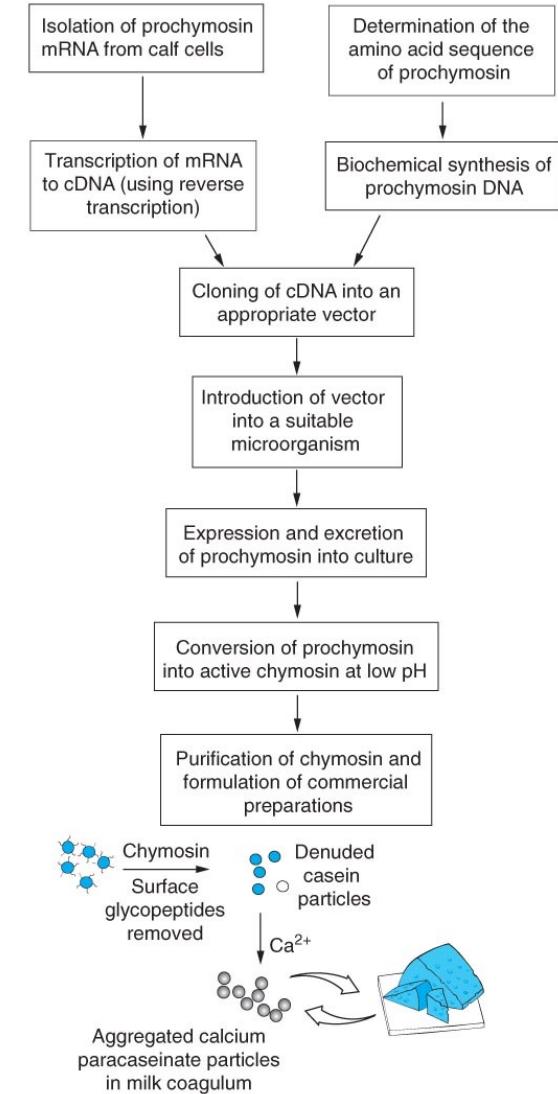
## CHAPTER 4

### Proteins as Products

PowerPoint® Lecture by:  
Melissa Rowland-Goldsmith  
Chapman University

# 4.1 Proteins as Biotechnology Products

- Use of proteins in manufacturing is a time-tested technology
  - Beer brewing and winemaking
  - Cheese making
- Recombinant DNA technology made it possible to produce specific proteins on demand
  - Enzymes
  - Hormones
  - Antibodies



# 4.1 Proteins as Biotechnology Products

**TABLE 4.1 SOME ENZYMES AND THEIR INDUSTRIAL APPLICATIONS**

Enzyme	Application
Amylases	Digest starch in fermentation and processing
Proteases	Digest proteins for detergents, meat/leather, cheese, brewing/baking, animal/human digestive aids
Lipases	Digest lipids (fats) in dairy and vegetable oil products
Pectinases	Digest enzymes in fruit juice/pulp
Lactases	Digest milk sugar
Glucose isomerase	Produce high-fructose syrups
Cellulases/hemicellulases	Produce animal feeds, fruit juices, brewing converters
Penicillin acylase	Produces penicillin

## 4.1 Proteins as Biotechnology Products

- **Biotech Drugs and other medical applications**
  - Produced through microbial fermentation or mammalian cell culture
  - Complicated and time-consuming process
  - Once the method is determined produce large batches of the protein products in bioreactors by growing the host cells that have been transformed to contain the therapeutic gene of interest
    - Cells are stimulated to produce the target proteins through precise culture conditions that include a balance of temp., oxygen, acidity, and other variables
    - At appropriate time the proteins are isolated from the cultures, tested at every step of purification and formulated into pharmaceutically active products

# 4.1 Proteins as Biotechnology Products

**TABLE 4.2 SOME PROTEIN-BASED PHARMACEUTICAL PRODUCTS (MOST PRODUCED AS RECOMBINANT PROTEINS)**

Protein	Application
Erythropoietins	Treatment of anemia
Interleukins 1, 2, 3, 4	Treatment of cancer, AIDS; radiation- or drug-induced bone marrow suppression
Monoclonal antibodies	Treatment of cancer, rheumatoid arthritis; used for diagnostic purposes
Interferons ( $\alpha$ , $\beta$ , $\gamma$ , including consensus)	Treatment of cancer, allergies, asthma, arthritis, and infectious disease
Colony-stimulating factors	Treatment of cancer, low blood cell count; adjuvant chemotherapy; AIDS therapy
Blood clotting factors	Treatment of hemophilia and related clotting disorders
Human growth factor	Treatment of growth deficiency in children
Epidermal growth factor	Treatment of wounds, skin ulcers, cancer
Insulin	Treatment of types 1 and 2 diabetes mellitus
Insulin-like growth factor	Treatment of type 1 diabetes mellitus
Tissue plasminogen factor	Treatment after heart attack, stroke
Tumor necrosis factor	Cancer treatment
Vaccines	Vaccination against hepatitis B, malaria, herpes

## 4.1 Proteins as Biotechnology Products

- **Applications of Proteins in Industry**
  - Food processing
  - Textiles and leather goods
  - Detergents
  - Bioremediation: treating pollution with proteins and clean up harmful wastes with proteins

## 4.2 Protein Structures

- **Proteins**
  - Are complex molecules built of chains of amino acids
  - Have specific molecular weights
  - Have electrical charge that causes them to interact with other molecules
    - *Hydrophilic* – water loving
    - *Hydrophobic* – water hating

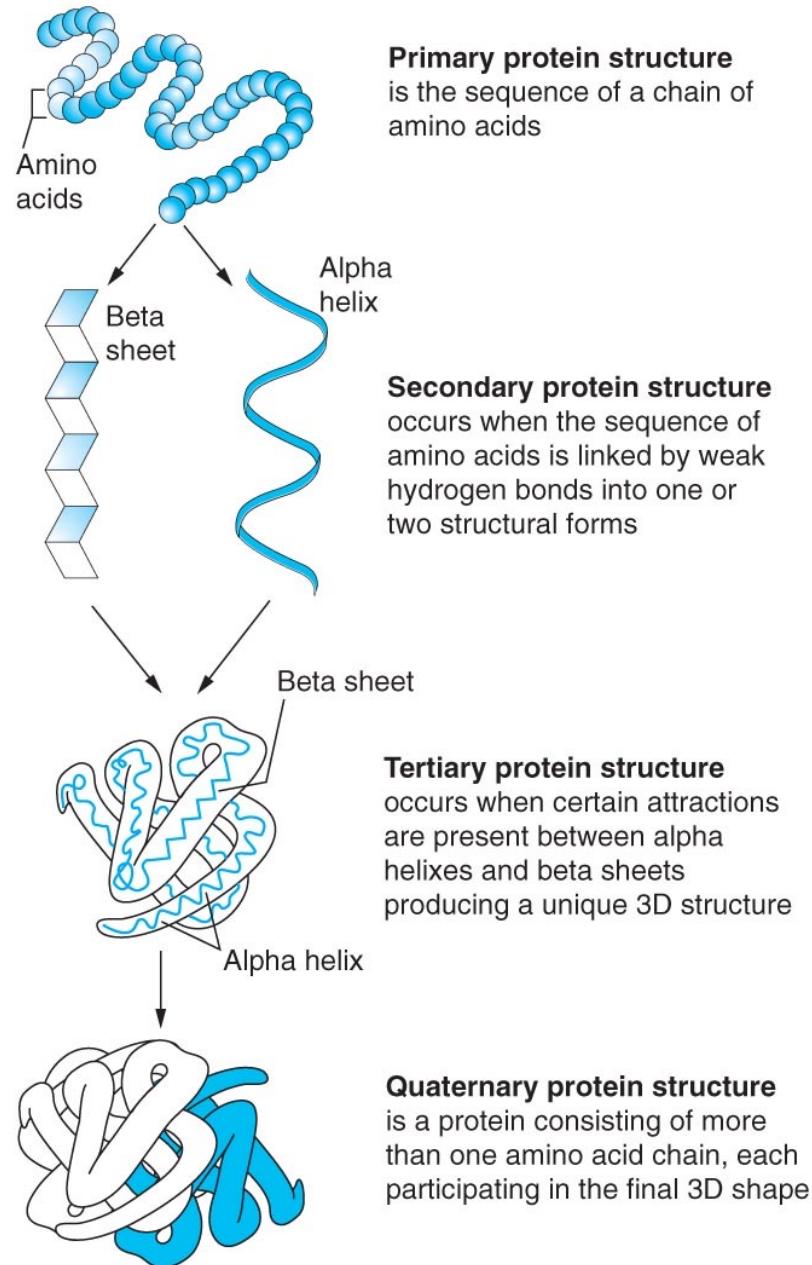
## 4.2 Protein Structures

- **Structural Arrangement – four levels**
  - **Primary structure** is the sequence in which amino acids are linked together
  - **Secondary structure** occurs when chains of amino acids fold or twist at specific points forming new shapes due to the formation of hydrogen bonds between hydrophobic amino acids
    - Most common shapes are alpha helices and beta sheets
      - Alpha helix = amino acids form right handed spiral so hydrogen bonds stabilize the structure linking an amino acid's nitrogen atom to the oxygen atom of another amino acid
      - Beta sheet = hydrogen bonds link nitrogen and oxygen atoms forming sheets that are parallel (chains run in same direction or anti-parallel (chains alternate in direction))

## 4.2 Protein Structures

- Structural Arrangement – four levels
  - **Tertiary structures** = 3 dimensional polypeptides and are formed when secondary structures are cross linked
    - Tertiary structure of protein determines its function
  - **Quaternary structures** are unique, globular, three-dimensional complexes built of several polypeptides
- Why are proteins so fragile?

## 4.2 Protein Structures



## 4.2 Protein Structures

- **Protein Folding**
  - The structure and function of a protein depends on protein folding
  - If protein is folded incorrectly, desired function of a protein is lost and a misfolded protein can be detrimental
  - 1951: two regular structures were described
    - **Alpha helices and beta sheets**
    - Structures are fragile; hydrogen bonds are easily broken
    - Incorrectly folded proteins can lead to diseases including Alzheimers, cystic fibrosis, mad cow, forms of cancer and even some heart attacks
    - \*\*\* a challenge of biotechnology is to understand and control the protein folding in the manufacturing process

## 4.2 Protein Structures

- **Glycosylation** – post-translational modification wherein carbohydrate units are added to specific locations on proteins
- More than 100 post-translational modifications occur within eukaryotic cells
- This change can affect the protein's activity by increasing:
  - its solubility
  - orient proteins into the membranes
  - extend the active life of a molecule in an organism

## 4.2 Protein Structures

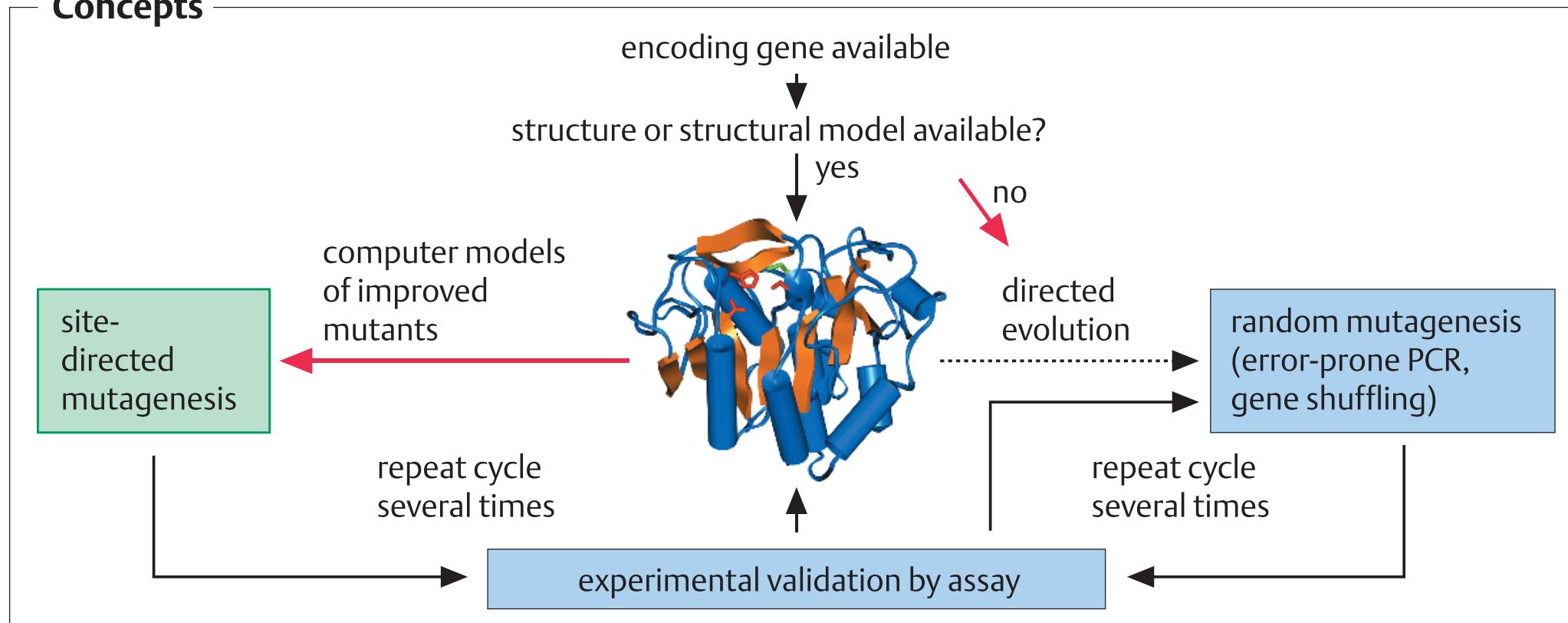
- **Protein Engineering**
  - Biotechnology is able to design and build new proteins in the lab
  - Might be possible to invent proteins that are tailored for specific applications

# Protein engineering

- Protein engineering describes the intentional alteration of a protein's amino acid sequence, usually with the aim of achieving either:
  1. A better understanding of the relationship between a protein's primary and higher-level structure, or its structure and function
  2. The development of a protein variant which, relative to the wild-type protein, displays some enhanced property in the context of its commercial use.
- Protein engineering is facilitated by bioinformatics and molecular biology, in particular **PCR**

# Protein engineering

## Concepts



# Protein engineering approaches

- There are three approaches to protein engineering:

## 1. Directed evolution

- Relies on iterative cycles of genetic diversity creation followed by selection or screening until the desired property is attained.
- The most common methods to generate genetic diversity include error-prone PCR and DNA recombination.

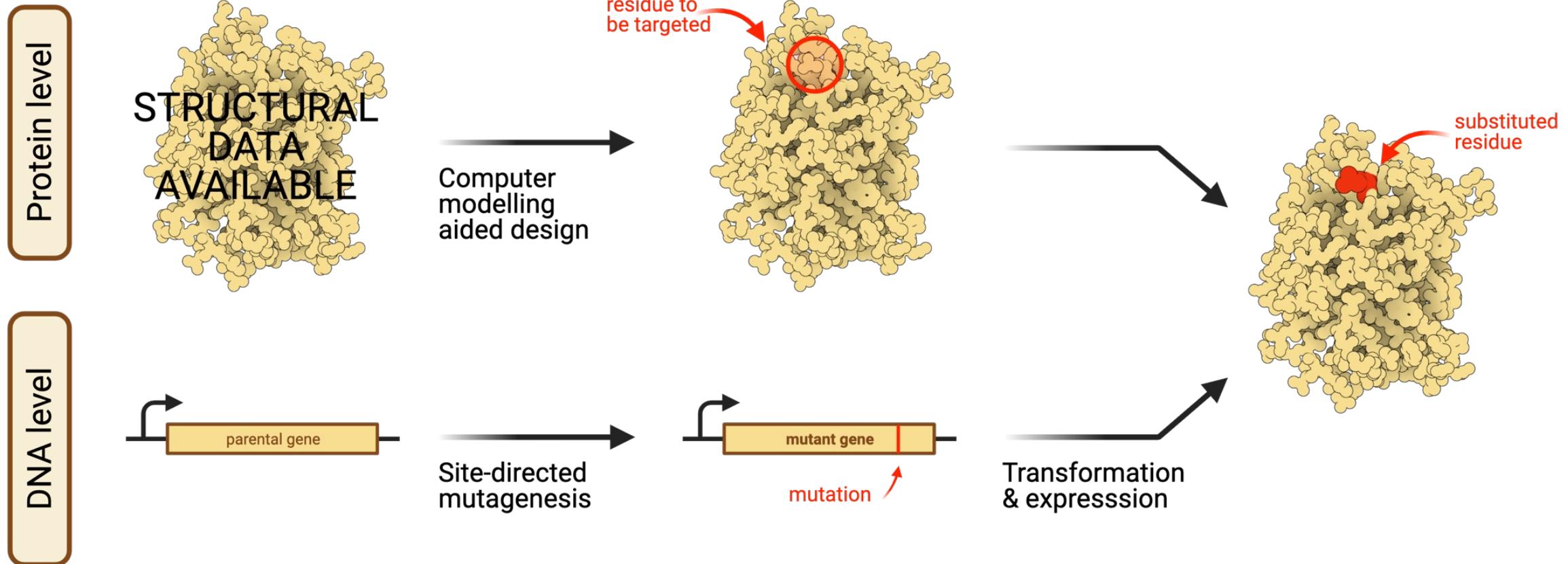
## 2. Rational design

- This is a knowledge-driven process, requiring *a priori* information about the protein such as its structure.
- The knowledge is used to make specific, targeted amino acid mutations.

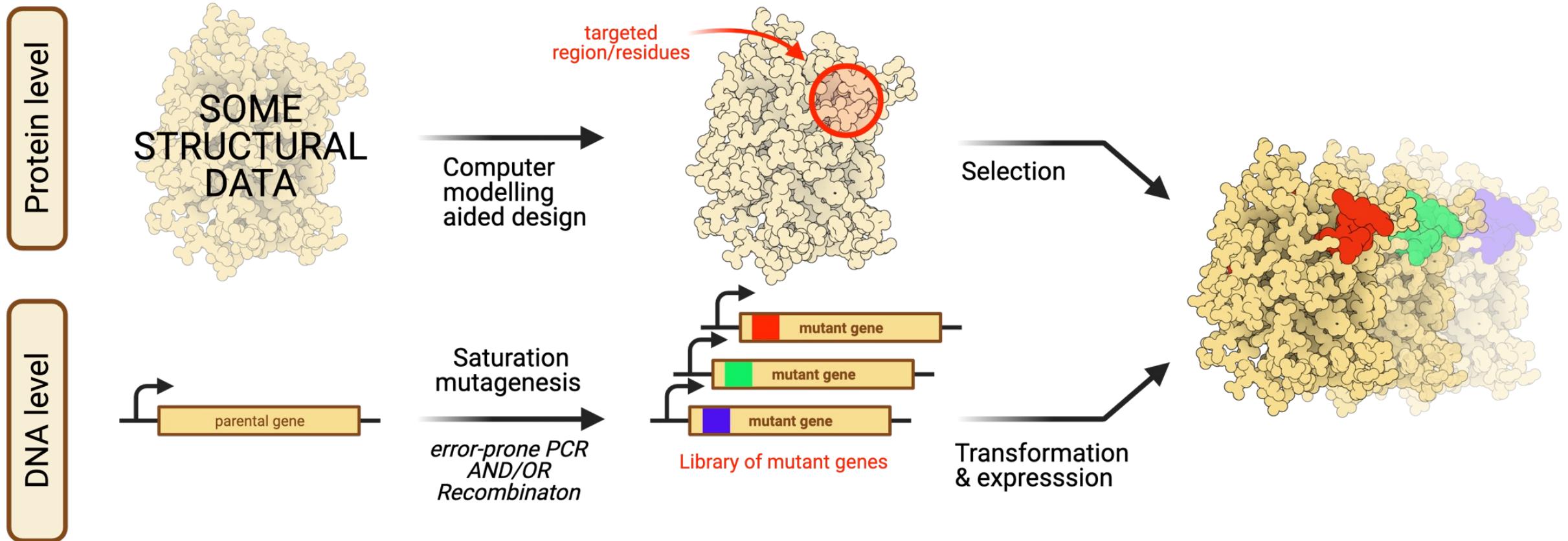
## 3. Semi-rational design:

- This is a combination of rational design and directed evolution.
- It targets specific residues for saturation mutagenesis or mutagenizing a specific domain/region that is suspected to have a crucial effect on the desired property.

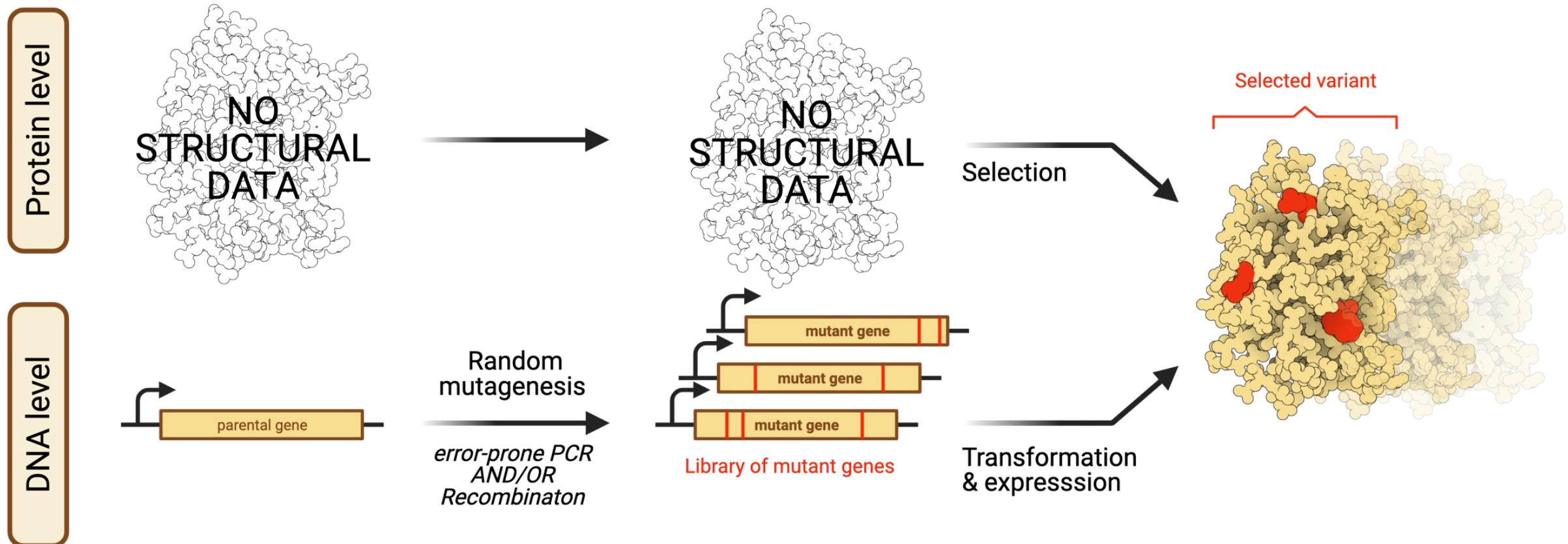
# RATIONAL DESIGN



# SEMI-RATIONAL DESIGN



# DIRECTED EVOLUTION



	Directed evolution	Rational design	Semi-rational design
<b>Parental gene</b>	A single gene or a group of homologous sequences	A single gene	A single gene
<b><i>A priori</i> knowledge requirement</b>	Not required	Required	Required
<b>Genetic diversity creation</b>	Random mutagenesis or DNA recombination	Focused mutagenesis	Focused mutagenesis
<b>Library size</b>	Large	Small	Small to medium
<b>Screening</b>	High to ultra high throughput	Low to high throughput	Low to high throughput
<b>Advantages</b>	<ul style="list-style-type: none"> <li>No prior knowledge of the enzyme structure and mechanism is required.</li> <li>Mutate the entire enzyme, and as such, it is possible to identify mutations distant to the active site that affect the enzymatic activity via allosteric interaction.</li> </ul>	<ul style="list-style-type: none"> <li>Small library size.</li> <li>Less time and effort on screening.</li> <li>Particularly advantageous when there is no high-throughput screening system available.</li> </ul>	<ul style="list-style-type: none"> <li>Library size is significantly reduced compared to directed evolution.</li> <li>A larger portion of the protein sequence space is explored compared to rational design.</li> </ul>
<b>Disadvantages</b>	<ul style="list-style-type: none"> <li>Large library size.</li> <li>Impossible to explore the full protein sequence space, even with the most powerful selection or screening method.</li> <li>Time consuming to develop an assay and to screen large library.</li> <li>Resource intensive.</li> </ul>	<ul style="list-style-type: none"> <li><i>A priori</i> knowledge is required.</li> <li>Mutations are mainly targeted at the active site.</li> </ul>	<ul style="list-style-type: none"> <li><i>A priori</i> knowledge is required.</li> <li>Mutations are mainly targeted at the active site.</li> </ul>

# Mutagenesis

- There are three approaches :

## **1. Focused mutagenesis**

- These methods insert mutations at specific sites or region within a gene.

## **2. Random mutagenesis**

- Introduce mutations at random positions along a target gene..

## **3. Recombination:**

- Larger gene segments are exchanged between a group of homologous genes, thereby creating chimeras.

# Site-directed mutagenesis

- Site-directed mutagenesis facilitates the introduction of specific predefined sequence alterations into a protein's backbone
- Site-directed mutagenesis changes the DNA sequence encoding the target protein to encode a different amino acid at the selected location
- Such alterations can include:
  - Insertion
  - Deletion
  - Replacement of either a single or multiple amino acid residues.

# Site-directed mutagenesis

- Useful for:
  1. Identification of residue(s) (or indeed sequence) is important in the context of some aspect of protein structure and/or function.
    - For Alanine (**alanine scanning**) is often used in this context as its side chain is non-bulky, non-charged, chemically unreactive and is compatible with the formation of various structural motifs
  2. Alteration of protein characteristics
    - For example, many modern detergent enzymes have been rendered oxidation resistant via the replacement of oxidation-sensitive methionine residues on their surface

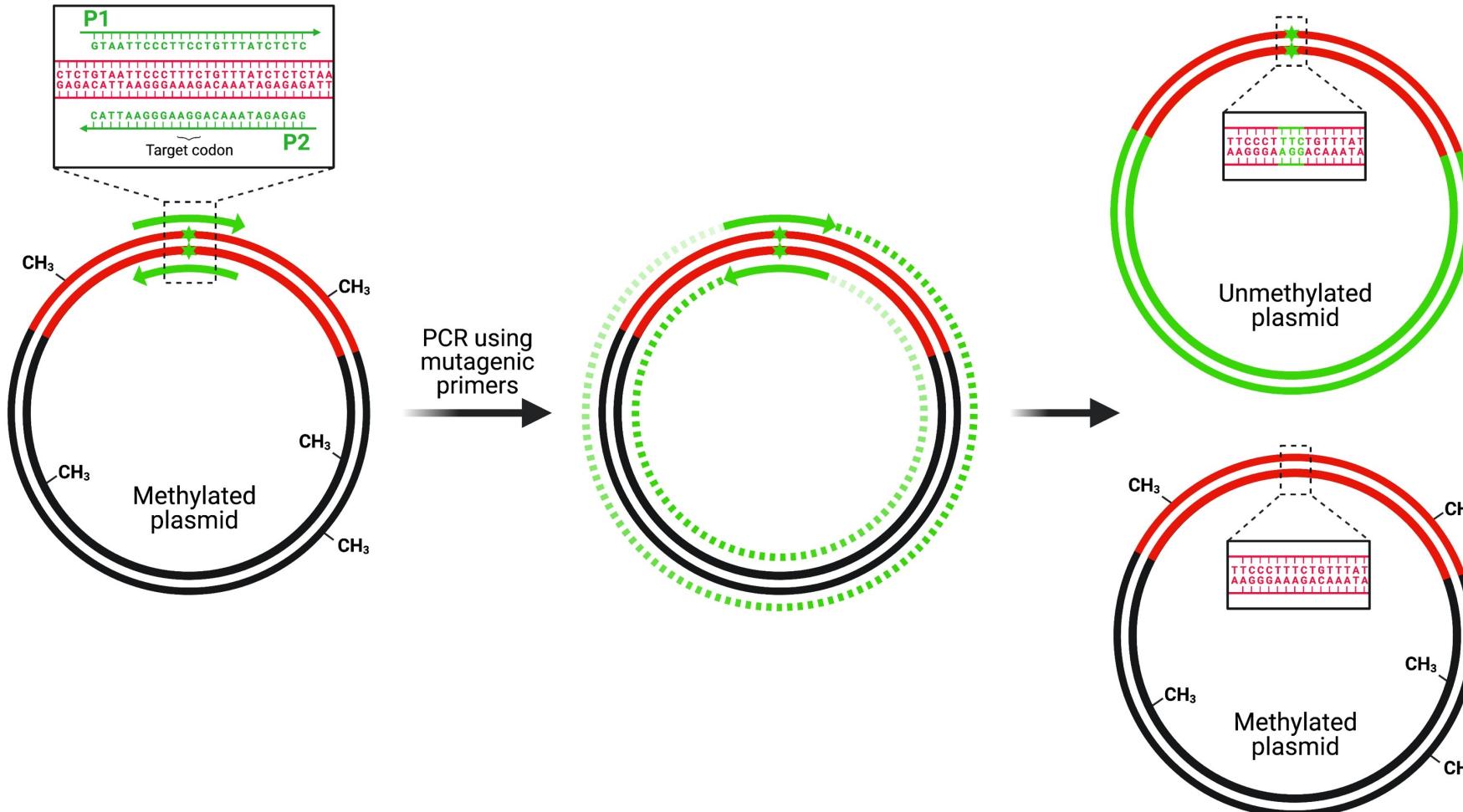
# Site-directed mutagenesis

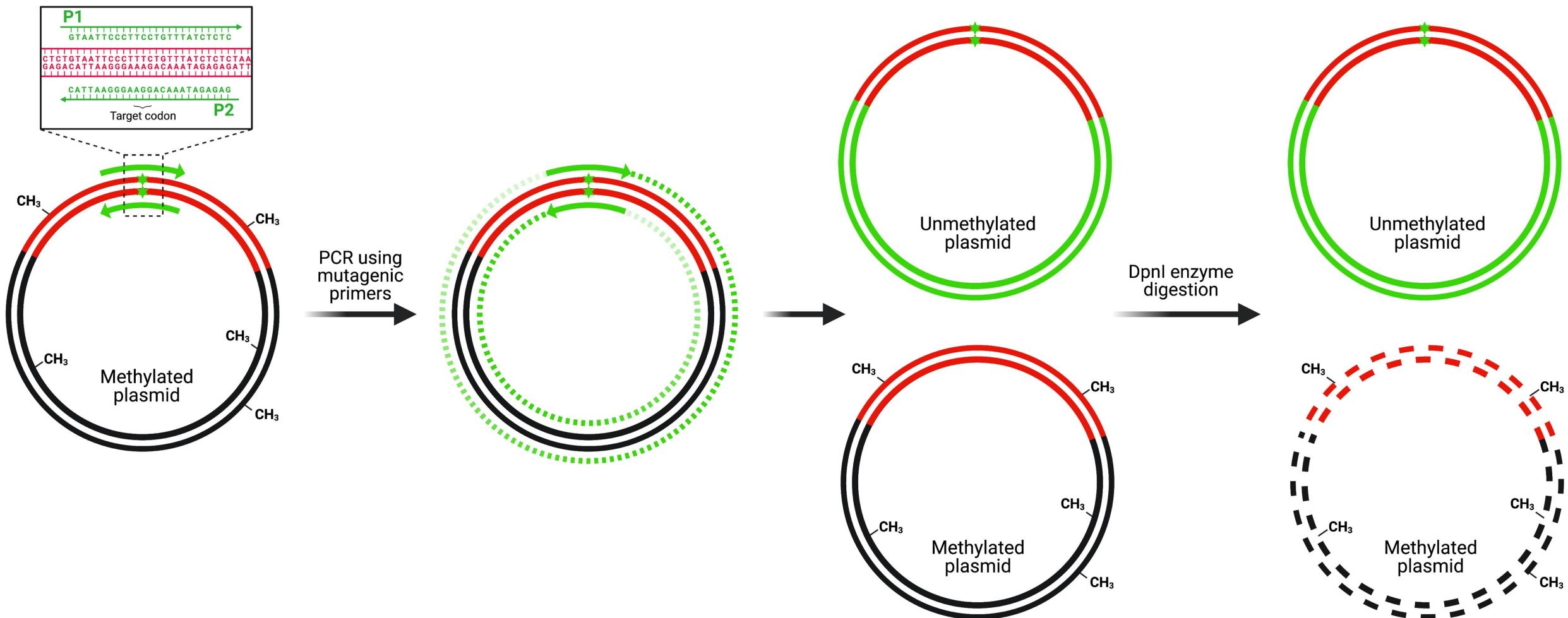
- As an example of site-directed mutagenesis, consider the replacement of methionine 222 with alanine to stabilize subtilisin for use in laundry detergent

...	218	219	220	221	222	223	224	225	...	amino acid numbering
...	AAC	GGT	ACG	TCA	ATG	GCA	TCT	CCG	...	original DNA sequence
...	Asn	Gly	Thr	Ser	Met	Ala	Ser	Pro	...	original protein encoded
...	AAC	GGT	ACG	TCA	gcG	GCA	TCT	CCG	...	modified DNA sequence
...	Asn	Gly	Thr	Ser	Ala	Ala	Ser	Pro	...	variant protein encoded

# Site-directed mutagenesis with OneClick

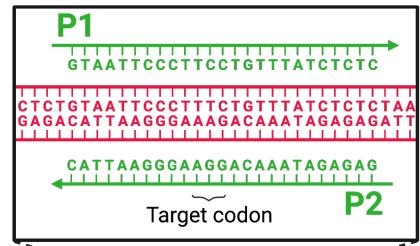
- OneClick is a free online tool for designing and setting up reactions for site-directed mutagenesis.
- The tool incorporates information of most DNA polymerases and plasmid systems.
- It provides flexibility in choosing the type of mutagenic primers to design, as well as the PCR modes
- OneClick webpage (<http://www.oneclick-mutagenesis.com/>)



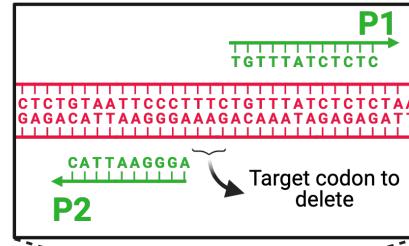


# Site-directed mutagenesis strategies

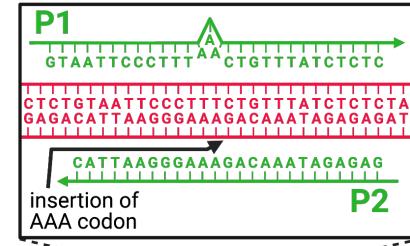
## Substitution



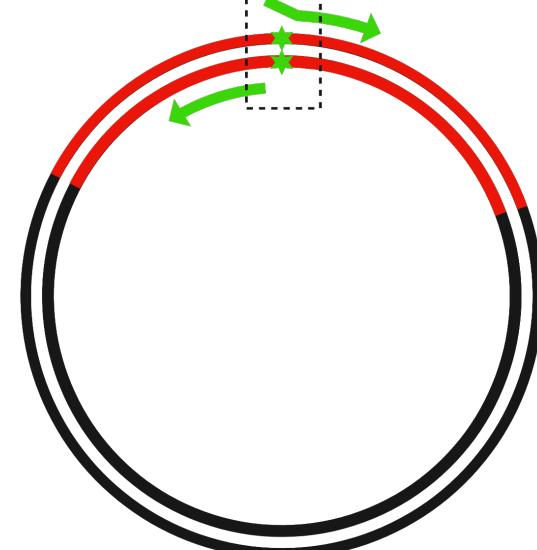
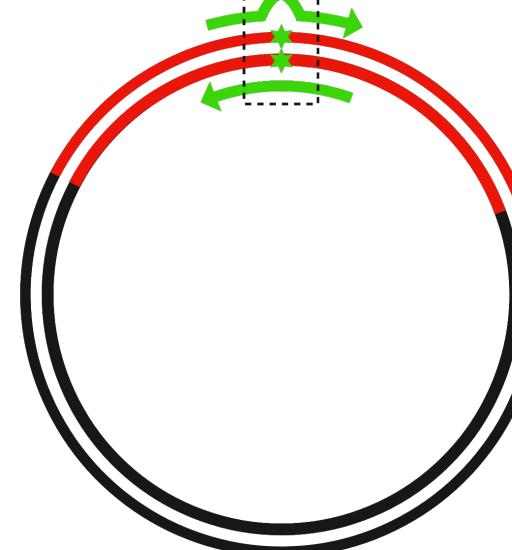
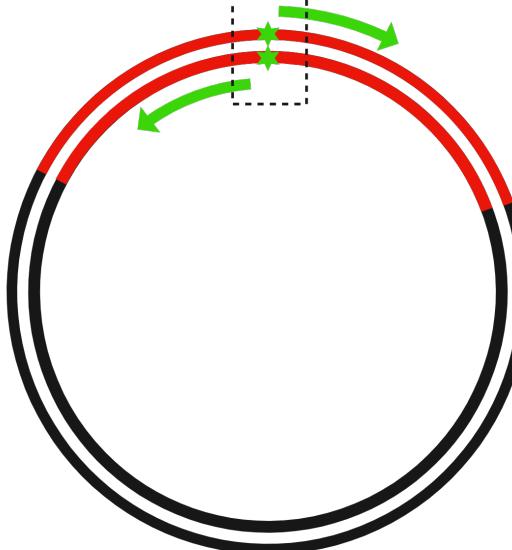
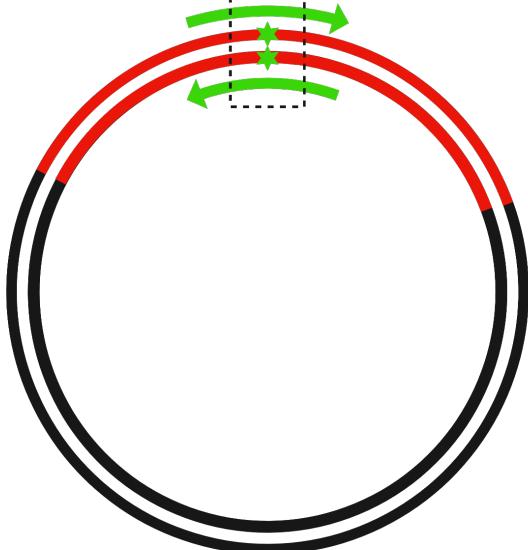
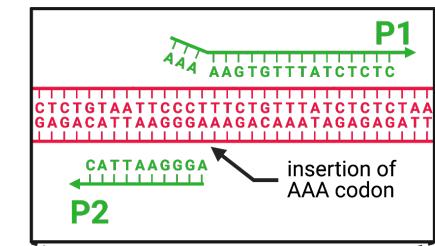
## Deletion



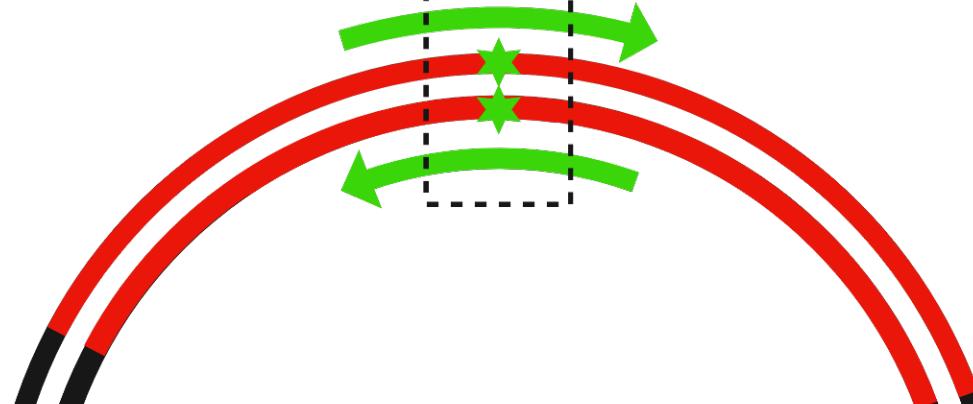
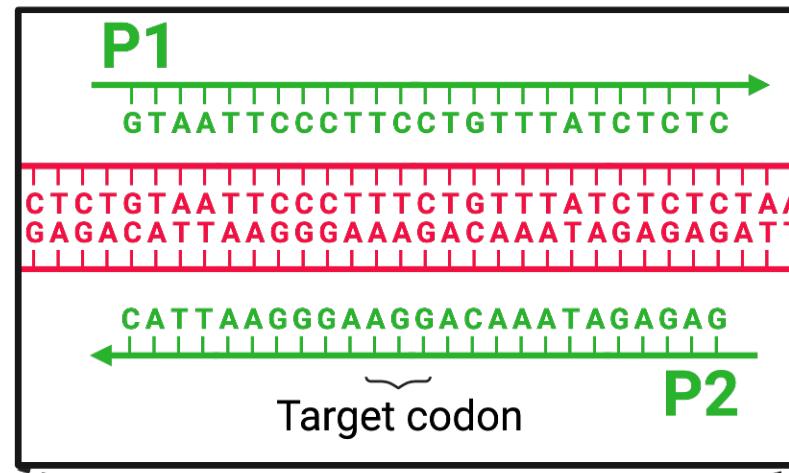
## Insertion



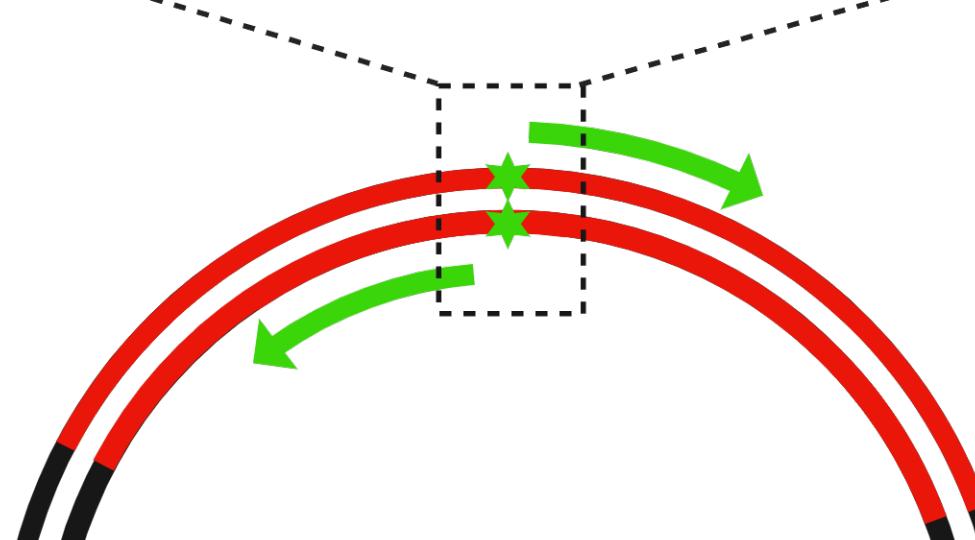
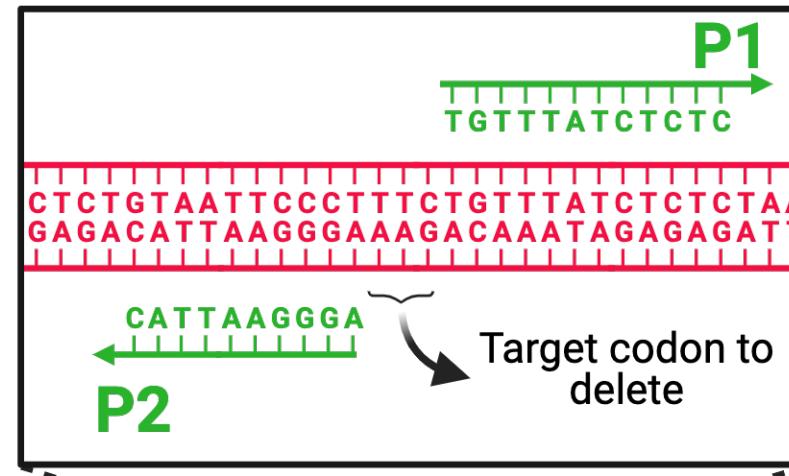
## Insertion



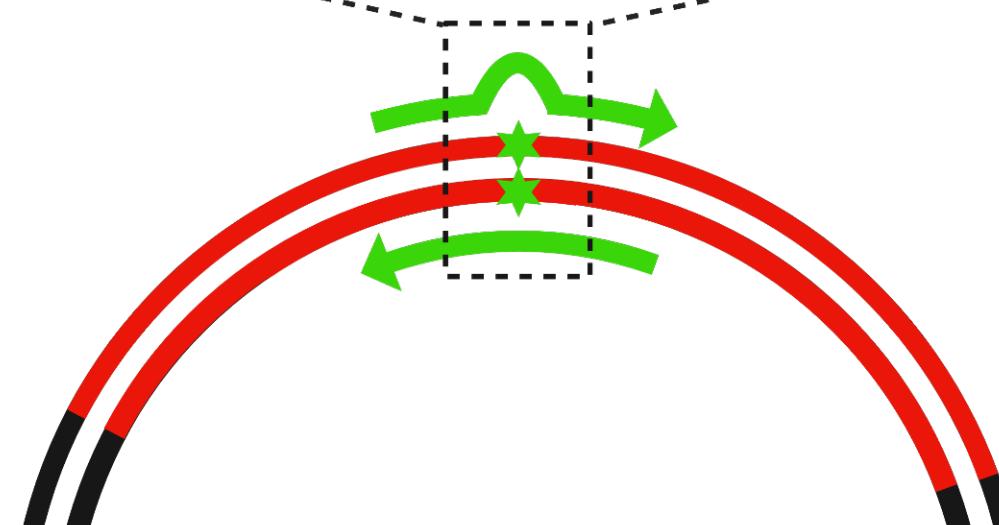
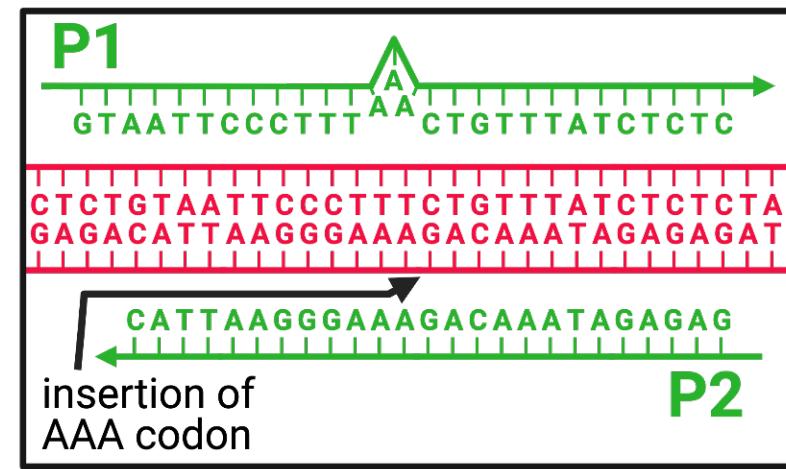
# Substitution



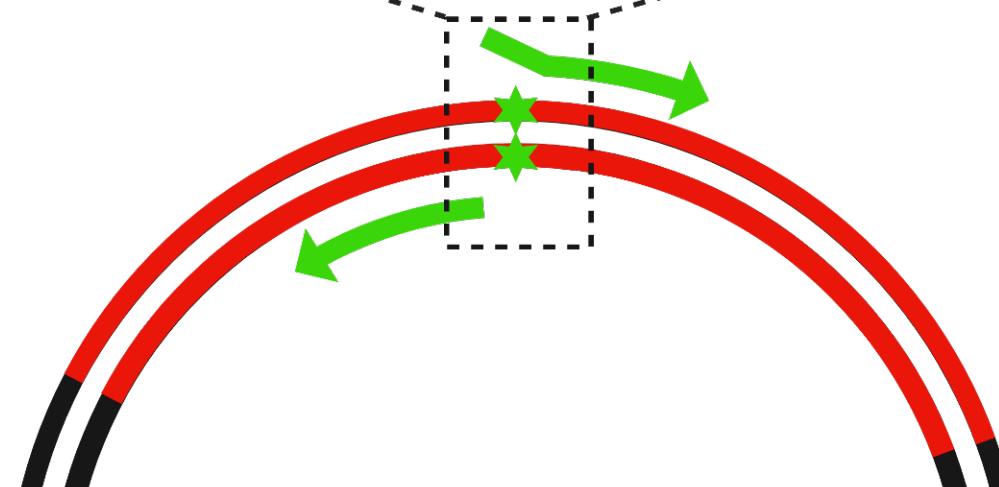
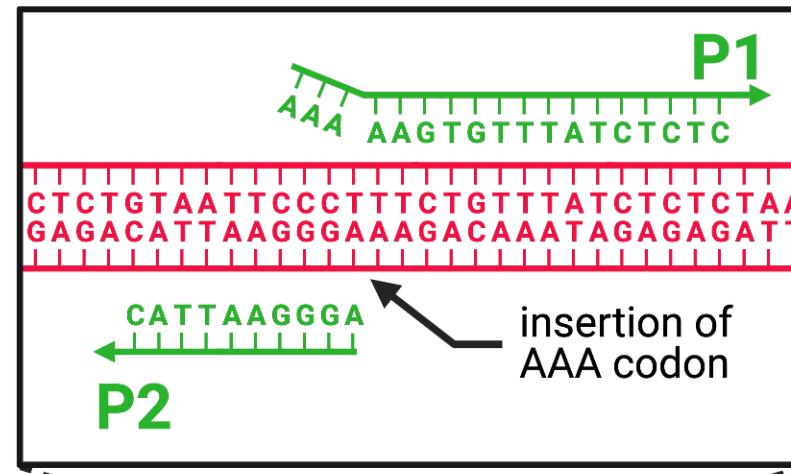
# Deletion



# Insertion



# Insertion



# Exhaustive search of all single amino acid substitutions

- The straightforward way to find a single amino acid substitution that improves a protein is to make all possible single amino acid substitution variants and test them.
- The advantage of such systematic screening is completeness.
- Testing every possibility ensures finding the best one. Other methods described below make incomplete sets of single-substitution variants, so the best possible variant might not be in the library.
- For example, libraries of single substitution variants created by error-prone PCR are incomplete and would likely not include the Ala190His variant.

# Site-saturation mutagenesis

- Site-saturation mutagenesis makes all nineteen possible amino-acid substitutions at a particular location.
- Involves the use of primers that encode multiple amino acids. These primers contain degenerate codons, which are mixtures of codons prepared by DNA synthesis
- Site-directed mutagenesis using degenerate codons creates a mixture of multiple protein variants.

one-letter code	nucleotides
N (any)	A, T, G, C
B (not A)	T, G, C
D (not C)	A, T, G
H (not G)	A, T, C
V (not T)	A, G, C
K (keto)	G, T
M (amino)	A, C
R (purine)	A, G
S (strong)	G, C
W (weak)	A, T
Y (pyrimidine)	C, T

# Degenerate codons encode multiple amino acids

wild type sequence

5'...GATTGCAGCTGCTTTCCACAATTCAGTATTGCCAGACACCGAGC...

primer for site-directed mutagenesis His103Val

5' GATTGCAGCTGCTTTCGTGAATTCAGTATTGCCAGAC 3'

		Second Base					
		U	C	A	G		
First Base	U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr Stop Stop	Cys Cys Stop Trp	U C A G	Third Base
	C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G	
	A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G	
	G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G	

one-letter code	nucleotides
N (any)	A, T, G, C
B (not A)	T, G, C
D (not C)	A, T, G
H (not G)	A, T, C
V (not T)	A, G, C
K (keto)	G, T
M (amino)	A, C
R (purine)	A, G
S (strong)	G, C
W (weak)	A, T
Y (pyrimidine)	C, T

# Degenerate codons encode multiple amino acids

wild type sequence

5'...GATTGCAGCTGCTTTCCACAATTCAGTATTGCCAGACACCGAGC...

primer for site-directed mutagenesis His103Val

5' GATTGCAGCTGCTTTCGTAATTCAGTTATTGCCAGAC 3'

primer mixture for site-saturation mutagenesis His103Xxx

5' GCTTTTCNNKAATTCAGTATTGCCAGAC 3'

		Second Base					
		U	C	A	G		
First Base	U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr Stop Stop	Cys Cys Stop Trp	U C A G	Third Base
	C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G	
	A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G	
	G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G	

one-letter code	nucleotides
N (any)	A, T, G, C
B (not A)	T, G, C
D (not C)	A, T, G
H (not G)	A, T, C
V (not T)	A, G, C
K (keto)	G, T
M (amino)	A, C
R (purine)	A, G
S (strong)	G, C
W (weak)	A, T
Y (pyrimidine)	C, T

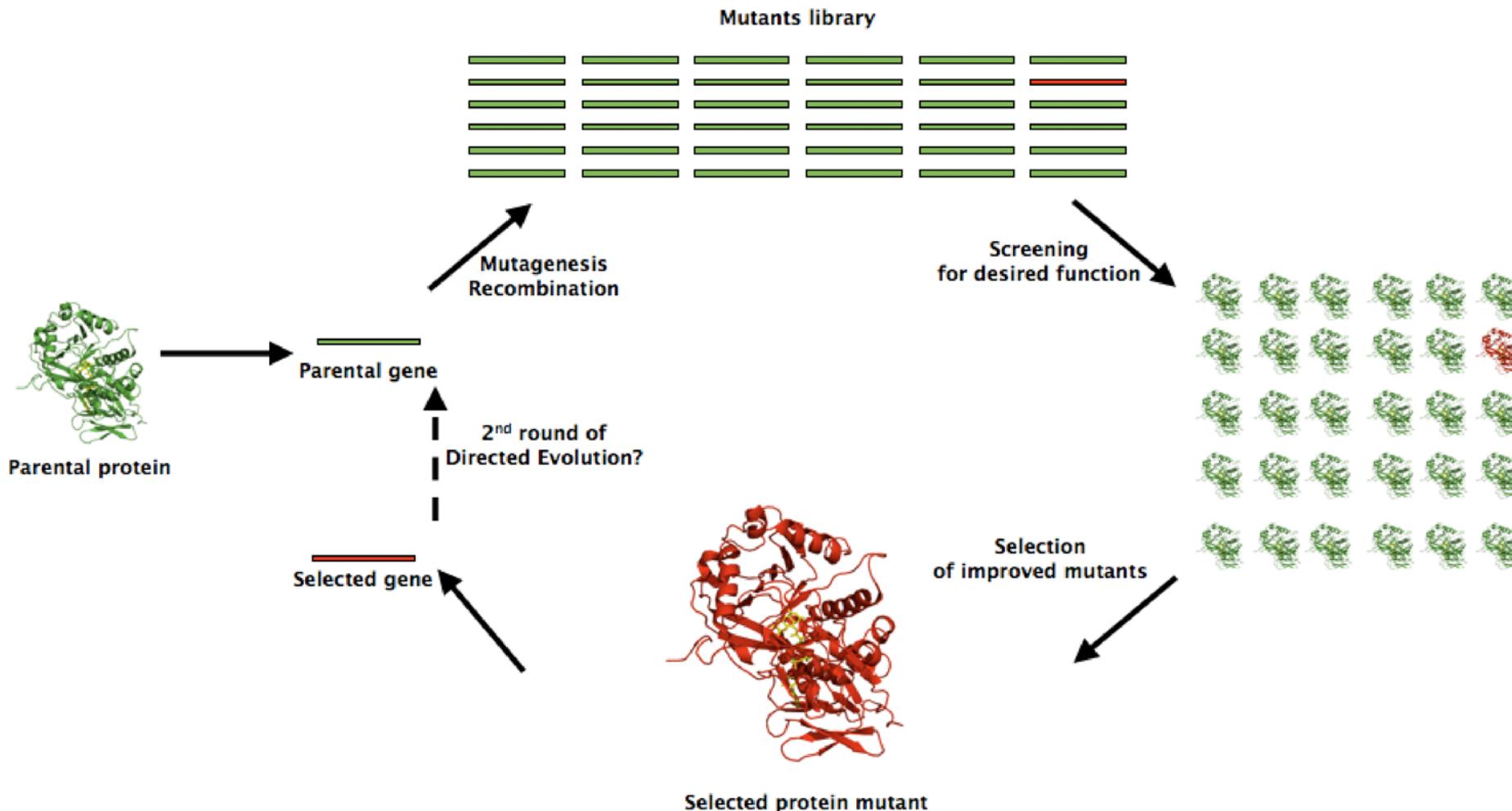
# Degenerate codons encode multiple amino acids

codon	nucleotide mixture	encoded amino acid	codon	nucleotide mixture	encoded amino acid	one-letter code	nucleotides
AAT		Asn		AAA	Lys	N (any)	A, T, G, C
AGT		Ser		ACA	Thr		T, G, C
ATT		Ile	VMA	GAA	Glu		A, T, G
TAT		Tyr		GCA	Ala		A, T, C
TGT		Cys		CAA	Gln		A, G, C
TTT		Phe		CCA	Pro		G, T
GAT		Asp		ATG	Met		A, C
GGT		Gly		TGG	Trp		A, G
GTT		Val					G, C
CAT		His					A, T
CGT		Arg				Y (pyrimidine)	C, T
CTT		Leu					

## Directed evolution

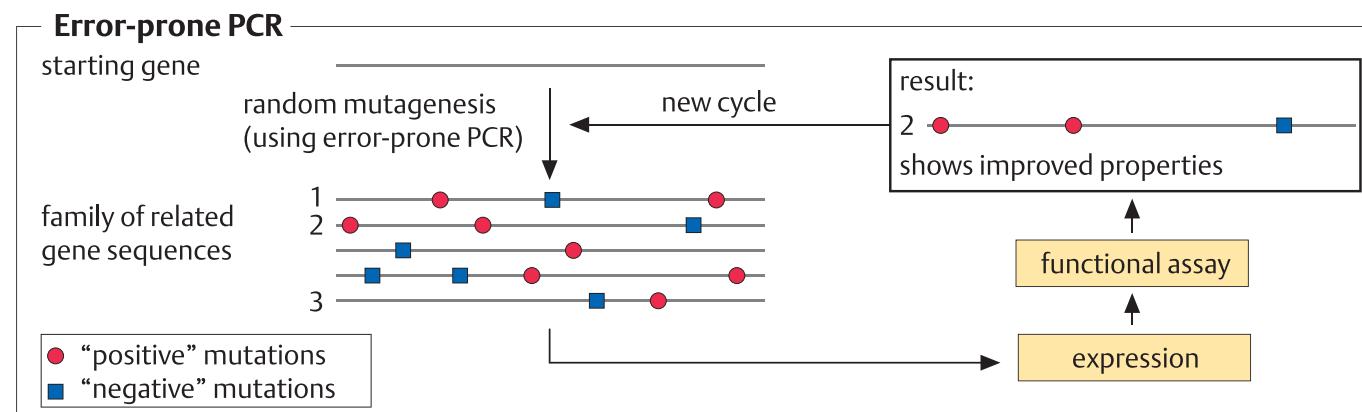
- The introduction of random mutations into the gene coding for the protein of interest, thus generating a large library of gene variants. This process is known as **diversification**.
- The library of variants is expressed and screened to identify variants displaying enhanced target characteristics
- The gene coding for the desired variant is then cloned and sequenced in order to identify the exact amino acid alterations present.
- Diversification is usually achieved in practice using a technique termed **error-prone PCR or gene shuffling**

# Directed evolution



# Error-prone PCR

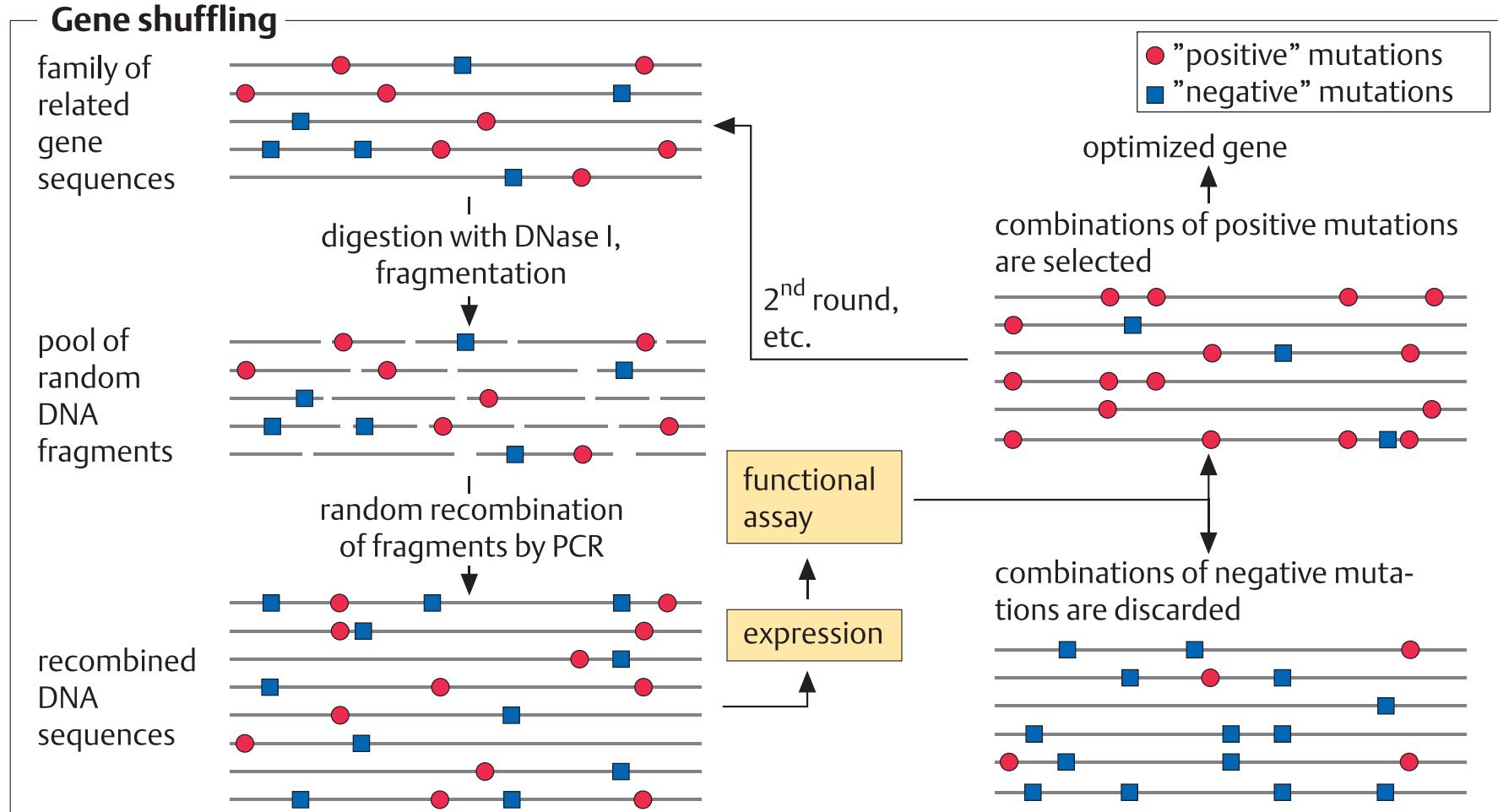
- In error-prone PCR reaction conditions are controlled in order to reduce replication fidelity by:
  1. The use of DNA polymerase such as Taq polymerase, which lacks proofreading capability.
  2. Replacement of the polymerase's natural cofactor ( $Mg^{2+}$ ) with  $Mn^{2+}$
  3. Including suboptimal relative concentrations of dNTPs in the reaction mixture



## Gene shuffling

- Gene shuffling facilitate the recombination of several individual beneficial mutations from different mutants into a single protein.
- The process entails fragmentation of several mutant genes, generally into lengths in the region of 50 base pairs.
- The fragment mix is then subject to PCR, but in the absence of a primer.
- Overlapping fragments will anneal with each other, allowing eventual extension to full gene sequence size.
- The recombinated genes can be cloned, expressed and screened for improved target characteristics.

# Gene shuffling



# Protein post-translational modification

- Polypeptides undergo covalent modification, either during or after their ribosomal synthesis is a process called **post-translational modification, PMT**).
- PTMs are characteristic particularly of eukaryotic proteins and are introduced by specific enzymes or enzyme pathways.
- Many occur at the site of a specific characteristic protein sequence (signature sequence) within the protein backbone.

# Protein post-translational modification

**Table 2.7** The more common forms of post-translational modifications that polypeptides may undergo.  
Refer to text for additional details.

Modification	Comment
Glycosylation	For some proteins glycosylation can increase solubility, influence biological half-life and/or biological activity
Proteolytic processing	Various proteins become biologically active only on their proteolytic cleavage (e.g. some blood factors)
Phosphorylation	Influences/regulates biological activity of various regulatory proteins including polypeptide hormones
Acetylation	Modulation of target protein activity
Acylation	May help some polypeptides interact with/anchor in biological membranes
Amidation	Influences biological activity/stability of some polypeptides
Sulfation	Influences biological activity of some neuropeptides and the proteolytic processing of some polypeptides
Hydroxylation	Important to the structural assembly of certain proteins
$\gamma$ -Carboxyglutamate formation	Important in allowing some blood proteins to bind calcium
ADP-ribosylation	Regulates biological activity of various proteins
Disulfide bond formation	Helps stabilize conformation of some proteins

## 4.3 Protein Production

- Proteins are valuable, complex and fragile products
  - Production of proteins is a long and painstaking process
  - Two major phases used in producing proteins
    - *Upstream processing* includes the actual expression of the protein in the cell
    - *Downstream processing* involves purification of the protein and verification of the function; a stable means of preserving the protein is also required

## 4.3 Protein Production

- Protein Expression: Upstream Processing
  - Selecting the cell to be used as a protein source
    - Bacteria
      - Fermentation processes are well understood
      - Cultured in large quantities in short period of time
      - Relatively easy to alter genetically
      - Can increase level of production of a bacterial protein by introducing additional copies of relevant gene to the host cell via recombinant DNA technology (Chapter 3) **How is the gene under control of expression?**
      - Most common bacterial species used is *E. coli*

## 4.3 Protein Production

- **Protein Expression: Upstream Processing**
  - Selecting the cell to be used as a protein source
    - Bacteria- majority of proteins synthesized naturally by *E coli* are intracellular
      - Most cases resultant foreign protein accumulates in cell's cytoplasm in form of insoluble clumps = **inclusion bodies** which must be purified from other proteins
    - Sometimes the genetically engineered *E coli* produce the desired protein in the form of a fusion protein
      - Target protein is fused to bacterial protein so an additional step is necessary to break the two apart
        - » Fused bacterial protein is usually an enzyme that will bind to its substrate and can be attached to a purification column

## 4.3 Protein Production

- **Protein Expression: Upstream Processing**
  - There are pros and cons to expressing proteins in bacteria
    - Prokaryotic cells are unable to carry out processes such as glycosylation

## 4.3 Protein Production

**TABLE 4.3 ADVANTAGES AND DISADVANTAGES OF RECOMBINANT PROTEIN PRODUCTION IN *E. COLI***

**Advantages**

*E. coli* genetics are well understood

Almost unlimited quantities of proteins can be generated

Fermentation technology is well understood

**Disadvantages**

Foreign proteins produced as inclusion bodies must be refolded

Proteins cannot be folded in ways needed for many proteins active in mammalian systems

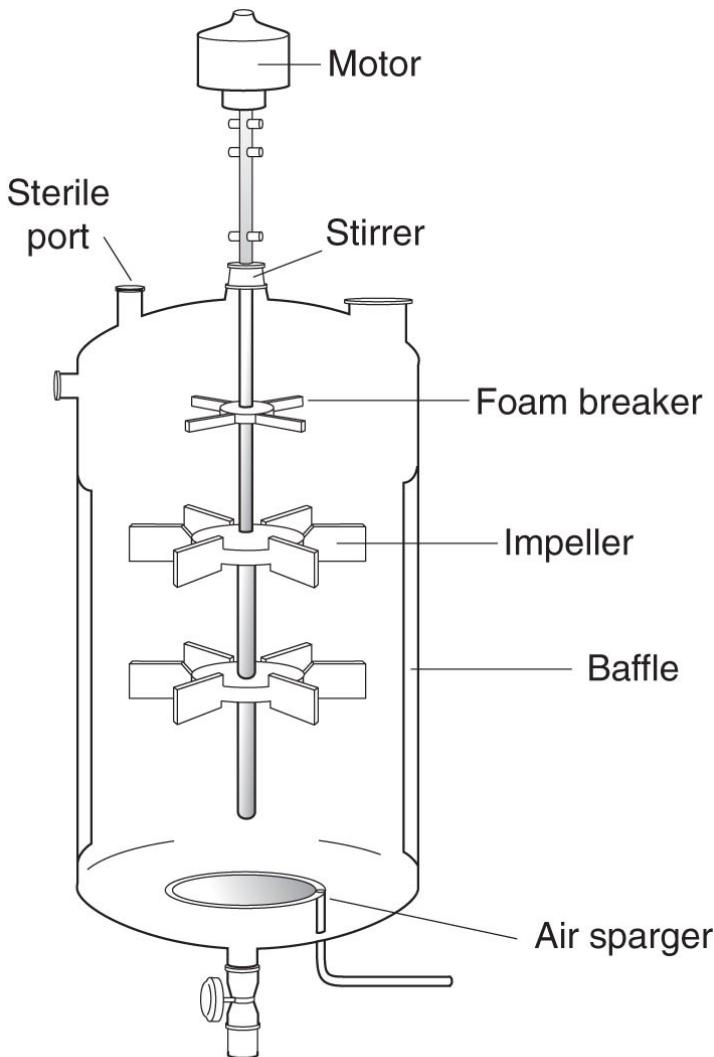
Some proteins are inactive in humans

## 4.3 Protein Production

- **Protein Expression: Upstream Processing**
  - Genetically engineered bacteria can be grown in large scale fermenters (anaerobic) or bioreactors (aerobic)
    - Computers monitor air in bioreactors keeping oxygen levels and temperature ideal for cell growth
    - When phase of cell growth is highest- the bacteria promoter must be activated to stimulate foreign gene expression
      - Must activate gene in recombinant organism after the organism has completed synthesizing important natural proteins needed for its metabolism

## 4.3 Protein Production

- Protein Expression: Upstream Processing



## 4.3 Protein Production

- **Protein Expression: Upstream Processing**

- Fungi

- Source of wide range of proteins used in products including animal feed and beer
    - Many species of fungi are used as hosts for engineered proteins
    - Fungi are eukaryotic and capable of posttranslational modifications to allow proper folding of proteins

## 4.3 Protein Production

- **Protein Expression: Upstream Processing**
  - Plants: used for protein expression
  - Plants can be genetically engineered to produce specific proteins that do not occur naturally
- Process encourages rapid growth and reproductive rates in plants
  - Tobacco plant- great choice for biotech protein production
  - Once genetic material is integrated a million new plant protein factories can fill the fields
  - Disadvantages using plants: not all proteins can be expressed in plants; have cell wall which makes purification difficult; process of glycosylation is slightly different from animal cells

## 4.3 Protein Production

- **Protein Expression: Upstream Processing**
  - Mammalian cell culture systems = challenging
    - Nutritional requirements are complex
    - Mammalian cells grow slowly
    - Easy to contaminate
    - Best choice for proteins destined to be used in humans

## 4.3 Protein Production

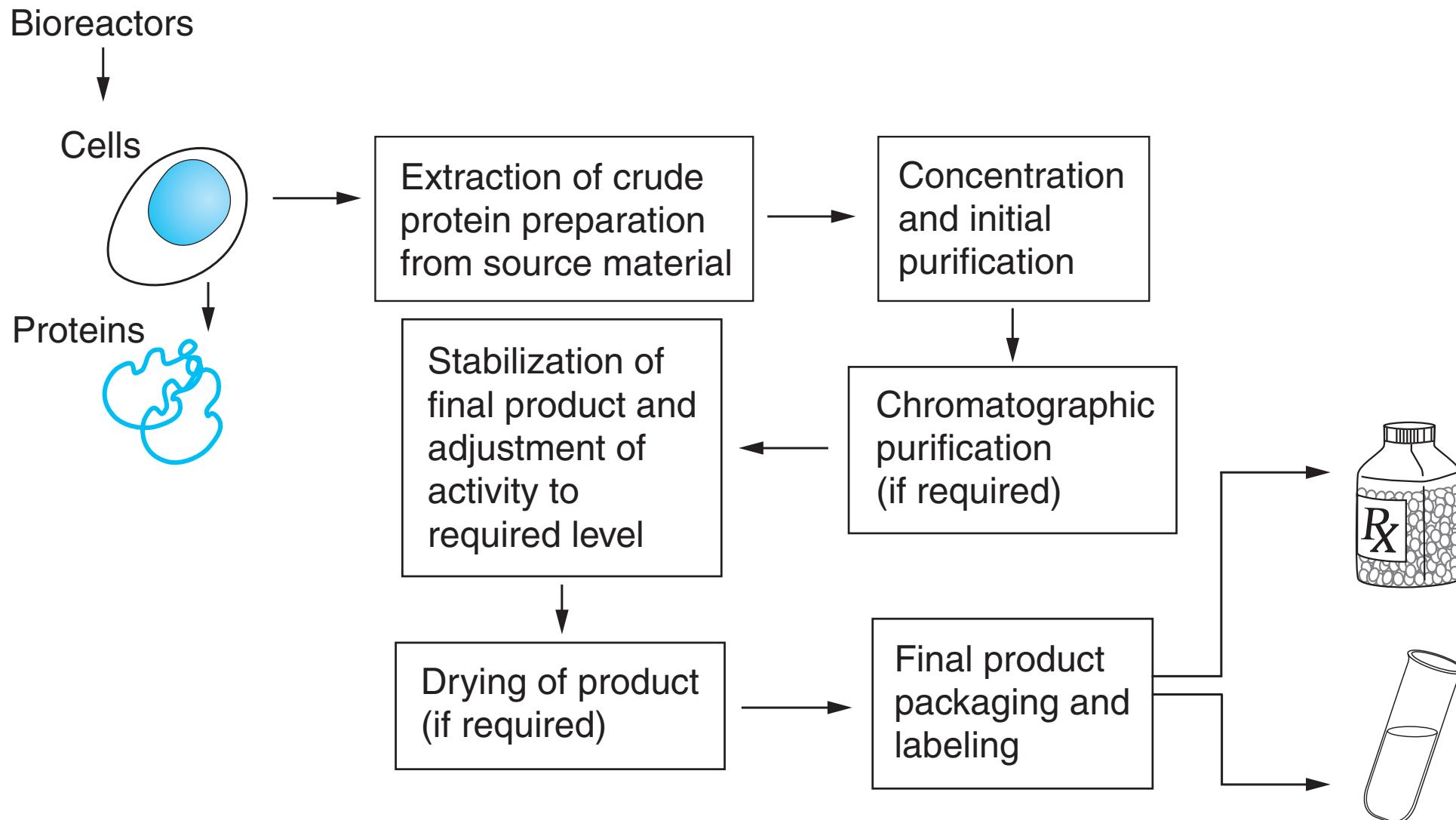
- **Protein Expression: Upstream Processing**
  - Animal bioreactor production systems
  - Use for monoclonal antibody production
    - Mice are injected with an antigen and mouse secretes the desired antibody and the antibody is then purified
  - What is an antibody? Proteins produced in reaction to antigens- invading viruses or bacteria
    - Antibodies combine with and neutralize an antigen protecting the organism
    - Production of antibodies is part of immune response that helps living things resist infectious disease

## 4.3 Protein Production

- **Protein Expression: Upstream Processing**

- Insect system
  - Baculoviruses (viruses that infect insects) are used as vehicles to insert mammalian DNA causing the desired proteins to be produced by insect cells
    - Sometimes the posttranslational modification of proteins is slightly different in insects than in mammals
    - Currently used when small quantities of proteins are needed in research

## 4.3 Protein production



## 4.3 Protein production

- **Downstream processing methods:**
  - Protein purification
    - Step 1: preparing an extract for purification
      - Protein must be harvested
  - Entire cell is harvested if protein is intracellular
    - Requires cell lysis to disrupt the cell wall and release the protein
      - Methods include freeze/thaw; detergents; mechanical methods mixture which releases the entire contents of the cell
    - After cells are ruptured organic alcohols and salts are added to mixture
      - agents increase the interactions between protein molecules to separate them from the mixture
  - Culture medium is collected if the protein is extracellular because it is excreted into the culture medium

## 4.3 Protein production

- **Downstream processing methods:**
  - Protein purification
    - Step 2: stabilizing proteins in solution
      - Maintain low temperature and proper pH in buffering solution
      - Add protease inhibitors and antimicrobials to prevent proteins from being digested
        - » Important to remove these additives later during purification process
      - Add additives to prevent foaming and shearing from destroying the protein
        - » These additives must be removed later

## 4.3 Protein production

- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Similarities between proteins allow them to be separated from other macromolecules
      - Differences between individual proteins are used to separate target proteins from others

## 4.3 Protein production

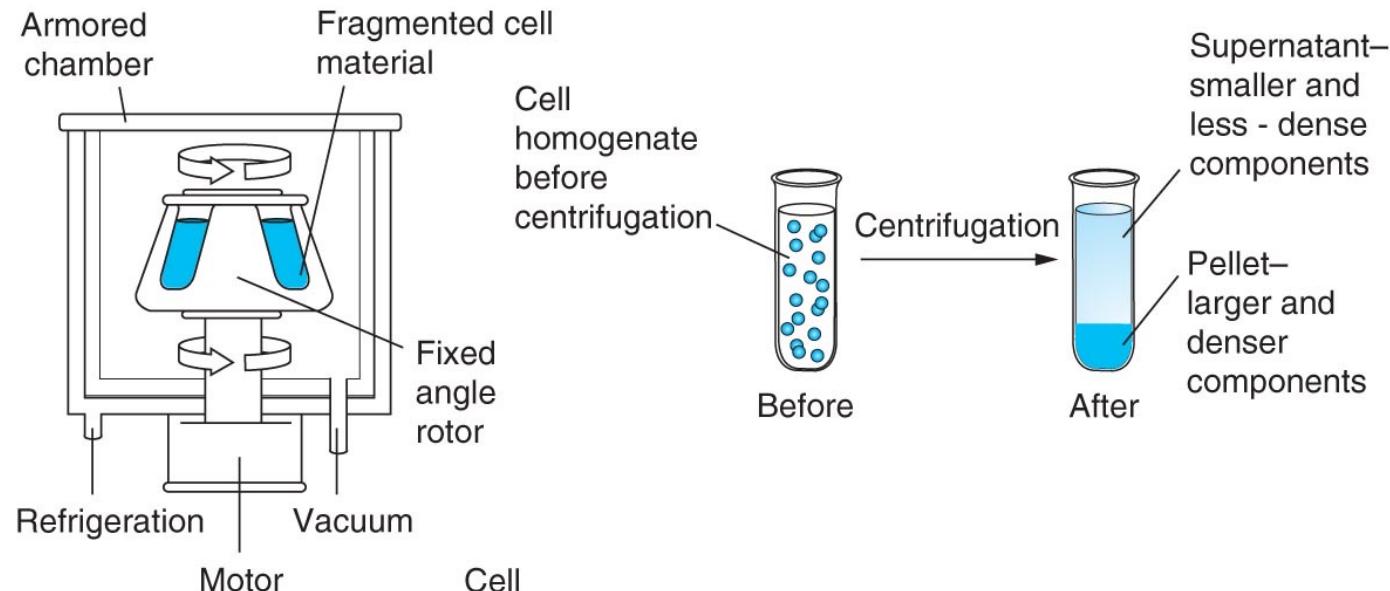
- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Methods used for protein separation
      - **Protein precipitation-** proteins quite often have hydrophilic amino acids on their surface that attract and interact with water molecules
        - » This characteristic is used as basis for separating proteins from other substances in extract
        - » Salts (ammonium sulfate) are added to mix to precipitate the proteins
        - » Solvents such as ethanol, isopropanol, acetone, diethyl ether can also cause protein precipitation by removing water from between the protein molecules

## 4.3 Protein production

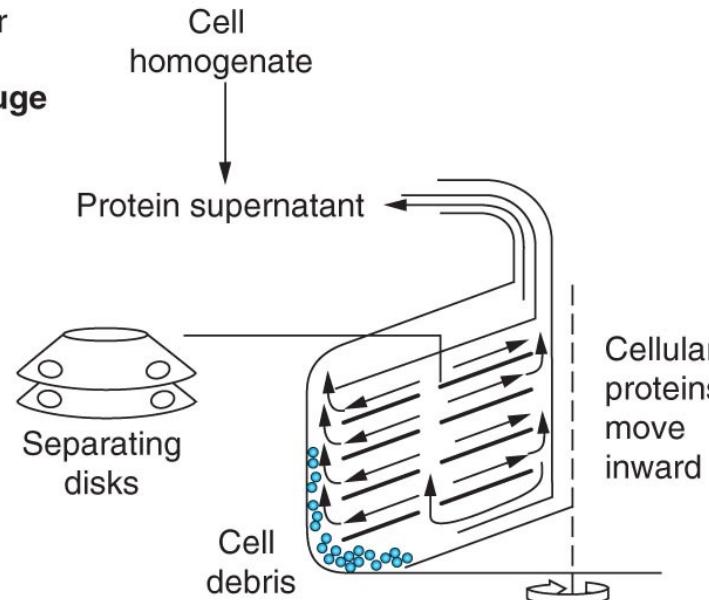
- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Methods used for protein separation
      - **Size based filtration separation method**
        - » Centrifugation - separates samples by spinning them at high speed
        - » Proteins can be isolated in single layer or separated from heavier cell components

## 4.3 Protein production

(a) Small - volume fixed angle centrifuge



(b) Batch centrifuge



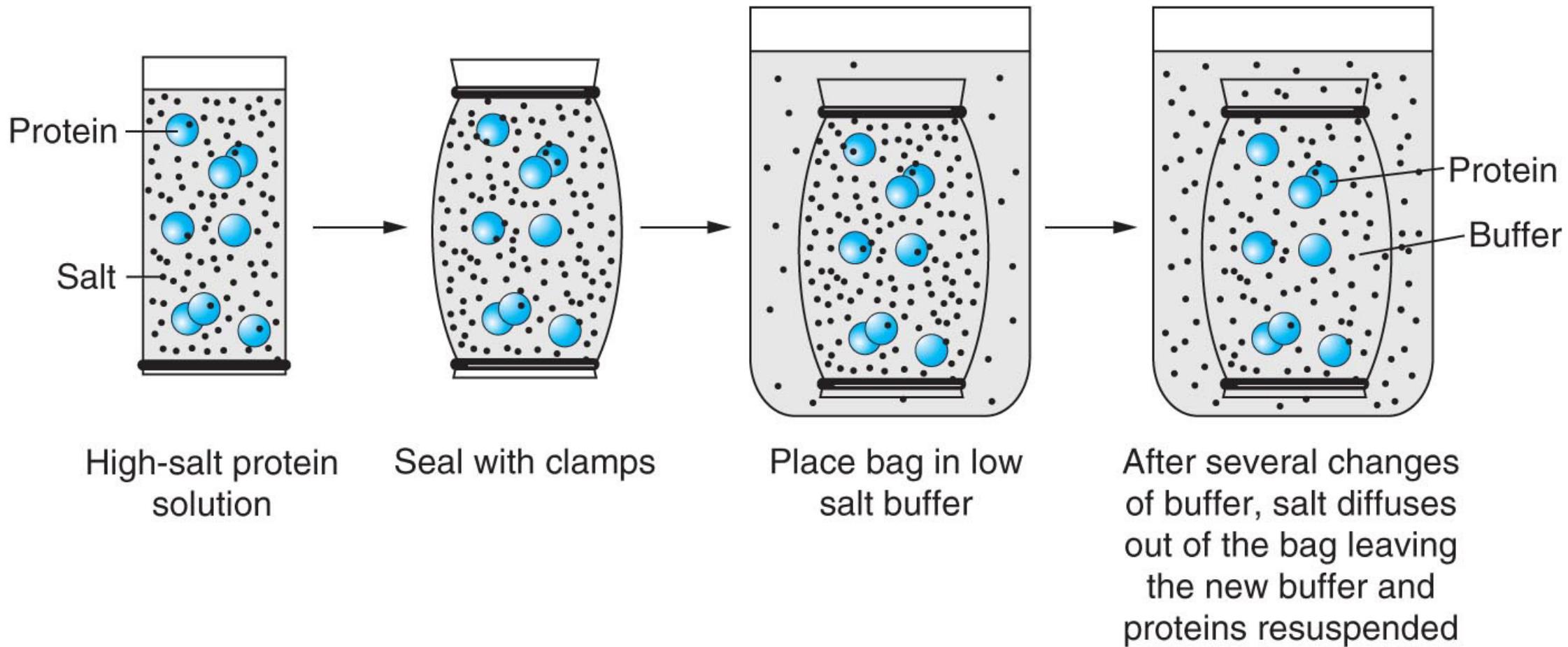
## 4.3 Protein production

- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Methods used for protein separation
      - **Size based filtration separation method**
        - » Membrane filtration - thin membranes of nylon or other engineered substances with varying pore sizes are used to filter out all cellular debris
          - \*\*\* This is a faster method than centrifugation
            - a. microfiltration removes precipitates and bacteria
            - b. ultrafiltration uses filters that catch molecules such as proteins and nucleic acids
          - \* some ultrafiltration processes can separate large proteins from small proteins
        - » Problem of membrane filtration: easy to clog

## 4.3 Protein production

- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Methods used for protein separation
      - **Size based filtration separation method**
        - » Diafiltration and dialysis - rely on chemical concept of equilibrium (migration of dissolved substances from areas of higher concentration to areas of lower concentration)
          - a. dialysis - some molecules can pass through semi-permeable membrane and others can't due to large size
            - dialysis is used to remove smaller salts, solvents, and other additives used in earlier process of protein purification
            - salts are replaced with buffering agents to stabilize proteins during rest of protein purification process

## 4.3 Protein production



## 4.3 Protein production

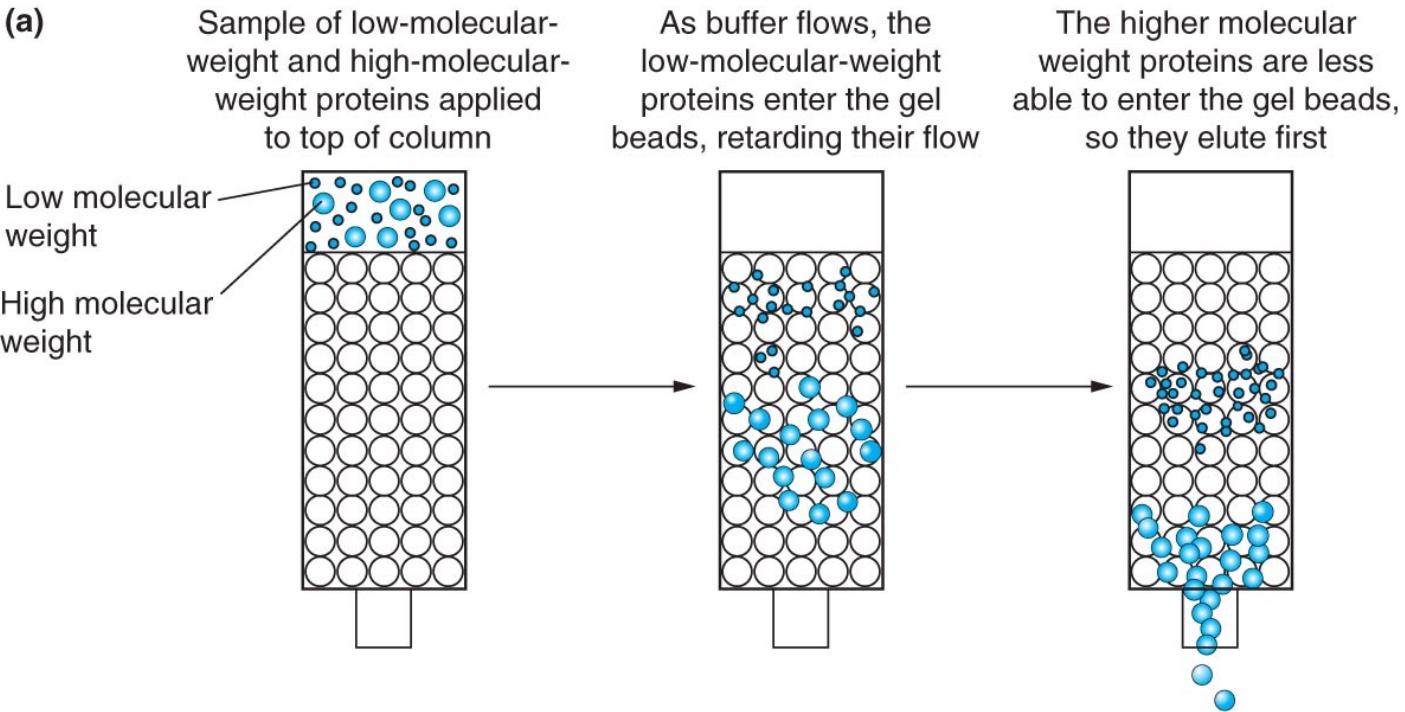
- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Methods used for protein separation
    - **Chromatography** - way to sort proteins by size or by how they bind to or separate from other substances
      - A. Long glass tubes are filled with resin beads and buffer
      - B. Then protein extract is added and flows through resin beads
      - C. Depending on resin used proteins either stick to beads or passes through glass column while beads act as filtration system

## 4.3 Protein production

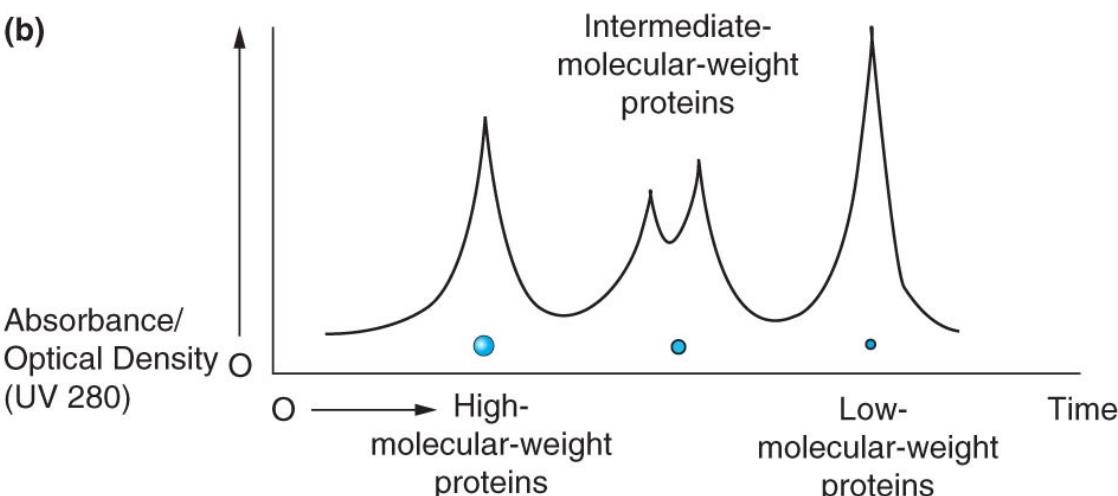
- Downstream processing methods:
  - Protein purification
    - Step 3: separating components in extract
      - Methods used for protein separation
      - **Size exclusion chromatography** - uses gel beads as filtering system so large protein molecules work their way around gel beads and small protein molecules pass through more slowly because they are able to enter the gel beads retarding their flow
        - Gels are available in many pore sizes and correct gel for separation depends on molec. weight of proteins and contaminants being separated
        - Method requires large volumes
        - Method makes preliminary separations

## 4.3 Protein production

(a)



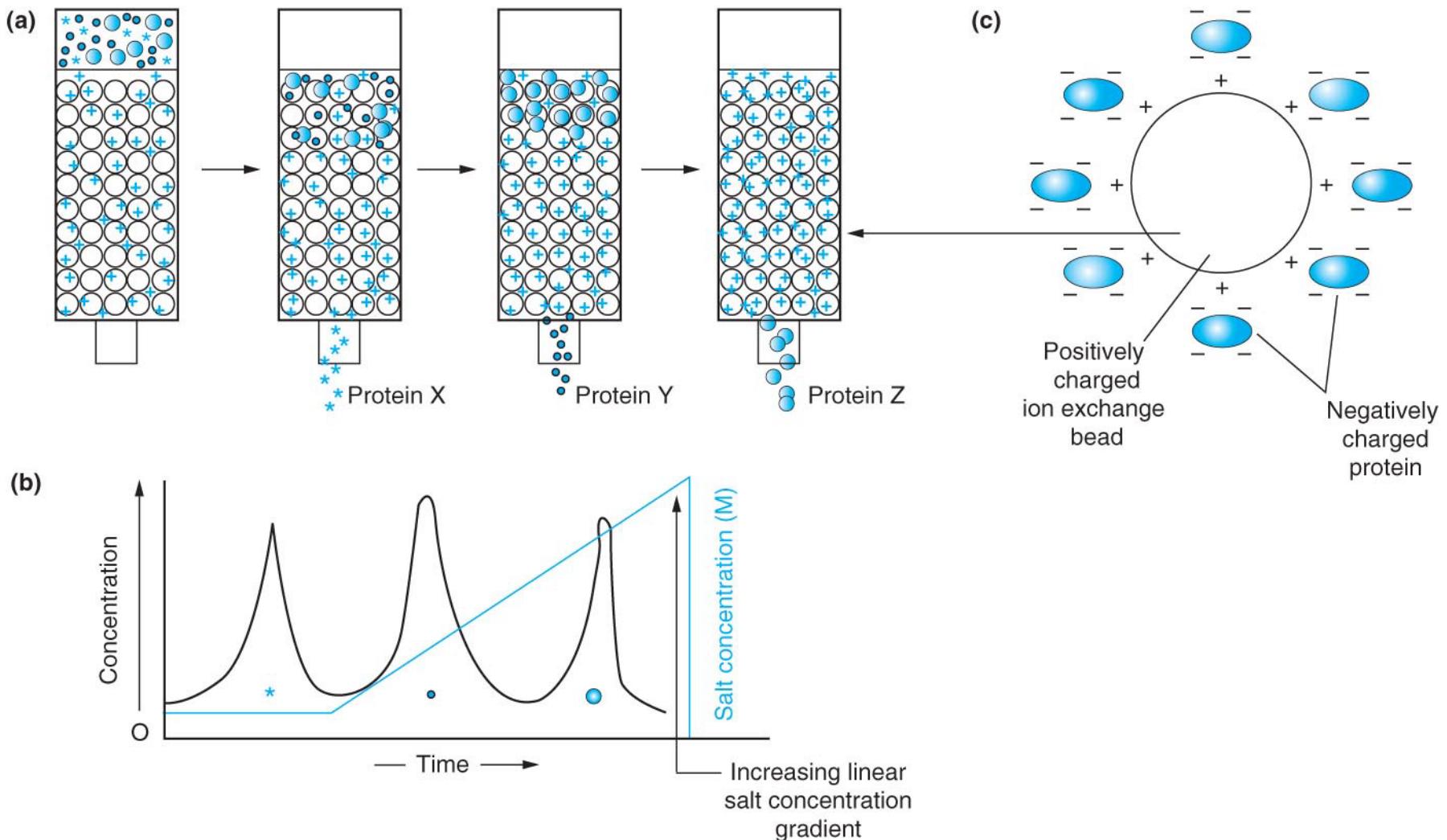
(b)



## 4.3 Protein production

- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Methods used for protein separation
    - **Ion exchange chromatography-** relies on electrostatic charge to bind proteins to resin beads in the column
      - When charged proteins bind to resin then the contaminants pass through and out of the column
      - Proteins can then be released from column by changing electrostatic charge by increasing salt concentrations of salt solutions
        - » Anion exchange resin is positively charged
        - » Cation exchange resin is negatively charged

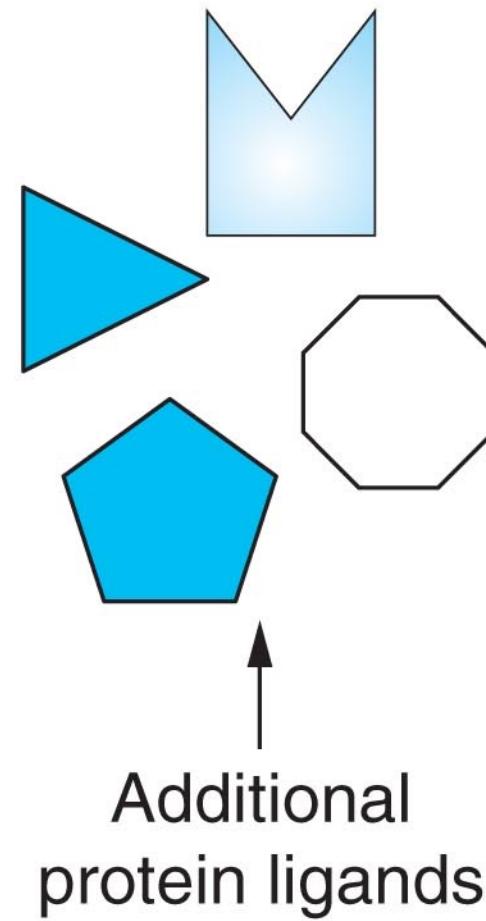
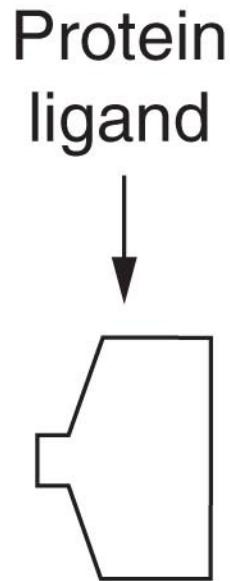
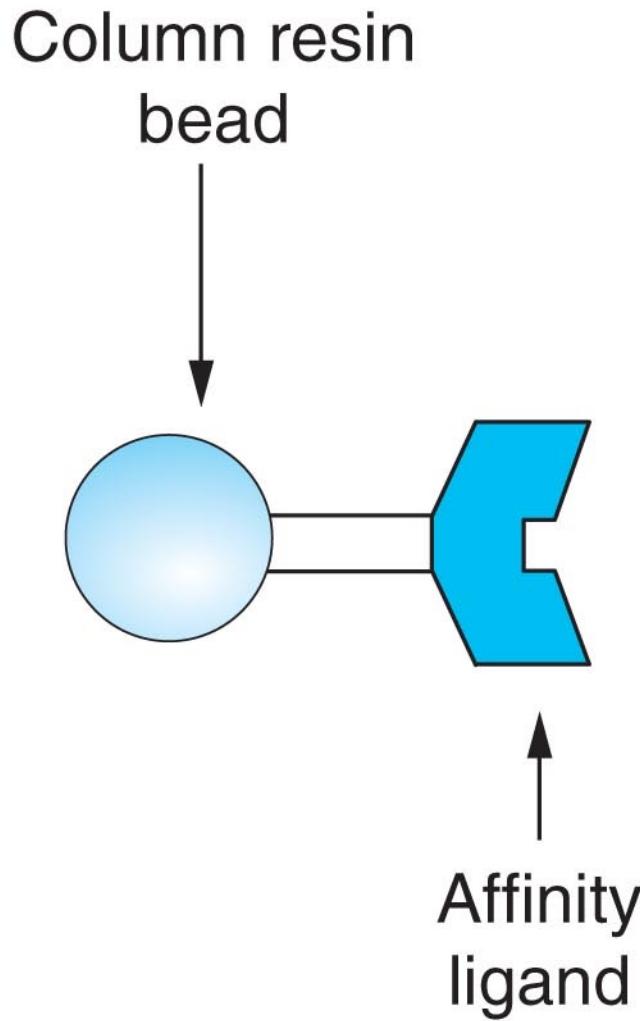
## 4.3 Protein production



## 4.3 Protein production

- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Methods used for protein separation
    - **Affinity chromatography-** relies on ability of most proteins to bind specifically and reversibly to ligands- small molecules that bind to a particular large molecule in a protein (lock: key model)
      - After proteins have bound to resin, buffer solution is used to wash out unbound molecules
      - Then special buffers are used to break ligand bonds of the retained protein
        - » Quite often used for purification of fusion proteins
    - \* this technique shortens purification process by reducing number of steps

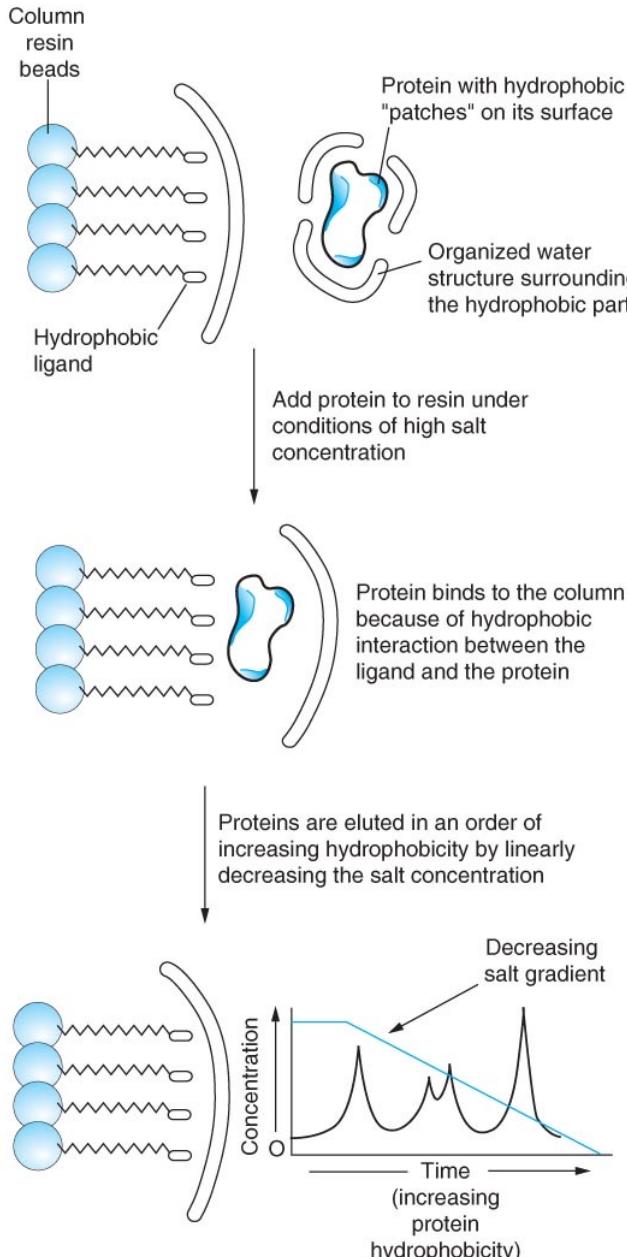
## 4.3 Protein production



## 4.3 Protein production

- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Methods used for protein separation
    - **Hydrophobic interaction chromatography**
      - Proteins are sorted on basis of their repulsion of water
        - » Column beads are coated with hydrophobic molecules and the hydrophobic amino acids in a protein are attracted to similar chemicals in the beads

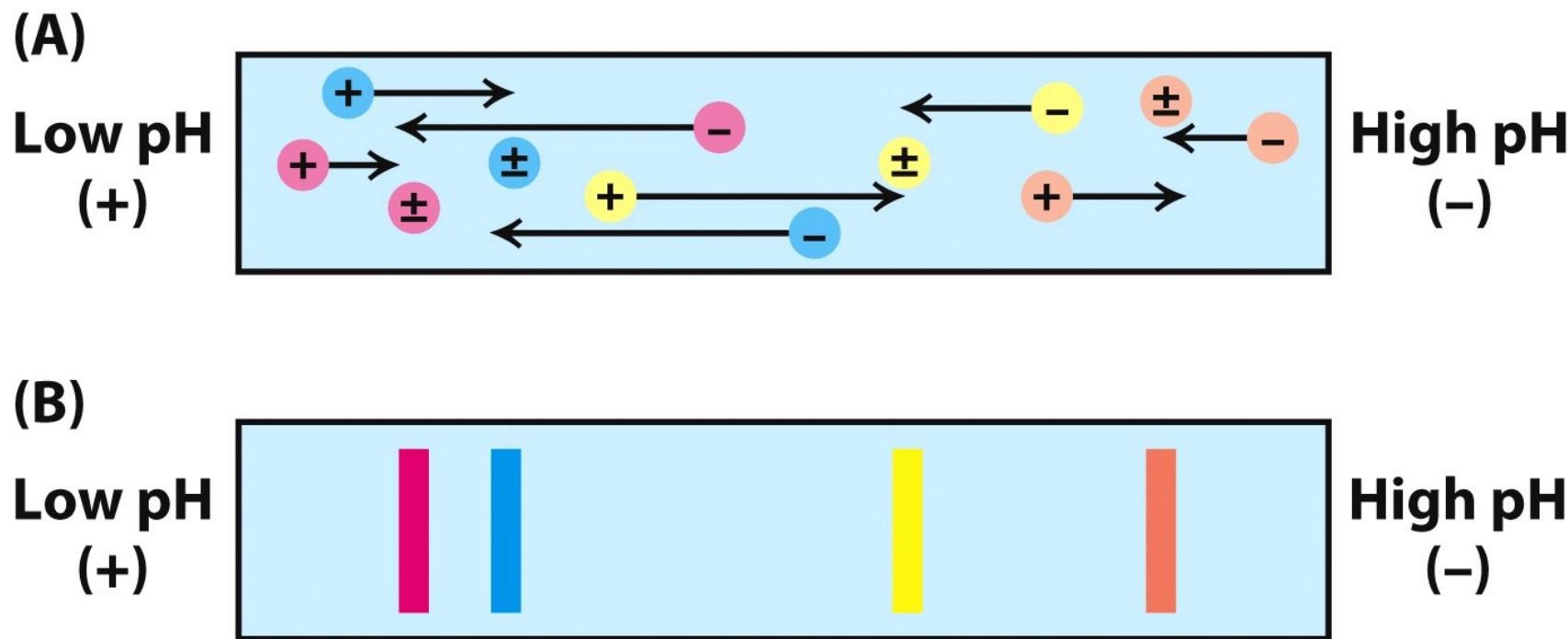
## 4.3 Protein production



## 4.3 Protein production

- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Methods used for protein separation
    - **Isoelectric focusing**
      - Each protein has a specific number of charged amino acids on its surface in specific places so due to this unique combination of charged groups each protein has a unique electronic signature = **ISOELECTRIC POINT**
        - » Isoelectric point = where charges on protein match pH of solution
        - » Used to separate similar proteins from each other
        - » Is first dimension of 2 dimensional gel electrophoresis to separate proteins based on electrical charge and size

# Isoelectric focusing

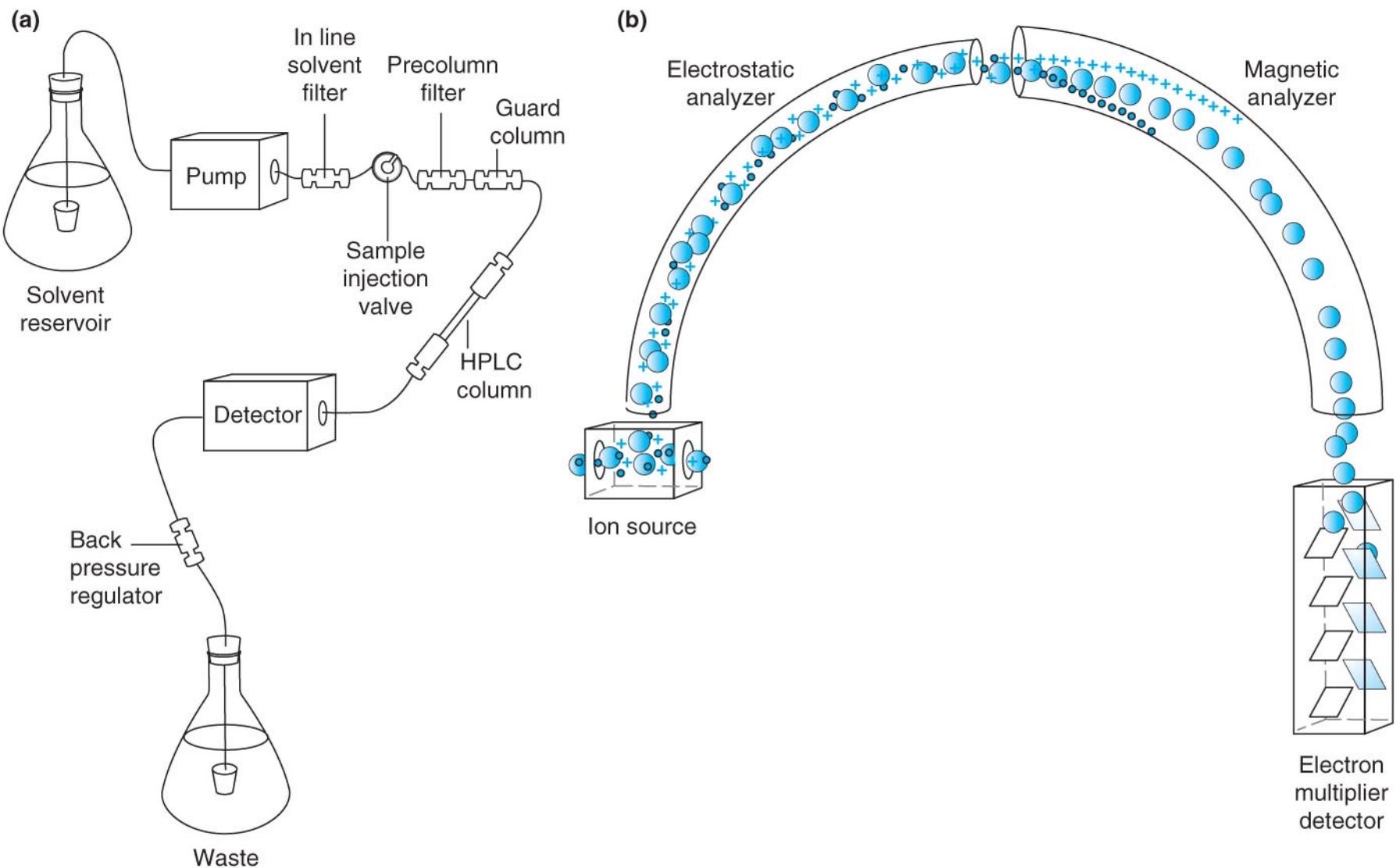


**Figure 3.11**  
*Biochemistry, Seventh Edition*  
© 2012 W. H. Freeman and Company

## 4.3 Protein production

- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Analytical Methods used for protein separation
        - » **High performance liquid chromatography (HPLC)**
        - » Use greater pressure to force the extract through the column in shorter time than other chromatography methods
        - » Limitations: less protein is separated
        - » Useful for analytical labs instead of mass production settings

## 4.3 Protein production



## 4.3 Protein production

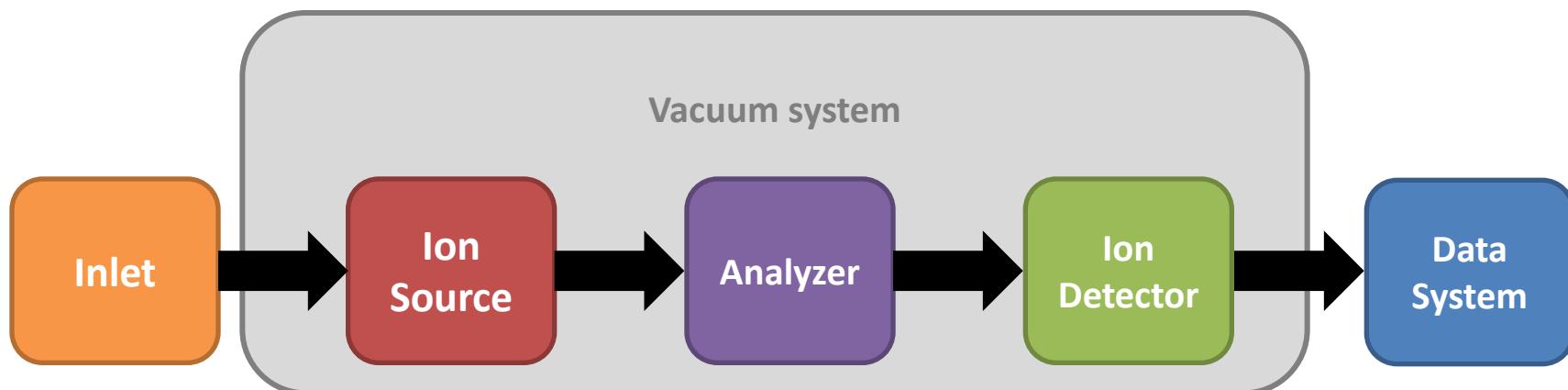
- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Analytical Methods used for protein separation
        - » **Mass spectrometry**- sensitive method used to identify small differences between proteins
        - » Used on the outflow of HPLC systems
        - » Protein sequencing - to do this technique first the protein is digested into smaller peptide fragments and analyzed via mass spec to determine the amino acid sequence

## 4.3 Protein production

- **Downstream processing methods:**
  - Protein purification
    - Step 3: separating components in extract
      - Analytical Methods used for protein separation
        - » **Mass spectrometry**- sensitive method used to identify small differences between proteins
        - » Used on the outflow of HPLC systems
        - » Protein sequencing - to do this technique first the protein is digested into smaller peptide fragments and analyzed via mass spec to determine the amino acid sequence

# Mass spectrometry

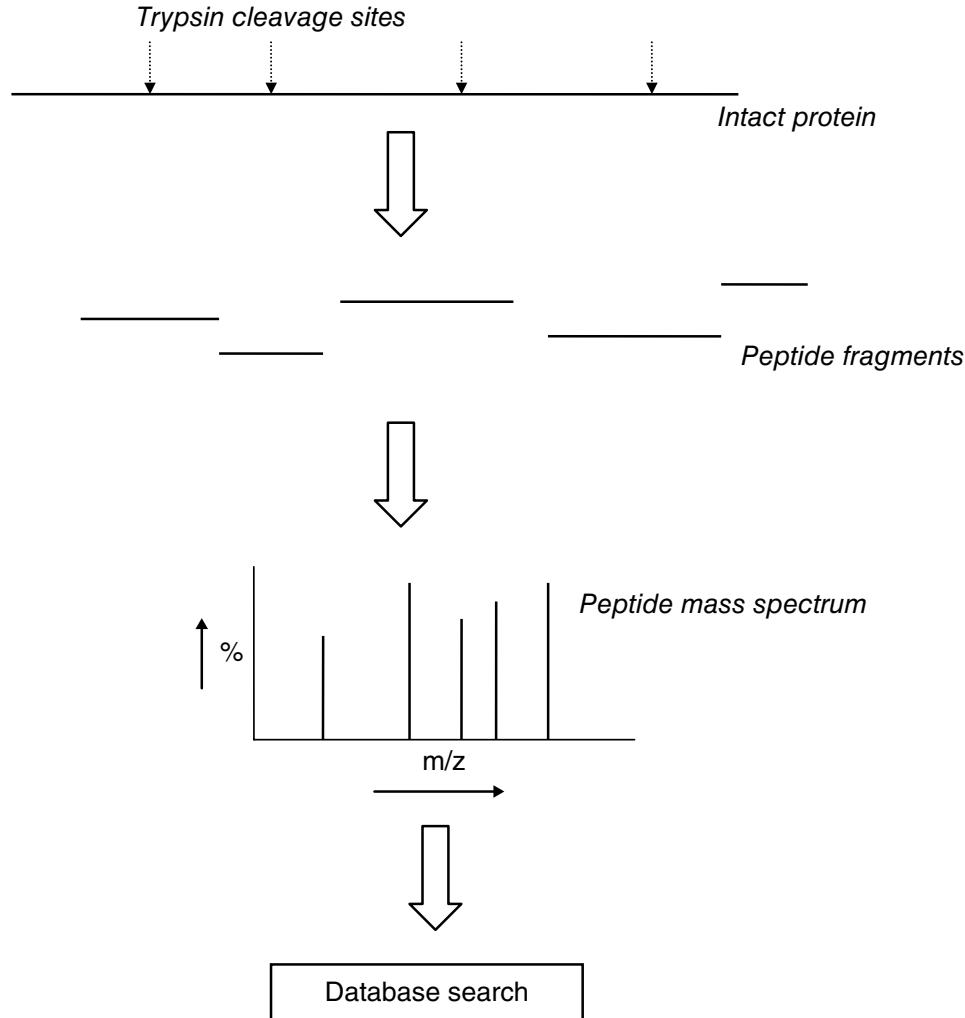
- MS is the analytical technique most intimately associated with proteomics.
- MS separates a mixture of (vaporized and ionized) analytes on the basis of their mass to charge ratio



# Mass spectrometry

- MS is used to:
  1. Determine protein mass
  2. Generate partial or full amino acid sequence data for a protein
  3. Quantify the amount of protein present in a sample
  4. Detect and identify protein PTMs
  5. Detect protein modification such as oxidation, deamidation or proteolysis
  6. Provide some information on protein structural detail.

# MS-based protein identification



## 4.3 Protein production

- **Verification**
  - important to verify that target protein is not lost during each stage of protein purification
    - **SDS PAGE (polyacrylamide gel electrophoresis)**
      - Detergent called SDS (sodium dodecyl sulfate) is added to protein mixture and mixture is heated
      - Sulfate charges are evenly distributed along denatured protein making separation dependent on size and not charge of protein
      - After treatment with SDS protein sample is loaded onto PAGE where it forms a single band at a specific location on the gel depending on its molecular size and mass

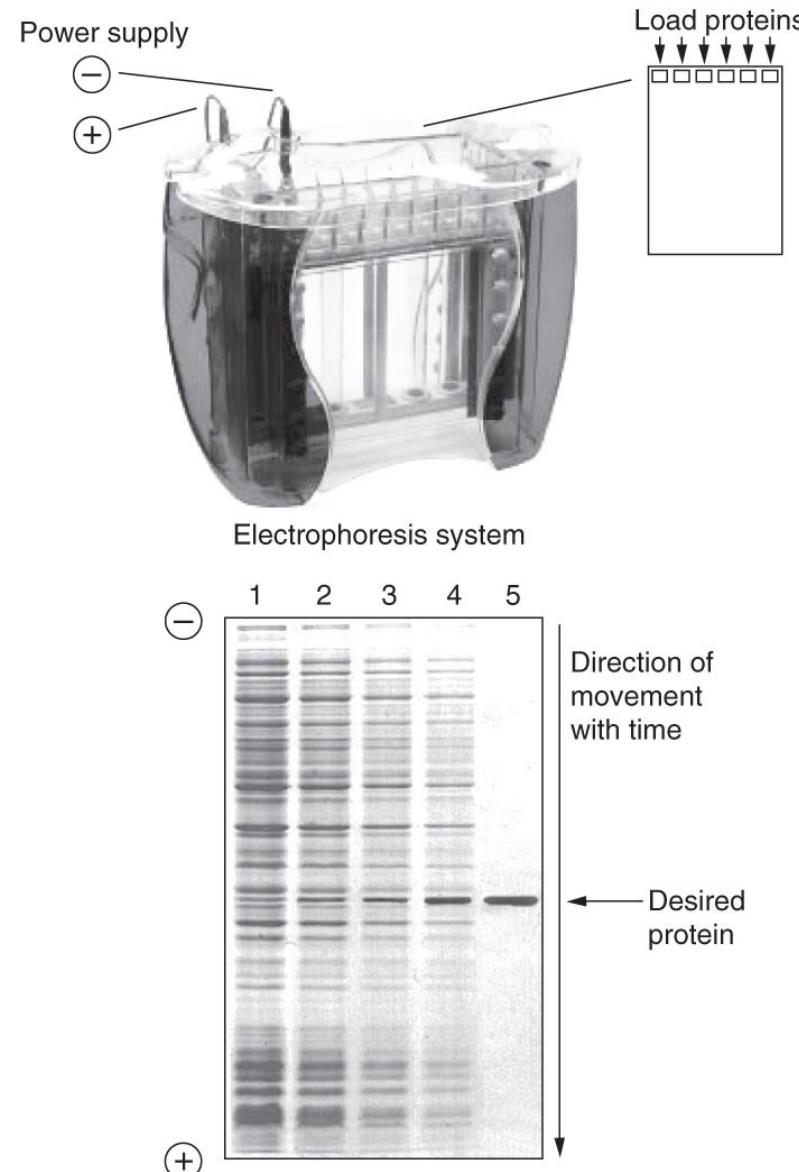
## 4.3 Protein production

- **Verification**
  - important to verify that target protein is not lost during each stage of protein purification
    - **SDS PAGE (polyacrylamide gel electrophoresis)**
    - How do you visualize the proteins after running the SDS/PAGE?
      - Coomassie stain - dye that combines with proteins so a colored band results
        - » Compare stained sample with known size marker to estimate size of protein
        - » During purification process the colored band representing the protein should become increasingly intense showing that the protein is not lost during the procedure

## 4.3 Protein production

- Based on the picture, knowing that lane 1 represents the protein marker, which lane represents the most concentrated and pure protein from the many purification steps? Explain your answer.
- Based on the photo of the gel, how many purification steps took place? In your answer mention some of the possible procedures that were used to ultimately purify the protein.

Work in groups of two to answer both questions.



## 4.3 Protein production

- **Verification**
- Specific method for proteins separated by SDS/PAGE= **Western blotting**
  - Procedure:
    - a. Proteins are transferred from the SDS/PAGE to a nitrocellulose membrane using electrical current
    - b. All sites on membrane are blocked so that the antibody will not bind to them nonspecifically
    - c. Primary antibody is added and incubated with the membrane
      - If there are any antibodies present that are directed against one or more of the blotted antigens (proteins on the membrane) those antibodies will bind to the protein and the other antibodies will be washed away at end of incubation
    - d. To detect antibodies that have bound to protein a secondary anti-immunoglobulin antibody coupled to a reporter group (alkaline phosphatase) are added to the membrane
    - e. Excess secondary antibody is washed off the membrane
    - f. Substrate is added that precipitates upon reaction with conjugate resulting in visible band where primary antibody is bound to the protein

## 4.3 Protein production

- **Verification**
- **Enzyme linked immunosorbent assay (ELISA) method**
  - Used to detect specific proteins
  - Needs 2 antibodies
  - Procedure takes place on a multiwell plate using affinity chromatography
    - a. Antibody 1 to capture the unique protein is plated on multiwell ELISA plate and then the protein is added to the plate and then there are several washes and blocking steps
    - b. Antibody 2 attached to an enzyme is added and addition of substrate allows colored reaction to occur if antibody has bound to the protein

## 4.3 Protein production

- **Preserving Proteins** - once protein of interest is isolated, collected and purified it must be saved to preserve its activity
- Ways to preserve protein activity
  - **Lyophilization (freeze drying)**
    - Maintains protein structure
    - Many freeze dried proteins can be stored at room temp for long periods of time

## 4.3 Protein production

- **Post purification analysis methods**
  - Protein sequencing - determine the protein's primary amino acid sequence
    - Mass spec method is used in which peptide masses are identified by their unique signatures
      - it is possible to identify many unique proteins in just a short period of time
  - X-ray crystallography - used to determine the complex tertiary and quaternary structures of proteins

## 4.4 Proteomics

- **Proteomics** - new scientific discipline dedicated to understanding complex relationship of disease and protein expression
  - Proteomes are compared between healthy and diseased states
    - Variations of protein expression are correlated to onset or progression of a disease
  - Goal is to discover protein markers to be used in diagnostic methods and development of targeted drugs

## 4.4 Proteomics

### Techniques used for proteomics

1. Two dimensional gel electrophoresis separates proteins followed by mass spec to verify the protein's identify
  - Use amino acid sequencing to characterize the protein
2. Protein microarray- set of proteins immobilized on a surface coated with a reagent that will indicate binding by color change
  - Functional protein microarrays have recently been applied to discovery of protein interactions; protein-protein; protein-lipid; protein-DNA; protein-drug; and protein-protein interactions

# Two-dimensional electrophoresis

