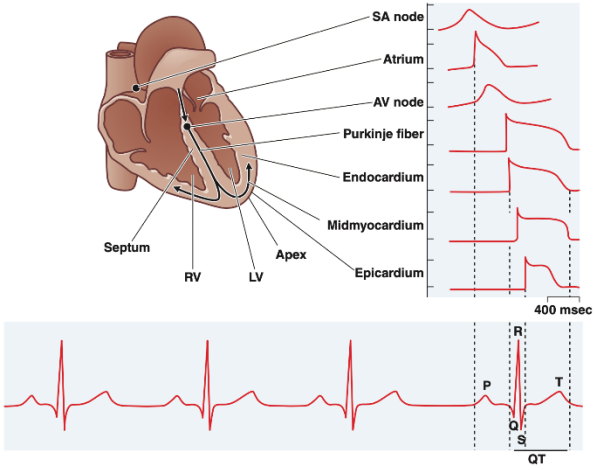
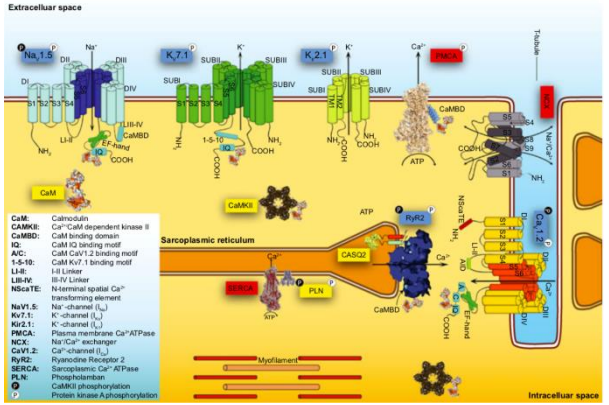
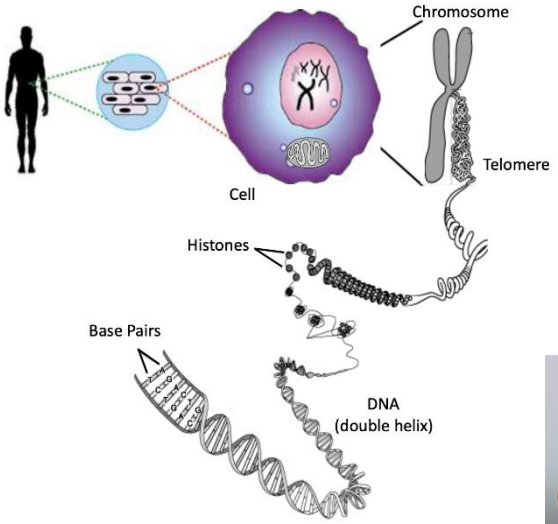
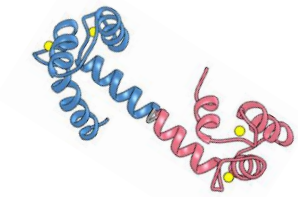
Nucleic Acids and Molecular Biology *Part II*

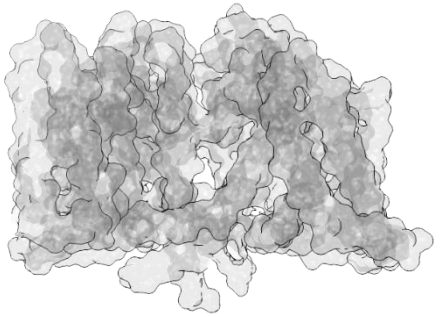
Biochemistry Boot Camp 2023 !





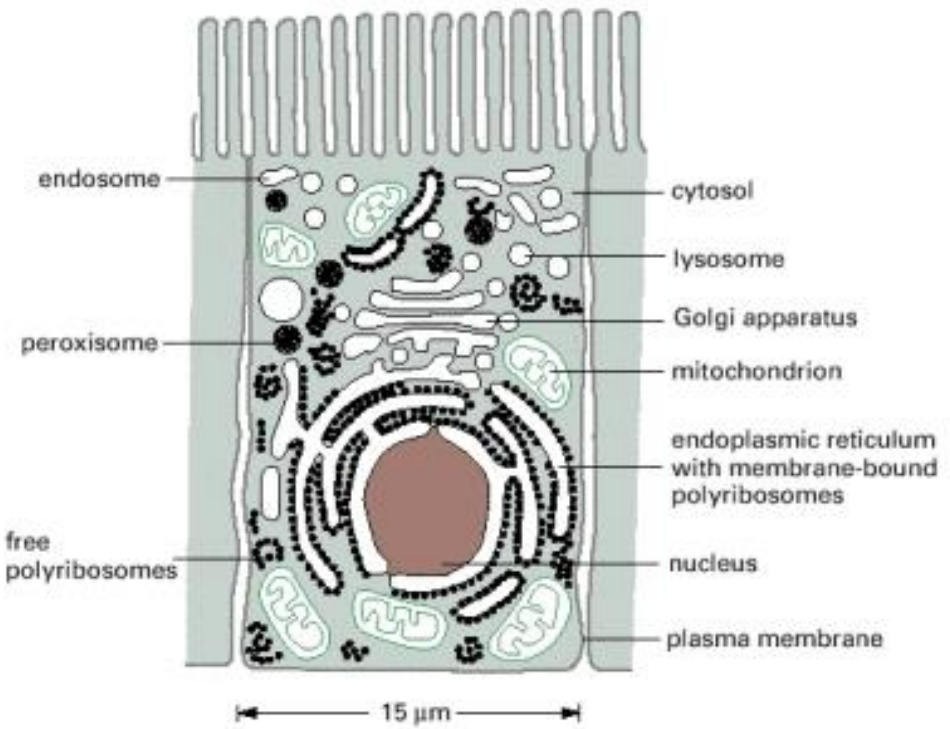






Review of Intro Biology

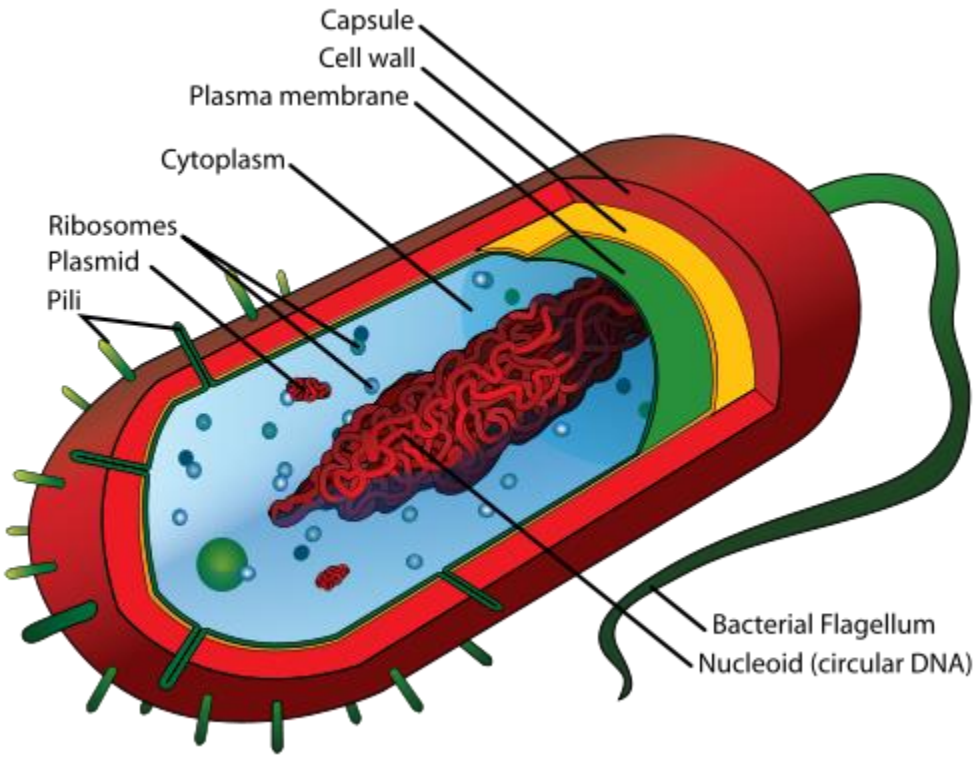
- Parts of a eukaryotic animal cell
- Has a nucleus where DNA is stored
- Membrane-bound organelles



Alberts, *et al*. *Molecular Biology of the Cell,* 4th Edition.

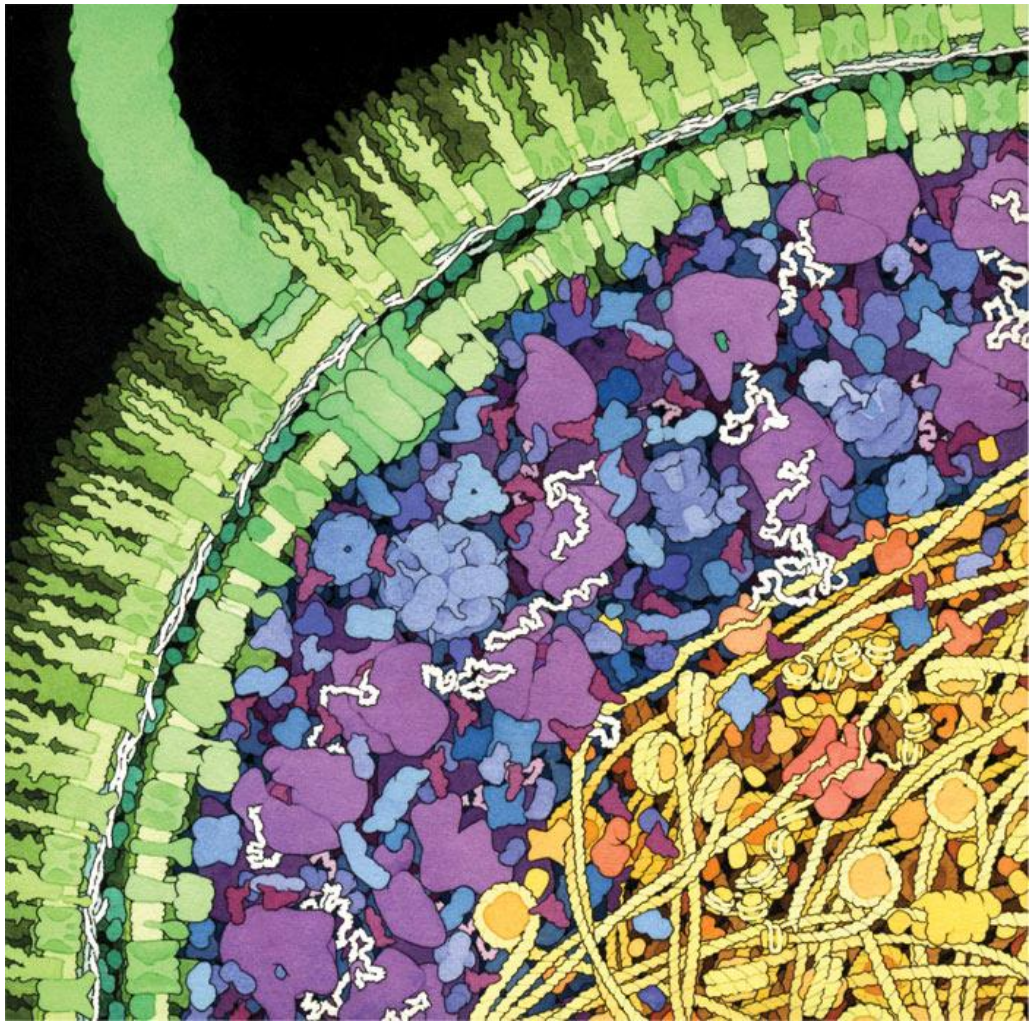
Review of Intro Biology

- Parts of a prokaryotic bacterial cell
- No nucleus: DNA is not linear but circular (no ends)
- No organelles, but ribosomes, etc. exist in the cytoplasm



Source: Wikipedia, "Bacterial Cell Structure."

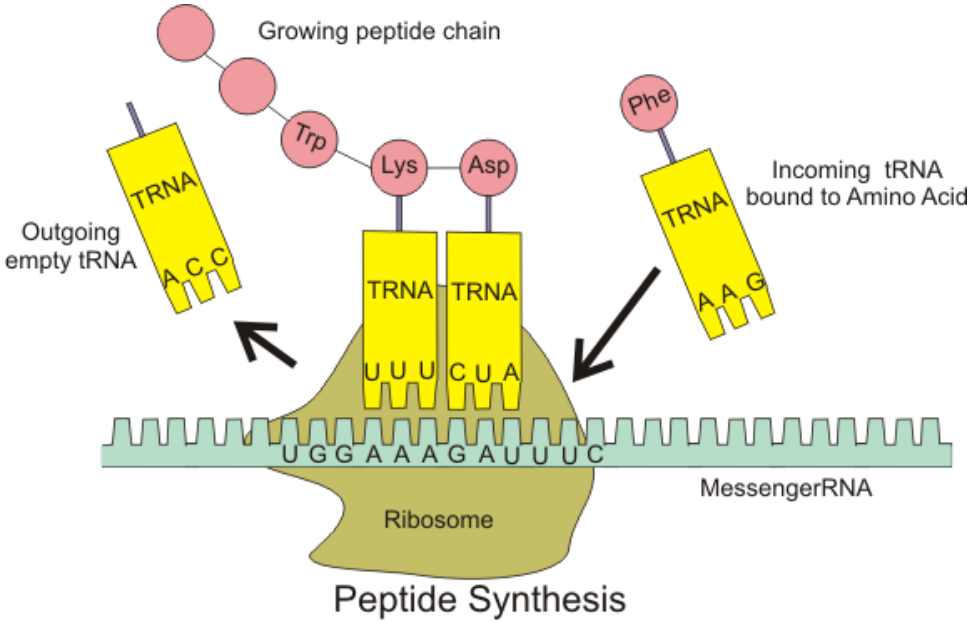
It’s Crowded in There!



Source: Goodsell, D. http://mgl.sripps.edu/people/goodsell/illustration/public/

Central Dogma

- DNA → mRNA "Transcription"
  - Synthesized RNA Polymerase
  - RNA formed from 5' to 3'
- mRNA → Protein "Translation"
  - Synthesized by ribosome
  - New proteins formed from NT to CT



Trick: Reading the DNA in the "standard way", one can easily identify the codons for peptide synthesis.

Source: Wikipedia, "Ribosome"

Genetic Code

nonpolar | polar | basic | acidic | (stop codon)

Standard genetic code

1st base	2nd base								3rd base
	U		C		A		G		
U	UUU	(Phe/F) Phenylalanine	UCU	(Ser/S) Serine	UAU	(Tyr/Y) Tyrosine	UGU	(Cys/C) Cysteine	U
	UUC		UCC		UAC		UGC		C
	UUA	(Leu/L) Leucine	UCA		UAA	Stop (*Ochre*)	UGA	Stop (*Opal*)	A
	UUG		UCG		UAG	Stop (*Amber*)	UGG	(Trp/W) Tryptophan	G
C	CUU		CCU	(Pro/P) Proline	CAU	(His/H) Histidine	CGU	(Arg/R) Arginine	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	(Gln/Q) Glutamine	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	(Ile/I) Isoleucine	ACU	(Thr/T) Threonine	AAU	(Asn/N) Asparagine	AGU	(Ser/S) Serine	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	(Lys/K) Lysine	AGA	(Arg/R) Arginine	A
	AUG[A]	(Met/M) Methionine	ACG		AAG		AGG		G
G	GUU	(Val/V) Valine	GCU	(Ala/A) Alanine	GAU	(Asp/D) Aspartic acid	GGU	(Gly/G) Glycine	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	(Glu/E) Glutamic acid	GGA		A
	GUG		GCG		GAG		GGG		G

Source: Wikipedia, “Genetic Code”

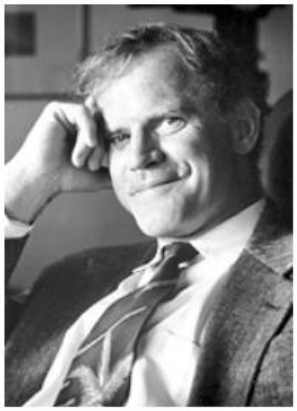
Why is this Useful?

- Site-Directed Mutagenesis

- Good Primers:
  - T_m > 78 °C (2 mM $MgCl_2$, 50 mM NaCl)
  - GC content > 40%
  - No secondary structure (< 50 bp)
  - End with G or C

The Nobel Prize in Chemistry 1993
Kary B. Mullis, Michael Smith

The Nobel Prize in Chemistry 1993
Nobel Prize Award Ceremony
Kary B. Mullis
Michael Smith



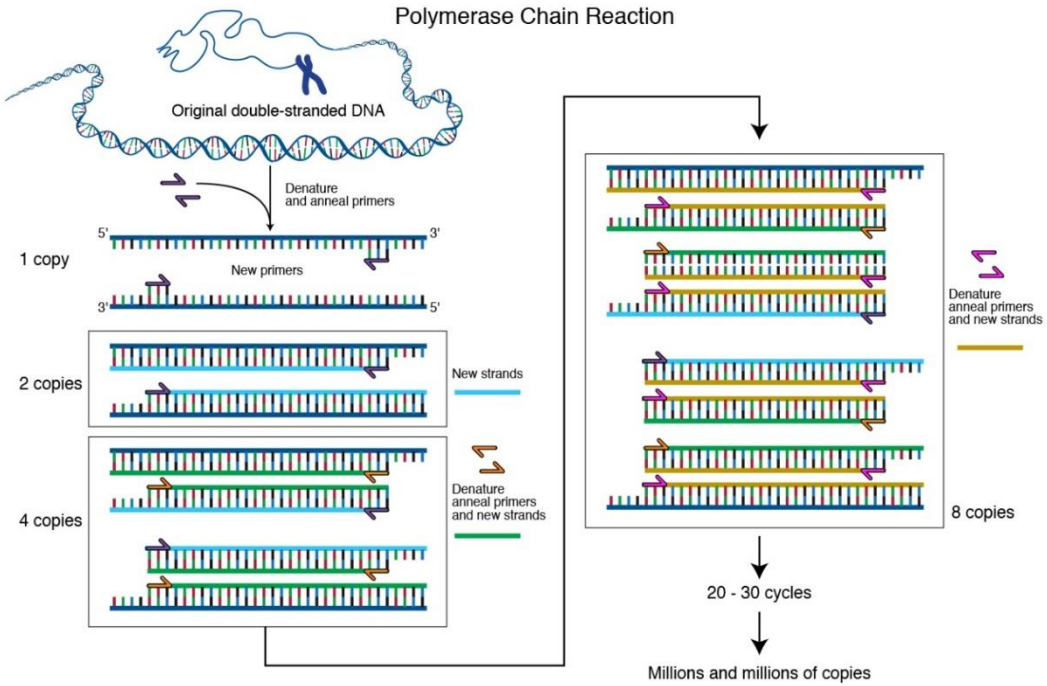
Kary B. Mullis Michael Smith

The Nobel Prize in Chemistry 1993 was awarded *"for contributions to the developments of methods within DNA-based chemistry"* jointly with one half to Kary B. Mullis *"for his invention of the polymerase chain reaction (PCR) method"* and with one half to Michael Smith *"for his fundamental contributions to the establishment of oligonucleotide-based, site-directed mutagenesis and its development for protein studies"*.

Polymerase Chain Reaction (PCR)

3 main steps

1. denature
2. anneal
3. polymerase activity



- **Key consideration:** Temperature for primer annealing (computational tools)!

https://www.genome.gov/genetics-glossary/Polymerase-Chain-Reaction

Example: Designing Primers

http://www.genomics.agilent.com/primerDesignProgram.jsp

described in QuikChange manuals, this program calculates/designs the appropriate primer sequences with the optimal melting temperature. Read Help for more information about the program.

Expand Help

1. It is recommended that you clear this form prior to loading each new sequence:

Clear Input

2. Select QuikChange® mutagenesis kit that you are using:

QuikChange® II

Help in choosing a mutagenesis kit

3. Find your DNA sequence by pressing

Load a file from your hard-drive:

No file selected Browse_

Or, paste plain text or FASTA-formatted DNA sequence in the box below:

atagatatagacagacatacatcagacatatacaggaggaggatacacag
acatcatacagcaga

4. Load it.

Upload Now - or - Upload Translated

or specify a DNA region to translate from to Upload Translated Region

5. Select up to seven nucleotides that you want to change

DNA: change nucleotide(s) to:	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7
	g	-*-	-*-	-*-	-*-	-*-	-*-

or

Delete a region between two checked nucleotides (Note: two checked nucleotides will not be deleted)

or

Insert between two checked nucleotides

1 a	2 t	3 a	4 g	5 a	6 t	7 a	8 t	9 a
10 g	11 a	12 c	13 a	14 g	15 a	16 c	17 a	18 t
19 a	20 c	21 a	22 t	23 c	24 a	25 g	26 a	27 c
28 a	29 t	30 a	31 t	32 a	33 c	34 a	35 g	36 g
37 a	38 g	39 g	40 a	41 g	42 g	43 a	44 t	45 a
46 c	47 a	48 c	49 a	50 g	51 a	52 c	53 a	54 t
55 c	56 a	57 t	58 a	59 c	60 a	61 g	62 c	63 a
64 g	65 a							

Finally, obtain your designed primer sequences.

Design Primers

Clear this form to load a new sequence:

Clear Input

or

Delete a region between two checked nucleotides (Note: two checked nucleotides will not be deleted)

or

Insert between two checked nucleotides

1 a	2 t	3 a	4 g	5 a	6 t	7 a	8 t	9 a
10 g	11 a	12 c	13 a	14 g	15 a	16 c	17 a	18 t
19 a	20 c	21 a	22 t	23 c	24 a	25 g	26 a	27 c
28 a	29 t	30 a	31 t	32 a	33 c	34 a	35 g	36 g
37 a	38 g	39 g	40 a	41 g	42 g	43 a	44 t	45 a
46 c	47 a	48 c	49 a	50 g	51 a	52 c	53 a	54 t
55 c	56 a	57 t	58 a	59 c	60 a	61 g	62 c	63 a
64 g	65 a							

Finally, obtain your designed primer sequences.

Design Primers

Clear this form to load a new sequence:

Clear Input

Primer sequences:

Primer Name	Primer Sequence (5' to 3')
	5'-atcctcctcctgtatatgtcggatgtatgtctgtctatatc-3'
a24c_	5'-gatatagacagacatacatccgacatatacaggaggaggat-3'

Oligonucleotide information:

Primer Name	Length (nt.)	Tm	Duplex Energy at 68 °C	Energy Cost of Mismatches
	41	78.60°C	-45.55 kcal/mole	3.05%
a24c_	41	78.60°C	-49.02 kcal/mole	3.83%

Primer-template duplexes:

Primer Name	Primer-Template Duplex
	gatatagacagacatacatcagacatatacaggaggaggataca
	||||||||||||||||||| |||||||||||||||||||
	3'-ctatatctgtctgtatgtaggctgtatatgtcctcctccta-5'
	5'-gatatagacagacatacatccgacatatacaggaggaggat-3'
a24c_	||||||||||||||||||| |||||||||||||||||||
	???ctatatctgtctgtatgtagtctgtatatgtcctcctcctatgt

General Primer Design Principles

- **PCR Steps:** Denature (95 °C), anneal (60 °C), extend (70 °C)

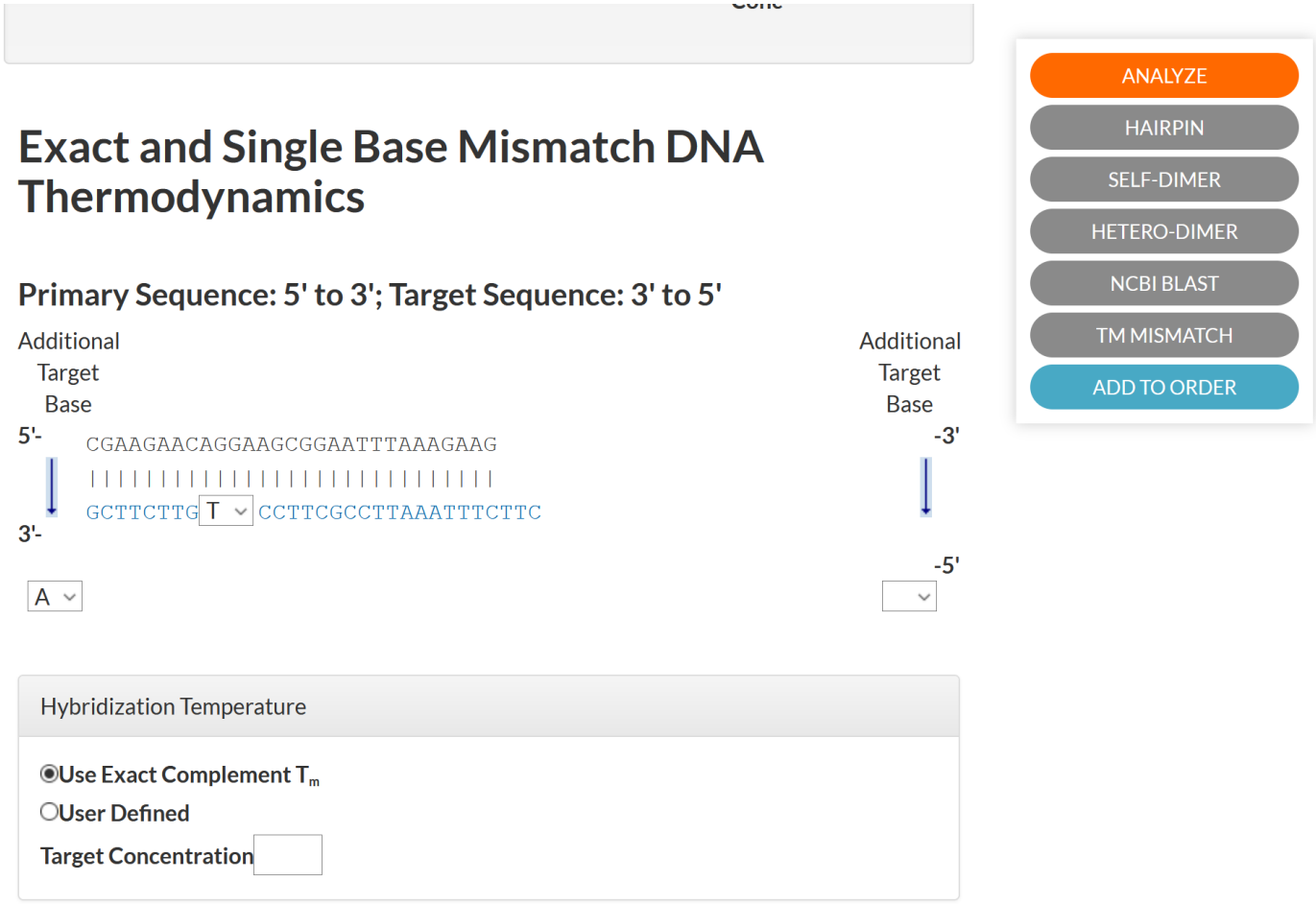
- Considerations:
  - **Melting Temperature:** Should be 52-58 °C
  - **GC Content:** 40-60%
  - **Length:** ~30 bp (but longer can be okay)
  - **Secondary Structure:** Avoid if possible

- Lots of software exists (some costs $$$). For more information (some trial and error here):

https://goo.gl/4EwMG3 (Life Technologies)

http://www.premierbiosoft.com/tech_notes/PCR_Primer_Design.html

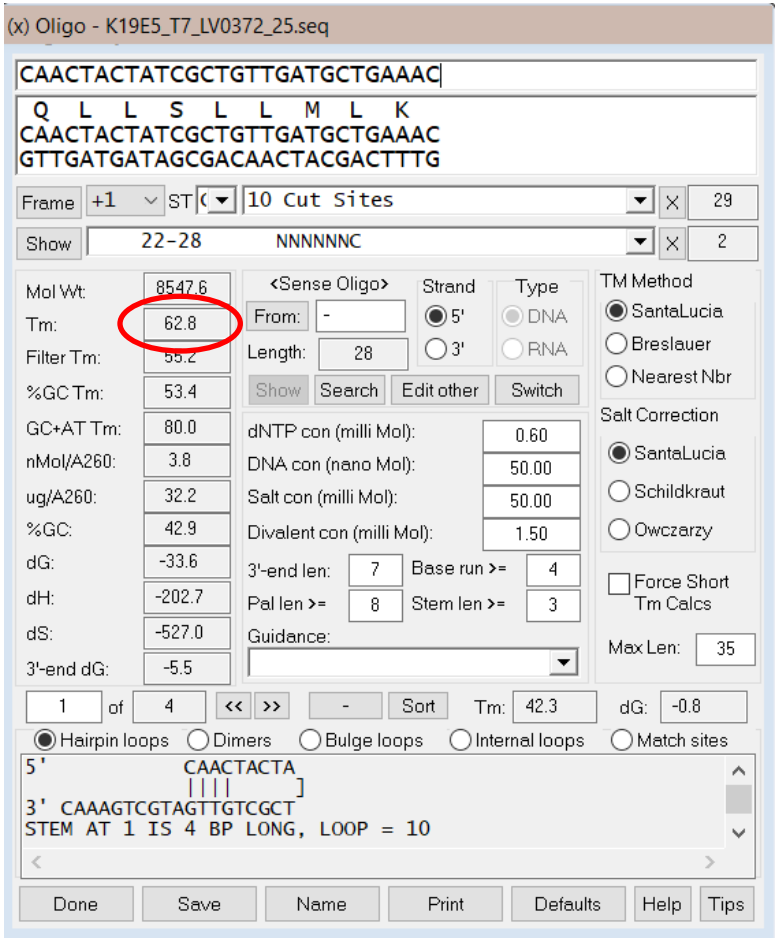
Example: General Approach



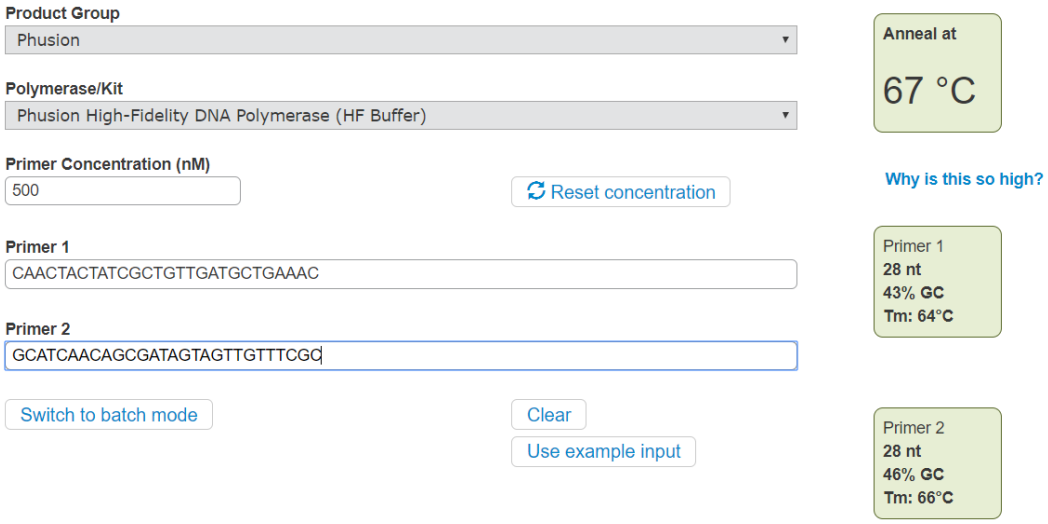
IDT OligoAnalyzer Mismatch Mode – estimate Tm cost of non-matching bases

Other Calculators: GeneRunner

- Download (free) from http://generunner.net/
- Analysis → Oligo brings up window to the right
- Suggested options shown to the right
  - SantaLucia temperature & salt recommended
  - Check your specific dNTP, DNA, Mg^{2+} and salt concentrations

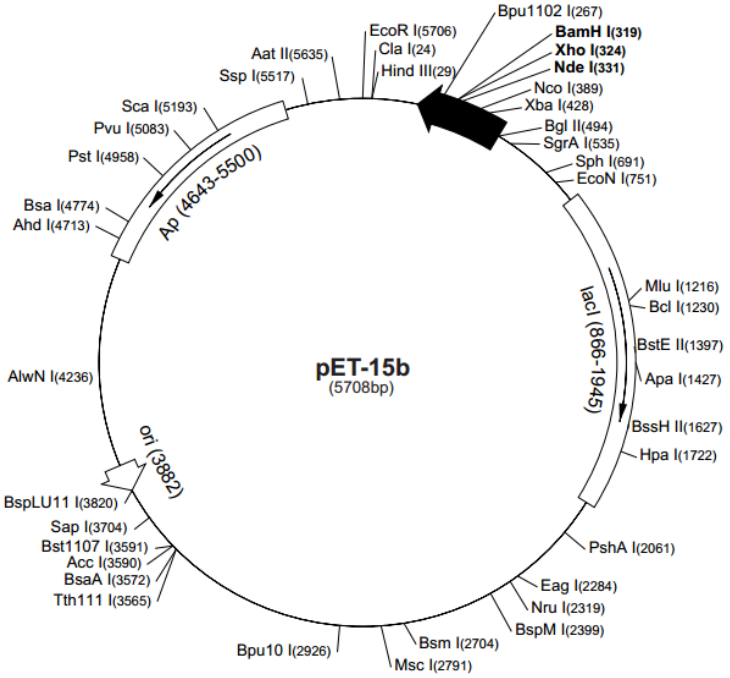


Other Calculators: Phusion Calculator



- Some polymerases have their own calculator specific to buffers, recommendations
- Phusion Polymerase found at NEB website (https://tmcalculator.neb.com/#!/main)

Protein and Expression Plasmid (pET-15b)

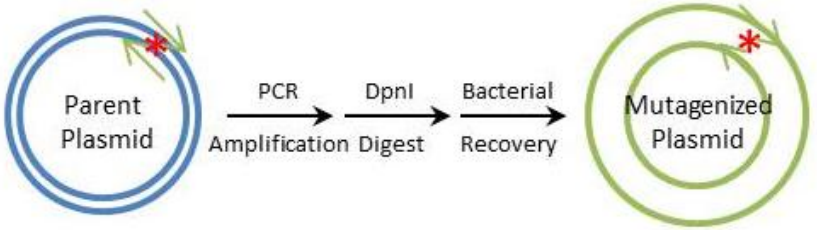


WT GB3 DNA/Protein Sequence:

```
(-6) L  *  E  G  D  I  H  M  Q  Y  K  L  V  I  N  G  K  T  L  K    13
1    ctttaagaaggagatatacatatgcagtacaaattagttatcaatggtaaaacattgaaa  60

14   G  E  T  T  T  K  A  V  D  A  E  T  A  E  K  A  F  K  Q  Y    33
61   ggcgaaacaactactaaagctgttgatgctgaaactgcagaaaaagctttcaaacaatac  120

34   A  N  D  N  G  V  D  G  V  W  T  Y  D  D  A  T  K  T  F  T    53
121  gctaacgacaacggtgttgacggtgtttggacttacgacgatgcgactaagacctttaca  180

54   V  T  E  *  D  P  A  A  N  K  A
181  gttactgaataggatccggctgctaacaaagcc  213
```

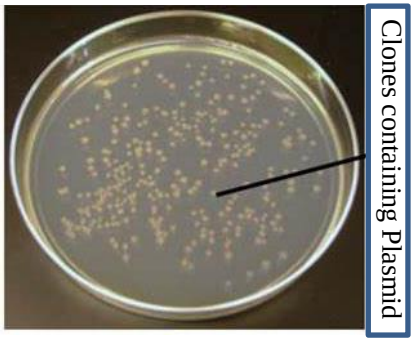
Successful Mutagenesis Primers

Primer Name	Primer Sequence (5' to 3')
K19I-Fwd	CAACTACTATCGCTGTTGATGCTGAAAC
K19I-Rev	GCATCAACAGCGATAGTAGTTGTTTCGC



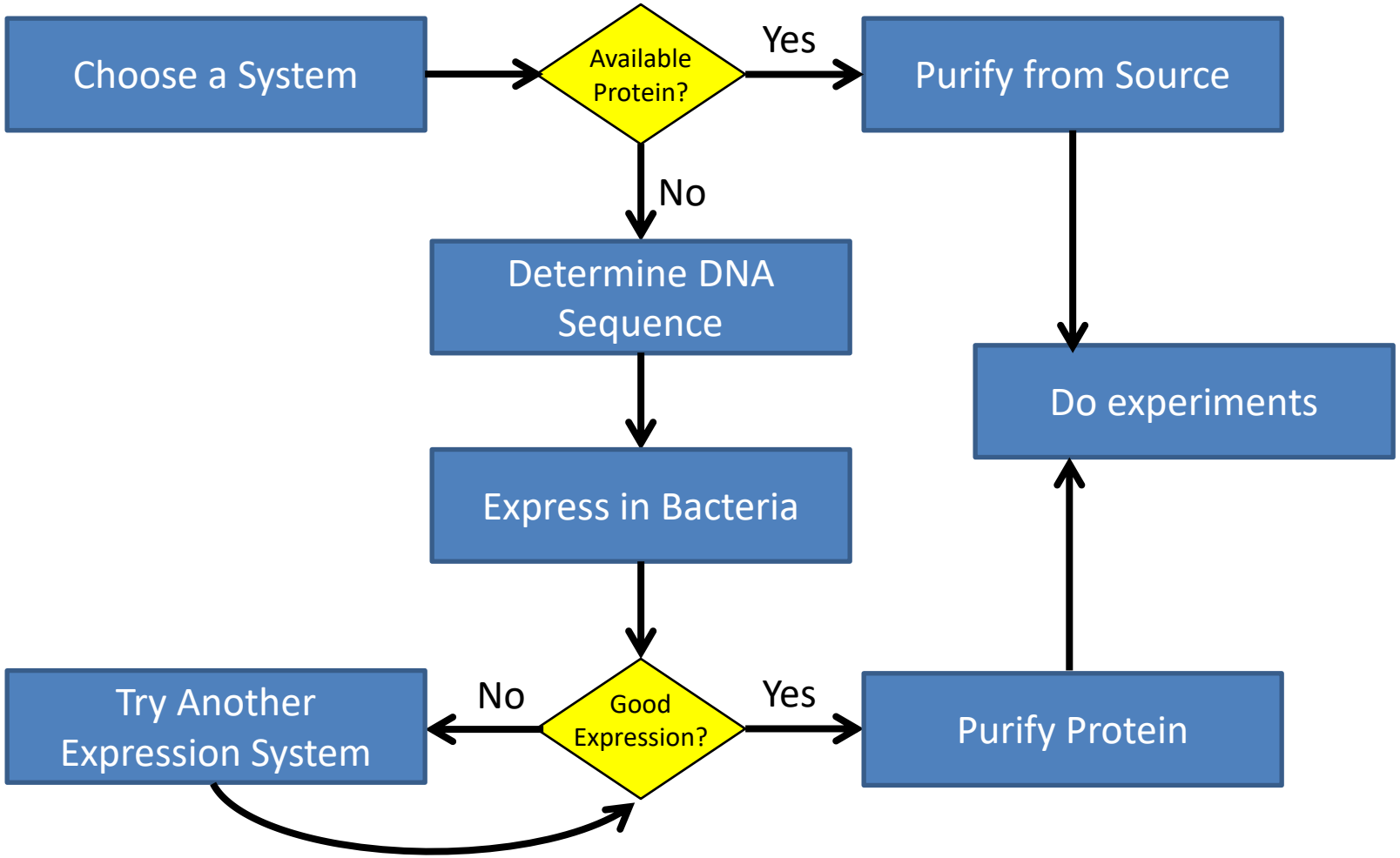
Description	Primer Duplex
K19I -Fwd/ Template	5′-CAACTACTATCGCTGTTGATGCTGAAAC ||||||||| |||||||||||||| 3′-TAACTTTCCGCTTTGTTGATGATTTCGACAACTACGACTTTGACGT
K19I -Rev/ Template	CATTGAAAGGCGAAACAACTACTAAAGCTGTTGATGCTGAAACTGCAGAAA ||||||||||||||| |||||||||| CGCTTTGTTGATGATAGCGACAACTACG
K19I -Fwd/ K19I -Rev	CAACTACTATCGCTGTTGATGCTGAAAC CGCTTTGTTGATGATAGCGACAACTACG

After Transformation:



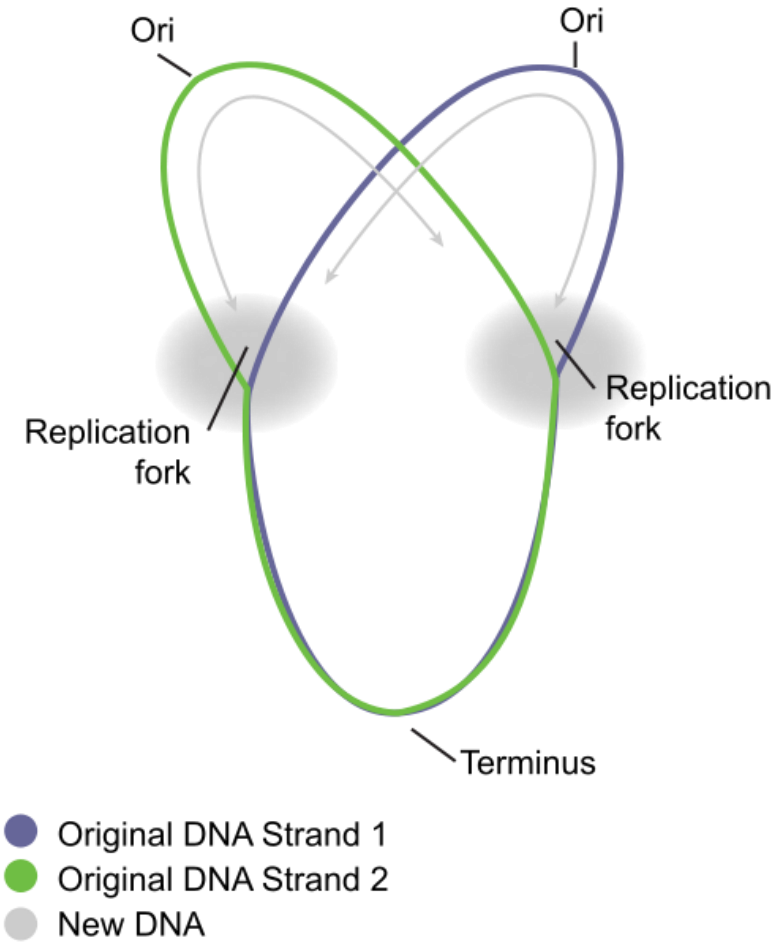
- Annealing temperature used: 53 °C

Recombinant Protein Production



Bacterial DNA: Features

- Chromosome is *circular*
- Replication starts at the *origin of replication* (Ori, TTATCCACA)
- **Plasmid:** *Any* circular DNA in the bacterial cell can be replicated if it has an Ori



Source: Wikipedia, "Circular Bacterial Chromosome"

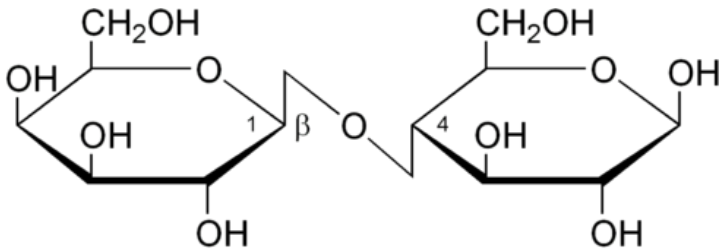
The Lactose (lac) Operon

- **Idea:** Bacteria only want to produce proteins if they are needed
- Why metabolize lactose (hard) when glucose (easy) is available?
- **Operon:** A set of genes (proteins) under the control of other genes in the cell

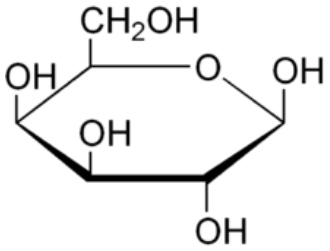
The Lactose (lac) Operon

Proteins:

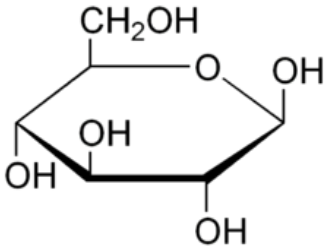
- **lacI** (lac repressor): binds at operator when no lac present; prevents binding of RNA polymerase at promoter
- **lacZ** (β-galactosidase): converts Lac in to Gal and Glc by hydrolyzing glycosidic linkage
- **lacY** (β-galactoside permease): Pumps Lac into the cell



Lactose



D-galactose

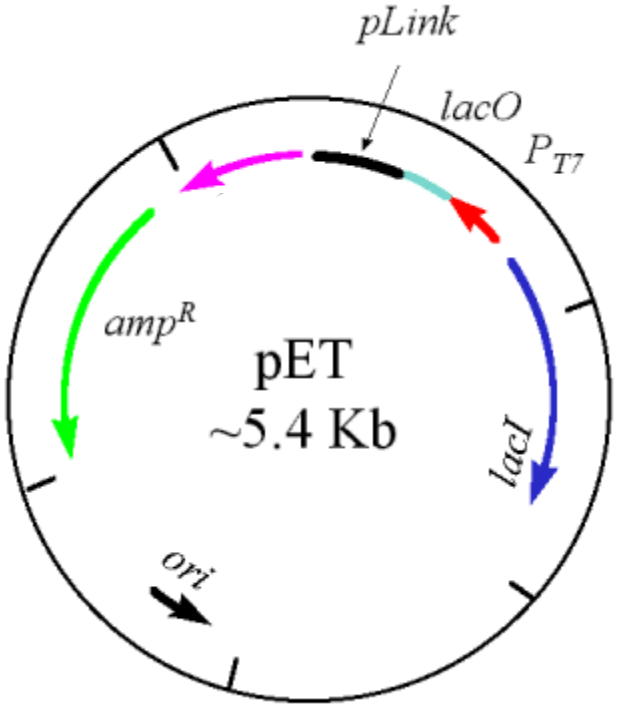


D-glucose

Source: Wikipedia, "Lac Operon"

Bacterial Expression Vectors

- pET Plasmid Genes
  - Origin of replication
  - Lac repressor (lacI)
  - RNA Pol promoter (P_{T7})
  - Lac Operator (lacO)
  - Polylinker – where your DNA sequence goes (pLink)
  - Ampicillin resistance (amp^R)



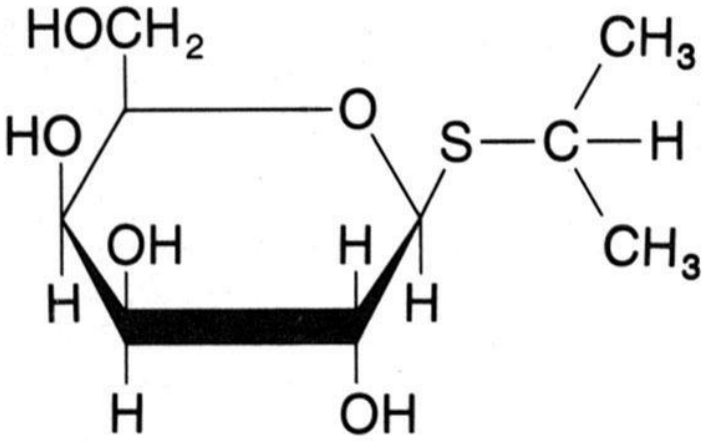
- Is this plasmid persistent?

Source: Mike Blaber, BCH5425 Course Notes

Inducible Expression

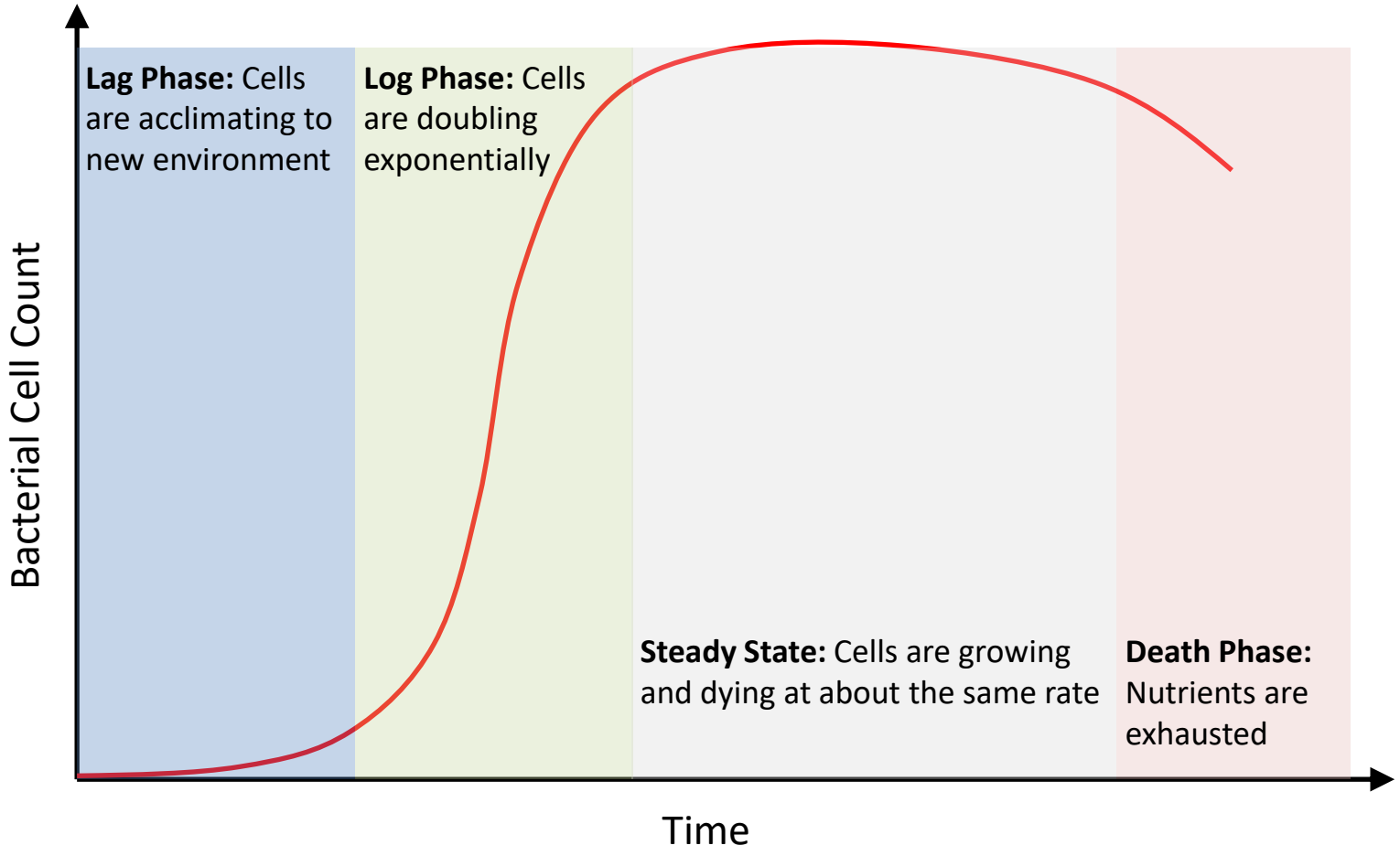
- **IPTG:** Turns on protein expression without being hydrolyzed

- Protein expression can be switched on when desired



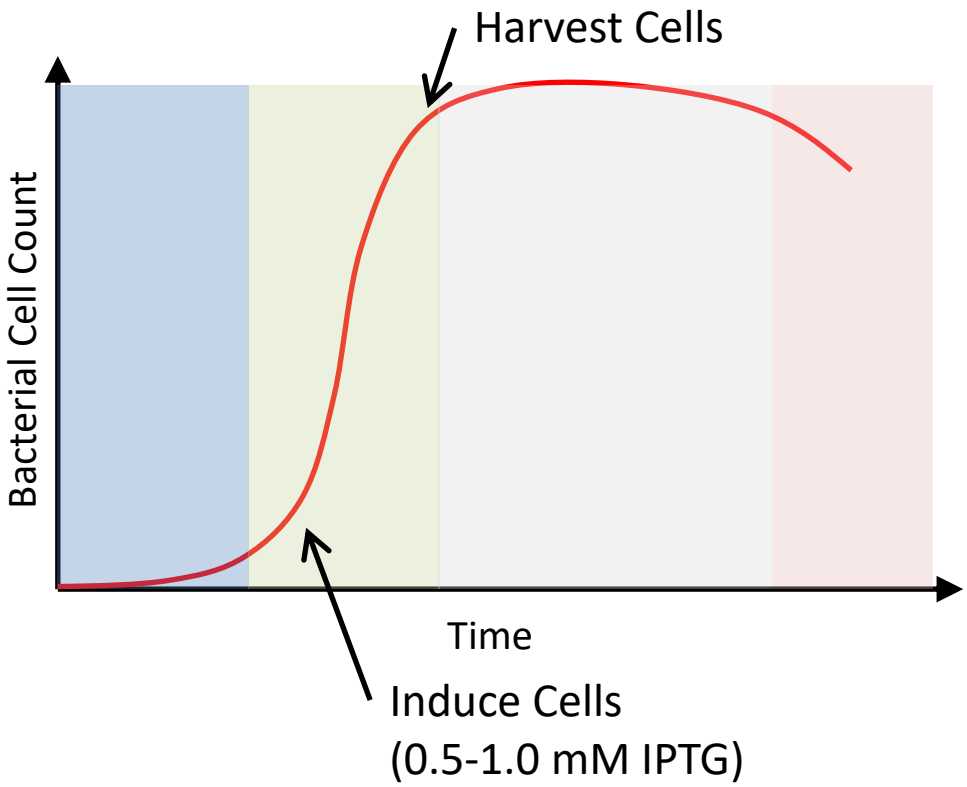
Isopropyl-β-D-thiogalactoside
(IPTG)

When Should I Induce?

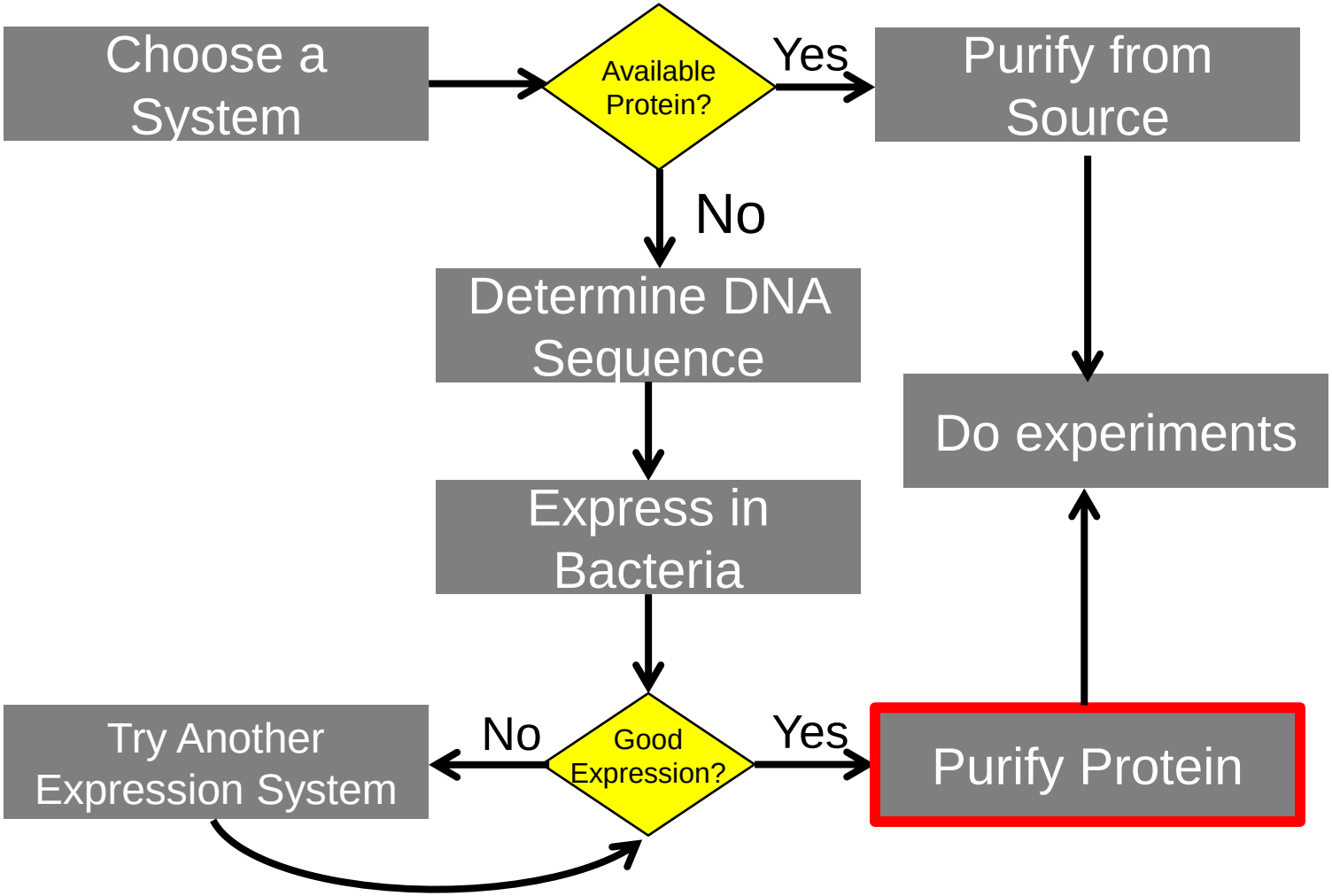


When Should I Induce?

- Protein expression is greatest during log phase
- Inducing at lag phase may unnecessarily cripple your cells
- Typically, induce at an OD_{600} of 0.5-0.6
- Always follow your lab's protocols!



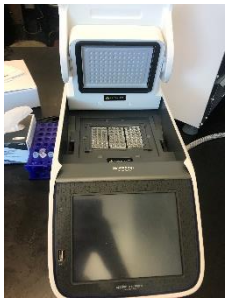
How to obtain sample?



Overview: Recombinant Protein Production

Obtain or build plasmid

molecular biology



Transform into cells

e. coli cells, or others such as insect cell lines

Induce cells

starts production of protein

Harvest cells

spin down with centrifuge, collect cell pellet

Overview: Recombinant Protein Production

Resuspend cells

select buffer, salt, and protease inhibitor(s)

Homogonize / dounce / vortex

mixing /

Lyse cells

sonication, French press

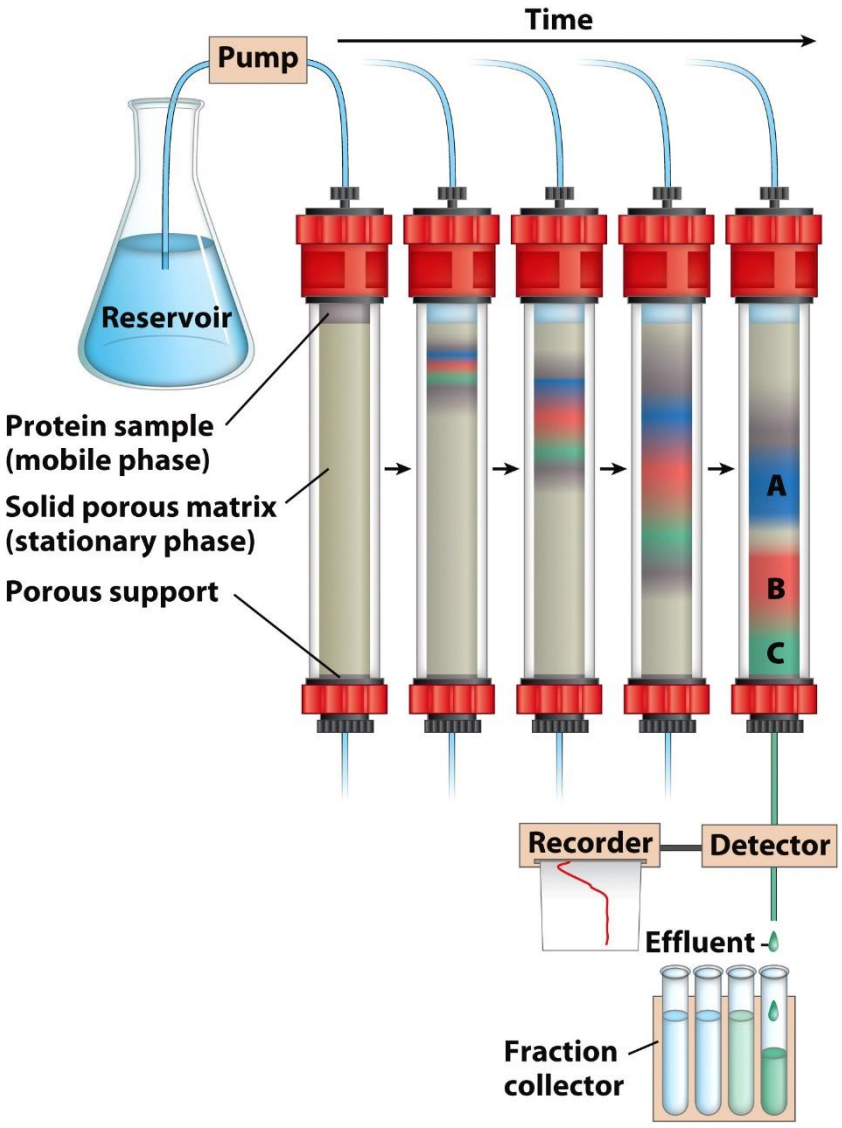
Clarify lysate

centrifuge, syringe filter

Ready for chromatorgraphy

A mixture of proteins can be separated

- Recombinant protein production or purification from endogenous tissue
- Separation relies on differences in physical and chemical properties
  - Charge
  - Size
  - Affinity for a ligand
  - Solubility
  - Hydrophobicity
  - Thermal stability
- Chromatography is commonly used for preparative separation



Column Chromatography / Van Deemter eq

$$HETP = A + \frac{B}{u} + (C_s + C_m) \cdot u$$

Height Equivilent of Theroetical Plate = a measure of the resolving power of the column [m]

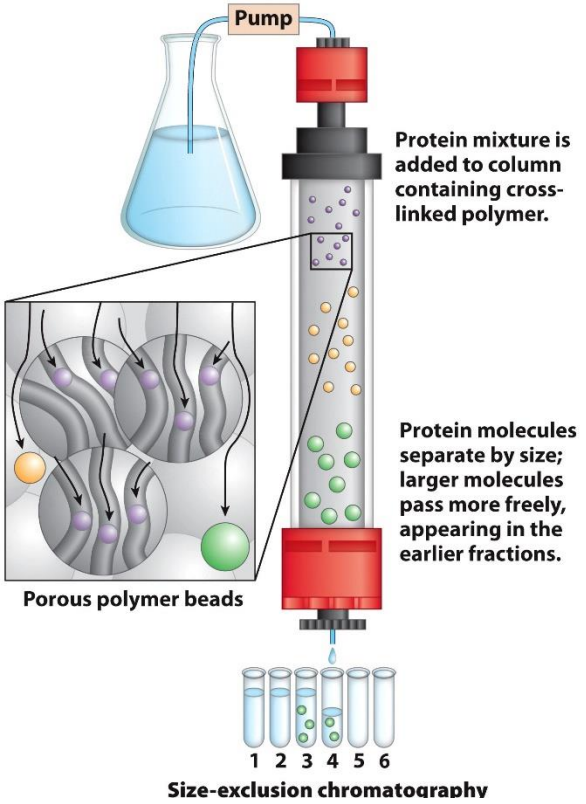
A = Eddy-diffusion parameter, related to channeling through a non-ideal packing [m]

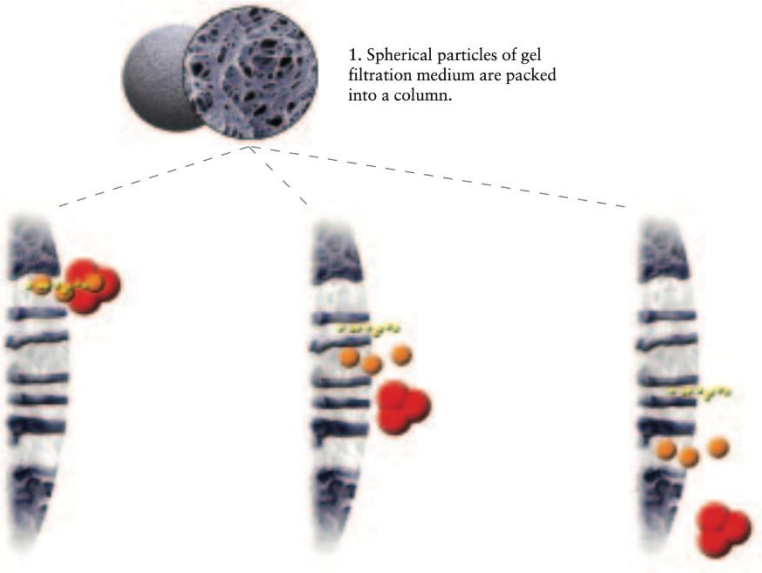
B = diffusion coefficient of the eluting particles in the longitudinal direction, resulting in dispersion [$m^2\ s^{-1}$]

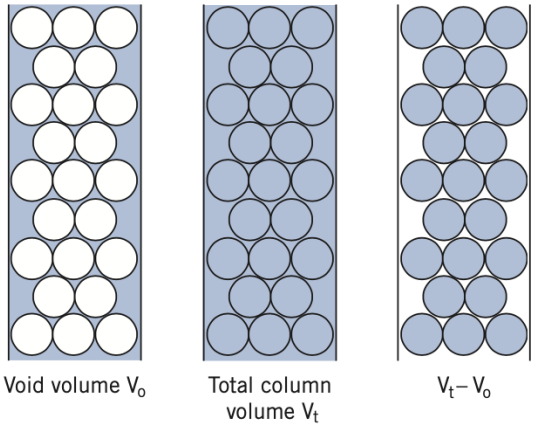
C = Resistance to mass transfer coefficient of the analyte between mobile and stationary phase [s]

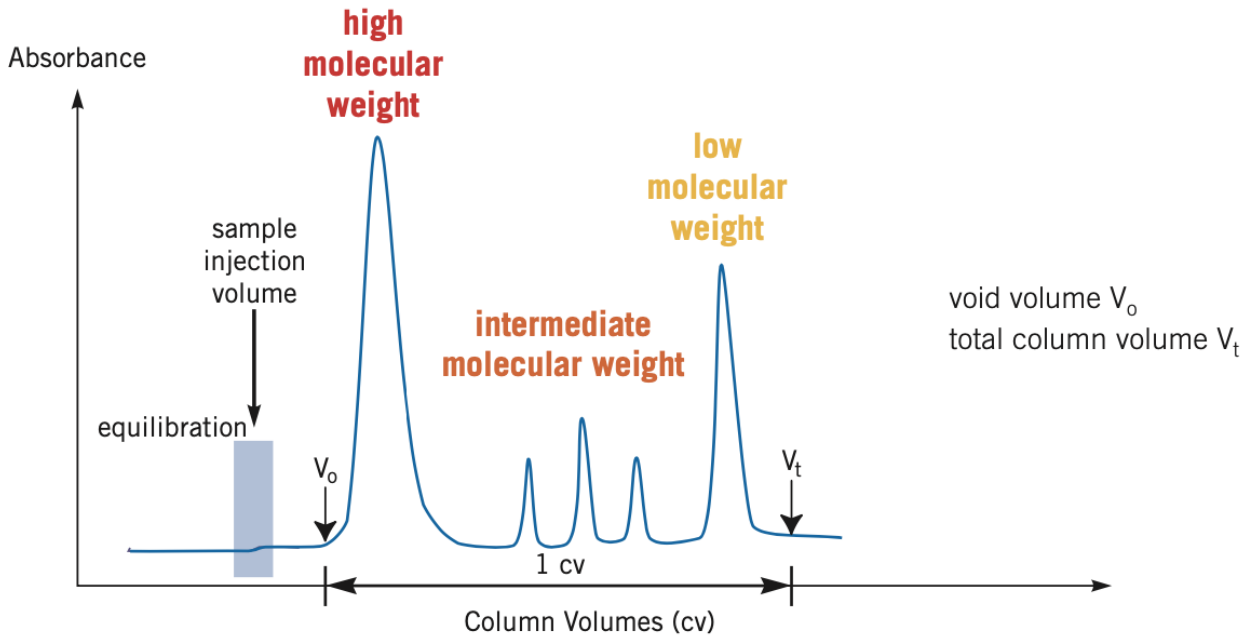
u = speed [$m\ s^{-1}$]

Separation by Size

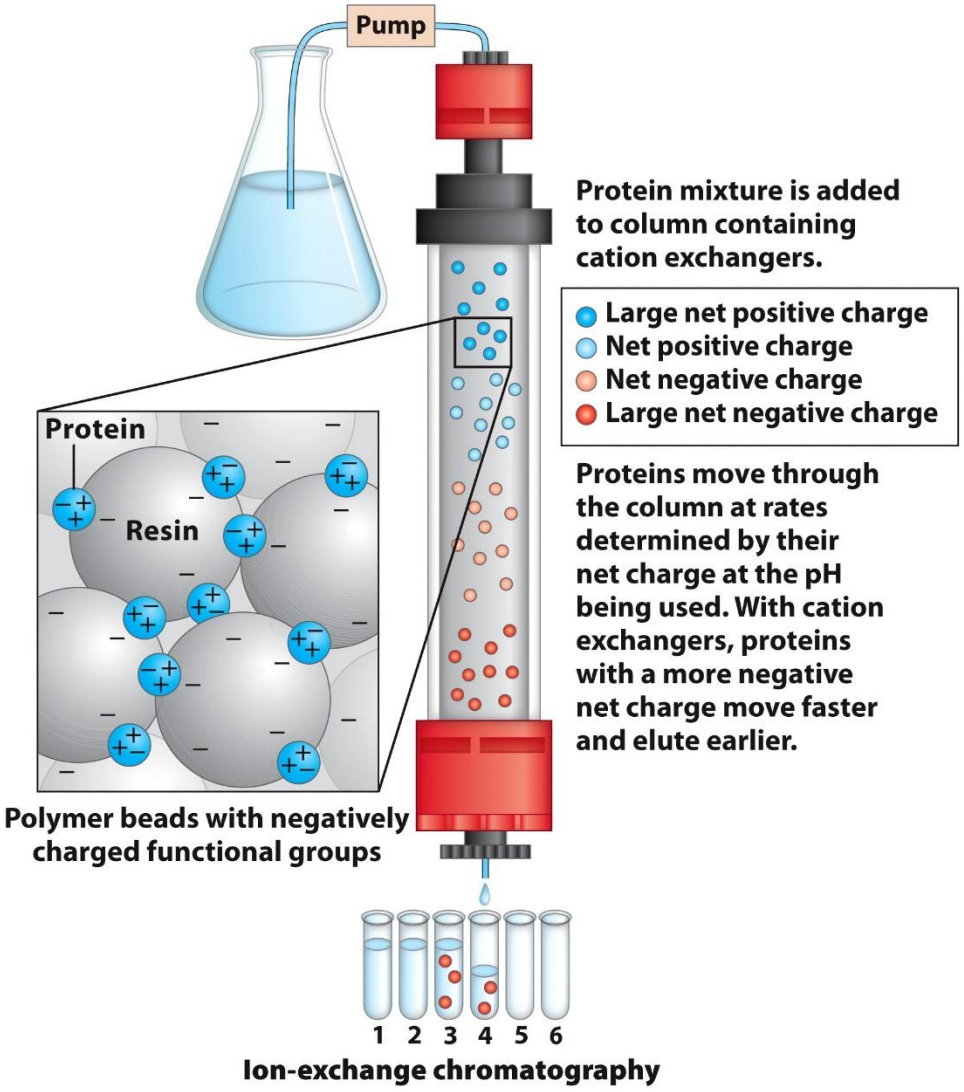




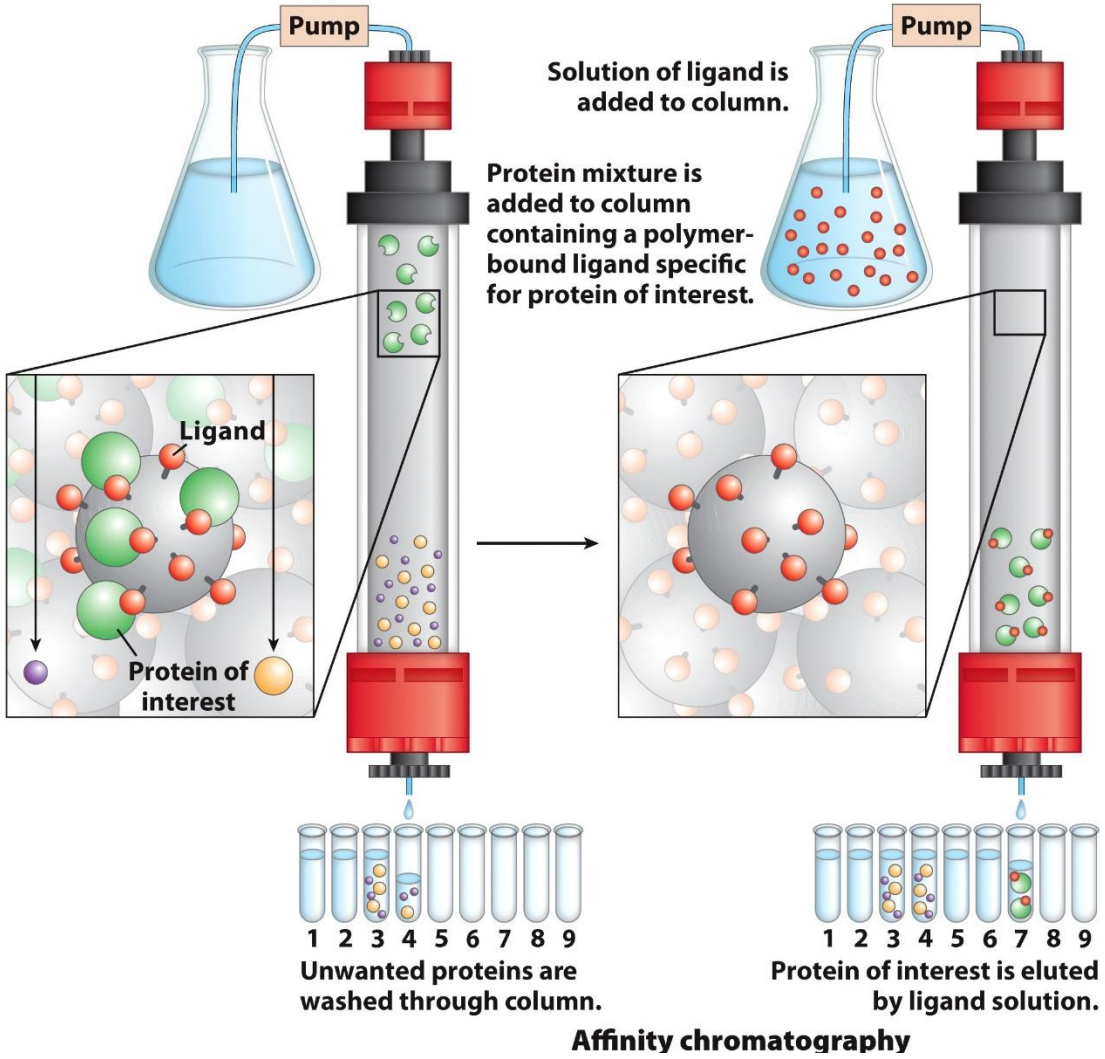




Separation by Charge



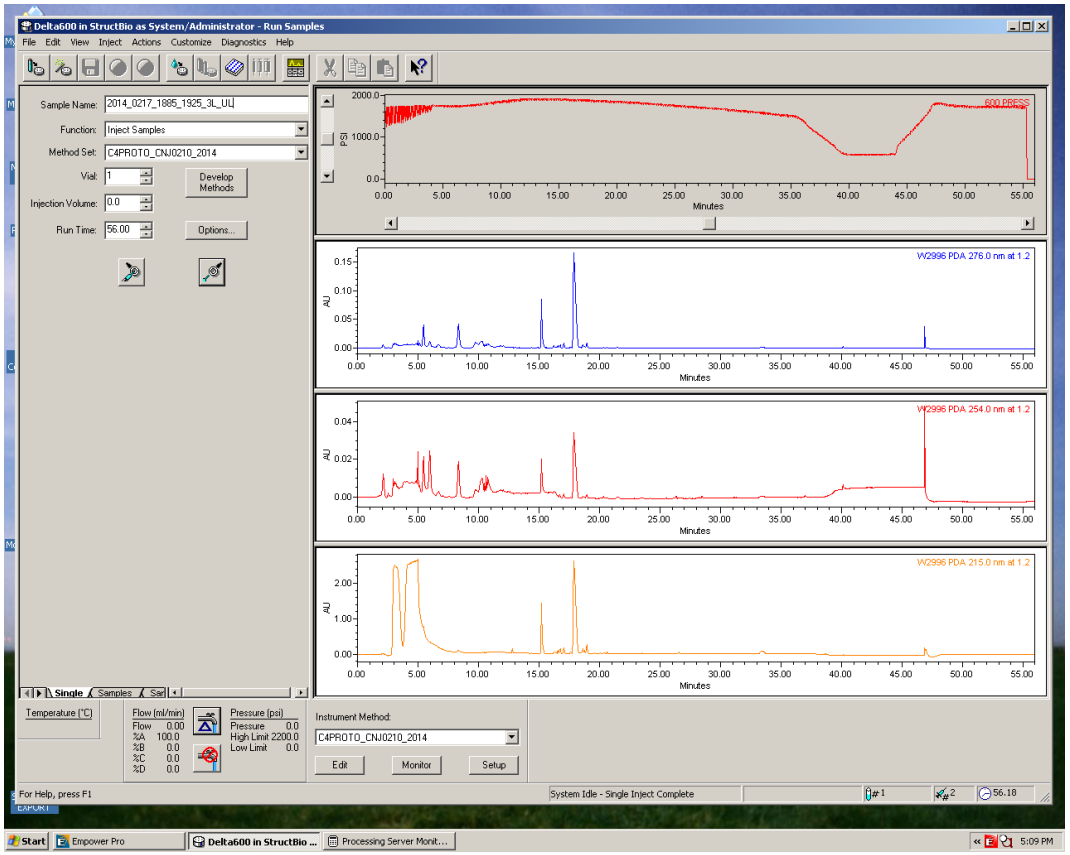
Separation by Affinity



Separation by Hydrophobicity

Reverse phase chromatography

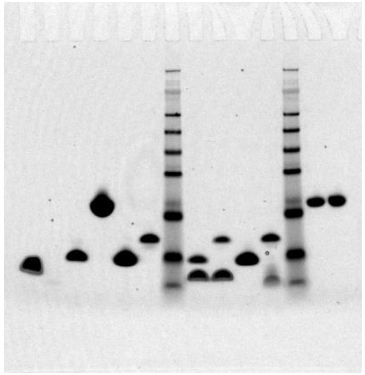
- stationary phase = Carbon polymer
- mobile phase = often water

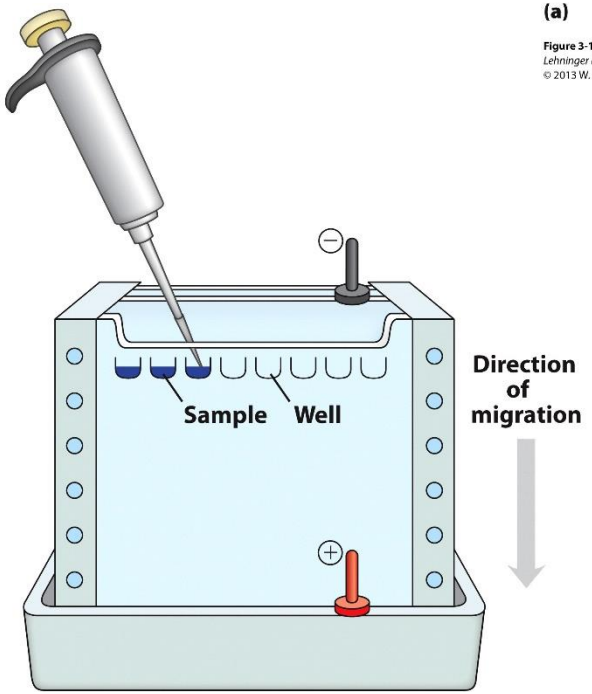


Where is the Protein & How Clean is it?

Electrophoresis

- Electric field pulls proteins by charge
- Gel matrix hinders mobility
  - proteins sorted by size and shape





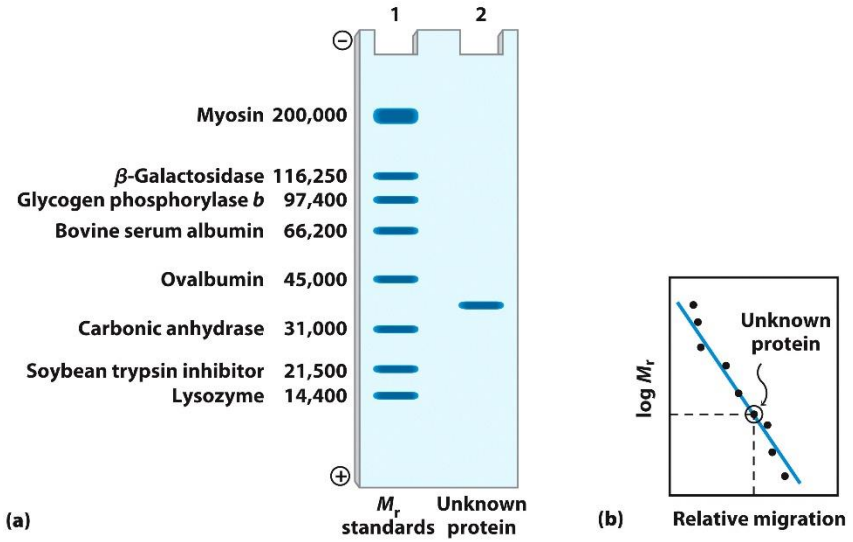


Figure 3-19
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

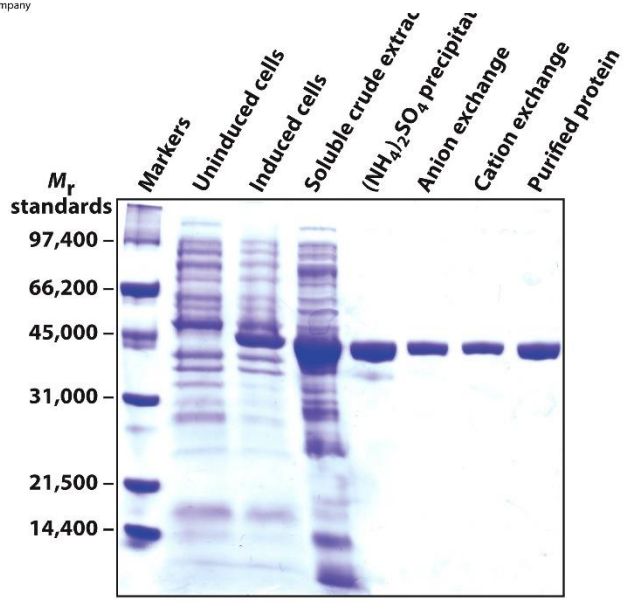
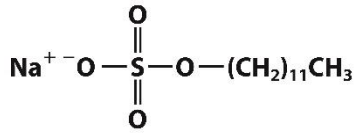


Figure 3-18b
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

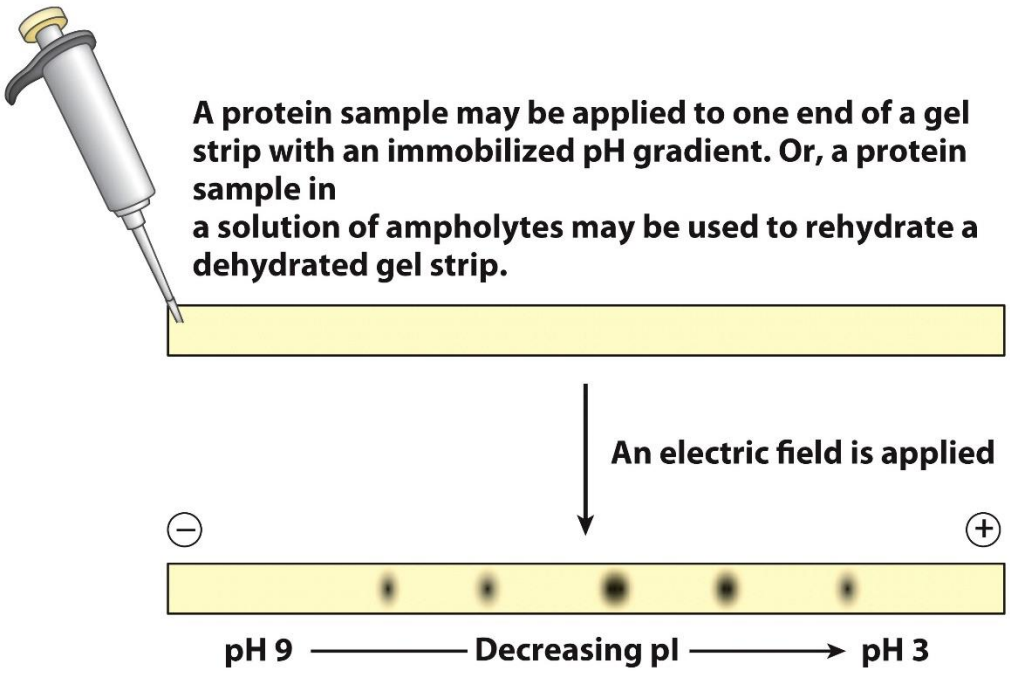
SDS PAGE: Molecular Weight

- SDS – sodium dodecyl sulfate – a detergent

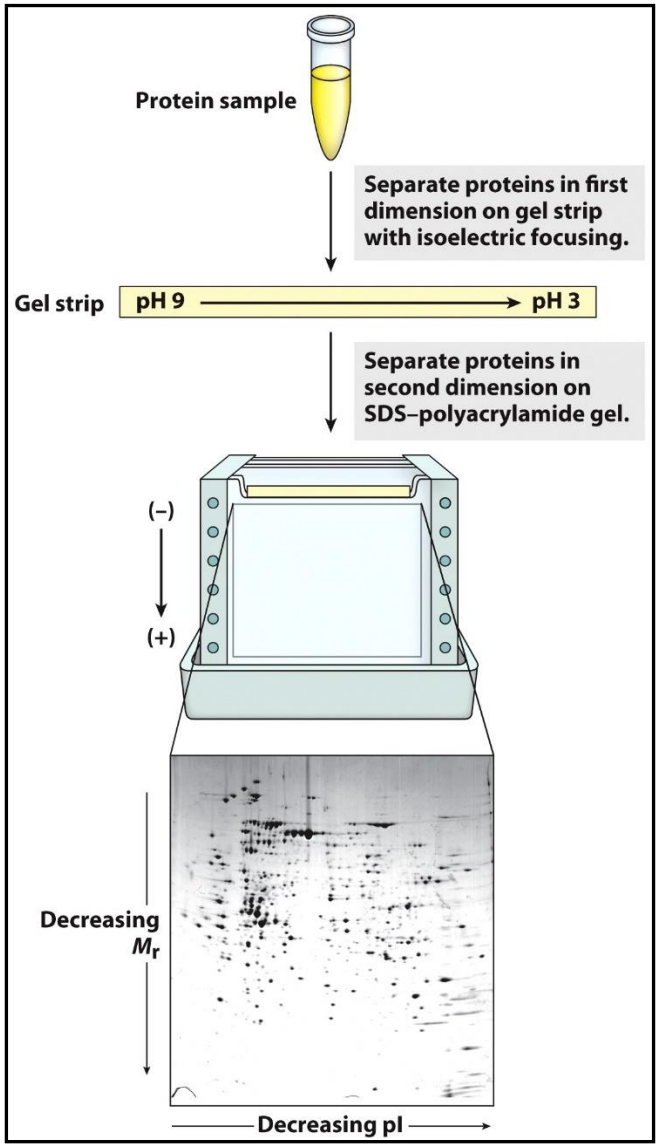


- SDS micelles bind to and unfold all the proteins
  - SDS gives all proteins an uniformly negative charge
  - The native shape of proteins does not matter
  - Rate of movement will only depend on size: small proteins will move faster

Isoelectric Focusing: pI of a protein



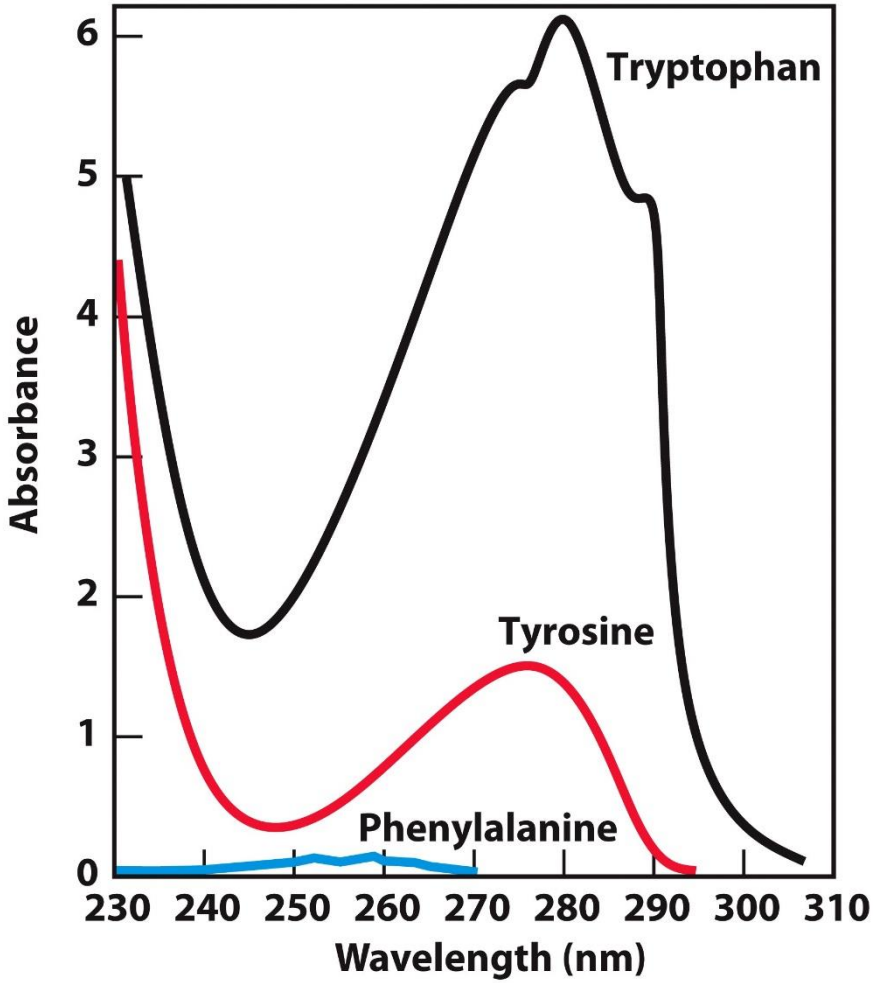
Isoelectric focusing + SDS-PAGE



2D electrophoresis:

Spectroscopic Detection of Aromatic Amino Acids

- Aromatic amino acids absorb light in the UV region
- Proteins - UV absorbance max 275–280 nm
- W and Y = strongest chromophores
- Beers law: $A = \varepsilon \cdot c \cdot l$



Protein Sequencing

- Protein sequence - essential for many biochemical investigations
- sequence typically determined by DNA sequence
- Edman Degradation (historical method)
  - Successive rounds of N-terminal modification, cleavage, and identification
  - Can be used to identify protein with known sequence
- Mass Spectrometry (Modern method)
  - MALDI MS and ESI MS can precisely identify the mass of a peptide, and thus the amino acid sequence
  - Can be used to determine post-translational modifications