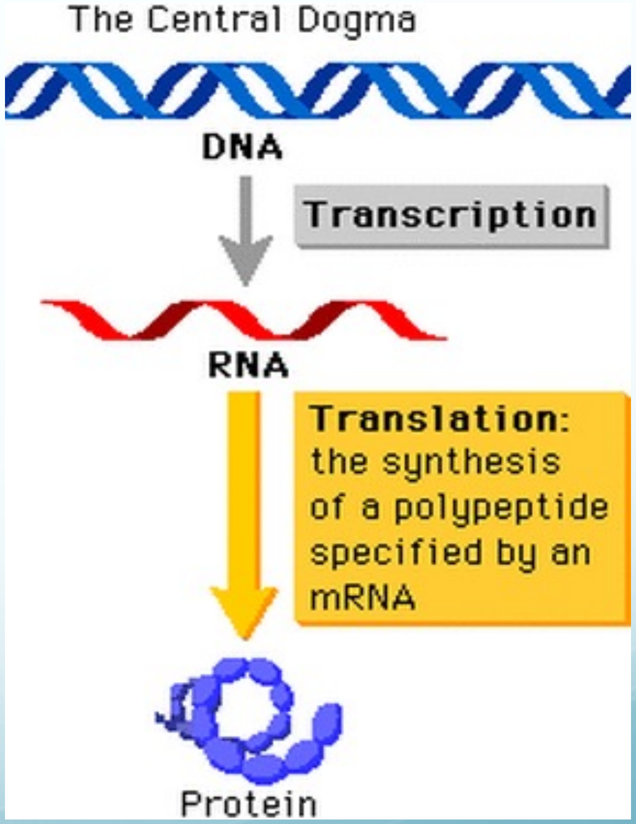
The *lac* operon in *E. coli*

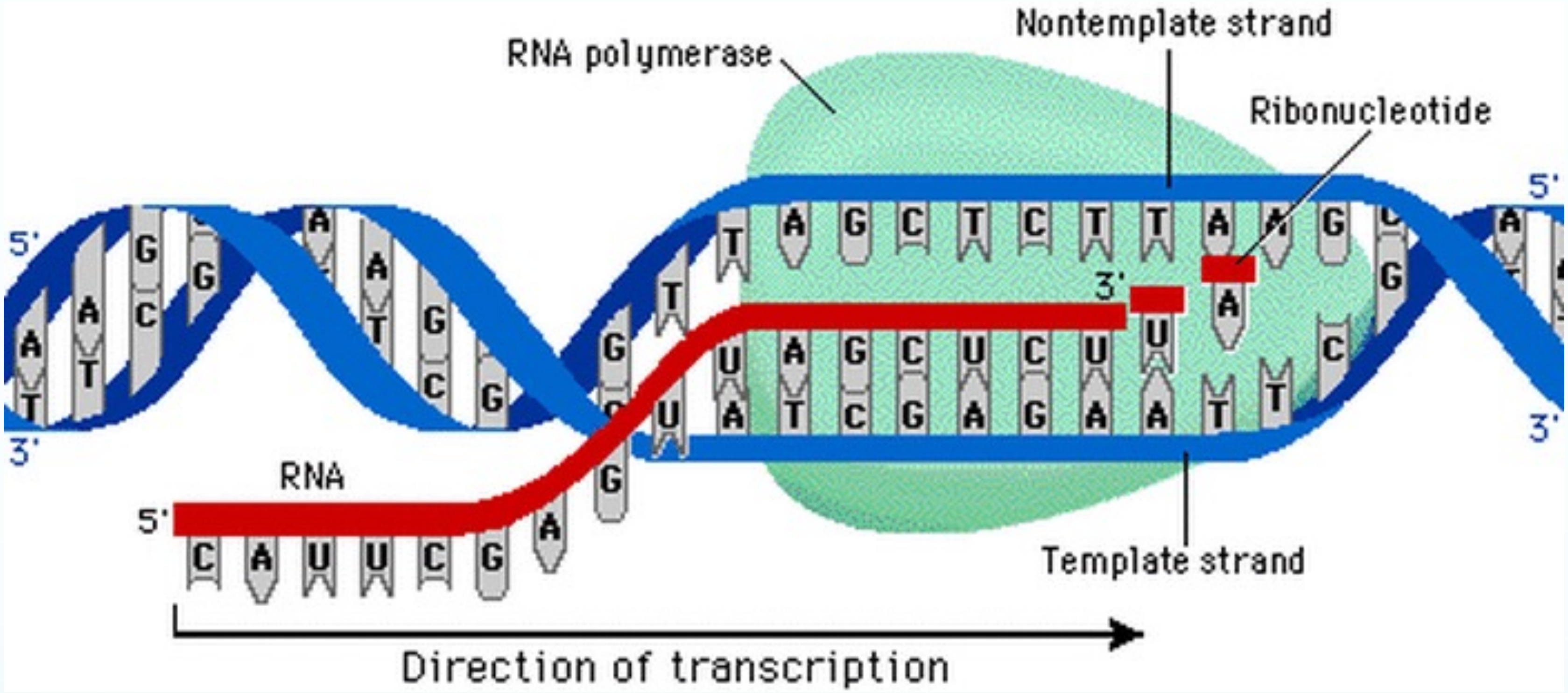
Matthew Macauley
Clemson University

Gene expression

- Gene expression is a process that takes gene info and creates a functional gene product (e.g., a protein).
- Gene Expression is a 2-step process:
  1) transcription of genes (messenger RNA synthesis)
  2) translation of genes (protein synthesis)
- DNA consists of bases A, C, G, T.
- RNA consists of bases A, C, G, U.
- Proteins are long chains of amino acids.
- Gene expression is used by all known life forms.

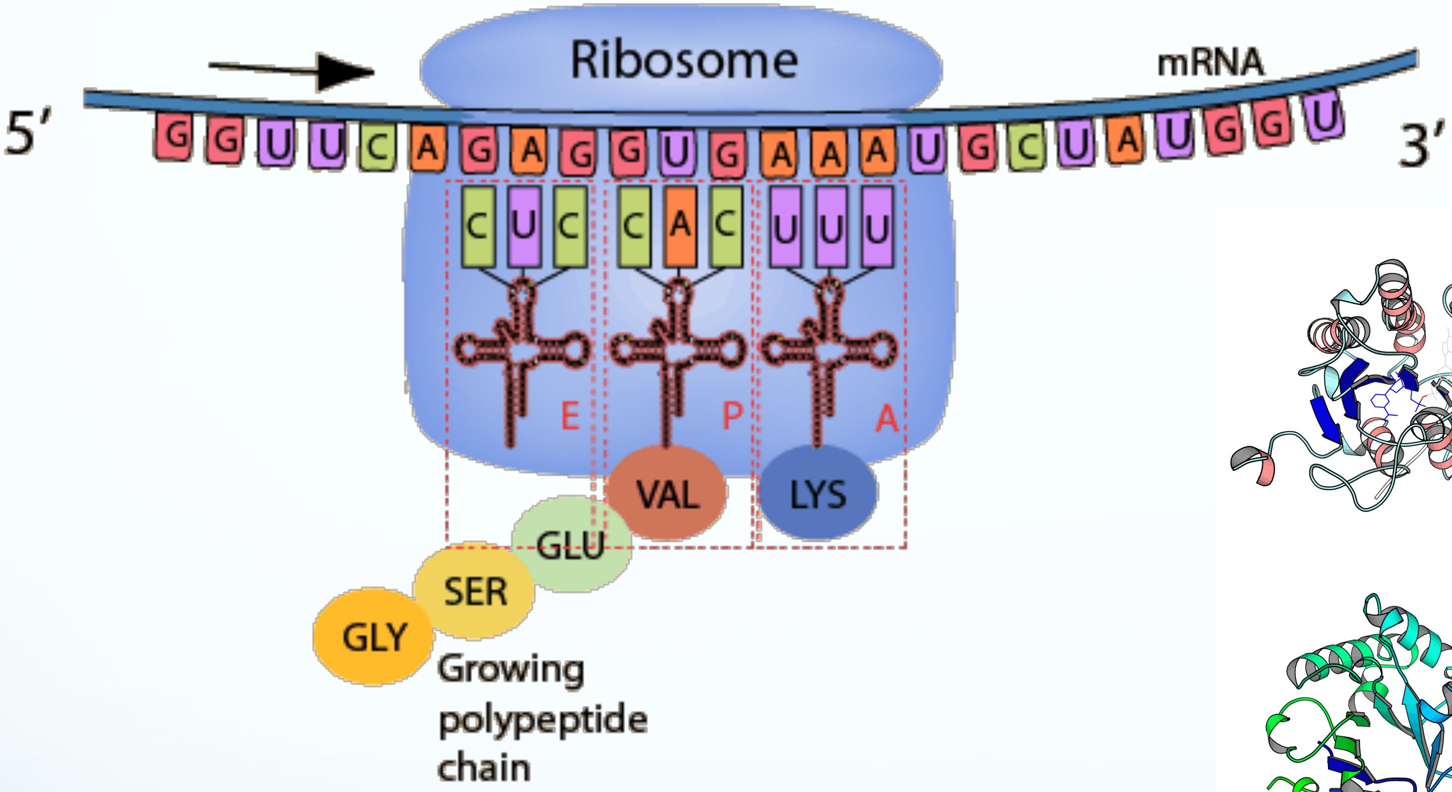


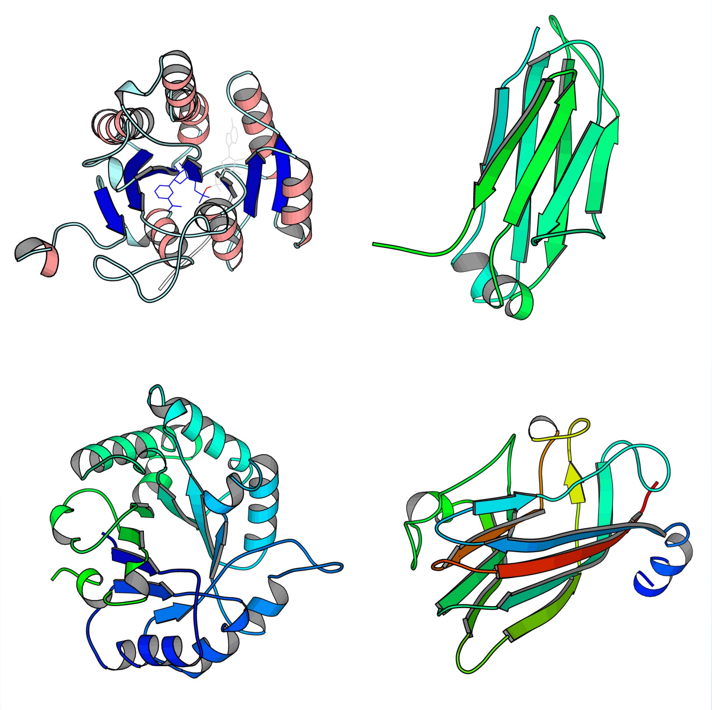
Transcription



- Transcription occurs inside the cell nucleus.
- RNA polymerase uses the helicase enzyme to binds to DNA, “unzipping” it to read it.
- DNA is copied into mRNA.
- Segments of RNA not needed for protein coding are removed.

Translation





- During translation, the mRNA is read by ribosomes.
- Each triple of RNA bases codes for an amino acid.
- The result is a protein: a long chain of amino acids.
- Proteins fold into a 3-D shape which determine their function

Gene expression

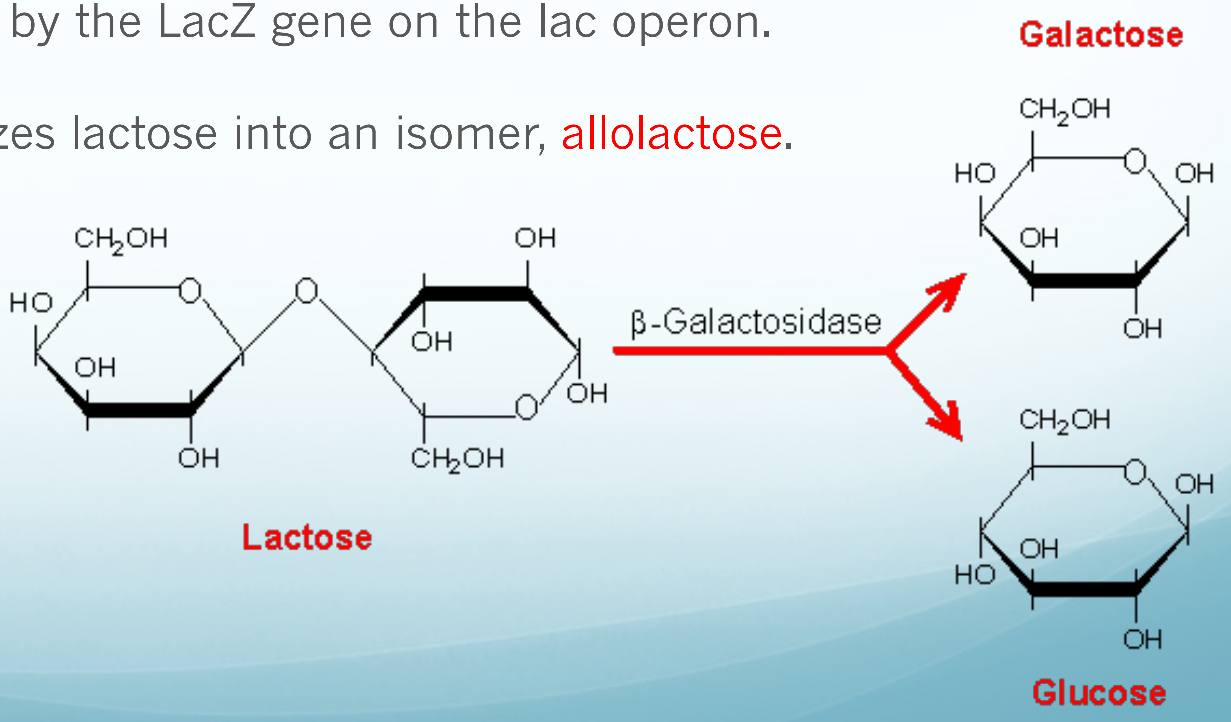
- The expression level is the rate at which a gene is being expressed.
- Housekeeping genes are continuously expressed, as they are essential for basic life processes.
- Regulated genes are expressed only under certain outside factors (environmental, physiological, etc.). Expression is controlled by the cell.
- It is easiest and most efficient to control gene regulation by affecting transcription.
- One way to block transcription is for repressor proteins bind to the DNA or RNA.
- **Goal**: Understand the complex cell behaviors of gene regulation, which is the process of turning on/off certain genes depending on the requirements of the organism.

The *lac* operon in *E. coli*

- An operon is a region of DNA that contains a cluster of genes that are transcribed together.
- *Escherichia coli* is a bacterium in the gut of mammals and birds. Its genome has been sequenced and its physiology is well-understood.
- The lactose (*lac*) operon controls the transport and metabolism of lactose in *E. coli*.
- The *lac* operon was discovered by Francois Jacob and Jacques Monod in 1961, which earned them the Nobel Prize.
- The *lac* operon was the first operon discovered and is the most widely studied mechanism of gene regulation.
- The *lac* operon is used as a "test system" for models of gene regulation.
- DNA replication and gene expression were all studied in *E. coli* before they were studied in eukaryotic cells.

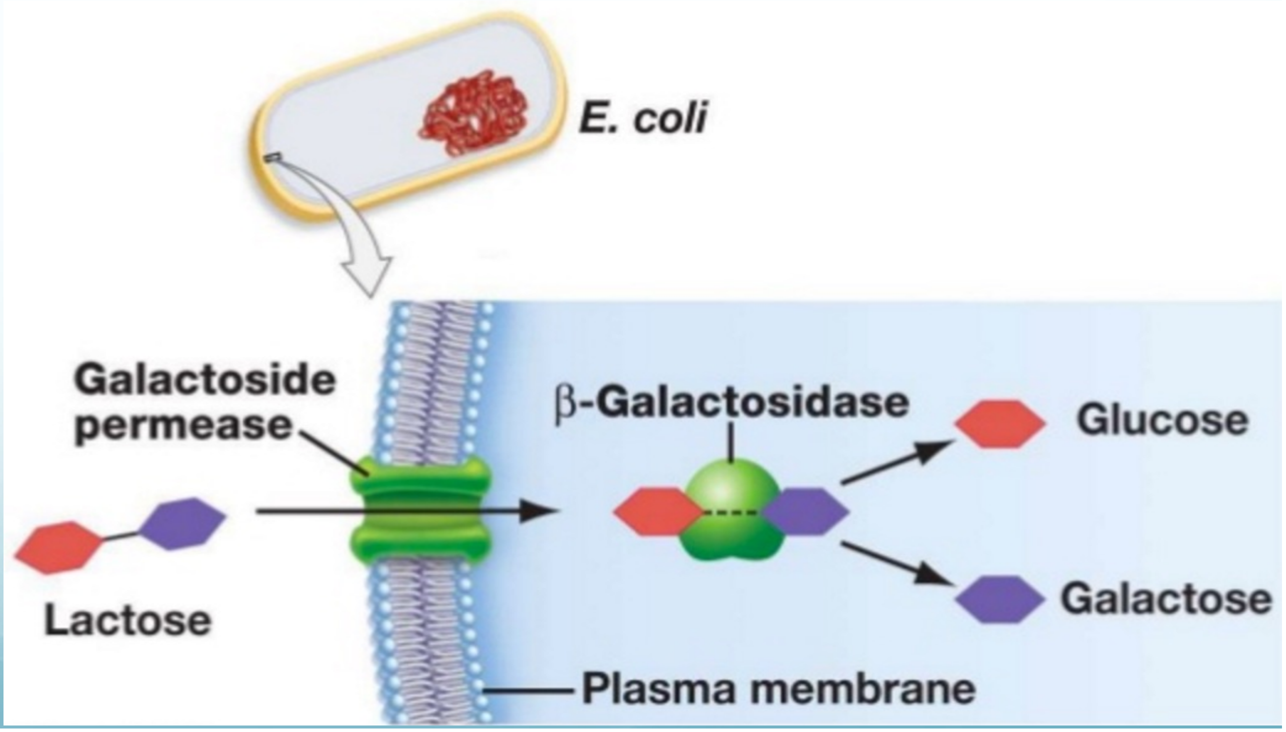
Lactose and β–galactosidase

- When a host consumes milk, *E. coli* is exposed to lactose (milk sugar).
- Lactose consists of one glucose sugar linked to one galactose sugar.
- If both glucose and lactose are available, then glucose is the preferred energy source.
- Before lactose can be used as energy, the β–galactosidase enzyme is needed to break it down. (Analogy: "*molecular scissors*".)
- β–galactosidase is encoded by the LacZ gene on the lac operon.
- β–galactosidase also catalyzes lactose into an isomer, allolactose.

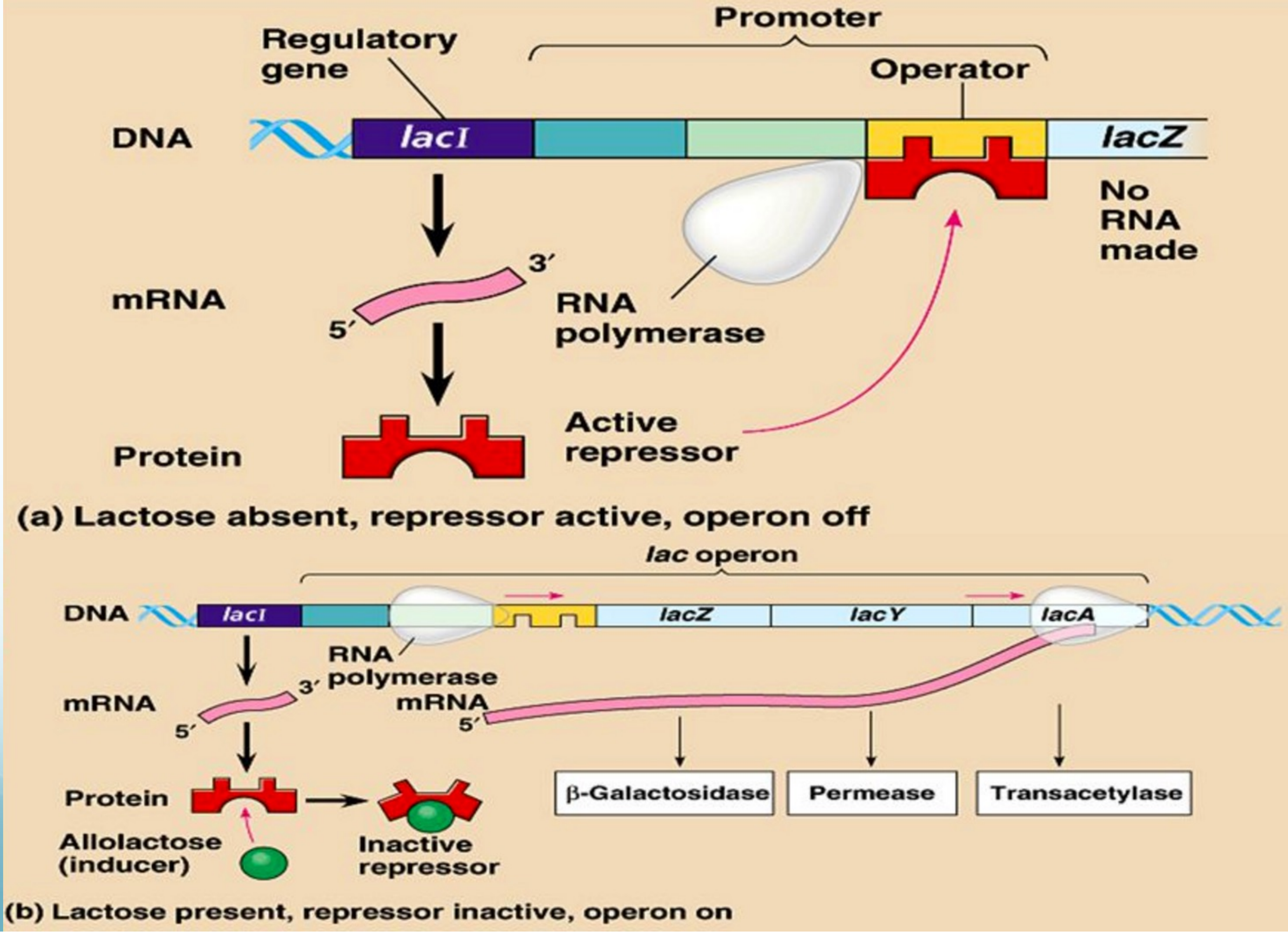


Transporter protein

- To bring lactose into the cell, a transport protein, called *lac* permease, is required. (Analogy: "*doggie door*".)
- This protein is encoded by the LacY gene on the *lac* operon.
- If lactose is not present, then neither of the following are produced:
  1) β–galactosidase (LacZ gene)
  2) *lac* permease (LacY gene)
- In this case, the *lac* operon is OFF.



The *lac* operon



(a) Lactose absent, repressor active, operon off

(b) Lactose present, repressor inactive, operon on

With lactose and no glucose

- Lactose is brought into the cell by the *lac* permease transporter protein
- β−galactosidase breaks up lactose into glucose and galactose..
- β−galactosidase also converts lactose into allolactose.
- Allolactose binds to the *lac* repressor protein, preventing it from binding to the operator region of the genome.
- Transcription begins: mRNA encoding the *lac* genes is produced.
- Lac proteins are produced, and more lactose is brought into the cell. (The operon is ON.)
- Eventually, all lactose is used up, so there will be no more allolactose.
- The *lac* repressor can now bind to the operator, so mRNA transcription stops. (The operon has turned itself OFF.)

Catabolite repression

- The cellular mechanism catabolite repression turns the *lac* operon OFF when both glucose and lactose are present.

- The *lac* operon promoter region has 2 binding sites:
  - One for RNA polymerase (this “unzips” and reads the DNA)
  - One for the CAP-cAMP complex. This is a complex of two molecules: catabolite activator protein (CAP), and the cyclic AMP receptor protein (cAMP, or *crp*).

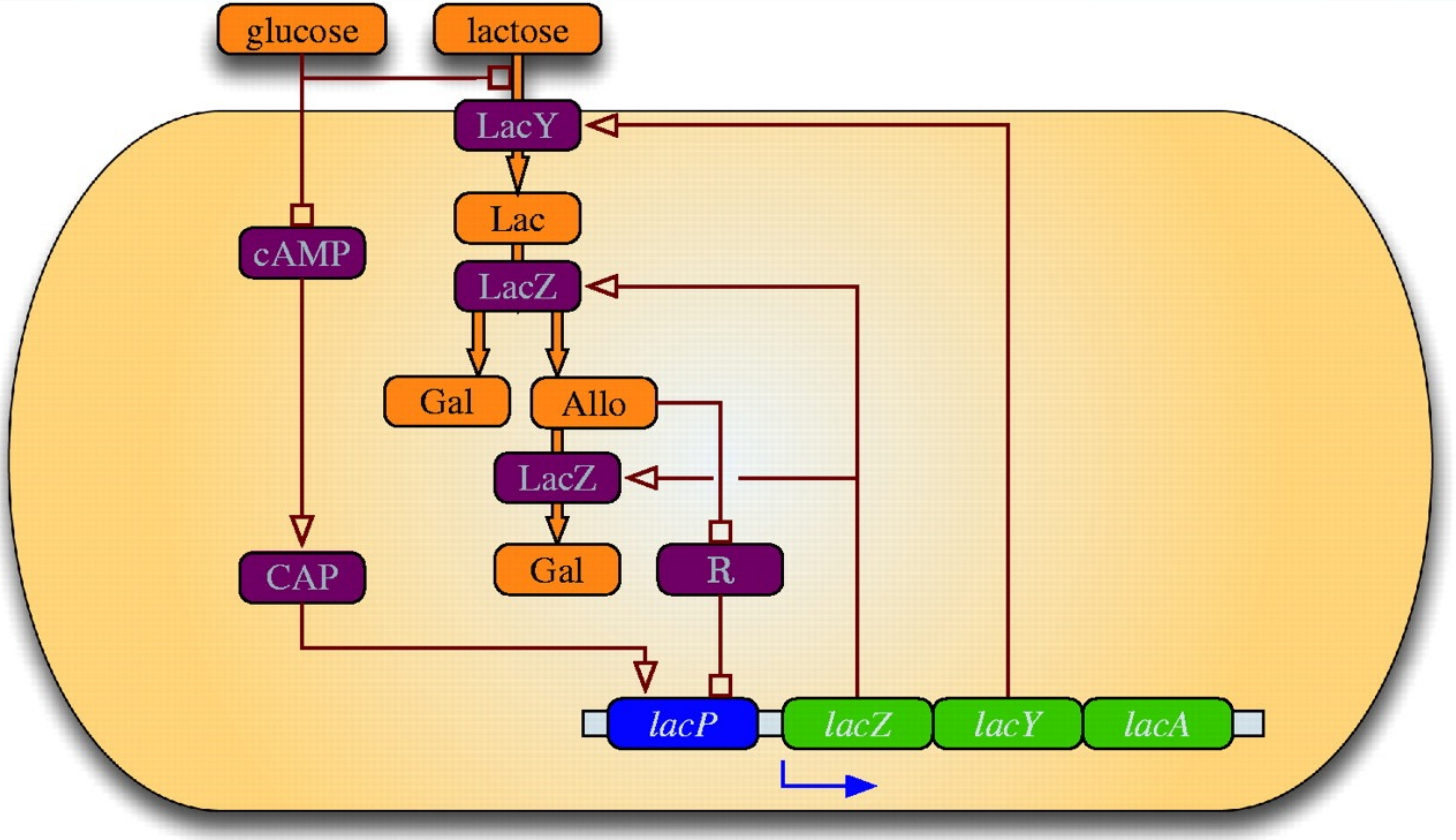
- Binding of the CAP-cAMP complex is required for transcription for the *lac* operon. (Analogy: “*the key that starts the engine*”.)

- Intracellular glucose causes the cAMP concentration to decrease.

- When cAMP levels get too low, so do CAP-cAMP complex levels.

- Without the CAP-cAMP complex, the promoter is inactivated, and the *lac* operon is OFF.

Lac operon gene regulatory network



Modeling the *lac* operon

- Models of molecular networks like the *lac* operon have variables that represent concentration levels of the key gene products and substrates.
- Any reasonable model should be able to capture the following:
  - No lactose: operon is OFF
  - Lactose and glucose: operon is OFF
  - Lactose and no glucose: operon is ON.
- Molecular concentrations are highly nonlinear, and they depend on complex biochemical reactions.
- There are also other features such as time delays, dilution, degradation, and bistability that modeling frameworks should be able to handle.
- We will see two very different approaches to modeling the *lac* operon:
  - Delay differential equations (quantitative continuous framework).
  - Boolean model (qualitative discrete framework).
- Both have their pros and cons, which we will discuss.