Boolean models of the lac operon

(1) The lac operon regulatory network

Gene expression: A process that takes gene info & synthesizes (= creates) a functional gene product (e.g., a protein).

Some genes code for proteins

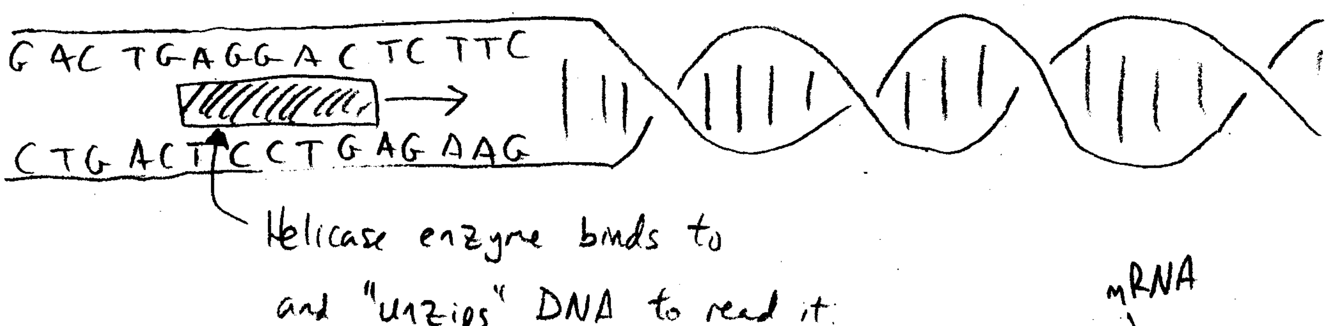
Others (e.g., rRNA, tRNA) code for functional RNA

Summary: Gene expression is a 2-step process

(i) transcription of genes (messenger RNA synthesis)

(ii) translation of genes (protein synthesis).

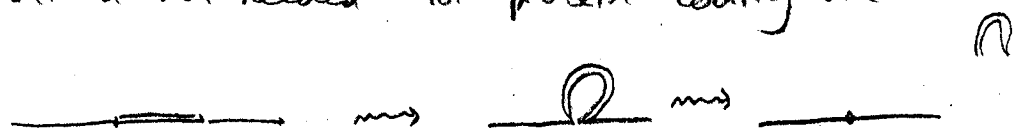
Transcription: DNA is copied into mRNA (inside cell nucleus)



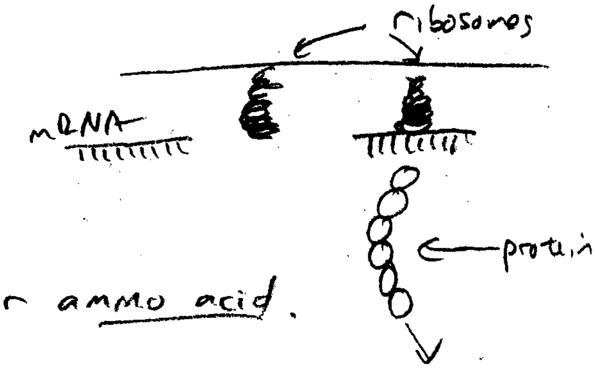
Output: A complementary strand of mRNA ...CUGACUCCUGAGAAG (mRNA)

Note: DNA consists of basepairs A, C, G, T. RNA of A, C, G, U

Segments of RNA strand not needed for protein coding are removed:



RNA then leaves the nucleus.



<u>Translation</u>: RNA read by ribosomes.

Each triple of RNA bases code for an <u>ammo acid</u>.

A protein is a chain of ammo acids.

| GUG | CAU | CUG | ACU | CCU | GAG | AAG | ← mRNA strand

↓ ↓ ↓ ↓ ↓ ↓ ↓

—V—H—L—T—P—E—K— ← protein.

The protein then folds into a 3D shape

The <u>expression level</u> is the rate at which the gene is being expressed.

<u>Housekeeping genes</u> are continuously expressed, as they are essential for basic life processes.

<u>Regulated genes</u> are expressed only under certain outside factors (environmental, physiological). Expression is controlled by the cell.

Easiest to control gene regulation by affecting transcription.

Certain repressor proteins bind to sites on DNA or RNA.

Goal: Understand this complex cell behavior.

Mathematical models need to reflect dependencies between a system components.

* Types of models

Deterministic vs. stochastic

Dynamic vs. static

↳ time-continuous vs. time-discrete

Continuous models (e.g., DEs) need knowledge of interactions rates concentrations (constants). These models are quantitative

Discrete-time, discrete-space models are called algebraic models. They are qualitative.

Boolean models (state space = $\{0, 1\}$) were introduced in 1969 by S. Kauffman for gene regulatory network models.

A classic example: the lactose (lac) operon — controls transport & metabolism of lactose in E. coli.

One of the most widely studied mechanisms of gene regulation

Used as a "test system" for models of gene regulation.

E. coli: Bacterium in the intestines of mammals & birds.
(affected by animal's diet!).

DNA replication & gene expression all studied in E. coli before they were studied in Eukaryotic cells.

Physiology well-understood, genome has been sequenced.

Host consumes milk ⇒ E. coli exposed to lactose (milk sugar)

Lactose consists of one glucose sugar linked to one galactose sugar.

Glucose can be used as an energy source

So can galalactose, but an enzyme is needed.

- To transport sugars in/out of cell, sugar transport proteins are required.

- Once imported, specific enzymes act on the sugar. Either:
  - Use it to make a cellular molecule (anabolism)
  - break it down to harvest energy

Glucose is the preferred energy source for all cells.

Interactions between sugars, transport proteins, enzymes are specialized.

Ex: E. coli: won't make lactose transport protein if no lactose present

won't make enzyme to break lactose → glucose / → galactose either.

Cells only make most proteins when needed: inducible proteins.

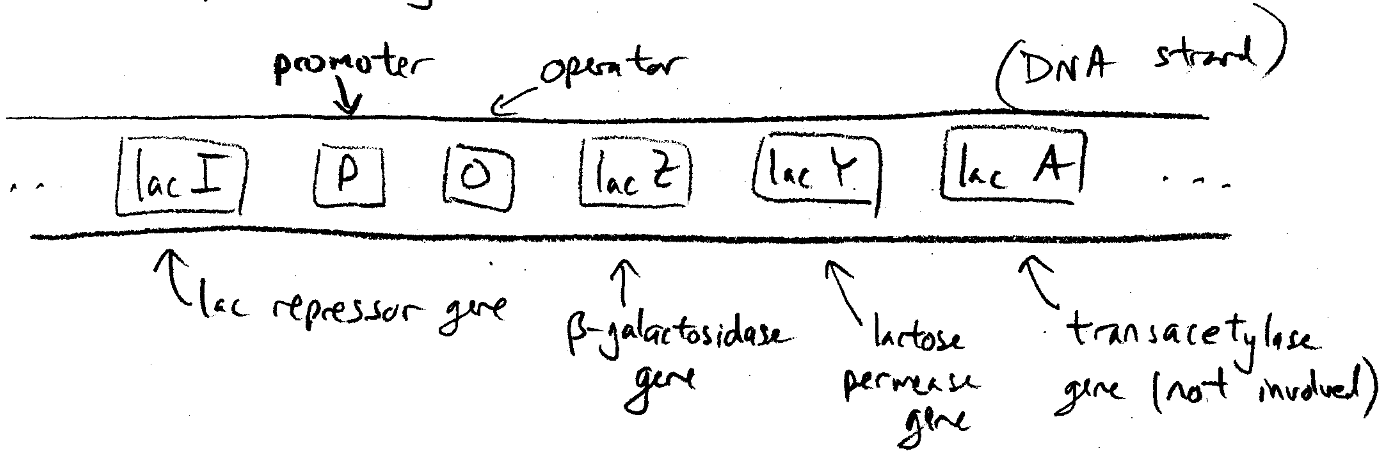
Inducible genes are regulated genes.

Ex: lac operon:

lactose needs:
- lac permease (transporter protein)
- β-galactosidase (enzyme)

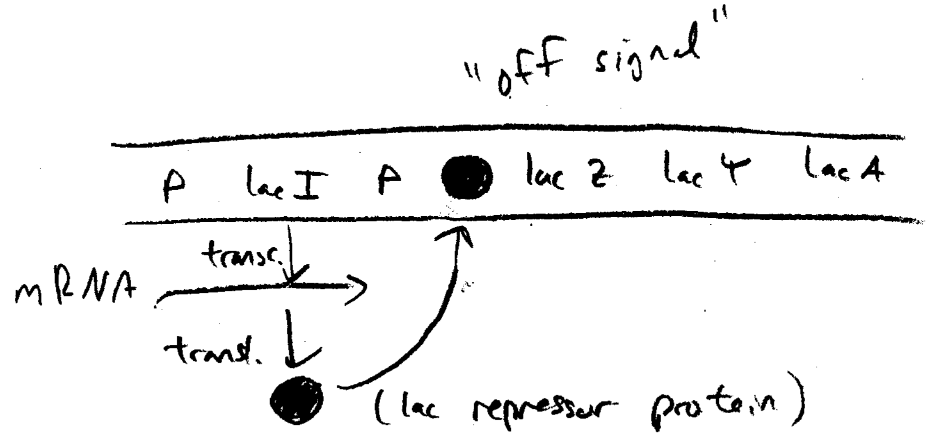
to be utilized.

These are produced by the lac operon:



All three of these genes (lac I, lac Z, lac A) encoded by a single mRNA strand.

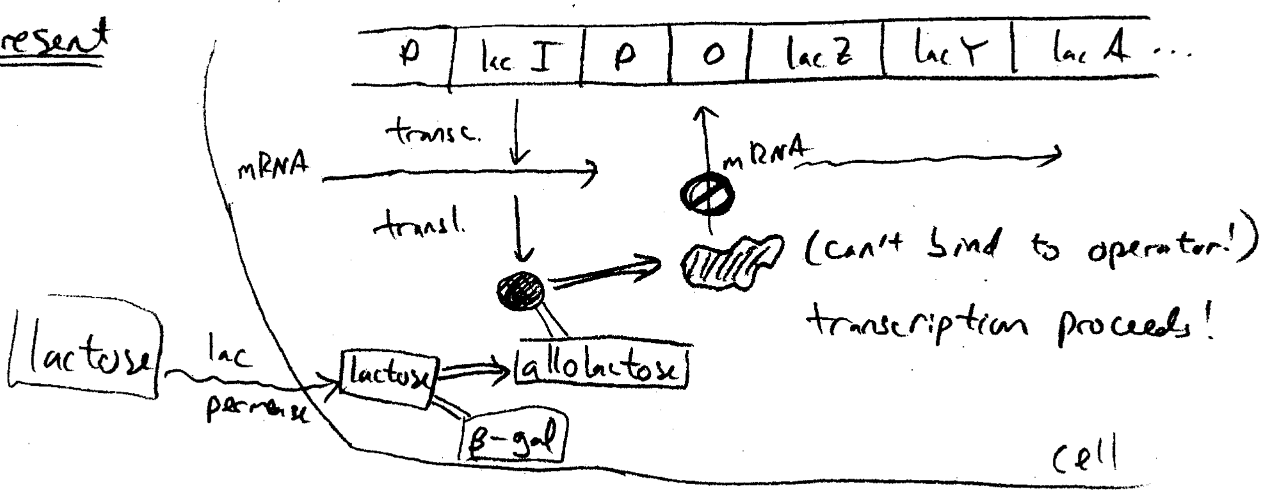
<u>No lactose present</u>



Lac I gene codes for the lac repressor protein which binds to the operator (controlling region) and prevents mRNA transcription.

This "turns off" the operon

<u>Lactose present</u>



<u>Summary</u>

- Lactose is brought into cell by lac permease protein
- β-galactosidase converts lactose into allolactose
- allolactose binds to lac repressor, so it can't bind to operator
- transcription continues! mRNA encoding lac genes is produced.

- lac proteins are produced, more lactose is brought into cell. (operon is on)
- Eventually, all lactose is used up, so there will be no more allolactose.
- The lac repressor binds to operator, mRNA transcription stops. (the operon has turned itself off).

Case 3 Glucose and lactose are both available:

Glucose (preferred energy source) is used up first.

This is controlled by mechanism of catabolite repression.

Without it, the steps outlined in the "lactose present" case would take place (lac repressor protein doesn't bind, transcription occurs)

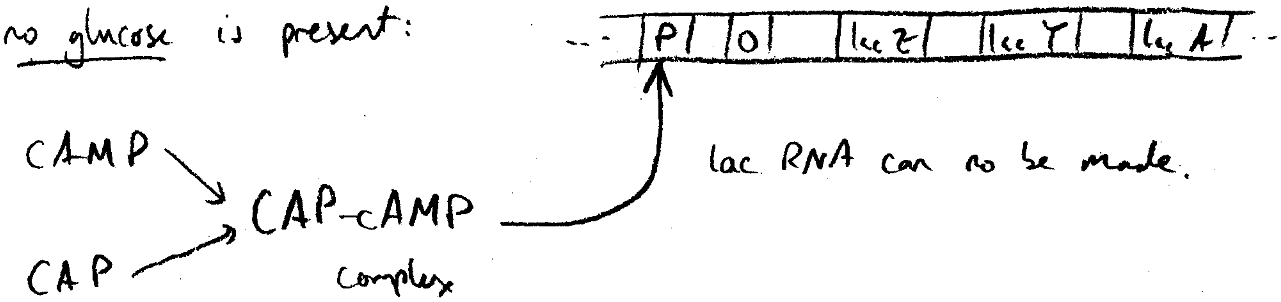
- Catabolite repression represses this process.
- Needs DNA binding protein, "catabolite activator protein" (CAP).

  Product of cAMP receptor protein (crp) gene
- cAMP produced by adenylate cyclase (enzyme).

  ↳ if and only if no glucose present!

(8)

So if no glucose is present:

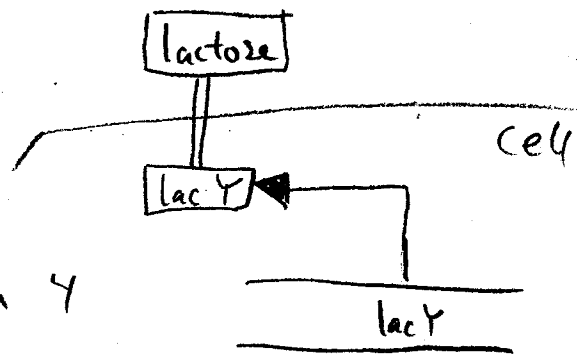


If glucose is present, E. coli uses it. Keeps cAMP levels low.

When glucose is used up, cAMP levels rise.

The CAP-cAMP facilitates attachment of RNA polymerase.

Network interactions & motifs (Summary)

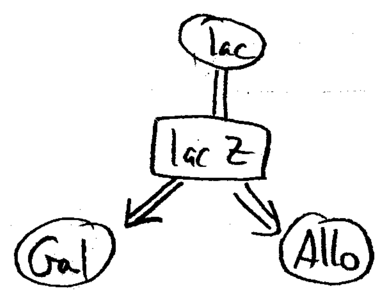


- Lac Y expression creates lac permease, transports lactose into cell.

- Lac Z expression creates β-galactosidase from 4 Lac Z polypeptides. β-gal converts:

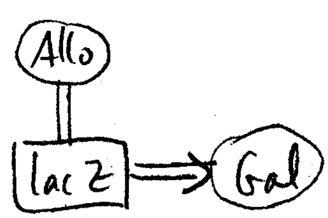
  * lactose into allolactose (Allo)

  or * lactose into glucose & galactose (Gal)



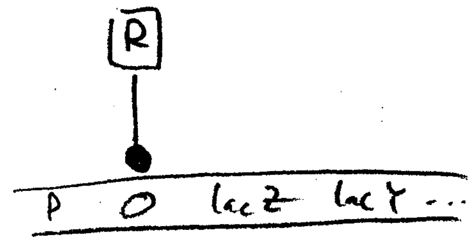
- β-gal also converts

  * allolactose into glucose & galactose



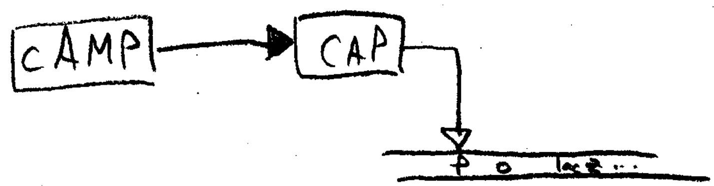
- Allolactose binds to repressor protein (R), inhibiting it:

- When not bound by Allolactose, R binds to lac operon genes, preventing transcription

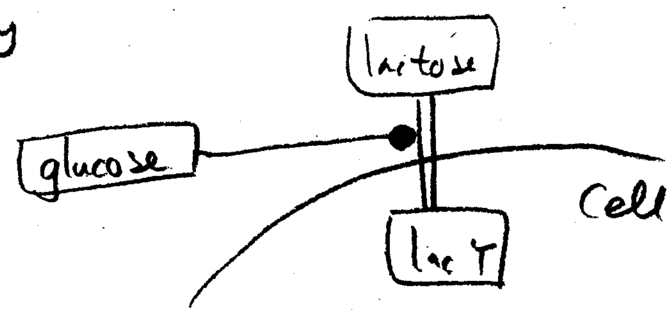


- External glucose inhibits production of cAMP

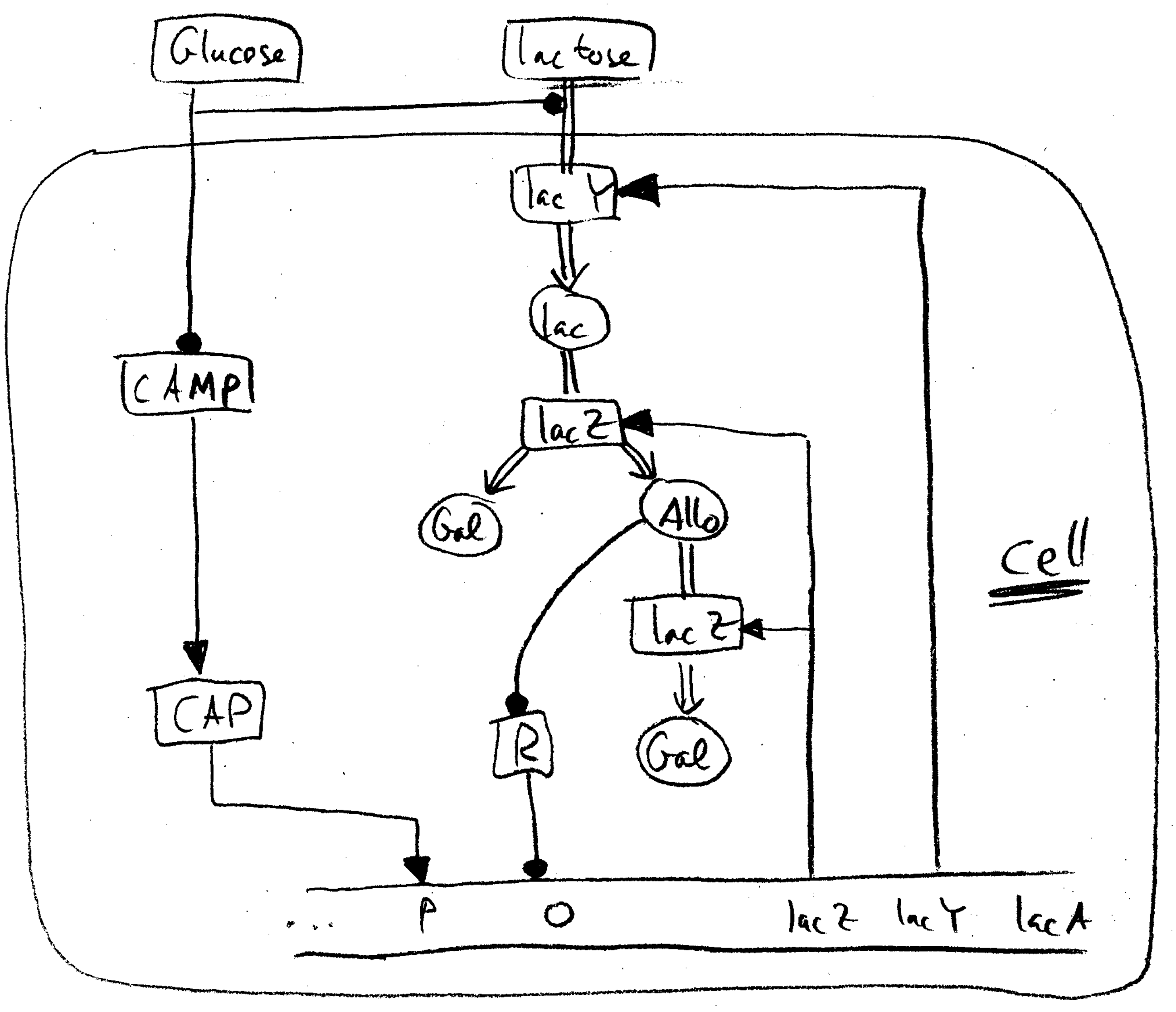
- When cAMP binds to CRP protein (CAP), it forms CAP-cAMP complex, activating the lac operon:



- External glucose inhibits uptake by permease proteins



We can put these motifs together for a "cartoon" of the lac operon regulatory mechanism:



Next: How to design a mathematical model that can reflect its qualitative behavior.