Biosphere

Communities and Ecosystems

Populations

Organisms

Cells

Pathways and Systems

Proteins

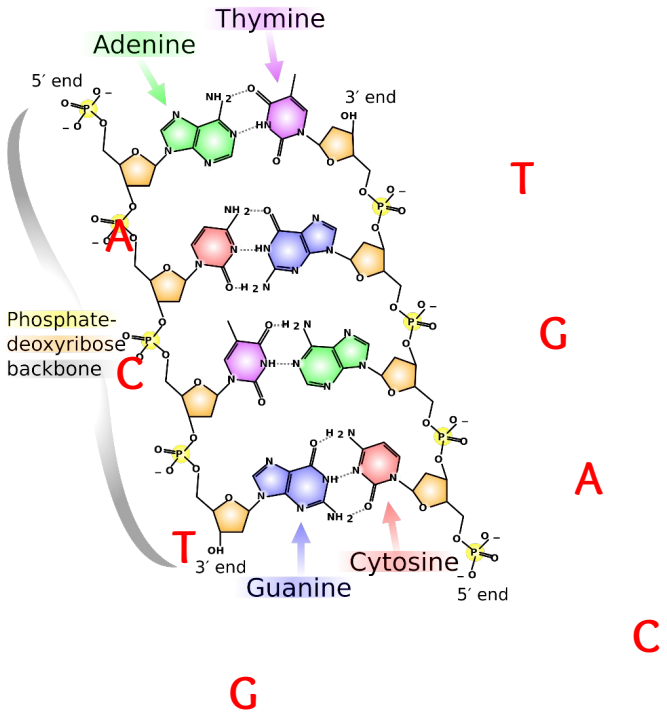
Genes

Genomes

Groupings

Building blocks

The DNA sequence of a gene




```
5´ - ATGCGTTACTTCGAAATGGCAACCCACTCGGGGACTTCCTCCAACGGTTGA- 3´
3´ - TACGCAATGAAGCTTTACCGTTGGGTGAGCCCCTGAAGGAGGTTGCCAACT- 5´
```

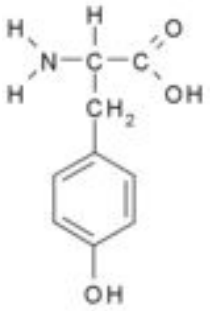

http://en.wikipedia.org/wiki/File:DNA_chemical_structure.svg (Madeleine Price Ball)

DNA to protein

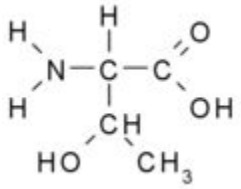
DNA is read in triplets

ATG CGT TAC TTC GAA ATG GCA ACC CAC TCG GGG ACT TCC TCC AAC GGT TGA

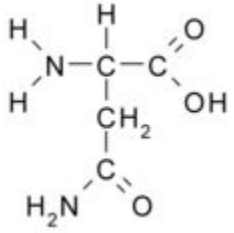
M A Y F E M A T H S G T S S N G *



Tyrosine



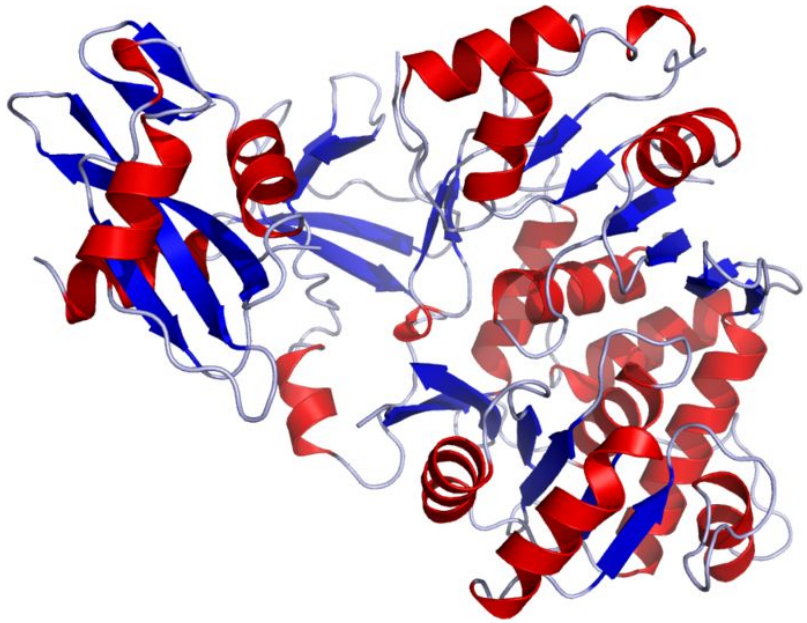
Threonine



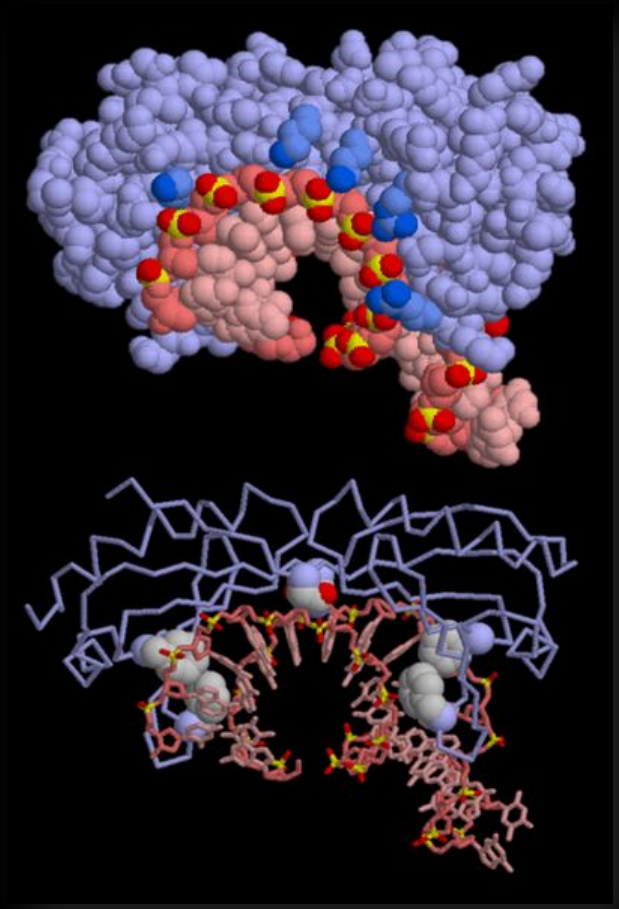
Asparagine

Protein sequence and structure

M A Y F E M A T H S G T S S N G *



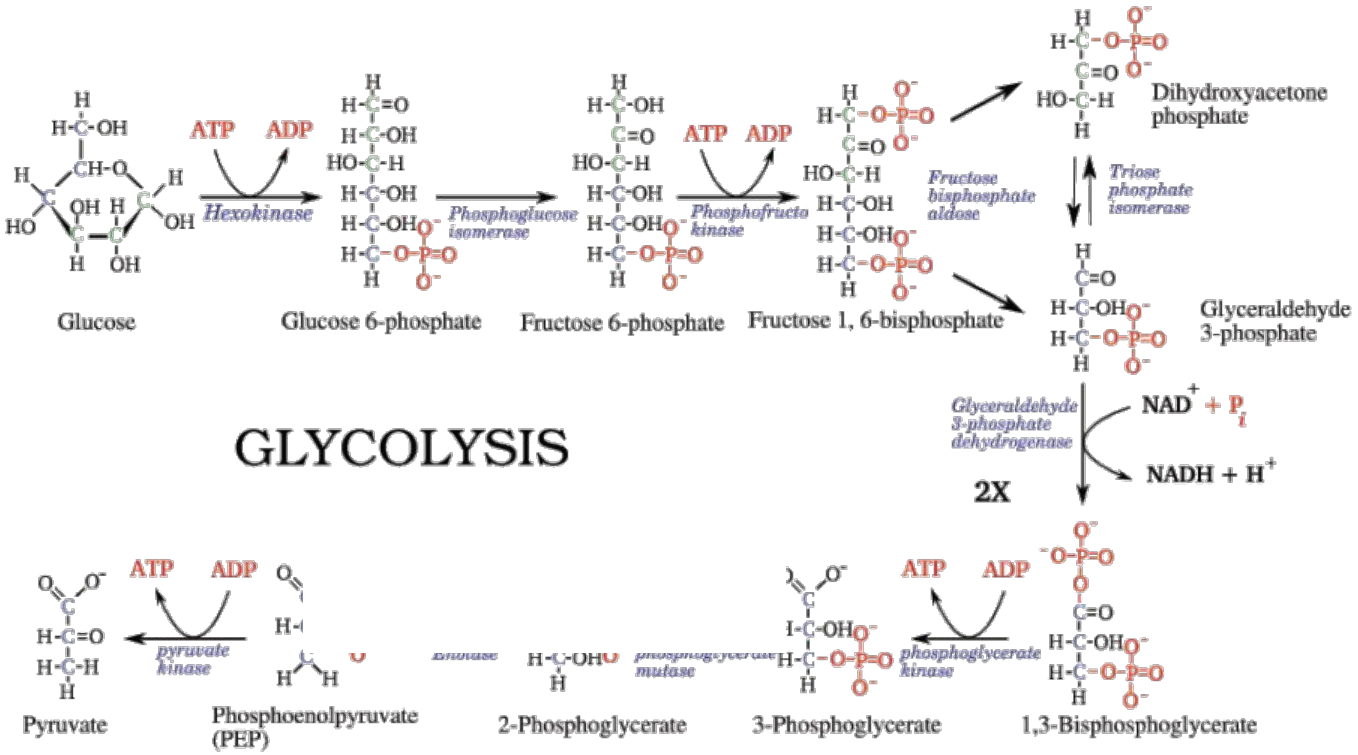
A DNA-protein complex



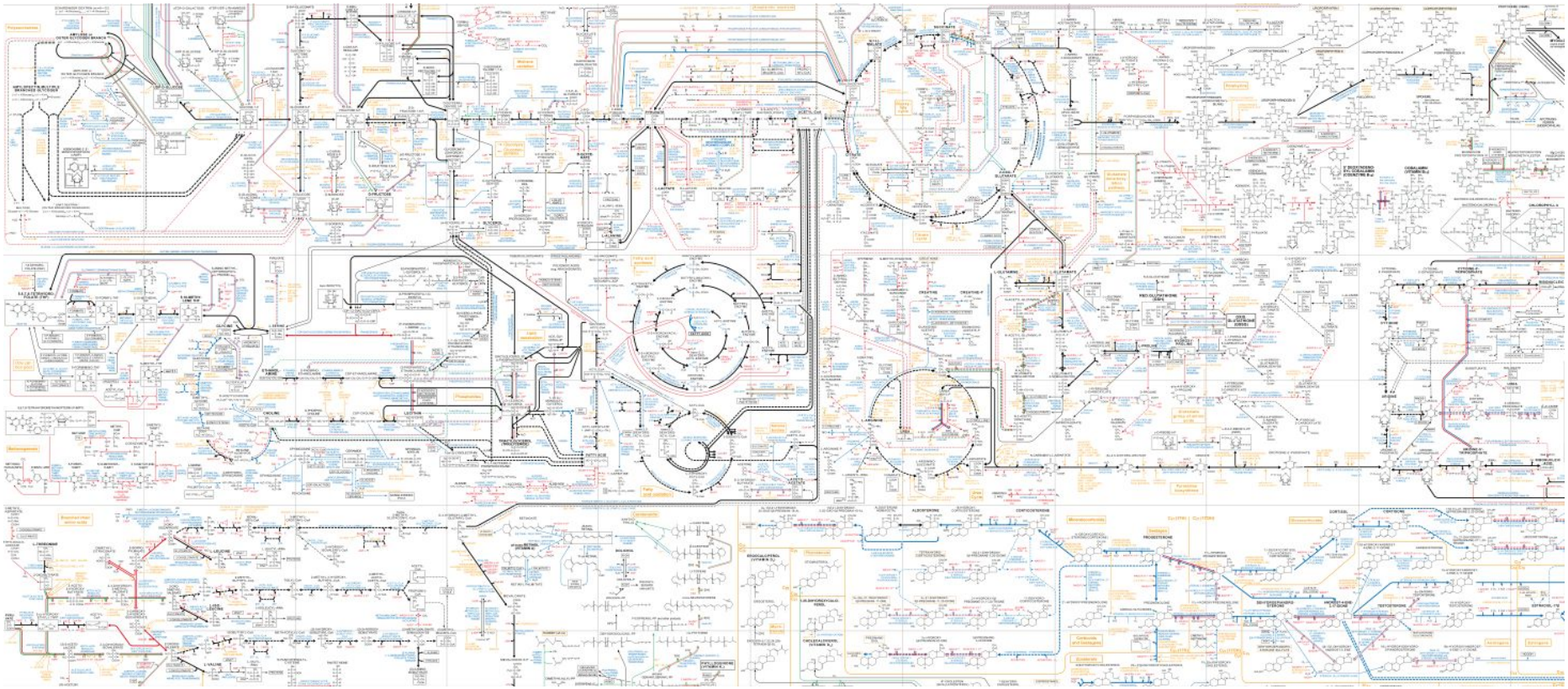
DNA-binding protein
("TATA-box binding protein")

DNA (note the recurring pattern;
yellow = phosphate)

Metabolism – Proteins working together

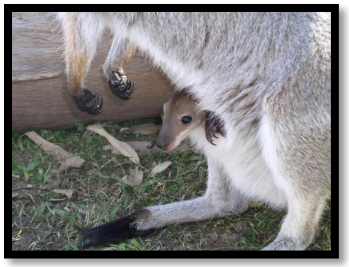


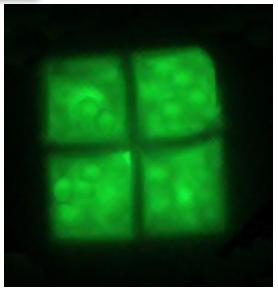
http://www.opencurriculum.org/5363/powering-the-cell-cellular-respiration-and-glycolysis/



Pathways (metabolism
+ self-replication
+ signalling)

=





Populations

Communities and Ecosystems

Overview

1. All living organisms have several key essential **properties**

2. Life can be viewed as a **hierarchical structure** with many levels of organization from **genome** (including genomic elements) to the **biosphere**

3. The levels we cannot observe with the naked eye are as (or more) **diverse** as the levels we can observe

02: Pathways & Central Dogma

CSCI4181/6802 Bioinformatics Algorithms

Finlay Maguire (finlay.maguire@dal.ca)

02a: Central Dogma

CSCI4181/6802 Bioinformatics Algorithms

Finlay Maguire (finlay.maguire@dal.ca)

Overview

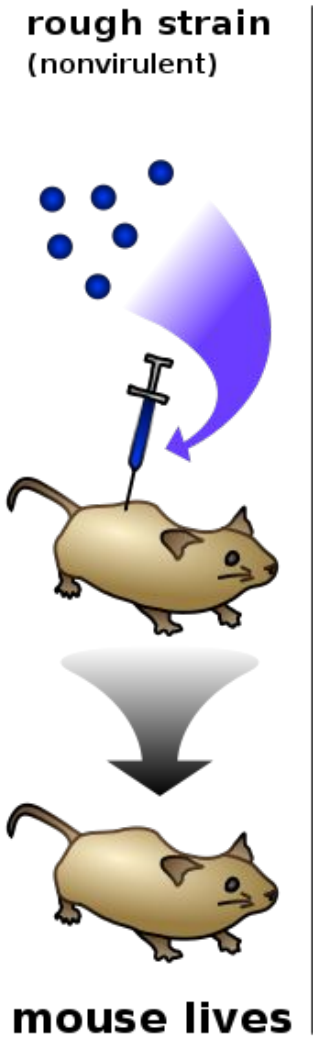
Essential processes for copying and interpreting biological information:

1. REPLICATION – the synthesis of a new DNA molecule from an existing template

2. TRANSCRIPTION – synthesis of an RNA molecule using a DNA template

3. TRANSLATION – synthesis of protein using an RNA template

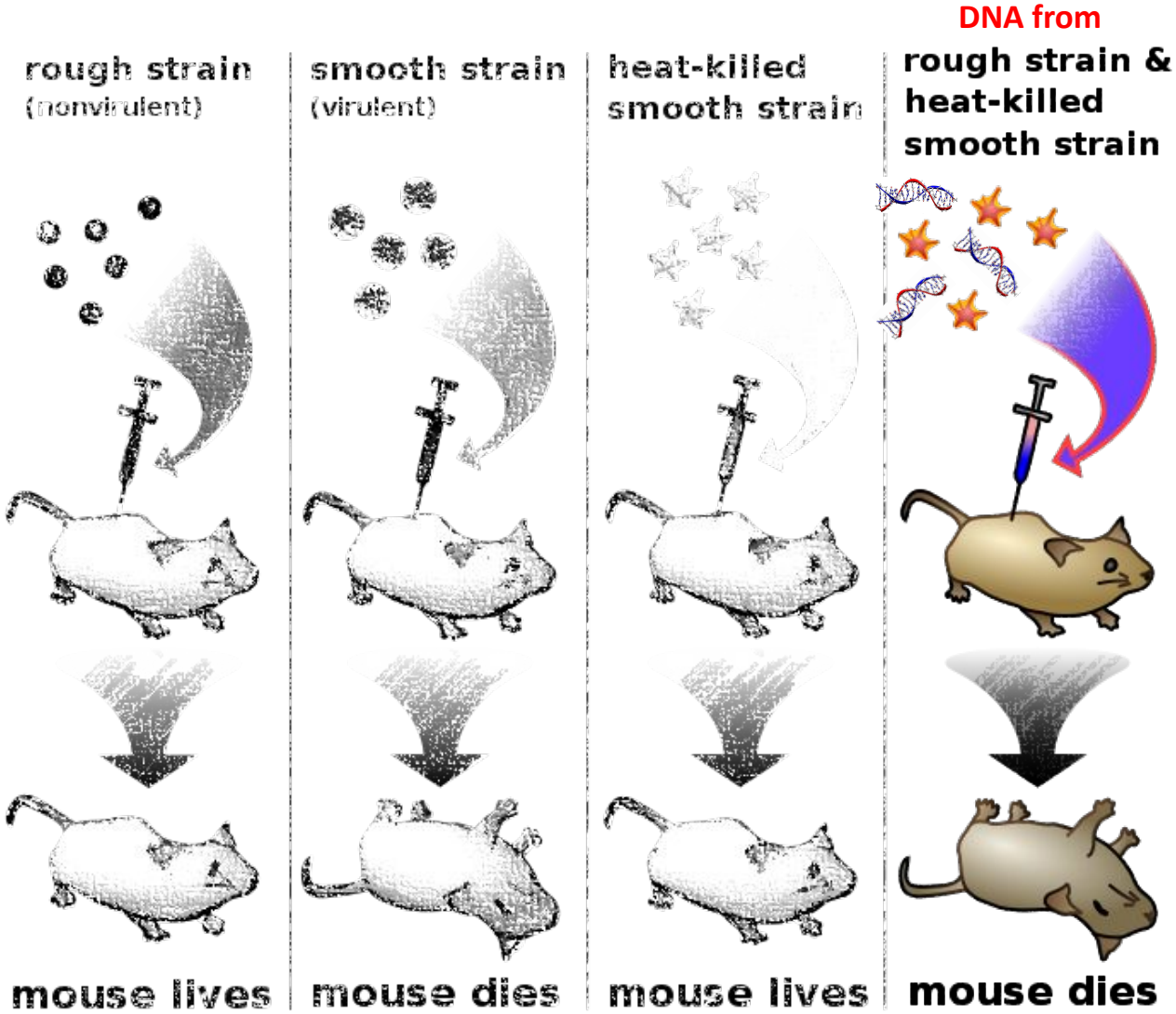
Griffith's experiment – something's moving between organisms



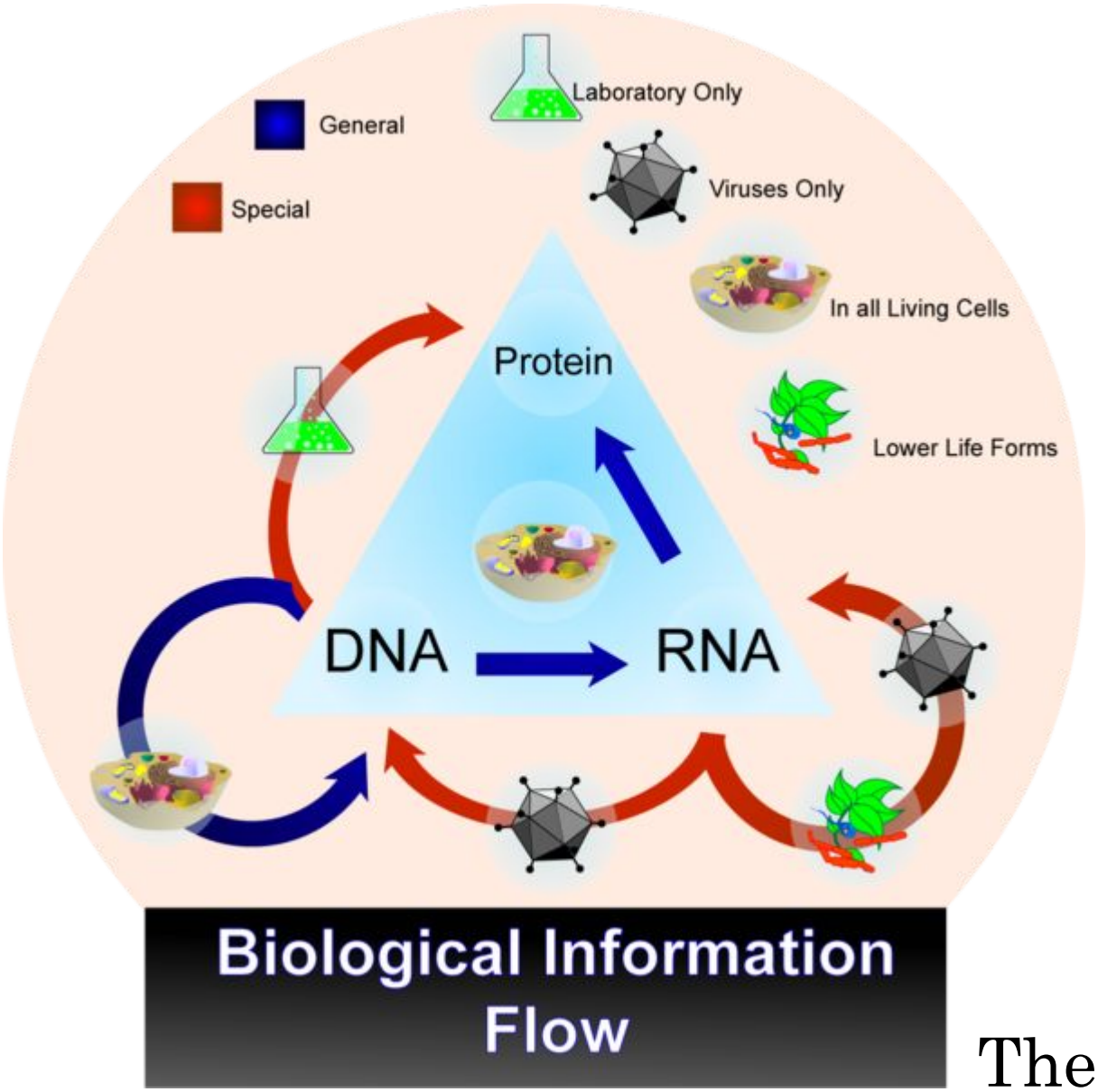
Frederick Griffith (1928) *J Hygiene*

Image: Madeleine Price Ball, Wikimedia Commons

Avery-MacLeod-McCarty experiment – that something is DNA

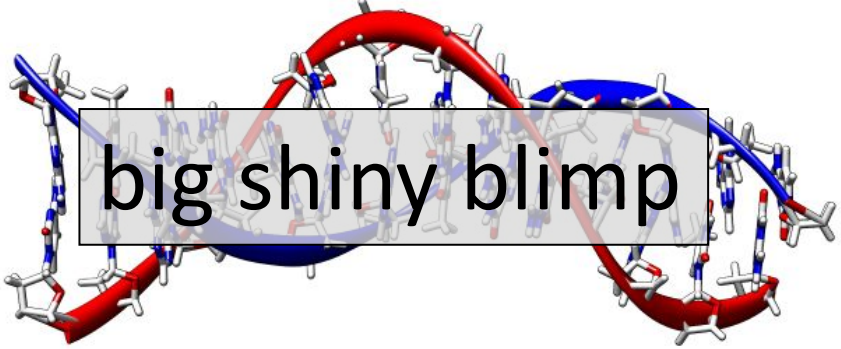


Avery et al. (1944) *J Exp Med*



The 'Central Dogma' – From Information to Function

Mike Jones - http://commons.wikimedia.org/wiki/File:CDMB2.png



Replication

big shiny blimp
big shiny blimp
big shiny blimp
big shiny blimp
big shiny blimp

Transcription

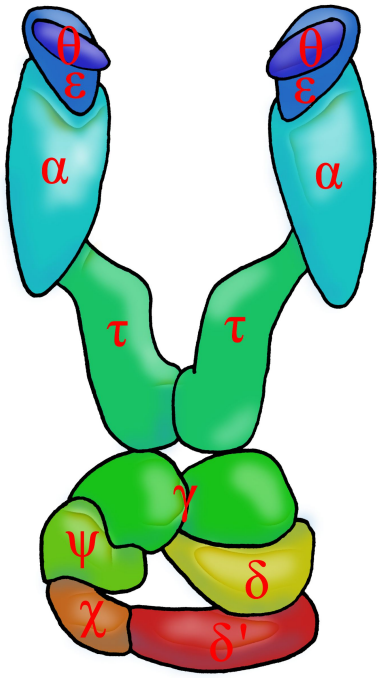
BIG SHINY BLIMP

Translation

-... .. --. ... . -. -.-- -... .-.. .. -- .--.

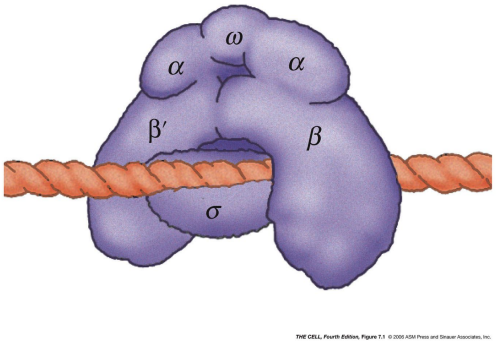
Key steps and commonalities

All three processes are carried out by **multi-protein complexes** (sometimes with extra bits thrown in)

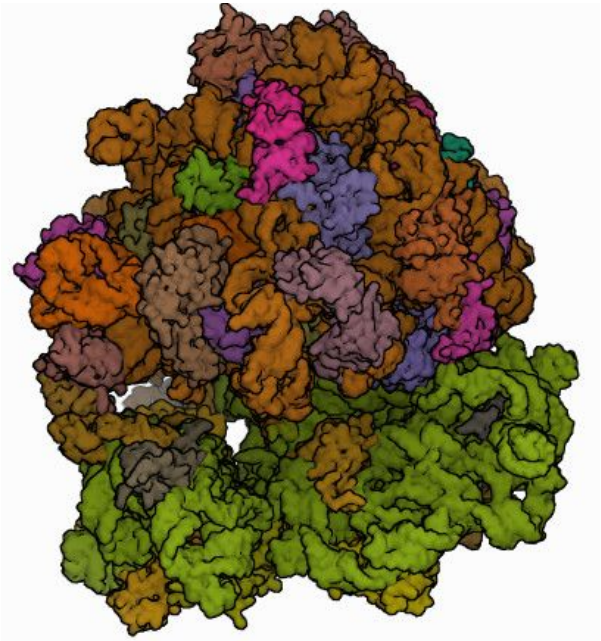


DNA polymerase III

https://en.wikipedia.org/wiki/DNA_polymerase_III_holoenzyme



RNA polymerase



Ribosome

https://www.rcsb.org/structure/5V93

Key steps and commonalities

All processes and phases are **regulated** and have three phases:

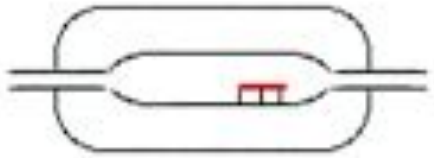
https://www.pnas.org/content/98/1/5/tab-figures-data

Key steps and commonalities

All processes and phases are **regulated** and have three phases:

- Initiation

Let's make RNA!



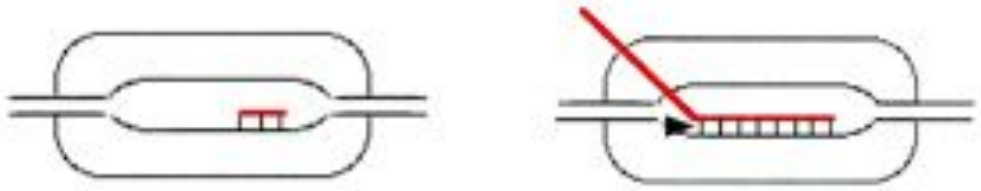
https://www.pnas.org/content/98/1/5/tab-figures-data

Key steps and commonalities

All processes and phases are **regulated** and have three phases:

- Initiation
- Elongation

Let's make RNA!



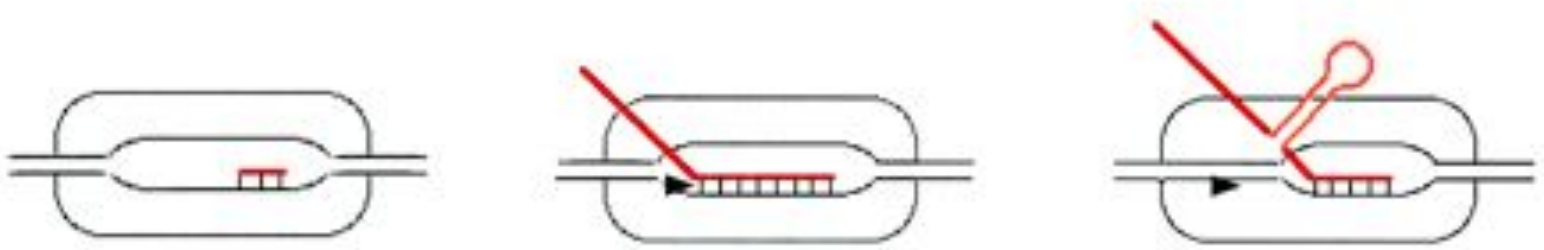
https://www.pnas.org/content/98/1/5/tab-figures-data

Key steps and commonalities

All processes and phases are **regulated** and have three phases:

- Initiation
- Elongation
- Termination

Let's make RNA!



https://www.pnas.org/content/98/1/5/tab-figures-data

Key steps and commonalities

Processes in **eukaryotes** tend to be more complex than those in **prokaryotes**

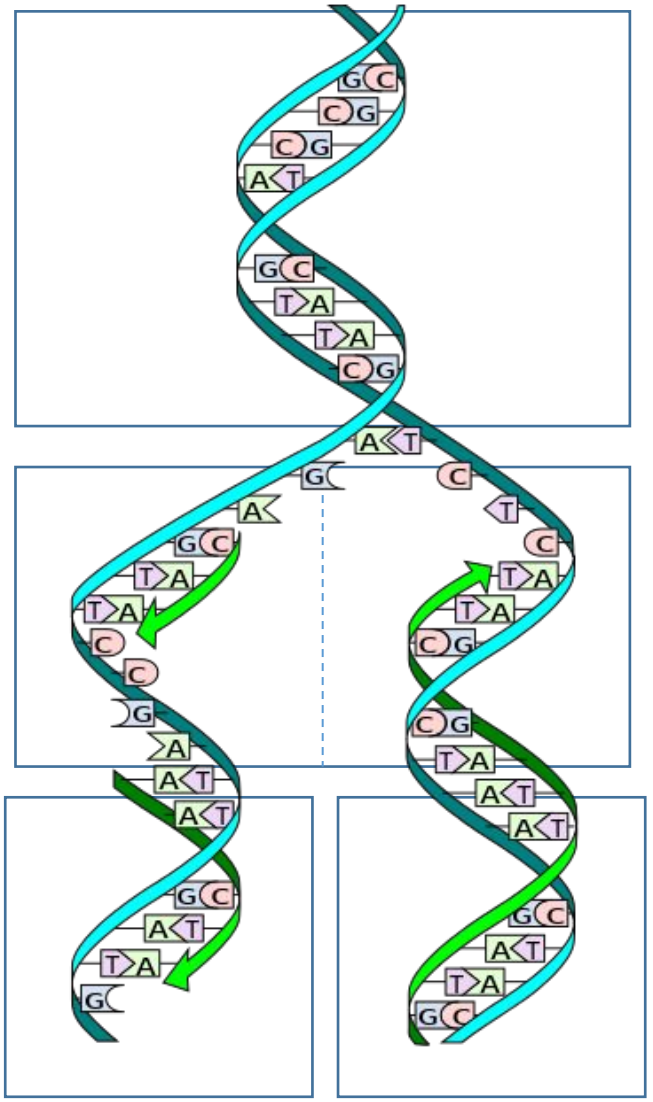
REPLICATION:
from DNA to more DNA

The replication process

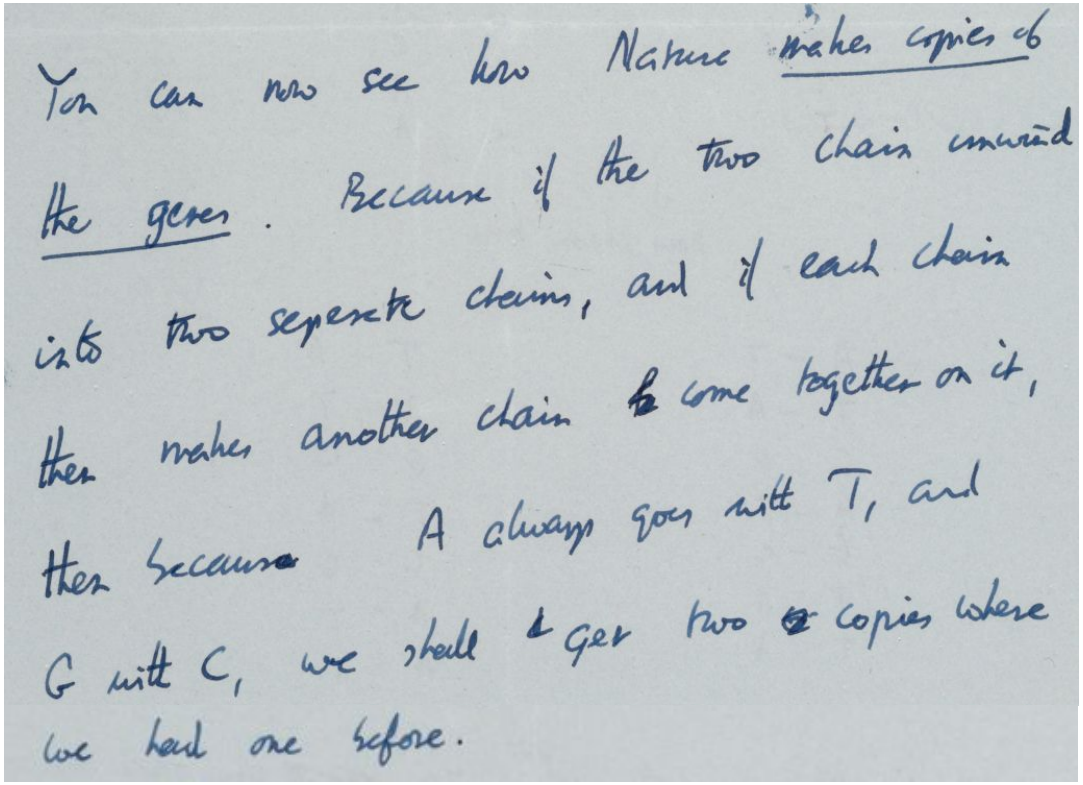
(1) DNA is UNWOUND

(2) A copy of EACH STRAND is made independently

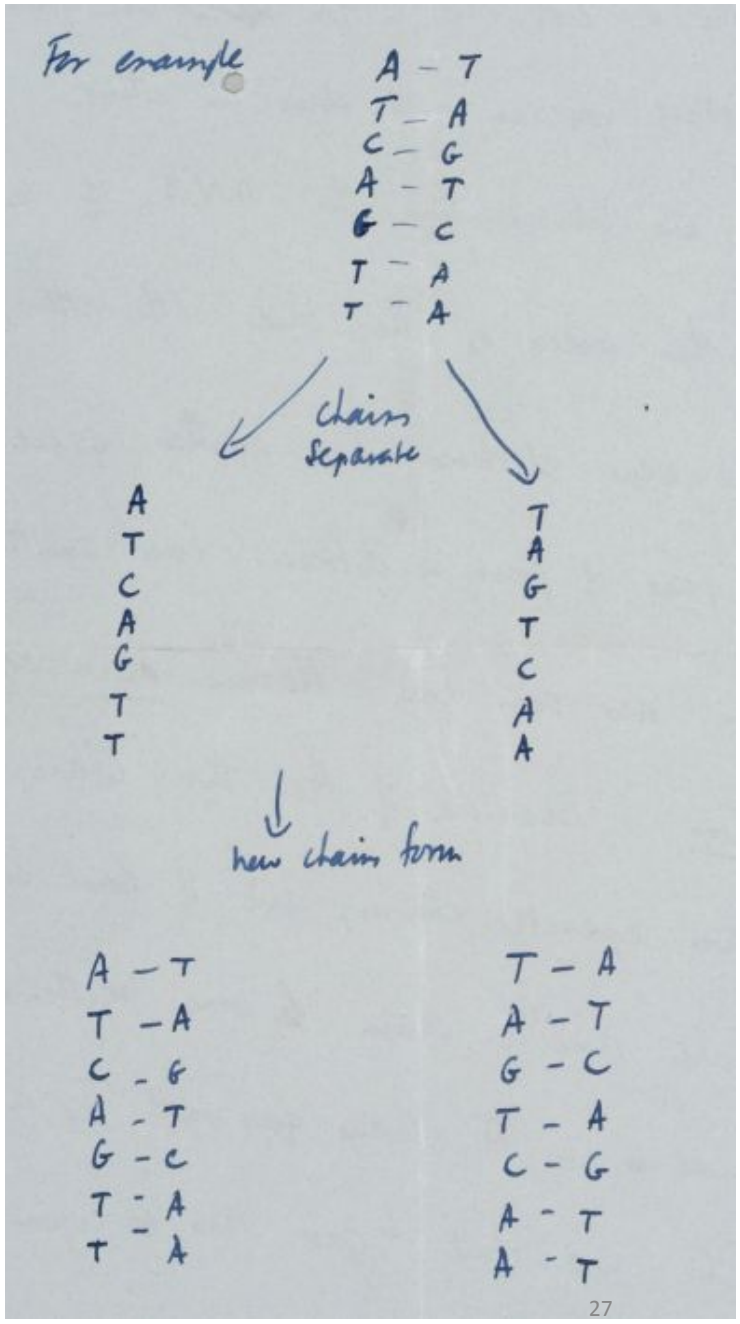
(3) Each copy is packaged into a cell



http://en.wikipedia.org/wiki/File:DNA_replication_split.svg


You can now see how Nature makes copies of the genes. Because if the two chain unwind into two seperate chains, and if each chain then makes another chain to come together on it, then because A always goes with T, and G with C, we shall get two copies where we had one before.

From Francis Crick's letter to his son Michael, 1953
$5.3M at auction

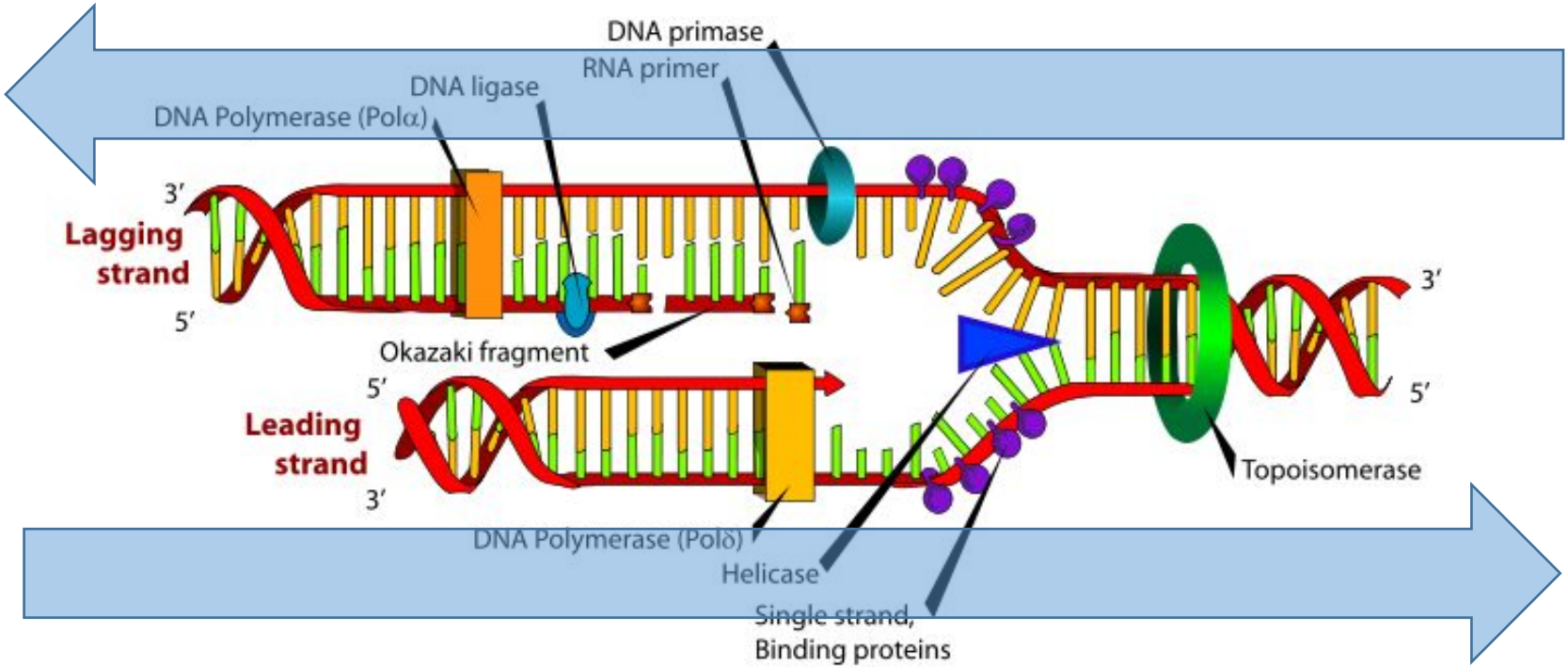


http://voices.nationalgeographic.com/2013/04/11/francis-cricks-letter-to-son-describing-dna-auctioned/

Replication proceeds in the 5′ – 3′ direction

But DNA is antiparallel (the two strands point in opposite directions)!

So replication proceeds *differently* on each strand (leading strand is *way* easier)



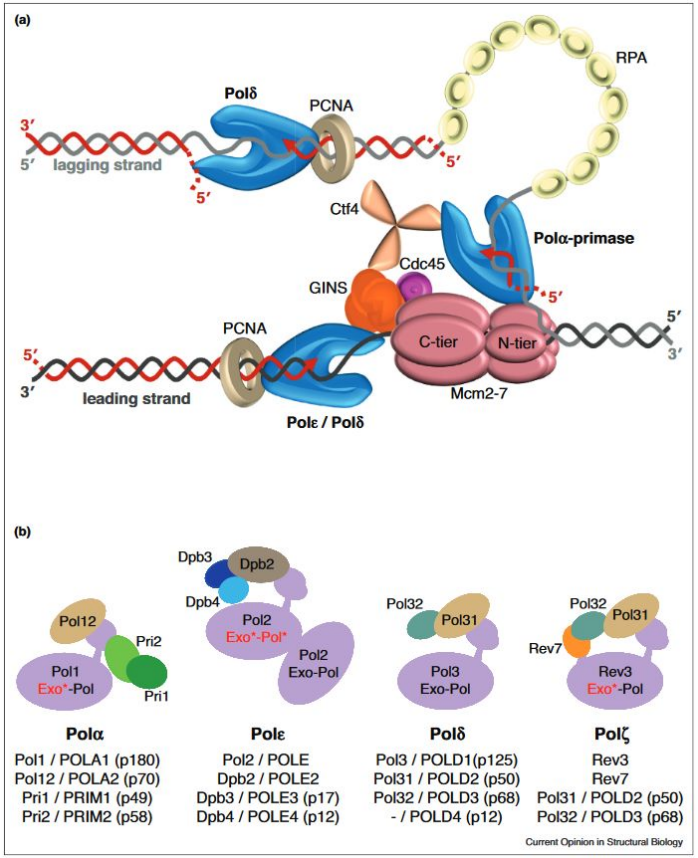
http://en.wikipedia.org/wiki/File:DNA_replication_en.svg

DNA polymerase III (bacteria): >6 different subunit types

http://www2.mrc-lmb.cam.ac.uk/groups/mlamers/Site/Home.html

Eukaryotes:
At least 14 different DNA polymerase complexes

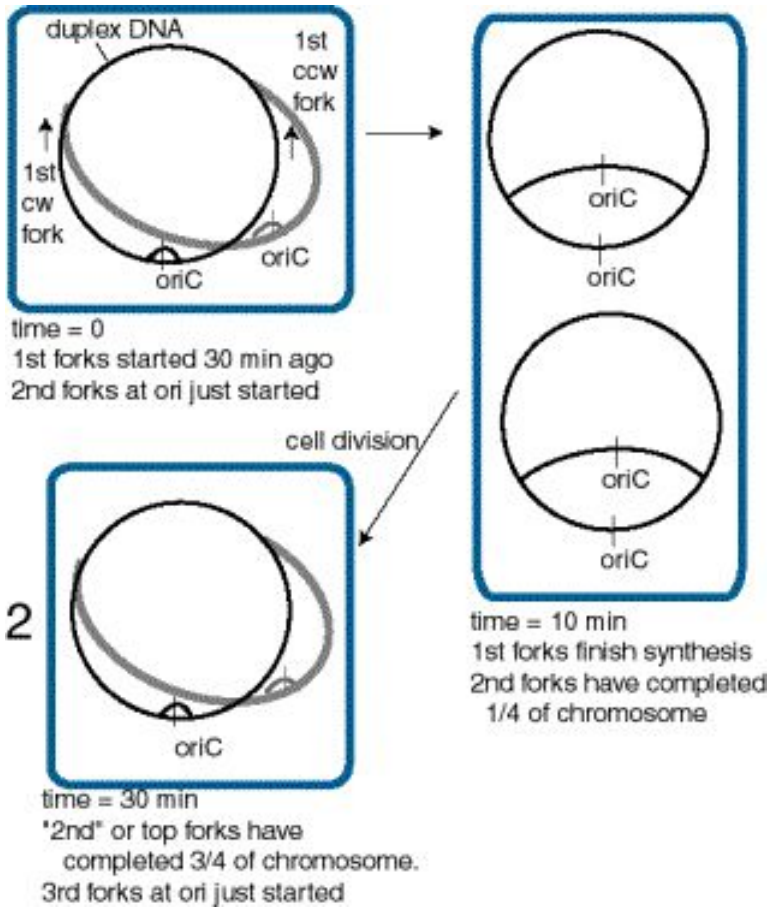
DNA polymerase	Function
α	DNA replication/priming
β	Base excision repair
γ	Mitochondrial DNA replication
δ	Chromosomal replication/excision repair
ε	Chromosomal replication/repair
ζ	REV3: error-prone bypass synthesis
η^*	RAD30: error-free bypass of UV-induced CPDs
θ	DNA repair
ι^*	RAD30B: bypass synthesis
κ^*	DinB: bypass synthesis
λ	Base excision repair
μ	Non-homologous end joining
σ	Sister chromatid cohesion
REV1*	Deoxycytidyl transferase



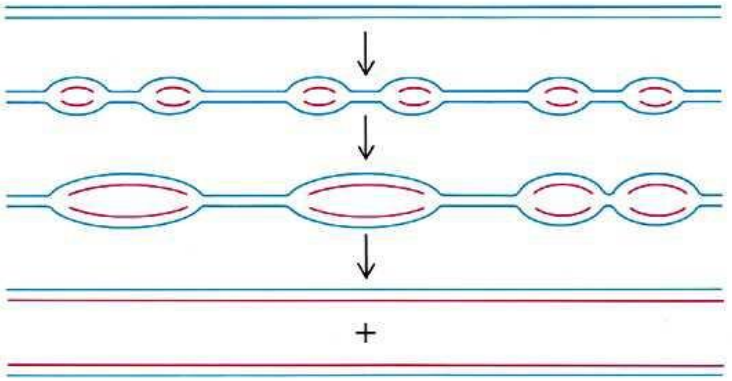
Carty et al. (2003) *Biochemical Society Transactions*

Jain et al. (2018) *Current Opinion in Structural Biology*

Replication



E. coli: ~40 minutes



Humans: Hours

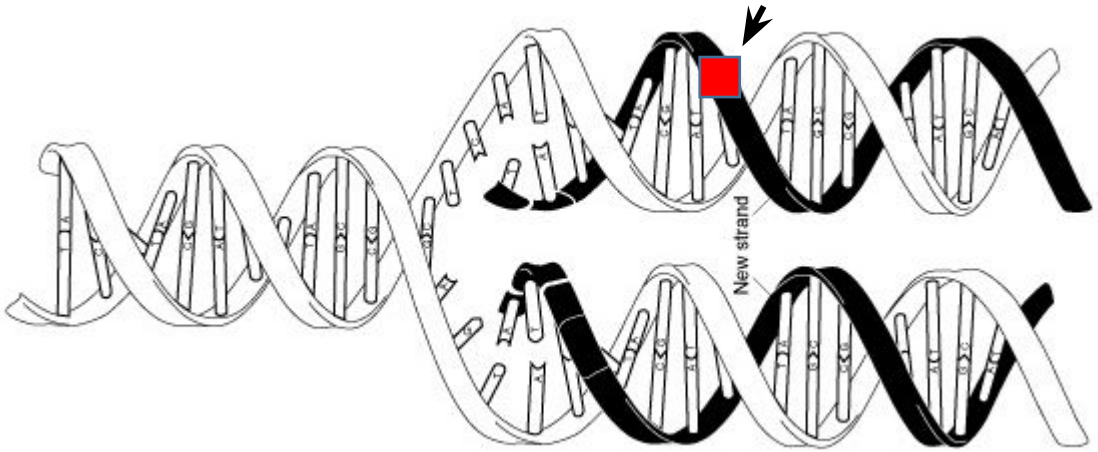
http://www.personal.psu.edu/rch8/workmg/Replication2Ch6_files/image049.png

http://www.mun.ca/biochem/courses/3107/images/Stryer/Stryer_F37-13.jpg

Mutations

When DNA polymerase makes an error that is not caught by the 'proofreading' mechanism, a *mutation* results

G inserted where A is supposed to go (oops!)



Fidelity

~ 10^{-8} error rate (varies by species)

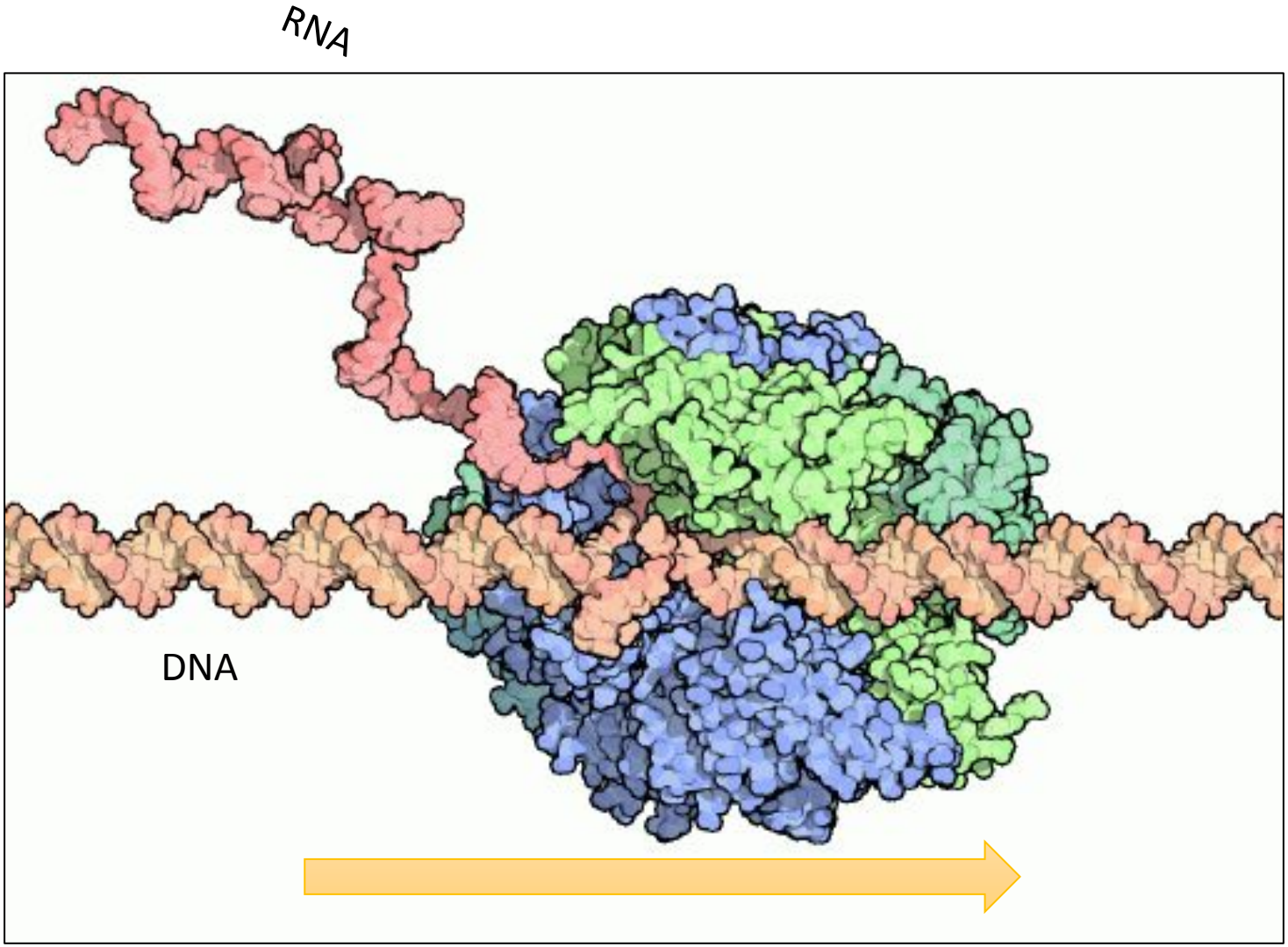
Without proofreading, error rate ~ 10^{-5}

Some viruses: 10^{-3}

TRANSCRIPTION

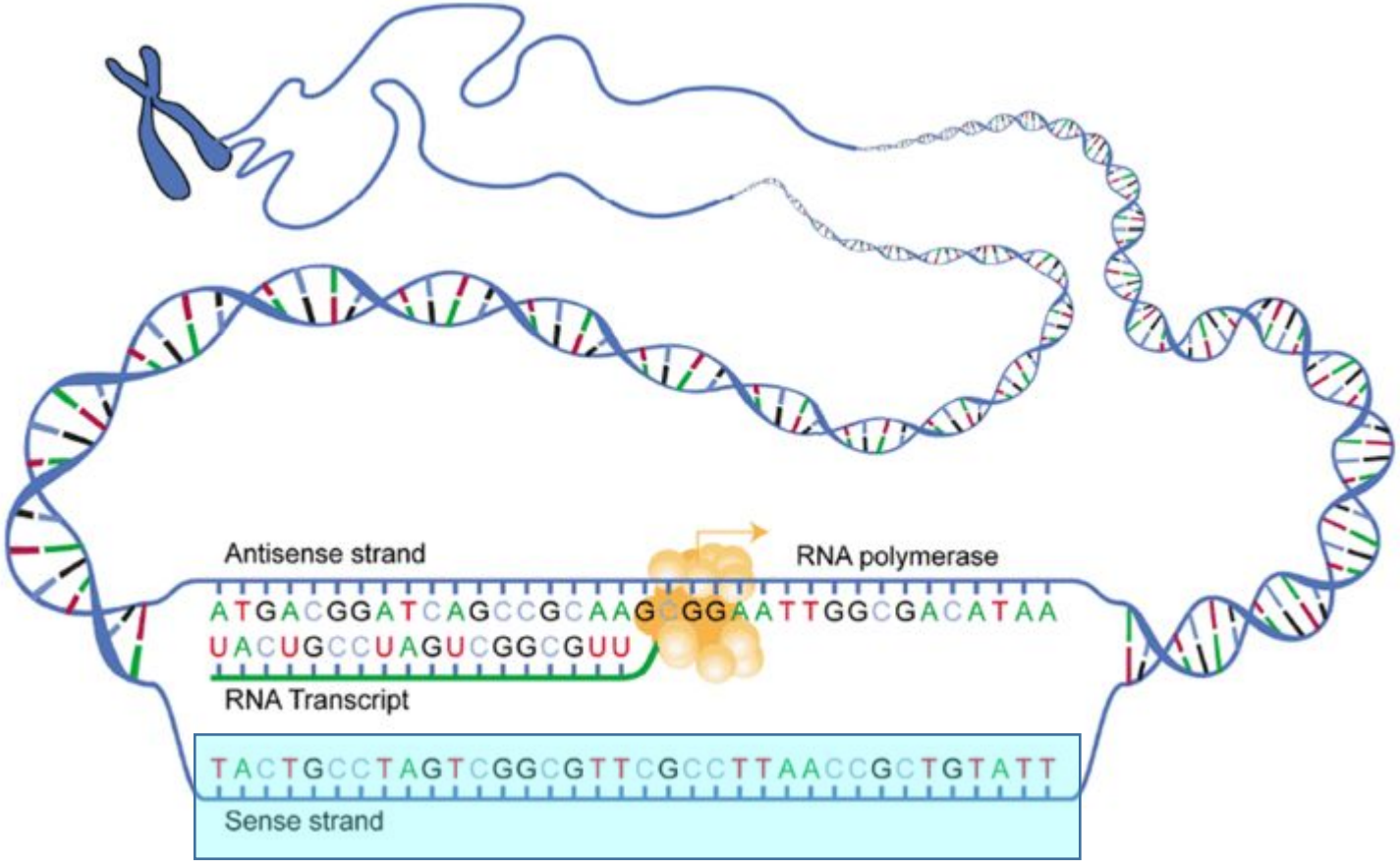
: from DNA to RNA

What transcription looks like



http://www.rcsb.org/pdb/101/motm.do?momID=40

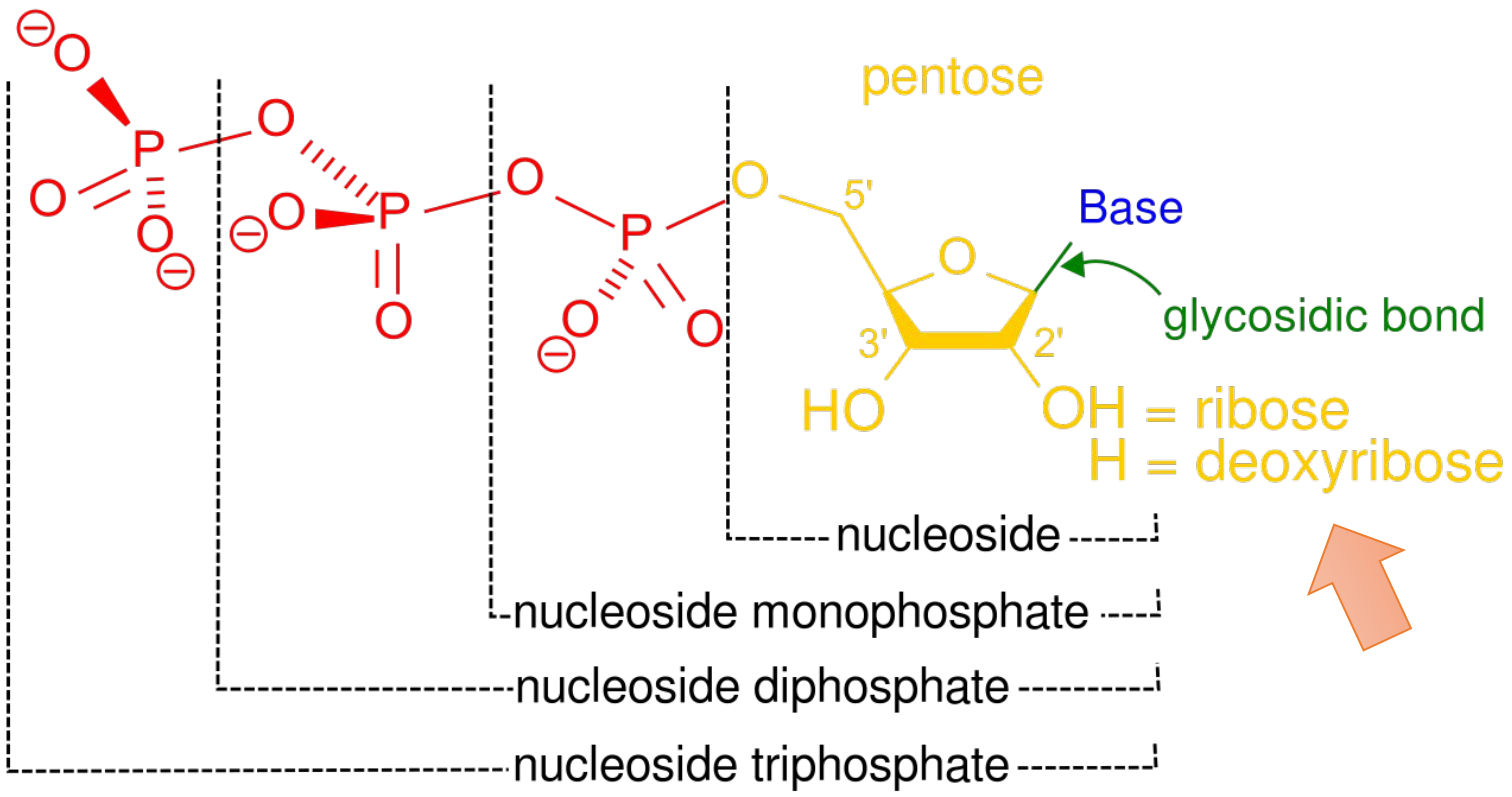
Transcription



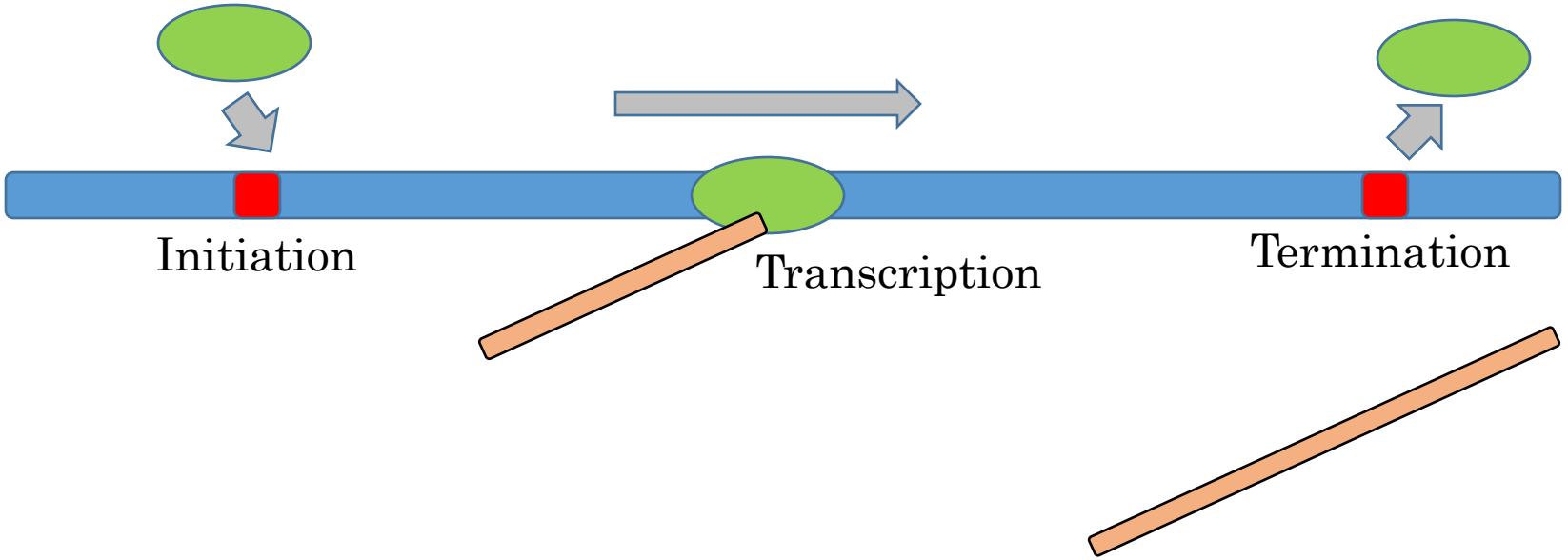
(This is the one we typically write out)

http://commons.wikimedia.org/wiki/File:DNA_transcription.svg

One silly oxygen, a whole world of difference



http://en.wikipedia.org/wiki/File:Nucleotides_1.svg

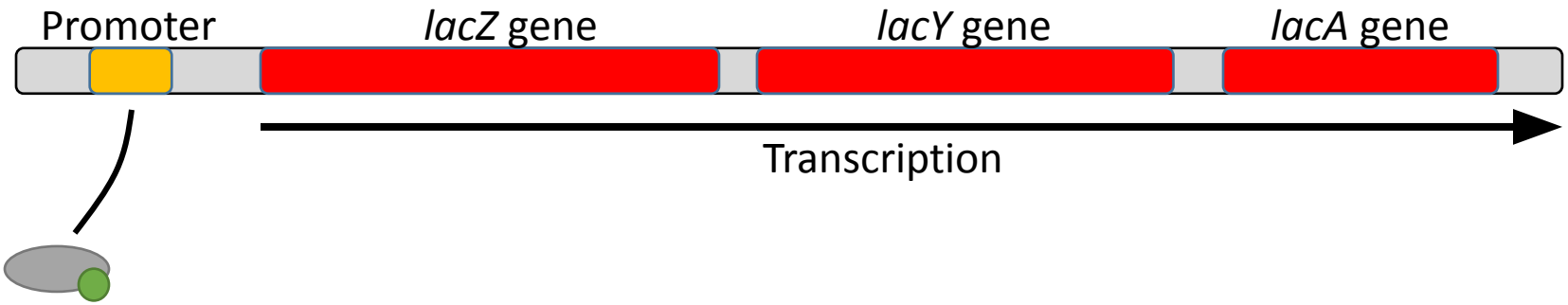

Initiation
Transcription
Termination

Regulating transcription: the promoter

lac operon in *E. coli* – breakdown of lactose sugar (simplified)

operon = *set of genes transcribed together but translated independently*

RNA polymerase must recognize a stretch of DNA upstream of the genes to be transcribed – the **promoter**



RNA polymerase with sigma factor

Sigma factor binding sites in front of different genes

```
tttatgttgctttttgtaaacagattaacacctcgtcaaaatcctgctattctgcccgtTgcggtactgggcatttaccc
atcgtggtaagacctgccgggatttagttgcaaatttttcaacattttatacactacgaaAaccatcgcgaaagcgagttt
gttcgtgaatttacaggcgttagatttacatacatttgtgaatgtatgtaccatagcacgAcgataatataaacgcagcaa
aaagtcagttaatgtaatgcctcctactgaccaaagaatacttgcacttaaggttcagtaTaaaagggcatgataatttac
ttacgaggttttaattctgcctctttcaacccgcgtcaaaataaaacagtagaatattaaTctttttttgtgtttatgtgc
gatttcaaattggtcaatggtcaaaagttaataaacccattgctgcgtttatattatcgtCgtgctatggtacatacattc
cgcaaagctgaccgcacaaaaggggagtgcttttctgtgcttagcggttagaatagtctcAtgactatatctggagttgac
aataaccacactgtgaatgttgtctttaatcaattgtaagtgcatgtaaaataccactttAgagttagtcagtatcttcct
atactaaacaaaactgccaataccctacatttaacgcttatgccacatattattaacatCctacaaggagaacaaaagca
aaaaattaaagcgcaagattgttggtttttgcgtgatggtgaccgggcagcctaaaggctAtccttaaccagggagctgat
aacgtaaaaaatcgttgcgcaatcgtggatttttaccctgctttgtttttataatggtgcGcactttatatccagaaaaa
ttcagtgataattatcacatttcaattgcacattaatggatattctttaataatctcgcgAcgtttctttatgataaataa
caacaacggttcagtgataattatcacatttcaattgcacattaatggatattctttaatAatctcgcgacgtttctttat
ttcaattgcacattaatggatattctttaataatctcgcgacgtttctttatgataaataAtaatcaaattgataaaatca
ttgttatctagttgtgcaaaacatgctaatgtagccaccaaatcatactacaatttattaActgttagctataatggcgaa
ttcaaaaatgtttgtttttcacgcgctttacagcccgaaaaggccggaagatacttgcccGcaacgaagattccttcataa
agcttgcgtcaatgggcaaggtgggcttgcatttgcttaatagaaaggcgttaataggcaAaacgaaatgaaacgaaagtt
tcgttttatttctttttctccattgaactttcagtttcttttctatagattttaatcaaCgaaagacatcaccaagtgaa
atcgatttgataatggaaacgcattagccgaatcggcaaaaattggttaccttacatctcAtcgaaaacacggaggaagta
taggtaagagcttagatcaggtgattgccctttgtttatgagggtgttgtaatccatgtcGttgttgcatttgtaagggca
gataagaatgttttagcaatctctttctgtcatgaatccatggcagtgaccatactaatgGtgactgccattgatggaggg
taagaaactaatattagacgtaaatattgaaatttttatatttttttcttatttaggctttGcatttggcaaaattttgagg
agaaactaatattagacgtaaatattgaaatttttatatttttttcttatttaggctttgcAtttggcaaaattttgaggca
```

???

(more on this in the machine-learning section)

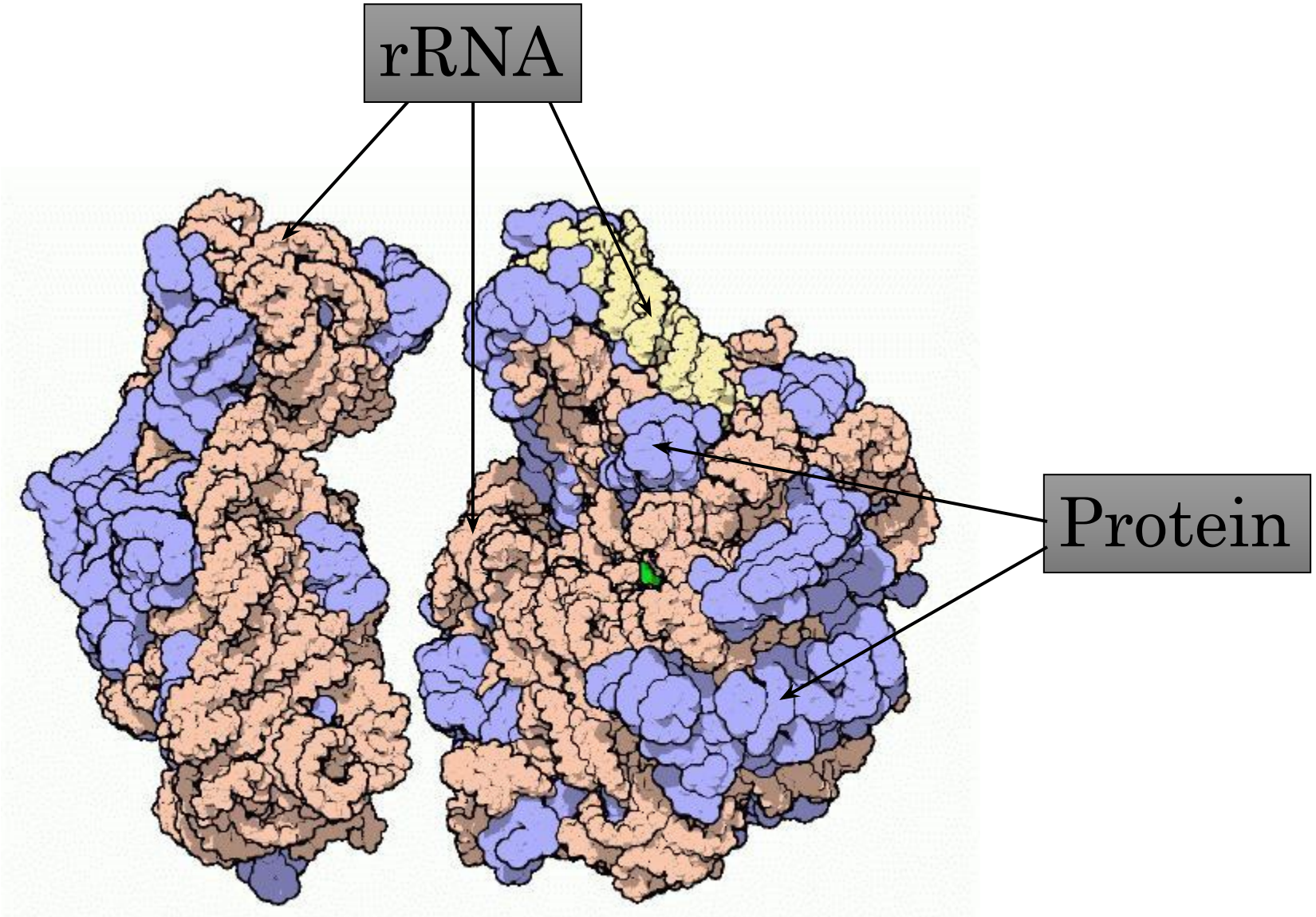
Extracted from RegulonDB (http://regulondb.ccg.unam.mx/)

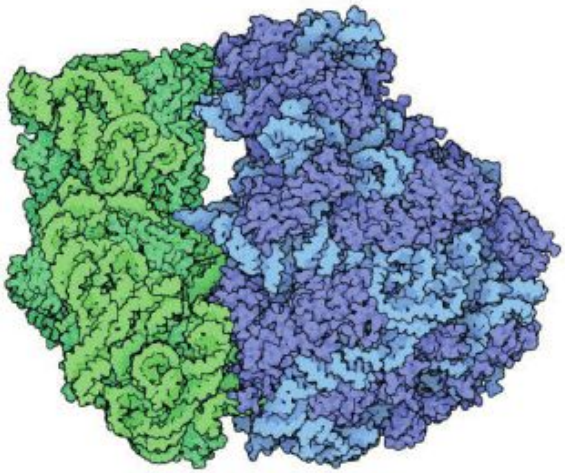
The fate of transcripts

- Messenger RNA (mRNA) – ultimately translated to protein
- Ribosomal RNA (rRNA) – important to translation
- Transfer RNA (tRNA) – also important to translation
- About 20 other types at last count!

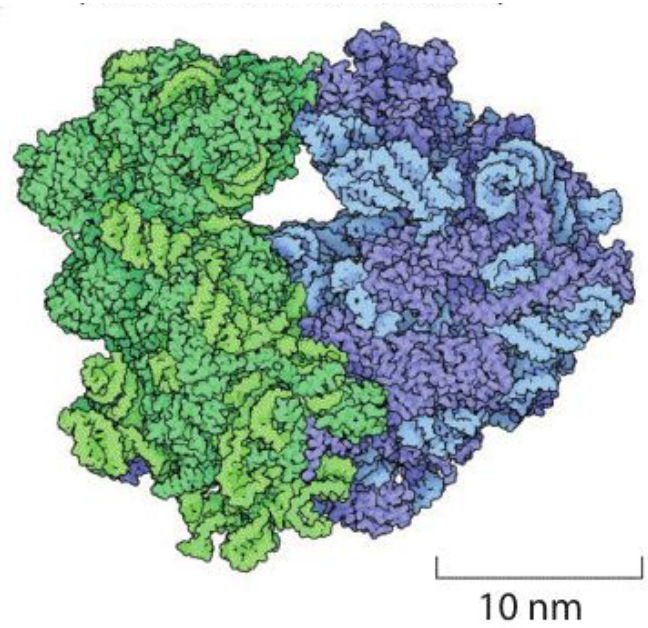
TRANSLATION:
from mRNA to protein

Key player: the ribosome



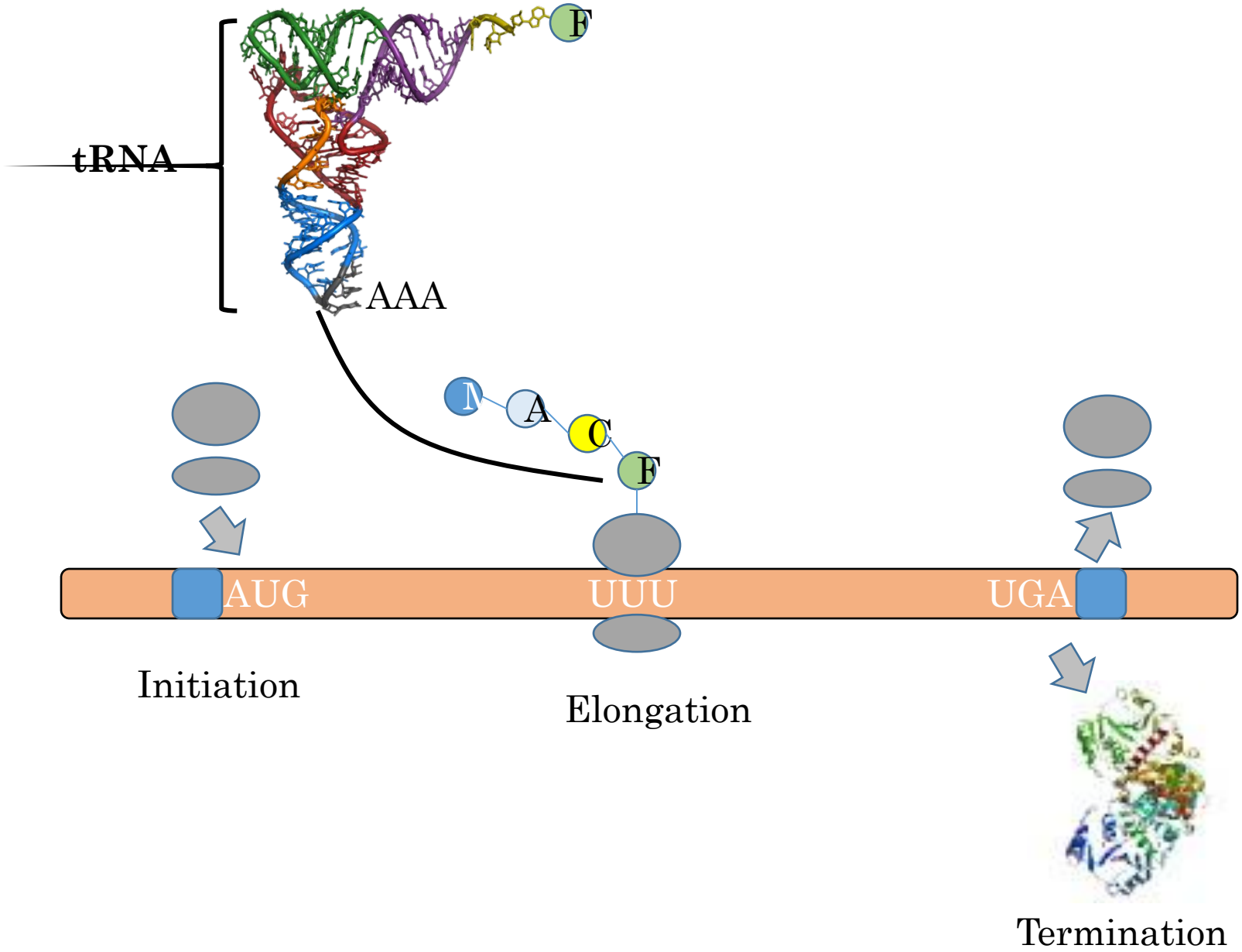


Prokaryotic ribosome
~45 proteins
3 rRNAs



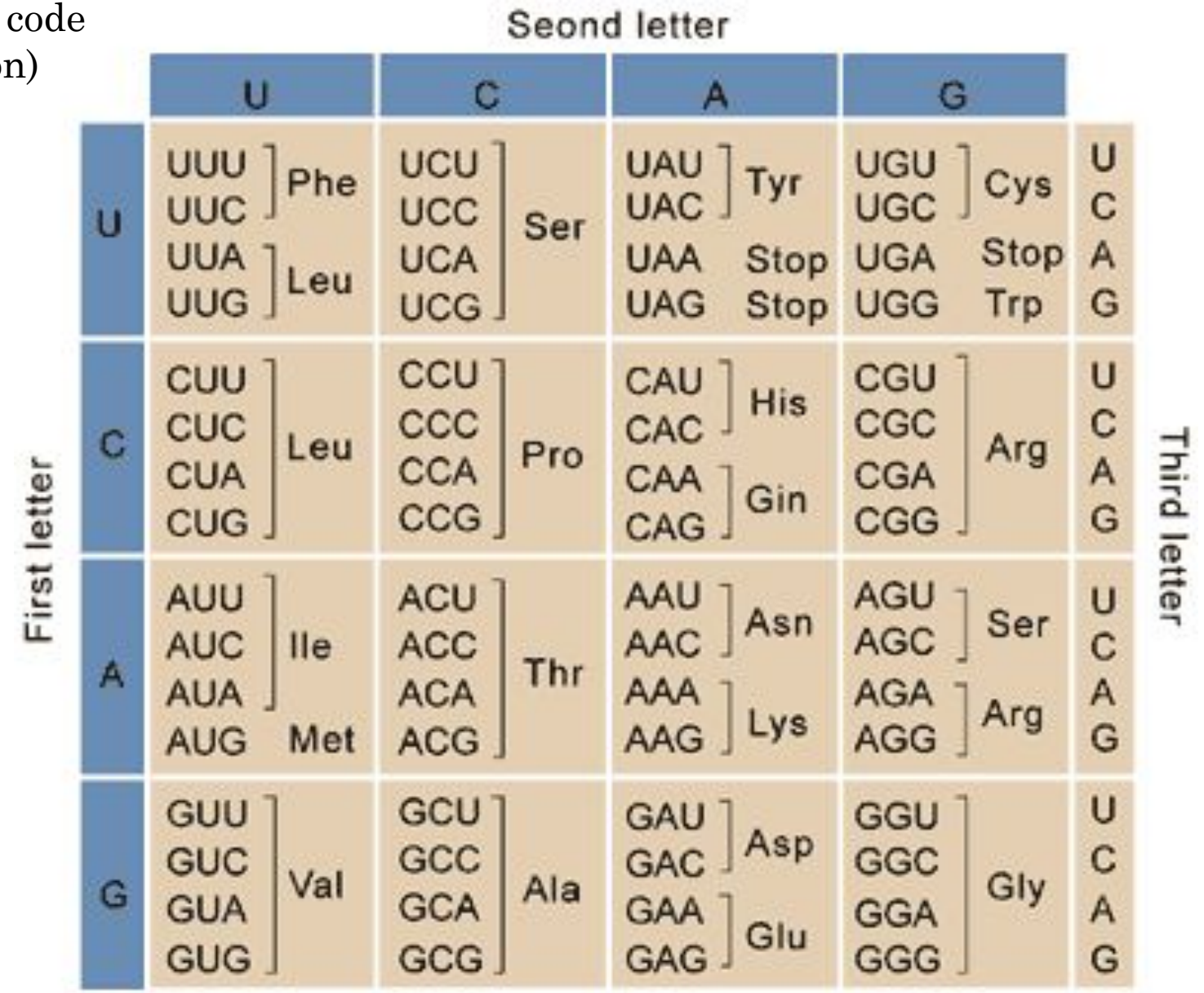
Eukaryotic ribosome
~80 proteins
4 RNAs

http://book.bionumbers.org/how-big-are-the-molecular-machines-of-the-central-dogma/

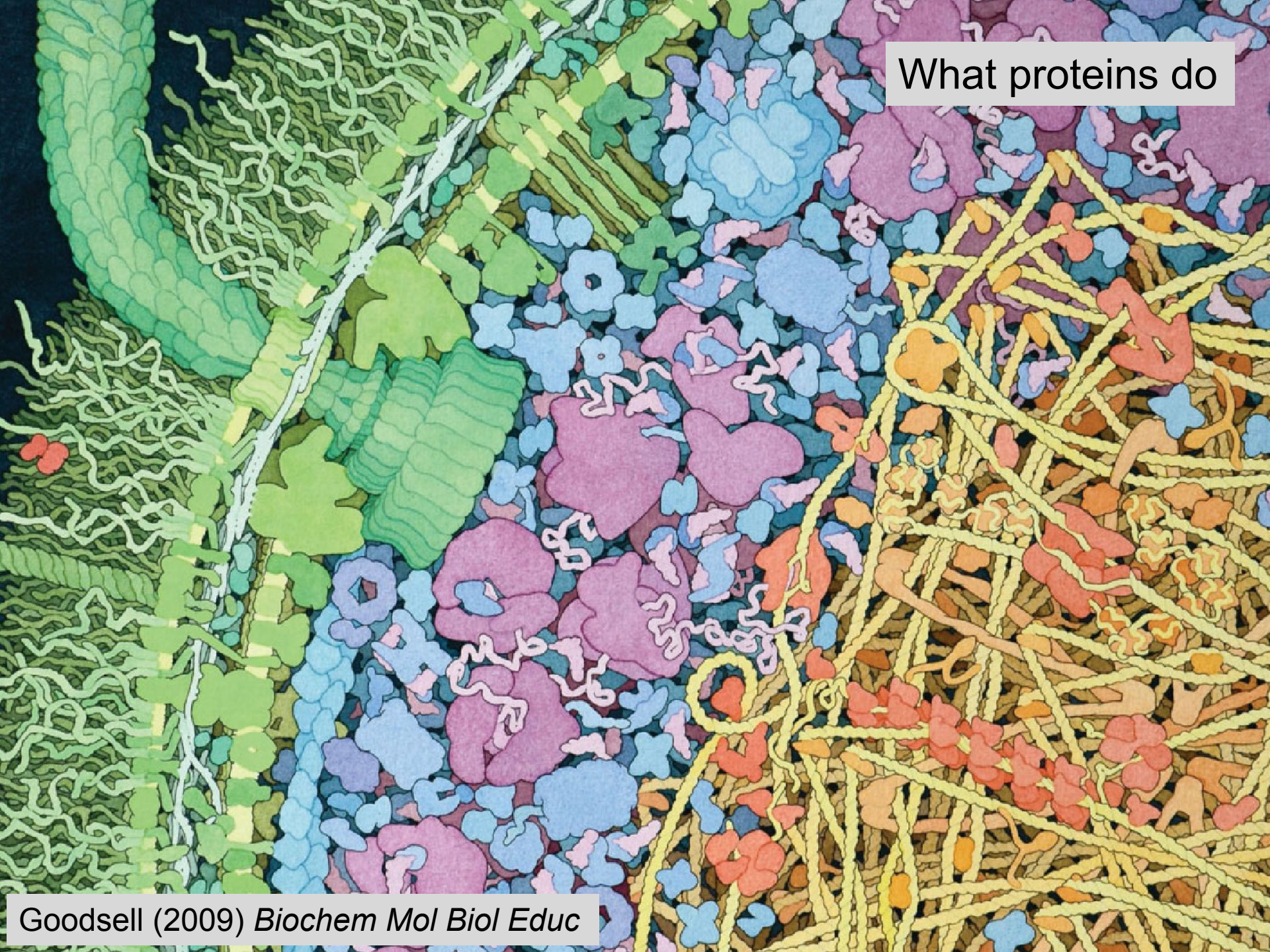


http://commons.wikimedia.org/wiki/File:TRNA-Phe_yeast_1ehz.png

The triplet genetic code (standard version)



http://upload.wikimedia.org/wikipedia/commons/2/21/06_chart_pu3.png



Goodsell (2009) *Biochem Mol Biol Educ*

Summary

Essential processes for copying and interpreting biological information:

1. REPLICATION – the synthesis of a new DNA molecule from an existing template

2. TRANSCRIPTION – synthesis of an RNA molecule using a DNA template

3. TRANSLATION – synthesis of protein using an RNA template

02b: Pathways

CSCI4181/6802 Bioinformatics Algorithms

Finlay Maguire (finlay.maguire@dal.ca)

Overview

1. Human metabolism and phenylketonuria
2. Mechanisms of penicillin / ampicillin resistance
3. Metabolism can be complex!

DISCLAIMER: This is a very, very simplified view of how metabolism works!!

(for example, it's not always about proteins!)

Two example systems

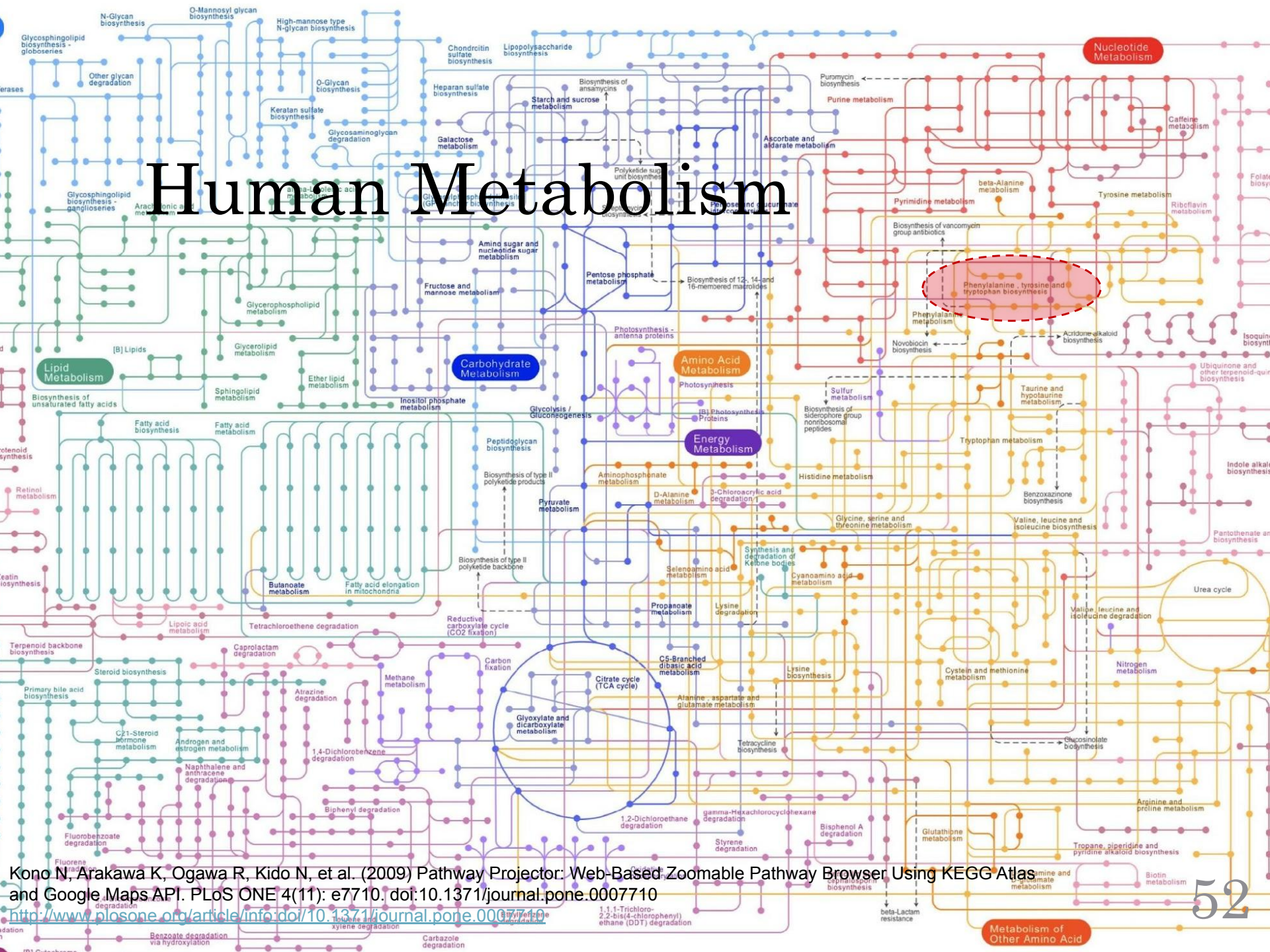
Phenylketonuria

A **very important** metabolic process breaks due to mutation(s) in a critical gene, with severe health consequences

Beta-lactam antibiotic resistance

Bacteria acquire a new and **very useful** function that protects them against a toxic compound

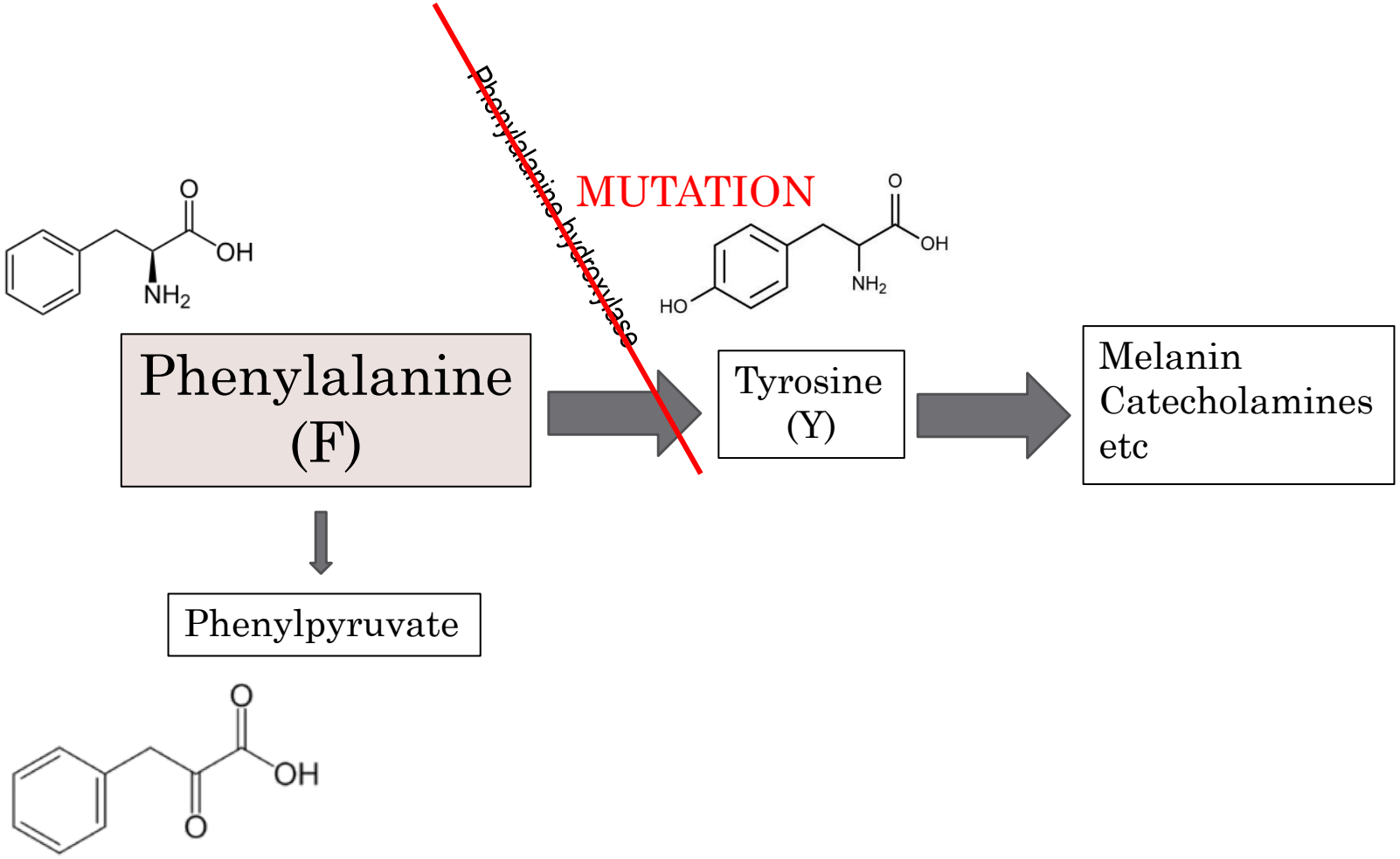
Human Metabolism



Kono N, Arakawa K, Ogawa R, Kido N, et al. (2009) Pathway Projector: Web-Based Zoomable Pathway Browser Using KEGG Atlas and Google Maps API. PLoS ONE 4(11): e7710. doi:10.1371/journal.pone.0007710
http://www.plosone.org/article/info:doi/10.1371/journal.pone.0007710

Photo: R. Beiko

Phenylketonuria (PKU)



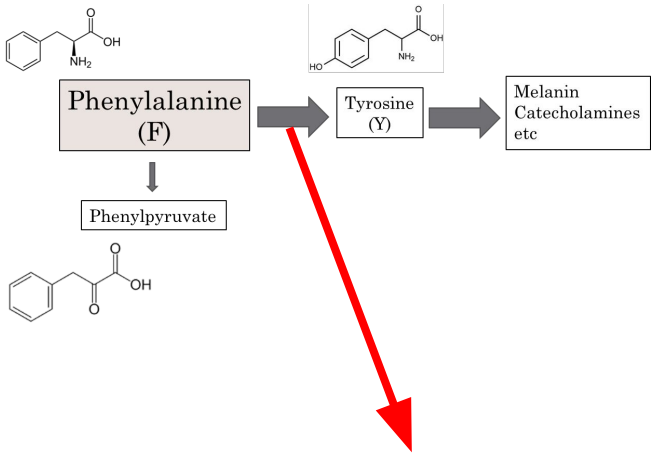
en.wikipedia.org (user NEUROtiker)

Consequences

Too much phenylalanine and phenylpyruvate:

- Blocks the production of neurotransmitters (DOPA, serotonin, GABA)
- Interferes with energy production (pyruvate transport is blocked)
- Leads to impaired brain function

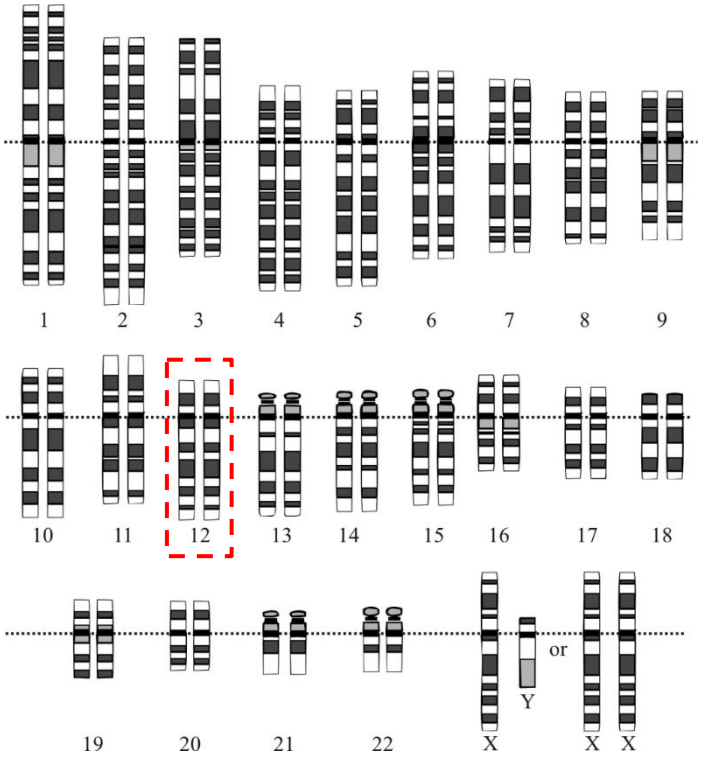
How to break a gene

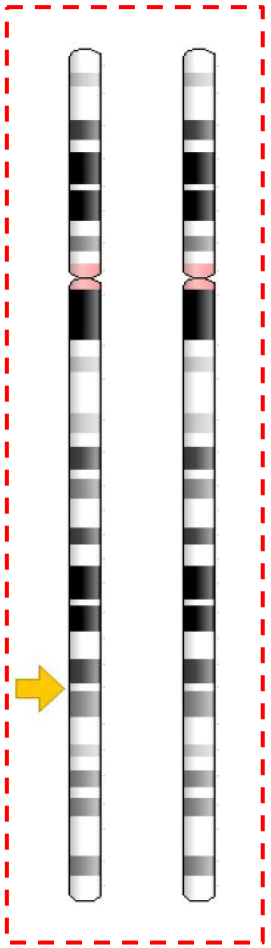


Phenylalanine hydroxylase (*pah*) gene

Chen et al. (2018) *Sci Rep*

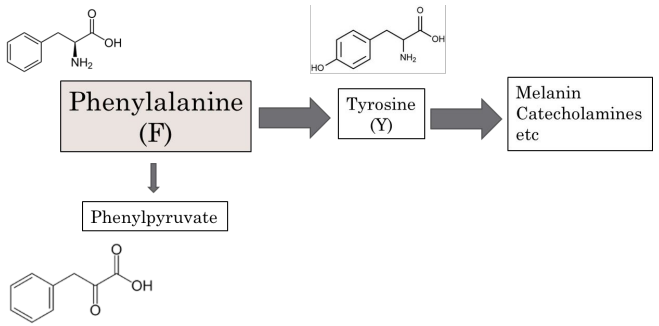
Where is the *pah* gene?

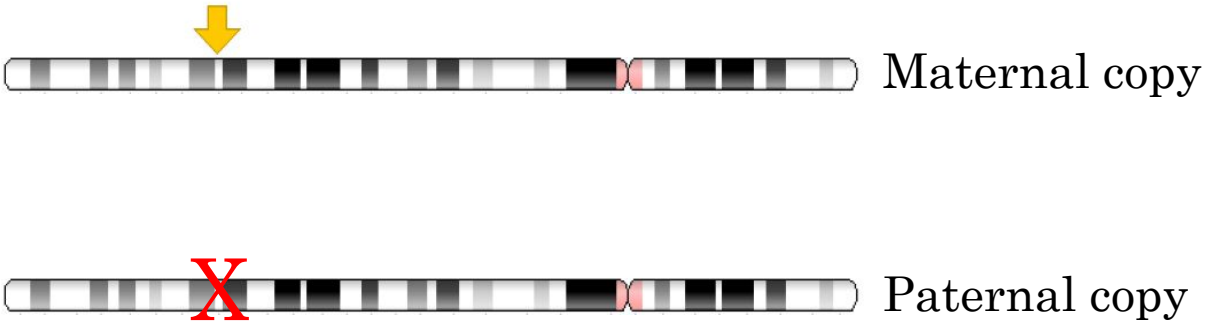




https://en.wikipedia.org/wiki/Human_genome#/media/File:Karyotype.png
Shahid (2020) *Int J Innov Sci Res Tech*

Yes mutations, no PKU

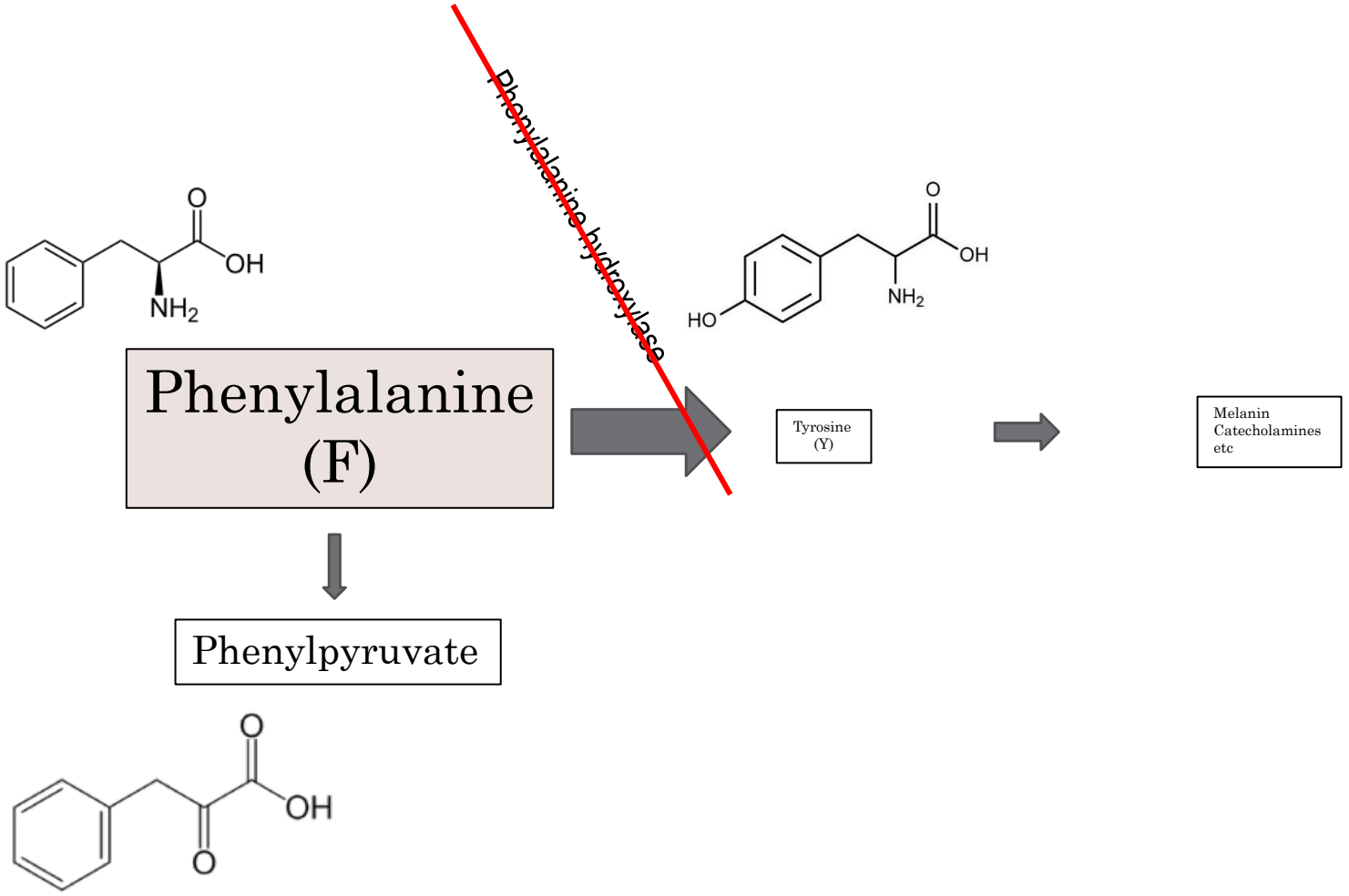




One copy is *good enough* –
if you have one "healthy" copy you don't have PKU

This is how two non-affected parents can have offspring with PKU

Treating PKU



en.wikipedia.org (user NEUROtiker)

Treating PKU

No phenylalanine-containing foods

No aspartame

http://bcbiologypku.wikispaces.com/

PKU is completely treatable, **if** diagnosed at birth

No excess phenylalanine = no physiological consequences

Treatment must be initiated as soon as possible after birth!

See http://www.pahdb.mcgill.ca/images/pku.gif: siblings born before and after PKU screening at birth

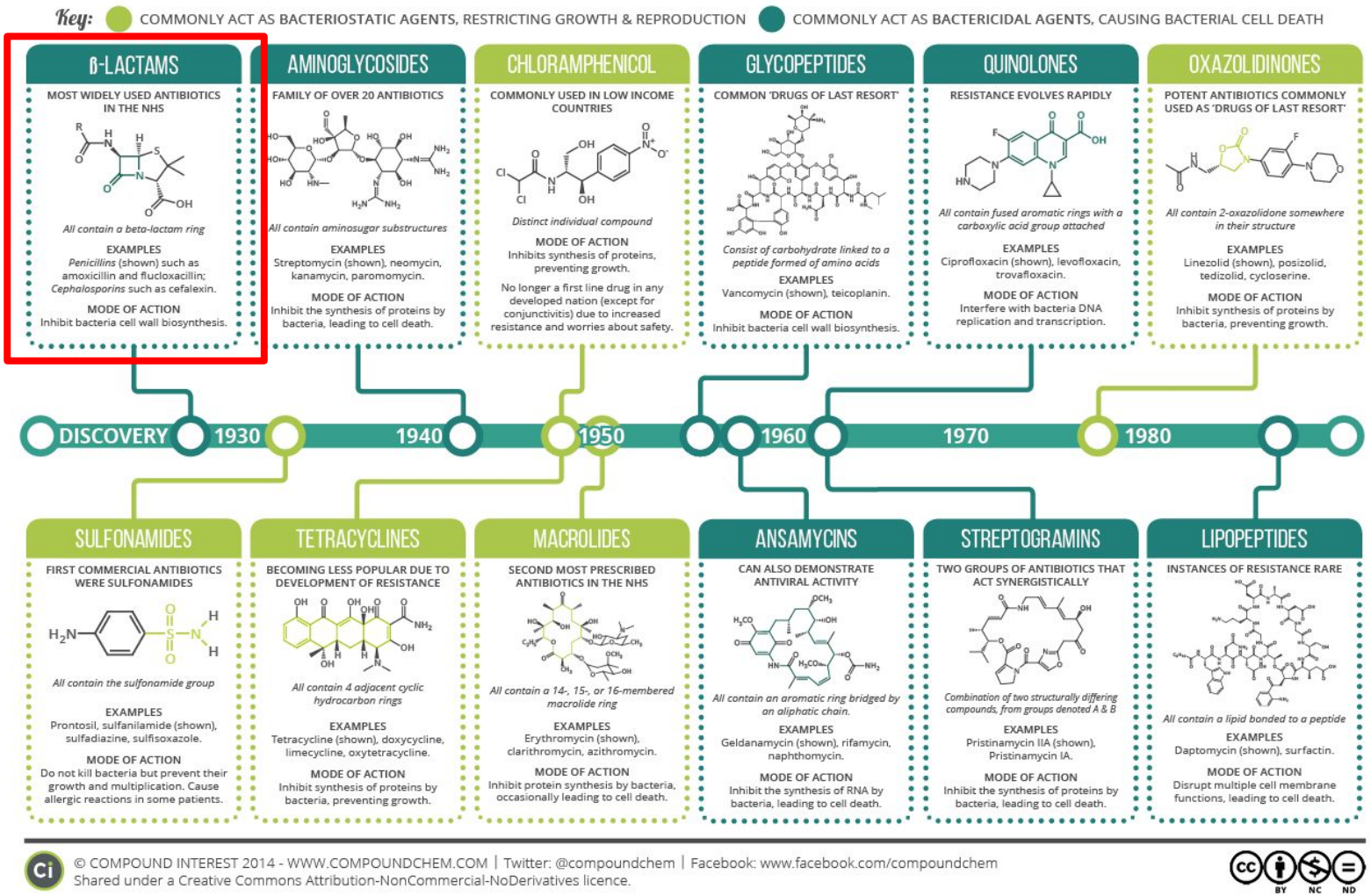
Incidence of PKU:

Region / Country		Incidence of PKU
Asian Populations	China	1 : 17,000
	Japan	1 : 125,000
	Turkey	1 : 2,600
	Yemenite Jews (in Israel)	1 : 5,300
	Scotland	1 : 5,300
	Czechoslovakia	1 : 7,000
	Hungary	1 : 11,000
	Denmark	1 : 12,000
European Populations	France	1 : 13,500
	Norway	1 : 14,500
	United Kingdom	1 : 14,300
	Italy	1 : 17,000
	Canada	1 : 22,000
	Finland	1 : 200,000
Arabic Populations		Up to 1 : 6,000
Oceania	Australia	1: 10,000

Williams et al. (2008) *Clin Biochem Rev*

http://commons.wikimedia.org/wiki/File:Penicillin_cures_gonorrhea.jpg

DIFFERENT CLASSES OF ANTIBIOTICS - AN OVERVIEW



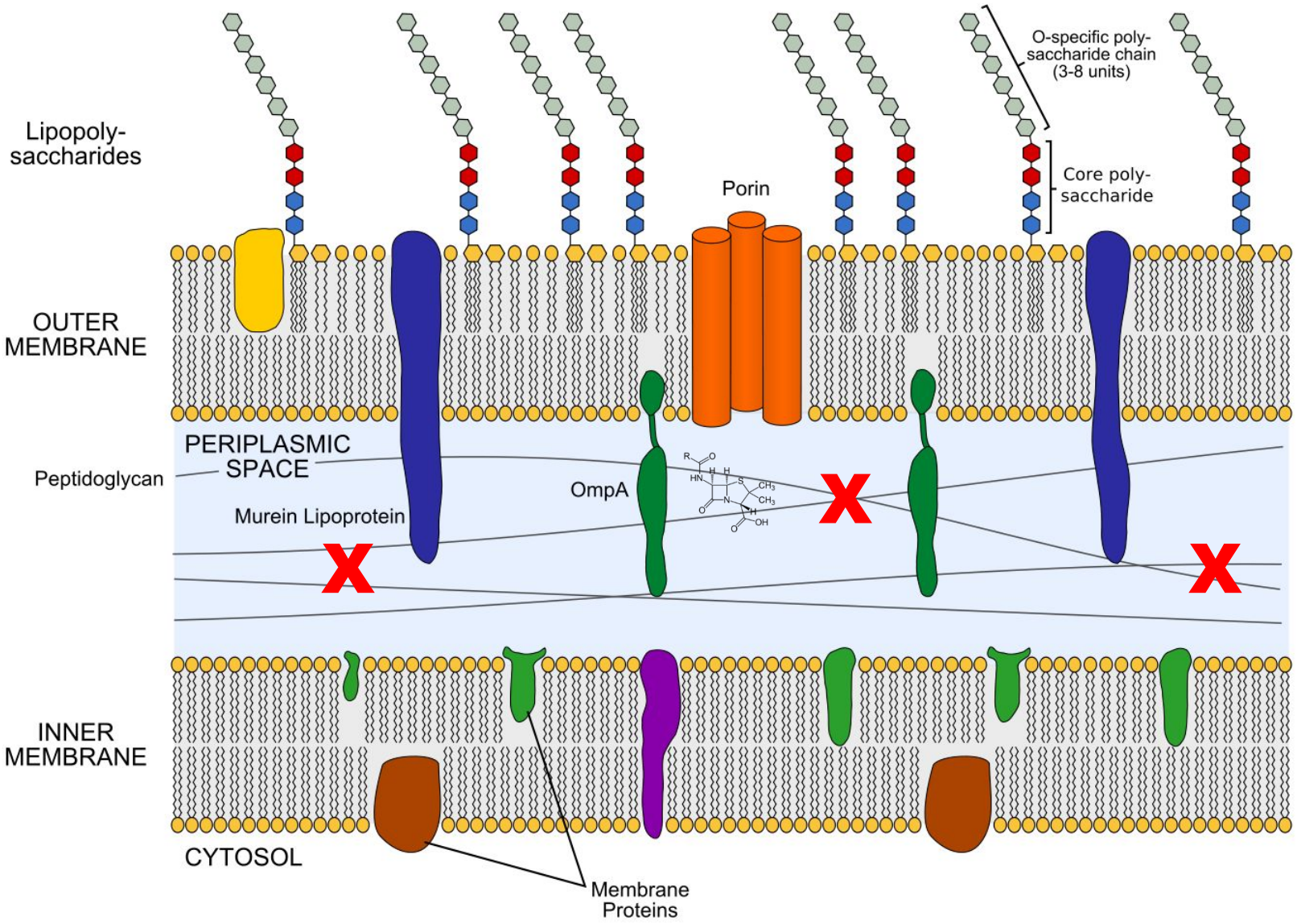
https://www.compoundchem.com/2014/09/08/antibiotics/

Quick disclaimer

- There are many enzymes that can break down beta-lactams

- I'll be referring to a couple of very similar but non-identical genes in a couple of different organisms

The bacterial cell wall

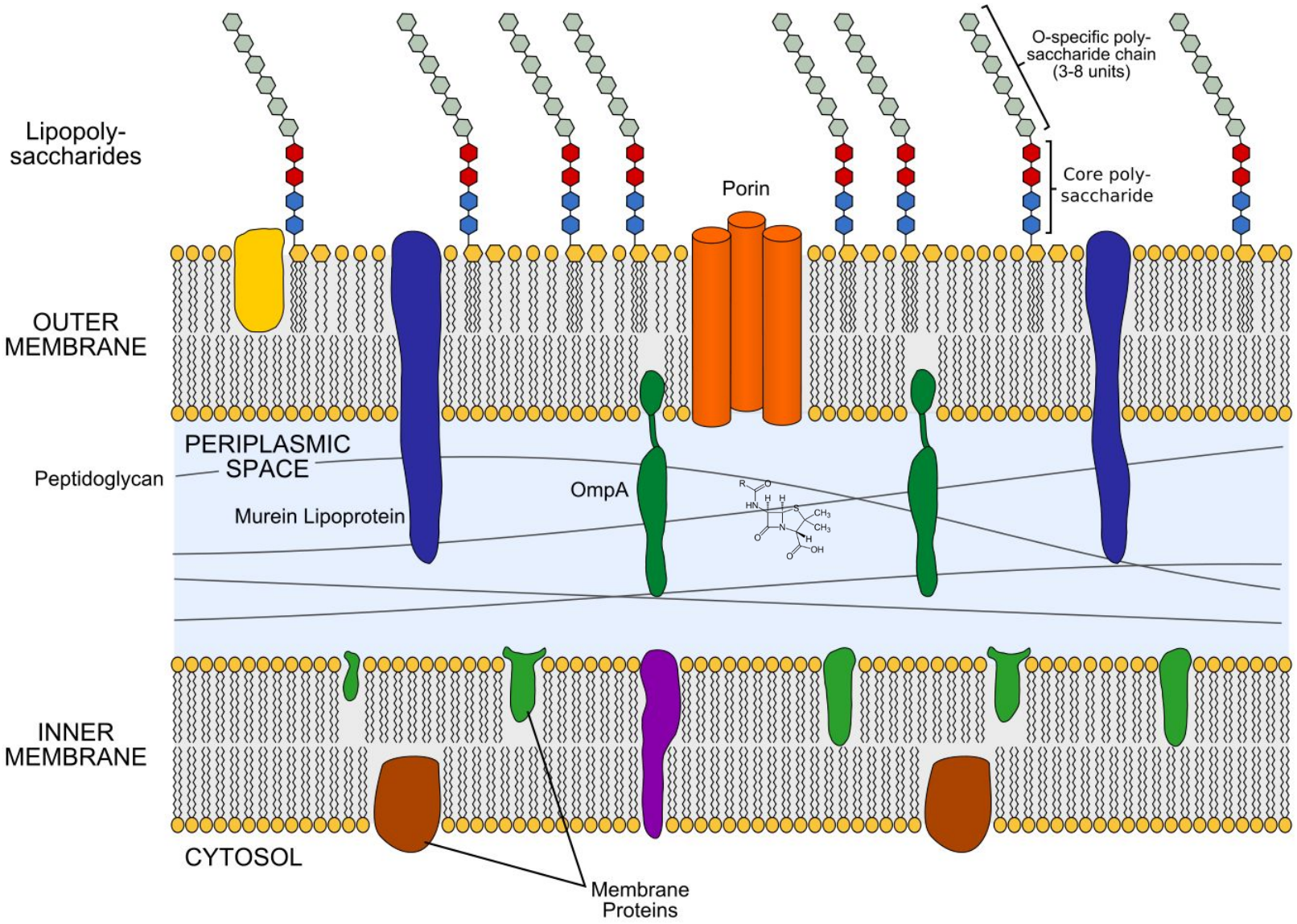


β-lactam antibiotics (such as penicillin) interfere with cell wall assembly

http://upload.wikimedia.org/wikipedia/commons/thumb/8/8b/Gram_negative_cell_wall.svg/2000px-Gram_negative_cell_wall.svg.png

http://www.hls.utas.edu.au/teaching/micro/mma.fb/mma.fb.10.html

The bacterial cell wall

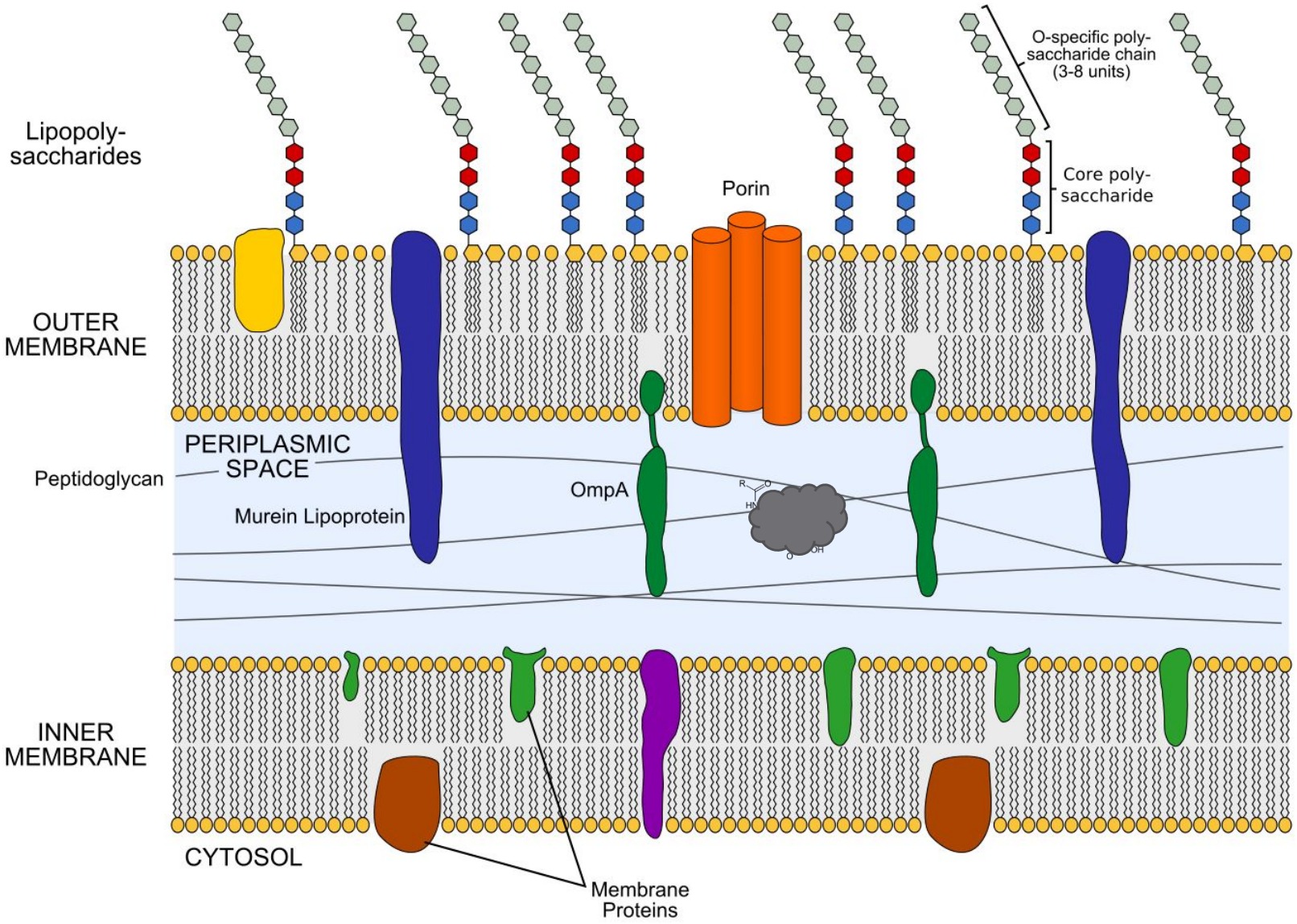


Cells can defend themselves with a class of proteins called **β-lactamases**

http://upload.wikimedia.org/wikipedia/commons/thumb/8/8b/Gram_negative_cell_wall.svg/2000px-Gram_negative_cell_wall.svg.png

http://www.hls.utas.edu.au/teaching/micro/mma.fb/mma.fb.10.html

The bacterial cell wall

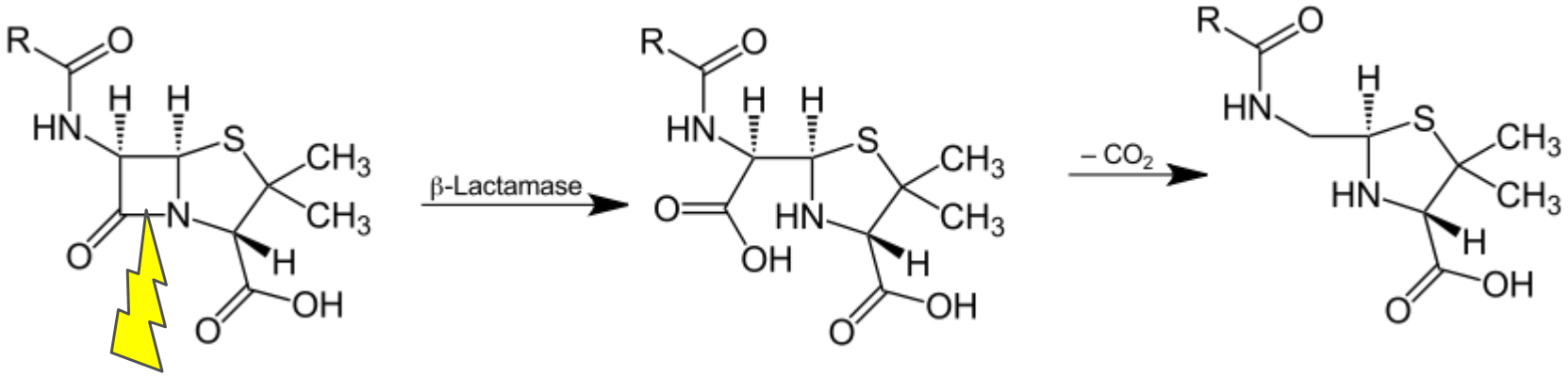


Cells can defend themselves with a class of proteins called **β-lactamases**

http://upload.wikimedia.org/wikipedia/commons/thumb/8/8b/Gram_negative_cell_wall.svg/2000px-Gram_negative_cell_wall.svg.png

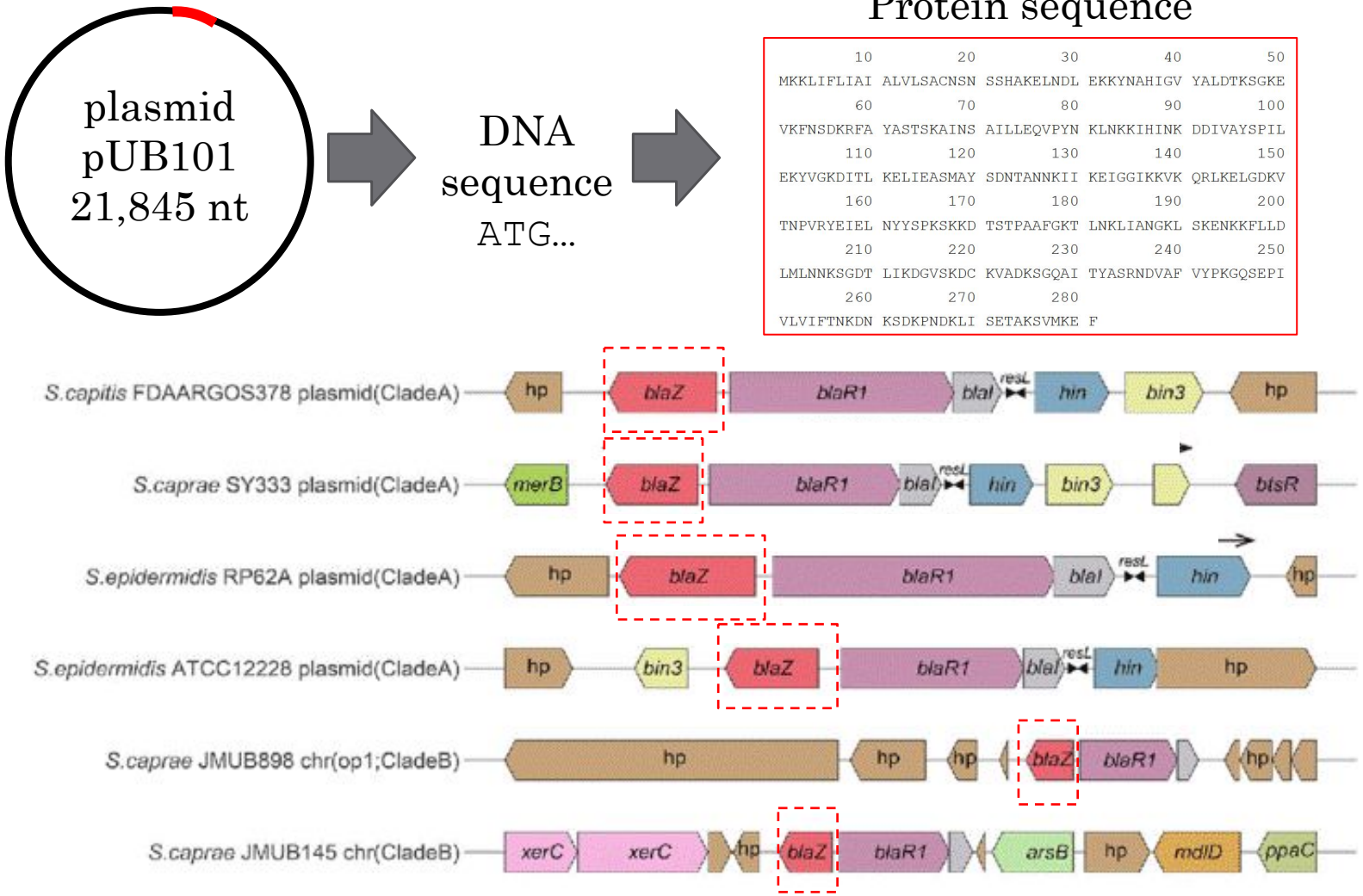
http://www.hls.utas.edu.au/teaching/micro/mma.fb/mma.fb.10.html

Breaking down beta-lactams



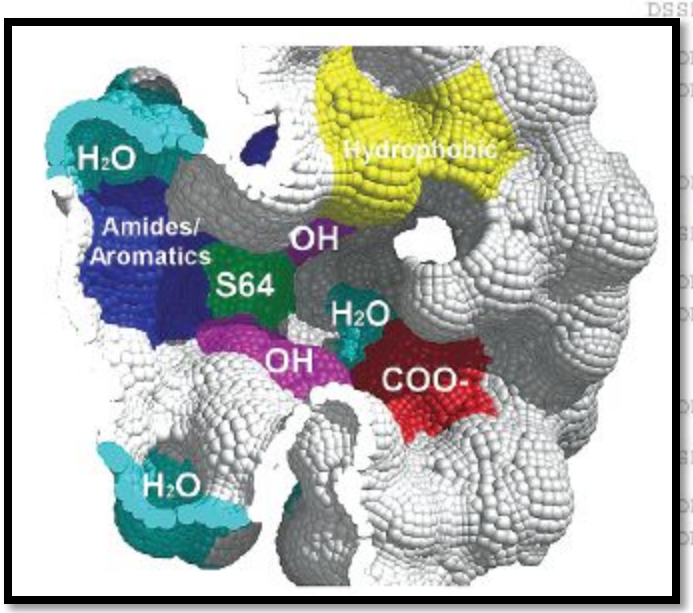
http://en.wikipedia.org/wiki/File:Lactamase_Application_V.1.svg

The *blaZ* gene

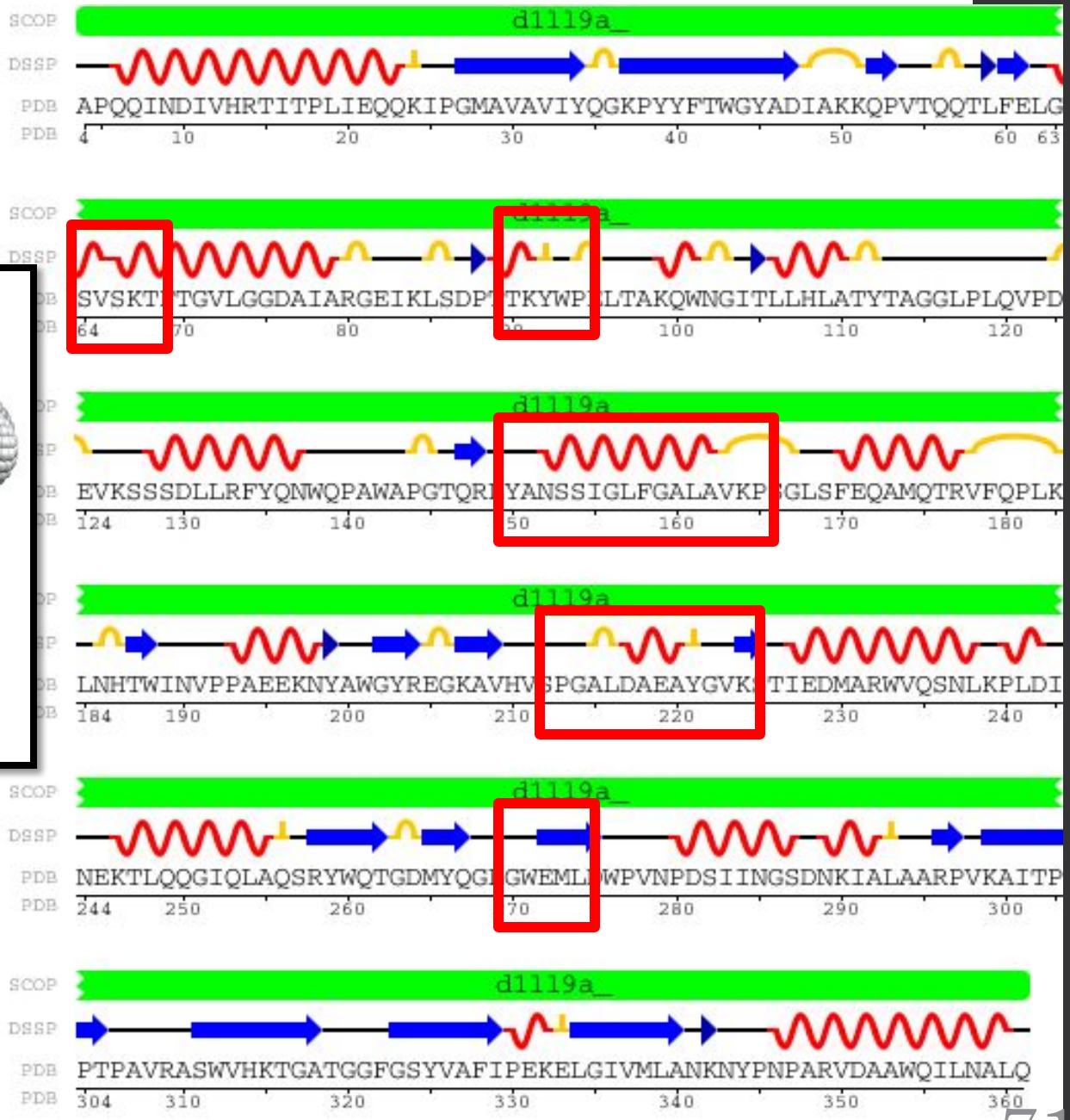


Sun et al. (2020) *Front Cell Infect Microbiol*

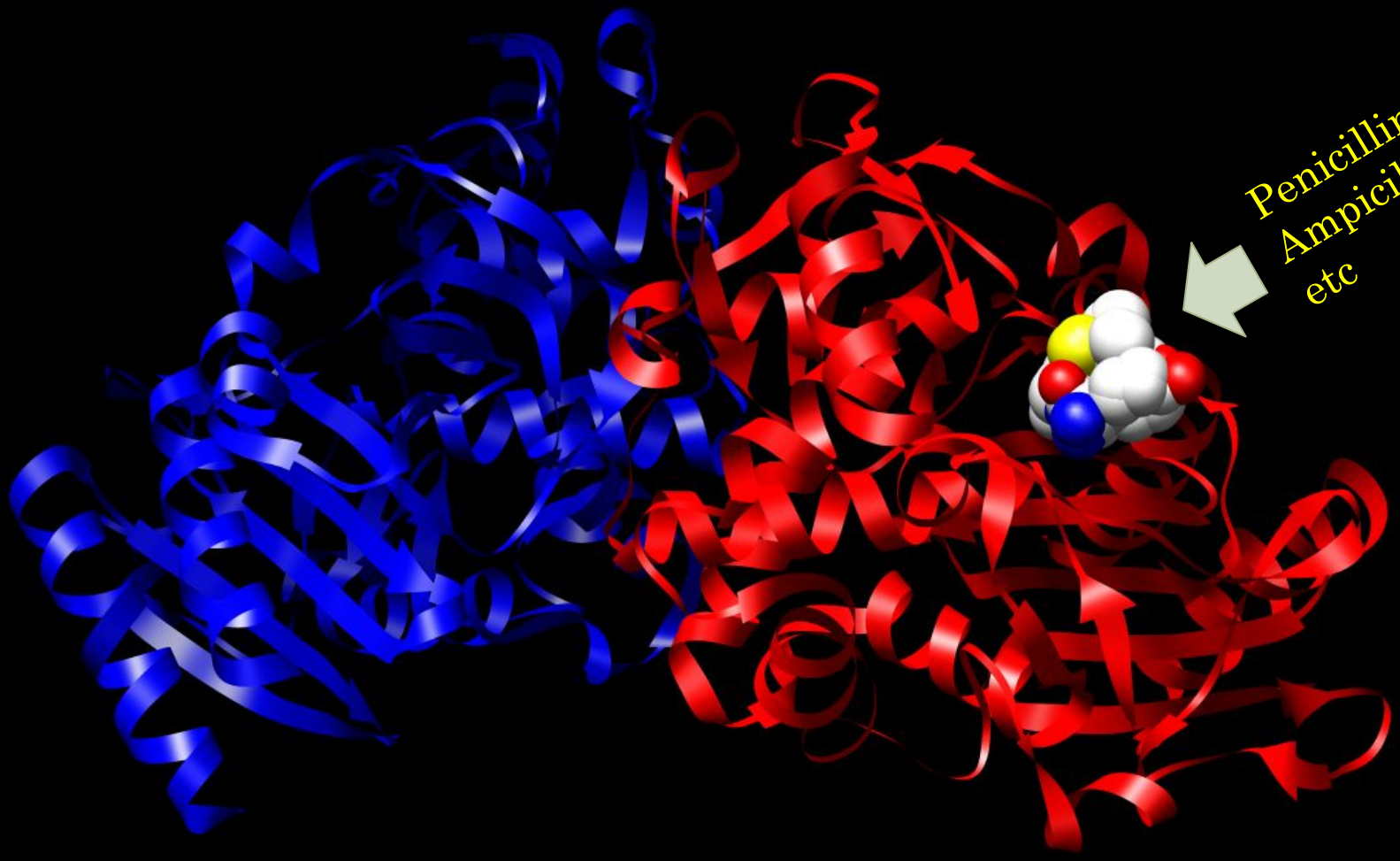
Functional sites



Powers and Sholchet, *J Med Chem* 2002



http://www.rcsb.org/pdb/explore/remediatedSequence.do?structureId=1LL9

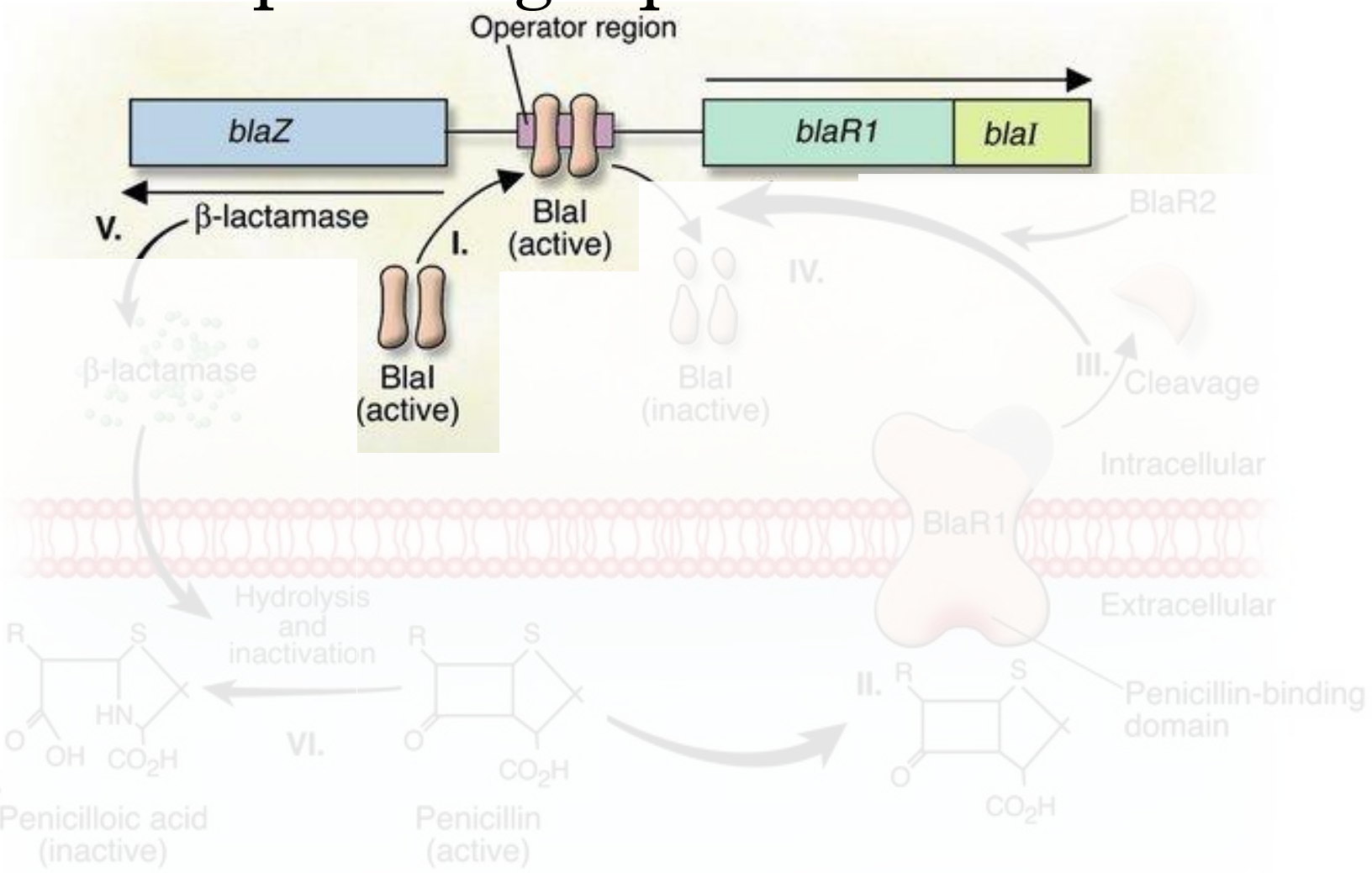


Binding antibiotics

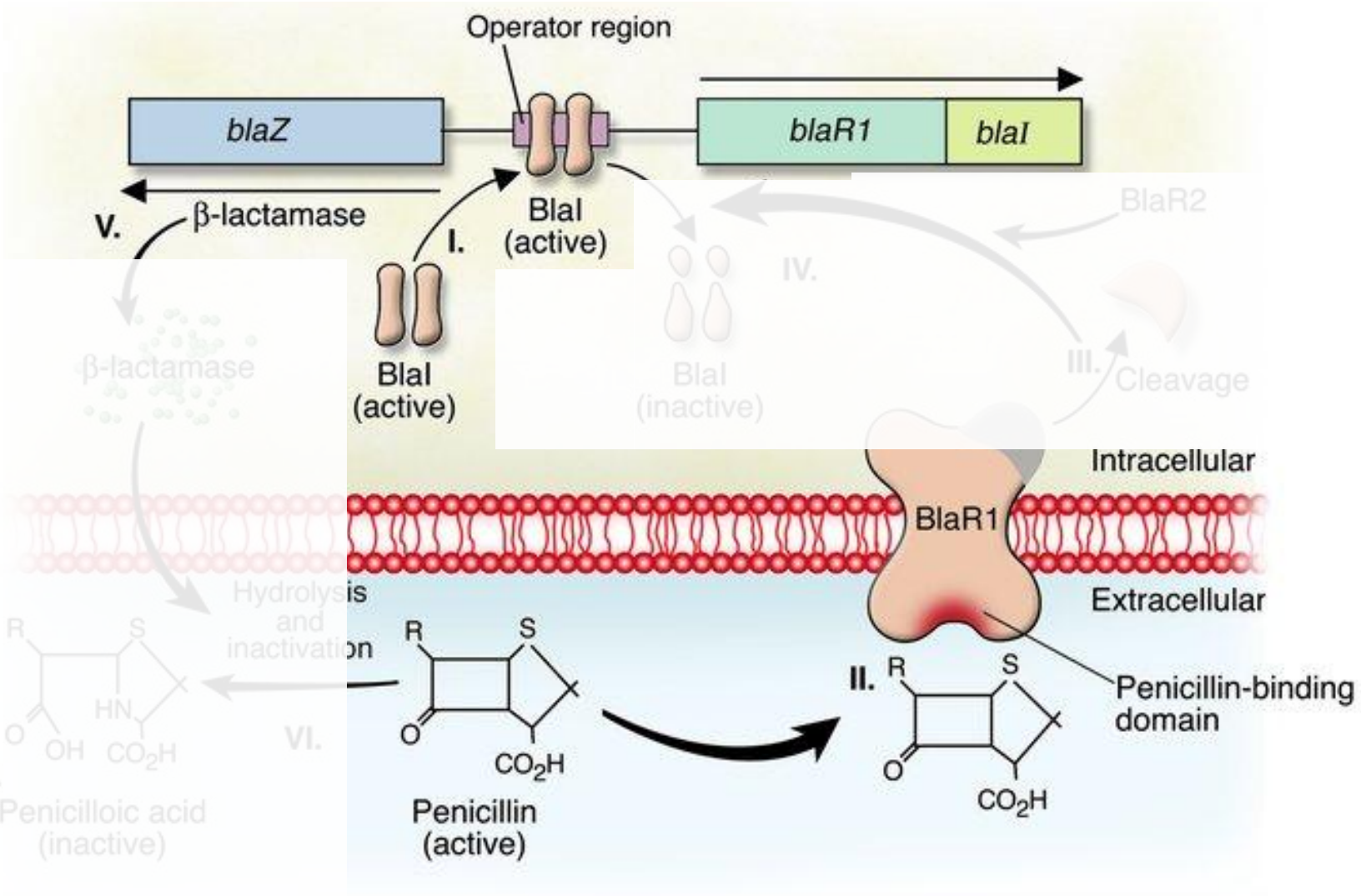
Self-defence in several easy steps

Lowy (2013) *J Clin Invest*

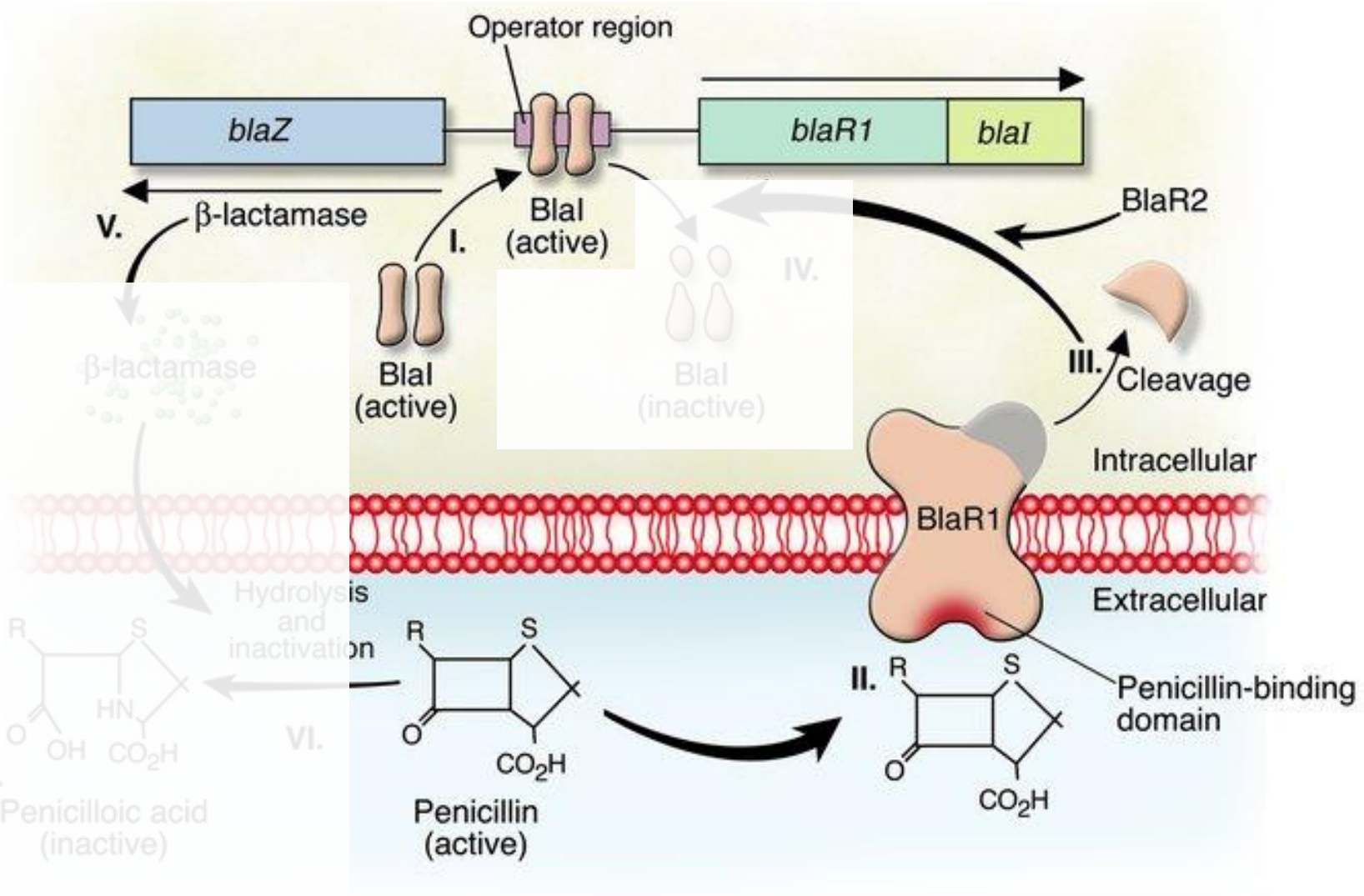
Safe and sound – BlaI keeps things quiet



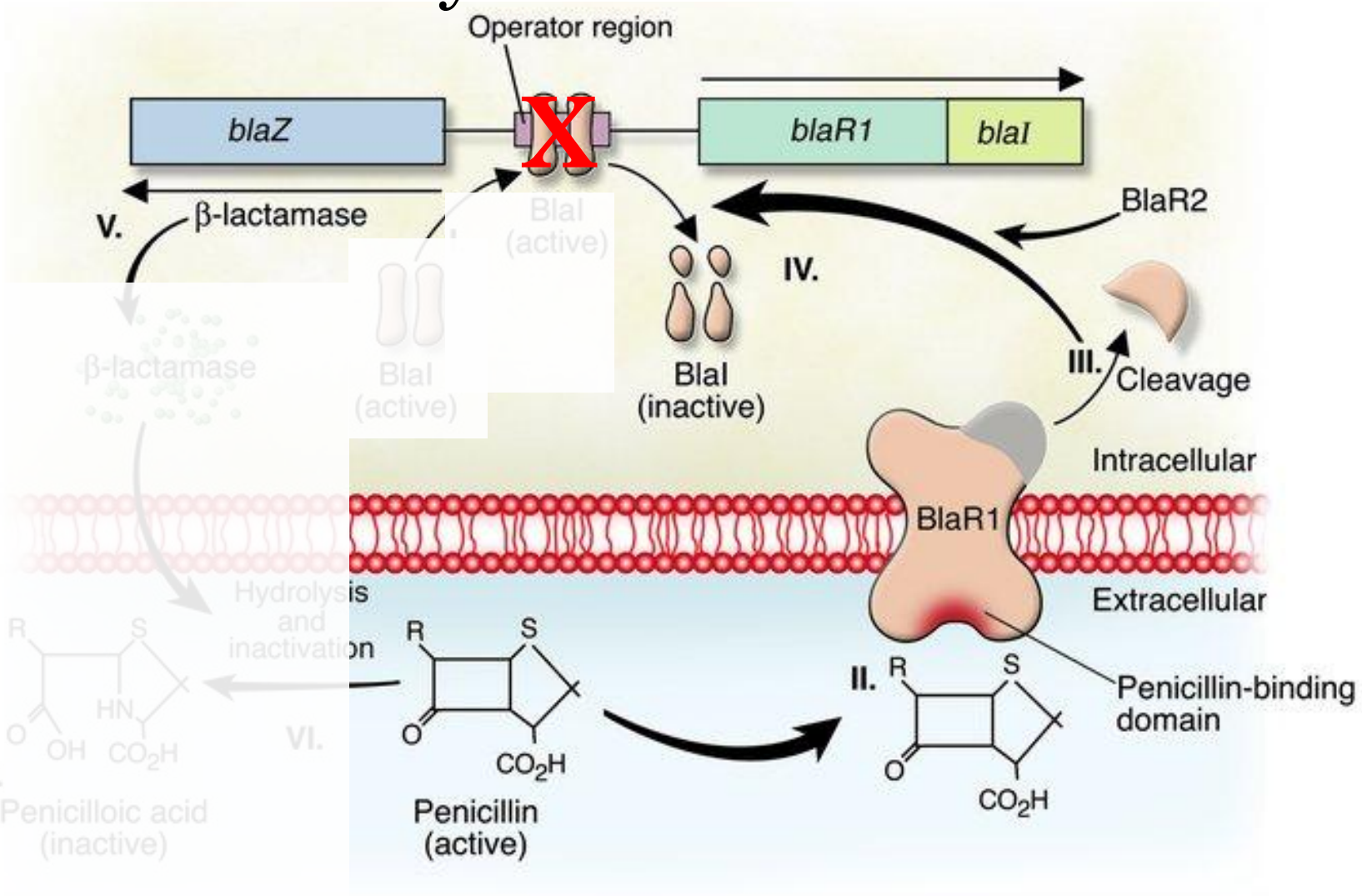
Danger!
Penicillin is in the air



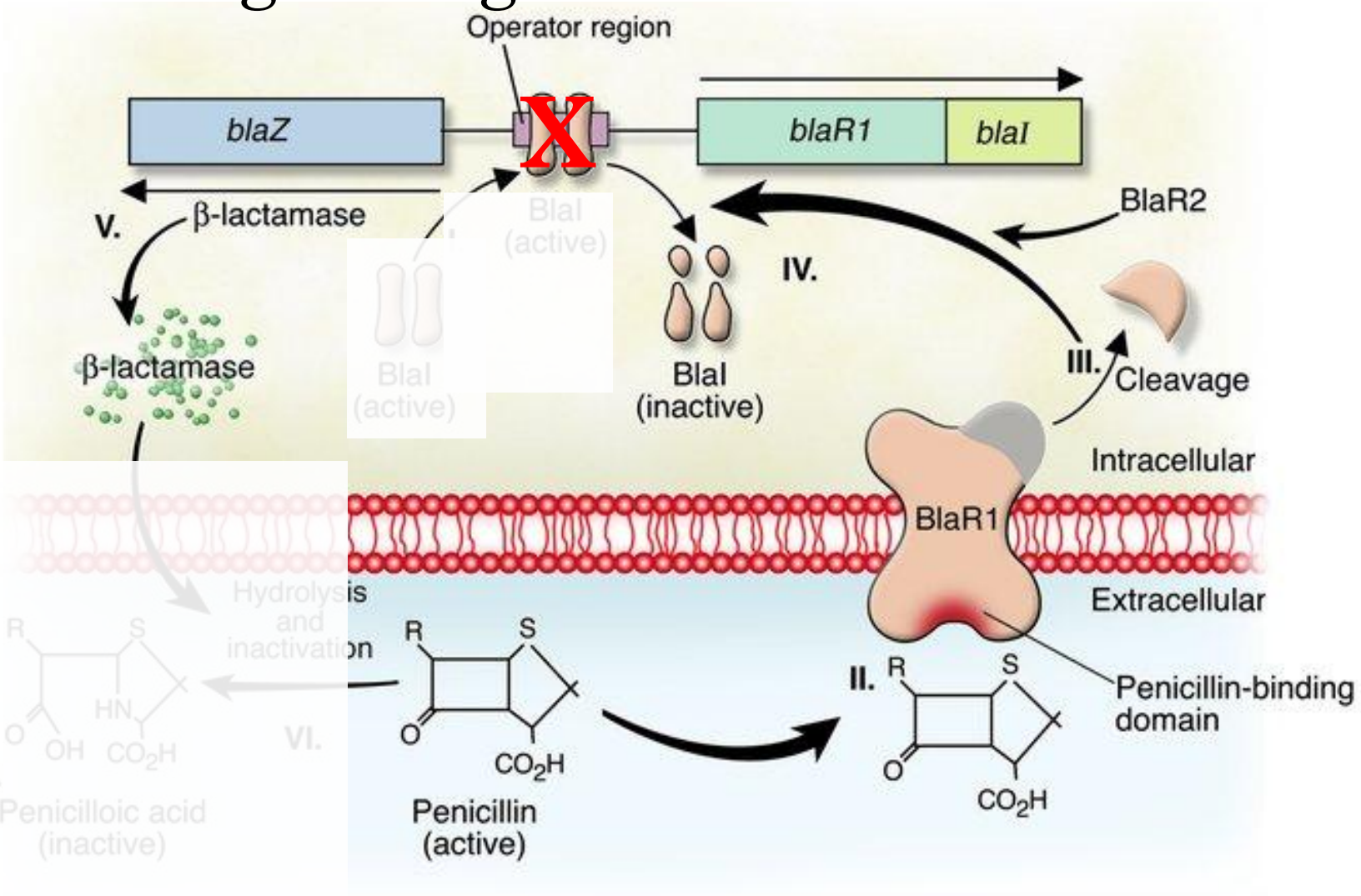
Oh snap!


Operator region
blaZ
blaR1
blaI
β-lactamase
V.
I.
BlaI (active)
BlaR2
III.
Cleavage
BlaI (active)
BlaI (inactive)
IV.
Intracellular
BlaR1
Extracellular
Penicillin-binding domain
II.
Penicillin (active)
Penicilloic acid (inactive)
VI.
VII.

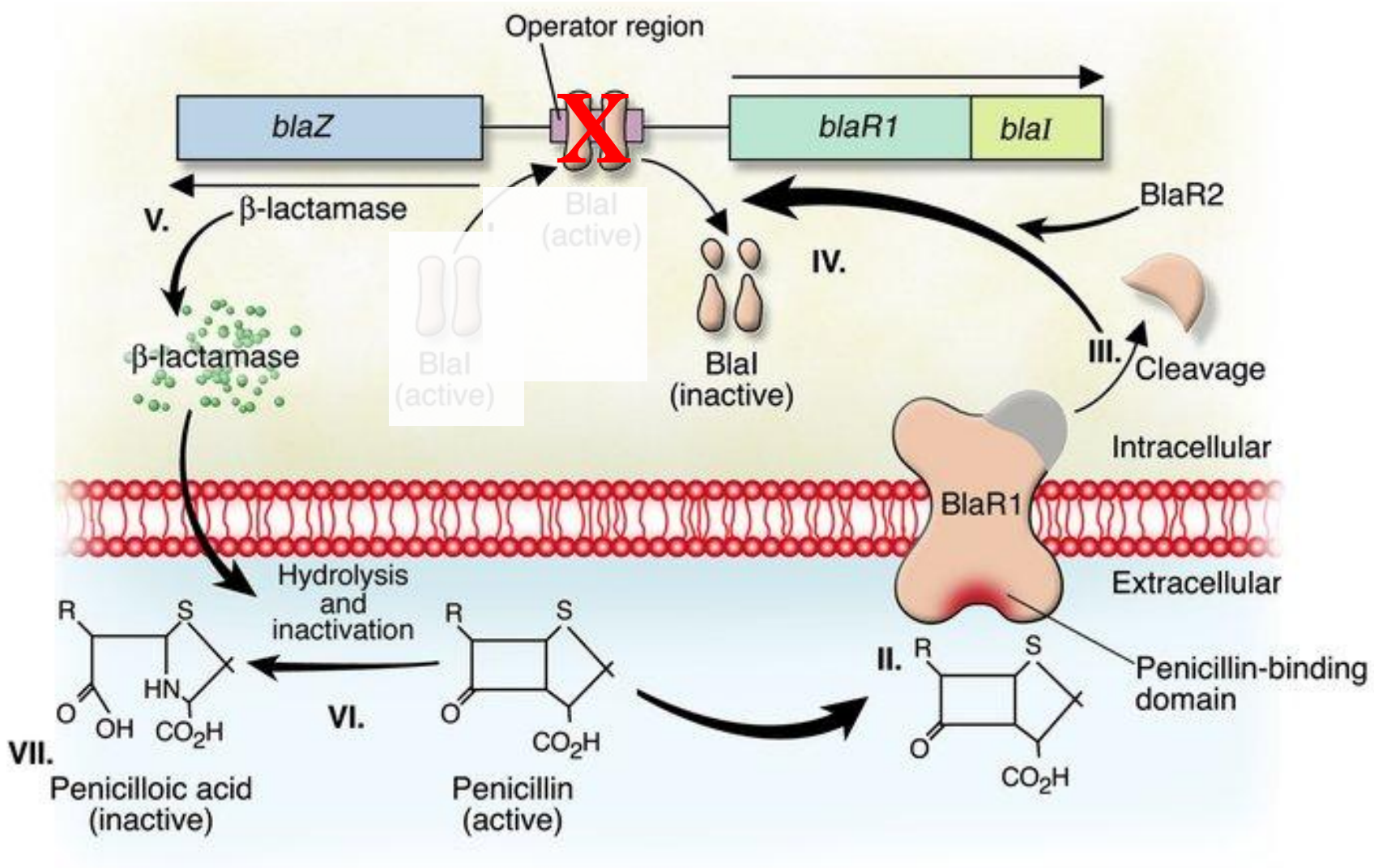
Destroy the repressor, activate the system



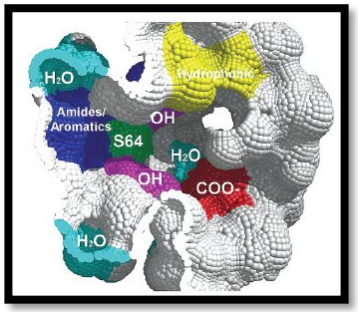
Transcribe and translate
- The beginning of the end



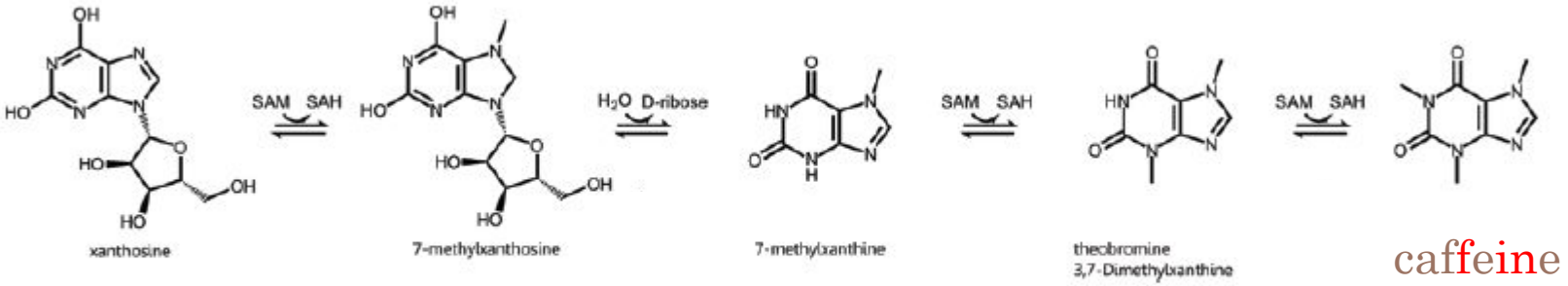
Export and destroy



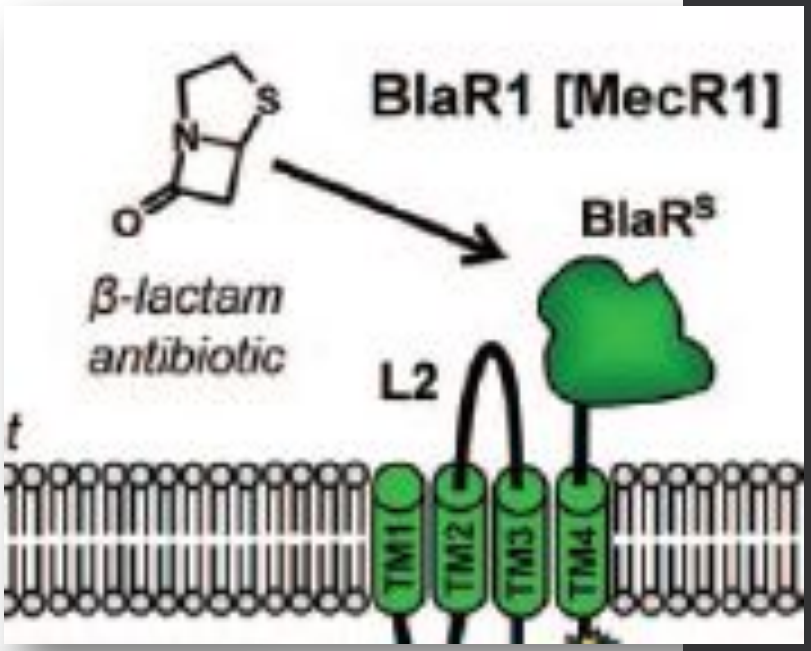
- Different parts of a protein can play different roles
  - Binding small molecules such as antibiotics, food / energy molecules, etc.
  - Contact with other proteins or DNA
  - Assembly
  - Internal stability once assembled



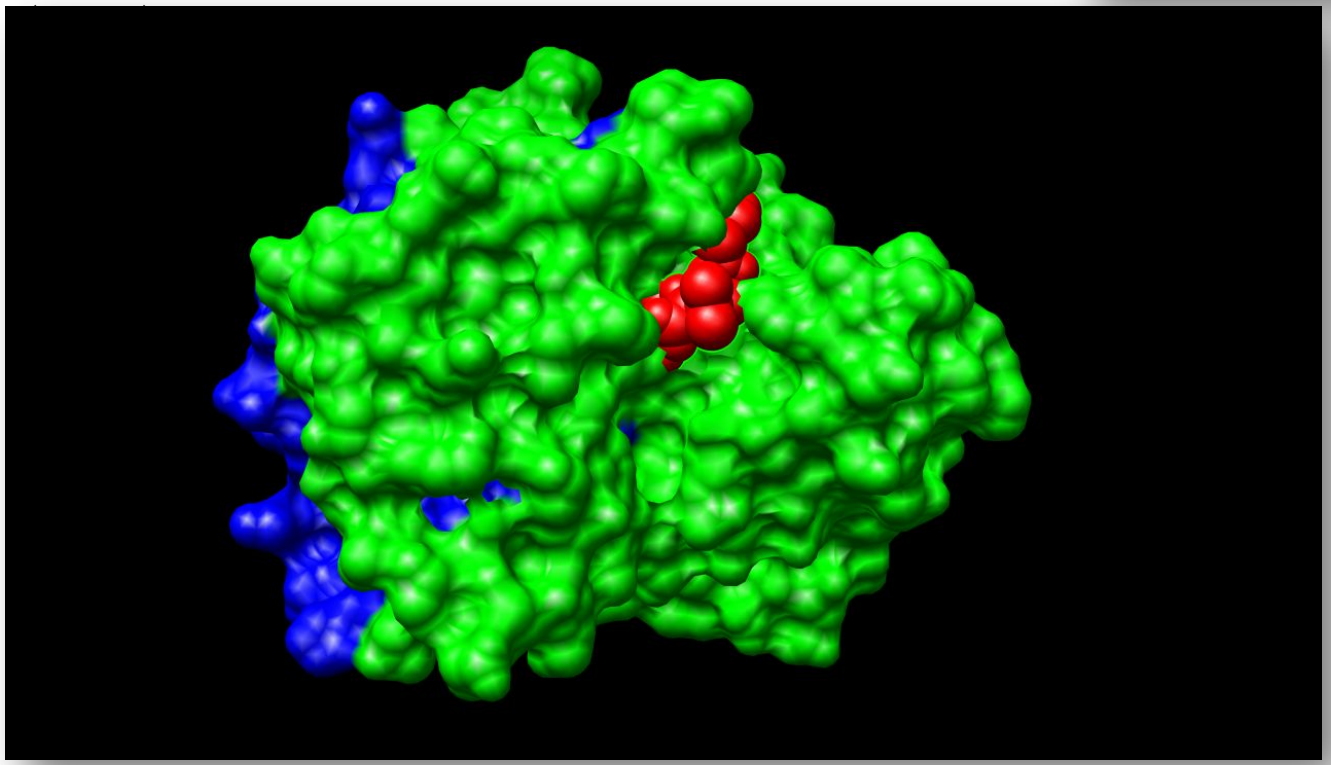
- These functions impose different *constraints* on protein sequence and structure, and affect how gene and protein sequences can evolve
  - *'Ultraconserved' elements*

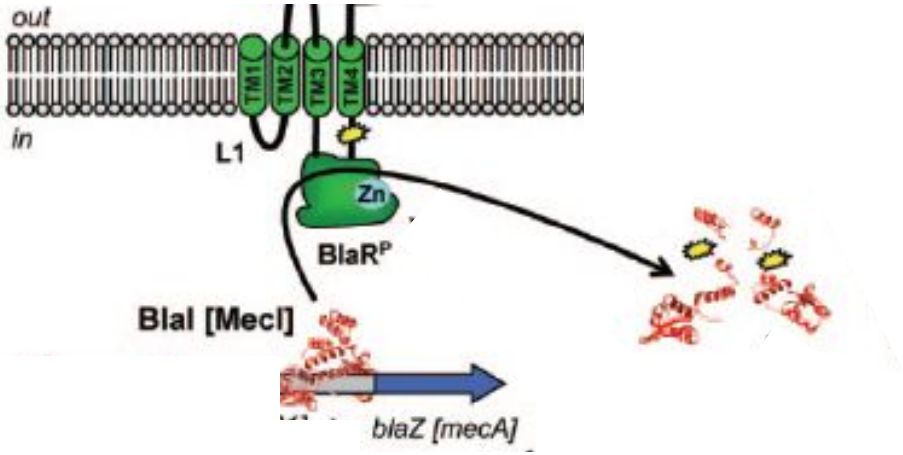


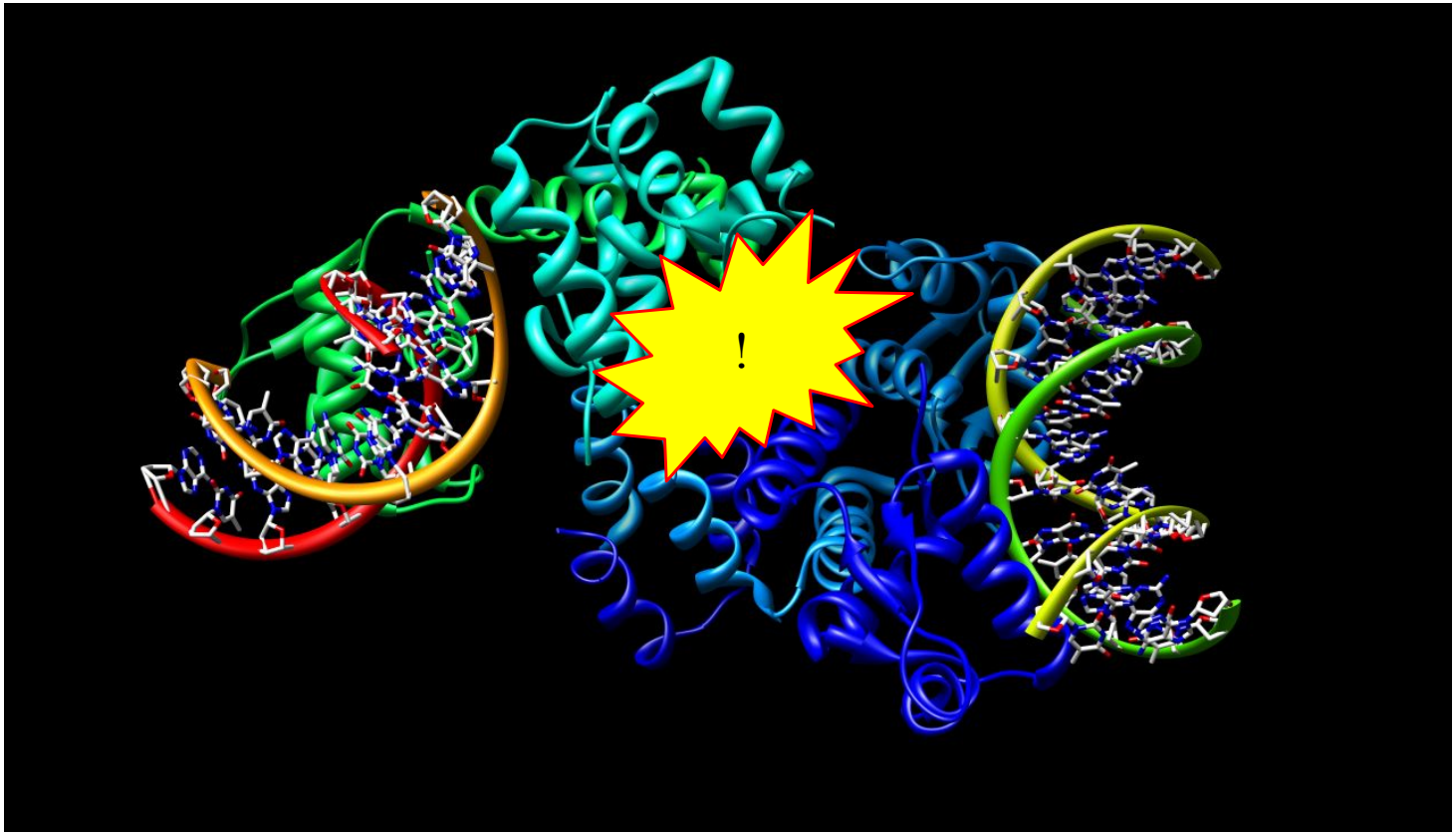
http://parts.igem.org/Part:BBa_K801077

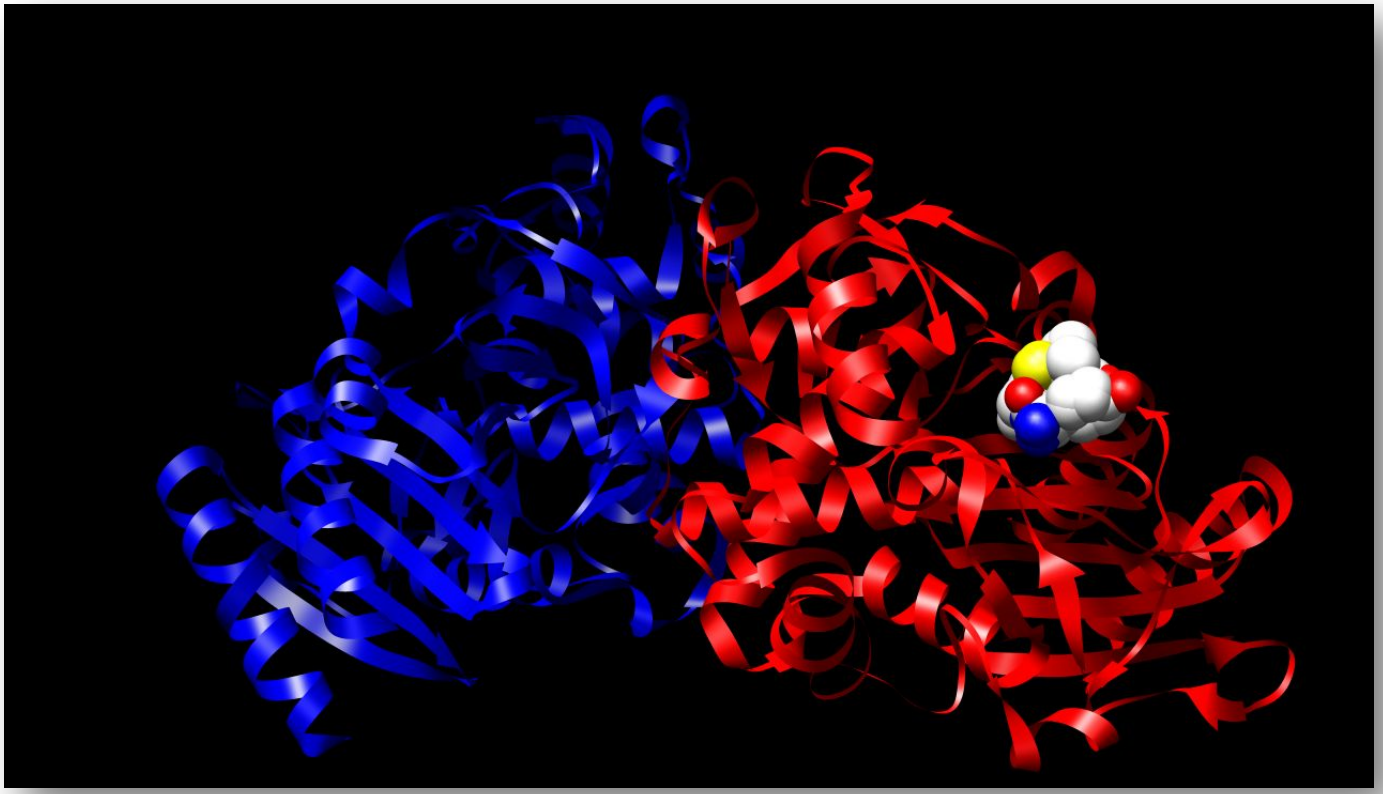


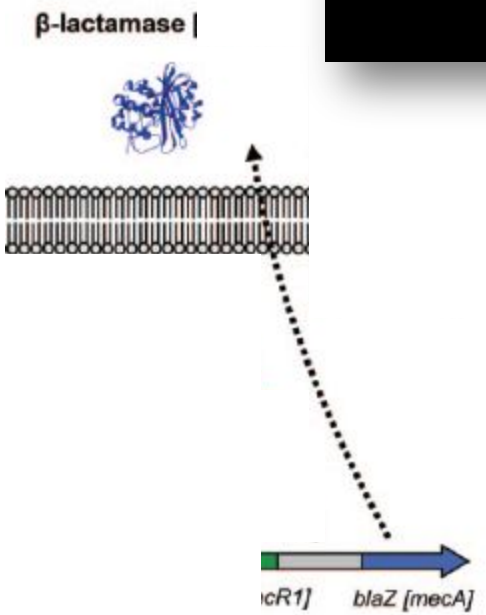
Wilke et al., *J Biol Chem*



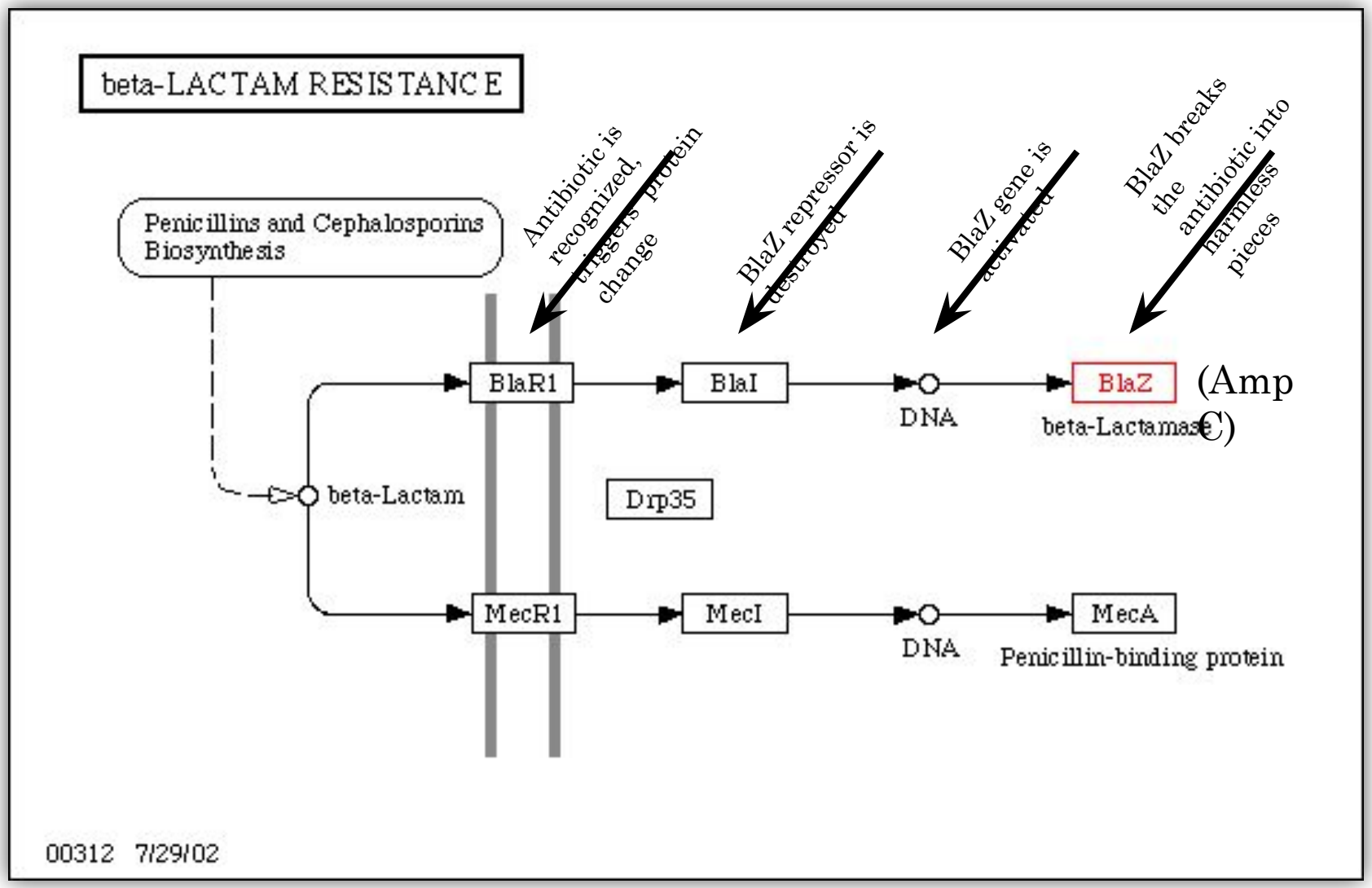

out
in
TM1
TM2
TM3
TM4
L1
Zn
BlaR^P
BlaI [MecI]
blaZ [mecA]


!




β-lactamase |
cR1]
blaZ [mecA]

Mobilization of Penicillin Defence



http://www.genome.jp/dbget-bin/show_pathway?map00312+3.5.2.6