

Overview of molecular microbiology

Plan

1. biochemical aspects
2. mechanistic aspects
- 3. regulatory aspects**
4. genomic aspects
5. physical aspects
6. comparison to eukaryotes

❖ dependence on the environment:

- exponential growth:

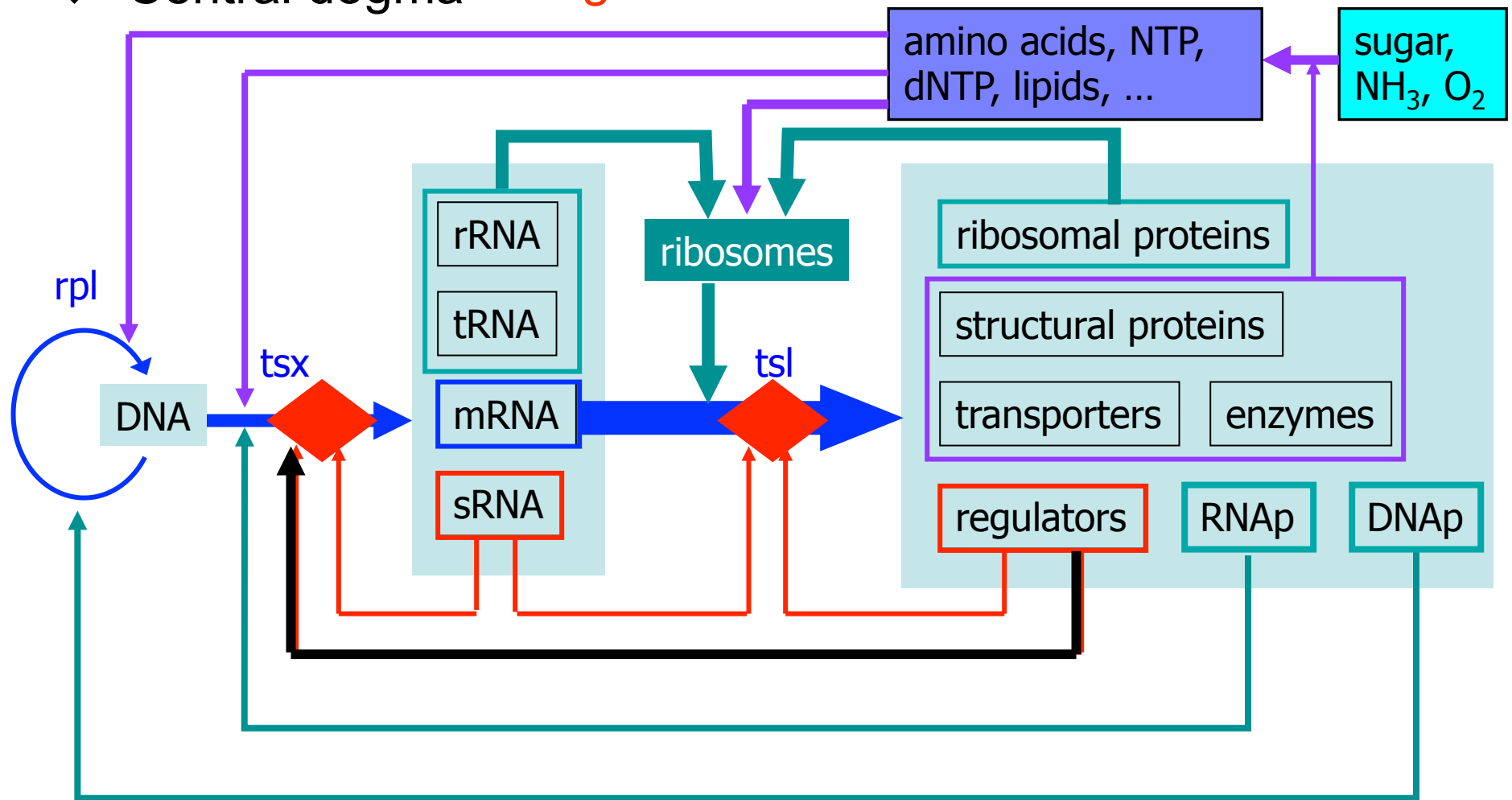
	minimal medium	rich medium
energy cost (J/g)	560	85
doubling time	2 - 3 hr	≥ 20 min

→ bacteria can sense the environment and adjust their “growth program” according to environmental conditions

- coping with stressful conditions:
 - motility: flagella synthesis and chemotaxis
 - osmotic response: porin/channel synthesis
 - heat shock response: chaperons
 - quorum sensing, biofilms, bacterial community
 - SOS response (e.g., to DNA damage)
- non-growth condition
 - stationary phase (*E. coli* can be dormant for > 10 years)
 - sporulation (e.g., *B. subtilis*)
 - Exchange of genetic material: transformation (competence) and conjugation

need regulation to maintain optimal growth

❖ Central dogma + regulation

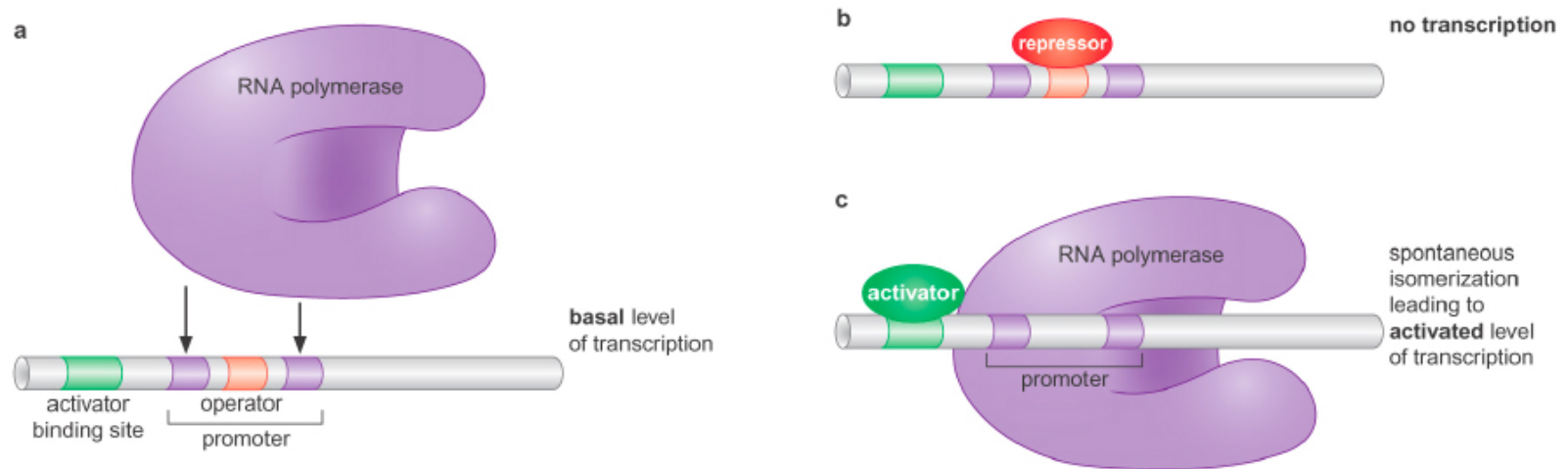


- **tsx** initiation control by transcription factors (TF)
- **tsl** initiation control by sRNA and RNA-binding proteins
- **tsx** termination control by sRNA and anti-terminators
- control of mRNA and protein degradation

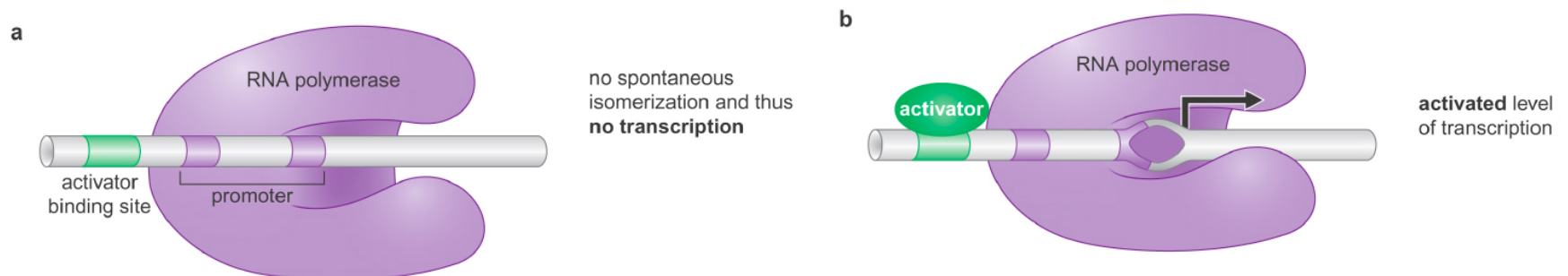
coupled to environmental signals; coord growth program

❖ transcriptional initiation control

- modulation of RNAP-promoter affinity via activators and repressors



- allosteric activation of RNAP

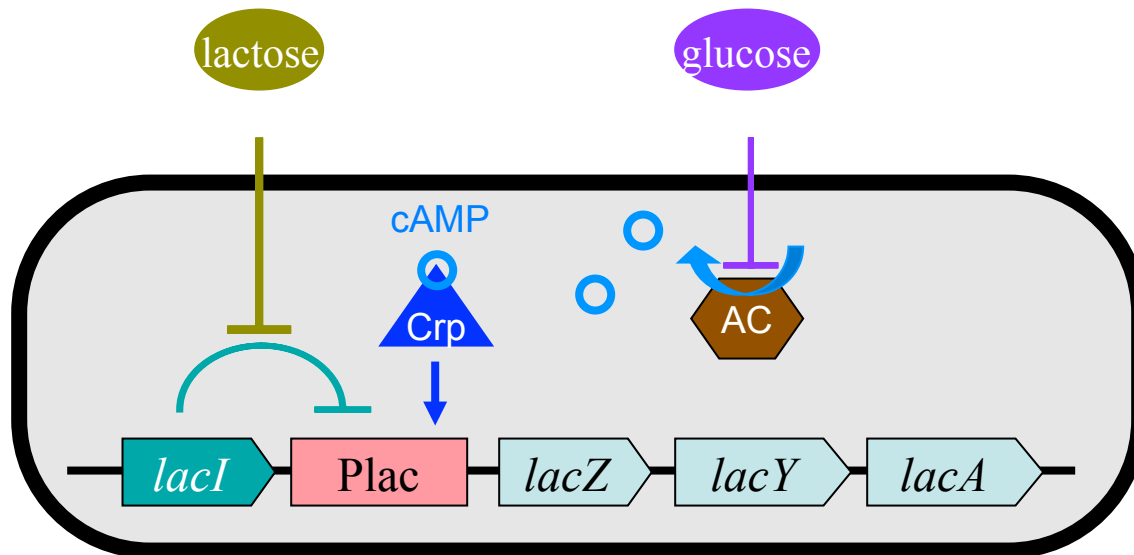


➔ net result: rate of tsx init dependent on cellular conc of activators/repressors controlled by, e.g., environmental inducer molecules

Regulation of the *lac*-operon of *E. coli*

Physiology:

- *lac*-operon: utilization of lactose
 - LacZ: digestive enzyme; LacY: lactose transporter; LacA: acetyl-xfer
- repressed by the **Lac Repressor** (encoded by *lacI*)
- repression alleviated by allo-lactose (minor by-product of lactose metabolism)
- activated by the global regulator **Crp** (aka **CAP**) requires the inducer **cAMP**
- cAMP synthesized endogenously by **Adenylate Cyclase** (encoded by *cyaA*)
- activity of **AC** repressed by **glucose** uptake

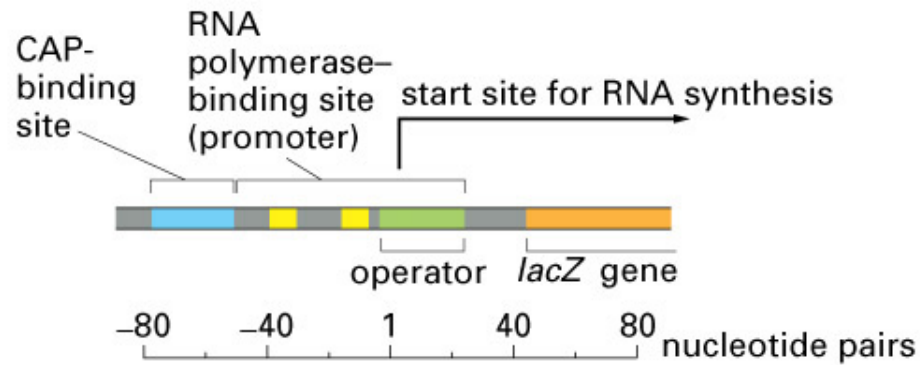


qualitative behavior:

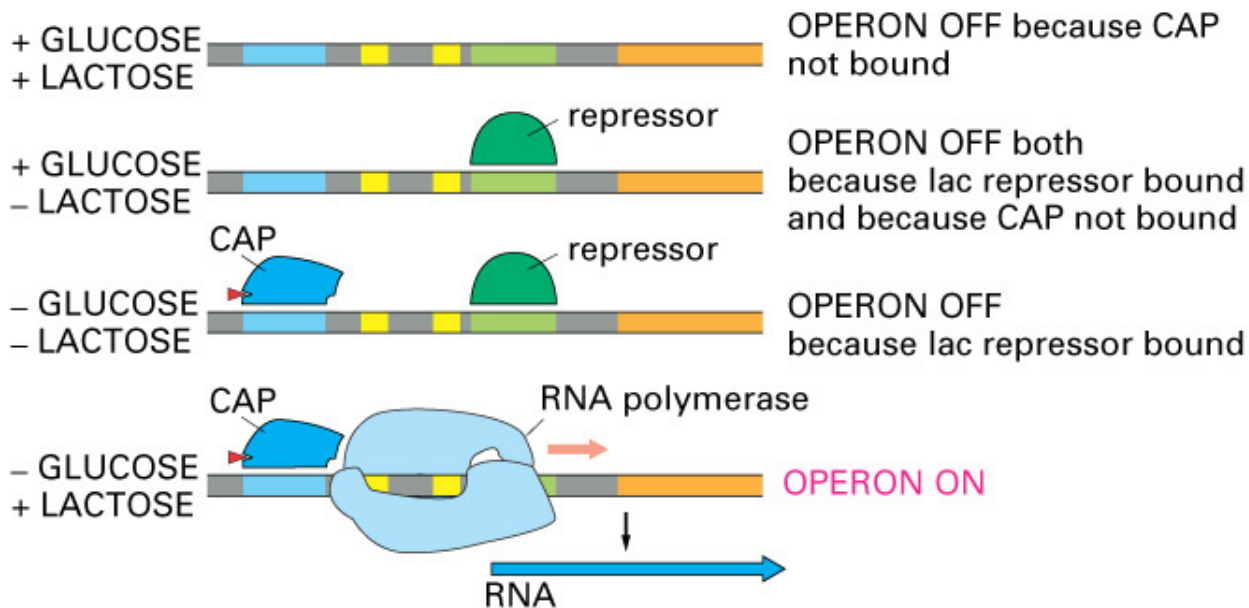
lac	glucose	expression
low	high	OFF
low	low	OFF
high	high	OFF
high	low	ON

Function: expression ONLY in the presence of lactose **AND** absence of glucose

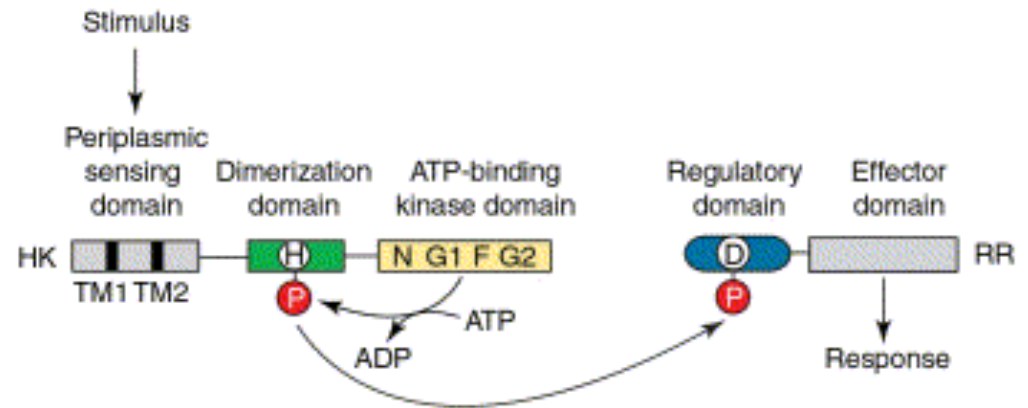
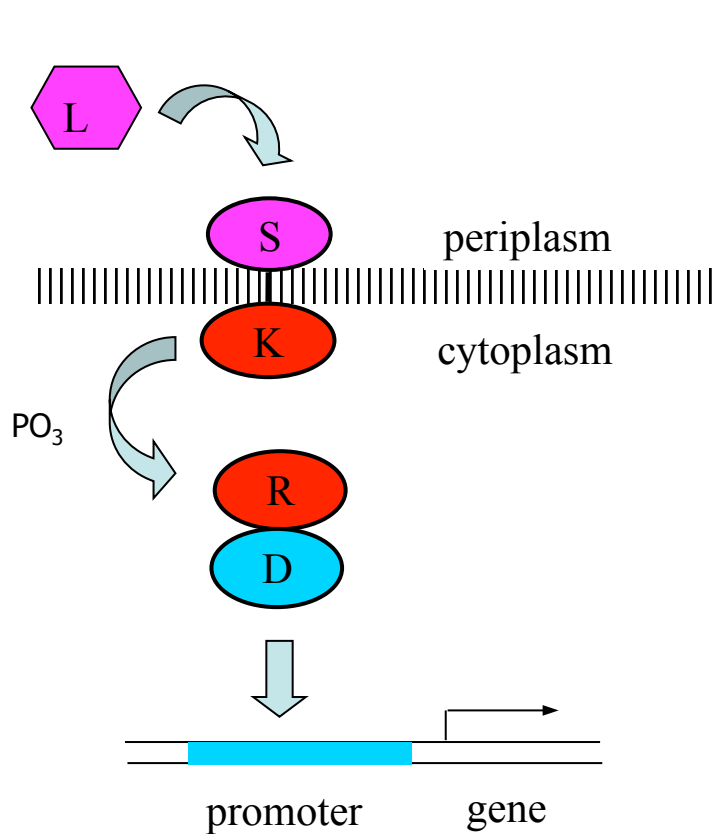
Molecular implementation of regulatory function



lactose	glucose	expression
low	high	OFF
low	low	OFF
high	high	OFF
high	low	ON



❖ two-component signaling systems



Notable examples in *E. coli* control osmoregulation (EnvZ/OmpR system) and the response to chemicals (CheA/CheY system)

Regulation does not necessarily rely on proteins. RNAs play very important roles in the regulation of translation and degradation but also transcription

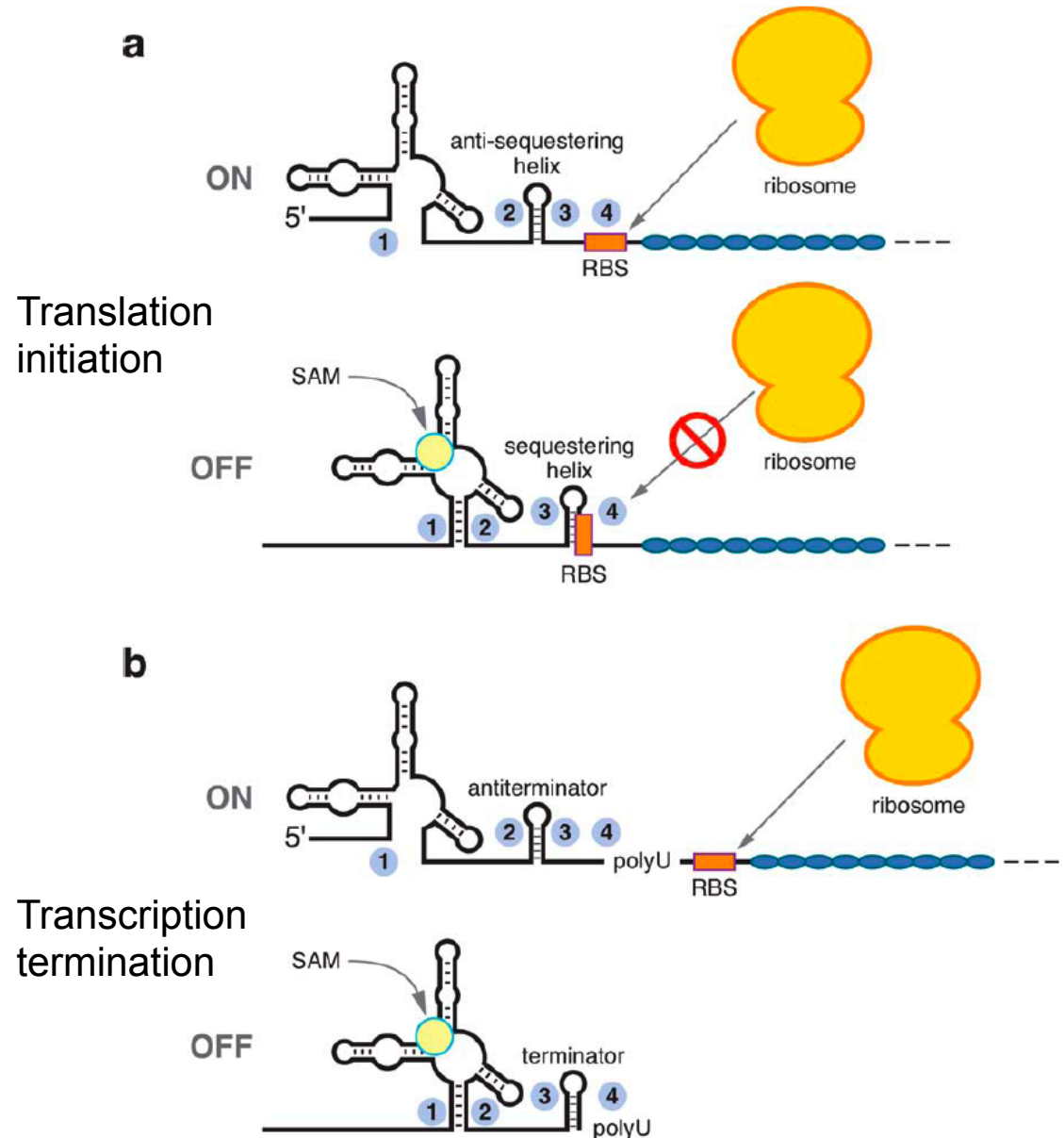
Schematic view of a riboswitch

Regulatory RNAs that control gene expression by binding metabolites (more recently tRNAs or metal ions or respond to temperature). They are typically located in non-coding regions of mRNAs.

See RR Breaker Cold Spring Harbor Perspectives in Biol. 2012 for review.

Database RFAM to check if your favorite RNA is a riboswitch of some sort:

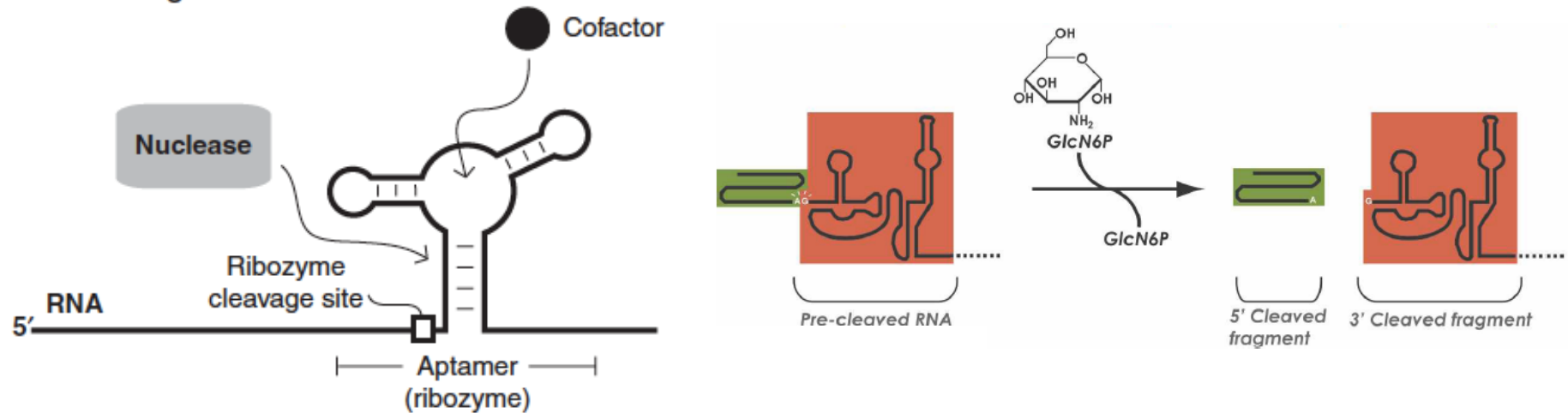
<http://www.sanger.ac.uk/resources/databases/rfam.html>



A riboswitch-ribozyme

(RNA molecules capable of catalyzing biochemical reactions, as protein enzymes)

F Self-cleaving



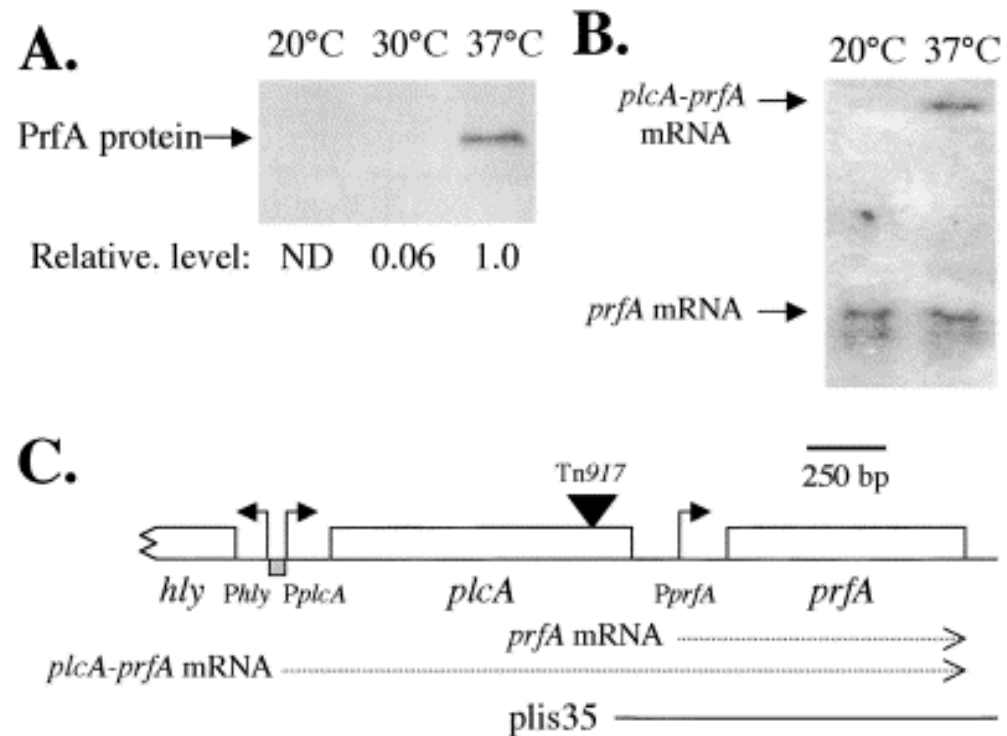
B. subtilis glmS gene

(see Collins et al. Genes & Dev 2007 for details)

Regulation in the RNA (pre-protein) world?

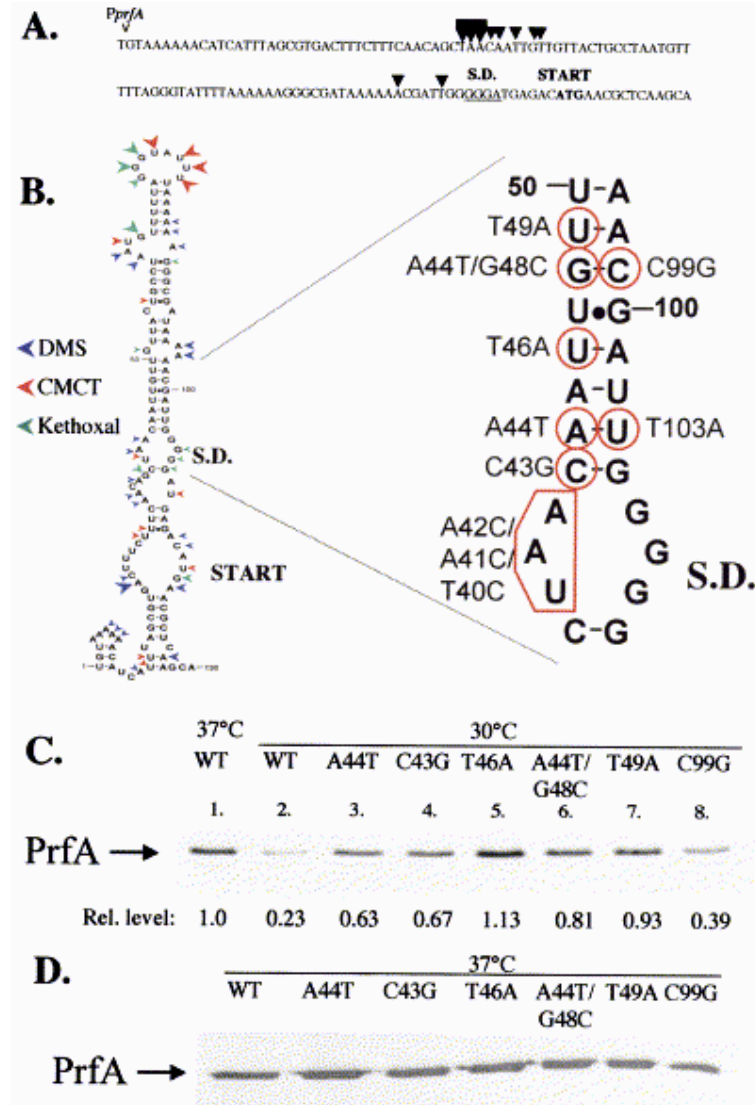
A RNA thermosensor

PrfA is the master regulator of the virulence of *Listeria*, the bacterium responsible for listeriosis.



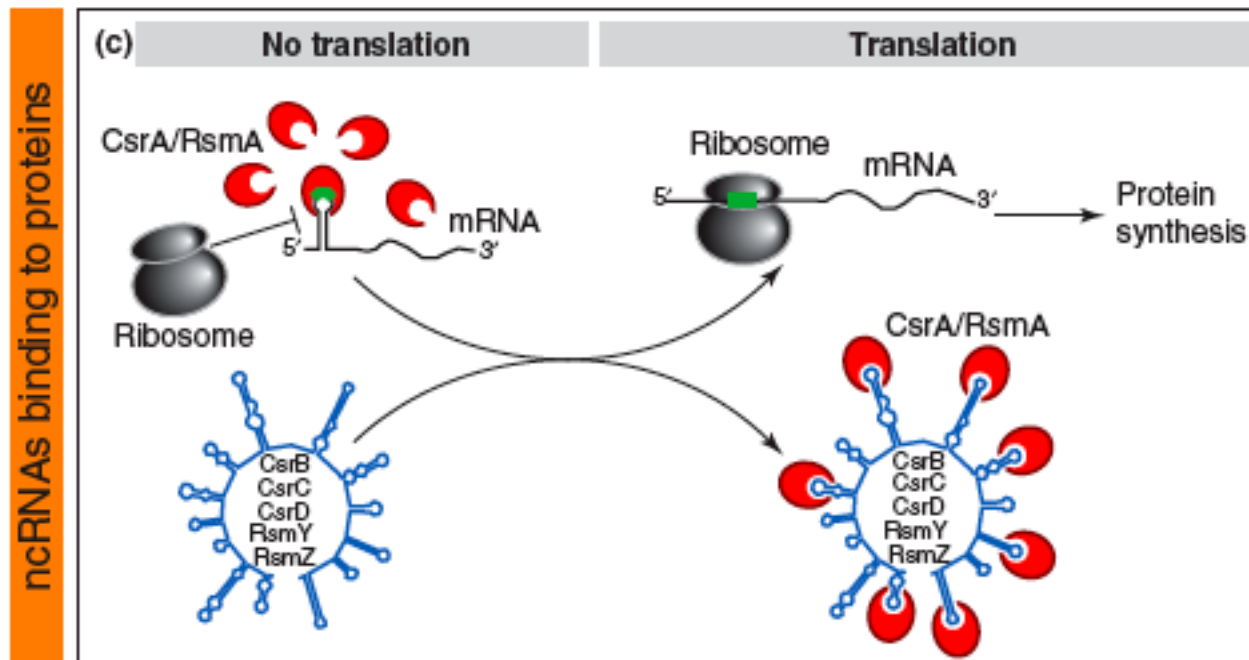
Johansson et al., Cell 2002

The effect is due to a stem, stable at low temperatures and sequestering SD, and getting unstable around $\approx 37^\circ$



Johansson et al., Cell 2002

Protein-interacting ncRNAs

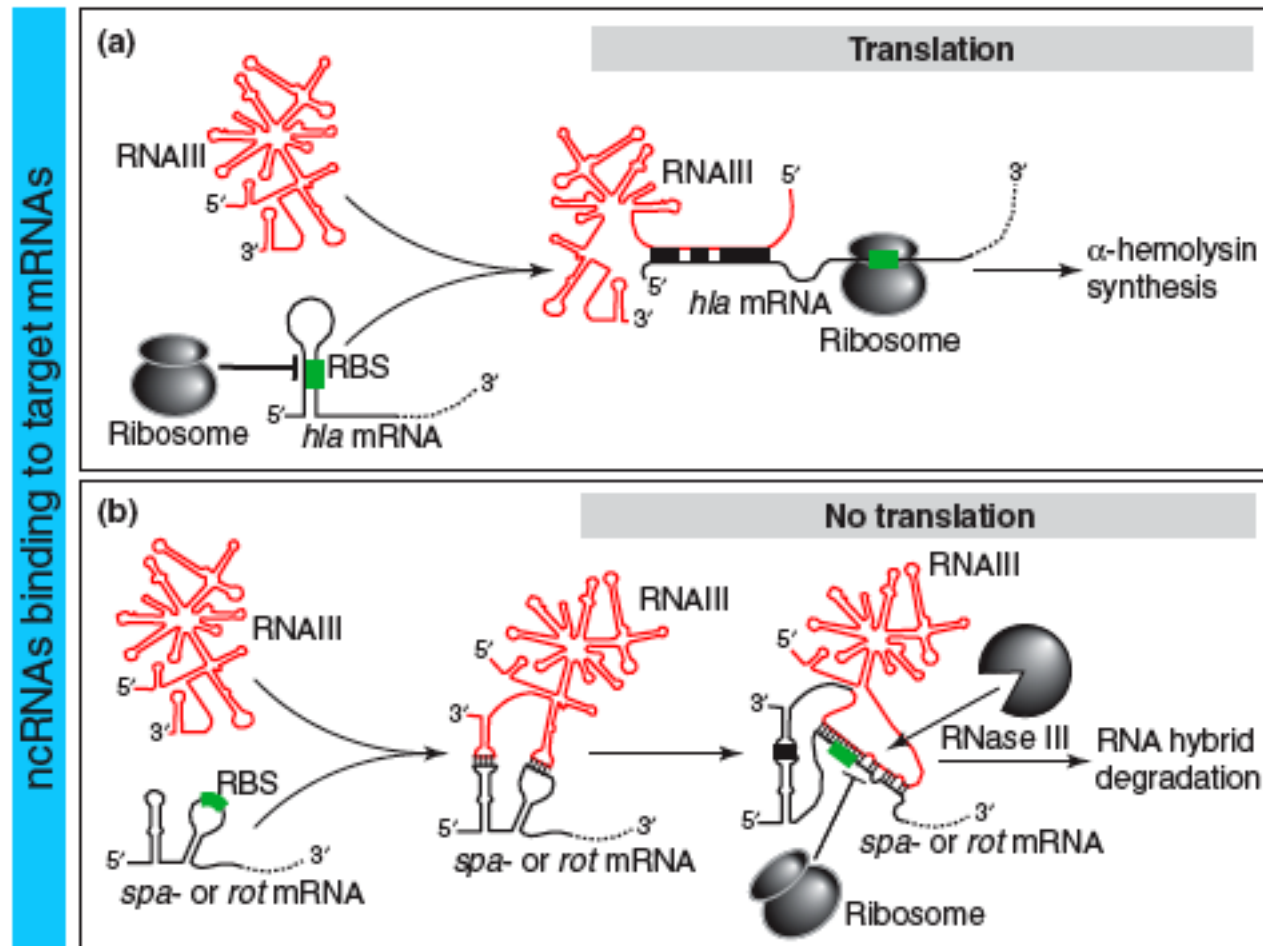


T. Romeo, Mol. Microb., 1998

mRNA-interacting ncRNAs

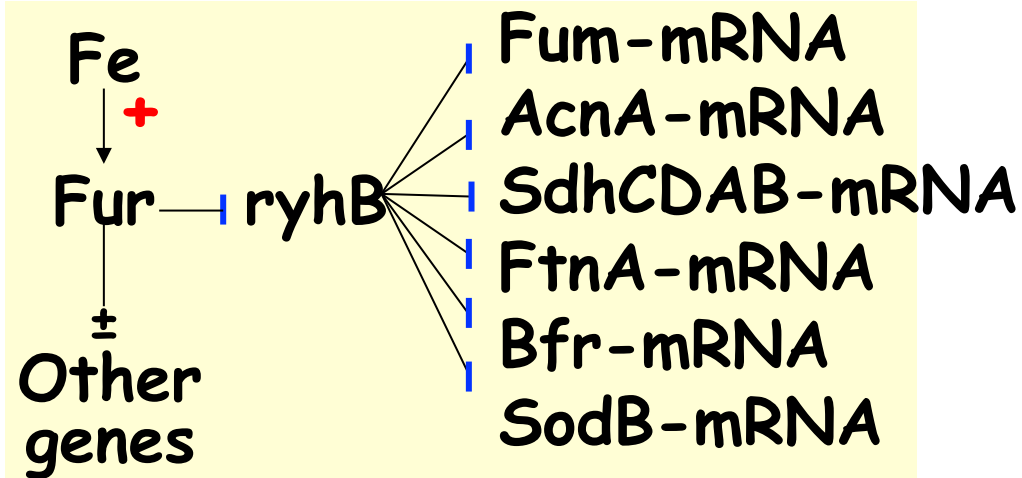
RNAIII is a 514nt long transcript of *S.aureus*.

The ncRNA regulates a set of genes involved in virulence. The expression of RNAIII is regulated by the agrA/C system.



RyhB regulates iron metabolism in *E. coli*

(Massé & Gottesman, 2002, 2003)



RyhB targets multiple mRNAs coding for iron-using and/or iron-storing proteins

RyhB expression is repressed by Fur (Fe^{2+} cofactor).

C. RyhB

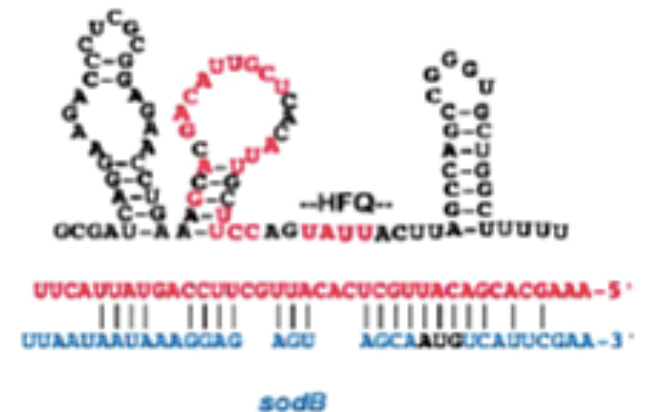
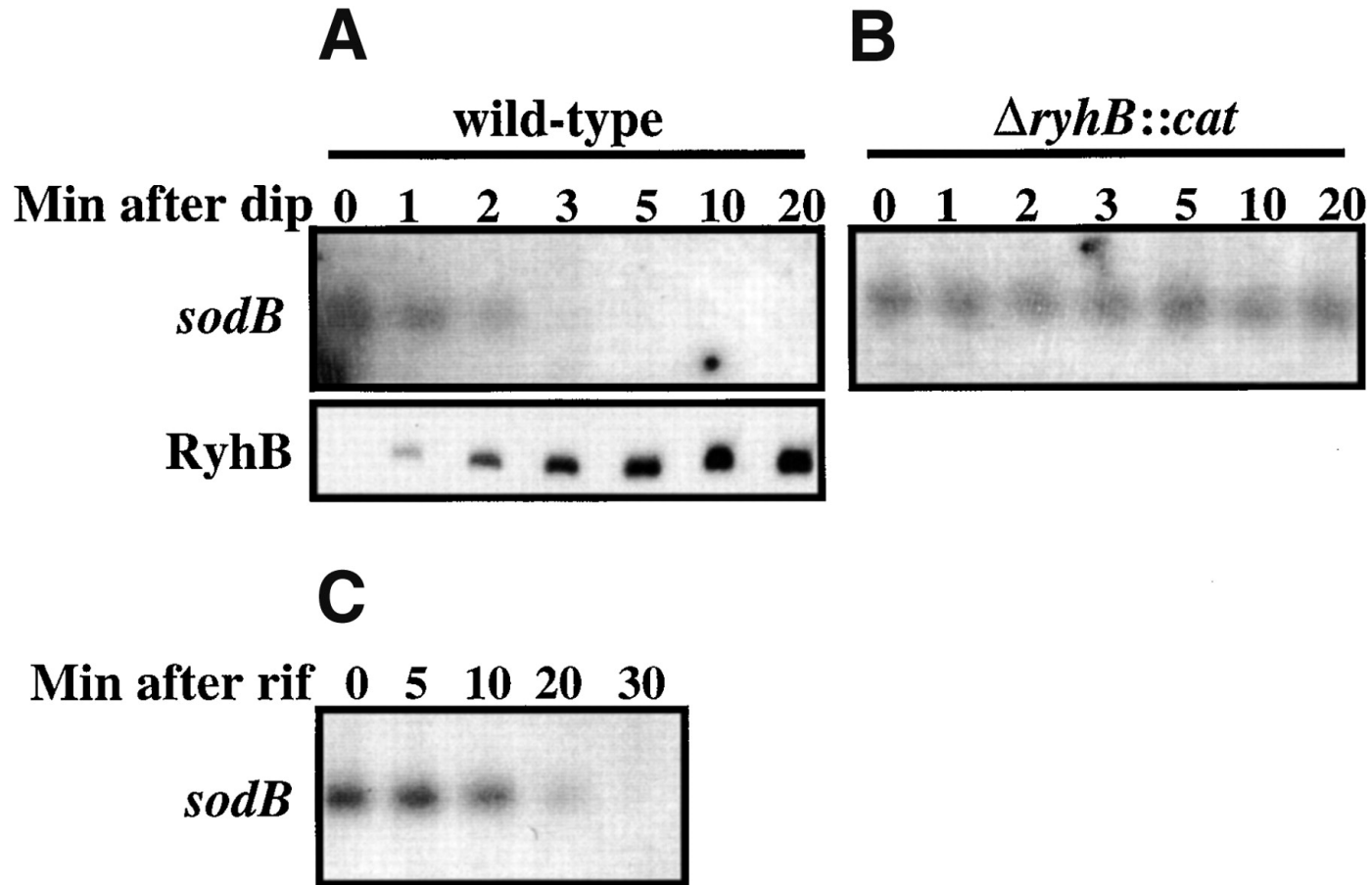
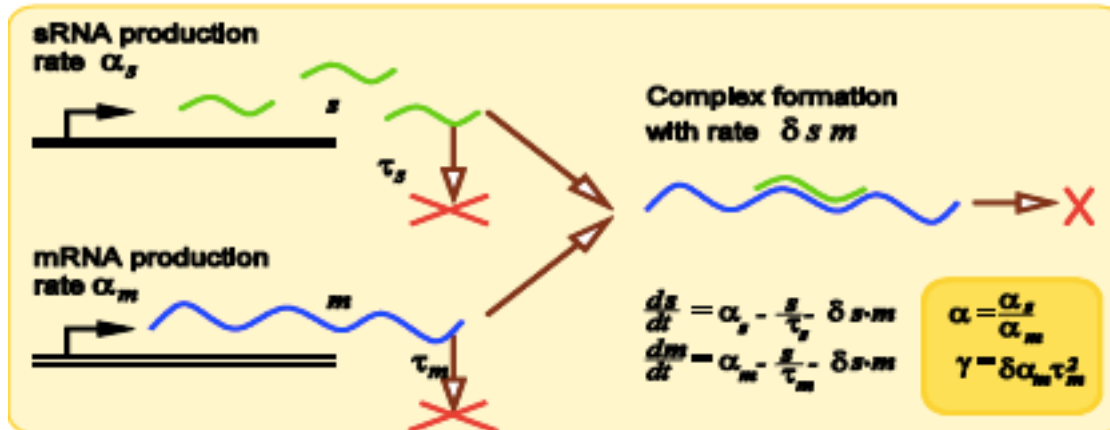


Figure 1. Degradation of full-length *sodB* mRNA



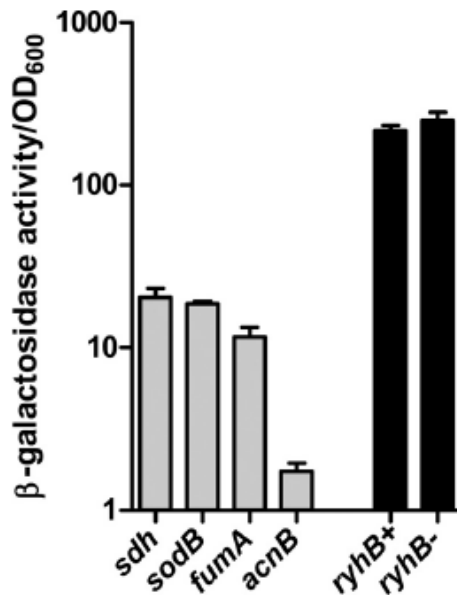
Eric Masse et al. *Genes Dev.* 2003; 17: 2374-2383

Modeling of the RyhB regulatory system

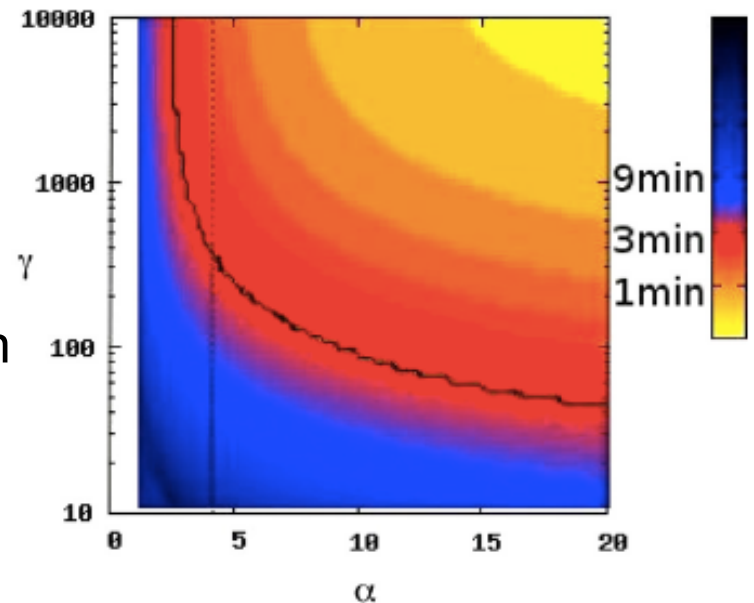


Mitarai et al.
Phys Biol. 2007;
PNAS 2010

α is determined by measuring the activity of promoters



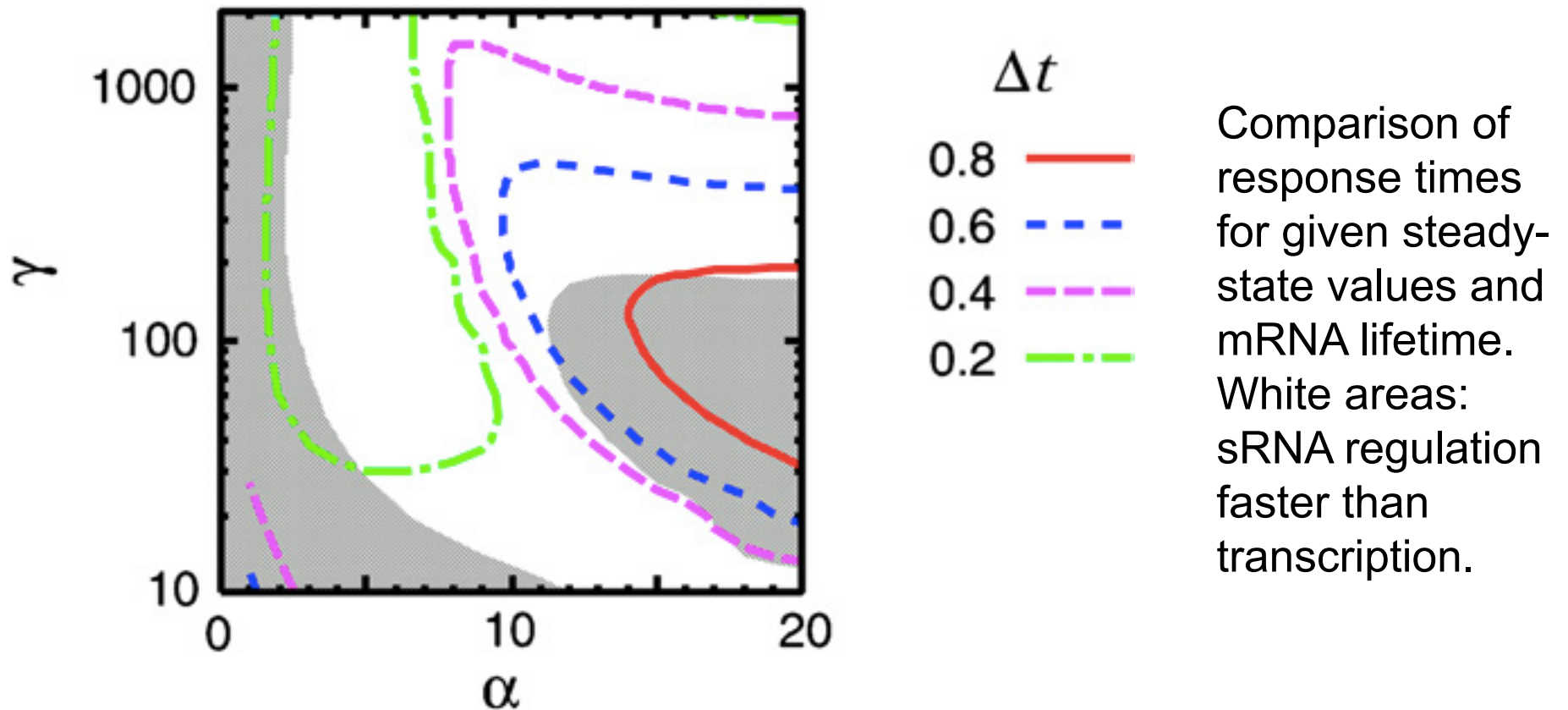
γ is determined by looking for a 3mins degradation time in the results of the model



τ 's are measured by the mRNA and the sRNA lifetimes

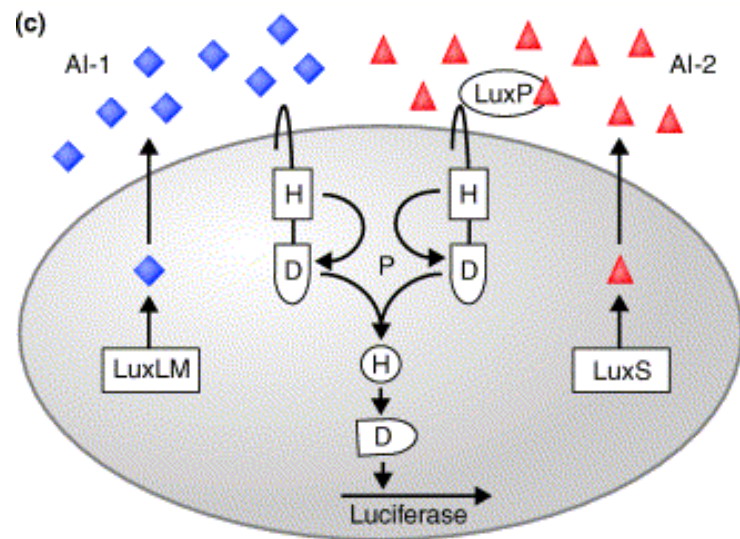
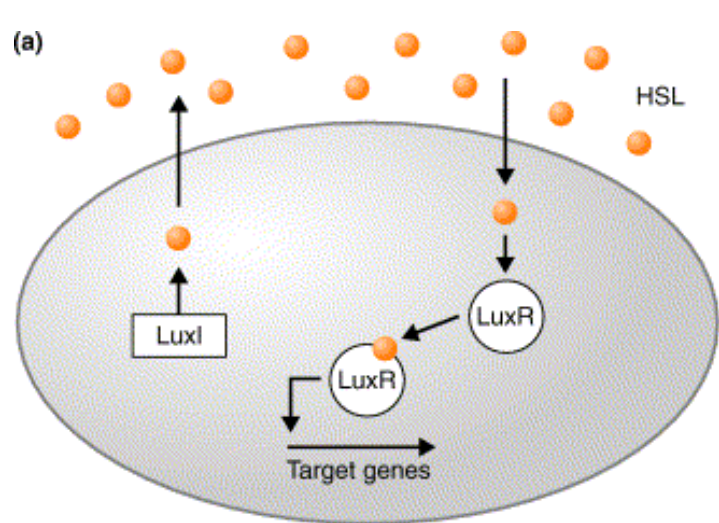
Comparison to pure transcription

Time for degradation by sRNA vs a purely transcriptional regulation



Mitarai et al. Phys Biol. 2007

Environmental conditions that bacteria must adapt to can be actively modified by the action of bacteria themselves



– density-dependent gene expression

– can detect multiple signaling molecules
– potential for complex language

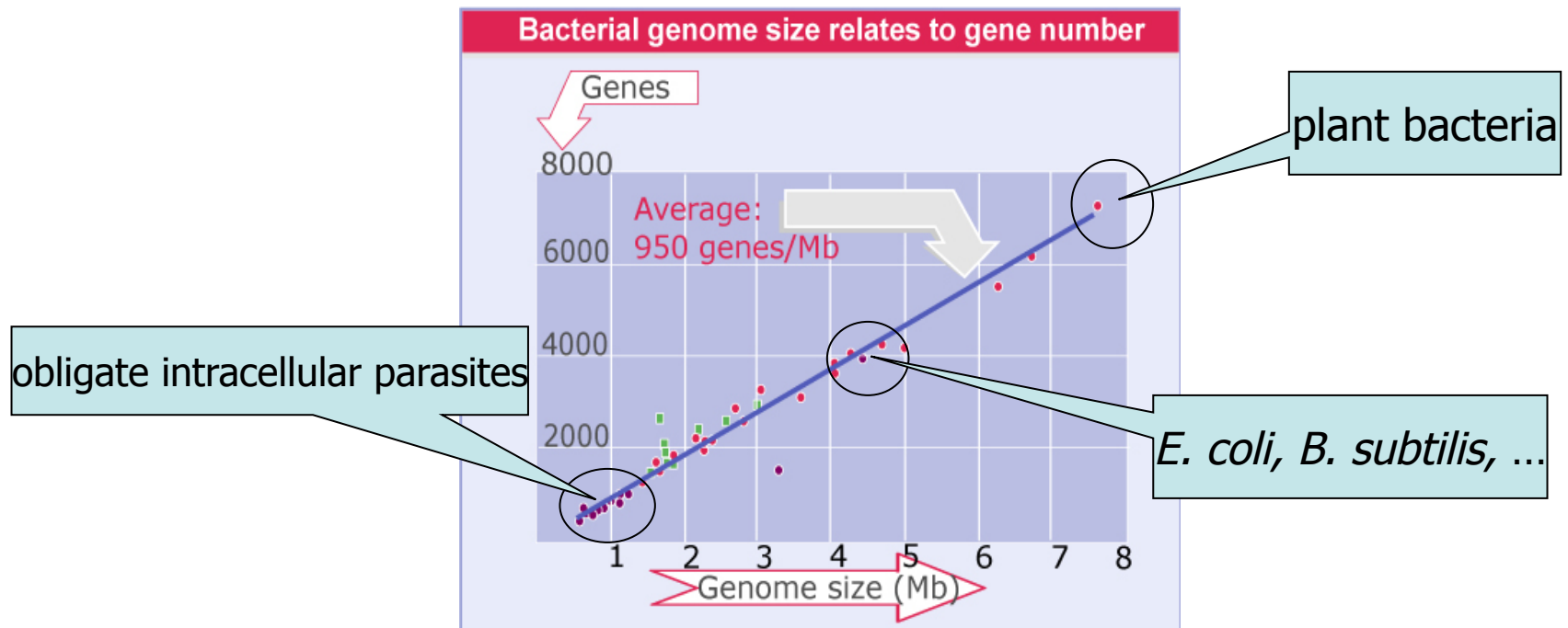
quorum sensing: inter-cellular communication

Overview of molecular microbiology

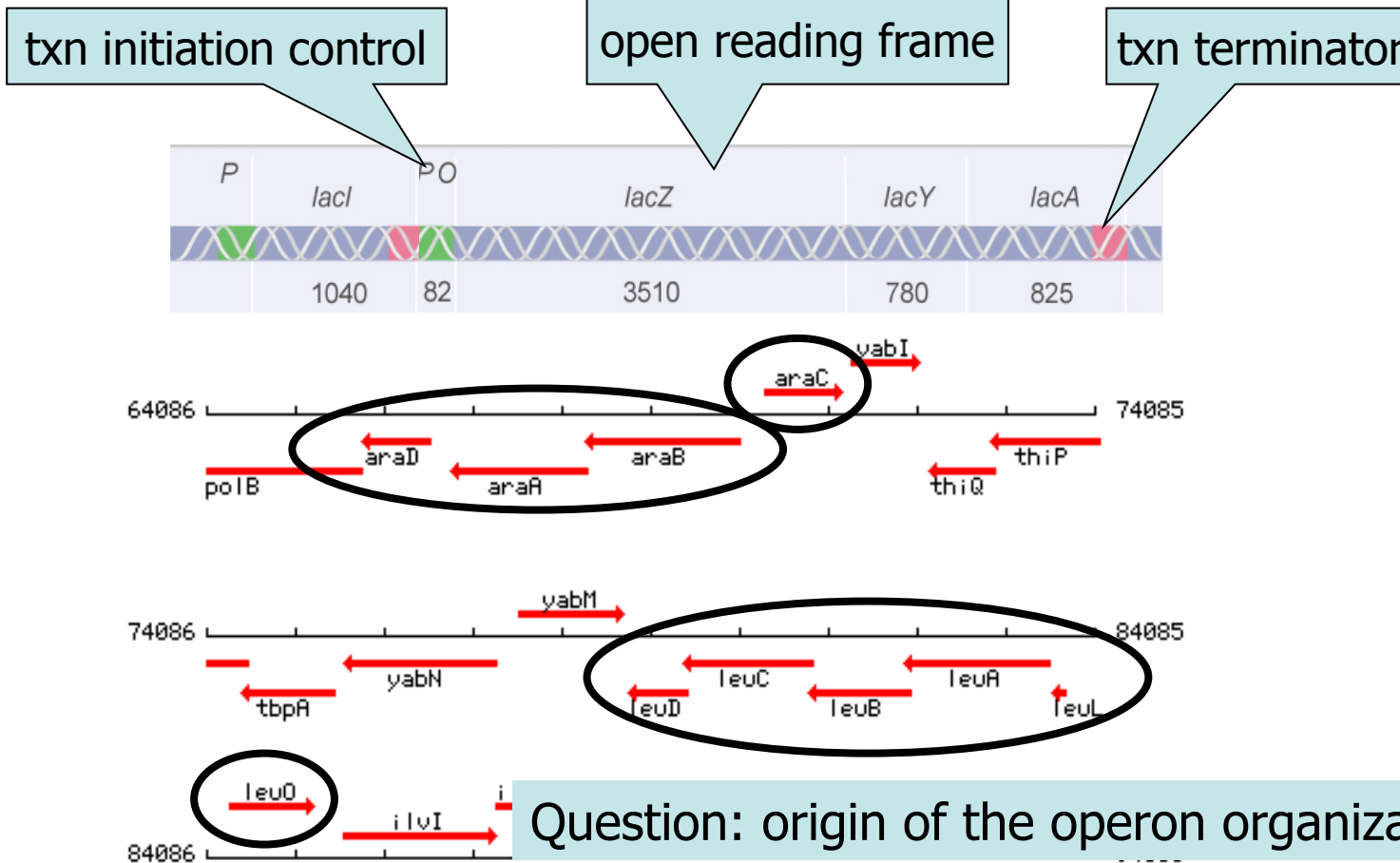
Plan

1. biochemical aspects
2. mechanistic aspects
3. regulatory aspects
4. **genomic aspects**
5. physical aspects
6. comparison to eukaryotes

- proteins:
 - amino acid chain; typically ~300 a.a. in length
 - each protein encoded by a “gene” (~1000 bases)
- bacterial genomes: circular chromosome(s)
 - [E. coli K-12 MG1655](#): 4.6Mb, ~4400 genes; 2/3 “identified”
 - [4 different strains of E. coli](#): 4.6 – 5.4 Mb (common: 3.8Mb)
 - [range for 150+ bacteria](#): 0.1 – 10 Mb

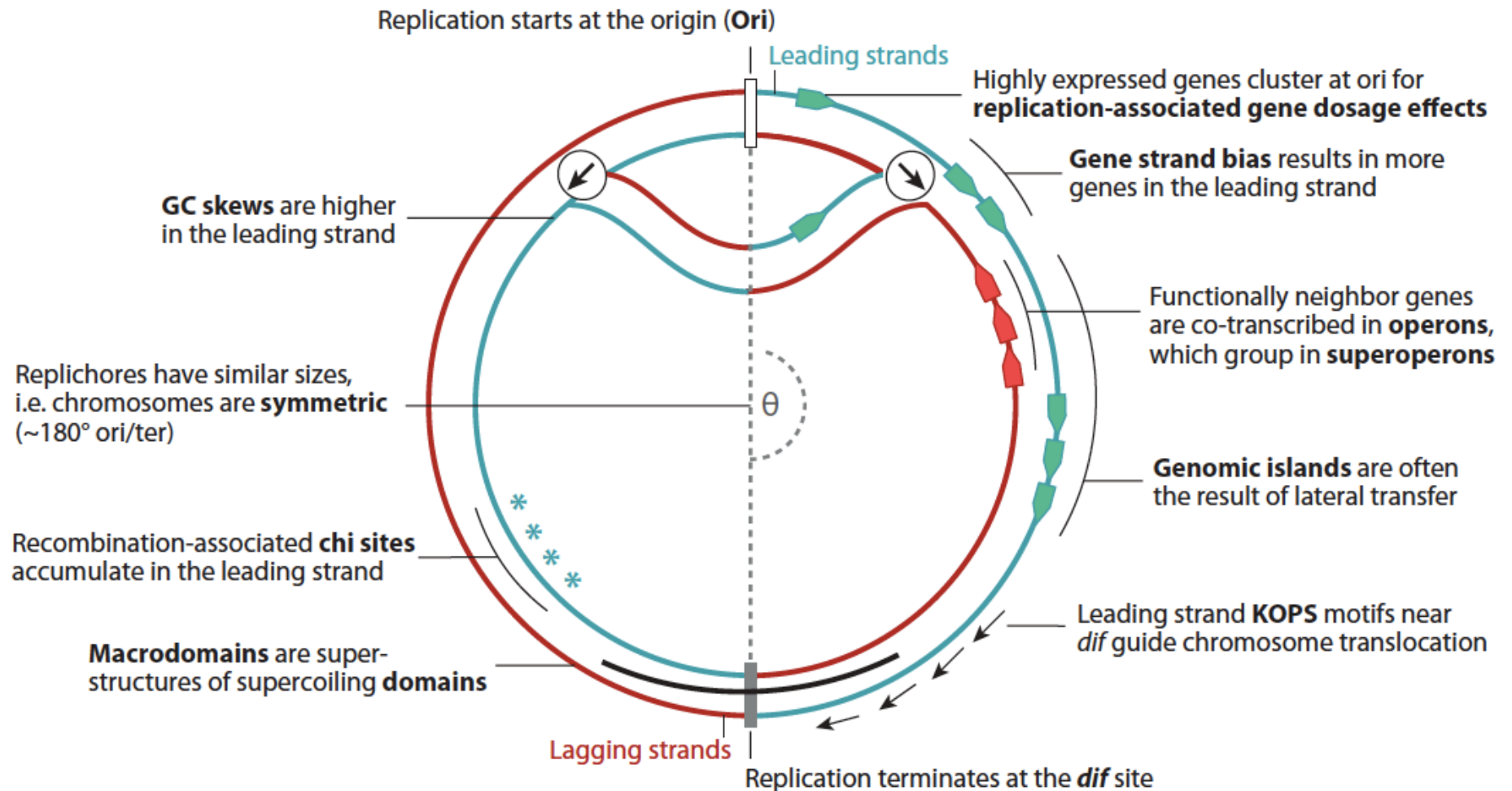


- approx 85% of the *E. coli* genome codes for proteins
- operon structure
 - genes coding for proteins that function in the same pathway may be located adjacent to one another and controlled as a single unit that is transcribed into a polycistronic mRNA called “operon”

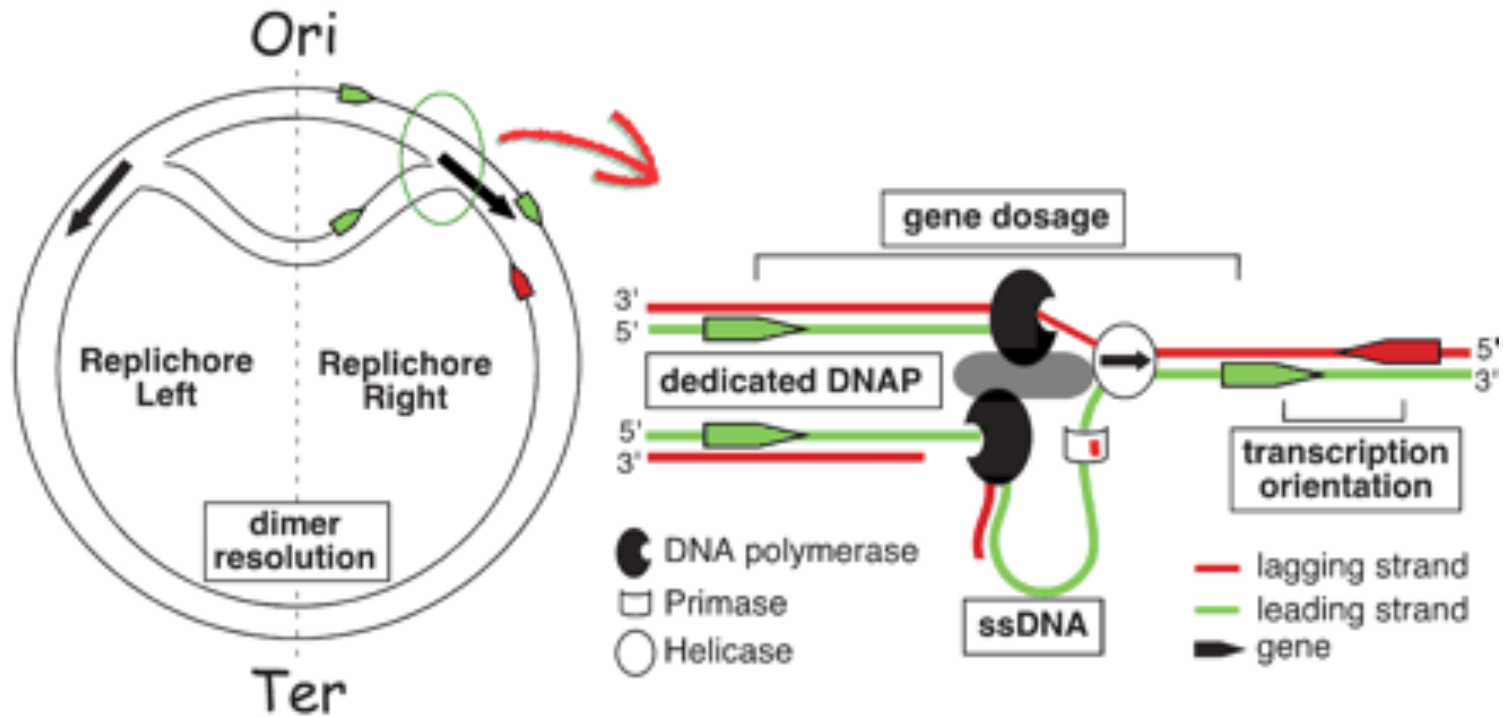


- regulator of an operon is often located nearby (and divergently transcribed)
- different operons involved in the same function can be located far away

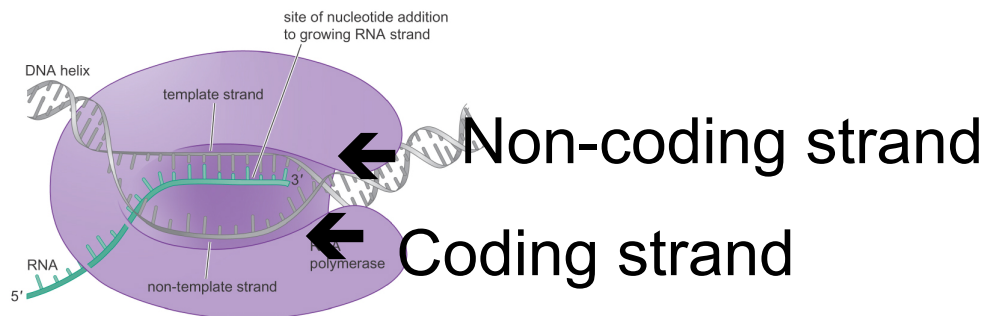
The organization of the bacterial genome



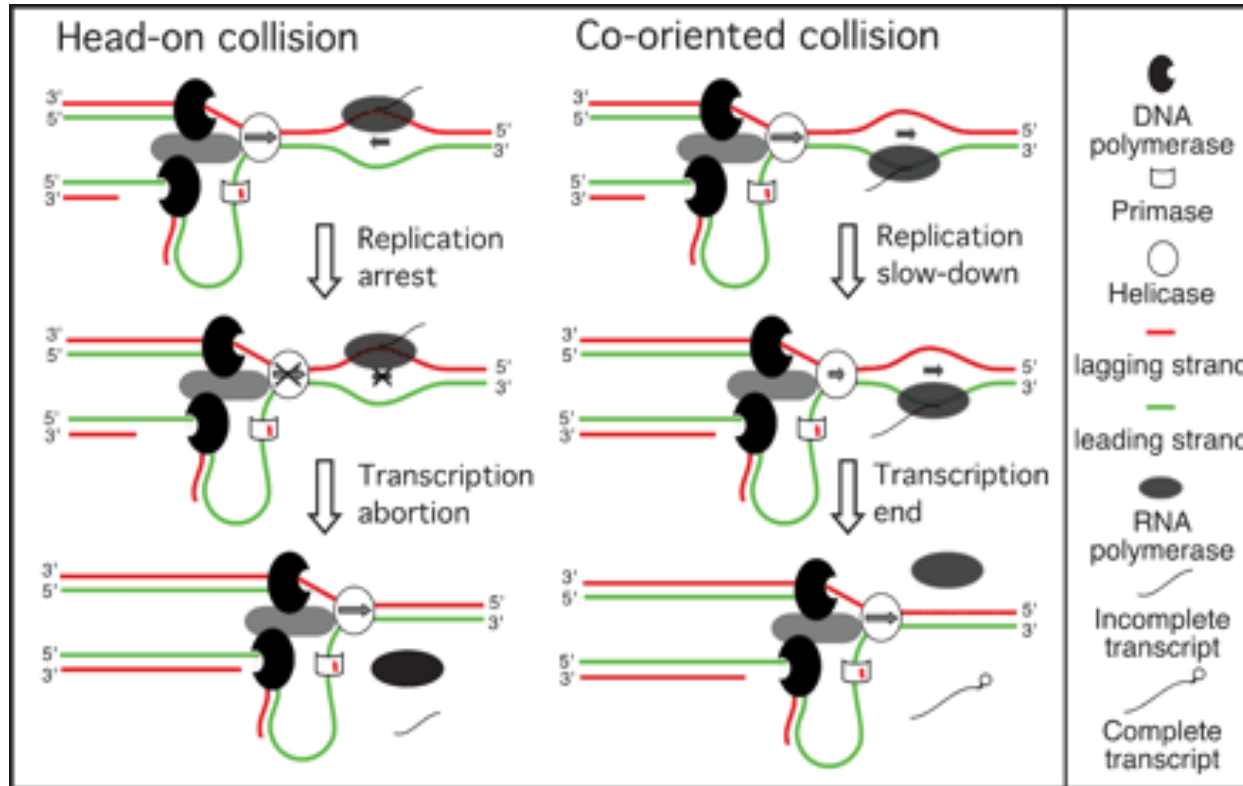
Interferences between transcription and replication



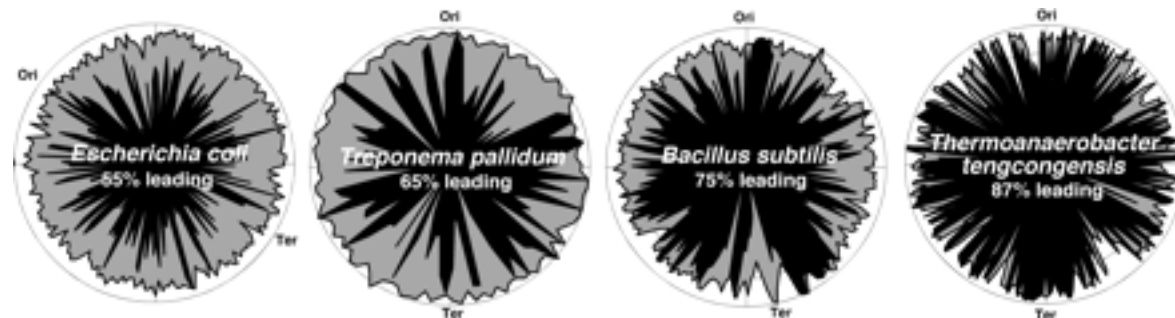
Both DNA and RNA are polymerized in the 5'→3' direction!



Effects of collisions and gene strand bias

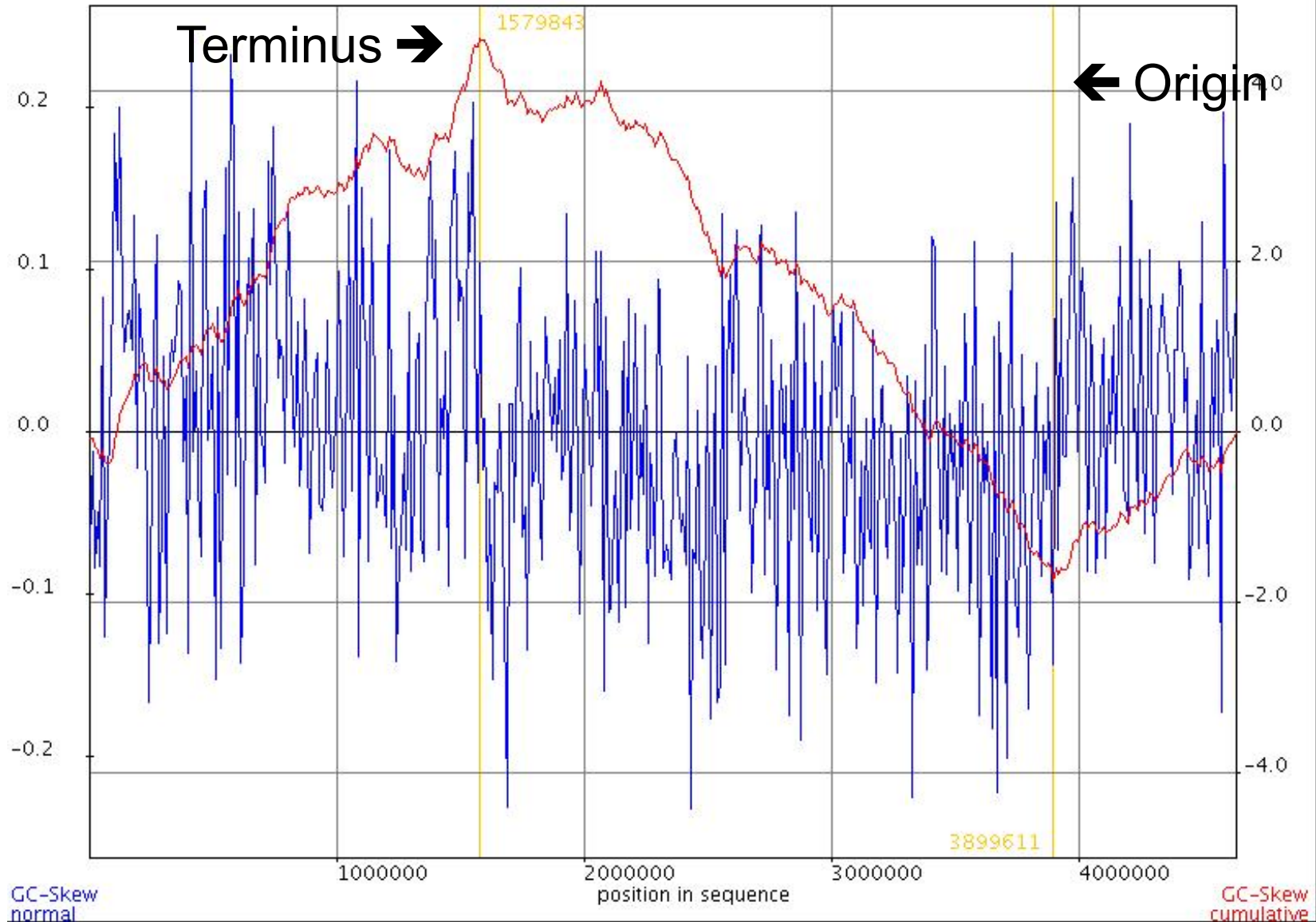


Collisions are inevitable: replication advances at $\approx 1\text{Kbp/s}$ and transcription at $\approx 50\text{bp/s}$



G/C skew on leading vs lagging strands

GC-skew plot for sequence ID: U00096.2 Desc: Escherichia coli K-12 MG



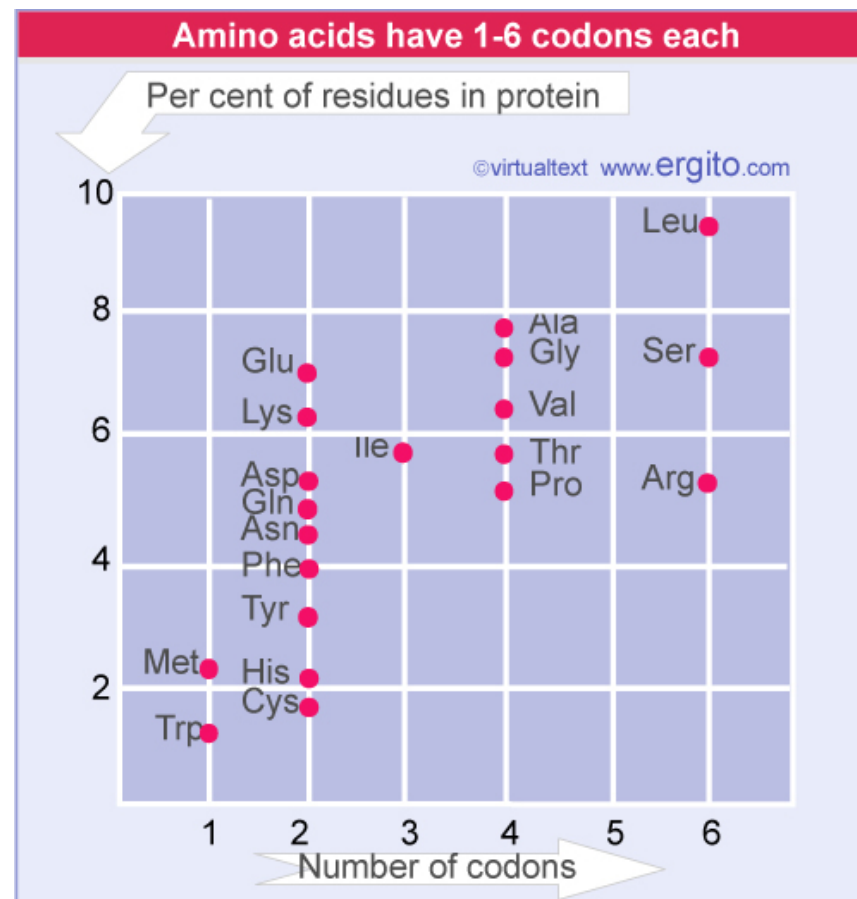
- other genomic elements and features
 - transfer RNA: 61 different types, 74-95 nt in length

The genetic code is triplet

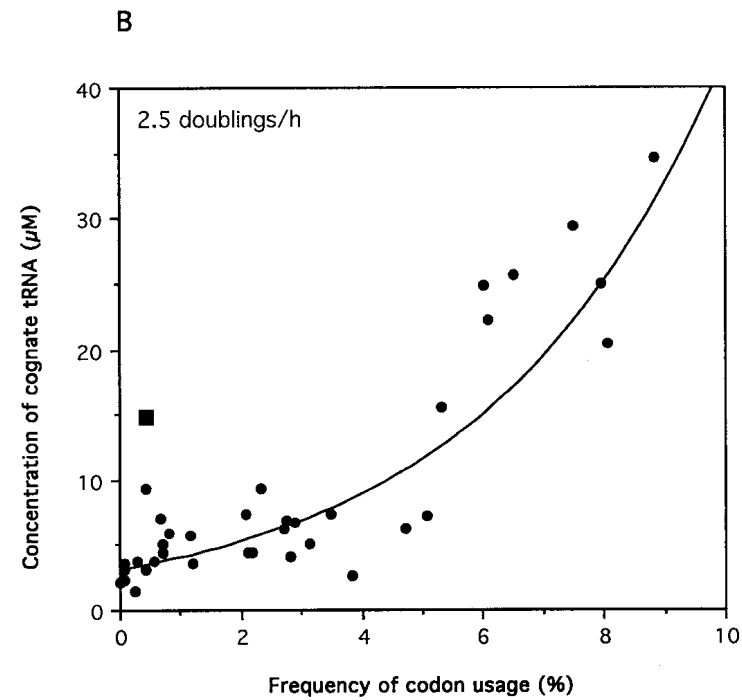
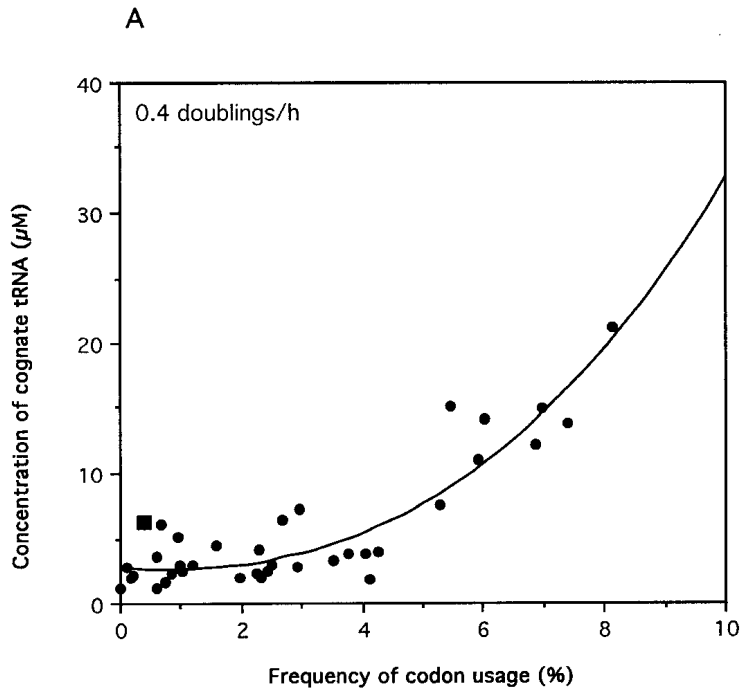
First base Second base

©virtualtext www.ergito.com

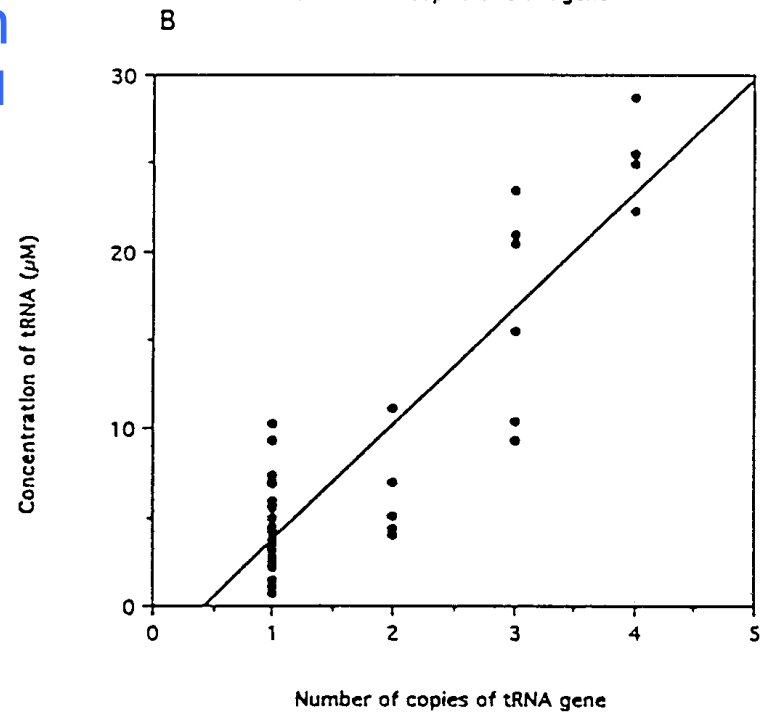
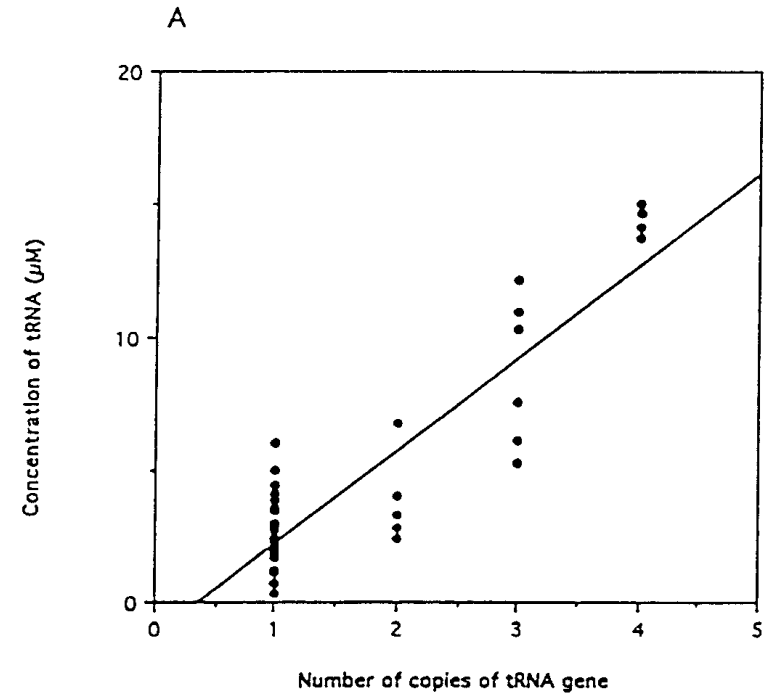
	U	C	A	G
U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA } STOP UAG }	UGU } Cys UGC } UGA } STOP UGG } Trp
C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }
A	AUU } AUC } Ile AUA } AUG } Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }
G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }



- ➔ codons not equally represented: 86 tRNA genes in *E. coli* MG1655
- ➔ codons not equally used in coding sequences [= codon bias]
 - (strong bias especially in highly expressed genes)



Codon bias
and cognate
tRNA
concentration
(Dong et al. J. Mol
Biol., 1996)



- plasticity of the genome

- mutation ≠ passive decay of genome

- o *Deinococcus radiodurans*: Withstands 1.5 Mrads, dessication, starvation, UV light, hydrogen peroxide; grows well at 6 krad/h

- o mutator strain: can increase mutation rate 1000-fold

- base-pair substitution (BPS): 10^{-9} /base/replication

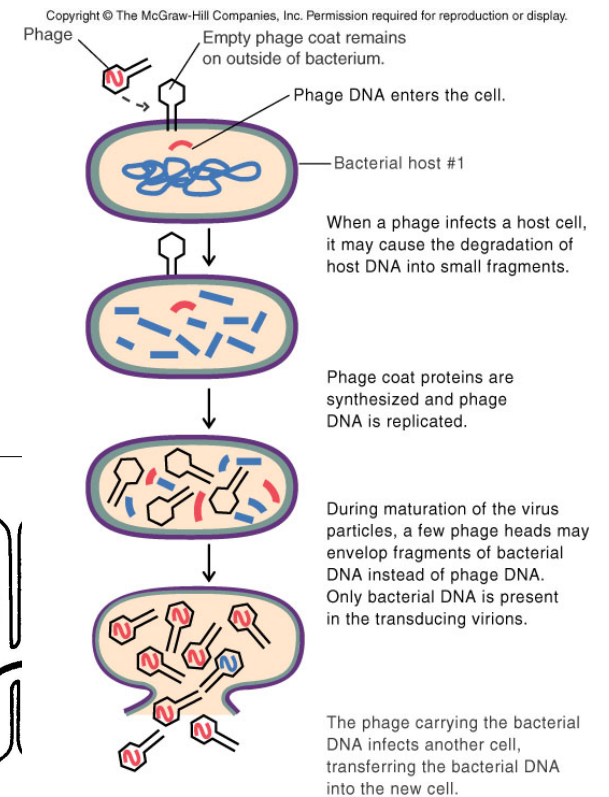
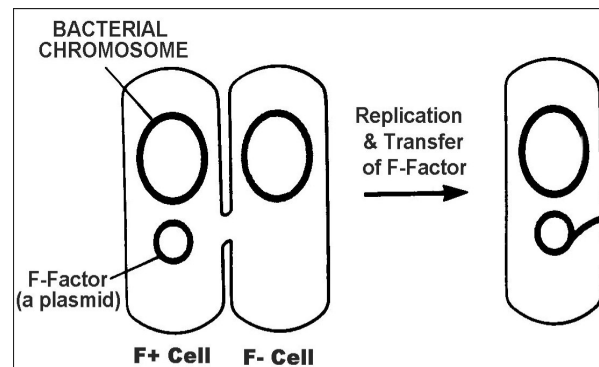
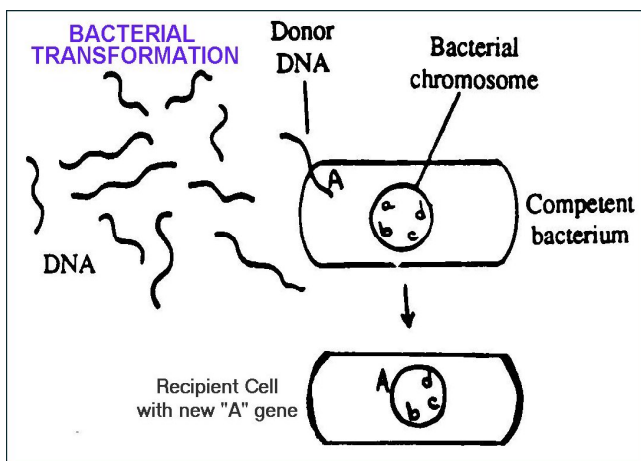
- insertions and deletions: ~1/3 of BPS rate

- horizontal transfer

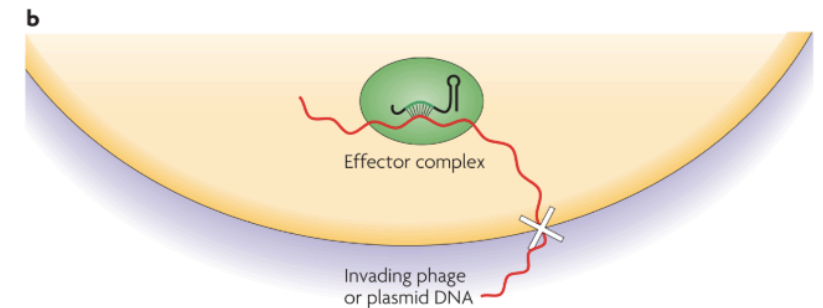
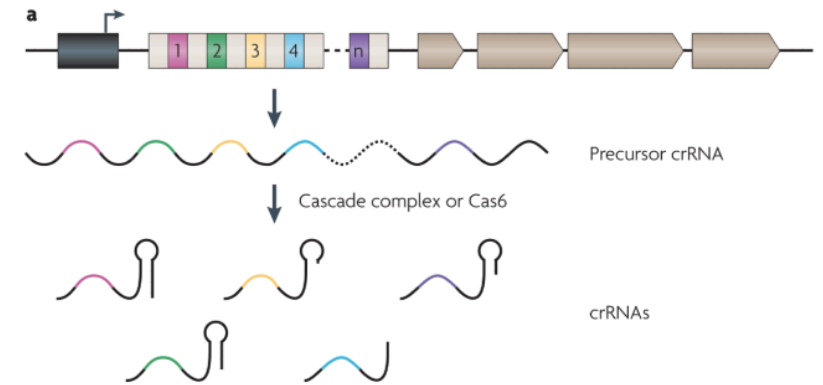
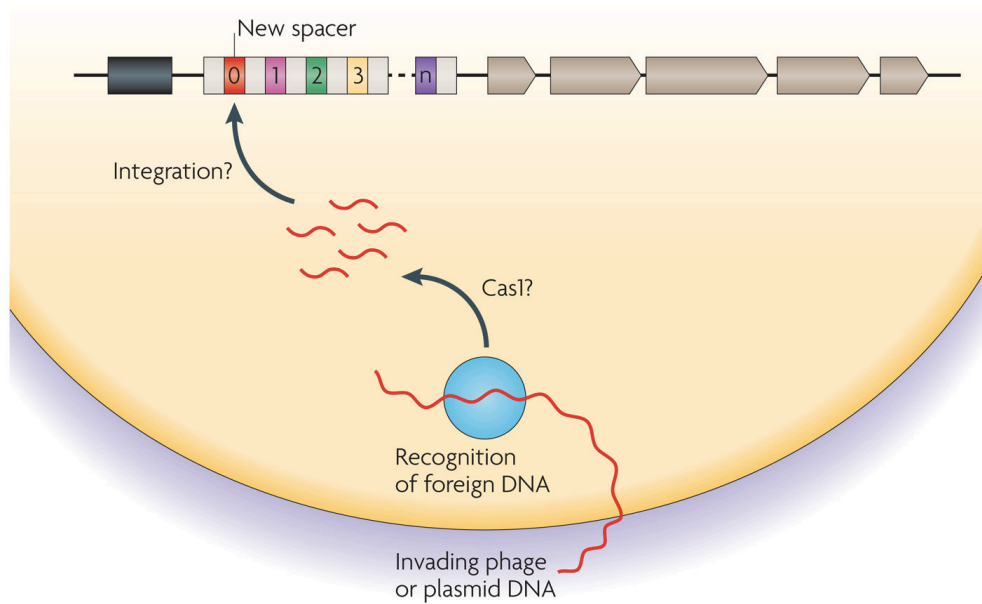
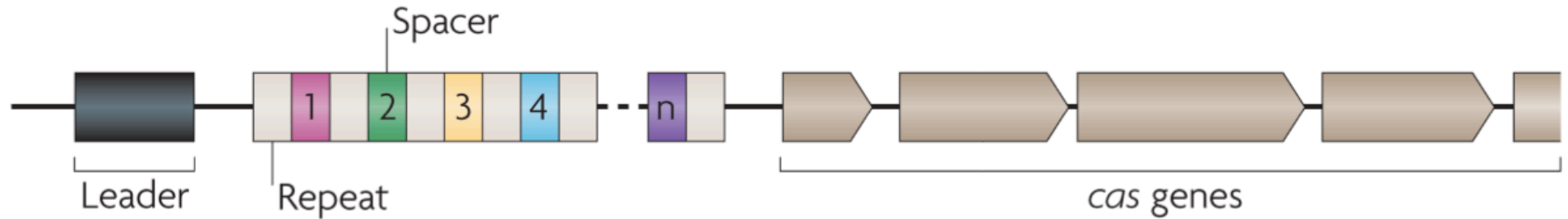
- competence

- conjugation

- transduction



CRISPR loci: a bacterial adaptive immune system

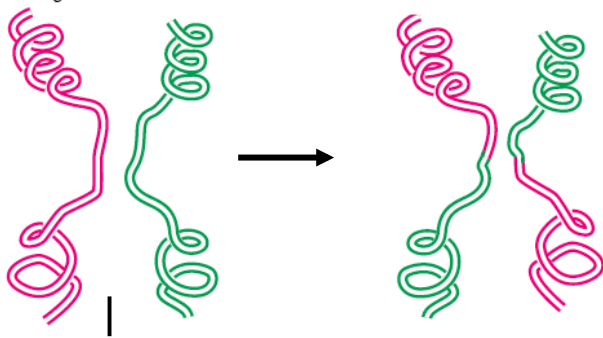


Self vs non-self?

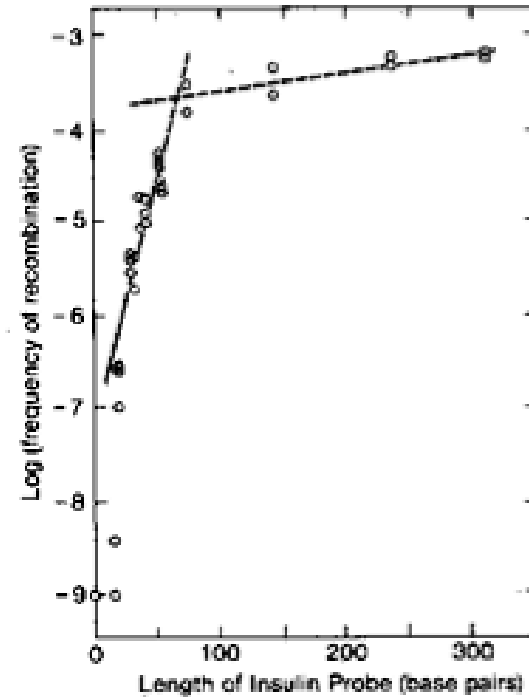


Marraffini & Sontheimer, Nat. Rev. Gen. 2010

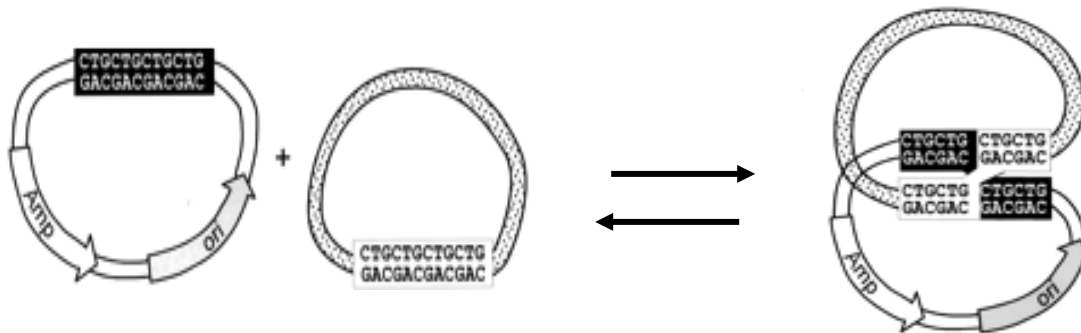
➤ homologous recombination



requires reasonably long regions of homology



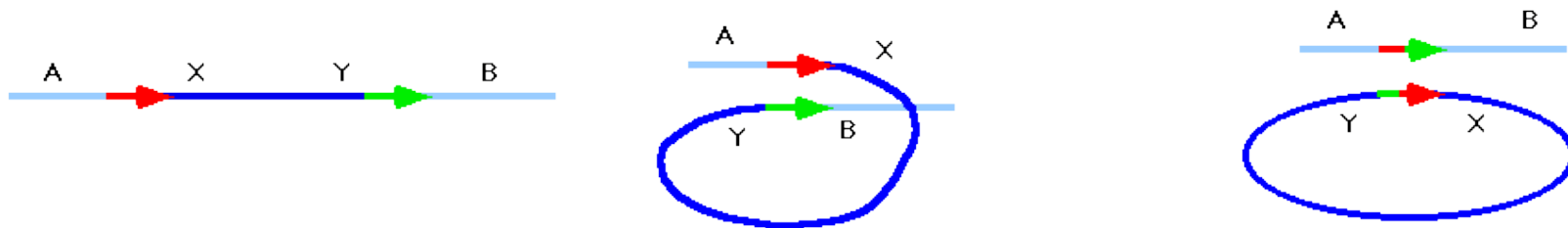
➤ site-specific recombination



[Watt et al, 85]

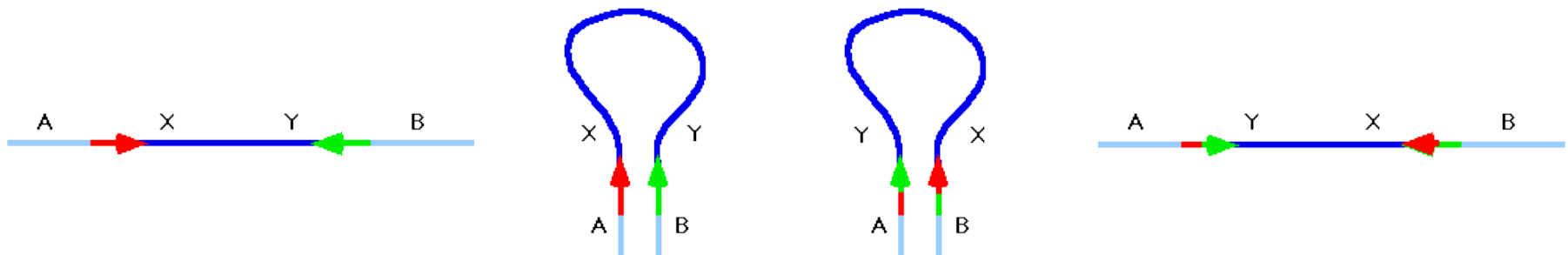
only occurs for specific sequences; requires special proteins

➤ consequence of recombination



direct repeat (e.g., insertion sequences)

- recombination releases material between repeats as circular molecule [may be transferred between cells]
- reverse process: chromosomal integration



inverted repeat:

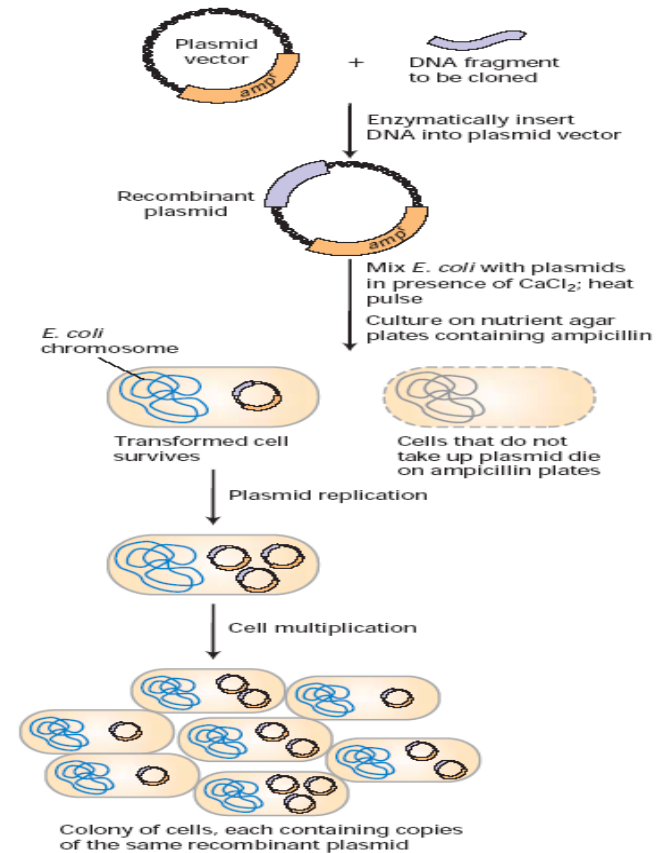
- sequence inversion (important mechanism of the evolution of gene regulation)

- plasmids:
 - types:

plasmids	F, P1	R1	pSC101	ColE1
copy #	1-2	3-5	~10	~30
length	90 kb	102 kb	6.5 kb	7.2 kb

- maintenance: provides a service to the host
 - e.g., poison-antidote system, antibiotic marker (genetic engineering)

Fluctuations and control of plasmid copy number (see Wong-Ng et al, Phys Rev E, 2010)



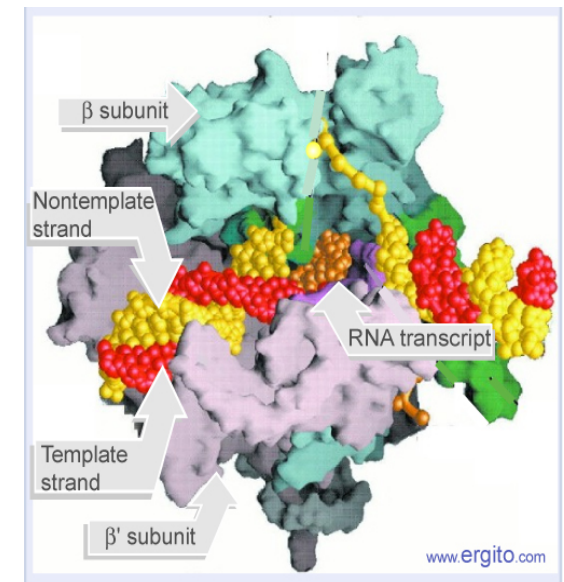
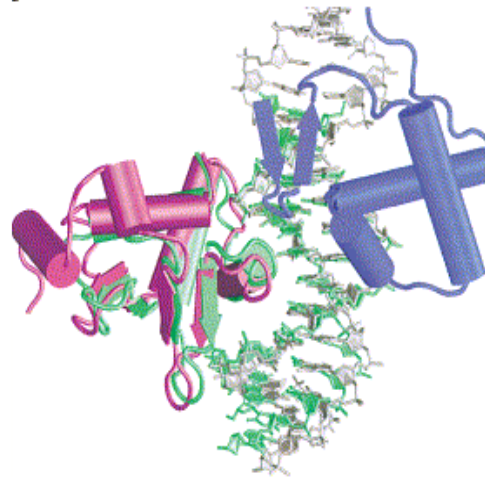
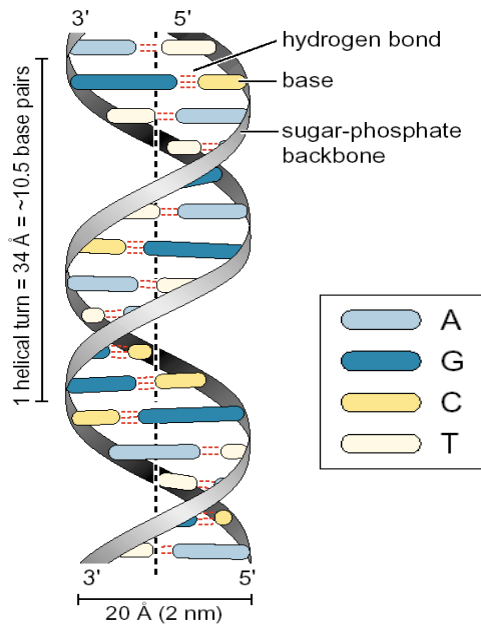
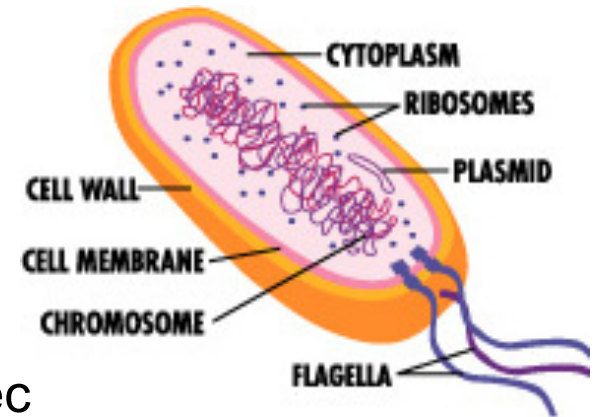
Overview of molecular microbiology

Plan

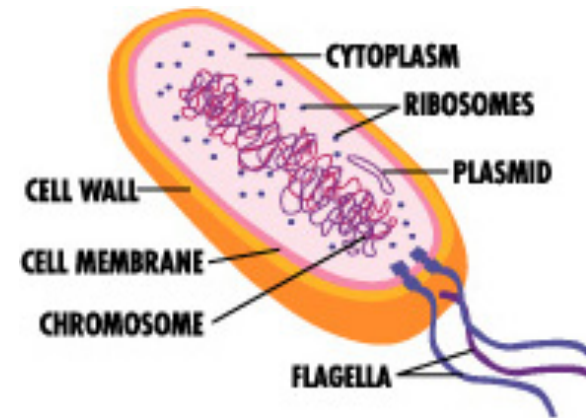
1. biochemical aspects
2. mechanistic aspects
3. regulatory aspects
4. genomic aspects
5. physical aspects
6. comparison to eukaryotes

- dimensions

- DNA: 2 nm x 2 nm x 3.4 nm/turn
- small proteins: (few nm)³ or ~10nt
- protein complexes, (10-20 nm)³ or 30 ~ 60nt
- cell size: 1 μm² x 3 μm
- concentration: 1 molecule/cell ~ 1nM
- intracellular diffusivity of protein: ~10 μm²/sec



- abundance
 - ribosomes: ~ 20,000 (52 proteins each)
 - RNAP ~ 1,000 (a few pct available)
 - proteins: 2×10^6 (TF: 10 ~ 1,000 / type)
 - mRNA: small fraction of RNA
0.1 ~ 100 copies/cell;
peaked at 2 ~ 3 copies / cell

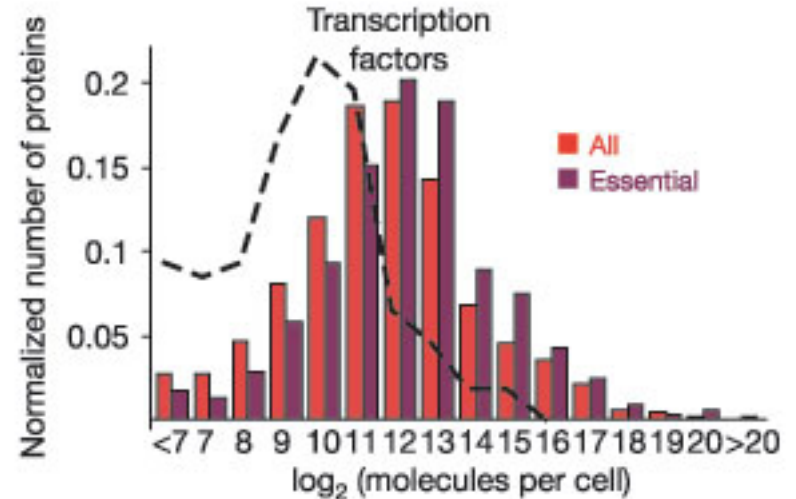


25% of bacterial dry mass is concerned with gene expression

Component	Dry Cell Mass (%)	Molecules /cell	Different types	Copies of each type
Wall	10	1	1	1
Membrane	10	2	2	1
DNA	1.5	1	1	1
mRNA	1	1,500	600	2-3
tRNA	3	200,000	60	>3,000
rRNA	16	38,000	2	19,000
Ribosomal proteins	9	10^6	52	19,000
Soluble proteins	46	2.0×10^6	1,850	>1,000
Small molecules	3	7.5×10^6	800	

@virtualtext www.ergito.com

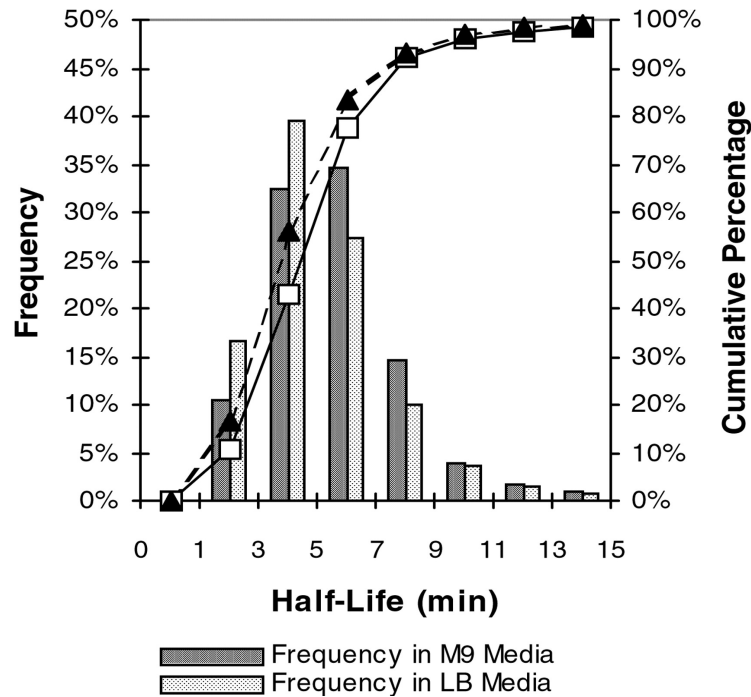
b



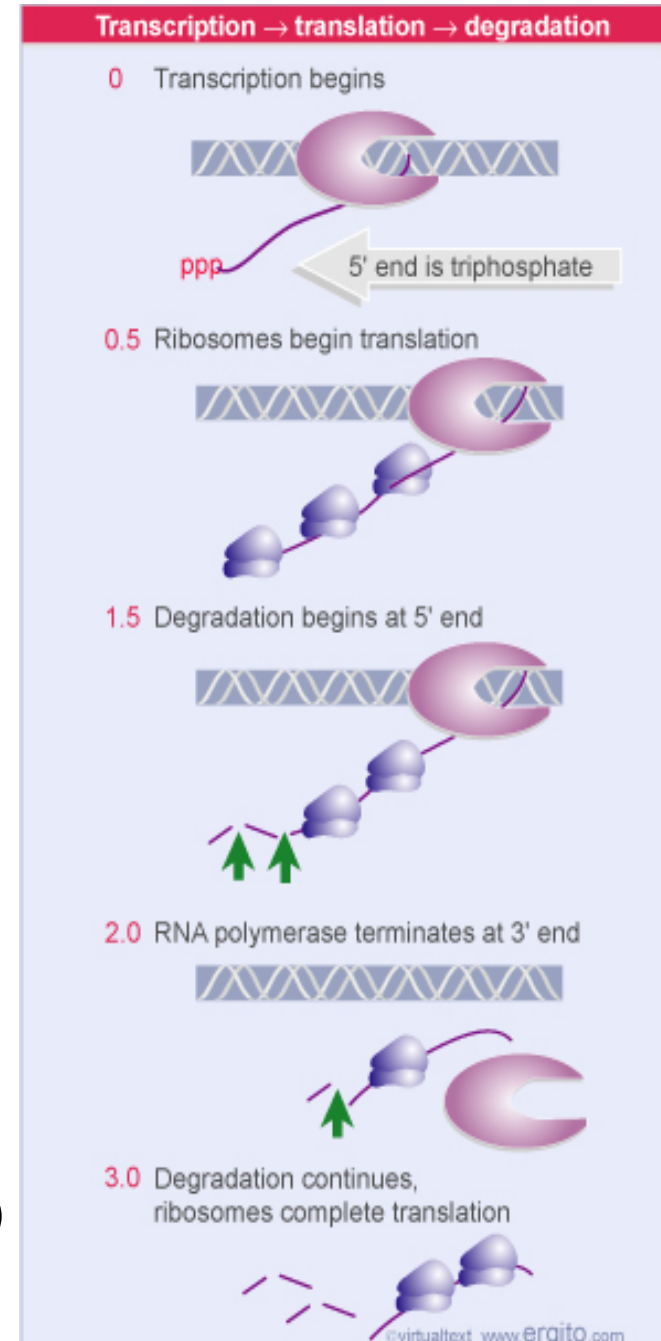
Warning: many of these numbers are growth-rate dependent!

- rates

- transcription: elongation ~40 nt/s
- translation: ~ 15 aa/s
- transcription-translation coupling: infrequently translated mRNA cleaved
- mRNA half-life: < 5 min



- protein half-life: from cell-doubling time (passive decay) down to a few min (active proteolysis)



Overview of molecular microbiology

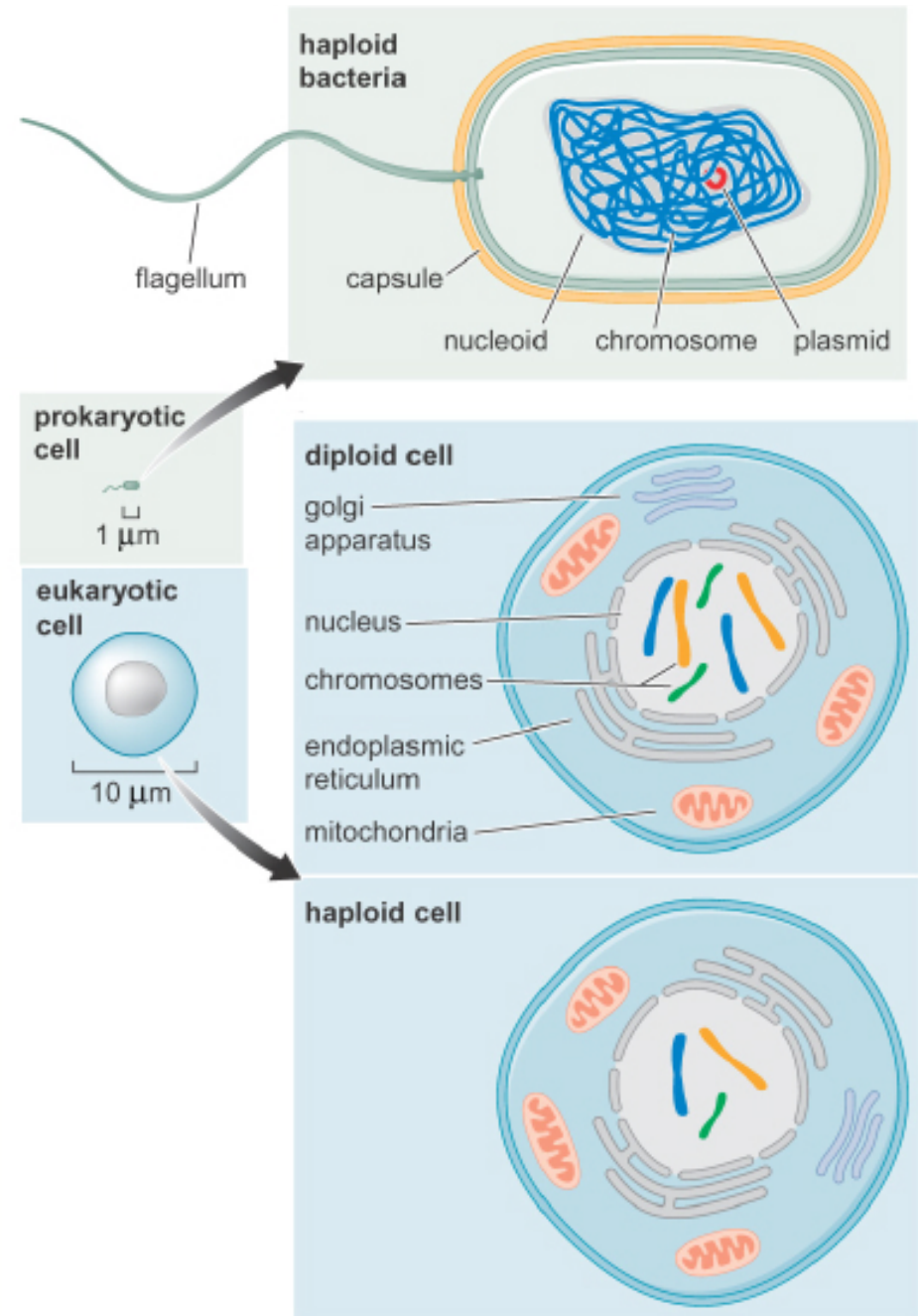
Plan

1. biochemical aspects
2. mechanistic aspects
3. regulatory aspects
4. genomic aspects
5. physical aspects
6. comparison to eukaryotes

❖ size and structure

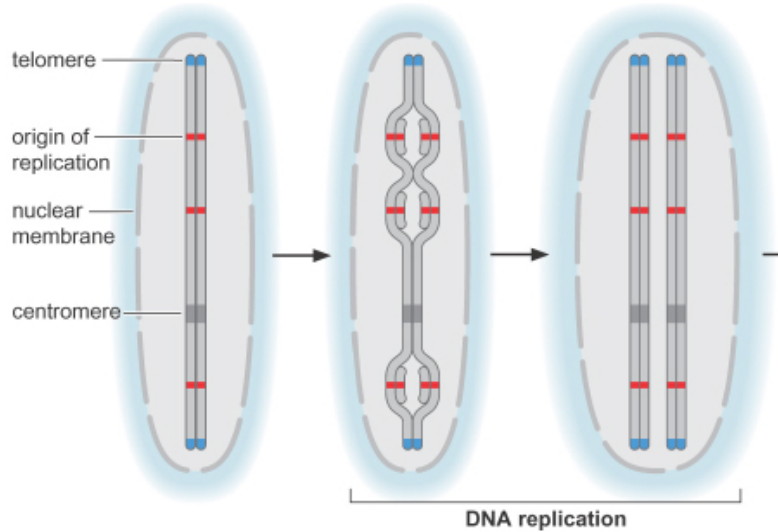
eukaryotes:

- cells ~1000x larger
- DNA in nucleus
- linear chromosomes
- haploids and diploids
- organelles
- vesicles
- microtubule network
- ...

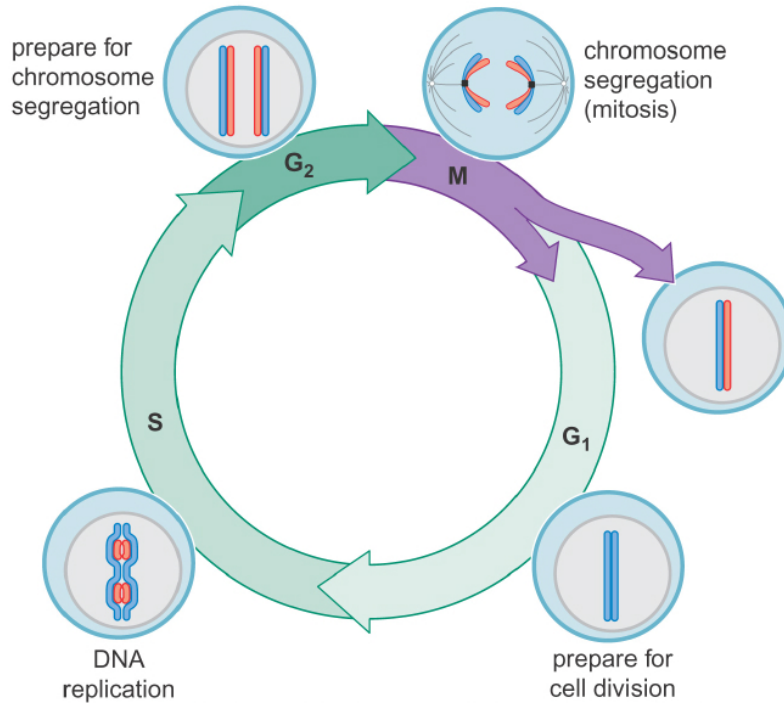


❖ Chromosome

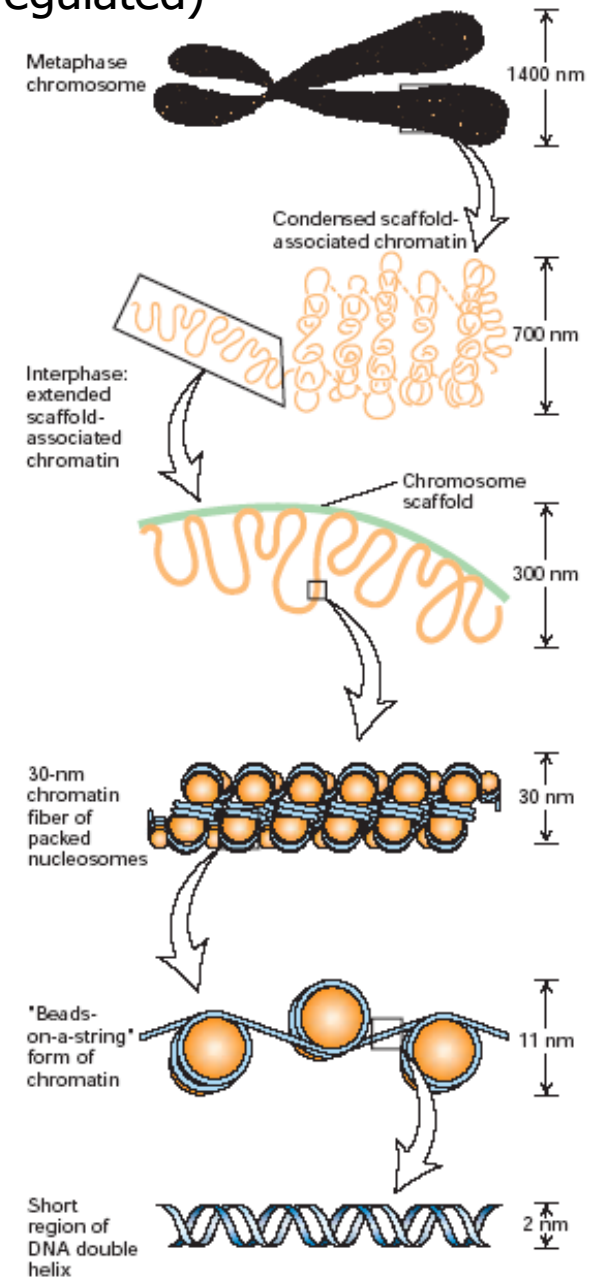
duplication



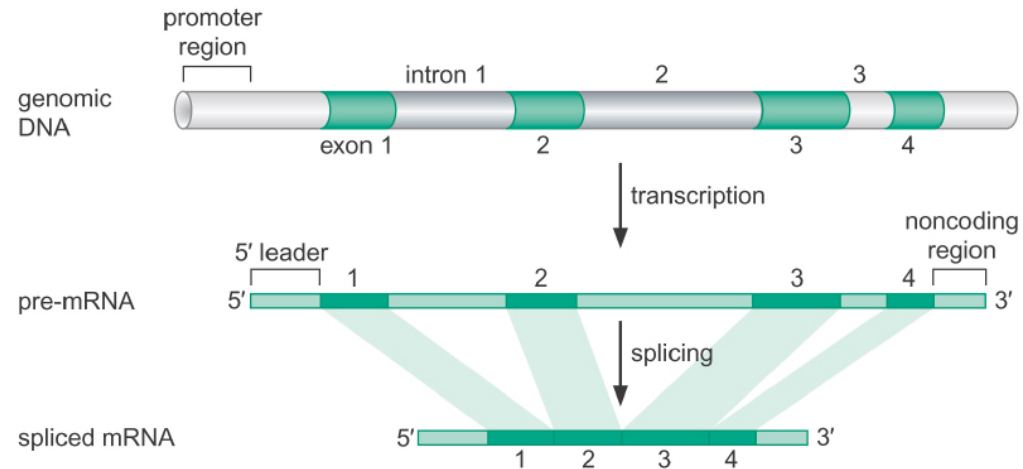
cell cycle



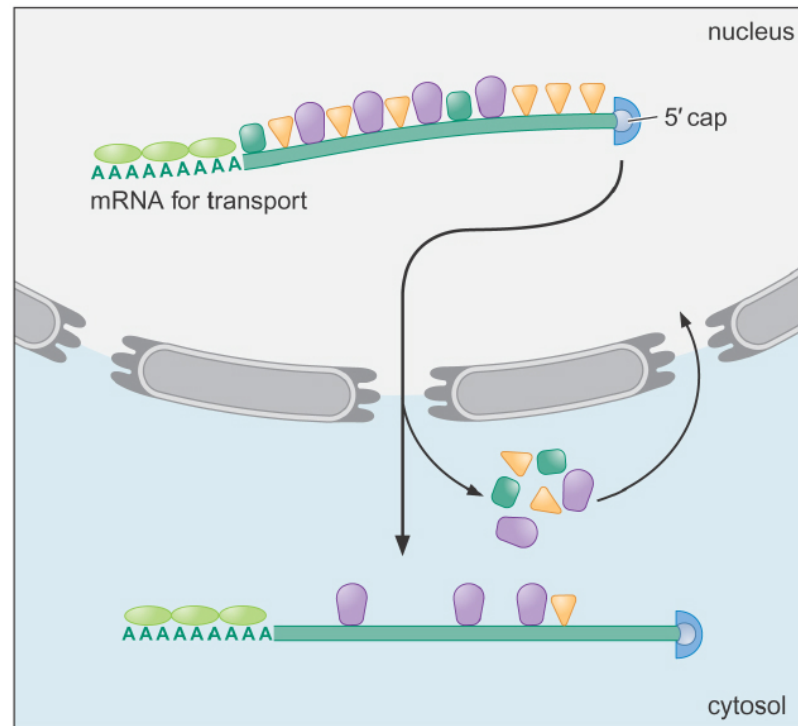
chromatin structure (regulated)



❖ RNA splicing and transport



only RNA with appropriate proteins bound are selected for transport out of the nucleus



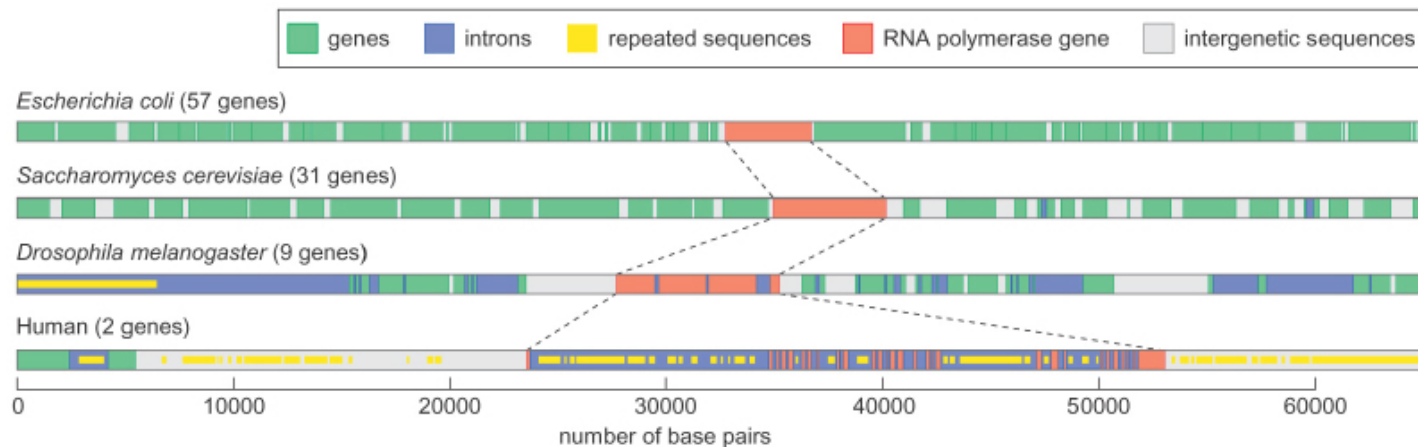
❖ genome and organization

- genome size

Organism	Genome length	No. genes
M. genitalium	0.5 Mb	500
E. coli	4.5 Mb	4,000
Yeast	12 Mb	6,000
Human	3,000 Mb	35,000
Rice	500 Mb	50,000
Lilly	90,000 Mb	?

- organization (human)

- multiple replication origins
- large intergenic separation: $3\text{Gb}/30,000\text{ genes} = 100\text{kb}$ (mostly transposable elements)



❖ Gene regulation in eukaryotes

- control of transcriptional initiation
 - direct activation by recruitment of RNAP
 - activation/repression by modifying chromatin structure
- control of entry of regulators into nucleus
- control of RNA splicing (e.g., alternative splicing)
- localization of mRNA
- control of mRNA life-time
- control of mRNA translation
- ubiquitination system to tag protein for degradation
- ...