REGULATION OF GENE EXPRESSION

THE REGULATION OF GENE EXPRESSION IN PROKARYOTES

Ho Huynh Thuy Duong

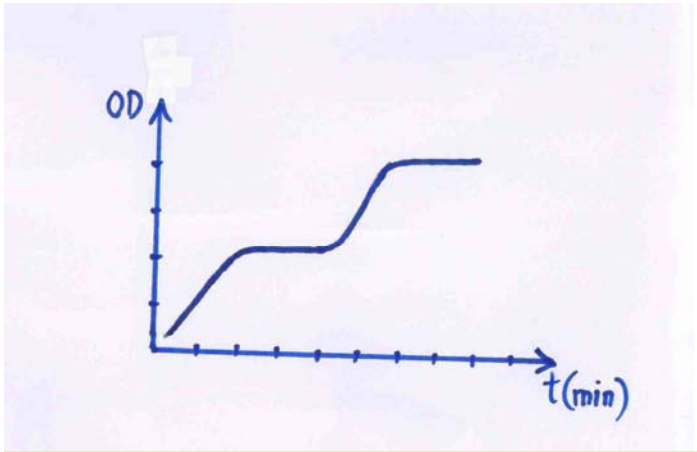
University of Science

REGULATION OF GENE EXPRESSION IN PROKARYOTES AND EUKARYOTES

Genes are expressed through transcription and translation, but what decide which gene, when, where and how it is expressed ?

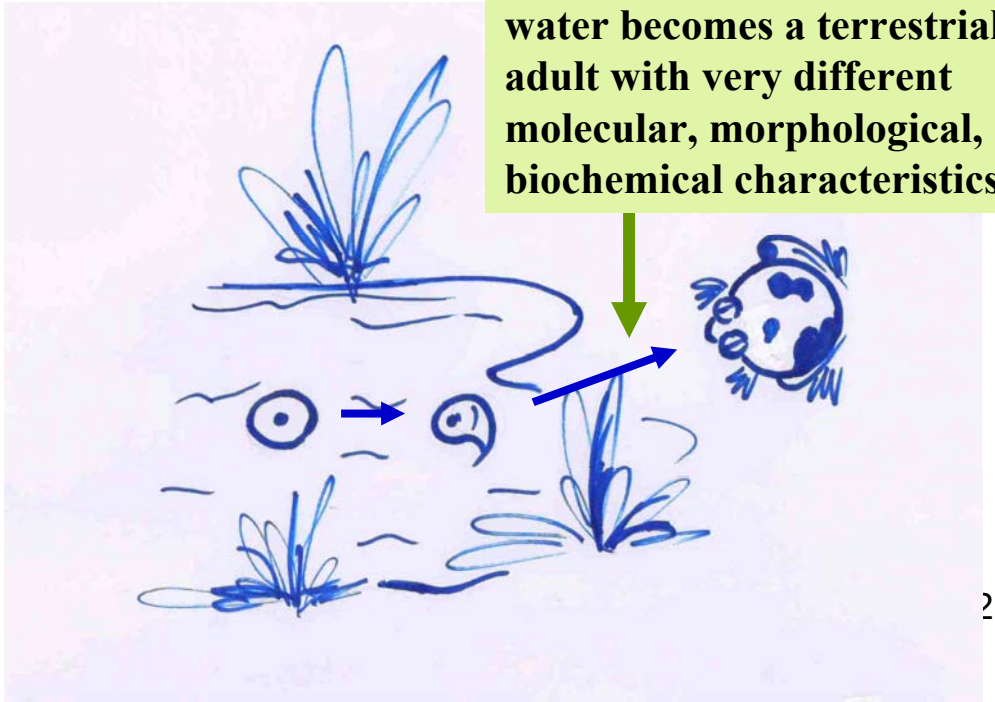
→ The expression of a gene (or a part of the genome) can be regulated in many ways depending on cell organization and needs of the organism

Examples concerning the regulation of gene expression in a bacterium and an animal



***E. coli* is grown in medium containing glucose and lactose. Cell density is measured according to culture time as OD (Optical density) value. Results are shown in the picture above**

Metamorphosis : The transition period where a larvae living in water becomes a terrestrial adult with very different molecular, morphological, and biochemical characteristics



THE REGULATION OF GENE EXPRESSION IN PROKARYOTES

- A prokaryote, as unicellular organism, is totally controlled by environmental changes
  - ➔ It has to respond as rapidly as possible to these changes to survive while saving energy
  - ➔ It uses mechanisms allowing quick adaptation to new environmental conditions
- The regulation of gene expression in prokaryotes intervenes at some levels during gene expression :

Since transcription, translation and RNA degradation in prokaryotes are coupled, regulation mainly acts at transcription level

In a few cases, a translational control can be made through : (1) Different degradation rates of mRNAS, (2) different efficiencies of translation initiation in different genes, (3) different efficiencies of translation rate due to different conformation of the mRNA (existence of secondary structures which slow down ribosome movement, …)

IN THIS TOPIC, WE WILL FOCUS ON TRANSCRIPTIONAL CONTROL OF GENE EXPRESSION IN PROKARYOTES

TRANSCRIPTIONAL CONTROL OF GENE EXPRESSION IN PROKARYOTES

ꕤ The two well studied main mechanisms of transcriptional control of gene expression are:

1. The operons : genes involved in a metabolic pathway are regrouped into a gene cluster controlled by common regulatory sequences and proteins. The expression of these genes are then rapid and synchronized. The operon model was developed by François Jacob and Jacques Monod (1961)

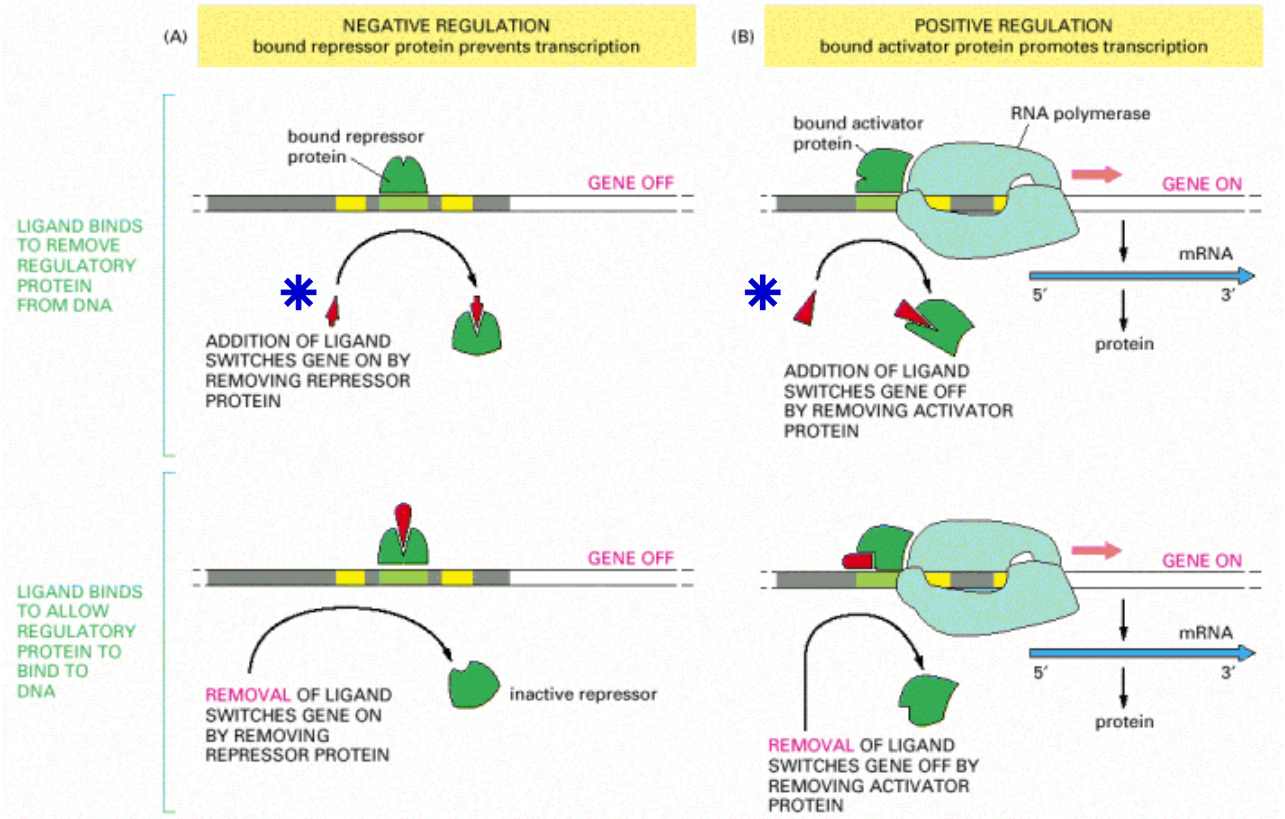
2. The cascades of gene expression : Under some environmental conditions, expression of a first set of genes can be "switch on", and one or more of the products of this first gene set will "switch on" a second gene set. This event could be repeated many times to mobilize wider gene sets to achieve a special metabolic pathway.

ꕤ In all organisms, structural genes can be classified into two groups :

1. Constitutive genes, also called "housekeeping" genes : encoding RNA and proteins having basal vital functions such as rRNA, ribosomal proteins, proteins of cellular respiratory system, … These genes are mostly expressed continually and with a stable amount.

2. Inducible genes : encoding proteins necessary for the survival of the organism in changing environment. These genes must be rapidly "switch on" or "off" depending on the temporary needs of the organism for their products.

POSITIVE AND NEGATIVE CONTROL OF GENE EXPRESSION



"Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC."

Gene expression can be positively or negatively controlled. In positive control, binding of activator protein triggers the transcription whereas in negative control, binding of repressor protein inhibits the transcription.

✱ ***Ligands which bind to the activators to "switch on" gene expression in positive control are called inducers ; those binding to the repressors and "switching off" gene expression are called co-repressors. Inducers and co-repressors are known as effectors***

THE OPERON

ꕤ The purposes of the regulation of gene expression in prokaryotes are remarkably well served by the use of operons : (1) all genes of an operon are coordinateley expressed → the metabolic pathway controlled by this operon can be regulated very fast, (2) there is energy saving as the same set of regulatory sequences and proteins is used for all structural genes of the operon.

ꕤ One of the most important challenges for prokaryotes is to adapt their metabolic processes to available environmental nutritive sources.

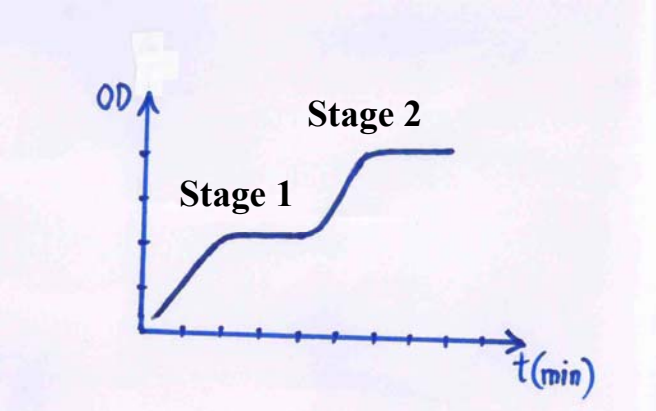
Depending on the metabolic pathways,

- In catabolic pathway (degradation of macromolecules into structural units), when a substrate to be degraded is present, the operon is "switched on". These operons are characterized as inducible

- In anabolic pathway (synthesis of macromolecules from small ones), when a product needed by the cell is present, the corresponding operon is "switched off". These operons are considered as repressible operons.

ꕤ An operon is composed of : regulatory sequences (promoter, operator, other sequences, …), structural genes, regulatory gene (promoter + coding sequence of a regulatory protein)

THE LACTOSE OPERON (*LAC* OPERON)



E. coli **is grown in medium containing glucose and lactose. Cell density is measured according to culture time as OD (Optical density) value. Results are shown in the picture above**

Q : What is the meaning of this two-stages growing pattern ?

A : In stage 1, bacteria grow using glucose as carbon source. When glucose is totally consumed, bacteria will stop growing (first plateau). After this lag phase, bacteria grow again (stage 2) using lactose until this second sugar is also finished (second plateau)

Q : Why is glucose preferred to lactose ?

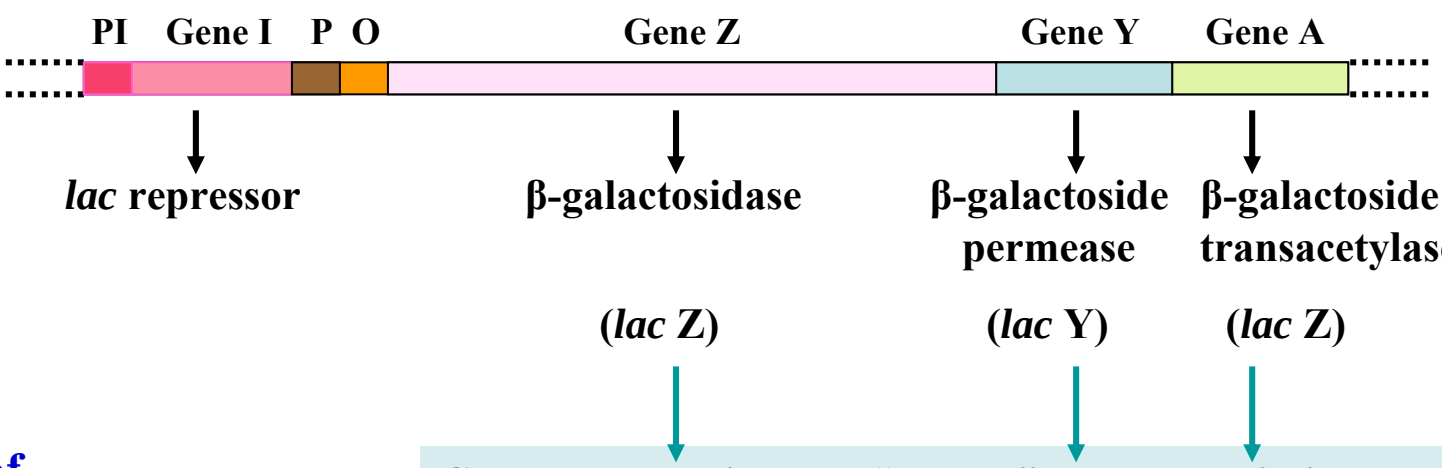
A : Glucose is preferentially used to other sugars because it is the most efficient energy source and maybe because bacterial metabolism was well adapted to this most ancient carbon source since the beginning of the Evolution on Earth.

Q : What does mean the first plateau ?

A : That is the period where bacteria switch from glucose metabolism to lactose metabolism by inducing expression of the lac operon

THE *LAC* OPERON (continued)

Structure of the *lac* operon



Functions of the enzymes controlled by the *lac* operon

"Adapted from Turner. et al. 1997. Instant Notes in Molecular Biology, p.180, fig 1. BIOS Scientific Publishers Ltd"

The *lac* operon is a negative inducible operon, composed of :

1. Regulatory sequences : (1) the operator (O) which binds the repressor protein, (2) the promoter (P) containing two binding sites, one for the RNA polymerase, the other for CAP-cAMP complex

2. Structural genes involved in lactose metabolism : gene Z, Y and A

3. Regulator gene : the promoter (P_I) and the coding sequence (gene I)

THE *LAC* OPERON (continued)

ꕤ To save energy, the *lac* operon is "switched off" in a lactose-free medium.

Under this condition, a small amount of *lac* repressor is produced and binds to the operator. Since the operator and promoter regions overlap, binding of *lac* repressor to the operator prevents the binding of RNA polymerase → no transcription.

Actually, the *lac* operon is never completely inhibited and a very weak activity is always observed.

ꕤ If lactose is added to the medium, lactose forms a complex with *lac* repressor, causes a conformation change of the repressor and its release from the operator. RNA polymerase can then be recruited to the promoter and initiates the transcription of the three structural genes.

Q : Why, in the previous example, *lac* operon is not induced in Stage 1 when lactose is already present ?

A : Because of a phenomenon called catabolite repression or glucose effect to ensure the preferential use of glucose to any other sugars.

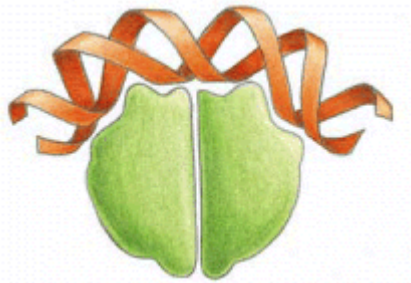
THE CATABOLITE REPRESSION

The concentration of glucose in the cell affect the activity of an enzyme called adenylate cyclase (adenylcyclase). At high glucose concentration, adenylate cyclase activity is inhibited. The decrease of glucose concentration activates this enzyme which transforms ATP (Adenosine 5'-triphosphate) into cAMP (3',5'-cyclic Adenosine monophosphate).

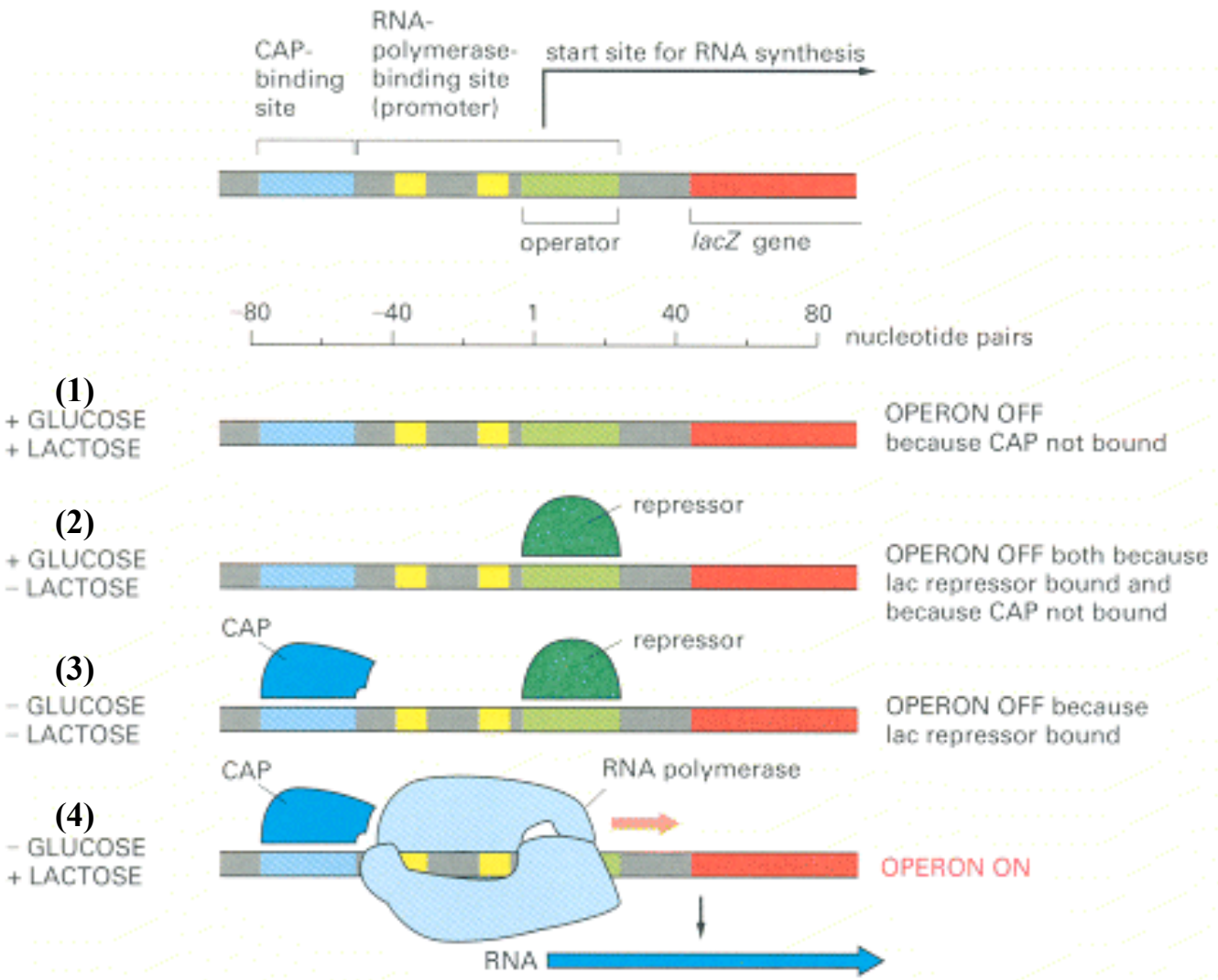
Increasing amount of cAMP then associates with CAP (Catabolite Activator Protein). CAP-cAMP complex binds to its binding site and induces expression of the *lac* operon

WHY IS CAP-CAMP COMPLEX NECESSARY TO TRANSCRIPTION INITIATION IN *LAC* OPERON ?

***Lac* promoter has a -35 sequence differring from the consensus sequence for strong promoters. This does not favorize the binding of RNA polymerase. Binding of the CAP-cAMP complex to lac promoter induces DNA bending, thus helps to recruit and stabilize the binding of RNA polymerase to the promoter.**



THE *LAC* OPERON – A SUMMARY

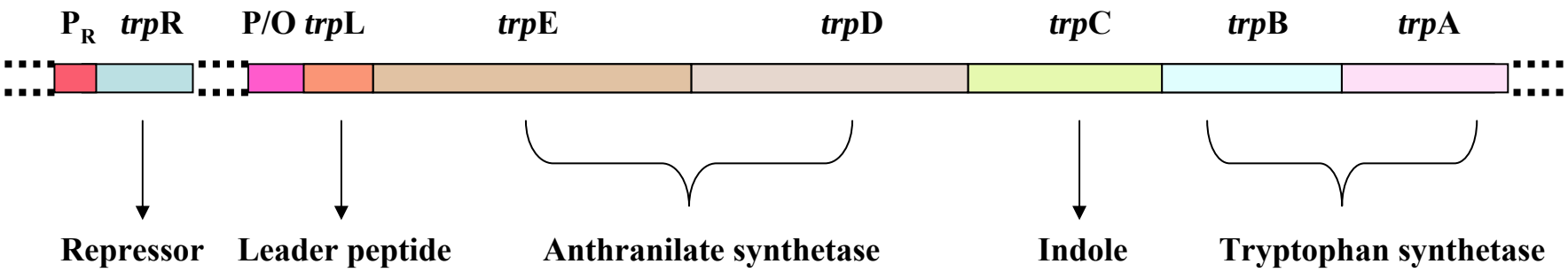


(1) : Actually, a very little amount of *lac* mRNA is produced

(2), (3) : no *lac* mRNA produced

(4) : Abundant production of *lac* mRNA

THE TRYPTOPHAN OPERON – *OPERON ORGANIZATION*



Functions of the enzymes belonging to the tryptophan operon

"Adapted from Turner. et al. 1997. Instant Notes in Molecular Biology, p.184, fig 1. BIOS Scientific Publishers Ltd"

The tryptophan (*trp*) operon is a negative repressible operon, composed of :

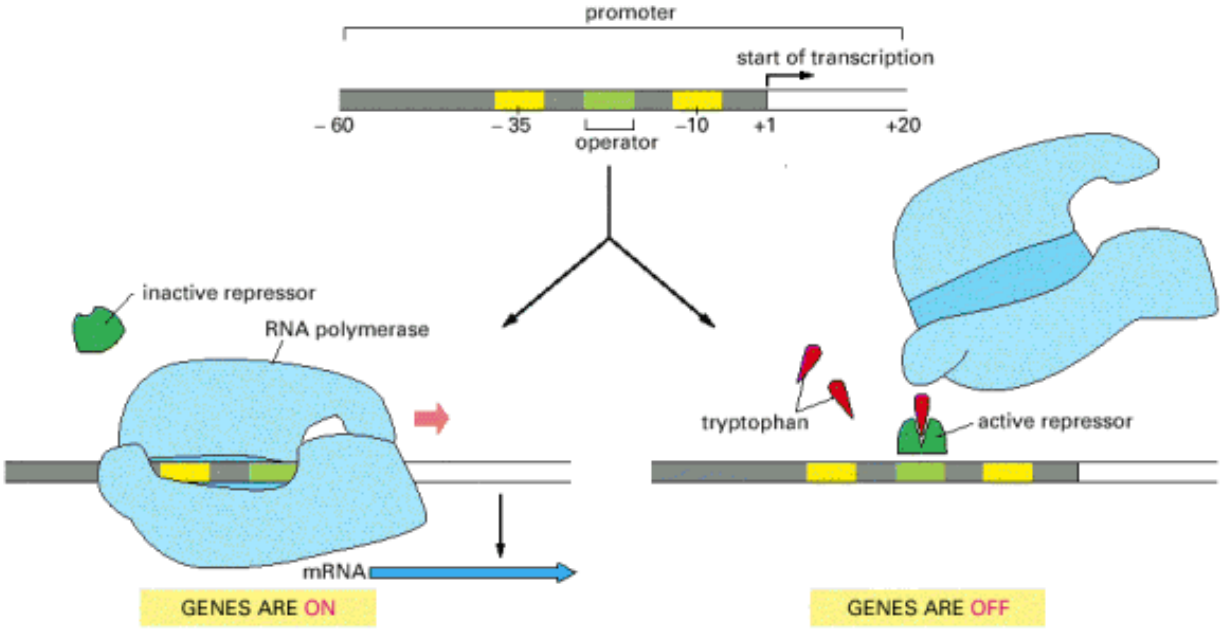
1. Regulatory sequences : the operator lies inside the promoter region

2. Structural genes include *trp*E, D, C, B, A involved in the synthesis of tryptophan

3. Regulatory gene : the coding sequence (*trp*R) for the repressor and its promoter (P_R) is not adjacent to the operon.

A special gene, *trp*L, encodes the Leader peptide which underlies a regulation mechnism called "attenuation"

THE *TRP* OPERON – *REPRESSION*

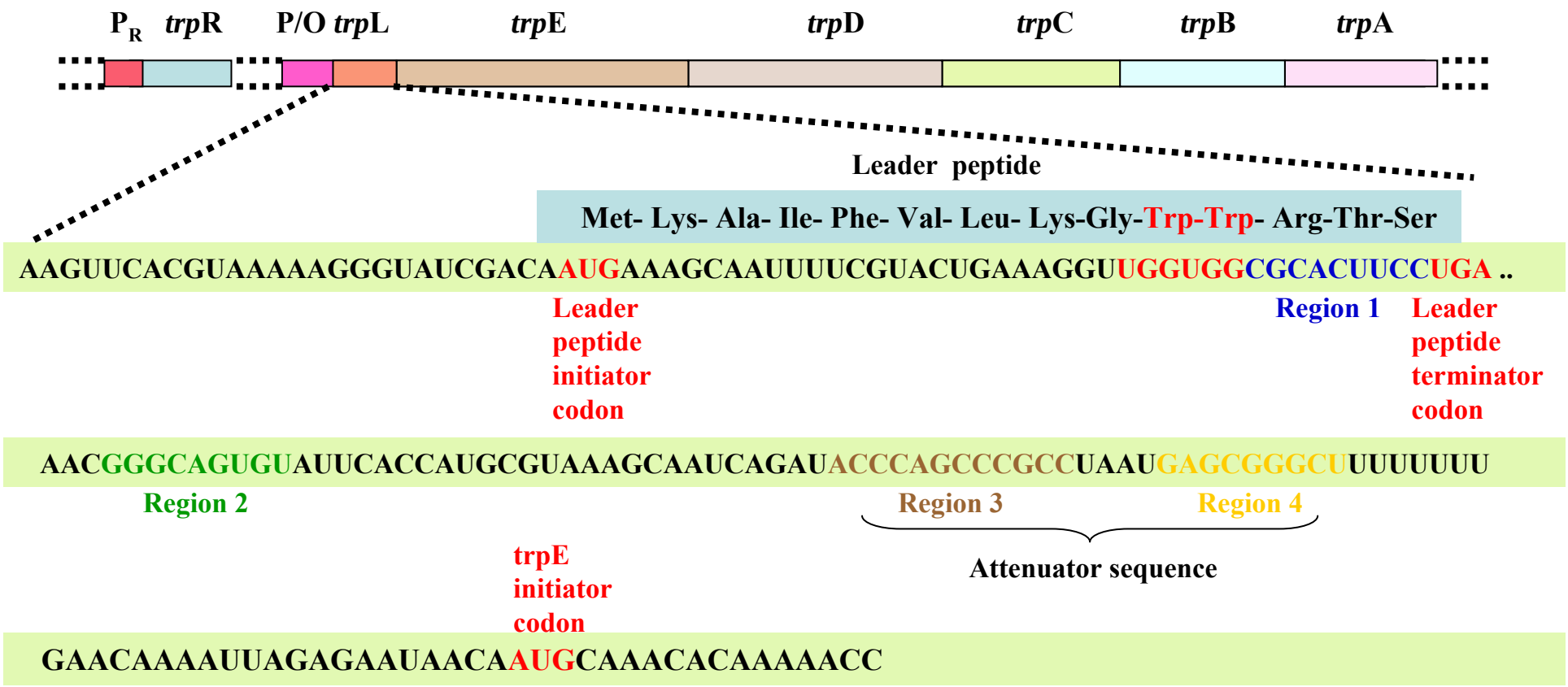


The transcriptional regulation of the *trp* operon is similar to that of the *lac* operon. The difference lies in the nature of effectors. In lac operon, effector is an inducer which inactivates the repressor whereas in trp operon effector is a co-repressor which activates the repressor protein.

In tryptophan starvation condition, the inactive repressor can not bind to the operator, RNA polymerase is recruited to the promoter and initiates the transcription of trp operon. In the presence of tryptophan, tryptophan binds to and activates the repressor which in its turn binds to the operator and blocks the promoter, inhibiting the transcription iniiation.

THIS MECHANISM REGULATES THE *TRP* OPERON BY 70 TIMES BUT THE *TRP* OPERON IS ACTUALLY REGULATED BY 700 TIMES ! ?

THE *TRP* OPERON - *ATTENUATION*



"Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.505, fig 16.20. Benjamin Cummings., CSHL Press"

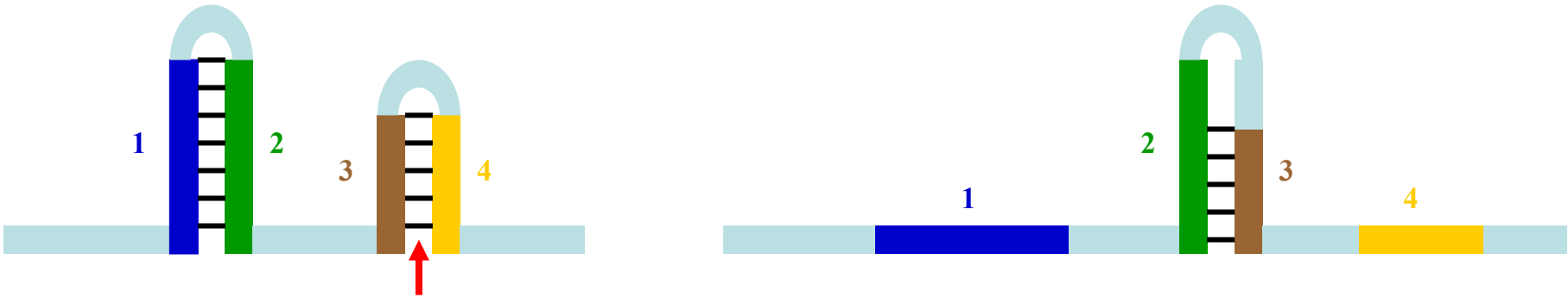
Complementary to repression, the trp operon is regulated by another mechanism called attenuation which regulates the expression by 10 times more.

Attenuation concerns a region upstream of the structural genes, called *trp*L

THE *TRP* OPERON – *ATTENUATION (continued)*

- **The *trp*L region is composed of : (1) a sequence encoding the leader peptide, (2) four regions 1, 2, 3, 4 which can basepair two by two ; the pairing of region 3 and 4 forms a hairpin structure called attenuator which is a transcription-termination structure.**

- **Attenuation is a regulation mechanism based on the simultaneous occurrence of transcription and translation in prokaryotes. The *trp*L mRNA in progress of being transcribed begins already to be translated.**

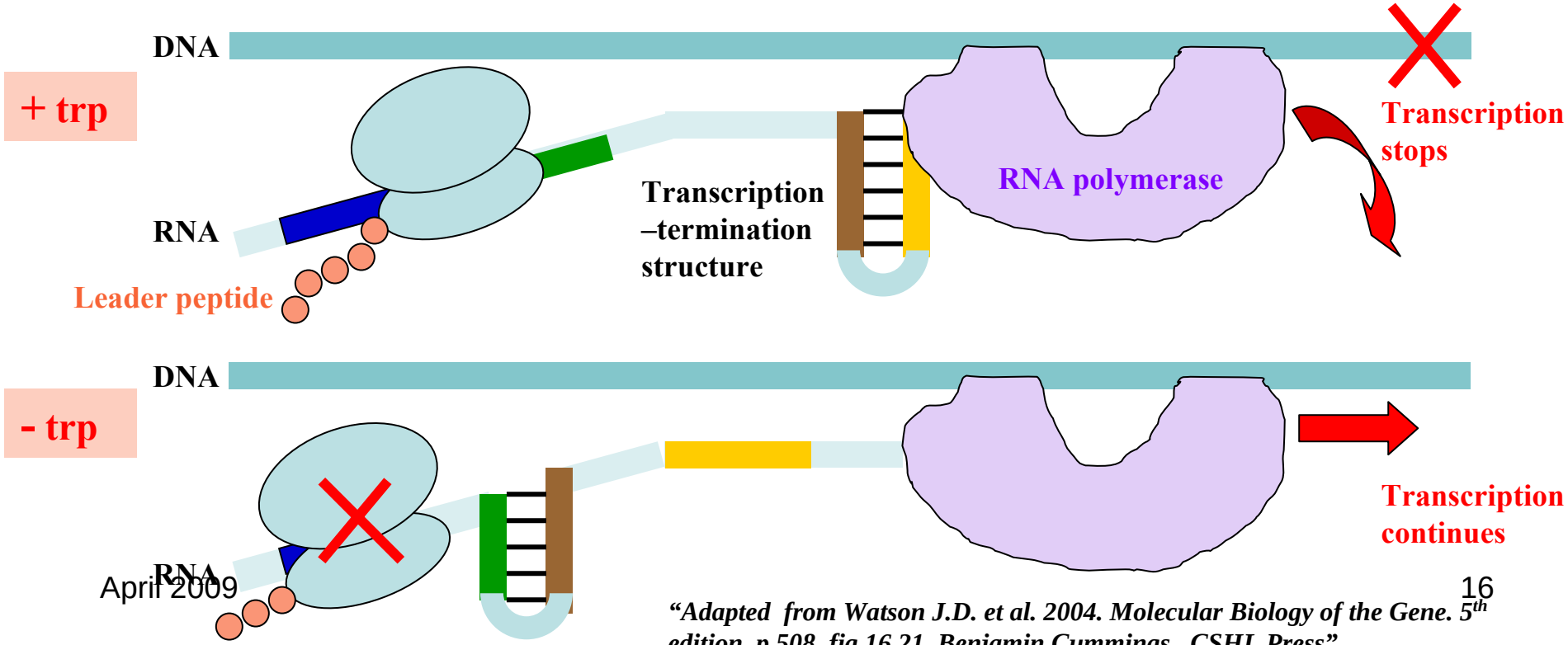


"Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.508, fig 16.21. Benjamin Cummings., CSHL Press"

THE *TRP* OPERON – *ATTENUATION (continued)*

- **When tryptophan is present (+ trp), the transcribed region 2 is blocked by a moving ribosome. The region 3 when transcribed will then basepair with region 4 forming the transcription-termination hairpin. The RNA polymerase stops the transcription before reaching the first structural gene.**

In the absence of tryptophan (- trp), ribosome translating *trp*L mRNA is stopped at the two trp codons, waiting for ^{trp}tRNAs. The transcribed region 2 is not blocked can basepair with transcribed region 3. This 2:3 pairing prevents the formation of 3:4 pairing. RNA polymerase can then continue the transcription of the downstream structural genes.



"Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.508, fig 16.21. Benjamin Cummings., CSHL Press"

SOME OPERONS REGULATED BY ATTENUATION

OPERON	LEADER SEQUENCE
Tryptophan	Met-Lys-Ala-Ile-Phe-Val-Leu-Lys-Gly-Trp-Trp-Arg-Thr-Ser
Phenylalanine	Met-Lys-His-Ile-Pro-Phe-Phe-Phe-Ala-Phe-Phe-PheThr-Phe-Pro
Histidine	Met-Thr-Arg-Val-Gln-Phe-Lys-His-His-His-His-His-His-His-Pro-Asp
Threonine	Met-Lys-Arg-Ile-Ser-Thr-Thr-Ile-Thr-Thr-Thr-Ile-Thr-Ile-Thr-Thr-Gln-Asn-Gly-Ala-Gly
Leucine	Met-Ser-His-Ile-Val-Arg-Phe-Thr-Gly-Leu-Leu-Leu-Leu-Asn-Ala-Phe-Ile-Val-Arg-Gly-Arg-Pro-Val-Gly-Gly-Ile-Gln-His

"Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.507, table 16.1. Benjamin Cummings., CSHL Press"

CASCADE REGULATION OF GENE EXPRESSION

A mechanism also ensuring a rapid and energy-saving regulation of gene expression in prokaryotes is the use of different σ factors. σ factor is a component of the RNA polymerase holoenzyme which specifically recognizes the promoter.

Each σ factor directs the RNA polymerase to a defined set of promoters. These promoters control the expression of groups of genes involving in a special metabolic activity of the cell.

Two examples illustrating the regulation of gene expression through different σ factors

- **Use of alternative σ factors by *E. coli* for self-adaptation to new environment**
- **Use of alternative σ factors by SPO1 bacteriophage during the infection process**

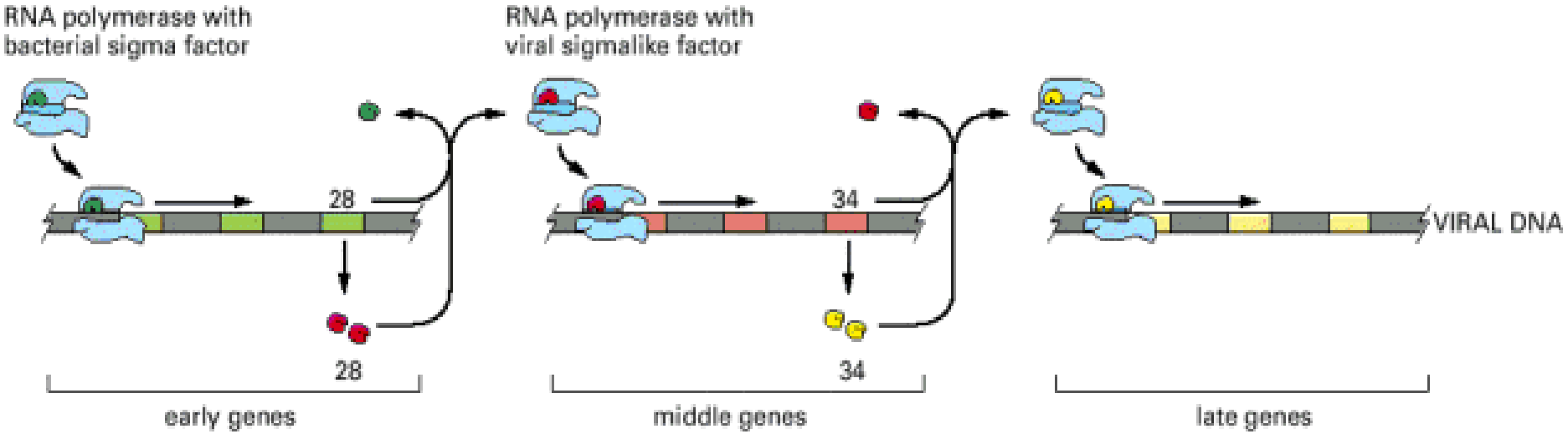
USE OF ALTERNATIVE σ FACTORS IN *E. COLI*

- **In normal conditions, RNA polymerase holoenzyme containing σ^{70}, the most common σ factor, transcribes a set of general promoters recognized by σ^{70}. When *E. coli* cells encounter a heat shock due to rising environmental temperature, a new σ factor, σ^{32}, is synthesized in large amount and replaces σ^{70} to direct RNA polymerase to a set of heat-shock gene promoters. The products of these genes protect the cell against harmful effects of heat shock. The increasing concentration of σ32 is due to : (1) enhanced translation of σ^{32} mRNA, (2) Stabilization of σ^{32} protein.**

- **Other σ factors are alternatively used in different circumstances to express different sets of genes :**

FACTOR	GENE SET	CONSENSUS SEQUENCES	
		-35	-10
σ^{70}, the major σ	Exponential growth	TTGACAT	TATAAT
Σ^{38}, the second most important σ	Stationary growth, response to general stresses	-24	-12
σ^{32}	Heat shock	TCTCNCCCTTGAA	CCCCATNTA
σ^{28}	Mobility-flagellar formation	CTAAA	CCGATAT
σ^{54}	Nitrogen metabolism	CTGGNA	TTGCA

ALTERNATIVE σ FACTORS INTERVENING IN THE INFECTION PROCESS BY BACTERIOPHAGE



SPO1 phage infecting *Bacillus subtilis* has three sets of genes – the early, middle and late genes which express at different time during the phage infection process.

The phage early genes are expressed by bacterial RNA polymerase bearing the bacterial σ factor. One of the early gene products is the phage factor σ^{28}. The σ^{28} factor will replace the bacterial σ factor to direct RNA polymerase to promoters of the phage middle genes. Among the middle gene products, there is σ^{34} factor. In its turn, σ^{34} factor participates in the expression of the late genes of the phage

TRANSLATIONAL REGULATION OF GENE EXPRESSION

ᔓ Prokaryotic mRNAs are usually polycistronic ; this means that one mRNA contains coding sequences for several genes, e.g the lac mRNA

In this example, even if all three structural genes of the lac operon are switch "on" and "off" together, they are not translated with the same efficiency. An *E. coli* cell grown in medium with lactose will contains about 3,000 molecules of β-galactosidase, 1,500 molecules of β-galactoside permease, 600 molecules of β-galactoside transacetylase.

ᔓ In prokaryotes, control of gene expression at the translational level is based on some mechanisms :

1. Different efficiencies of translational initiation due to the sequences surrounding theATG start codon

2. Different efficiencies of translational elongation due to secondary structures formed inside the mRNA molecule

3. Different degradation rates of the mRNAs

SUMMARY

The regulation of gene expression in prokaryotes provides the best survival opportunities to the organism by rapid and synchronized switch of gene transcription.

Mechanisms ensuring rapid and synchronized gene expression include :

- **Regulation by operons : Operons are composed of many protein-encoding genes (structural genes) involved in a metabolic pathway and regulatory sequences common for those genes. Depending on the operon types, the transcription of structural genes can be switched on with the presence (catabolic operon such as lac operon) or switched off (anabolic operon such as trp operon).**

There are additional regulation mechanisms including attenuation, in the case of trp operon and catabolite repression as found in lac operon. These additional mechanisms enhance the effect of operon regulation.

- **Cascade regulation : different σ factors are used to control different sets of genes required for special conditions. An example concerns the replacement of σ^{70} which is the initiation factor used to transcribe genes in normal conditions by σ^{32} which drives RNA polymerase to heat-shock protein-encoding genes in response to heat-shock stress.**

Due to the simultaneous occuring of transcription and translation in prokaryotes, the main regulation level of gene expression lies at the transcription initiation. Nevertheless, regulation sometimes can intervene in a few cases at the translational level.