

TÉCNICAS INTELIGENTES EN BIOINFORMÁTICA

Modelización de la regulación génica en procariontas El Operón Lactosa

Mario de J. Pérez Jiménez
Grupo de investigación en Computación Natural
Dpto. Ciencias de la Computación e Inteligencia Artificial
Universidad de Sevilla

Máster Universitario en Lógica, Computación e Inteligencia Artificial

Curso 2014-15



Introducción

Las células son estructuras muy bien organizadas.

Cantidad de proteínas que produce un gen activo por unidad de tiempo:

- ★ Varía según las necesidades de la célula.
- ★ Las bacterias son muy dependientes de las condiciones de su entorno: deben responder rápidamente a esos cambios para poder sobrevivir.
- ★ En *E. coli* hay proteínas muy escasas: se encuentran sólo decenas.
- ★ En *E. coli* hay proteínas muy abundantes: se encuentran miles de copias.
- ★ En una célula de un *mamífero* pueden existir 10^{10} proteínas pertenecientes a unos 10.000/20.000 tipos diferentes.
- ★ La *diferenciación celular* en organismos multicelulares se fundamenta en la regulación de los procesos de expresión de los genes.
- ★ **Genes constitutivos:** se expresan continuamente.
- ★ **Genes regulables:** se expresan de acuerdo con las necesidades de la célula.

Regulación de genes

Síntesis de proteínas: alto coste energético.

- ★ Hay que **regular** este proceso.

La regulación se puede dar a nivel de la transcripción o de la traducción

- ★ Lo más directo es regular la tasa de transcripción del gen en moléculas de RNAm.
- ★ La regulación suele llevarla a cabo una proteína (sintetizada por un **gen** denominado **regulador**).

Células procariontas: un RNAm puede codificar varios genes (**policistrónicos**).

Células eucariotas: un RNAm sólo codifica un gen (**monocistrónicos**).

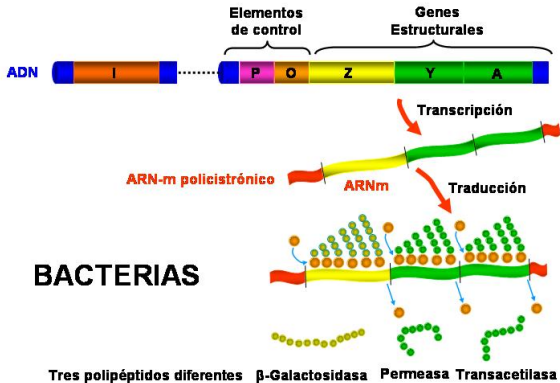
Concepto de operón

Operón: secuencia de nucleótidos en el DNA que equivale a una **unidad de transcripción**.

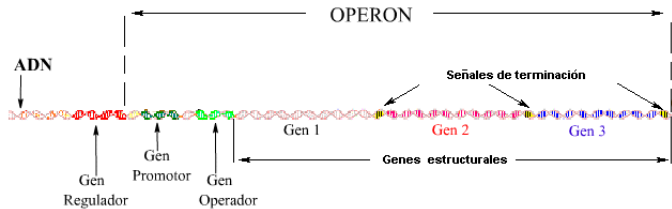
- ★ Conjunto de genes cuya expresión está controlada por un único RNAm y está regulada por los mismos elementos de control (promotor y operador) y los mismos genes reguladores.

Elementos que constituyen un operón:

- **Genes estructurales:** genes cuya expresión está regulada. Los operones bacterianos suelen contener varios (**policistronicos**). Los operones eucarióticos suelen contener uno sólo (**monocistronicos**).
- **Operador (O):** elemento de control (región del DNA con una secuencia que es reconocida por la proteína reguladora). Se encuentra inmediatamente antes de los genes estructurales.
- **Promotor (P):** elemento de control (región del DNA que es reconocida por la RNA polimerasa para comenzar la transcripción). Se encuentra inmediatamente antes del operador.
- **Gen regulador:** secuencia de DNA que codifica la proteína reguladora. Suele estar situado “cerca” de los genes estructurales del operón pero no inmediatamente al lado.
- **Proteína reguladora:** proteína codificada por el gen regulador que suele unirse al operador del operon.
- **Inductor:** molécula cuya presencia induce la expresión de los genes estructurales.



ESQUEMA SIMPLIFICADO DEL SISTEMA OPERON



Sistemas de regulación de genes

El proceso de activación de un operón tiene asociado un **gen regulador**.

★ **Regulación positiva:** Para la expresión, es necesaria una proteína reguladora que active el promotor.

- **Regulación positiva por inducción:** Por defecto, la proteína reguladora está inactiva. Se activa cuando se le une un **inductor** (operón lactosa).
- **Regulación positiva por represión:** Por defecto, la proteína reguladora está activa. Se inactiva cuando se le une un **represor**.

★ **Regulación negativa:** El sistema se expresa, a menos que sea desconectado por una proteína represora.

- **Regulación negativa con efectos inductores:** Por defecto, la proteína represora está activa. Se inactiva cuando se le une un **inductor** (operón lactosa).
- **Regulación negativa con efectos represores:** Por defecto, la proteína represora está inactiva. Se activa cuando se le une un **correpresor** (operón triptófano).

Promotores de un operón

El inicio de la transcripción puede estar regulado por una proteína represora.

- ★ Ejemplo: la síntesis del triptófano en la bacteria *E. coli*.
 - La proteína represora necesita al triptófano para unirse al operador e inhibir el proceso.

Promotores de un operón:

- ★ **Fuerte**: en ausencia del represor, la tasa de unión de la RNA-polimerasa es muy alta.
- ★ **Débil**: en ausencia del represor, no son capaces de iniciar la transcripción por sí solos de forma efectiva.
 - Necesitan una proteína activadora.

El Operón Triptófano

Tiene cinco **genes estructurales**: trpE-trpD-trpC-trpB-trpA (¡en este orden!).

- ★ Las proteínas codificadas por estos genes actúan en la ruta metabólica de **síntesis del triptófano** en el mismo orden en el que se encuentran los genes en el cromosoma.

El **promotor** y el **operador** están al lado de los genes estructurales

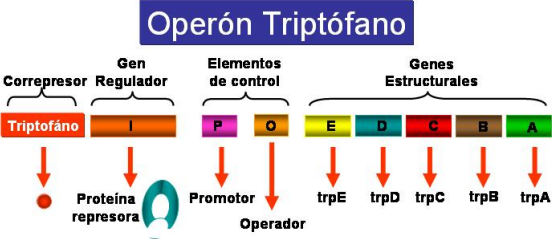
- ★ Primero el promotor, después el operador y, a continuación, los genes estructurales.

El **gen regulador** (trpR) codifica para la proteína reguladora.

- ★ Es un represor y se encuentra en otra región del cromosoma bacteriano, no muy lejos del operón.

Un **correpresor**: el propio triptófano (activa a la proteína represora).

Estructura del Operón Tryptófano

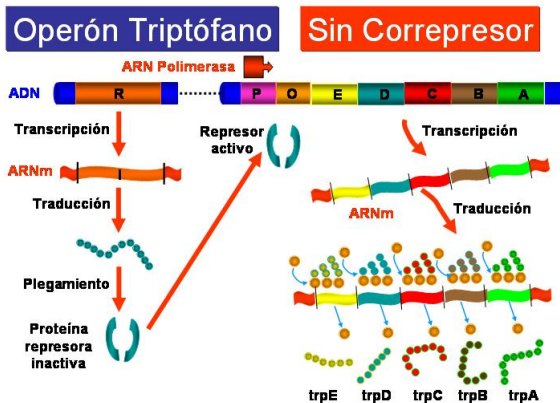


SITUACIÓN EN EL CROMOSOMA BACTERIANO



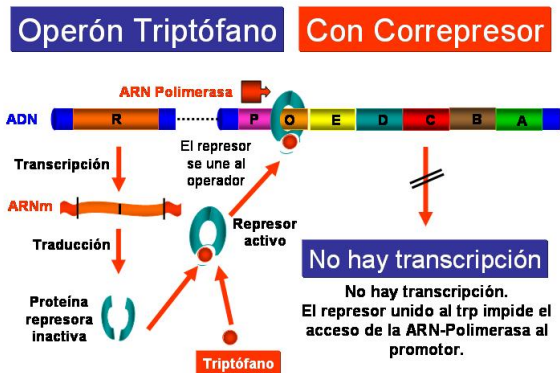
En **ausencia del triptófano** (o cuando hay muy poco):

- La proteína reguladora no es capaz de unirse al operador.
- Entonces la RNA-polimerasa se puede unir a la región promotora transcribiendo los genes del operón.



En presencia del triptófano:

- Éste se une a la proteína represora cambiando su conformación, propiciando la unión a la región operadora e impidiendo la acción de la RNA-polimerasa.
- No hay transcripción de los genes del operón triptófono.



El Operón Lactosa

Los estudios genéticos sobre la utilización de la lactosa como fuente alimenticia, permitieron describir el **primer modelo de regulación génica**.

Los **genes estructurales** del operón lactosa son los siguientes:

- El gen z^+ : codifica la enzima **β -galactosidasa (Lac Z)** que cataliza la hidrólisis de la lactosa en glucosa más galactosa.
- El gen y^+ : codifica la enzima **Lac Y** que transporta lactosa al interior de la célula a través de la membrana plasmática.
- El gen a^+ : codifica la enzima **Lac A** que no parece estar relacionada con el metabolismo de la lactosa.

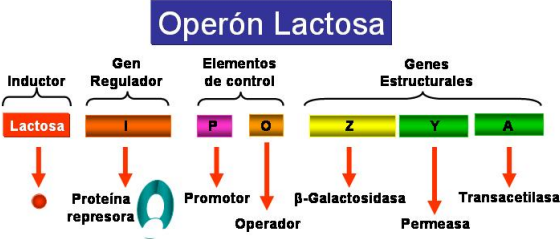
Promotor y operador: adyacentes a los genes estructurales

El **gen regulador** es el gen i que codifica la proteína reguladora (en este caso, represora): **LacI** compuesta de 300 aminoácidos.

El **inductor** del sistema es la lactosa (en realidad es la alolactosa, un derivado)

- La alolactosa inhibe la acción de la proteína represora uniéndose a una región del LacI.

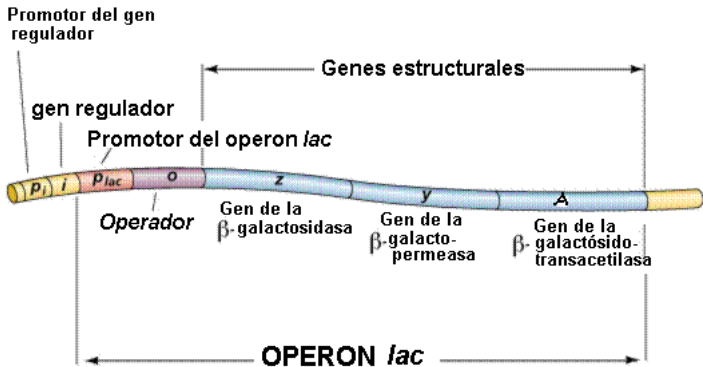
Estructura del Operón Lactosa (I)



SITUACIÓN EN EL CROMOSOMA BACTERIANO



Estructura del Operón Lactosa (II)



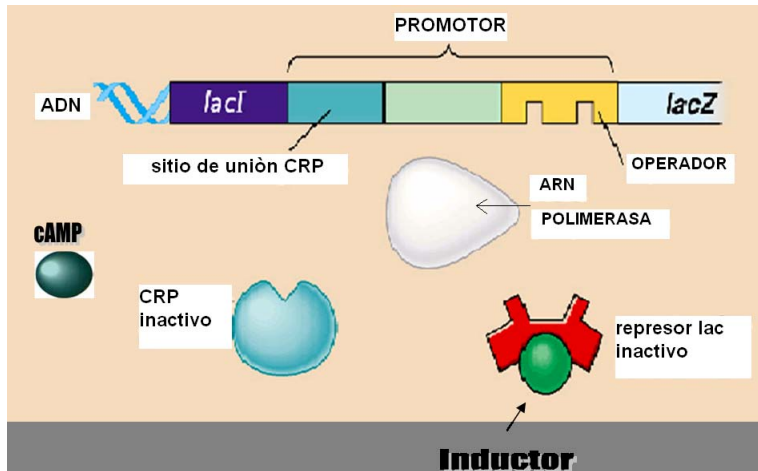
El Operón Lactosa. Represión

- ▶ La proteína represora del operón lactosa se llama Lacl.
- ▶ Parte de la molécula Lacl se une al operador impidiendo la unión de la RNA-polimerasa al promotor.
- ▶ El **inductor** del sistema es la alolactosa, un derivado de la lactosa: *represor del represor*.
 - ★ La alolactosa inhibe la acción de la proteína represora: se une a una región del Lacl cambiando su conformación e impidiendo que Lacl se una al operador.

El Operón Lactosa. Activación

- ▶ La inhibición del represor LacI es condición necesaria pero no suficiente para la transcripción de los genes estructurales.
- ▶ El complejo proteínico CRP-cAMP y su dímero CRP-cAMP₂: activador que propicia la unión de la RNA-polimerasa al promotor.
- ▶ La presencia de glucosa en el medio inhibe la síntesis de dicho complejo proteínico: **represión catabólica**.
 - La molécula EIIA~P:
 - ★ Ayuda al transporte de la glucosa por el interior de la bacteria (en ese transporte, la glucosa es fosforilada).
 - ★ Activa la enzima AC que, a su vez, regula la producción de cAMP.
 - ★ Inhibe a la enzima LacY que transporta lactosa al interior de la bacteria.
 - En medios ricos en glucosa, los niveles de cAMP son bajos.
 - Cuando no hay glucosa en el medio, la molécula EIIA~P está libre y se sintetiza cAMP.
 - Cuando la E. coli crece en un medio sin glucosa pero rico en lactosa, los niveles de cAMP son altos: El cAMP se une al receptor CRP y el complejo proteínico CRP-cAMP₂ propicia la unión de la RNA-polimerasa al promotor, estimulando la transcripción de los genes estructurales.

Estructura del Operón Lactosa (III)



Operón Lactosa en E. coli

Descubierto por F. Jacob y J. Monod, 1961 (Premio Nobel, 1965).

Mediante la degradación de la lactosa, las bacterias obtienen energía.

La bacteria E. coli es capaz de seleccionar la fuente de energía cuando tiene varias a su alcance.

- En un medio que contiene glucosa, la E. coli prefiere a ésta como fuente de energía.
- Los operones que producen las enzimas necesarias para obtener energía de otros azúcares, están bloqueados.

El Operón lactosa está sujeto simultáneamente a una:

- **Regulación positiva por inducción:** el complejo proteínico (CRP-cAMP₂) activa la transcripción de los genes estructurales.
- **Regulación negativa con efecto inductor:** la proteína reguladora Lac I es un represor que impide la expresión de los genes estructurales. Esa proteína se inhibe en presencia del inductor (la alolactosa).

Operón Triptófano versus Operón Lactosa

Operón Triptófano:

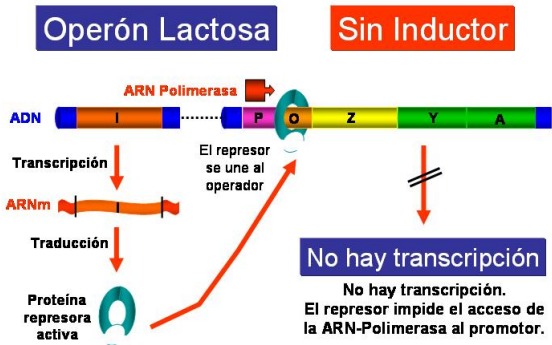
- La proteína represora sólo se unirá al operador cuando esté “activada” por un correpresor (triptófano).
- El promotor es fuerte ya que la expresión de los genes estructurales es alta en ausencia del represor “activado” .

Operón Lactosa:

- ★ La proteína represora sólo se puede unir al operador en ausencia de lactosa.
- ★ El promotor es débil ya que la transcripción está regulado por un represor (Lac I) y un complejo activador (CRP-cAMP₂).
- ★ La presencia de **glucosa** en el medio inhibe la síntesis del complejo activador.

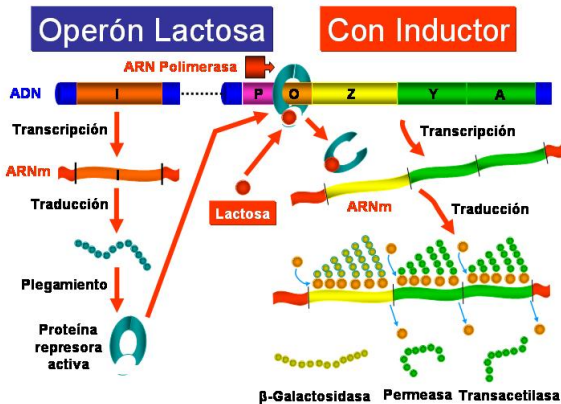
En ausencia de lactosa:

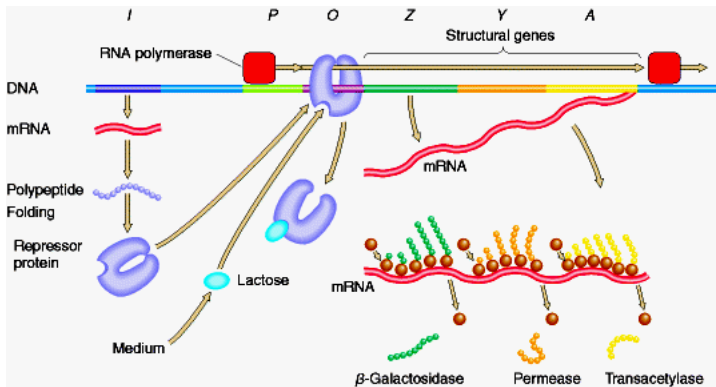
- La proteína represora se encuentra unida al operador e impide la unión de la RNAP a la región promotora.
- No se transcriben los genes estructurales (ahorro energético).



En presencia de lactosa:

- Ésta se une a la proteína reguladora que cambia su conformación y se suelta de la región operadora.
- La RNA-polimerasa se une a la región promotora y se transcriben los genes estructurales.





Transcripción del Operón Lactosa en E. Coli

▶ Caso 1: Ausencia de glucosa y lactosa.

- ★ Existirán muchos complejos CRP-cAMP₂ que asistirán a la RNA-polimerasa para que se una al promotor pero **NO se activará la transcripción** ya que el represor Lac I impedirá esa unión.

▶ Caso 2: Presencia de glucosa y ausencia de lactosa.

- ★ Existirá un bajo número de complejos CRP-cAMP₂ pero **NO se activará la transcripción** ya que el represor Lac I impedirá la unión del complejo al operador.

▶ Caso 3: Presencia de glucosa y presencia de lactosa.

- ★ Existirá un bajo número de complejos CRP-cAMP₂ pero ahora **SÍ** se activará la transcripción ya que el represor Lac I está inhibido por la presencia de la lactosa. Se producirá, pues, un **nivel bajo de transcripción** (en este caso, no se necesita degradar lactosa para obtener energía).

▶ Caso 4: Ausencia de glucosa y presencia de lactosa.

- ★ Existirán muchos complejos CRP-cAMP₂ que asistirán a la RNA-polimerasa para que se una al promotor del operón y **SÍ** se activará la transcripción ya que el represor Lac I está inhibido. Se producirá un **nivel alto de transcripción**.

Operón Lactosa en E. coli: Modelo basado en sistemas P (I)

En una bacteria E. coli existen dos regiones relevantes:

- La *superficie celular de la bacteria* que contiene un conjunto de proteínas que controlan el movimiento de las moléculas y detectan señales.
- El *lumen bacteriano* o interior acuoso de la bacteria donde una serie de proteínas están involucradas en procesos celulares específicos.

En nuestro modelo vamos a considerar tres membranas:

- La primera está etiquetada con **s** y representa la superficie celular. Los objetos que describen las entidades moleculares asociadas a la membrana *s* se localizarán en esta región, así como las reglas que especifican los procesos de selección de sustancia del entorno y de señalización (receptor-sígnal).
- La segunda está etiquetada con **c**, está contenida en la membrana *s* y describe el medio acuoso del interior de la bacteria. Contiene los objetos y cadenas que especifican las proteínas y otras entidades moleculares situadas en dicho medio. Las reglas describen las interacciones que tienen lugar dentro de la bacteria.
- La tercera membrana está etiquetada por **e**, contiene a la membrana *s* y describe el medio externo (cultivo). Dependiendo de los recursos de azúcares, la E. coli exhibirá un comportamiento diferente. En nuestro modelo usaremos esta región para codificar esos recursos a través de la multiplicidad de ciertos objetos.

Operón Lactosa en E. coli: Modelo basado en sistemas P (II)

$$\Pi_{Lac} = (\Gamma, \{e, s, c\}, [[[]_c]_s]_e, \mathcal{M}_e, \mathcal{M}_s, \mathcal{M}_c, \mathcal{R}_e, \mathcal{R}_s, \mathcal{R}_c)$$

en donde:

- ★ El alfabeto Γ contiene los objetos que representan las entidades moleculares, proteínas, lugares del DNA y RNA involucrados en el sistema del operón lactosa: $\Gamma = \Gamma_{prot} \cup \Gamma_{dna} \cup \Gamma_{ma}$.
- ★ Las etiquetas $\{e, s, c\}$ identifican los tipos de compartimentos definidos por la estructura de membranas en Π_{Lac} : entorno, superficie celular y citoplasma.
- ★ La estructura de membranas es lineal y consta de tres membranas: la primera representa al entorno (**e**), la segunda representa a la superficie celular (**s**), y la tercera al citoplasma (**c**).
- ★ Los multiconjuntos iniciales \mathcal{M}_e , \mathcal{M}_s y \mathcal{M}_c son parámetros de nuestro sistema de especificación Π_{Lac} .
- ★ Los conjuntos de reglas \mathcal{R}_e , \mathcal{R}_s y \mathcal{R}_c , están asociados a las regiones relevantes. Estas reglas describen las interacciones que tienen lugar en los compartimentos especificados.

$$\mathcal{R}_e = \{r_4, r_8\}$$

$$\mathcal{R}_s = \{r_2, r_3, r_5, r_6, r_9, r_{11}, r_{18}, r_{20}, r_{22}\}$$

$$\mathcal{R}_c = \{r_1, r_7, r_{10}, r_{12}, r_{13}, r_{14}, r_{15}, r_{16}, r_{17}, r_{19}, r_{21}, r_{23}, \dots, r_{55}\}$$

Entidades moleculares

Las proteínas y enzimas involucradas en la selección de azúcares del medio y en su metabolismo, serán representadas por **símbolos** de un alfabeto: Γ_{prot} .

- Por ejemplo, suponemos que las moléculas EIICB y β -galactosidasa son entidades sin estructura interna (en el modelo).

El alfabeto Γ_{prot} es el siguiente conjunto:

{EIIA, EIIA~P, EIICB, EIICB-EIIA~P, EIICB~P, Gluc, Gluc~P, EIICB~P-Gluc, Lact, LacY, Lact-LacY, AC, AC-EIIA, AC-EIIA~P, ATP, AC-EIIA~P-ATP, LacY-EIIA, β -Galac, Lact, β -Galac-Lact, Allolact, Lacl, Lacl-Alloct, CRP, CRP-cAMP, CRP-cAMP₂, RNAP}

Symbol	Molecule
EICB EIICB~P	Glucose Transporter Enzyme IICB and its phosphorylated state
EIIA EIIA~P	Glucose Transporter Enzyme IIA and its phosphorylated state
EIICB-EIIA~P	Complex transporter enzymes IICB and phosphorilated IIA
Gluc Gluc~P	Glucose and its phosphorylated state
EIICB~P-Gluc	Complex transporter IICB phosphorylated and glucose
Lact	Lactose
LacY	Lactose permease
Lact-LacY	Complex lactose permease and lactose
LacY-EIIA	Complex lactose permease and transporter enzyme IIA
ATP cAMP	Adenosine triphosphate and cyclic adenosine monophosphate
AC	Adelynate Cyclase
AC-EIIA AC-EIIA~P AC-EIIA~P-ATP	ATP and AC complexes with and ATP
β -Galac β -Galac-Lact	β -Galactosidase and its complex with lactose
Allolact	Allolactose, the inducer
LacI LacI-Alloct	The lac repressor and its complex with allolactose
CRP CRP-cAMP CRP-cAMP ₂	cAMP Receptor Protein, the activator, its complex with cAMP and its dimer
RNAP	RNA polymerase

Representación de genes, RNAm y sitios

- Los genes que intervienen aparecen en una estructura lineal en donde el orden es esencial (orden en el que los genes son expresados).
- La estructura lineal de los genes y del RNAm serán especificadas a través de cadenas.
- Los sitios relevantes en el operón lactosa y en el RNAm serán representados usando símbolos de los alfabetos Γ_{dna} y Γ_{rna} , resp.

- El alfabeto Γ_{dna} es el siguiente conjunto:

$\{cap, cap^{CRP-cAMP_2}, op, op^{LacI}, lacZ_s, lacZ_m, lacZ_e, lacY_s, lacY_m, lacY_e, lacA_s, lacA_m, lacA_e\}$

- El alfabeto Γ_{rna} es el siguiente conjunto:

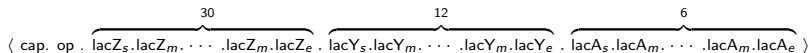
$\{Rib, \overline{op}, \overline{lacZ_s}, \overline{lacZ_m}, \overline{lacZ_e}, \overline{lacY_s}, \overline{lacY_m}, \overline{lacY_e}, \overline{lacA_s}, \overline{lacA_m}, \overline{lacA_e}\}$

Símbolos que indican los sitios relevantes

Symbol	Site
cap	Free CAP site where the activator CRP-cAMP ₂ binds
cap ^{CRP-cAMP₂}	CAP site occupied by the activator CRP-cAMP ₂ binds
op	Free operator site where the repressor LacI binds
op ^{LacI}	Operator site occupied by the repressor LacI binds
lacZ _s lacY _s lacA _s	Sites marking the start point of the lacZ, lacY and lacA gene respectively
lacZ _m lacY _m lacA _m	Sites located in the middle of the lacZ, lacY and lacA gene respectively
lacZ _e lacY _e	Sites located in the final point of the lacZ and lacY gene respectively.
lacA _e	Site marking the end of gene lacA which coincides with the transcription termination site of the lac operon
Rib	Ribosome Site
op	Site marking the starting point of the mRNA transcript
$\overline{\text{lacZ}}_s$ $\overline{\text{lacZ}}_e$ $\overline{\text{lacY}}_s$ $\overline{\text{lacY}}_e$ $\overline{\text{lacA}}_s$ $\overline{\text{lacA}}_e$	Sites marking the beginning and end of the reading frames in the mRNA for the genes lacZ, lacY and lacA respectively
$\overline{\text{lacZ}}_m$ $\overline{\text{lacY}}_m$ $\overline{\text{lacA}}_m$	Sites located in the middle of reading frames of lacZ, lacY and lacA gene respectively

Cadena que representa el operón lactosa en E. coli

La siguiente cadena representa el operón lactosa en E. coli:



La RNA polimerasa ocupa alrededor de 100 nucleótidos. Por ello, cada símbolo $\langle \text{lacZ}_i \rangle$, $\langle \text{lacY}_i \rangle$ y $\langle \text{lacA}_i \rangle$ ($i = s, m, e$) representa una sucesión de 100 nucleótidos del correspondiente gen.

Se tienen 30 sitios lacZ, 12 lacY y 6 lacA que representan los 3000, 1200 y 600 nucleótidos de los correspondientes genes.

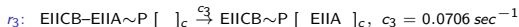
El lugar de enlace de CAP y el operador son representados por $\langle \text{cap} \rangle$ y $\langle \text{op} \rangle$, respectivamente.

Reglas que describen el transporte glucosa

La entrada de glucosa desde el entorno al citoplasma celular comienza con el reclutamiento de la proteína EIIA~P del citoplasma por parte de la proteína transmembrana EIICB (regla r_1): este proceso es reversible (regla r_2).

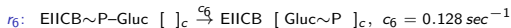
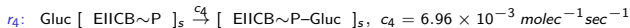


La proteína EIIA~P en la piel produce la fosforilización de EIICB y envía la molécula EIIA al citoplasma (regla r_3).



La proteína EIICB~P en la superficie celular enlaza con la glucosa transportándola a la la piel (regla r_4): este proceso es reversible (regla r_5).

Finalmente, la glucosa es transportada al interior de la bacteria (regla r_6), siendo fosforilada, Gluc~P.



Reglas que describen el transporte de lactosa

La entrada de lactosa desde el entorno al citoplasma celular comienza la presencia de la permeasa LacY en la superficie celular. Aunque LacY es sintetizada en el citoplasma, ella se puede mover a la superficie celular (regla r_7).

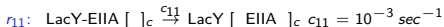


La lactosa presente en el entorno enlaza con la permeasa Lac Y (regla r_8). Seguidamente, la lactosa es transportada Por Lac Y al citoplasma (regla r_9).



El transporte de la lactosa desde el exterior al citoplasma es **inhibido** por EIIA que enlaza a LacY produciendo el complejo LacY-EIIA (regla r_{10}): este proceso es reversible (regla r_{11})

La lactosa no puede enlazar con el complejo LacY-EIIA que impide el transporte de la lactosa al interior de la célula.



La **presencia de glucosa** en el entorno:

- ★ Propicia la producción de EIIA en el citoplasma.
- ★ Las moléculas de EIIA inhibe la permeasa Lac Y.
- ★ Se inhibe la entrada de lactosa al sistema.



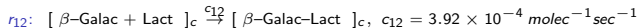
Reglas que describen la actividad de la β -Galactosidasa

La β -galactosidasa hidroliza la lactosa en glucosa y galactosa.

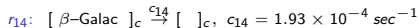
La producción de glucosa a partir de lactosa no es crucial en la regulación de la expresión de los genes en el operón lactosa (no lo consideramos en el modelo).

La producción de la alolactosa es un paso importante en la regulación ya que actúa como inductor.

- ★ La interacción entre la β -galactosidasa y la lactosa produce un complejo en el citoplasma (regla r_{12}).
- ★ El complejo anterior produce alolactosa en el citoplasma (regla r_{13}).



La maquinaria celular degrada la β -galactosidasa en el citoplasma (regla r_{14}).



Reglas que describen la actividad de la Alolactosa

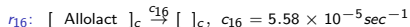
La Alolactosa es una señal de la presencia de lactosa en el entorno.

Actúa como inductor del sistema de regulación de genes del operón lactosa mediante la inhibición del represor, Lacl, con una interacción directa (regla r_{15}).



Cuando la alolactosa se une al represor, éste cambia de conformación impidiendo su enlace con el operador del sistema.

La alolactosa se degrada también en el citoplasma (regla r_{16}).

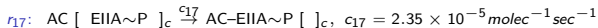


Reglas que describen la actividad de la Adenylate Cyclasa

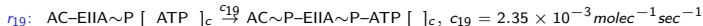
La Adenylate Cyclase (AC) regula la producción de las moléculas de cAMP

★ La concentración de cAMP es inversamente proporcional a la concentración de glucosa en el entorno.

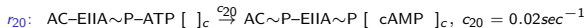
Para ello, la AC recluta del citoplasma a moléculas EIIA~P (regla r_{17}): este proceso es reversible (regla r_{18}).



El complejo AC-EIIA~P recluta ATP del citoplasma fosforilando AC (regla r_{19}).

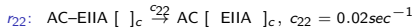
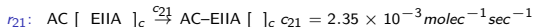


El complejo AC~P-EIIA~P-ATP produce cAMP en el citoplasma (regla r_{20}).



El número de moléculas de EIIA en el citoplasma es un indicador de la presencia de glucosa en el entorno.

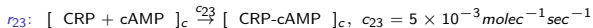
Mecanismo de represión catabólica: El enlace de EIIA con AC en la membrana celular, inhibe la producción de cAMP (regla r_{21}): este proceso es reversible (regla r_{22}).



La represión de la actividad de AC impide la producción de cAMP y del activador CRP-cAMP₂; es decir, se inhibe la activación del operón lactosa.

Reglas que describen la formación del activador CRP-cAMP₂

La proteína CRP interactúa con cAMP en el citoplasma produciendo el complejo CRP-cAMP (regla r_{23}). Éste formará un dímero CRP-cAMP₂ (regla r_{24}).



El dímero CRP-cAMP₂ permite incrementar la tasa de transcripción de los genes codificados en el operón lactosa (propiciando la unión de la RNA-polimerasa al promotor).

Reglas que describen la activación y represión del operón lactosa

La tasa de transcripción del operón lactosa viene determinada por el estado del *lac operon switch*.

Este switch lo integra:

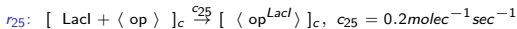
- ★ El CAP site, $\langle \text{cap} \rangle$, donde enlaza el activador.
- ★ El operador, $\langle \text{op} \rangle$, en donde enlaza el represor.

El switch tiene cuatro posibles configuraciones que dependen de la ocupación del CAP site y del operador:

- ★ $\langle \text{cap.op} \rangle$, $\langle \text{cap.op}^{\text{LacI}} \rangle$, $\langle \text{cap}^{\text{CRP-cAMP}_2}.\text{op} \rangle$ y $\langle \text{cap}^{\text{CRP-cAMP}_2}.\text{op}^{\text{LacI}} \rangle$.

Estas configuraciones se alcanzan mediante enlaces y disociación de las reglas de factores de transcripción.

En ausencia de lactosa, LacI se activará y enlazará con el operador inhibiendo la transcripción del operón lactosa (regla r_{25}). Ocasionalmente LacI se cae del operador permitiendo la producción basal de proteínas (regla r_{26}).



En ausencia de glucosa, hay un gran número de moléculas cAMP que producirán muchos activadores, CRP-cAMP₂ que enlazarán con el promotor (regla r_{27}). Ocasionalmente, CRP-cAMP₂ puede caer del promotor (regla r_{28}).



Reglas que describen el comienzo de la transcripción en el operón lactosa

Primer paso: el enlace de la RNA polimerasa (RNAP) al lac operon switch.

La afinidad entre el lac operon switch y la RNAP depende de la configuración del switch.

- Por una parte, cuando el CAP site está libre, $\langle \text{cap} \rangle$, la RNAP rara vez enlaza con el promotor y se produce una tasa de transcripción basal (rule r_{29}).
- Por otra parte, cuando el activador CRP-cAMP_2 ocupa el CAP site, $\langle \text{cap}^{\text{CRP-cAMP}_2} \rangle$, se produce un incremento en la tasa de transcripción de aproximadamente 40 veces (regla r_{30}).

$$r_{29}: [\text{RNAP} + \langle \text{cap} \rangle]_c \xrightarrow{c_{29}} [\langle \text{cap.RNAP} \rangle]_c, c_{29} = 5 \times 10^{-4} \text{ molec}^{-1} \text{ sec}^{-1}$$

$$r_{30}: [\text{RNAP} + \langle \text{cap}^{\text{CRP-cAMP}_2} \rangle]_c \xrightarrow{c_{30}} [\langle \text{cap}^{\text{CRP-cAMP}_2} . \text{RNAP} \rangle]_c, c_{30} = 2 \times 10^{-2} \text{ molec}^{-1} \text{ sec}^{-1}$$

La RNAP inicia la transcripción produciendo los nucleótidos complementarios del operador $\langle \overline{\text{op}} \rangle$ (regla r_{31}). El sitio $\langle \overline{\text{op}} \rangle$ marca el comienzo de la transcripción de un mRNA.

$$r_{31}: [\langle \text{RNAP.op} \rangle]_c \xrightarrow{c_{31}} [\langle \text{op. } \overline{\text{op}} . \text{RNAP} \rangle]_c, c_{31} = 2 \text{ sec}^{-1}$$

Tras aplicar la regla r_{31} la subcadena $\langle \text{op} \rangle$ queda libre y otra RNAP puede comenzar la transcripción (incluso antes de que la primera RNAP acabe de transcribir el operón).

Por tanto, estamos simulando la transcripción por diferentes RNAP como un proceso concurrente.

Reglas que describen la **elongación** del RNAm en el operón lactosa

Al principio de la elongación del RNAm, la RNAP se mueve a lo largo del gen lacZ transcribiéndolo en RNAm.

- ★ En primer lugar, transcribe $\langle \text{lacZ}_s \rangle$ y une los complementarios $\langle \overline{\text{lacZ}_s} \rangle$ al RNAm (regla r_{32}).

$$r_{32}: [\langle \overline{\text{op. RNAP. lacZ}_s} \rangle]_c \xrightarrow{c_{32}} [\langle \text{lacZ}_s.\overline{\text{op. lacZ}_s}.\text{RNAP} \rangle]_c, c_{32} = 2\text{sec}^{-1}$$

$\langle \overline{\text{lacZ}_s} \rangle$ representa el RBS para este gen. Una vez producido este sitio, la traducción puede comenzar, incluso antes de que finalice la transcripción. En nuestra simulación, ambos procesos tienen lugar en paralelo.

Durante la transcripción se produce una cadena complementaria de RNAm, descrita por la producción de $\langle \overline{\text{lacZ}_m} \rangle$ que se une al RNAm representado por la subcadena $\langle \overline{\text{op.w}} \rangle$, $w \in \Gamma_{ma}^*$, y $\langle \text{lacZ}_m \rangle$ se queda atrás (regla r_{33}).

$$r_{33}: [\langle \overline{\text{op.w.RNAP.lacZ}_m} \rangle]_c \xrightarrow{c_{33}} [\langle \text{lacZ}_m.\overline{\text{op.w. lacZ}_m}.\text{RNAP} \rangle]_c, c_{33} = 2\text{sec}^{-1}$$

Cuando la RNAP alcanza el final del gen LacZ une $\langle \overline{\text{lacZ}_e} \rangle$ a la subcadena $\langle \overline{\text{op.w}} \rangle$, $w \in \Gamma_{ma}^*$, que describe el crecimiento del RNAm (regla r_{34}).

$$r_{34}: [\langle \overline{\text{op.w.RNAP.lacZ}_e} \rangle]_c \xrightarrow{c_{34}} [\langle \text{lacZ}_e.\overline{\text{op.w. lacZ}_e}.\text{RNAP} \rangle]_c, c_{34} = 2\text{sec}^{-1}$$

$\langle \overline{\text{lacZ}_e} \rangle$ representa un sitio de final de la traducción. Cuando los ribosomas alcanzan este sitio, se disocian dejando libre la proteína β -galactosidasa codificada por lacZ.

Las siguientes reglas describen la transcripción de los genes lacY y lacA, de forma similar al caso del gen lacZ.

$$r_{35}: [\langle \overline{\text{op.w. RNAP. lacY}_s} \rangle]_c \xrightarrow{c_{35}} [\langle \text{lacY}_s.\overline{\text{op.w. lacY}_s}.\text{RNAP} \rangle]_c, c_{35} = 2\text{sec}^{-1}$$

$$r_{36}: [\langle \overline{\text{op.w. RNAP. lacY}_m} \rangle]_c \xrightarrow{c_{36}} [\langle \text{lacY}_m.\overline{\text{op.w. lacY}_m}.\text{RNAP} \rangle]_c, c_{36} = 2\text{sec}^{-1}$$

$$r_{37}: [\langle \overline{\text{op.w. RNAP. lacY}_e} \rangle]_c \xrightarrow{c_{37}} [\langle \text{lacY}_e.\overline{\text{op.w. lacY}_e}.\text{RNAP} \rangle]_c, c_{37} = 2\text{sec}^{-1}$$

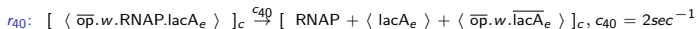
$$r_{38}: [\langle \overline{\text{op.w. RNAP. lacA}_s} \rangle]_c \xrightarrow{c_{38}} [\langle \text{lacA}_s.\overline{\text{op.w. lacA}_s}.\text{RNAP} \rangle]_c, c_{38} = 2\text{sec}^{-1}$$

$$r_{39}: [\langle \overline{\text{op.w. RNAP. lacA}_m} \rangle]_c \xrightarrow{c_{39}} [\langle \text{lacA}_m.\overline{\text{op.w. lacA}_m}.\text{RNAP} \rangle]_c, c_{39} = 2\text{sec}^{-1}$$

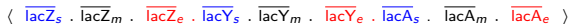
Regla que describe el final de la transcripción del operón lactosa

La transcripción del operón lactosa finaliza cuando la RNAP alcanza el sitio de terminación de la transcripción del último gen lac A: ese sitio está representado por la cadena $\langle \overline{\text{lacA}_e} \rangle$.

Cuando la RNAP alcanza este sitio, enlaza los ribonucleótidos, $\langle \overline{\text{lacA}_e} \rangle$ al RNAm $\langle \overline{\text{op.w.}} \rangle$, $w \in \Gamma_{ma}^*$, y lo disocia del operón, liberando el RNAm, $\langle \overline{\text{op.w.}} \cdot \overline{\text{lacA}_e} \rangle$ (regla r_{40}).



La molécula de RNAm producida se describe en nuestro modelo como sigue:



Los RBS de los genes lac Z, lac Y y lac A son $\langle \overline{\text{lacZ}_s} \rangle$, $\langle \overline{\text{lacY}_s} \rangle$ y $\langle \overline{\text{lacA}_s} \rangle$.

Los lugares de terminación de la traducción de lac Z, lac Y y lac A son $\langle \overline{\text{lacZ}_e} \rangle$, $\langle \overline{\text{lacY}_e} \rangle$ y $\langle \overline{\text{lacA}_e} \rangle$.

Reglas que describen el inicio de la traducción

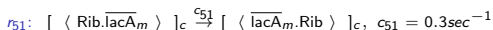
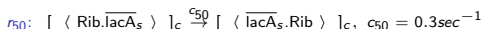
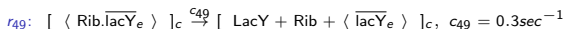
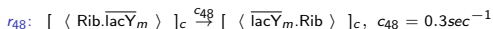
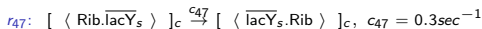
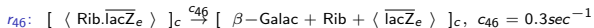
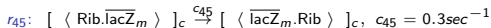
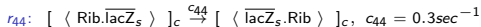
La traducción correspondiente a los genes *lac Z*, *lac Y* y *lac A* comienza cuando los ribosomas, Rib, reconocen los respectivos RBS, representados por $\langle \overline{\text{lacZ}}_s \rangle$, $\langle \overline{\text{lacY}}_s \rangle$ y $\langle \overline{\text{lacA}}_s \rangle$ (reglas r_{41} , r_{42} y r_{43}).



Reglas que describen la traducción y la disociación de los ribosomas

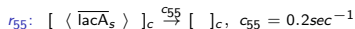
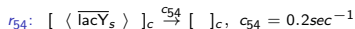
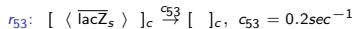
Durante la traducción, los ribosomas se mueven a lo largo de los sitios $\langle \overline{\text{lacZ}}_m \rangle$, $\langle \overline{\text{lacY}}_m \rangle$, $\langle \overline{\text{lacA}}_m \rangle$ del RNAm: elongación de la traducción (reglas r_{44} , r_{45} , r_{47} , r_{48} , r_{50} y r_{51}).

Cuando los ribosomas alcanzan los respectivos lugares de terminación, $\overline{\text{lacZ}}_e$, $\overline{\text{lacY}}_e$, $\overline{\text{lacA}}_e$, se disocian del RNAm dejando libre las proteínas β -galactosidase, Lac Y or Lac A, disociándose los ribosomas (reglas r_{46} , r_{49} y r_{52}).



Reglas que describen la degradación del RNAm

El RNAm se degrada cuando ciertas enzimas enlazan con los respectivos RBS, $\langle \overline{\text{lacZ}}_s \rangle$, $\langle \overline{\text{lacY}}_s \rangle$, $\langle \overline{\text{lacA}}_s \rangle$, y los eliminan. Con ello, evitan que nuevos ribosomas inicien la traducción (reglas r_{53} , r_{54} y r_{55}).



Parámetros del sistema: escenarios a estudiar

Parámetros de nuestro sistema de especificación $\mathcal{P}(\Pi_{Lac})$: los multiconjuntos iniciales y las constantes estocásticas asociadas a las reglas.

- En nuestro caso, las constantes estocásticas tendrán unos valores fijos (de un conjunto C).
- Los multiconjuntos iniciales constituyen los parámetros actuales del sistema: especificaremos el número de objetos y cadenas presentes inicialmente en cada compartimento.
- Analizaremos cuatro posibles escenarios iniciales en el entorno $\mathcal{M}_e^1, \mathcal{M}_e^2, \mathcal{M}_e^3, \mathcal{M}_e^4$.
 - ★ Ausencia de glucosa y de lactosa en el entorno: $\mathcal{M}_e^1 = \{\lambda, \lambda\}$.
 - ★ Abundante glucosa y ausencia de lactosa en el entorno: $\mathcal{M}_e^2 = \{\text{Gluc}^{300000}, \lambda\}$.
 - ★ Ausencia de glucosa y abundancia de lactosa en el entorno: $\mathcal{M}_e^3 = \{\text{Lact}^{300000}, \lambda\}$.
 - ★ Abundancia de glucosa y de lactosa en el entorno: $\mathcal{M}_e^4 = \{\text{Gluc}^{300000}, \text{Lact}^{300000}\}$.
- Analizaremos un sólo escenario en la superficie celular, dado por \mathcal{M}_s .
 - ★ $\mathcal{M}_s = \{\text{EIICB}^{2500} + \text{EIICB} \sim \text{P}^{15000} + \text{AC}^{10000} + \text{LacY}^{3000}, \lambda\}$.
- Analizaremos un sólo escenario en el citoplasma, dado por $\mathcal{M}_c = \{w_3, s_3\}$; en donde w_3 representa los objetos iniciales RNAP^{300} , Rib^{3000} , EIIA^{2000} , $\text{EIIA} \sim \text{P}^{13000}$, $\text{ATP}^{1000000}$, $\beta - \text{Galac}^{3000}$, LacI^{1500} , CRP^{10000} , y s_3 la cadena inicial que representa el operón lactosa:

$$\langle \text{cap. op.} \cdot \overbrace{\text{lacZ}_s \cdot \text{lacZ}_m \cdot \dots \cdot \text{lacZ}_m \cdot \text{lacZ}_e}^{30} \cdot \overbrace{\text{lacY}_s \cdot \text{lacY}_m \cdot \dots \cdot \text{lacY}_m \cdot \text{lacY}_e}^{12} \cdot \overbrace{\text{lacA}_s \cdot \text{lacA}_m \cdot \dots \cdot \text{lacA}_m \cdot \text{lacA}_e}^6 \rangle$$

Escenarios a estudiar

Los modelos de computación que corresponden a los distintos escenarios son:

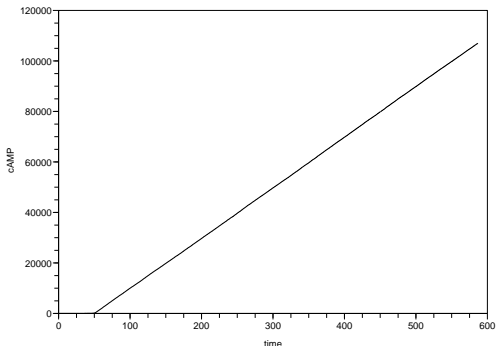
1. $\Pi_{Lac}^{(1)} = (\Pi_{Lac} ; \{M_e^1, M_s, M_c\}, C)$: Ausencia de glucosa y de lactosa.
2. $\Pi_{Lac}^{(2)} = (\Pi_{Lac} ; \{M_e^2, M_s, M_c\}, C)$: Abundante glucosa y ausencia de lactosa.
3. $\Pi_{Lac}^{(3)} = (\Pi_{Lac} ; \{M_e^3, M_s, M_c\}, C)$: Ausencia de glucosa y abundancia de lactosa.
4. $\Pi_{Lac}^{(4)} = (\Pi_{Lac} ; \{M_e^4, M_s, M_c\}, C)$: Abundancia de glucosa y de lactosa.

Análisis de la regulación de genes en el operón lactosa (I)

$\Pi_{Lac}^{(1)}$: comportamiento del sistema sin glucosa y sin lactosa en el entorno

Cuando la glucosa no está presente en el entorno, EIIA~P no es consumido y, por ello, se activa AC sobre la superficie celular.

Una vez activado el AC, se produce un elevado número de moléculas de cAMP, como muestra la figura siguiente.



cAMP enlaza con CRP dando lugar a CRP-cAMP que, por dimerización, produce el activador CRP-cAMP₂.

Análisis de la regulación de genes en el operón lactosa (II)

$\Pi_{Lac}^{(1)}$: comportamiento del sistema sin glucosa y sin lactosa en el entorno

El lugar CAP será ocupado por moléculas CRP-cAMP₂ que asistirán a la RNAP para que enlace con el promotor.

Al no haber lactosa en el entorno, no se producirá alolactosa en el citoplasma y, por tanto, el represor Lac I estará activo y colapsará el operador.

En este escenario, en nuestro modelo se obtiene el siguiente estado característico del operón lactosa:

```
cap-CRP-cAMP_2 RNAP op-LacI lacZ_s lacZ_m -op Rib
-lacZ_s -lacZ_m RNAP lacZ_m lacZ_m lacZ_m lacZ_m
lacZ_m lacZ_m lacZ_m -op Rib -lacZ_s -lacZ_m -lacZ_m
-lacZ_m -lacZ_m -lacZ_m -lacZ_m-lacZ_m -lacZ_m RNAP
lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m
lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m
lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_e
lacY_s lacY_m lacY_m lacY_m lacY_m lacY_m lacY_m lacY_m
lacY_m lacY_m lacY_m lacY_m lacY_e lacA_s lacA_m lacA_m
lacA_m lacA_m lacA_e
```

(1)

Obsérvese que el CRP-cAMP₂ está en el CAP site asistiendo a la RNAP para enlazar al promotor.

Una vez enlazado, la RNAP está lista para la transcripción siempre que el represor Lac I no bloquee al operador.

En resumen, la configuración del lac operon switch en este escenario es $\langle \text{cap}^{\text{CRP-cAMP}_2} \cdot \text{RNAP} \cdot \text{op}^{\text{LacI}} \rangle$

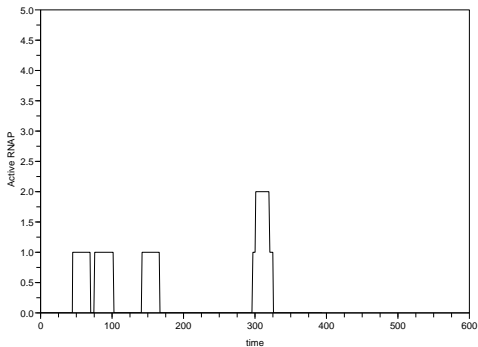
- ★ Se produce un leve incremento en la expresión de los genes codificados en el operón lactosa: **sin glucosa ni lactosa en el entorno, la bacteria procede a la producción basal de lactosa.**

Análisis de la regulación de genes en el operón lactosa (III)

$\Pi_{Lac}^{(2)}$: comportamiento del sistema con abundante glucosa y sin lactosa en el entorno

Al no haber lactosa en el entorno, no se producirá lactosa en el citoplasma y el represor Lac I estará activo enlazándose al operador del operón.

En este escenario la configuración del lac operon switch será $\langle \text{cap.op}^{LacI} \rangle$: **con glucosa pero sin lactosa en el entorno, la bacteria procede a la producción basal de lactosa**, como se puede observar en la siguiente figura.



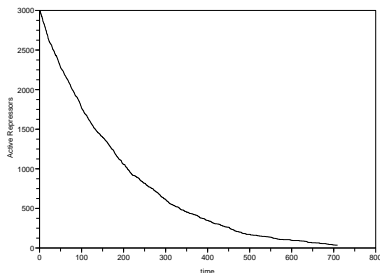
Al haber la glucosa en el entorno hay poca necesidad de metabolizar lactosa.

Análisis de la regulación de genes en el operón lactosa (IV)

$\Pi_{Lac}^{(3)}$: comportamiento del sistema sin glucosa pero con abundante lactosa en el entorno

En este escenario:

- La ausencia de glucosa permite que EIIA~P interactúe con AC para sintetizar un elevado número de cAMP y de activadores CRP-cAMP₂: el CAP será ocupado propiciando el reclutamiento de RNAP.
- La abundante lactosa en el entorno será transportada al citoplasma por LacY, expresada a nivel basal. La β -galactosidasa, prroducida en el citoplasma a nivel basal, hidroliza la lactosa en galactosa y glucosa, dando lugar a la alolactosa que actuará como inductor enlazando al represor LacI.
- En la siguiente figura se observa cómo los represores activos son inhibidos rápidamente.

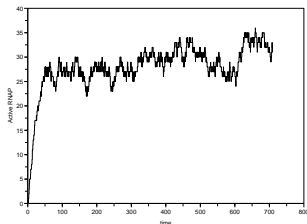


Análisis de la regulación de genes en el operón lactosa (V)

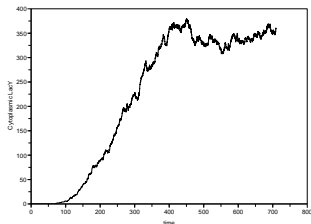
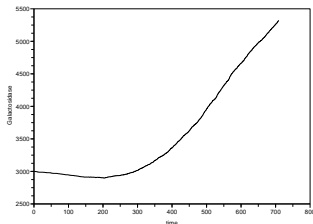
$\Pi_{Lac}^{(3)}$: comportamiento del sistema sin glucosa pero con abundante lactosa en el entorno

En este escenario: el operón lactosa será, a la vez, inducido (ningún represor se unirá al operador) y activado (el activador se enlazará al CAP site).

- La configuración del switch será $\langle \text{cap}^{\text{CRP-cAMP}}_{2.op} \rangle$ y los genes codificados en el operón se transcribirán masivamente, como se deduce del número de RNAP que están transcribiendo el operón.



- Esto provocará un incremento en el número de moléculas de β -galactosidasa y LacY.



Análisis de la regulación de genes en el operón lactosa (VI)

$\Pi_{Lac}^{(4)}$: comportamiento del sistema con abundante glucosa y abundante lactosa en el entorno

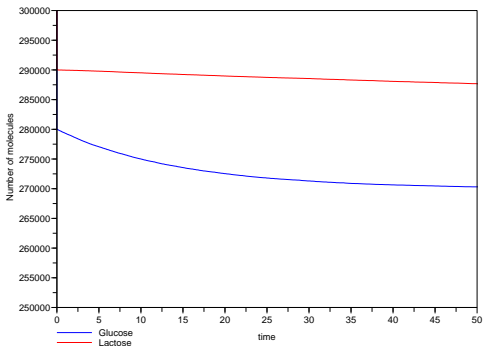
La presencia de glucosa produce un bajo número de activadores y el operón lactosa no será activado.

Con la presencia de lactosa, los genes codificados por el operón serán transcritos a muy bajos niveles.

También existirá poca absorción de lactosa desde el entorno y el LacY será inhibido.

El número de moléculas de glucosa en el entorno decrece mientras que el de la lactosa permanece casi constante

- Pocas moléculas de lactosa son transportadas al interior de la bacteria.

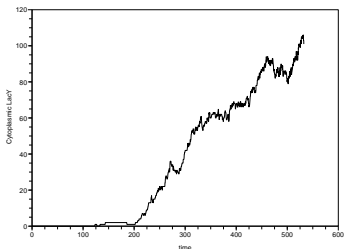
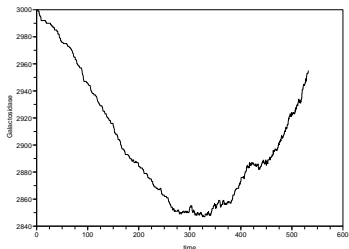


Análisis de la regulación de genes en el operón lactosa (VII)

$\Pi_{Lac}^{(4)}$: comportamiento del sistema con abundante glucosa y abundante lactosa en el entorno

No obstante, alguna lactosa estará presente en el citoplasma y producirá alolactosa, inhibiendo al represor permitiendo la transcripción a bajos niveles.

En estas condiciones, la configuración del lac operon switch será $\langle \text{cap.op} \rangle$. Ningún activador CRP-cAMP₂ se unirá al CAP site: no habrá transcripción o se producirá a bajos niveles como puede verse en las siguientes figuras (el número de moléculas de galactosidasa y LacY comienza lentamente).



Mientras la bacteria pueda metabolizar glucosa, tiene poca necesidad de metabolizar lactosa. No obstante, puesto que hay lactosa, la bacteria no ignorará completamente el suministro de azúcar.